National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, tests publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

NL-339 (r. 88/04)

Canada

THE UNIVERSITY OF ALBERTA

A Bi-spectral Approach to the Estimation of Total Cloud-Cover

by

David Allen Ball

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF Master of Science
IN
Meteorology

Geography

EDMONTON, ALBERTA
Fall 1987

ISBN 0-315-40906-1

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR David Allen Ball

TITLE OF THESIS A Bi-spectral Approach to the Estimation of Total Cloud-Cover

DEGREE FOR WHICH THESIS WAS PRESENTED Master of Science

YEAR THIS DEGREE GRANTED Fall 1987

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) ....

PERMANENT ADDRESS:

[illegible]

Edmonton, Alberta

[illegible]

DATED 19

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled A Bi-spectral Approach to the Estimation of Total Cloud-Cover submitted by David Allen Ball in partial fulfilment of the requirements for the degree of Master of Science in Meteorology.

ERReinelt

Supervisor

Date 24 August 1987

Abstract

Data from the NOAA-9 polar-orbiting satellite have been used to estimate the total cloud-cover within selected areas of the eastern Pacific Ocean during July of 1987. Each pixel scanned by the Advanced Very High Resolution Radiometer (AVHRR) exhibits a characteristic radiation signature that depends on whether cloud or ocean surfaces are being sensed. By determining the bi-variate frequency distribution of data from pairs of AVHRR channels 2, 3, and 4, it is possible to distinguish between areas of cloud and areas of ocean. A computer algorithm was designed to remove the effects of the ocean surface, thereby allowing fractional cloud-cover estimates to be made. In situations where the sample area was within 250km of the coast, the cloud-cover was overestimated, probably because of variable concentrations of surface plankton. Minor underestimates were produced in near-overcast cases because of incorrect selection of the 'best' estimate by the algorithm.

The use of data from AVHRR channel 3 has been found to be useful in delineating areas of low cloud over water. Its use, in combination with data from AVHRR channel 4 could allow this method to be used at night.

In general, this procedure provided estimates within 10% of those produced by visual inspection of satellite imagery. Operational use of this technique is limited, at present, but there is hope that new and more powerful computer hardware will allow it to be used more widely.

Acknowledgements

I would like to thank my supervisor, Dr. E.R. Reinelt for his quiet words of encouragement during the preparation of this thesis. I would also like to thank the other members of my examination comittee, Dr. E.P. Lozowski and Dr. G. Rostoker, both of whom showed a prodigious amount of patience on very short notice.

Special thanks go to Ron Goodson for his invaluable assistance during the development of the computer software used in this study.

And last, but certainly not least, thanks go to my wife, Tracy, who day after day manages to prove that tired old saying: behind every man...

Table of Contents

List of Tables

List of Figures

List of Plates

1. Introduction

The ability to forecast severe weather development and synoptic scale dynamics is greatly enhanced by using satellite imagery. In fact, operational applications of satellite data and imagery have been increasing since multi-spectral high resolution radiometers were developed in the early 1970's. Much of the information available from radiometric data, however, is not visually apparent, and requires computer processing to reveal it. Attempts have been made over the past decade to develop automated techniques that objectively analyze radiometric data. Despite promising results, many fail to completely identify pixels containing cloud, a fundamental problem that must be addressed by any automated procedure. A pixel (picture element) in this context is a single digital measurement.

Of particular importance to long-term climate models are accurate sea surface temperature (SST) estimates. For the most part, these are obtained using satellite radiometer data; although surface based observations are available, consistent large scale estimates of the world's ocean surface temperatures are more conveniently determined via satellite.

Early attempts to estimate SST's were made by constructing histograms showing the variance of satellite-measured temperatures across a given area (Smith et al., 1970). In the absence of cloud, these histograms are normally distributed, with the center of the peak providing the best estimate of the SST. The dispersion of the peak is produced by instrument noise. When cloud contaminates the field of view, the histogram shows a marked dispersion toward lower temperatures. Although reasonable SST estimates were produced, this technique had difficulty dealing with gradients in the sample area, as well as with low cloud that was thermally indistinguishable from the ocean surface.

The realization that a single radiometer channel is incapable, in most cases, of sensing the presence of all varieties of cloud within an area led to the development of bi-spectral cloud detection techniques that use visible and infra-red data to provide complementary information about cloud presence. These bi-spectral techniques provide much better cloud-cover estimates than single channel procedures. Maul, (1981), has developed a cloud detection technique for processing data from geostationary satellites (GOES), while Simmer et al., (1982), and Desbois et al., (1982), have developed methods for using METEOSAT data.

The basis for this study is a bi-spectral histogram technique developed by Phulpin et al., (1983), that produces total cloud cover estimates agreeing with those of visual nephanalysis in about 87% of the cases. Data from Advanced Very High Resolution Radiometer (AVHRR) channels 1, 2, and 4 (Table 1.1) aboard the NOAA-7 polar-orbiting satellite were used.

Data from pairs of these bands were partitioned into statistical bins: the visible data (AVHRR channels 1 and 2) in increment of approximately 6% albedo,and the infra-red data (AVHRR channel 4), in increments of approximately 0.3°K. Sample areas containing approximately 1250 pixels were selected. Using the band pair data as abscissa, bi-spectral histograms were constructed, having frequency as the ordinate. The construction is accomplished by counting the number of pixels that lie in each of the bins. As in the one-dimensional case, the histograms are normally distributed. Each peak within the field represents the thermal or reflectance characteristics of a stratified atmospheric layer (planetary surface or cloud). Fractional cloud-cover estimates for each layer were obtained by fitting a Gaussian surface to each of the peaks attributed to cloud, and subtracting it from the data. The cloud-cover estimate is the ratio of the number of points accounted for by extraction to the total number of points in the field of view. Peaks lying close to each other are assumed to be the result of gradients within the field of view, and their contributions are merged into a single estimate. Residual extraction points are ignored.

Table 1.1 AVHRR Channels

	Wavelength Interval (μm)	
Channel	NOAA-9	NOAA-10
1. Vis	0.55-0.68	0.55-0.68
2. Near IR	0.725-1.10	0.725-1.10
3. Vis/Thermal IR	3.55-3.93	3.55-3.93
4. Thermal IR	10.5-11.5	10.5-11.5
5. Thermal IR	11.5-12.5	Ch.4 repeated

Since each pixel scanned by the radiometer represents a view of either cloud or ocean surfaces, these residual points cannot be neglected. It is necessary, therefore, for the fitting procedure to be as accurate as possible.

In this study, a bi-spectral histogram procedure has been developed for estimating total cloud-cover, using high resolution satellite data obtained from the Advanced Very High Resolution Radiometer (AVHRR) aboard the NOAA-9 polar orbiting satellite. The best estimates produced by Phulpin et al., (1983) were obtained using a combination of AVHRR bands 2 and 4, which will be used in this study as well. Additionally, the correlation between bands 2 and 4 and AVHRR channel 3, the hybrid infra-red/visual channel, will be examined. Visual satellite based nephanalysis will be used to estimate the accuracy of the automated procedure.

2. The Satellite

2.1 Introduction

Satellite data used in this study were obtained from NOAA-9, a member of the TIROS series of sun-synchronous, polar orbiting satellites. Since Schwalb, (1978), provides a complete description of the design and operational characteristics of the TIROS satellites, only a brief summary is presented here.

NOAA-9 circumnavigates the earth in a near polar orbit, at a nominal height of 850 kilometers. The AVHRR on the satellite scans the earth with a mirror, rotating at 360 RPM perpendicular to the orbital path. Data are retained for points lying within ±55° of the satellite subpoint, limiting the influence of limb darkening and foreshortening on the data, ie. the apparent cross-scan reduction in the size of a pixel produced by the curvature of the Earth. The radiometer scans the Earth in five spectral bands (channel 4 data are repeated in channel 5 for NOAA-10), whose characteristics are summarized in Table 1.1. A total of 2048 samples are obtained per channel per earth scan at subpoint resolutions (the size of the pixel) approaching 1.1 kilometers. The resolution decreases away from the subpoint because of foreshortening and enlargement of the radiometers' field of view.

2.2 Data Calibration

No in-flight calibration of visible data (AVHRR channels 1-2) is performed; coefficients determined prior to launch, allow the conversion of raw digital data to percent albedo.

Brightness temperatures are determined using calibration coefficients, calculated in-flight, consisting of a view of space, having a near zero radiance, and a view of the radiometer housing — a designed blackbody. Four platinum resistance thermistors (PRT's) are built into the housing to measure its blackbody temperature (approximately 288 °K). The blackbody radiance and zero radiance of the space view allow a two-point calibration.

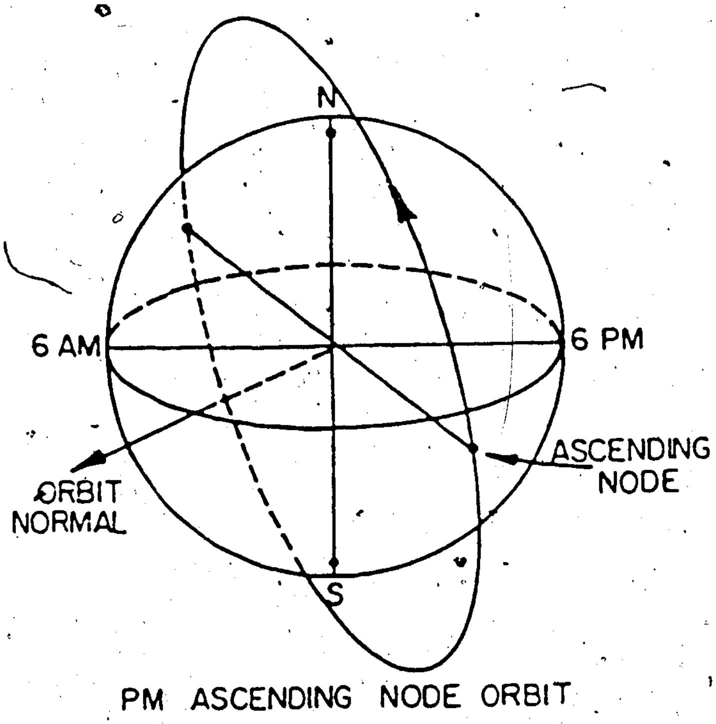


Figure 2.1 Geometry of a polar orbiting satellite on a "non-rotating" Earth. The orbital path of the satelllite is given by the large arrow. The direction of the radiometer scan is given by the small arrow. (after Reinelt et al., 1975).

The calibration procedure, for both the visible and infra-red channels, fully described by Lauritson et al. (1979), is not presented here. NOAA-9 and NOAA-10 calibration curves for channels 3 and 4 are given in Figure 2.2.

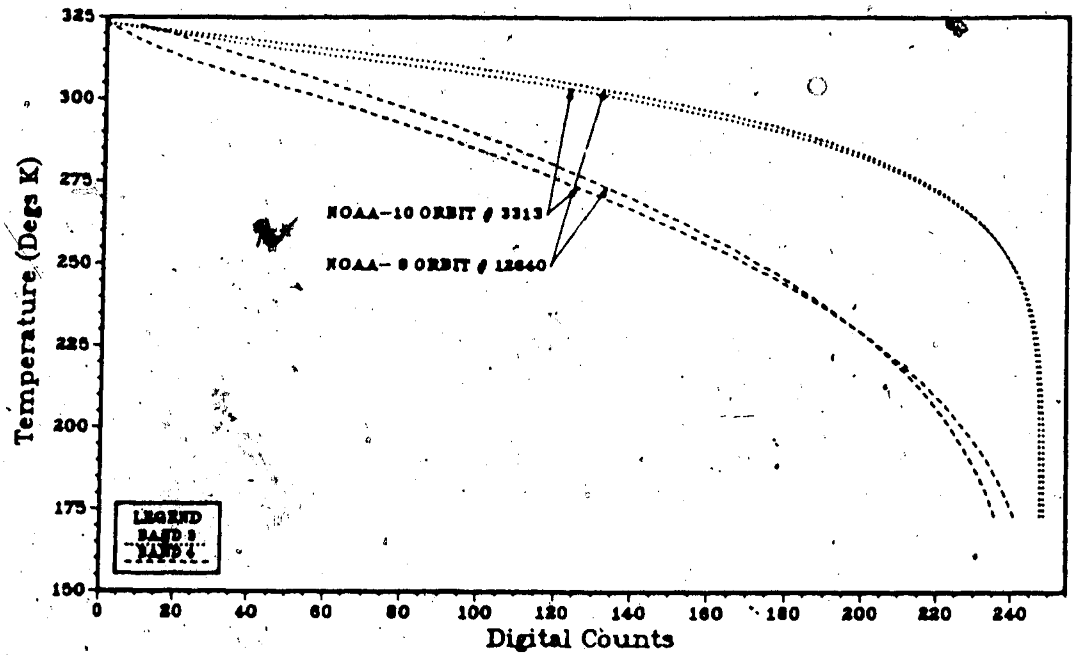


Figure 2.2 NOAA-9 and NOAA-10 infra-red calibration curves for Bands 3 and 4.

2.3 Band 3 Characteristics

Since AVHRR channel 3 is to play an important role in this study, some discussion of its spectral features is required. Although technically an infra-red channel, there is considerable reflection of incident solar radiation at channel 3 wavelengths, according to Bullas and Goodson, (1986). In the absence of solar radiation, data from band 3 can be interpreted in the same fashion as for band 4. The additional solar component causes the sensed object to appear warmer than it actually is. The degree of 'warming' depends on the albedo of the object. It is evident from Figure 2.2, that at temperatures greater than 0°K, small differences in temperature produce large changes in the measured digital data. The sensed Band 3 radiance of a pixel containing water can be markedly different from that of a pixel containing low cloud, such as stratus, despite that fact that the two surfaces may have very similar temperatures. The characteristics of Band 3 make it possible to distinguish areas of fog and low stratus from nearby stretches of cloud free ocean.

3. Data and Sample Area Selection

3.1 Cloud Identification Problems

Estimation errors are introduced by the Phulpin method because of the technique's inability to extract information about certain types of cloud. Thin cirrus, because of its transparency, exhibits highly variable reflectances and radiances; cumuliform clouds, lying to the rear of cold fronts (rear zone cloud) have variable thicknesses and sizes. Both produce non-normal histograms that are smeared across the field, making a Gaussian fit impossible. Clouds having subpixel sizes not resolved by the radiometer are not considered at all, despite their effects on the histogram field. Consider two pixels: one completely filled by cloud, having an albedo of 20-30%; the other a cloud-free ocean pixel, having an albedo of 1-2%. A small cloud filling 5% of the latter pixel will increase its albedo by 1-1.5% or 2-3.5 digital counts, for 8 bit AVHRR data (Bernstein, 1982). The same cloud will produce differences in brightness temperature, between channels 3 and 4, of 0.5-1.0 °K, or 1-2 digital counts (Maul and Sidran, 1973). Unresolved cloud clearly alters the sensed radiance and reflectivity on a scale sufficent to produce measurable changes in the shape and extent of a particular histogram.

