

THE RETRIEVAL OF **Surface Parameters**  
FROM SATELLITE BORNE INFRARED RADIOMETERS  
FOR THE STUDY OF CLIMATE

**Andrew Robert Harris**

*A thesis submitted in fulfillment of the requirements  
for the degree of PhD*

Mullard Space Science Laboratory  
University College London

*September 1991*

ProQuest Number: 10609303

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10609303

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

# Abstract

This thesis concerns the development and application of new infrared remote sensing techniques for measurement of climate-related variables. The nature of the climate system is discussed, and the need for global monitoring is noted, together with the suitability of satellite-based remote sensing for the task. Current applications of data from satellite-borne infrared radiometers are discussed, together with the attendant problems, particularly that of correction for the effects of the atmosphere on remotely-sensed thermal infrared temperatures. In addition, the monitoring of proxy indicators of climatic change, such as the areas of closed lakes, by remote sensing is seen as having great potential, despite the limited research to date.

The problem of accurate measurement of lake areas by the necessarily coarse resolution instruments which are capable of providing the required repeat coverage is addressed. An initial case study shows that lakes of order a few hundred km<sup>2</sup> can be measured to an accuracy of 1% with 1 km resolution data from the Advanced Very High Resolution Radiometer (AVHRR). A further study of a climatically-sensitive closed lake in Ethiopia demonstrates a qualitative relationship between the measured area cycle and climate records. It is noted that the accurate remote sensing of lake surface temperatures and tropical ocean surface temperatures, both important parameters for climate research, is difficult due to the problem of atmospheric correction. A new correction algorithm is developed which offers an improvement of a factor ~2 over conventional algorithms when applied to AVHRR data. Useful by-products of the algorithm are accurate atmospheric transmittance and total water vapour. Further developments of the techniques devised are suggested with a view to maximising the exploitation of both new and existing global datasets in order to provide the necessary long time series of accurate measurements required for climate research.

# Contents

<b>Abstract</b>	<b>2</b>
<b>Contents</b>	<b>3</b>
<b>List of Figures</b>	<b>6</b>
<b>List of Tables</b>	<b>9</b>
<b>Acknowledgements</b>	<b>10</b>
<b>Chapter 1 Climate and Remote Sensing</b>	<b>12</b>
1.1 The Nature of the Earth's Climate	13
1.1.1 <i>A definition of climate</i>	13
1.1.2 <i>Climate history</i>	15
1.1.3 <i>Some consequences of climate change</i>	17
1.2 The Climate System - a brief overview	23
1.2.1 <i>The overall energy budget</i>	23
1.2.2 <i>Atmospheric circulation</i>	27
1.2.3 <i>Rôle of the ocean, cryosphere and land in the climate system</i>	31
1.2.4 <i>Feedback mechanisms within the climate system</i>	33
1.3 The Need for Climate Monitoring and Prediction	34
1.3.1 <i>The variability of climate</i>	34
1.3.2 <i>Possible effects of climate change</i>	37
1.3.3 <i>The World Climate Research Programme</i>	38
1.3.4 <i>Climate modelling</i>	38
1.4 The use of Remote Sensing Data in Climate Research	41
1.4.1 <i>The use of proxy indicators as monitors of climatic change</i>	42
<b>Chapter 2 An Introduction to Infrared Radiometry from Space</b>	<b>44</b>
2.1 Overview	44
2.2 The Nature of Electromagnetic Radiation	45
2.3 Absorption Processes in the Atmosphere	50
2.4 Radiative Transfer	56
2.5 Atmospheric Correction	60



2.6 Radiometer Design with AVHRR as an example	70
2.6.1 <i>Telescope and receiver systems</i>	70
2.6.2 <i>Detector types</i>	73
2.6.3 <i>Detector cooling systems</i>	73
2.7 Calibration of the Radiometer Signal	74
2.7.1 <i>Need for on-board calibration of the system</i>	74
2.7.2 <i>Calibration schemes for thermal channels</i>	74
2.8 The Operation of the Along-Track Scanning Radiometer (ATSR)	77
2.9 Summary	79
 <b>Chapter 3 Remote Sensing of Lake Areas :</b>	
<b>Technique Development, Validation and Comparison</b>	<b>81</b>
3.1 Review of Lake Area Remote Sensing	81
3.1.1 <i>The importance of lake volume monitoring for climate research</i>	81
3.1.2 <i>Current algorithms for measuring lake areas</i>	84
3.1.3 <i>Lough Neagh as a case study for investigating and validating the methods</i>	86
3.2 The Development of a Simple Area Algorithm	88
3.2.1 <i>Introduction to the main problems</i>	88
3.2.2 <i>Approaches for extracting the basic pixel area</i>	92
3.2.3 <i>Idealised one-dimensional histogram</i>	92
3.2.4 <i>Real one-dimensional histograms</i>	94
3.2.5 <i>Problems caused by the broad land distribution</i>	94
3.2.6 <i>The nature of the edge pixel distribution</i>	97
3.2.7 <i>Simple lake area method assuming symmetrical edge pixel distribution</i>	97
3.3 Simulated results	100
3.3.1 <i>Study of the edge pixel distribution for a non-infinite coastline</i>	100
3.3.2 <i>Investigation of the precision of the LIC edge detection method</i>	104
3.4 Application of the histogram technique to AVHRR images	104
3.5 Conclusion	111
 <b>Chapter 4 Time Series Remote Sensing of a Climatically Sensitive Lake</b>	<b>112</b>
4.1 Introduction	112
4.2 Lake Abiyata as a Case Study	113
4.3 A Comparison of Area Measurement Techniques for a Closed Lake	114
4.4 Area Cycle of Lake Abiyata obtained from a Time Series of AVHRR Images	124

4.5 Discussion of Results	128
4.6 Conclusions	129
<b>Chapter 5 A New Technique for Atmospheric Correction and Total Water Vapour Retrieval</b>	<b>131</b>
5.1 Introduction	131
5.2 Accurate Transmittance and Total Water Vapour	134
5.3 Improved Atmospheric Correction	143
5.4 Limits to Accuracy	150
5.5 Applications and Conclusions	155
<b>Chapter 6 A First Validation Study of the New Technique for Atmospheric Correction</b>	<b>157</b>
6.1 Introduction	157
6.2 The Validation Dataset	158
6.3 Analysis Procedure for the AVHRR Data	159
6.3.1 <i>Declouding procedure</i>	159
6.3.2 <i>Calculation of gradient and retrieved temperature</i>	162
6.3.3 <i>Image rectification and geolocation</i>	164
6.4 Results obtained for the various algorithms	169
6.5 Discussion of results	172
6.6 Conclusions and Recommendations	183
<b>Chapter 7 Concluding Remarks and Recommendations</b>	<b>184</b>
7.1 Summary of work presented	184
7.2 Assessment of contributions with respect to original aims	186
7.3 Recommendations for further study	186
<b>References</b>	<b>189</b>
<b>Appendix A A Simple Numerical Model of an Infinite Coastline</b>	<b>196</b>
<b>Appendix B Geometric Correction of AVHRR Pixel Areas</b>	<b>201</b>

# Figures

## Chapter 1

1.1	Examples of climate change over the last century	14
1.2	Prehistoric climate and species records	16
1.3	Air temperature and CO <sub>2</sub> records from ice core measurements	18
1.4	Climate of the last 1000 years	19
1.5	Time series of SST isotherms illustrating ENSO event	20
1.6	Cave paintings in the Sahara	22
1.7	Solar and terrestrial Planck curves and atmospheric absorption	25
1.8	Simple heat budget illustrating greenhouse effect	26
1.9	Components of the earth's climate system	28
1.10	Distribution of shortwave and longwave fluxes wrt latitude	29
1.11	Hadley Cell circulation	30
1.12	Rise in concentrations of various greenhouse gases	36
1.13	Changes in global temperatures due to doubling of CO <sub>2</sub>	39

## Chapter 2

2.1	Solar and terrestrial Planck curves at top and bottom of atmosphere	47
2.2	Effects of absorption and scattering on incoming shortwave radiation	48
2.3	Full-disc Meteosat band 1 image	49
2.4	Curve of growth of an absorption line	53
2.5	Radiative transfer through slab of elemental thickness	58
2.6	Radiative transfer through slab of elemental pressure	61
2.7	Planck radiance between 260K and 310K at 10.8 and 12.0 $\mu$ m	63
2.8	Channel 4 atmospheric correction versus T <sub>4</sub> - T <sub>5</sub>	68
2.9	Absorption profiles for 3 different atmospheres	69
2.10	The AVHRR instrument	71
2.11	Non-Linear Correction for AVHRR Channels 4 and 5	75
2.12	The ATSR instrument	78

## Chapter 3

3.1	Areas of inland drainage	83
3.2	Image segmentation procedure	87
3.3	Idealised one-dimensional histogram	93
3.4	AVHRR Channel 2 image and brightness histogram	95

3.5	Idealised symmetrical histogram	98
3.6	Density-sliced TM image of Lough Neagh	102
3.7	Binned-up density-sliced TM image	103
3.8	Comparison of results from numerical model and binup analysis	105
3.9	Isoluminance contour for binned up TM image	106
3.10	Histograms of Lough Neagh for AVHRR Channels 1, 2 and 4	109
3.11	Edge warming effect on thermal image and histogram	110