The difficulties with cirrus can be removed if the automated technique is restricted to total cloud-cover estimation, regardless of the availability of information about individual cloud layers. Since the histograms within a given field are produced by either the sea or cloud, estimates can be made by removing the effects of the sea surface; the remaining points in the field must be produced by cloud. Low cloud can produce cloud peaks that are close to the sea-peak, creating a potential for confusion. Since surface and cirrus peaks should appear on opposite sides of the field, because of the disparity in their surface temperatures, there is little possibility of confusion.

Difficulties with rear zone and unresolved clouds can be removed by understanding the factors that influence the magnitude of the radiometer data and, to a certain extent, the histogram field. Analysis techniques relying on histograms depend on the relationship between individual field points, and not on their absolute magnitudes. As long as noise effects are uniform across the window area, the relationship between points is preserved. In an effort to ensure the uniformity of window noise, restrictions on the selection of sample areas, and minor corrections to the raw digital data have been imposed.

3.2 Data Correction and Window Selection

3.2.1 Visible Data

Spherical trigonometry (Smart, 1977) has been used to calculate local solar zenith angles; the following quantities are necessary: the latitude, ϕ, and longitude, λ, of the pixel, the solar declination, δ and right ascension, α, the total nutation in longitude, $\Delta\psi$, the obliquity of the ecliptic, ϵ, and the Greenwich mean sidereal time, GMST, of the observation.

Using equations from the scan by scan gridding procedure described by Reinelt et al (1975, pp.118-132), earth locations in scan line and pixel are converted to latitude and longitude. The remaining quantities were obtained from The Astronomical Almanac (1987).

These data are substituted into the version of the cosine formula given below:

$$\cos(z) = \sin(\phi)\sin(\delta) + \cos(\phi)\cos(\delta)\cos(H) \qquad 3.3.1$$

where H is the local solar hour angle which is given by:

$$H = \text{local apparent sidereal time} - \alpha \qquad 3.3.2$$

The local apparent sidereal time is given by:

$$LAST = GMST + \lambda + \Delta\psi\epsilon \qquad 3.3.3$$

The resulting local solar zenith angles are accurate to about 0.06 degrees.

The visible data (AVHRR channels 1-2) have been corrected so that local solar zenith angles are equal to zero across the window (a secant correction is applied; although this is not an exact correction, it does provide sufficient accuracy to make reflectance errors, produced by variable solar elevations, uniform across the window). Actual local solar zenith angles are limited to a maximum of 60°, ensuring stable sea surface reflectivities of 2-6%, regardless of the surface slope (Payne, 1972). Coastal waters are avoided because of increased reflectivities that are produced by suspended particulates, bottom reflection, and variable chlorophyll concentrations (Maul, 1981; Stewart, 1985). Measurements are not taken from the anti-sun side of large clouds because shadows often mask actual reflectivities within 1-2 pixels of the cloud (Bernstein, 1982).

3.2.2 Infra-red Data

According to Maul, (1981), reflection of incident solar radiation from the sea surface in the thermal infra-red is negligible, except in areas of specular reflection (sunglint). These areas are avoided by restricting the solar reflection angle, θ_r, to a maximum of 45° (Fig. 3.1). As stated previously, the effects of limb darkening and foreshortening on the data are relatively uniform for the range of cross-scan angles over which data are retained. Absorption and re-emission by water and carbon dioxide are assumed to be uniform for the small window areas selected, except in frontal areas, where sampling from two different air masses could take place. These areas have been avoided, removing any possible difficulties that might arise from rear zone clouds, although this is more a case of avoiding the problem than actually solving it.

3.2.3 Scattering

Molecular scattering is negligible at wavelengths above 1μm (Chahine et al, 1983), and despite the existence of a small contribution in band 2, no correction was applied. Scattering by atmospheric aerosols was assumed to be negligible when compared to the histogram partition size.

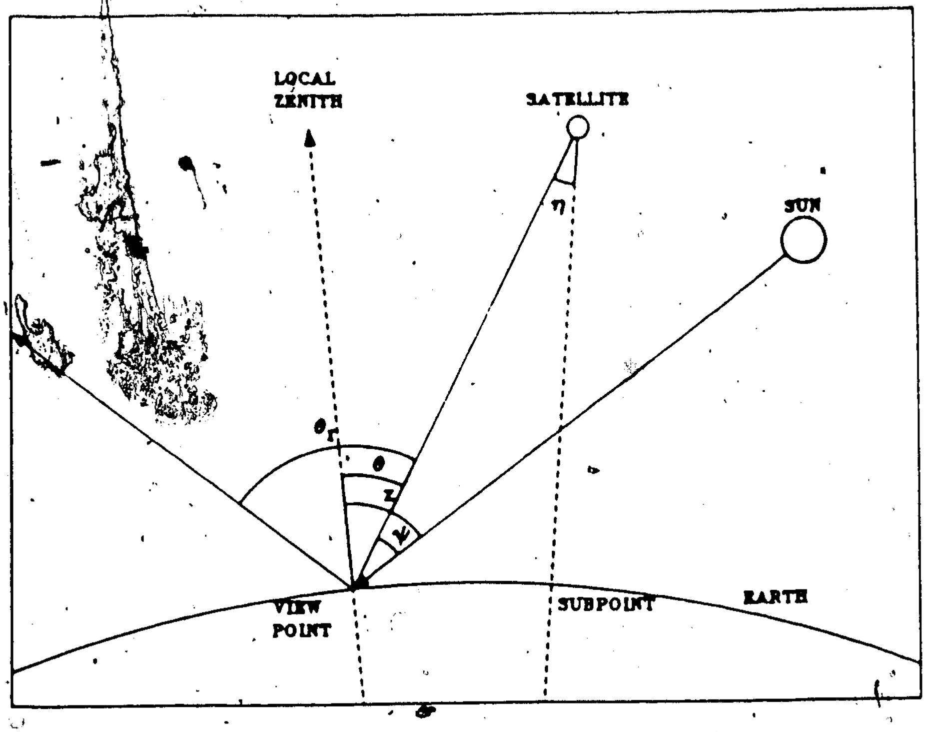


Figure 3.1 Satellite-pixel-sun geometry. Solar zenith angle, z, satellite zenith angle, θ, solar reflection angle, θ_r, scattering phase angle, ψ (not to be confused with the nutation in longitude, $\Delta\psi$, used in equation 3.3.3), and scan angle, η. After Bernstein, 1982.

4. Algorithm Development

4.1 Introduction

Forty sample windows were chosen from NOAA-9 orbit #13550. Latitudes and longitudes range from 32.5°N to 51.7°N and from 121.7°W to 133.2°W, respectively. The satellite crossed the equator at approximately 2339Z (Z = UTC, Coordinated Universal Time) on July 30, 1987 at a longitude of 122.2°W in a northbound trajectory that covered the west coast of North America and a section of the eastern Pacific Ocean.

The sample areas measured 40 scan lines by 40 pixels — large enough to provide data for the analysis, but small enough to limit the occurence of SST gradients in the field of view. The atmospheric abundances of water vapour and carbon dioxide across this small area are assumed to be constant, ensuring uniform absorption and re-emission. This size is also roughly equal to the sample size used by Phulpin et al., (1983). No data regarding coastal ocean conditions (ocean depth, salinity, plankton concentrations, etc.) were obtained prior to window selection, allowing the algorithm the freedom to estimate the total cloud cover without unnecessary bias.

Imagery for bands 2, 3, and 4 has been prepared showing the west coast of California in Plates 4.1 through 4.7, and the coastlines of Washington, and British Columbia, as far north as the Queen Charlotte Islands, in Plates 4.8 through 4.11. The images in Plates 4.1 and 4.8 are 2X enlargements showing the locations of the sample areas listed in Tables 4.1 and 4.2. The remaining photographs are 4X enlargements. The band 2 and band 4 images have been enhanced (Figure 4.1) to provide sufficient contrast between cloud and ocean surfaces to allow visual estimation of the total cloud-cover. No enhancement of the band 3 imagery has been applied because of its already high sensitivity to small differences in sensed radiation.

In the photographs, the lowest scan line and pixel are in the bottom right hand corner of the image; the maximum values are in the top left. The windows have been numbered in order of increasing scan line. The enlargement of the radiometer pixels away from the centre of the frame is evident.

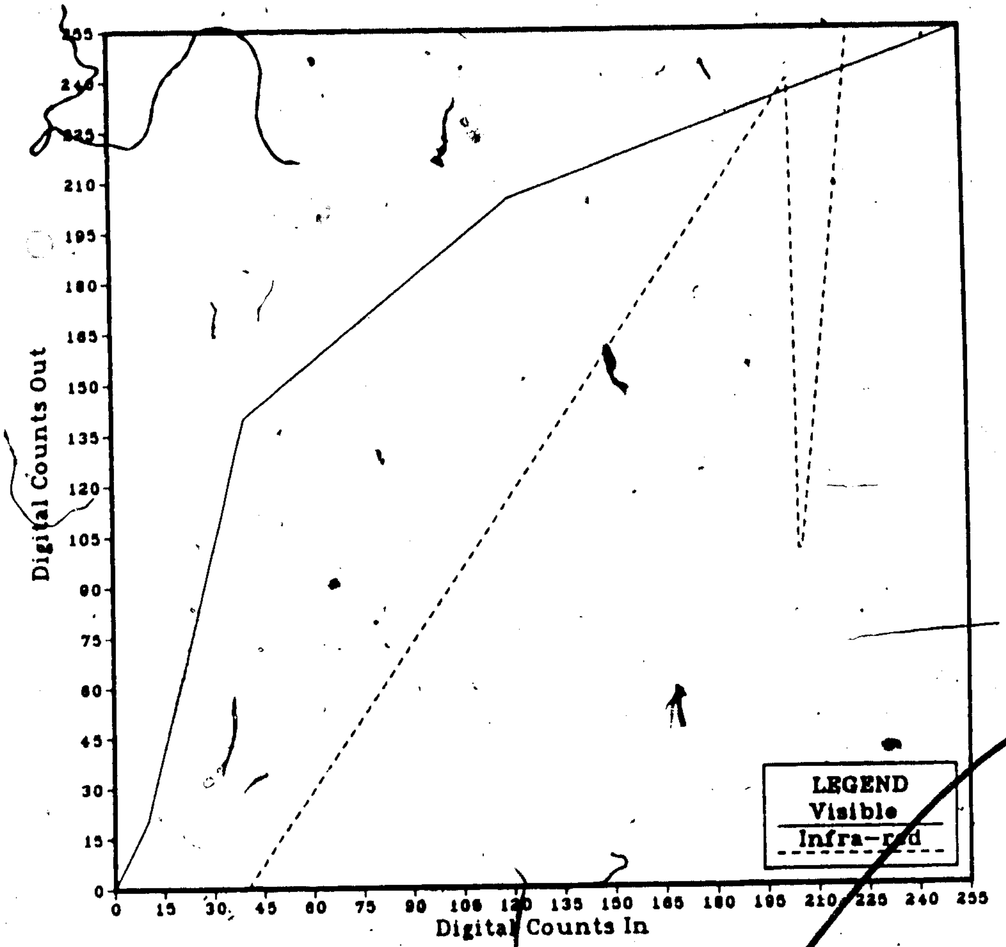


Figure 4.1 Enhancement curves for the visible and infra-red imagery in Plates 4.1-4.11. The values along the abscissa are the actual digital data retrieved by the satellite. The values along the ordinate axis are the same digital values transformed according to the enhancement curves. These values are used to map the digital data to grey scale during production of the imagery.

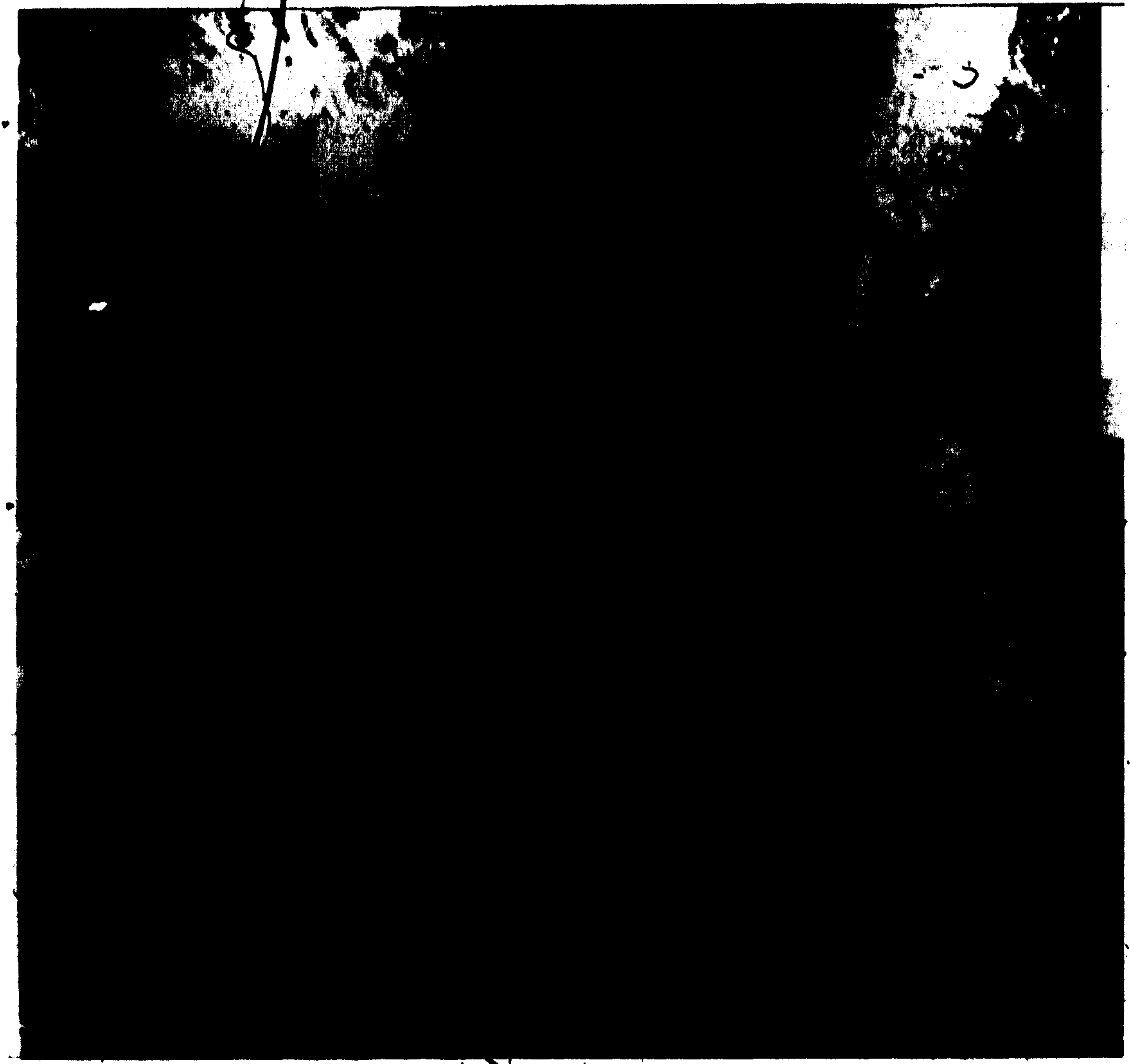


Plate 4.1 Band 2 image of the California coast, taken from NOAA-9 orbit #13550 showing the location of the 25 sample areas listed in Table 4.1. The image is enlarged 2X, and has been enhanced to increase the contrast between cloud and water. The scan line and pixel corresponding to the bottom right corner of the image are 350 and 128, respectively.