## Chapter 4

4.1	Level - area relationship and derived area variations of Lake Abiyata	115
4.2	False colour MSS image of Lake Abiyata	117
4.3	Resampled MSS image and AVHRR image	119
4.4	Close-fitting polygon over image and resultant histogram	120
4.5	Isoluminance contour for resampled MSS image	121
4.6	Calculation of effective active scan angle of MSS instrument	123
4.7	Example of thresholding out thin cloud with LIC method	126
4.8	Area cycle of Abiyata derived from AVHRR images	127

## Chapter 5

5.1	Illustration of radiance components observed by satellite	136
5.2	Planck radiance between 260K and 310K at 10.8 and 12.0 $\mu$ m	138
5.3	AVHRR channels 5 and 4 transmittance ratio versus $\Delta T_5/\Delta T_4$	139
5.4	AVHRR channel 4 and 5 transmittances versus $\Delta T_5/\Delta T_4$	140
5.5	Regressions of water vapour against $\Delta T_5/\Delta T_4$	142
5.6	Channel 4 atmospheric correction versus $T_4 - T_5$	146
5.7	Channel 5 transmittance versus channel 4 transmittance	148
5.8	Water vapour and temperature profiles for atmospheres 25 and 31	149
5.9	Channel 4 atmospheric correction versus $(T_4 - T_5)/R_{2,1}$	151
5.10	RMS residuals for final algorithm at nadir and 60°	152

## Chapter 6

6.1	Schematic of APOLLO declouding procedure	160
6.2	Angle subtended at centre of earth for known scan angle	167
6.3	X-dependent Y-distortion	168
6.4	Three geolocated AVHRR brightness temperature subscenes	170
6.5	Frequency histograms of results for the various algorithms	171
6.6	Algorithm - <i>in situ</i> differences versus expected correction	173
6.7	Retrieved temperatures versus <i>in situ</i> temperatures	174

6.8	<i>In situ</i> - $T_4$ versus <i>in situ</i> temperature	175
6.9	Frequency histograms of results with outliers removed	177
6.10	<i>In situ</i> - retrieved versus <i>in situ</i> - $T_4$ with outliers removed	179
6.11	New algorithm - conventional algorithm versus new algorithm - $T_4$	181
6.12	'Split-window' algorithms derived from global and regional data	182

## Appendix A

A.1	Diagram of straight coastline across a square pixel	197
-----	---	-----

## Appendix B

B.1	Angle subtended at centre of the earth	203
B.2	Spherical triangle formed by meridian, satellite track and equator	206

# Tables

## Chapter 3

3.1 Results for area of Lough Neagh obtained from AVHRR data	107
--	-----

## Chapter 5

5.1 NOAA-7 'split window' coefficients for various airmasses	145
5.2 NOAA-7 'new algorithm coefficients for various airmasses	153

# Acknowledgements

Firstly, I would like to thank Dr. Ian Mason for guiding my research efforts and providing necessary input of scientific rigour, and for his helpful comments on the manuscript. I would also like to thank the Head of the Remote Sensing Group at MSSL, Professor Chris Rapley, for being instrumental in ensuring that the work was completed on time. I would also like to thank the Director of MSSL, Professor Len Culhane, for showing concern and providing practical advice on thesis completion. David Palmer, as one of the originators of the LIC technique successfully employed in this thesis, is one whose help I gratefully acknowledge. I would also like to thank Gareth Davies and Simon Brown for taking over some of my duties in the final weeks of preparation, and Derek Hoyle for preparing one of the figures. Thanks also to the other members of the RSG for providing a stimulating environment to work in.

Other institutes have also contributed to my work, and I would like to thank several people in this respect.

I gratefully acknowledge Albin Zavody of RAL as supplier of the atmospheric transmission model used extensively in my work, and also for his valuable expertise.

I would like to thank NOAA as suppliers of the AVHRR data used in the closed lake study, and particularly the staff at Satellite Data Services Division in Washington who made me welcome and assisted in my data search.

The help of the overworked staff at the University of Dundee Satellite Station in providing the Lough Neagh AVHRR data is also acknowledged.

The provision of the Great Lakes dataset, including both buoy measurements and the AVHRR scenes, by George Irbe of CCAH, was a generous act which I gratefully acknowledge.

Finally, thanks are due to SERC who provided the funding for me.

**To my parents, who have supported me in everything I have ever done**



# Chapter 1

## Climate and Remote Sensing

The Earth's climate is a complex and dynamic system which has always had a fundamental effect on the life of the planet. In recent years, concern for the impact of variations in the climate over the globe has resulted in much research into all aspects of the system. Various international programmes have been set up to monitor changes using a variety of techniques, and also to model the climate system in an attempt to provide a means of predicting future trends under various scenarios. The measurement of climate-related variables is required in order to facilitate the development of reliable models, and also to determine current trends with a view to empirical prediction of future climatic conditions. Quantities such as air and sea temperatures and precipitation can be regarded as 'direct' measures of the Earth's climate, and these can be measured by conventional means using meteorological stations and ships. However, the coverage provided by such ground-based methods is very non-uniform, for example, the vast majority of measurements of sea surface temperature are provided by 'ships of opportunity' and hence are usually confined to areas of commercial shipping lanes. The advent of remote-sensing, however, particularly from satellites, has provided a new source of such data with the advantage of global coverage.

Remotely-sensed data can be used to provide direct measures of the Earth's climate, such as sea surface temperature maps derived from satellite-borne radiometers (eg. Llewellyn-Jones *et al*, 1984), and also provide the ability to study various phenomena which are related to climate. Quantities such as the sea-ice extent (eg. Laxon, 1989) and lake volume cycles (Mason *et al*, 1985) can serve as 'proxy indicators' of climate. However, some understanding of the problems associated with the use of remotely-sensed data, together with a review of current methods, is necessary before we can derive techniques to enable the extraction of physical quantities, and this is covered in Chapter 2. The development of techniques for measuring lake areas, which can serve as proxy indicators of climate change, is described in Chapter 3, and demonstrates that accuracy requirements can be met with existing data that is also capable of providing frequent global coverage. The application of a technique on a climatically-sensitive lake is shown in Chapter 4, and, although the precise interpretation of results is complex and beyond the scope of this

thesis, it provides a qualitative demonstration of the use of lake areas as proxy indicators. The development of a new algorithm for measurement of surface temperatures and atmospheric water vapour content from space-borne infrared radiometer data is described in Chapter 5, and a first validation of the new correction method is shown in Chapter 6, and demonstrates significant improvements over existing techniques.

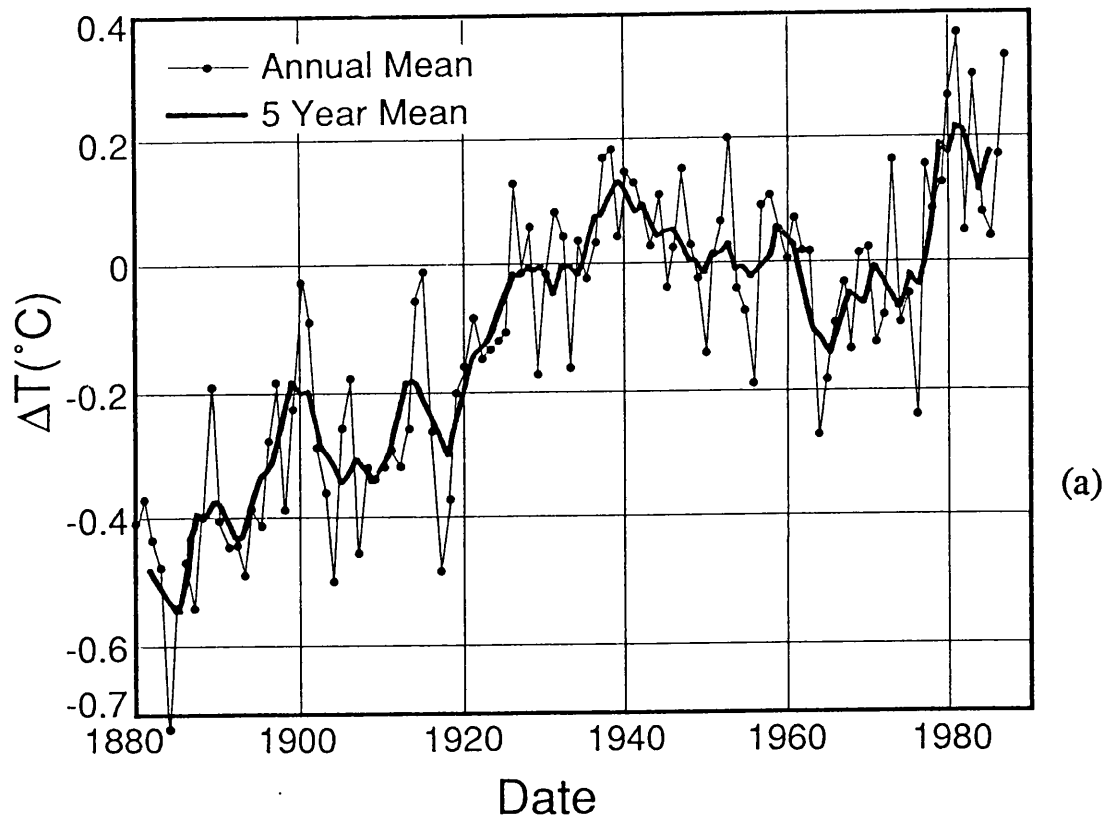
The principal aim of the work contained in this thesis can be summarised as the development of techniques to enable the extraction of physical quantities from remotely-sensed data that can contribute to climate research. Before this, however, a brief review of the nature of the climate system and attendant phenomena is presented in order to provide some background to the work.

## **1.1 The Nature of the Earth's Climate**

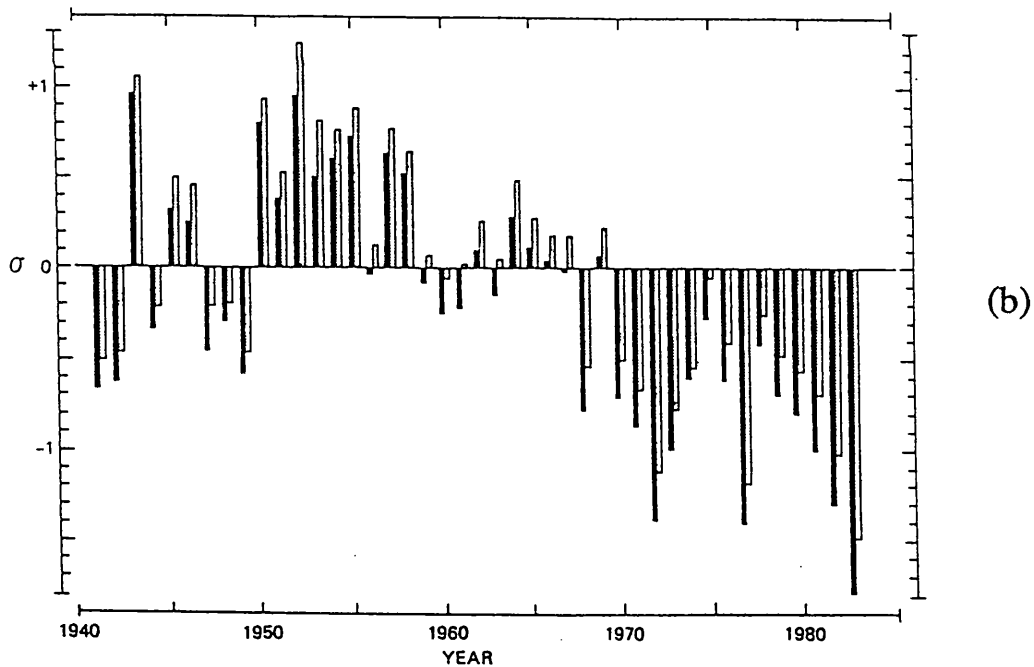
### **1.1.1 A Definition of Climate**

It is generally recognised that climate is in some sense the average of weather, its fluctuations, and its influence on the surface of the Earth (Leith, 1984). For the purposes of climate monitoring, this intuitive concept needs to be rendered more precise. Whereas weather is the state of the atmosphere at a given time and location, the climate can be considered to be the state of the atmosphere over a period of time and for a particular region, expressed in terms of average values and a measure of variability. Examples of the variations of two climate parameters (air temperature and rainfall) are shown in Figures 1.1(a) and 1.1(b). These are both time-averaged over the year; spatial averaging is over the globe and a region respectively. When monitoring long term climatic variations, it is often useful to remove the strong annual frequency component by taking the annual average, as in the above examples. Variations in the annual cycle itself can be measured by sampling at particular phases (eg. measuring the mean temperature for each month over a period of years). Note that although such descriptions of the climate often refer to the atmosphere alone since its effect is most easily noticed, the climate system (shown in Figure 1.9) includes components other than the atmosphere whose interactions with the atmosphere and each other determine the climate.

We can see that the annually and globally averaged air temperature is by no means an adequate description of the Earth's climate, for the seasonal variations are a recognisable part of what we are trying to define. Even if we allow the annual variability to be included, this still will not allow us to define the global climate, since



Rainfall in Subsaharan West Africa during 1941—83



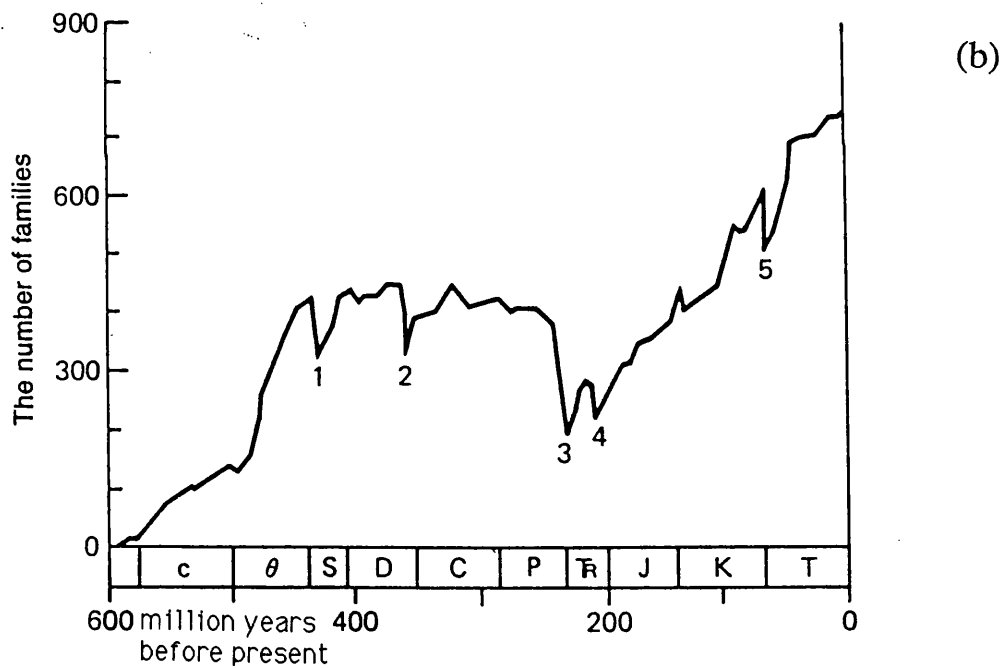
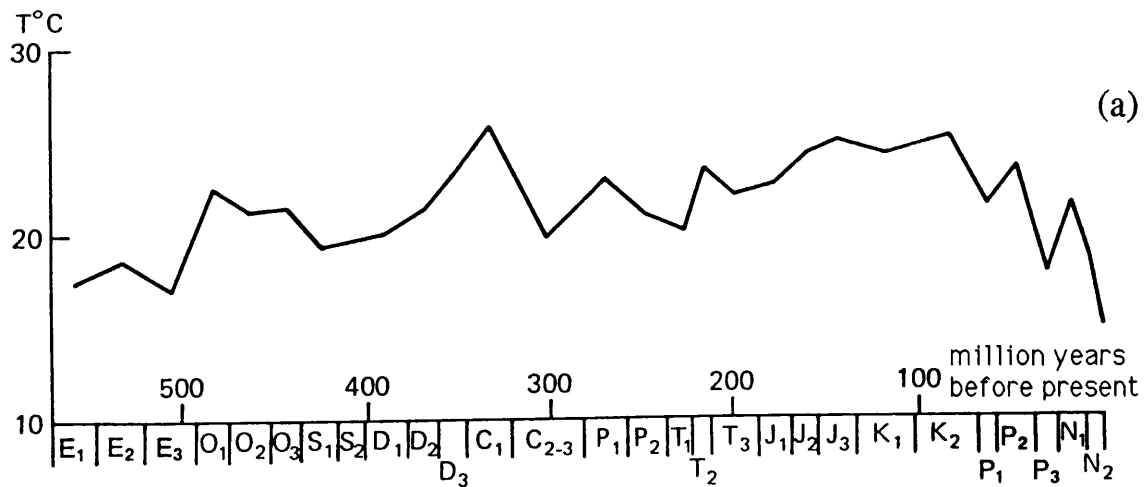
**Figure 1.1** The global mean air temperature record for the past century is shown in (a), taken from 'Global Climate Change', 1990, after Hansen and Lebedeff, 1987. A time series of the yearly average of the normalised rainfall departures for 14 - 20 SubSaharan stations that lie between 11° and 18° N and west of 10° E is shown in (b). The solid bars have a different base period (1941-74) from the others (1941-82) (from Lamb, 1985).

the local averages and annual cycles also play their part in what we would call the normal climate. For example, we expect the island of Rhodes to be sunny for over 300 days per year and would consider the earth's climate to be different if Rhodes had the sunshine record of Bognor Regis and vice-versa, even if the annual cycle and average for the globe would remain unaltered. We may also consider the interannual variations, eg. possibly induced by the 22 year solar activity cycle and other sources of 'natural' variability, to be part of the normal climatic state. We therefore need to decide what measure of variability to include in the definition of a particular climatic state before we can go on to detect climate change.

### 1.1.2 Climate History

Climate is a dynamic system subject to natural variations on all timescales. Prehistoric records indicate the variable nature of the Earth's climate over the Phanerozoic spanning the last 600 million years. The start of the Cambrian period at the beginning of the Paleozoic era represents the onset of skeletal animals and the vegetation of the land masses, thus estimates of the mean surface air temperature can be derived from fossil records (see Figure 1.2(a), Budyko *et al*, 1985). At the beginning of this period the solar luminosity was about 3% less than the present and the oxygen content of the atmosphere was approaching current levels (Budyko *et al*, 1988). Despite the lower levels of incoming solar energy, the temperature was some 15 degrees K higher than the present due to the atmospheric concentrations of CO<sub>2</sub> being 6 to 10 times higher than current levels; the physical implications of this will be discussed in section 1.2. The climate appears to be relatively stable throughout the entire period, although the onset of a cooling trend is apparent about 100 million years ago. There is also some evidence for a period of glaciation between the Carboniferous and Permian periods about 270 million years ago. Fossil records also catalogue species extinction events (Figure 1.2(b)) which may have been brought about by sudden climate change, the most famous of which, the extinction of the dinosaurs, occurred at the end of the Cretaceous period, marking the end of the Mesozoic era.