An outline of the cloud detection algorithm is presented in Appendix 2. Tables 4.3, 4.4, and 4.5 contain partial histogram fields for the (2,3), (2,4), and (3,4) channel sets, respectively. These data were extracted from a single sample window bounded by scan lines 500 and 539 and pixels 550 and 589 (see Tables 4.1 and 4.2). Figures 4.2, 4.3, and 4.4 are corresponding three-dimensional representations of the data from the tables. Determination of the total cloud-cover and its uncertainty is presented for this window so that the algorithm can be explained.

4.2 Data Partitioning

Prior to construction of the bi-spectral histogram field, frequency class intervals and their limits were determined. Individual pixel temperatures (bands 3 and 4) were rounded to the nearest degree, producing a 1°K class width. Dispersion of the peak along the visible axis, at low reflectances, was less than 1%. An albedo class width of 1% gave the sea-peak a one-dimensional appearance that allowed fitting with a Gaussian curve, rather than a surface. The lower class boundaries remained flexible so that rounding or truncation of the visible data would not produce artificial dispersion along the visible axis. For example, if, for a given sea-peak, a set of band 2 values ranging from 4.0% to 4.99% were obtained, simple truncation would produce an apparently one-dimensional peak (all values would be at 4% albedo). However, if the data ranged from 4.3% to 5.29%, truncation would produce a two-dimensional peak (values at 4% and 5%), even though the range of values was less than the 1% class size (5.29% - 4.3% = 0.99%). In order to overcome this difficulty, the minimum measured albedo is identified (4.3%) and a correction amounting to the truncation residual (0.3%) is subtracted from every albedo value in the field, reducing the sea-peak values to a range of 4.0-4.99%. Truncation can then take place without producing a component across the class boundary.

4.3 Sea Peak Identification

For histogram fields constructed using band 2 data, the sea-peak is assumed to display the lowest albedo in the 0-6% range (Payne, 1972; Phulpin et al, 1983). The total number of points within each reflectance class are added in this range; the sea-peak is assumed to be in the first class exhibiting a total greater than 3. The 3 point threshold is entirely arbitrary and is imposed to circumvent occasional anomalous points that cause the algorithm to mis-locate the sea-peak. For example, consider the case of a thin stratus layer lying close to the water,

Table 4.1 Total cloud-cover estimates and uncertainties for NOAA-9 orbit #13550, between 32.5° and 42.7° N, and 121.7° and 131.5°W. Since it is possible, in principle, to estimate the cloud-cover down to the pixel level, the uncertainty in the estimates was allowed 3 significant figures.

NOAA-9 Orbit # 13550
July 30, 1987 23:35:39Z

Window	Scan Line Range	Pixel Range	Cloud Cover
1	400-439	256-295	0.075±0.015
2	450-489	300-339	0.001±0.002
3	480-519	640-689	0.667±0.001
4	500-539	550-589	0.291±0.003
5	550-589	300-339	0.013±0.102
6	600-639	550-589	0.659±0.001
7	700-739	330-369	0.003±0.000
8	800-839	384-423	0.070±0.002
9	900-939	512-551	0.227±0.001
10	1000-1039	400-439	0.152±0.009
11	1000-1039	1010-1049	0.421±0.001
12	1050-1089	790-829	0.576±0.030
13	1070-1109	600-639	0.217±0.001
14	1150-1189	500-539	0.521±0.000
15	1150-1189	850-889	0.326±0.019
16	1200-1239	512-551	0.236±0.000
17	1200-1239	780-819	0.413±0.001
18	1250-1289	400-439	0.314±0.000
19	1300-1339	640-679	0.499±0.005
20	1300-1339	880-919	0.418±0.001
21	1350-1389	768-807	0.384±0.000
22	1400-1439	450-489	0.462±0.000
23	1400-1439	768-807	0.388±0.000
24	1500-1539	380-407	0.236±0.001
25	1520-1559	700-739	0.541±0.000

and covering 10% of the window. Assume that the sea and cloud peaks lie in adjacent reflectance classes (at 3 and 4%, respectively) and that a single anomalous point lies in the 2% reflectance class. A minimum class total of 3 points causes the algorithm to correctly identify the point in the 2% reflectance class as anomalous. The correct location of the sea-peak and a correct total cloud-cover estimate of 10% is then made. Without this minimum threshold value, the single point at 2% albedo is identified as the sea peak, and the total cloud-cover is estimated at 99.9%.

Table 4.2 Total cloud-cover estimates and uncertainties for NOAA-9 orbit #13550, between 46.1° and 51.7° N, and 128.8° and 133.2°W. Since it is possible, in principle, to measure the cloud-cover down to the pixel level, the uncertainty in the estimates was allowed 3 significant figures.

NOAA-9 Orbit # 13550
July 30, 1987 23:35:39Z

Window	Scan Line Range	Pixel Range	Cloud Cover
1	1840-1879	500-539	0.652±0.002
2	1900-1939	490-529	0.562±0.001
3	1900-1939	770-809	0.420±0.006
4	1950-1989	440-479	1.000±0.000
5	1950-1989	650-689	0.811±0.000
6	2000-2039	620-659	0.676±0.003
7	2000-2039	890-929	0.261±0.000
8	2050-2089	750-789	0.963±0.000
9	2100-2139	512-551	0.441±0.004
10	2100-2139	820-859	0.884±0.001
11	2200-2239	680-719	0.217±0.007
12	2200-2239	880-919	1.000±0.000
13	2250-2289	640-679	0.072±0.002
14	2300-2339	512-551	0.489±0.000
15	2400-2439	640-679	0.894±0.003

The sea-peak, for histogram fields using bands 3 and 4, corresponded to the local maximum exhibiting the highest band 4 temperature and the lowest band 3 temperature. The band 3 solar component, produced by reflection from the sea surface, was assumed to be small, causing little 'warming' of the band 3 temperature. Consequently, the temperatures measured by bands 3 and 4 were assumed to be within ±1°K, for the sea-peak.

4.4 Fitting of Curves and Surfaces to the Sea Peak

Total cloud-cover estimates were obtained by fitting a curve, in the case of the (2,3) and (2,4) channel sets, or a surface, in the case of the (3,4) channel set, to the sea-peak, and removing it. The one-dimensional curve is of the form:

$$f = \frac{N}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-\bar{x})^2}{2\sigma_x^2}} \qquad 4.4.1$$

where f is the predicted frequency, $\bar{x}$ is the mean value, σ_x^2, is the variance, and the predicted maximum (central) frequency of the Gaussian, f_{max}, is given by:

$$f_{max} = \frac{N}{\sqrt{2\pi\sigma_x^2}} \tag{4.4.2}$$

where N is the total number of pixels within the Gaussian limits. The two-dimensional surface is described by:

$$f = \frac{N}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}}\, e^{-\frac{Q^2}{2(1-r^2)}} \tag{4.4.3}$$

where

$$Q^2 = \frac{(x-\bar{x})^2}{\sigma_x^2} - \frac{2r(x-\bar{x})(y-\bar{y})}{\sigma_x\sigma_y} + \frac{(y-\bar{y})^2}{\sigma_y^2} \tag{4.4.4}$$

where f is the predicted frequency, $\bar{x}$, $\bar{y}$ are the mean values, σ_x^2, σ_y^2 are the variances, r is the correlation coefficient between the variates x and y, and the predicted maximum (central) frequency of the surface, f_{max}, is given by:

$$f_{max} = \frac{N}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} \tag{4.4.5}$$

Clearly, for the one-dimensional case, the mean, variance, and central curve frequency must be predicted, while for the two-dimensional case, the means, variances, correlation coefficient, and central surface frequency must be estimated.

Unlike Phulpin et al., (1983), the means and variances (and in the case of fields using bands 3 and 4, the correlation coefficient) of the sea cluster were free parameters, even though this increased the computation time; the sacrifice of computational speed was offset by the increased accuracy of the fit.

Surface gradients within the selection window produce additional peaks that are often superimposed, masking the actual parameters of each peak. In principle, it is possible to find many peaks clustered together, but the selected sample area was small enough so that the maximum number of close proximity peaks that was encountered was two. In cases of superposition, the variance of the primary peak is generally overestimated, while its central frequency is underestimated. Parameters for the secondary peak are nearly impossible to determine until extraction of the primary peak has been completed, although it should exhibit a variance comparable to that of the primary peak. To facilitate peak by peak extraction, only those bins exhibiting no evidence of superposition were used to determine the parameters of the 'best-fit' Gaussian. Criteria have been established to identify possible cases of this. All points along each histogram axis are expected to decrease uniformly away from the maximum; even slight point to point increases are an indication of the presence of an additional peak or multiple peaks. The actual central frequency of the 'best-fit' Gaussian was assumed to be greater than the maximum data value except where the actual peak center coincided with a class boundary. For this reason, estimates of the maximum cluster frequency were restricted to values greater than the measured value. Low estimates of this quantity cause the algorithm to compare values at the axis boundaries; the lowest value is omitted, and the Gaussian parameters are re-estimated. This process continues until a satisfactory estimate of the central frequency is obtained. As a final check on the success of the extraction, it should be noted that typical sea-peak variances range between 0.25 and 0.35.

Two methods were used to determine the Gaussian parameters of a given sea-peak: direct calculation and a modified least squares procedure (Appendix 1). Bins having unit frequencies do not contribute to the parameter estimation when the method of least squares is used, because of its logarithmic nature. So that unnecessary computation could be minimized, the lowest allowable bin frequency was greater than 1. Fitting by both methods was attempted; the method having the lowest residual mean sum of squares was selected to perform the extraction.

4.4.1 (2,3) Band Set Histograms

In Table 4.3, the sea-peak lies in the 3% reflectance class between 289°K and 293°K. Those points outside of this class are attributed to cloud; the ratio of these points to the total number in the field gives an initial estimate of the total cloud-cover (in this case, 466/1600 = 29.1%). The measured central frequency of the peak (at 290°K) is 745; this is the minimum allowed value for the predicted central frequency; anything less is rejected ~~by the~~ algorithm on

Table 4.3 Partial histogram field for the (2,3) band set, taken from NOAA-9 orbit 13550, between scan lines 500-539 and pixels 550-589.

Band 3 Temperature (°K)	Albedo (%) 2	3	4	5	6	7	8
289	0	172	0	0	0	0	0
290	0	745	1	0	0	0	0
291	0	153	3	0	0	0	0
292	0	51	22	3	0	0	0
293	0	12	29	1	1	2	0
194	0	1	28	5	3	1	0
295	0	0	18	8	3	4	3
296	0	0	2	8	2	4	3
297	0	0	1	3	3	3	4
298	0	0	0	1	5	3	3
Class Totals	0	1134	104	29	17	17	13

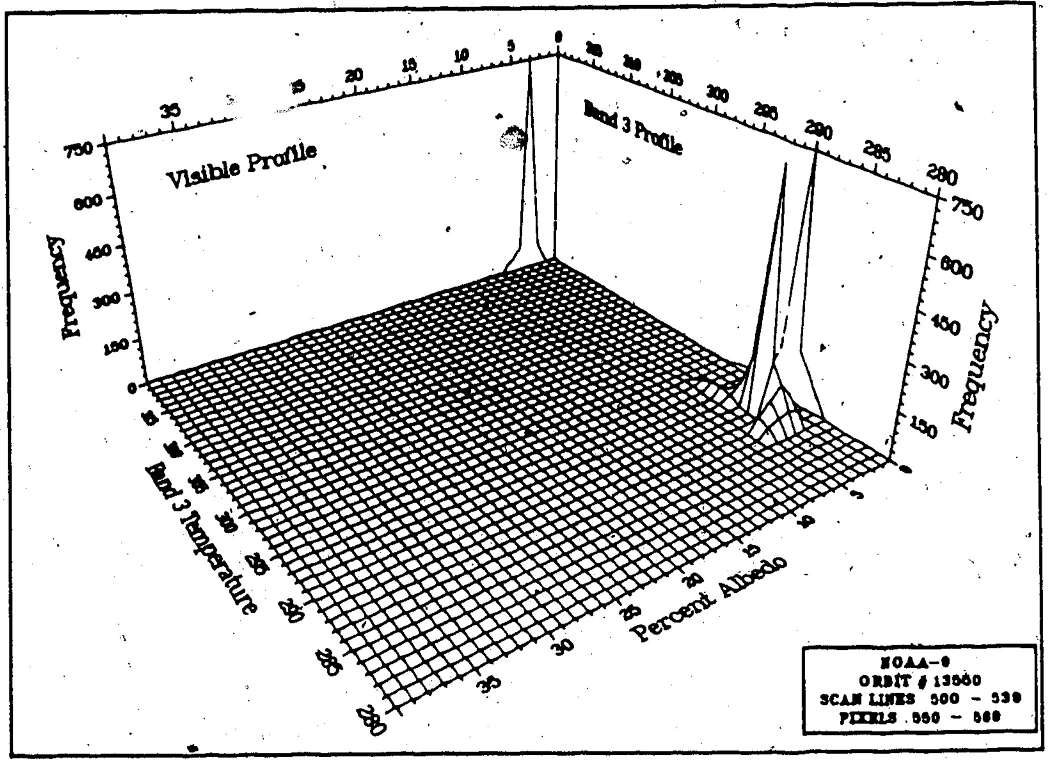


Figure 4.2 Histogram field for the (2,3) band set, taken from NOAA-9 orbit 13550, between scan lines 500-539 and pixels 550-589.

the basis of gradient contamination.

Table 4.4 Extraction of the sea-peak for the (2,3) band set from Table 4.3. The first extraction is performed by the method of least squares, the second, by direct calculation. Note the asymmetric residuals produced by the first extraction.

Parameters for 1st Extraction	Band 3 Temperature (°K)	Initial Data	Residuals
	289	172	0.0000
Mean = 289.98°K	290	745	0.0022
	291	153	0.0140
Variance = 0.328	292	51	49.5105
	293	12	11.9993
Central Freq. = 775.42	294	1	1.0000
	Class Totals	1134	62.52

Parameters for 2nd Extraction	Band 3 Temperature (°K)	Residuals from 1st Extraction	Residuals from 2nd Extraction
	289	0.0000	0.0000
Mean = 292.20°K	290	0.0022	0.0022
	291	0.0140	0.0000
Variance = 0.157	292	49.5105	0.0000
	293	11.9993	4.1312
Central Freq. = 61.92	294	1.0000	0.9881
	Class Totals	62.52	5.17

Both fitting techniques initially yield low estimates of the central frequency when all five points are used, indicating the presence of gradient contamination. This is supported by visual inspection which shows a marked broadening of the peak toward "warm" temperatures. According to the criteria established for dealing with this situation, the frequencies at 289°K and 293°K were compared and the lowest frequency (12) dropped. The parameters were then recalculated. For this particular case, the process was repeated until the range of the histogram was reduced to between 289°K and 291°K. The mean, variance, and central frequency, determined by direct calculation, were 289.98°K, 0.30, and 774.94, respectively (Table 4.5). For the method of least squares, the estimated values were 289.98°K, 0.33, and 775.42 (Table 4.4). Based on the residual mean sum of squares, the method of least squares was chosen by the algorithm to perform the first extraction. The least squares parameters were substituted into equation 4.4.1 and the predicted frequencies subtracted from the actual values at every point in the 3% reflectance class. Negative values were assumed to be equal to zero. The residuals show the presence of a secondary peak between 292°K and 293°K that is very narrow and asymmetric. A second extraction was performed by direct calculation (the method of least squares requires a minimum of 3 points to fit the curve). Note that the variance of the secondary peak (0.157) differs greatly from that of the primary (0.33). The residuals from the second extraction amount to only 5.17, or in terms of the total number of field points, 0.3%. The extraction process from beginning to end can be seen in Table 4.4.