The Cenozoic era began around 67 million years ago, and represents an increase in the rate of cooling (see Figure 1.2(a)). The polar ice caps formed for the first time about 14 million years ago and the Quaternary period covers the last 2 million years (the Pleistocene epoch) which have been marked by cycles of glacials and interglacials (Kukla, 1980), the epoch marking the 10,000 years since the last glaciation being known as the Holocene. The glacial cycles themselves are generally



**Figure 1.2** Mean surface air temperature changes in the Phanerozoic are shown in (a), and changes in marine animal families for the same period are shown in (b), with 1 - 5 marking epochs of mass extinctions. Both figures are taken from Budyko *et al*, 1988.

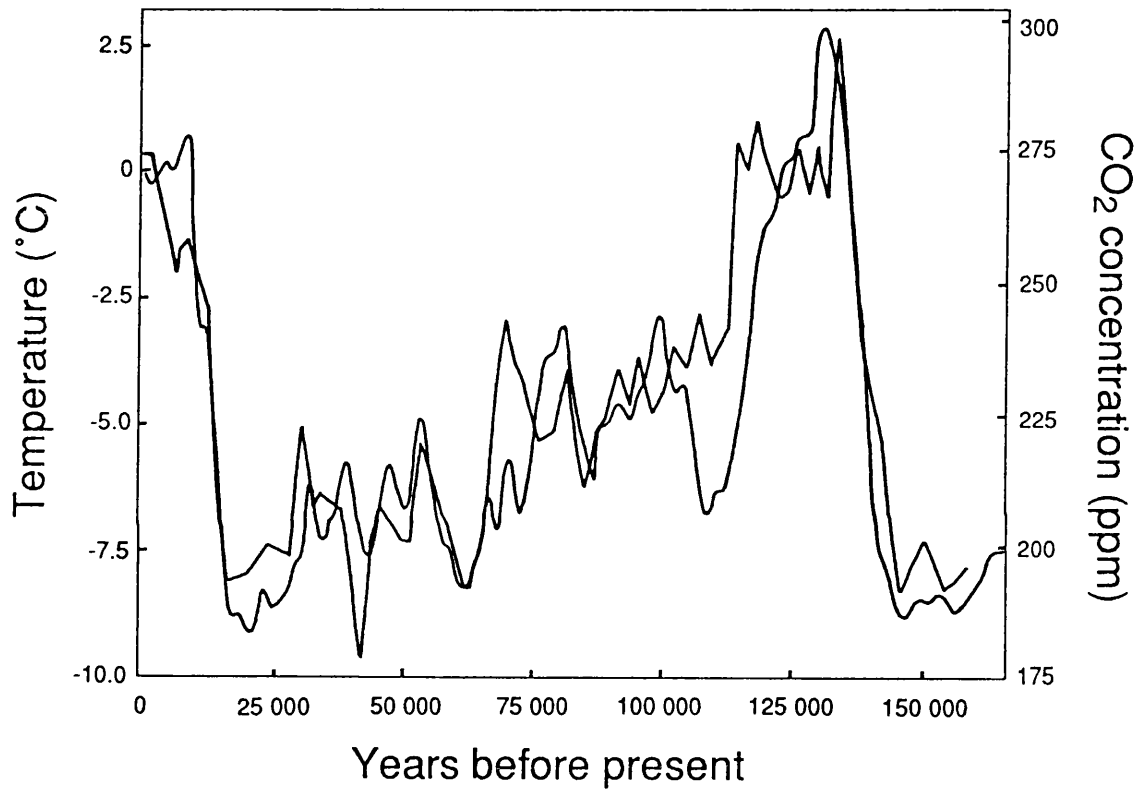
agreed to be related to changes in the orbital parameters of the earth, according to the theory first proposed by Milankovitch (Berger, 1980). The recurrent variation of the eccentricity of the earth's orbit causes a 0.2% variation in the amount of solar radiation intercepted by the earth; the climate system must considerably amplify these small flux changes in order to induce such differing climatic conditions as represented by glacials and interglacials.

The most reliable information about paleoclimates is derived from the most recent glacial cycle. The variation in the ratios of  $O^{18}$  to  $O^{16}$  contained in the  $H_2O$  molecules of ice cores permit an estimation of the air temperature over the ice and results, including the  $CO_2$  concentrations, from a 2000 metre deep ice core drilled at the Soviet Vostok base in Antarctica are shown in Figure 1.3 (Barnola *et al*, 1987). At the time of the last glacial maximum, 20,000 years ago, geological records have shown that ice sheets covered Canada (the Laurentide ice sheet) and large areas of Asia and Northern Europe, and sea level was almost 80m lower (Houghton and Morel, 1984). Climate models indicate that the sea ice extent was also much greater than at present (Untersteiner, 1984).

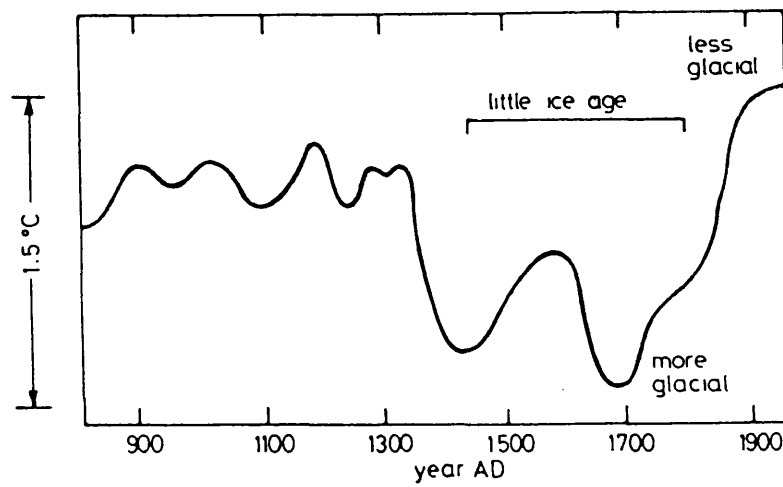
During the past 1000 years, the most prominent climatic feature has been the 'little ice age' (Figure 1.4). Although the mean temperature of the Northern Hemisphere was only 1 degree K lower than the present, sea ice cover and glaciers advanced considerably and winters in Europe were much more severe. During the present century, there have been some smaller climatic trends (see Figure 1.1(a)) but there is now increasing evidence to show that a global warming, induced by Man's activities, is beginning to be the dominant trend, although the increases detected so far are still not outside the 'natural variability' fluctuations observed over the last century or so. For example, there is an important short-term quasi-periodic variation in climate known as the Southern Oscillation (see Figure 1.5).

### **1.1.3 Some consequences of climate change**

Climate variation has a large direct impact on man's activities and on the economy of human populations, as demonstrated by the extremes of climate in recent years which have led to human tragedy. It is not necessary to do more than mention the droughts in Ethiopia and Sudan, or, at the other precipitation extreme, the floods in Bangladesh. Global warming raises the spectre of wider ranging effects such as climatic zone shifts affecting major grain production and indirect effects on low-lying countries due to a rise in sea level from thermal expansion of the ocean or the melting

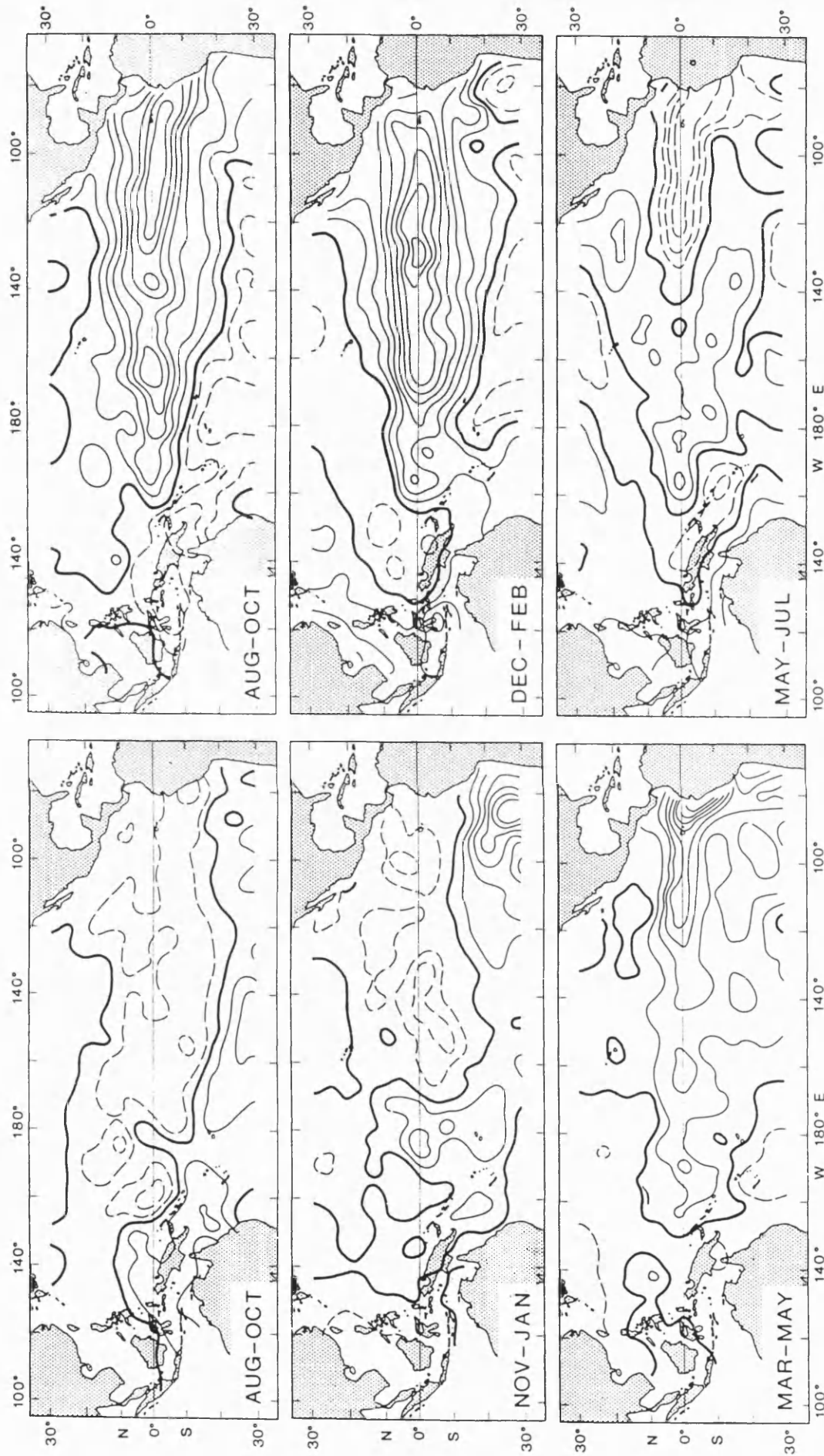


**Figure 1.3** Paleoclimatological record of temperature change with respect to modern climate and atmospheric CO<sub>2</sub> concentration over Antarctica during the last glacial cycle (from 'Global Climate Change', WCRP, 1990, after Barnola *et al*, 1987). The temperature record is the smoother curve.



**Figure 1.4** Climate of last 1000 years estimated from evidence relating to east European winters, and showing the occurrence of the 'Little Ice Age'. (From Houghton and Morel, 1984, after Lamb, 1966).





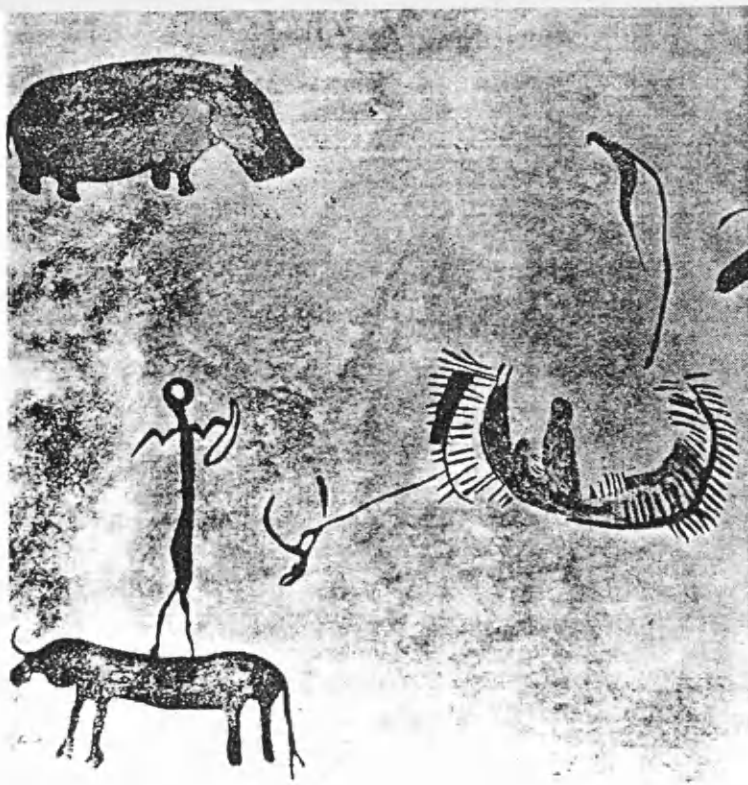
**Figure 1.5** Time series of SST isotherms showing the growth and decay of an El Niño event of the Southern Oscillation. (from Woods, 1984).

of large polar ice sheets.

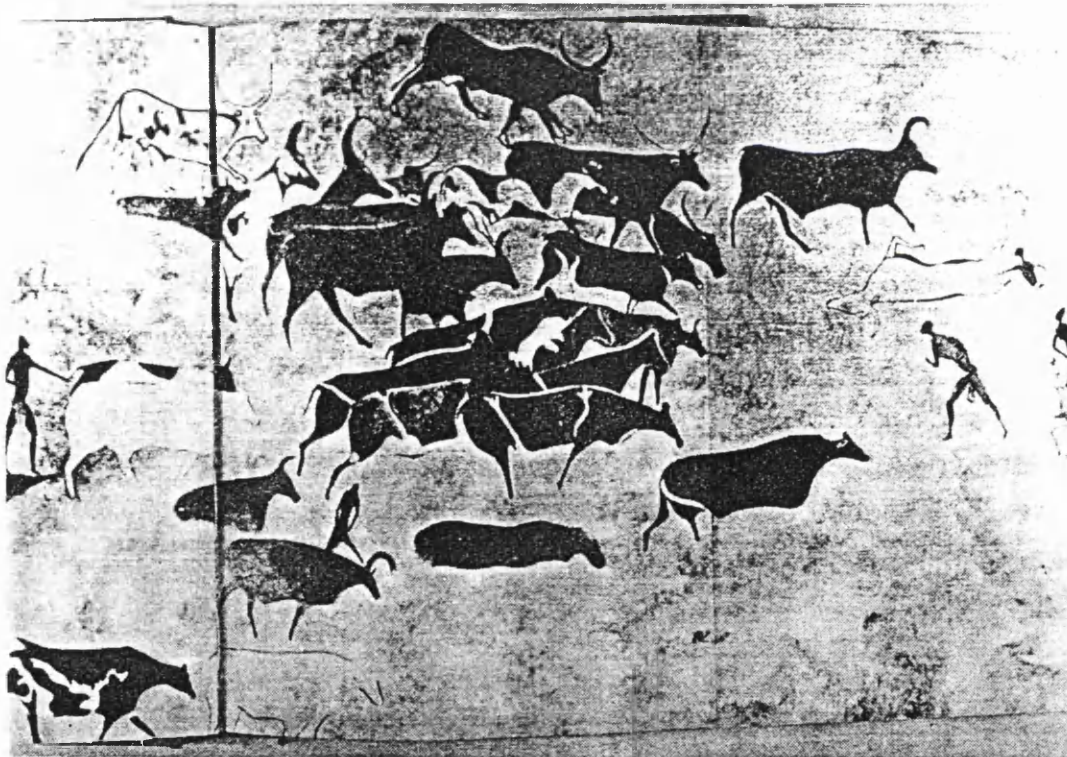
It is important to note the fundamental difference between the prospect of anthropogenically-induced climate change and the vast majority of 'natural' variations that have occurred in the past. In considering the geological evidence, we note that the previous long term changes in climate are not correlated with the species extinction records (see Figures 1.2(a) and 1.2(b)). Even the quite rapid climate change events such as the onset of the glacial cycles some 2 - 3 million years ago did not result in mass extinctions. Perhaps the most striking evidence for recent climatic change are the cave paintings found in Aounrheth in the Sahara (see Figure 1.6) dating from around 3500 BC which depict scenes such as a hippopotamus hunt with canoes, which imply a much wetter climate than today's in the region (Lamb, 1982). Imaging radars flown aboard the Shuttle have also revealed ancient water courses under the sand (Ford *et al*, 1983). As we will see in section 1.2, the existence of the desert zones is a fundamental consequence of the current atmospheric circulation pattern, thus this represents a significant change in the climatic regime. Therefore we see that a climate change spanning even as short a geological time as 10,000 years poses little threat to the human populations because mankind has the ability to adapt to different situations. However, the predicted rise in temperature associated with global warming is capable of reversing the two million years of cooling in just 100 years (Budyko *et al*, 1988) and, although mankind is unlikely to become extinct, the economic, social and ecological consequences are likely to be severe.

Current predictions not only concern a global rise in temperature but also substantial shifts in the different climate zones and associated weather patterns (WCRP, 1990). The frequency and locations of extreme events is likely to change (SWCC, 1990), and it is these that are the most destructive part of the current climate. Human societies are largely dependent on the current climatic regime for their economic and social well-being. Many agricultural policies are based around expected crop yields for the range of climate experienced during the recent past. Also, protective measures against extreme events such as flood barriers and dykes are built with the climatological record of the past century or so in mind. Asian societies have built a dependence on the annual occurrence of the monsoon rains for their economic and agricultural production. There are also a significant number of populations living in marginal areas of the world such as the Horn of Africa and the levels of the Indian sub-continent who are much more vulnerable than most to shifts in the climate zones.





(a)



(b)

**Figure 1.6** Parts of cave paintings of around 3500 BC from Aounrhet in the Sahara, illustrating (a) a hippopotamus hunt with canoes, clearly implying a much wetter climate than today's in the region, and (b) cattle herding. (From Lamb, 1982).

At this point it is helpful to consider some of the major physical processes that operate in the climate system. This exercise will help to put the nature of possible changes into a physical context, as well as providing additional background to the rest of the work.