It is interesting to note that had direct calculation been chosen to perform the first extraction, a symmetric secondary peak would have been exposed, having a variance (0.33) close to that of the primary (0.30) (Table 4.5). This suggests that direct calculation might have been a better choice to perform the first extraction, and points out one of the basic flaws of the algorithm: the fitting technique that is selected provides the 'best' fit to the data, not necessarily the most 'appropriate' fit.

The inclusion of the least squares technique might, at first glance, appear to be redundant, given the excellent results produced by direct calculation; however, it should be noted that the sea-peak generally exhibits truncation on the warm side to varying degrees, and direct calculation often does not perform well, while least squares does.

4.4.2 (2,4) Band Set Histograms

The histograms produced using a combination of bands 2 and 4 can be seen in Table 4.6 and Figure 4.3. The reflectance class totals are virtually the same as those for the previous example; any differences arise because the entire histogram field is presented. The sea-peak is

Table 4.5 Extraction of the sea-peak for the (2,3) band set from Table 4.3. Both extraction are performed by direct calculation. Note that the residuals produced by the first extraction are now symmetric.

Parameters for 1st Extraction	Band 3 Temperature (°K)	Initial Data	Residuals
	289	172	13.9446
Mean = 289.98°K	290	745	0.0000
	291	153	12.4436
Variance = 0.303	292	51	50.0554
	293	12	12.0000
Central Freq. = 774.94	294	1	1.0000
	Class Totals	1134	89.44

Parameters for 2nd Extraction	Band 3 Temperature (°K)	Residuals from 1st Extraction	Residuals from 2nd Extraction
	289	13.9446	13.9446
Mean = 291.99°K	290	0.0000	0.0000
	291	12.4436	0.9360
Variance = 0.328	292	50.0554	0.0000
	293	12.0000	0.9024
Central Freq. = 51.89	294	1.0000	0.8875
	Class Totals	89.44	16.67

Table 4.6 Partial histogram field for the (2,4) band set, taken from NOAA-9 orbit 13550, between scan lines 500-539 and pixels 550-589.

Band 4 Temperature (°K)	Albedo (%) 2	3	4	5	6	7	8
281	0	0	0	0	0	0	0
282	0	0	0	0	0	0	0
283	0	0	0	0	0	0	0
284	0	0	0	0	0	0	0
285	0	0	0	0	0	0	0
286	0	0	0	0	0	0	1
287	0	0	0	1	5	18	8
288	0	0	20	22	17	9	5
289	0	977	84	6	0	0	0
290	0	157	0	0	0	0	0
Class Totals	0	1134	104	29	23	27	13

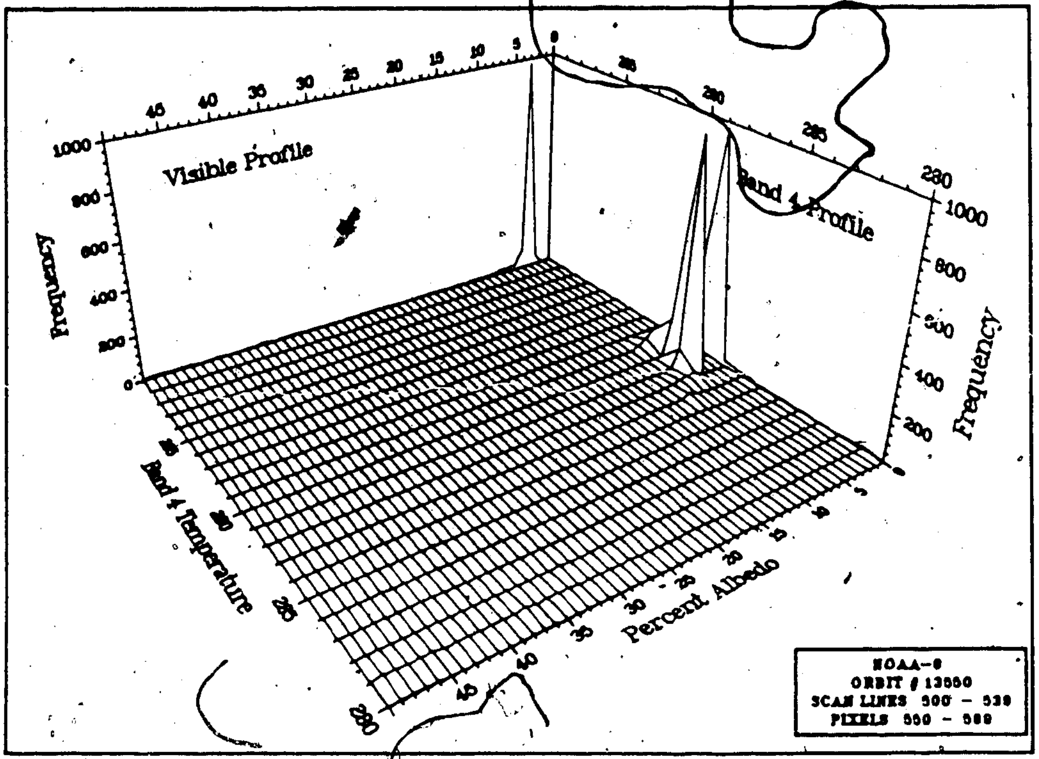


Figure 4.3 Histogram field for the (2,4) band set, taken from NOAA-9 orbit 13550, between scan lines 500-539 and pixels 550-589.

located at temperatures between 289°K and 290°K. The difference in dispersion between the band 4 peak and the band 3 peak of the previous example, is a result of the high sensitivity of band 3; detection of gradient contamination is impossible in this case. The initial cloud-cover estimate is the same as the previous one: 29.1%.

Table 4.7 Extraction of the sea-peak for the (2,4) band set from Table 4.4. The extraction is performed by direct calculation. Note the large residual impulse at 290°K.

Parameters for 1st Extraction	Band 4 Temperature (°K)	Initial Data	Residuals
Mean = 289.14°K	286	0	0.0000
	287	0	0.0000
Variance = 0.119	288	0	0.0000
	289	977	0.0000
Central Freq. = 775.42	290	157	98.6711
	Class Totals	1134	98.6711

Direct calculation of the Gaussian parameters yields initial estimates for the mean, variance, and maximum of 289.14°K, 0.12, and 1309.79, respectively. When substituted into 4.4.1, the resulting best-fit curve drastically overestimates the maximum value, and underestimates the other, causing a large impulse (98.67, or 6.2%) to be left at 290°K. The main reason for the poor fit would seem to be the variance, σ^2, which lies well outside the typical range of sea-peak values. No attempt is made to fit the curve by the method of least squares because there are insufficient data points. The extraction results can be seen in Table 4.7.

4.4.3 (3,4) Band Set Histograms

The histograms produced by the combination of bands 3 and 4 are presented in Table 4.8 and Figure 4.4. The narrow band 4 dispersion again causes analysis difficulties; no least squares fit is attempted because of this. The sea-peak lies at band 4 temperatures between

Table 4.8 Partial histogram field for the (3,4) band set, taken from NOAA-9 orbit 13550, between scan lines 500-539 and pixels 550-589.

Band 4 Temperature (°K)	Band 3 Temperature (°K) 289	290	291	292	293	294	295
281	0	0	0	0	0	0	0
282	0	0	0	0	0	0	0
283	0	0	0	0	0	0	0
284	0	0	0	0	0	0	0
285	0	0	0	0	0	0	0
286	0	0	0	0	0	0	1
287	0	0	0	0	0	1	1
288	0	0	1	2	11	11	20
289	164	604	148	74	34	26	15
290	8	143	7	0	0	0	0
Class Totals	172	747	156	76	45	38	37

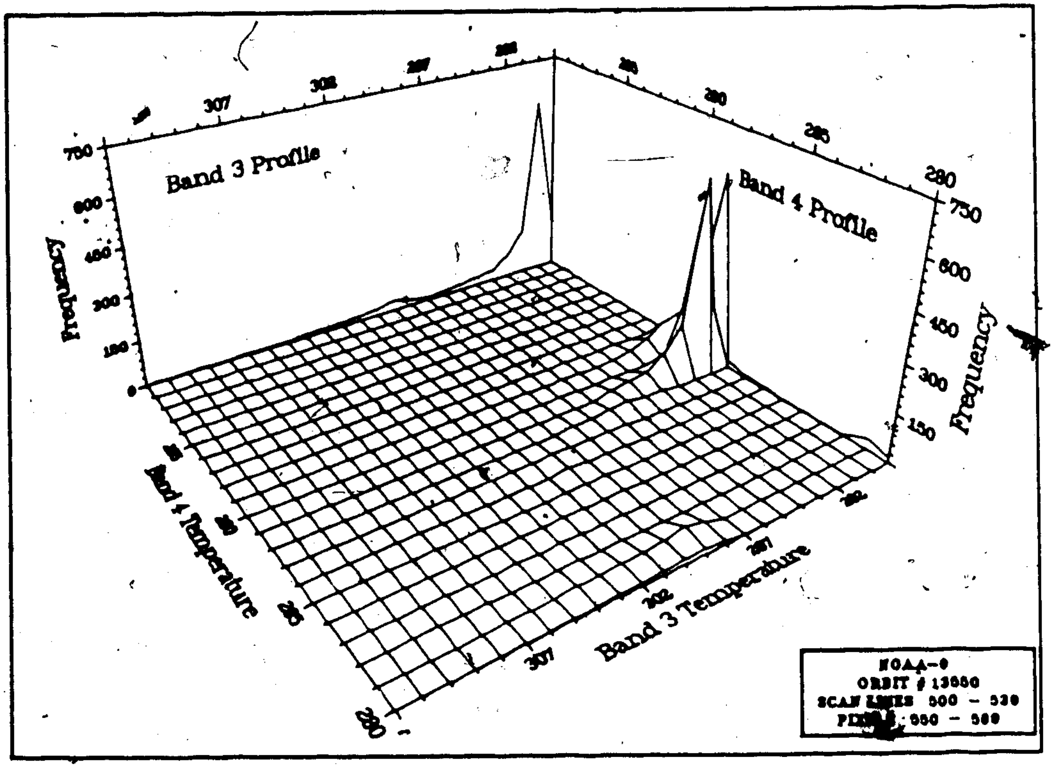


Figure 4.4 Histogram field for the (3,4) band set, taken from NOAA-9 orbit 13550, between scan lines 500-539 and pixels 550-589.

289°K and 290°K and at band 3 temperatures between 289°K and 296°K (not shown in Table 4.8). The initial cloud-cover estimate is 23.4% — slightly higher than the previous two values, but certainly a representative value. Initial estimates of the band 3 and band 4 means and variances, the correlation coefficient, and the maximum frequency are 290.20°K and 289.13°K, 0.74 and 0.12, -0.089, and 642.13, respectively. These parameters are substituted into equation 4.4.3. Extraction of the sea-peak can be seen in Table 4.9. The extraction produces an overestimate of some of the points and an underestimate of others, leaving several discrete impulses which the algorithm is unable to extract. The uncertainty in the estimate is ±14%, which is rather high, but it encompasses the estimates from the previous two examples. The main difficulty with the extraction seems to lie with the direct calculation method, which had to be relied on because of the narrow band 4 dispersion. The least squares fitting technique provides better estimates of the surface peak parameters, but three points in each direction are required to provide unambiguous results.

It should be noted that the limits of the sea-peaks are determined by scanning bin frequencies along the two axes only; in most cases this is sufficient to produce cloud-cover estimates that are comparable to those from the (2,3) and (2,4) channel sets. In several instances; however, thin stratus layers have produced cloud peaks well within the limits of the sea-peak. For this reason, off-axis bins are tested for abnormality by comparing their frequencies with those of surrounding bins and applying a distance correction. When such an anomaly is discerned, an attempt is made by the algorithm to reduce the limits of the sea peak to an acceptable range; failing this, a flag is set, and the operator is notified that extraction of the (3,4) surface is impossible.

4.5 Final Cloud Cover Estimates

No attempt was made to compare and contrast the accuracy of the estimates produced by the three channel sets since reliable surafce estimates could not be ensured. Rather, the estimate having the lowest uncertainty was selected as being the most reliable.

Table 4.9 Extraction of the sea-peak for the (3,4) band set from Table 4.5. The extraction is performed by direct calculation. Note the large residual peak left at band 3 temperature 291°K. Being one-dimensional, no extraction of this feature could be made.

Parameters for 1st Extraction	Band 4 Temperature (°K)	Band 3 Temperature (°K) 289	290	291	292	293	294	295
	281	0	0	0	0	0	0	0
B3	282	0	0	0	0	0	0	0
Mean = 290.19°K								
B4	283	0	0	0	0	0	0	0
Mean = 289.13°K								
B3 Var. = 0.747	284	0	0	0	0	0	0	0
B4 Var. = 0.116	285	0	0	0	0	0	0	0
Correlation	286	0	0	0	0	0	0	1
Coef. = -0.089	287	0	0	0	0	0	1	1
Central	288	0	0	1	2	11	11	20
Freq. = 642.13	289	164	604	148	74	34	26	15
	290	8	143	7	0	0	0	0
	Class Totals	172	747	156	76	45	38	37

Residuals after 1st Extraction	Band 4 Temperature (°K)	Band 3 Temperature (°K) 289	290	291	292	293	294	295
	281	0	0	0	0	0	0	0
	282	0	0	0	0	0	0	0
	283	0	0	0	0	0	0	0
	284	0	0	0	0	0	0	0
	285	0	0	0	0	0	0	0
	286	0	0	0	0	0	0	1
	287	0	0	0	0	0	1	1
	288	0	0	0	1.50	10.96	11	20
	289	0	29.92	0	2.62	30.67	25.96	15
	290	0	117.87	0	0	0	0	0
	Class Totals	0	147.79	0	4.12	41.63	37.96	37

The cloud-cover estimates in Table 4.1 and 4.2 will be compared to visual estimates taken from Plates 4.1 through 4.11. Possible causes of estimation differences will be discussed. In certain cases, windows exhibiting a common difficulty will be discussed together.