## 1.2 The Climate System - a brief overview

### 1.2.1 The overall energy budget

If we assume that, on average, the Earth is in thermal equilibrium, the rate at which it radiates energy away to space must exactly balance the rate at which it receives energy from all sources. We can therefore derive a simple expression describing this thermal equilibrium (Goody and Walker, 1972).

The Earth receives virtually all its energy from the Sun. The rate at which a planet absorbs solar energy depends on (1) the solar flux at appropriate distance from the Sun, (2) the albedo of the planet and (3) the cross-sectional area of the planet. This can be expressed as an equation :

$$\text{Energy absorbed} = \pi \times (\text{Radius})^2 \times \text{Solar flux} \times (1 - \text{albedo}) \quad (1.1)$$

We now consider the rate at which a planet radiates energy away into space. For a black body, the *Stefan-Boltzmann law* states that the amount of heat energy radiated by a surface of unit area in unit time is proportional to the fourth power of the temperature. Over the thermal infrared wavelength range, the Earth (atmosphere, clouds, ocean or land) is assessed to have an average emissivity close to 1, and its surface area is  $4\pi(\text{Radius})^2$  so we have :

$$\text{Energy radiated} = 4\pi \times (\text{Radius})^2 \times \sigma \times (\text{Temperature})^4 \quad (1.2)$$

Equating 1.2 and 1.1 and rearranging terms we obtain an expression for the temperature :

$$\text{Temperature} = \sqrt[4]{\frac{\text{Solar flux} \times (1 - \text{albedo})}{4 \times \sigma}} \quad (1.3)$$

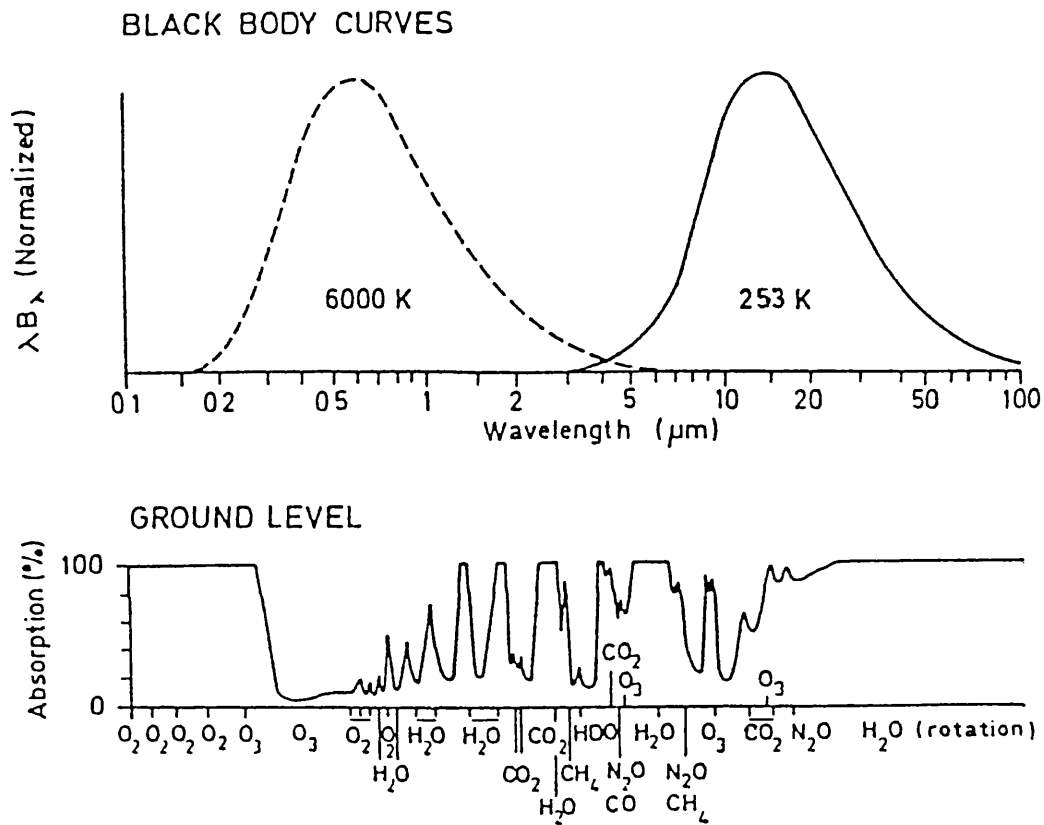
The temperature calculated in this manner is called the *effective temperature* ( $T_e$ ). For the Earth,  $T_e = 253$  K.

It is important to distinguish between the calculated effective temperature of a planet and the actual temperature of the planetary surface. If a planet has a substantial atmosphere, then this can absorb much of the thermal radiation emitted from the planetary surface before the radiation penetrates into outer space and re-emit some of this energy to the surface again. Thus the radiation which finally escapes has been emitted from a higher (and cooler) layer in the atmosphere, and the surface may be substantially warmer. The effective temperature of the planet is thus the temperature of this emitting region, and lower levels may have higher temperatures. On Earth the average surface temperature is 288 K. This phenomenon is known as the greenhouse effect.

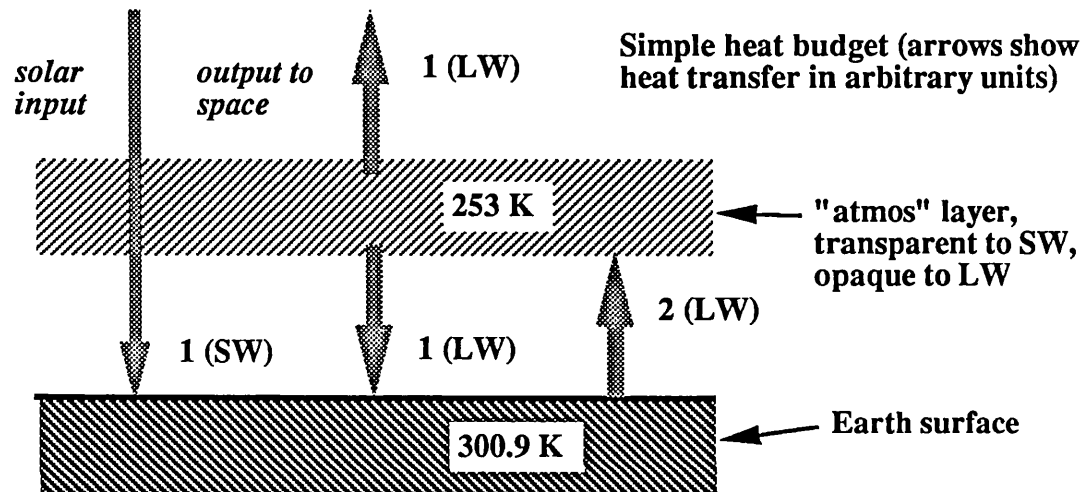
The extent to which the temperature of the surface can differ from the effective temperature depends on the mass of the atmosphere and its constituents. The important property is the difference in opacity of the gas to electromagnetic radiation of different wavelengths. If we compare the wavelengths of incoming solar radiation and thermal infrared emission from the planet, we find that there is virtually no overlap of the Planck curves (Figure 1.7). This is to be expected, since by Wien's Displacement law the wavelength of the peak of the Planck curve is inversely proportional to the temperature of the emitting body. Since the effective temperature of the Earth is  $\sim 1/24$  that of the Sun, the peak wavelength of terrestrial radiation is 24 times greater than solar radiation.

The atmosphere has different absorption characteristics for short wavelength solar radiation and long wavelength terrestrial heat radiation. It can be seen from Figure 1.7 that the atmosphere is reasonably transparent in the visible region of the electromagnetic spectrum, and much of the solar radiation can pass through unhindered. Thermal infrared radiation, however, is strongly absorbed by minor constituents of the atmosphere, particularly water vapour, carbon dioxide, methane and chlorofluorocarbons (the 'Greenhouse' gases), so the atmosphere is largely opaque to planetary heat radiation (Houghton, 1987).

For illustration we now consider a simple model (Figure 1.8), with the surface of the Earth being a black body in both the short-wave range of the incoming solar radiation and the long-wave range of the outgoing thermal radiation, and the atmosphere being represented by a single isothermal layer which is transparent in the short wave range and a black body in the long wavelength range. We can see that although the incoming solar radiation and outgoing terrestrial radiation balance exactly, the earth's surface receives twice the energy it would receive if there were no atmosphere. Thus,



**Figure 1.7** At the top are the Planck curves that show the proportion of energy radiated at each wavelength by the Earth and the solar radiation incident on the Earth. At the bottom is the percentage of the radiation at each wavelength that is absorbed in the atmosphere. Absorption is strong on the average for planetary radiation but weak on the average for the solar radiation. (From Houghton, 1987).



**Figure 1.8** Simple heat budget model illustrating the principle of the greenhouse effect by using an isothermal layer that is transparent to short wave radiation but opaque to long wave radiation to represent the atmosphere.

in order to maintain the energy balance, it must radiate twice as much and in order to do this, the surface temperature is increased to 300.9 K. The actual case is more complicated, with heat exchange taking place between the different elemental layers in the atmosphere. This situation can be described by the theory of radiative transfer (see Chapter 2), and results in a temperature profile in the atmosphere. If the rate of decrease in temperature is greater than the adiabatic lapse rate (the decrease in temperature due to expansion resulting from the decrease in pressure), as is the case in the troposphere, then convection takes place, resulting in a linear temperature profile (Houghton, 1987).

### **1.2.2 Atmospheric circulation**

There are several components of the climate system which interact with each other in a variety of ways (see Section 1.2.3). These are the atmosphere, ocean, land, cryosphere and biosphere (see Figure 1.9). The energy required to drive the interactions between these elements is provided by the Sun. The majority of the solar heat input is absorbed at low latitudes, usually by the ocean (Figure 1.10).

Figure 1.10 also shows that the outgoing thermal radiation is much more evenly distributed around the globe than the solar heat input. Both the ocean and atmosphere serve to redistribute the heat to higher latitudes. The general circulation of the atmosphere is driven by the latitude gradient of heat input. At low latitudes, the warm ocean and land release energy to the atmosphere by evaporative transfer of latent heat, conduction (sensible heat transfer) and radiation of thermal infrared radiation which is mostly absorbed by the various atmospheric constituents. The air near the ground is heated fastest and expands, resulting in convection. Since the logarithm of the pressure decreases with height at a rate inversely proportional to the ground temperature (Houghton, 1987), the proportional pressure difference between low latitudes and higher latitudes increases with height. The resulting imbalance of forces causes air to flow from a region of higher temperature and pressure to a region of lower temperature and pressure. When the air reaches the region of lower temperature, it cools and falls convectively, thus resulting in an increase of air mass near the surface of the cool region. A rise in pressure results and the consequential pressure difference between the low temperature and high temperature regions at this low level in the atmosphere causes a flow of cool air from the low temperature region to the high temperature region, where it is heated again. Such a circulation pattern is known as an Equatorial Hadley Cell (Figure 1.11). Here, the strong surface winds flowing towards the equator are the trade winds.



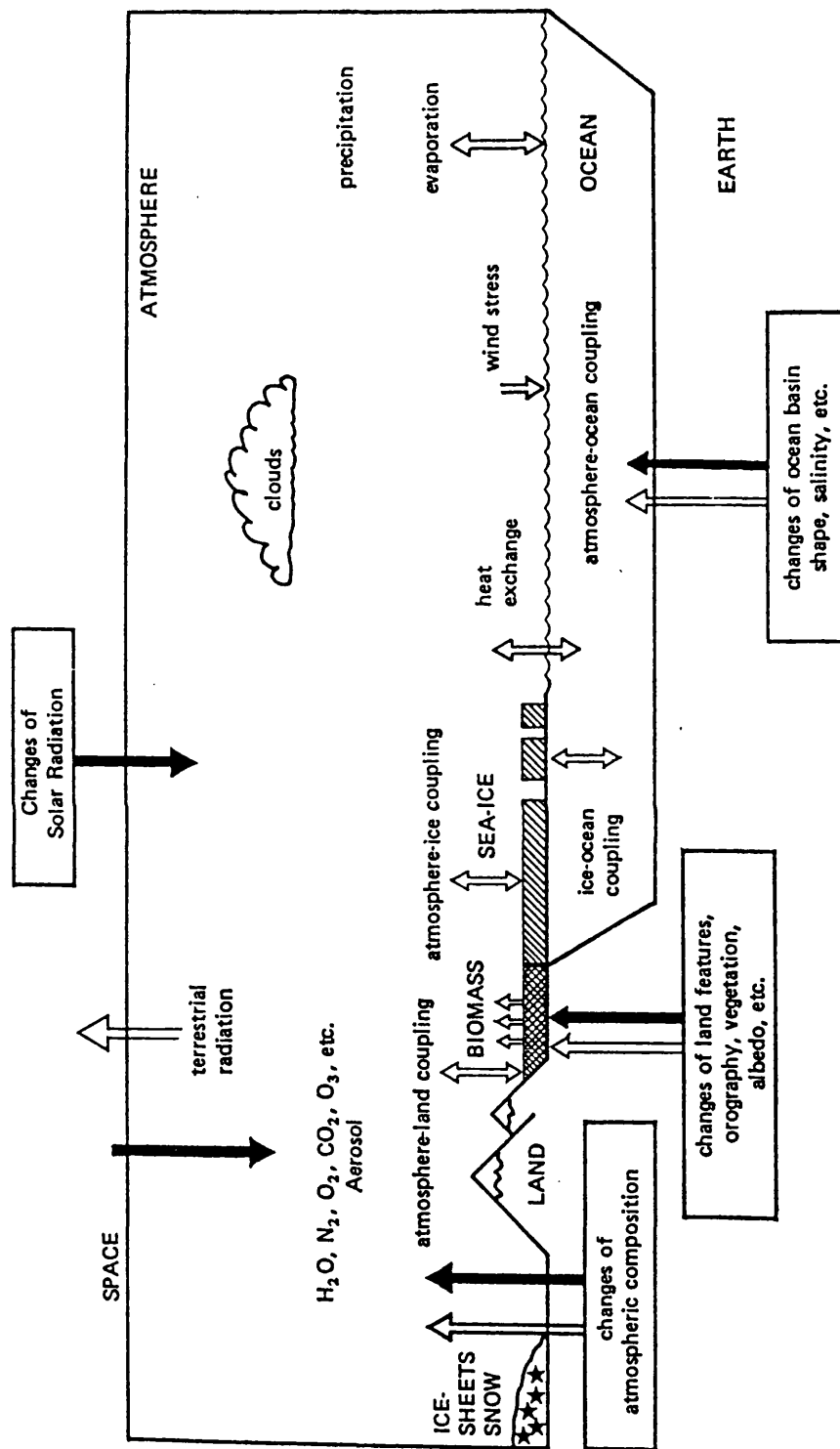
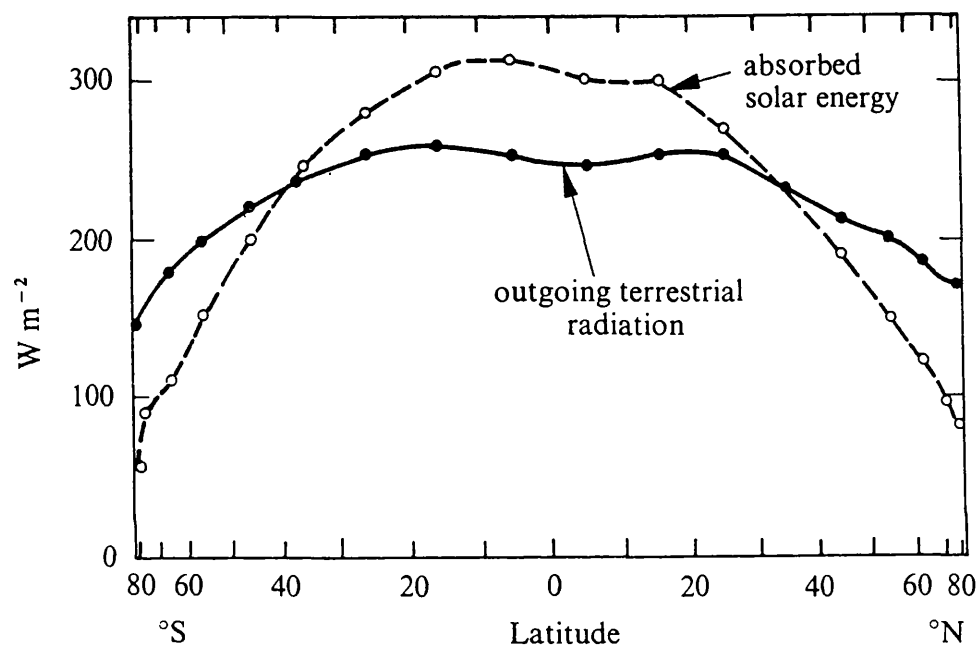
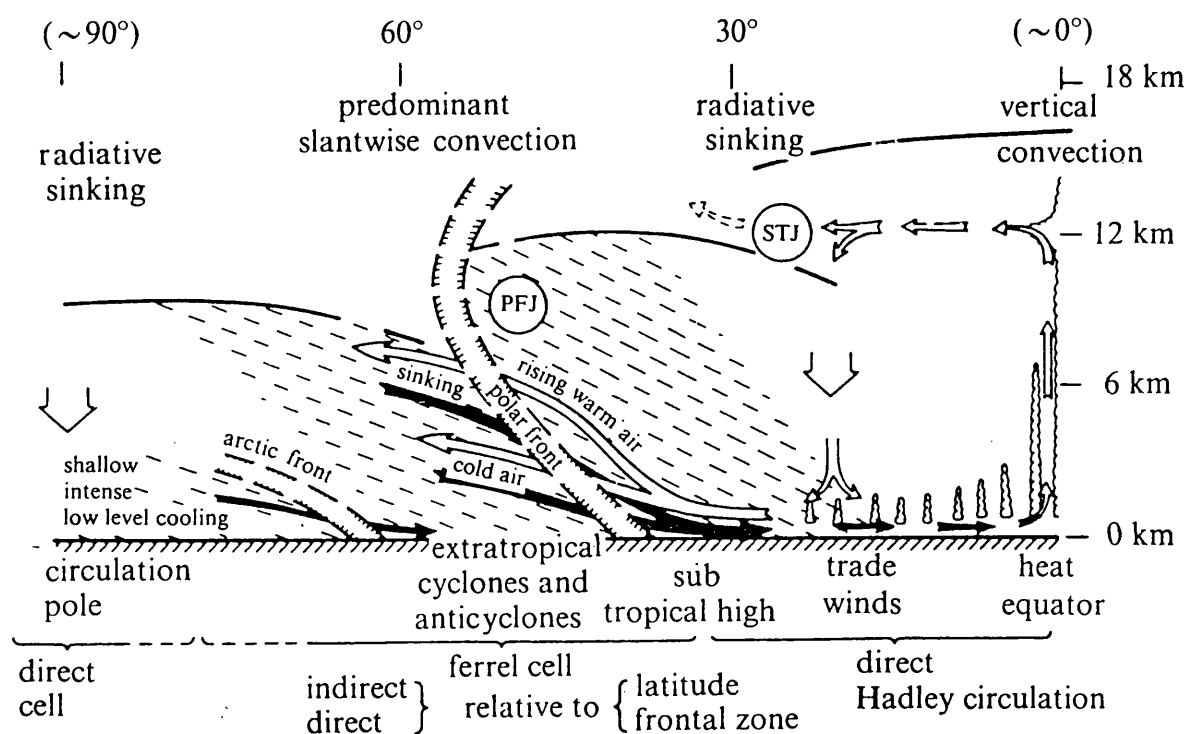


Figure 1.9 Schematic illustration of the components of the earth's climate system. (Houghton and Morel, 1984).