4.6 Case Studies

4.6.1 Introduction

The atmosphere in the vicinity of the sample areas is dominated by regions of both cumulus and stratocumulus. The band 4 images (Plates 4.4, 4.7, and 4.11) show little contrast between the ocean surface and the cloud tops, suggesting low level cloud. Upon closer examination, Plate 4.4 also shows SST gradients in the vicinity of the coastline (San Fransisco Bay is in the top right corner of the image). The band 2 images (Plates 4.1, 4.2, and 4.5) show a number of large cells that are probably stratocumulus, judging by their brightness and cloud-form texture. Also present are large patches of indistinct cloud at scales close to that of the pixel. The extent of these areas is difficult to determine. The band 3 images (Plates 4.3, 4.6, and 4.10) show a marked contrast between the water and cloud; the ocean surface is white, as would be expected, and the clouds range in shade from black to light-grey — indicative of stratiform cloud. The texture of the cloud tops suggests that they are convective.

The northern images also contain areas of mid and high level cloud (Plates 4.8-4.11). Cyclogenesis is taking place in the center of the image where a small comma cloud may be seen.

4.6.2 Plates 4.2-4.4

Examination of the satellite imagery in Plates 4.2-4.4, shows five sample areas within 250 kilometers of the California coast (numbers 1,2,5, 7, and 8 in Table 4.1), that exhibit no evidence of cloud-cover. Results from the algorithm, however, give the total cloud-cover at between 0.1% and 7.5%. These windows lie in a region having variable concentrations of both phytoplankton and zooplankton (Walton Smith and Kalber, 1974). These bioforms contain chlorophyll active photosynthesis, a process that is driven by absorption of solar radiation. Some of this absorption takes takes place in band 2, where there is an absorption peak at 0.675μm (Robinson, 1985). The phytoplankton also scatter incident solar radiation in much the same way as atmospheric aerosols. For the most part, this scattering is constant, except in

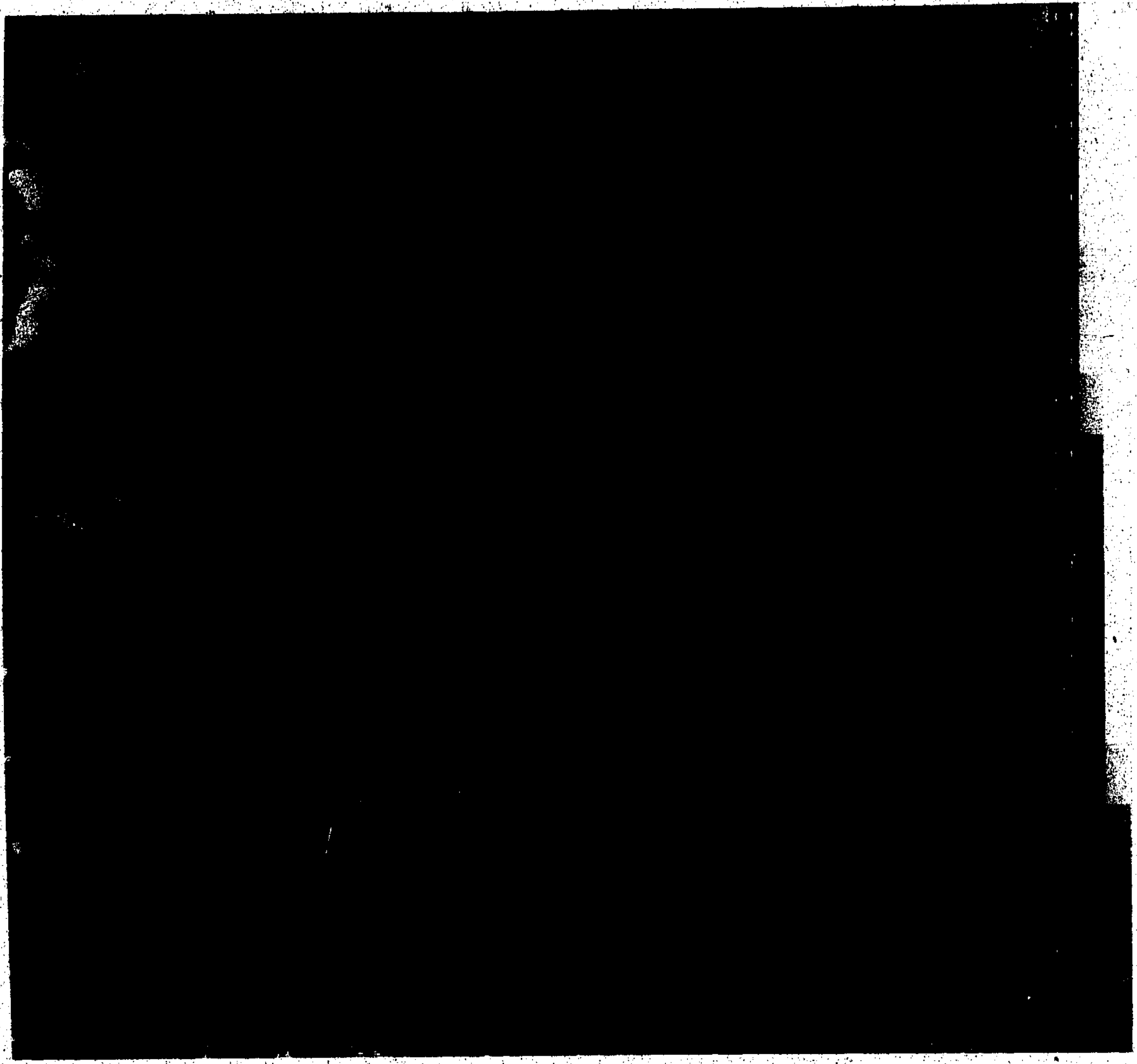


Plate 4.2 Band 2 image of the California coast, taken from NOAA-9 orbit #13550. The image is enlarged 4X, and has been enhanced to increase the contrast between cloud and water. The first nine sample areas in Table 4.1 are shown.

the vicinity of the absorption peak (0.675μm), where a slight decrease occurs (Robinson, 1985). The most likely cause of these slightly high cloud-cover estimates, is dispersion of the sea-peak, along the visible axis, that is greater than the 1% class width. This could be produced by ocean surface plankton variability. The presence of the SST gradients in the vicinity adds credence to this view. Stewart, (1985), comments on the strong correlation between SST gradients and chlorophyll concentrations along the California coast. The fact that the sample areas fall in an area of SST variability could be an indication of the presence of variable plankton and chlorophyll concentrations near the ocean surface.

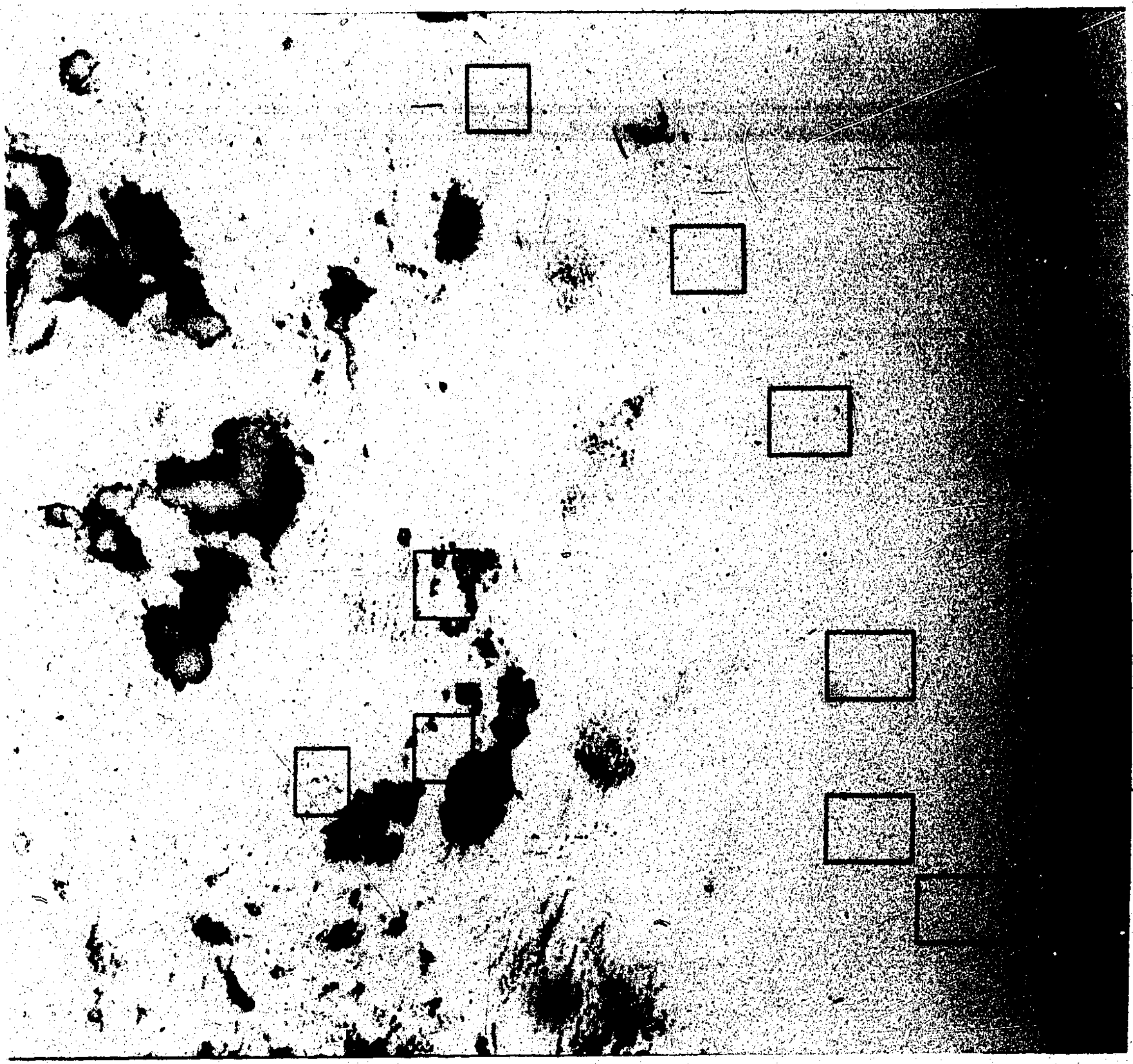


Plate 4.3 Band 3 image of the California coast, taken from NOAA-9 orbit #13550. The image is enlarged 4X. No enhancement has been applied. Cloud ranges from black (stratus/fog) to white (cirrus). The ocean surface is white. The first nine samples areas Table 4.1 are shown.

Windows 3, 4, and 6 in Table 4.1 are in an area containing several large stratocumulus cells. At first glance, the estimates of 66.7%, 29.1%, and 65.9% would seem to be high; however, close scrutiny of the band 2 and band 3 images indicates the presence of a considerable amount of the small, pixel-sized, cumulus. These clouds are spread extensively through sample window 3 and 6 and, to a lesser degree, through window 4. The algorithm estimates do not seem unreasonable in light of this.

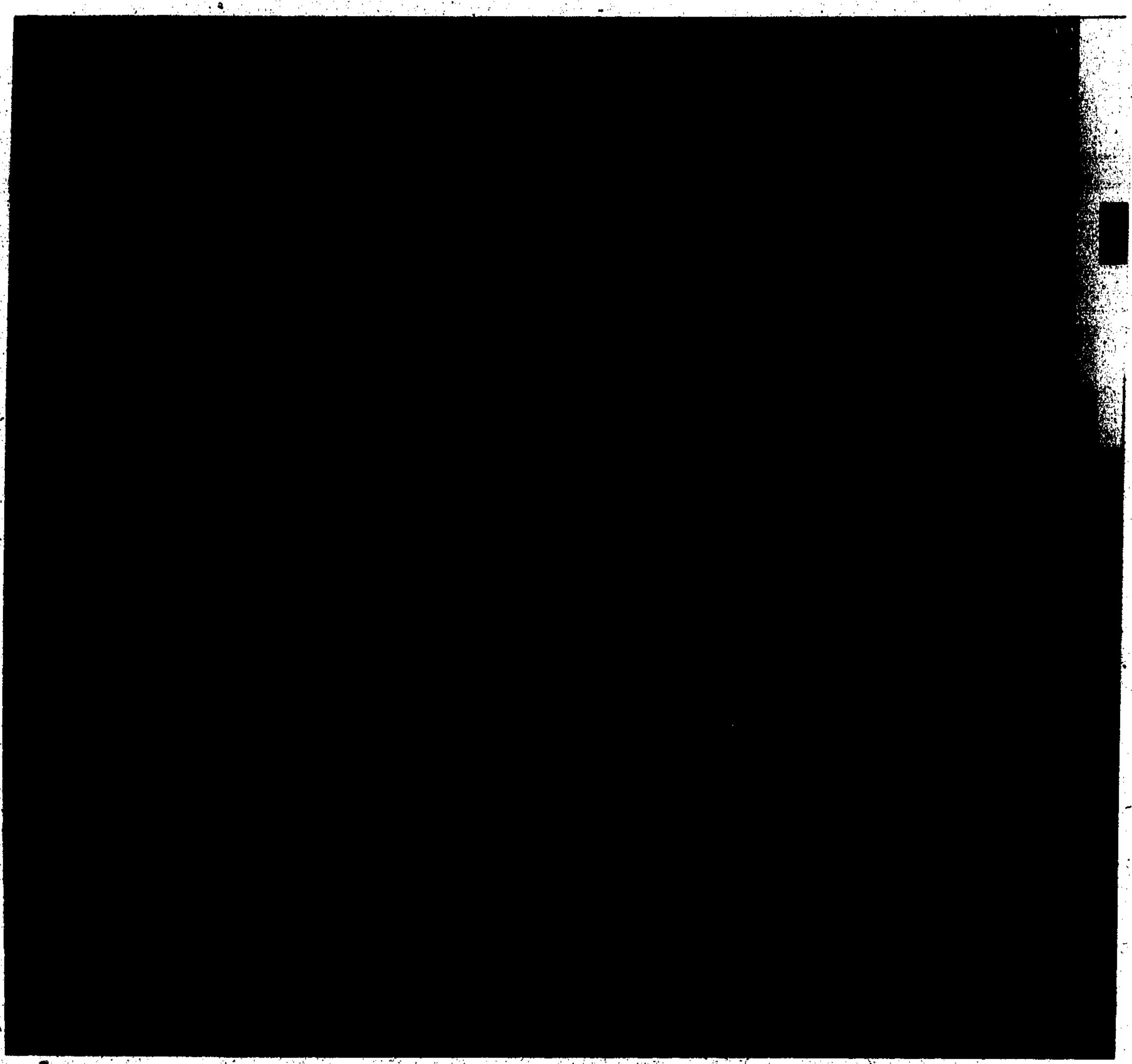


Plate 4.4 Band 4 image of the California coast, taken from NOAA-9 orbit #13550. The image is enlarged 4X, and has been enhanced to contrast cloud and water surfaces. Cloud ranges from grey (stratus/fog) to white (cirrus). The ocean surface is black. The first nine sample areas in Table 4.1 are shown.

The cloud-cover estimate for window 9 (the top sample in Plate 4.2-4.4) also appears reasonable, even though little or no evidence for the presence of cloud is seen in Plates 4.2 and 4.4. The band 3 image (Plate 4.3) does show some indications of the presence of the small cumuliform cloud mentioned earlier.

Plate 4.5 Band ~~2~~ image of the California coast, taken from NOAA-9 orbit #13550. The image is enlarged 4X, and has been enhanced to increase the contrast between cloud and water. Samples 10-25 from Table 4.1 are shown.

4.6.3 Plates 4.5-4.7

For the most part, the total cloud-cover estimates for sample areas 10-25 agree with visual inspection of the imagery in Plates 4.5 through 4.7, with a couple of notable exceptions.

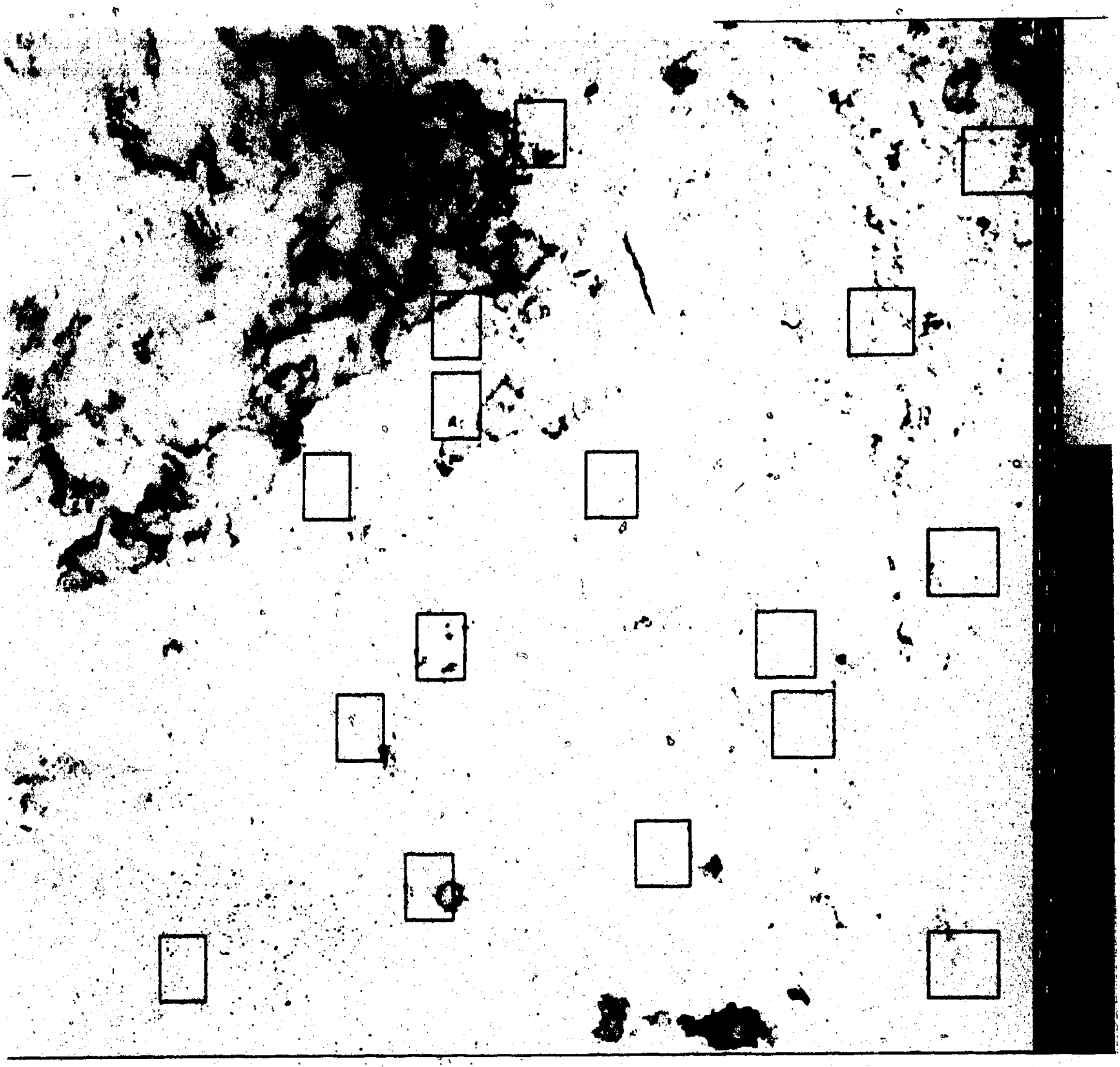


Plate 4.6 Band 3 image of the California coast ,taken from NOAA-9 orbit #13550. The image is enlarged 4X. No enhancement has been applied. Cloud ranges from black (stratus/fog) to white (cirrus). The ocean surface is white. Samples 10-25 from Table 4.1 are shown.

The cloud-cover estimate of 15.2% for sample area 10, appears to be low. Visual inspection of the band 2 and band 3 images suggests a total cloud-cover closer to 30%. Once again, there is evidence of SST gradients in the sample area (Plate 4.7), but it is doubtful that these would account for the low cloud-cover estimate. A possible shortfall in the estimate could be produced, if the associated cloud peak were to be dispersed beyond its lower class boundary, into the reflectance class containing the sea-peak. This effect could be caused by the presence of sub-pixel sized cloud in the window. These were shown in Chapter 3 to produce noticeable changes in the sensed temperature and albedo of a given pixel. If a cloud were to have a peak in the reflectance class adjacent to the one containing the sea-peak, and if it did not entirely fill the pixel, the contribution of the sea surface to the measured pixel reflectance could produce lowered albedos. For example, if the cloud had an albedo of 4%, and if it filled half of the pixel, the reflectance contribution of the cloud to the total reflectance of the pixel would be 2%. At the same time, the sea surface contribution would amount to 1.5%, if the reflectance of the sea were 3%. The combined reflectance would be 3.5%. If simple truncation were used to partition the data, this pixel would fall in the 3% reflectance class, along with the sea-peak. The algorithm would then consider it a part of the sea surface, and a lower cloud-cover estimate would result.

4.6.4 Plates 4.9-4.11

Total cloud-cover estimates from Table 4.2 agree very well with visual inspection of the imagery in Plates 4.9, 4.10, and 4.11, except for samples 4,5, and 12, whose estimates appear to be very high. Estimates for samples 7 and 13 yield surprisingly good results. Both areas contain small regions of very tenous stratiform cloud and the algorithm shows its accuracy by retrieving representative estimates for its extent.

Areas 4,5, and 12 have total cloud-cover estimates that range from 81.1% to 100%. The visible imagery (Plate 4.9) does not indicate cloud-covers that come close to these values. The infra-red photographs (Plates 4.10 and 4.11) however, do show the presence of a thermally indistinct layer close to the sea surface.

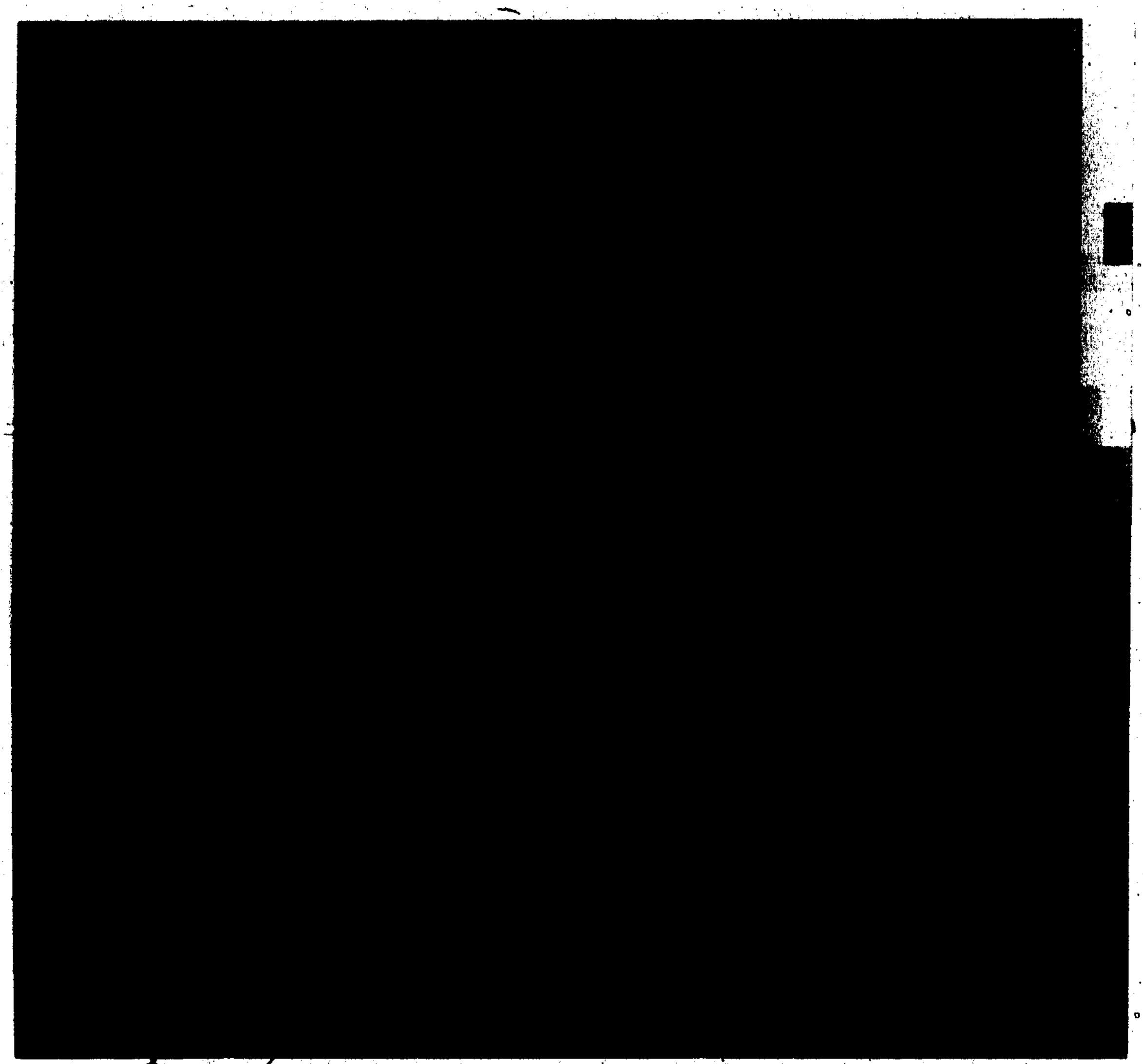


Plate 4.7 Band 4 image of the California coast, taken from NOAA-9 orbit #13550. The image is enlarged 4X, and has been enhanced to contrast cloud and water surfaces. Cloud ranges from grey (stratus/fog) to white (cirrus). The ocean surface is black. Samples 10-25 from Table 4.1 are shown.

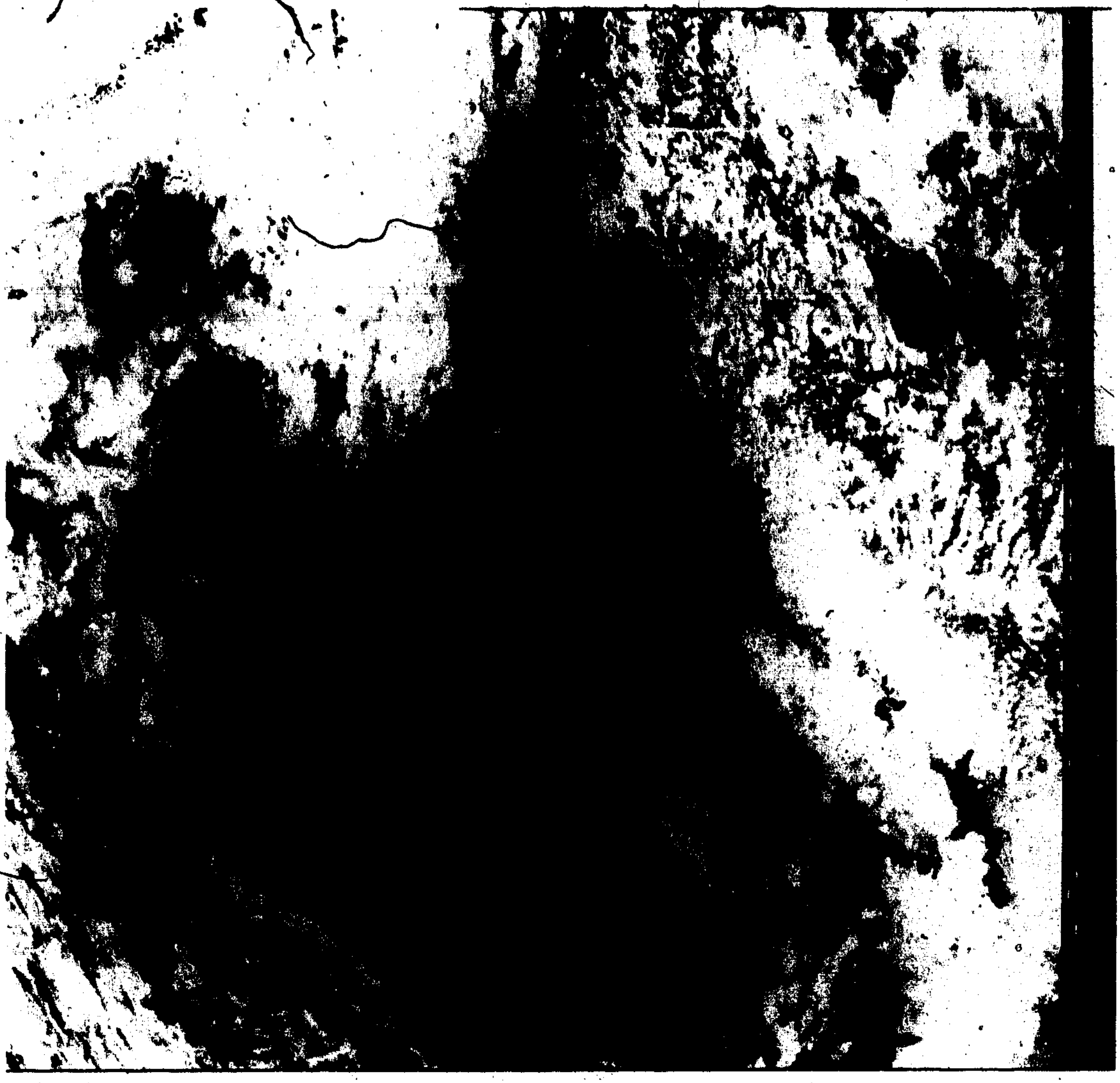


Plate 4.8 Band 2 image of the west coast of the United States and British Columbia, taken from NOAA-9 orbit #13550, showing the location of the 15 sample areas listed in Table 4.2. The image is enlarged 2X, and has been enhanced to increase the contrast between cloud and water. The scan line and pixel corresponding to the bottom right corner of the image are 1750 and 128, respectively.

Plate 4.9 Band 2 image of the west coast of the United States and British Columbia, taken from NOAA-9 orbit #13550. The image is enlar[illegible] 4X, and has been enhanced to increase the contrast between cloud and water.

For the case of sample 12, a small sea-peak lies in the 3% reflectance class, amounting to approximately 9% of the sample area; the initial estimate of the total cloud-cover is 91% plus or minus a finite uncertainty. However, no sea-peak was identified in the (3,4) band set histogram field. The resulting cloud-cover estimate returned by the algorithm, is 100%; no uncertainty is produced because no extraction of the sea-peak takes place. Since the algorithm is designed to select the cloud-cover estimate having the lowest uncertainty, the estimate from the (3,4) band set is chosen (Table 4.2). Since the tenuous cloud layer is identified on the band 3 and band 4 images, the choice of the (3,4) cloud-cover estimate does not seem unreasonable. A similar situation during the estimation process for sample 4 probably took place.

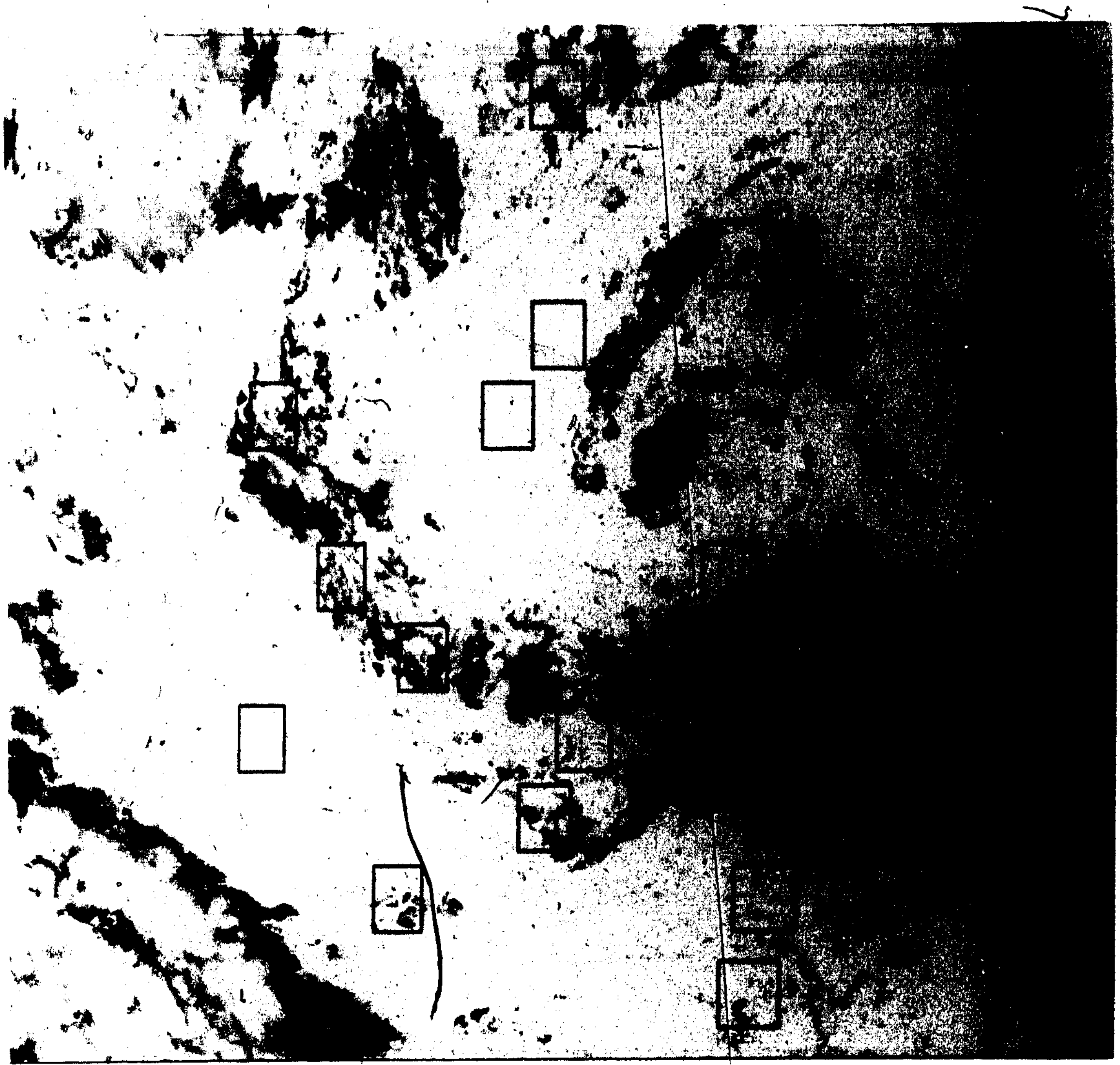


Plate 4.10 Band 3 image of the west coast of the United States and British Columbia, taken from NOAA-9 orbit #13550. The image is enlarged 4X, and has been enhanced to increase the contrast between cloud and water. Cloud is black (stratus/fog) to white (cirrus), the ocean surface, white.

The high estimate for sample 5 arises because of differences in the extraction accuracy between histograms using visible data and those using pure infra-red data; extraction of the sea peak performed using a curve is generally more accurate than for extraction performed using a surface. A small sea-peak was identified in the histogram fields for all three band sets; the lowest residuals were undoubtedly produced by extraction from the (2,3) and (2,4) histograms, despite the fact that the (3,4) histograms probably produced a better cloud-cover estimate. Since the algorithm's choice of the best estimate is dependent on the size of the residuals, a possibly less accurate estimate would have been selected.

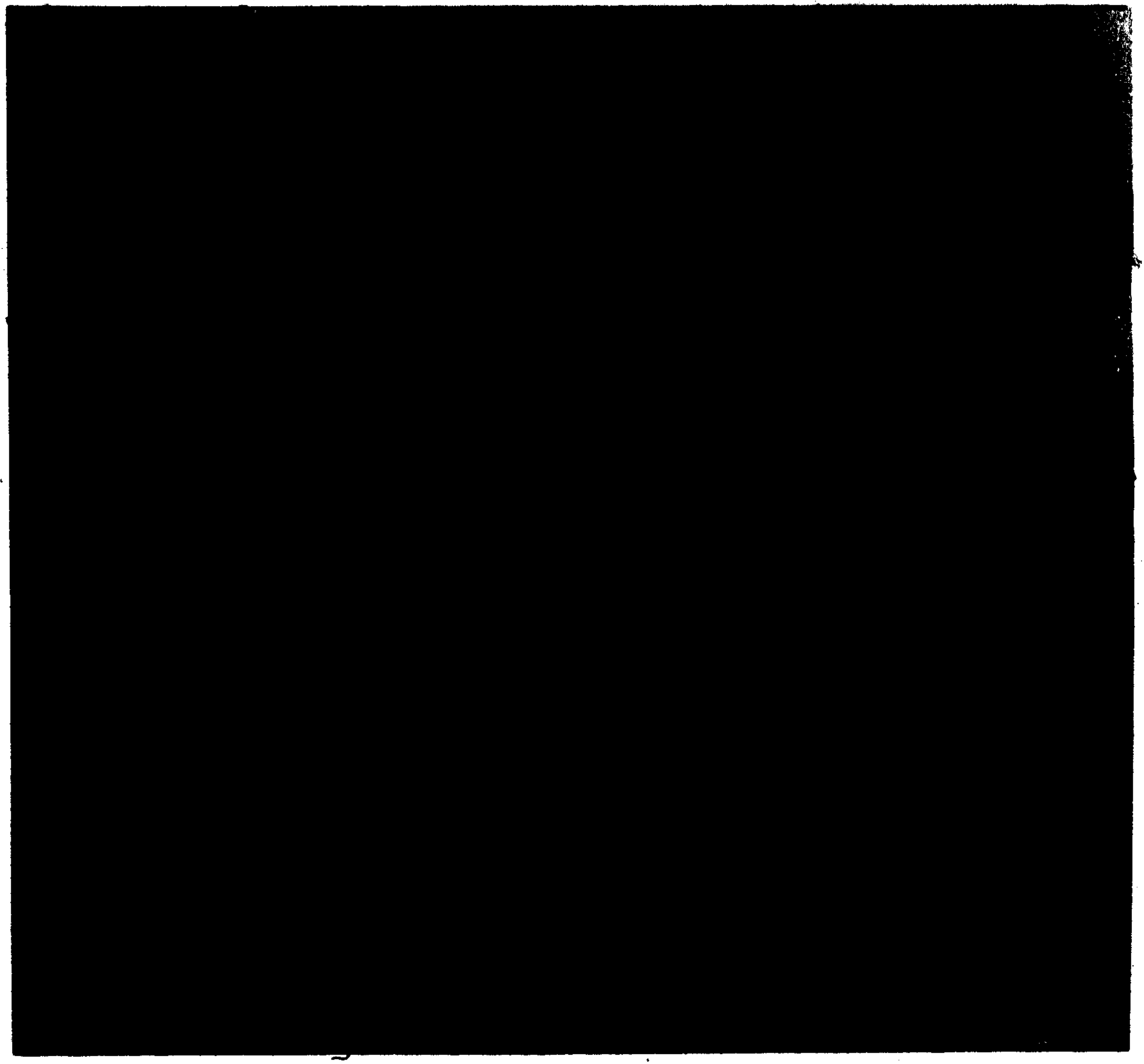


Plate 4.11 Band 4 image of the west coast of the United States and British Columbia, taken from NOAA-9 orbit #13550. The image is enlarged 4X, and has been enhanced to increase the contrast between cloud and water. Cloud is grey (stratus/fog) to white (cirrus), the ocean surface black.

5. Conclusions

The automatic cloud detection algorithm, as it has been applied here, produces estimates of the total cloud-cover that are accurate to ±10% with those of human visual nephanalysis based on satellite imagery. Unfortunately, because of a lack of on-the-spot cloud verification, such as ship reports, little can be said regarding the overall accuracy of the method. It is doubtful that ship reports, had they been available, would have contained cloud-cover estimates having accuracies better than ±10%.

Several problems have become apparent regarding its application. It appears to be very sensitive to small changes in sea-surface reflectance; increased dispersion of the sea and adjacent cloud peaks, beyond their respective class limits, can be produced by variable concentrations of surface plankton as far as 250km from shore, causing noticeable changes in the cloud-cover estimates. Errors in sea-peak identification, and poor selection of the best cloud-cover estimates, produce noticable inconsistencies between algorithm estimates and those from visual nephanalysis. A slightly wider reflectance class width would have been useful in dealing with this problem. The band 4 class width could also have been a little narrower. Had this been done, the dispersion of the peak along the band 4 axis would have been larger, allowing more accurate extraction of the sea-peak. The effectiveness of the (3,4) band set histograms would also have been increased.

Probably the most significant result; however, is the way in which band 3 data were able to complement, and even enhance the cloud detection accuracy. Its overall sensitivity to small changes in sensed radiation provided good cloud-cover estimates in areas where band 4 alone failed. The potential use of this band, with band 4 data, for nighttime cloud detection is al[illegible] present. In the absence of solar contamination, temperatures measured by band 3 should equal those of band 4 regardless of the nature of the surface. Some way of identifying the sea-peak will have to be developed, since the ±1°K peak temperature difference criterion used here is not valid at night.

In general, bi-spectral techniques such as this show a good deal of promise for detecting areas of cloud over water. Information regarding the extent of individual cloud layers is available by extracting the cloud peaks in the same fashion that the sea-peak was extracted here. It also has the potential to provide cloud classification information, provided the correct criteria can be developed and built into the algorithm.

Unfortunately, it may be difficult at present to implement these techniques for real-time operational use, because of practical limitations, such as the high-priority requirements for imagery by meteorologists having to meet rigid forecast dead-lines. It is to be hoped, however, that such limitations will be removed in the near future, when increased computing power and new hardware will make it possible to produce multiple copies of data sets and imagery in a routine manner.

Bibliography

The Astronomical Almanac, 1987: U.S. Printing Office, Washington, D.C./Her Majesty's Stationary Office, London.

Bernstein, R.L., 1982: Sea surface temperature estimation using the NOAA-6 Satellite Advanced Very High Resolution Radiometer, *J. Geophys. Res.*, 87, 9455-9465.

Bullas, John, and Goodson, Ron, 1986: Detection of Fog and Stratus Using Band 3 Imagery, *First Workshop on Operational Meteorology*, Atmospheric Environment Service/Canadian Meteorological and Oceanographic Society, 237-244.

Chahine, Moustafa T., McCleese, Daniel J., Rosenkranz, Philip W., and Staelin, David H., 1983: Interactions Mechanisms Within the Atmosphere, *Manual of Remote Sensing, 2nd Edition*, American Society of Photogrammetry, Falls Church, Virginia, 165-230.

Desbois, M., Seze, G., Szejwach, W., 1982: Automatic Classification of clouds on Meteosat imagery: Application to high-level clouds, *J. Appl. Meteor.*, 21, 401-412.

Essenwanger, O., 1976: *Applied Statistics in Atmospheric Science, Part A: Frequencies and Curve Fitting*, Elsevier Scientific Publishing Company, Amsterdam.

Greaves, B.T., 1981: *Satellite Measured Changes in Thermal Fields of an Arctic Island*, MSc Thesis, University of Alberta, Edmonton, 189pp.

Isaak, S., and Manougian, M.N., 1976: *Basic Concepts of Linear Algebra*, W.W. Norton and Company, New York.

Lauritson, S., Nelson, G.J., and Porto, F.W., 1979: *Data Extraction and Calibration of TIROS-N/NOAA Radiometers*, NOAA Tech. Mem., NESS 107, Washington, D.C., 90pp.

Maul, George A., 1981: Application of GOES Visible-Infrared Data to Quantifying Mesoscale Ocean Surface Temperatures, *J. Geophys. Res.*, **86**, 8007-8021.

Maul, George A., and Sidran, M., 1973: Atmospheric effects on Ocean Surface Temperature Sensing from the NOAA Satellite Scanning Radiometer, *J. Geophys. Res.*, **78**, 1909-1916.

Payne, Richard E., 1972: Albedo of the Sea Surface, *J. Atmos. Sci.*, **29**, Part 2, 959-970.

Phulpin, T., Derrien, M., and Brard, A., 1983: A Two-Dimensional Histogram Procedure to Analyze Cloud Cover from NOAA Satellite High-Resolution Imagery, *J. Appl. Meteorol.*, **22**, 1332-1345.

Reinelt, E.R., Hof, P., Oracheski, D., and Broszkowski, J., 1975: *Research Studies of Numerical Enhancement of APT Radiometer Data for Application to Arctic Weather and Ice Prediction*, Final Report, DDS(AES) Contract OSV4-0183, University of Alberta, Edmonton, 204pp.

Robinson, I.S., 1985: *Satellite Oceanography*, John Wiley & Sons, Rexdale, Ontario.

Ruff, I., and Gruber, A., 1975: *Graphical Relations Between a Satellite and a Point Viewed Perpendicular to the Satellite Velocity Vector (Side Scan)*, NOAA Tech. Mem., NESS 65, Washington, D.C., 14pp.

Schwalb, A., 1979: *The TIROS-N/NOAA A-G Satellite Series*, NOAA Tech. Mem., NESS 95, Washington, D.C., 75pp.

Smart, W.M., 1977: *Spherical Astronomy, 6th Edition*, Revised by Green, R.M., Cambridge University Press, Cambridge.

Smith, W.L., Rao, P.K., Koffler, R., and Curtis, W.P., 1970: The determination of sea surface temperature from satellite high resolution infrared window radiation measurements, *Mon. Weather Rev.*, **98**, 604-611.

Stewart, Robert H., 1985: *Methods of Satellite Oceanography*, University of California Press, Berkeley, California.

Walton Smith, F.G. and Kalber, F.A.,1974: *Handbook of Marine Science, Vol 2*, CRC Press, Cleveland.

Appendix 1

Fitting Gaussian Curves and Surfaces using the Method of Least Squares.

Consider a gaussian distribution of the form:

$$f = \frac{N}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-\bar{x})^2}{2\sigma_x^2}} \qquad (1)$$

where N is the number of sample points, $\bar{x}$ is the sample mean, σ_x^2, is the sample variance, and f is the predicted frequency. It is possible to reduce equation (1) to a second order polynomial in x by taking the natural logarithm of both sides. The result is of the form:

$$\ln(f) = a_0 - \frac{(x-\bar{x})^2}{2\sigma_x^2} \qquad (2)$$

which reduces to:

$$\ln(f) = a_0 + a_1x + a_2x^2 \qquad (3)$$

where:

$$a_0 = \ln\left[\frac{N}{\sqrt{2\pi\sigma_x^2}}\right] \qquad (4)$$

$$a_1 = \frac{\bar{x}}{\sigma_x^2} \qquad (5)$$

$$a_2 = -\frac{1}{2\sigma_x^2} \qquad (6)$$

The parameters for the 'best-fit' gaussian for the one-dimensional case are then given by:

$$\sigma_x^2 = \frac{1}{2a_2}, a_2<0 \qquad (7)$$

$$\bar{x} = a_1\sigma_x^2 \qquad (8)$$

$$f_{max} = \frac{N}{\sqrt{2\pi\sigma_x^2}} e^{(a_0 - a_2x^2)} \qquad (9)$$

where f_{max} is the central frequency of the gaussian curve.

The coefficients a_0-a_2 in equations 7-9, are determined using the method of least squares by forming the following normal equations:

$$\Sigma\ln(f) = a_0N + a_1\Sigma x + a_2\Sigma x^2 \qquad (10)$$

$$\Sigma x\ln(f) = a_0\Sigma x + a_1\Sigma x^2 + a_2\Sigma x^3 \qquad (11)$$

$$\Sigma x^2\ln(f) = a_0\Sigma x^2 + a_1\Sigma x^3 + a_2\Sigma x^4 \qquad (12)$$

and solving the system of equations using Gauss-Jordan elimination (Isaac and Manougian, 1976).

The bi-variate case is entirely analogous to the uni-variate case, except that diagonalization of a 6x6 matrix is required to determine the normal coefficients.

Consider a bi-variate Gaussian distribution of the form:

$$f = \frac{N}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} e^{-\frac{Q^2}{2(1-r^2)}} \qquad (13)$$

where

$$Q^2 = \frac{(x-\bar{x})^2}{\sigma_x^2} - \frac{2r(x-\bar{x})(y-\bar{y})}{\sigma_x\sigma_y} + \frac{(y-\bar{y})^2}{\sigma_y^2} \qquad (14)$$

N represents the total number of points in the sample, $\bar{x}$, $\bar{y}$ and σ_x^2, σ_y^2 are the mean values and variances of the variates x and y respectively, r is the correlation coefficient, and f is the predicted frequency of the gaussian. Taking the natural logarithm of both sides produces a second order surface in x and y of the form:

$$\ln(f) = a_0 - \frac{Q^2}{2(1-r^2)} \tag{15}$$

which, after some algebra, reduces to:

$$\ln(f) = a_0 + a_1 x + a_2 y + a_3 xy + a_4 x^2 + a_5 y^2 \tag{16}$$

where:

$$a_0 = \ln\left[\frac{N}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}}\right] \tag{17}$$

$$a_1 = \frac{1}{(1-r^2)}\left[\frac{\overline{x}}{\sigma_x^2} - \frac{r\overline{y}}{\sigma_x\sigma_y}\right] \tag{18}$$

$$a_2 = \frac{1}{(1-r^2)}\left[-\frac{r\overline{x}}{\sigma_x\sigma_y} + \frac{\overline{y}}{\sigma_y^2}\right] \tag{19}$$

$$a_3 = \frac{r}{(1-r^2)\sigma_x\sigma_y} \tag{20}$$

$$a_4 = -\frac{1}{2(1-r^2)\sigma_x^2} \tag{21}$$

$$a_5 = -\frac{1}{2(1-r^2)\sigma_y^2} \tag{22}$$

Solution of equations 17-22 yields the gaussian parameters in the form:

$$r = -\frac{a_3}{2a_5}\sqrt{\frac{a_5}{a_4}} \tag{23}$$

$$\sigma_x^2 = -\frac{1}{2(1-r^2)a_4}, a_4<0 \tag{24}$$

$$\sigma_y^2 = -\frac{1}{2(1-r^2)a_5}, a_5<0 \tag{25}$$

$$\bar{x} = a_1\sigma_x^2 + a_1 r\sigma_x\sigma_y \tag{26}$$

$$\bar{y} = a_2\sigma_y^2 + a_1 r\sigma_x\sigma_y \tag{27}$$

$$f_{max} = \frac{N}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} = e^{(a_0 - a_4\bar{x}^2 - a_5\bar{y}^2 - a_3\overline{xy})} \tag{28}$$

where f_{max} is the central frequency of the surface.

The normal equations, used to determine the coefficients a_0-a_5 are of the form:

$$\Sigma\ln(f) = a_0N + a_1\Sigma x + a_2\Sigma y + a_3\Sigma xy + a_4\Sigma x^2 + a_5\Sigma y^2 \tag{29}$$

$$\Sigma x\ln(f) = a_0\Sigma x + a_1\Sigma x^2 + a_2\Sigma xy + a_3\Sigma x^2y + a_4\Sigma x^3 + a_5\Sigma xy^2 \tag{30}$$

$$\Sigma y\ln(f) = a_0\Sigma y + a_1\Sigma xy + a_2\Sigma y^2 + a_3\Sigma xy^2 + a_4\Sigma x^2y + a_5\Sigma y^3 \tag{31}$$

$$\Sigma xy\ln(f) = a_0\Sigma xy + a_1\Sigma x^2y + a_2\Sigma xy^2 + a_3\Sigma x^2y^2 + a_4\Sigma x^3y + a_5\Sigma xy^3 \tag{32}$$

$$\Sigma x^2\ln(f) = a_0\Sigma x^2 + a_1\Sigma x^3 + a_2\Sigma x^2y + a_3\Sigma x^3y + a_4\Sigma x^4 + a_5\Sigma x^2y^2 \tag{33}$$

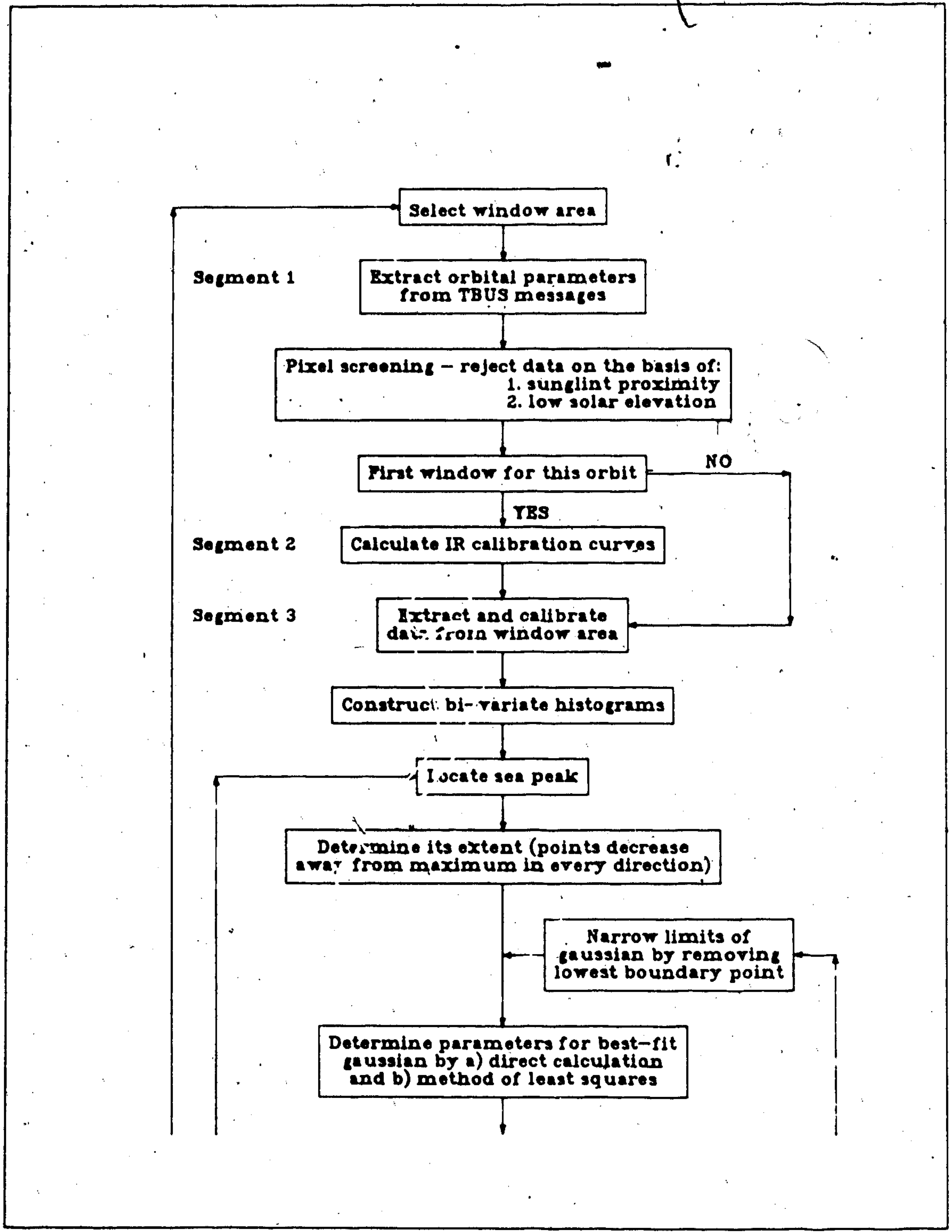
$$\Sigma y^2\ln(f) = a_0\Sigma y^2 + a_1\Sigma xy^2 + a_2\Sigma y^3 + a_3\Sigma xy^3 + a_4\Sigma x^2y^2 + a_5\Sigma y^4 \tag{34}$$

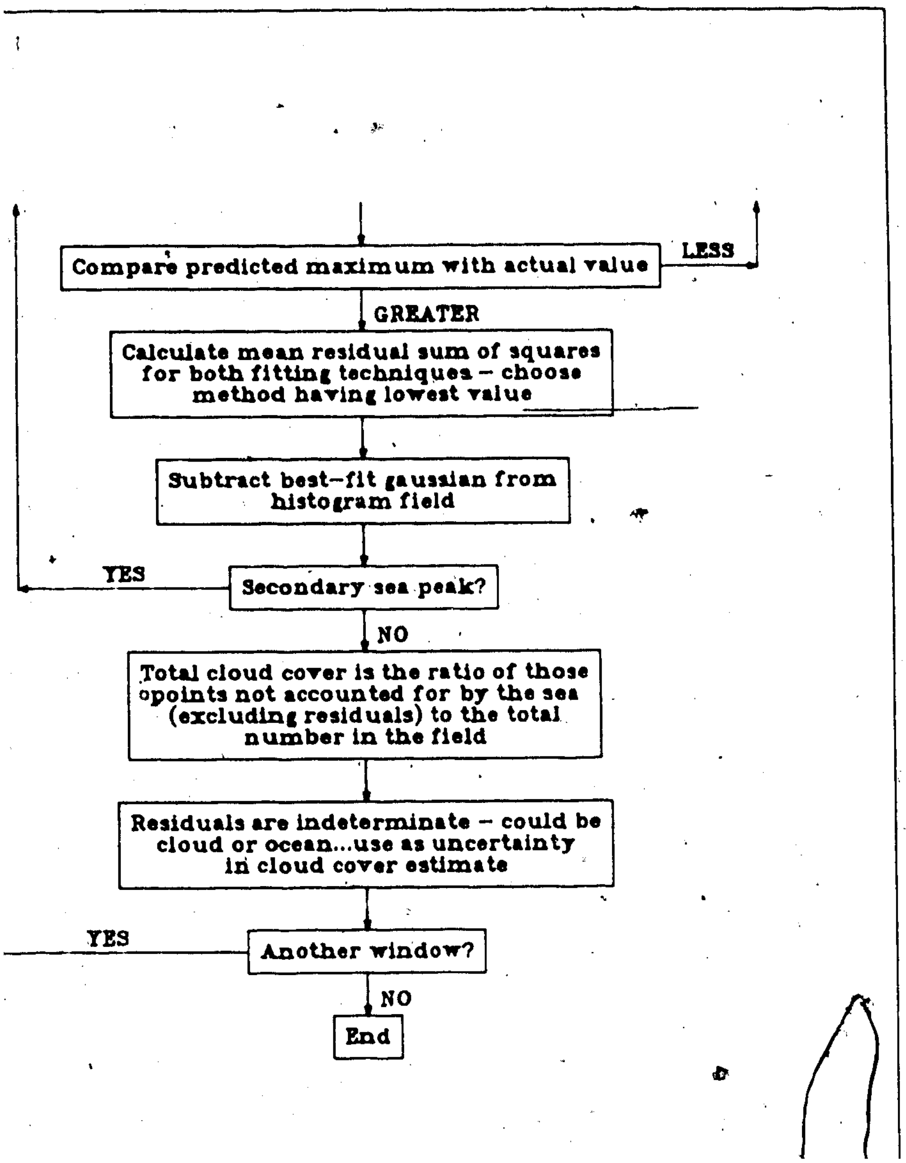
Zero is an unacceptable data value because of the logarithmic nature of the least squares technique; however, little error is introduced if zero values are assumed to equal one. This assumption is only required during analysis of the band (3,4) histograms; it arises because the off-axis frequencies are not examined during initial determinion of the extent of the sea peak, and zero bin frequencies are often present.

As in any fitting procedure, a minimum of three points are required in each direction to uniquely determine the best fit curve or surface. Certain gaussian parameters will become arbitrary if fewer than three points are used, generally leading to non-recoverable computer errors.

Appendix 2

A Schematic Diagram of the Total Cloud-cover Estimation Algorithm.




Compare predicted maximum with actual value
LESS
GREATER
Calculate mean residual sum of squares for both fitting techniques – choose method having lowest value
Subtract best-fit gaussian from histogram field
YES
Secondary sea peak?
NO
Total cloud cover is the ratio of those points not accounted for by the sea (excluding residuals) to the total number in the field
Residuals are indeterminate – could be cloud or ocean...use as uncertainty in cloud cover estimate
YES
Another window?
NO
End