**Figure 1.10** Average components of the earth's radiation budget as deduced from satellite observations, 1962 - 66. (From Houghton, 1987).



**Figure 1.11** Schematic features of the atmospheric circulation in winter, showing Hadley circulation in the tropics, sloping convection at mid and high latitudes and various frontal systems. (From Houghton, 1987).

The Hadley circulation creates distinct zones :

- The Inter-Tropical Convergence Zone (ITCZ) which has convective cloud formation and high precipitation, creating the environment for the tropical rainforests.
- The sub-tropical desert zones, where the air is gradually sinking, thus little condensation or precipitation results.

The actual situation is complicated by, for example, the increased moisture and cloud cover in the tropics, the uneven distribution of land over the globe and the seasonal variations in solar radiation incidence angle due to the axial tilt of the Earth. Heat is, however, transported from the equator to the mid-latitudes by this mechanism, where other processes then affect the circulation and complete the transport to higher latitudes. These other processes are much more complex, and it is felt inappropriate to go into them in any more detail beyond mentioning one of the main reasons for their complexity, which is the Coriolis force. The Coriolis force results from the air's motion being in a rotating frame of reference rather than an inertial one, and thus it appears to be deflected to the right to an earth-bound observer, and is responsible for the strong East - West component in the major trade winds of the tropics and the circulating nature of the high and low pressure systems in mid-latitudes. These winds in turn drive the tropical ocean circulation which can be deflected by the continental land masses and other winds; this process is also important for heat transfer to high latitudes.

### **1.2.3 Rôle of the ocean, cryosphere and land in the climate system**

The other main elements of the climate system are the ocean, cryosphere and land. Their rôles can be briefly summarised as follows.

There are several important points to note about the ocean as regards the climate system. These are:

- The ocean has a very low albedo in the visible and near infrared and therefore absorbs the majority of solar radiation incident upon it.
- Solar radiation penetrates to tens of metres depth, and so the ocean has a large heat capacity readily available for the storage of solar energy.

- The ocean exchanges heat with the atmosphere via the processes of evaporative heat transfer, conduction and thermal infrared radiation.
- However, this heat release is relatively slow, owing to slow upward transport of heat within the ocean and a fairly small ocean/air temperature difference. Typically, changes in SST have time constants of several days.
- Hence heat can be transported by ocean currents before release. For example, it is well-known that the Gulf Stream ameliorates the West European climate.
- The transport of heat from the equator to the poles by the ocean currents is known to be of the same order as that transported by atmospheric circulation processes (Woods, 1984).

There are a number of important points to note about the cryosphere as regards the climate system. These are:

- Ice has a very high albedo in the visible and very near infrared and consequently most of the incident solar radiation is reflected back into space, little energy is absorbed by the ice.
- Sea-ice has the effect of insulating the air from the warmer ocean, thus inhibiting both evaporation and heat transfer (Untersteiner, 1984).
- A further consequence of this is that stable high pressure areas form above the sea ice (eg. in Northern Hemisphere winter) which inhibit atmospheric circulation.

There are several important properties of the land masses which have a bearing on the processes of the climate system. These are:

- The land typically has a relatively low albedo thus the majority of incident solar radiation is absorbed.
- Land is opaque and has a relatively low diffusivity, thus heat from incident solar radiation is stored only in the surface layer. This leads to the temperature of the land rising rapidly during the day, thus providing the heat source necessary for convective processes. Similarly, the heat is given up rapidly at night, leading to a

cooling of the atmosphere near the ground. Work performed in Chapter 3 demonstrates the large diurnal range for land temperatures.

- The relief of various land surfaces forces air to rise and expand, resulting in a cooling from the adiabatic process. This cooling leads to an increase in the relative humidity and, when saturation is reached, precipitation can occur. An example of this is that we expect that precipitation is more likely to occur over highland regions than the lowlands.
- The irregular distribution of land breaks up the basic convection cells and modifies the ocean currents, leading to a non-uniform transport of heat from the equator to the poles.

#### **1.2.4 Feedback mechanisms within the climate system**

The interactions between the various elements of the climate system are generally complex, and there are various feedback mechanisms which operate within it. The most important ones affect the amount of energy flowing into and out of the climate system and several examples are described below.

##### ***The ice - albedo feedback mechanism (a positive feedback)***

- Ice has a high albedo and therefore reflects most of the incoming solar radiation incident upon it.
- Therefore an increase in ice extent results in less radiation being absorbed by the Earth's surface.
- Therefore the temperature of the Earth's surface is lowered.
- Since the temperature has decreased, more ice forms.
- The reverse occurs when the ice cover decreases.

##### ***The sea-ice - heat transport mechanism (positive feedback)***

- Sea-ice forms and insulates the atmosphere from the warm sea.
- The atmosphere cools, leading to a stable pool of cold air which inhibits the convective transport of heat, thus the convection cells do not extend so far towards the poles.
- The inhibition of the convective heat transport process leads to a cooling of the atmosphere adjacent to the sea-ice, resulting in the formation of more sea-ice.
- This process also couples with the ice - albedo feedback mechanism.

### ***The water vapour - radiation feedback mechanism (positive feedback)***

- An increase in the H<sub>2</sub>O content of the atmosphere leads to more thermal infrared radiation emitted from the Earth's surface being trapped by the atmosphere (increased greenhouse effect, WCRP 1990, also see Chapter 2).
- The temperature of the globe rises, leading to more evaporation of water into the atmosphere.

### ***The cloud - radiation feedback mechanism (negative and positive feedback)***

- An increase in cloud cover results from an increase in the amount of water vapour in the atmosphere.
- The relatively high albedo of clouds means that the majority of incoming solar radiation is reflected back out into space (Webster and Stephens, 1984).
- This results in a decrease in the amount of solar radiation absorbed at the Earth's surface, leading to a decrease in temperature.
- Thus the amount of water evaporation decreases, leading to fewer clouds being formed.
- This mechanism acts as a negative feedback component on the water vapour - radiation feedback mechanism described above.
- Some positive feedback also results from the blanketing effect of the cloud on the outgoing thermal radiation emitted from the Earth's surface (eg. Webster and Stephens, 1984, Manschke, 1989, Cusbasch and Cess, 1990), as in the water vapour/radiation feedback.

## **1.3 The Need for Climate Monitoring and Prediction**

### **1.3.1 The variability of climate**

Having briefly reviewed the physical processes that operate in the climate system, we can now review the prospects for climate change. As mentioned previously, there are natural quasi-periodic and other variations in the climate, the most notable of which is the El Niño - Southern Oscillation. Such events are a subject of much research, particularly with respect to their modelling and prediction since they are thought to be responsible for extreme events (WCDP, 1990), but here we are concerned primarily with long term variability that have the potential to produce changes in the global environment. There are two main causes for such variations in the climate. The first

is variation in the amount of solar radiation that is absorbed by the Earth's surface. There are various possible events which could lead to a change in this quantity :

- Variations in the input of radiation from the Sun (Milankovitch variability) which is on a reasonably long time scale of ~100,000 years. It is this variability that is thought to be the driving factor for the ice ages via the ice-albedo feedback mechanism and reduction of the CO<sub>2</sub> content in the atmosphere (eg. Lamb, 1972).
- Shorter term variations in solar radiation output (for example, the 22 year solar activity cycle) which are thought possibly to have some influence on the weather extremes and average annual temperature (eg. Schuurmans, 1980).
- The blocking of solar radiation by aerosols produced by natural events such as volcanic eruptions. For example, El Chichón in 1982 deposited aerosols such as sulphuric acid droplets into the stratosphere that affected the amount of received solar radiation for several years (eg. Budyko *et al*, 1988).
- The blocking of solar radiation by aerosols released into the atmosphere by Man's activities, such as the burning of fossil fuels and forest clearing for agriculture (Berger, 1980).

The second major cause of climatic change is variation in the amount of outgoing thermal infrared radiation being absorbed by the atmosphere ie. variation in the strength of the greenhouse effect. Processes which can lead to this type of change include:

- The release of the so-called greenhouse gases into the atmosphere (see Figure 1.12), for example:
  - CO<sub>2</sub> produced by the burning of forests and fossil fuels
  - Chloro-Fluoro-Carbons (CFC's) released from aerosols, fridges and plastic foams, which are very stable compounds and consequently highly persistent. Although only present in small quantities, CFC's are powerful greenhouse gases and contribute an estimated 20 % of the total increase in the greenhouse effect to date (Watson *et al*, 1990).
  - Methane (CH<sub>4</sub>) released by waste management and agricultural activities.
- The inhibition of the CO<sub>2</sub> removal process by destroying biomass, in particular



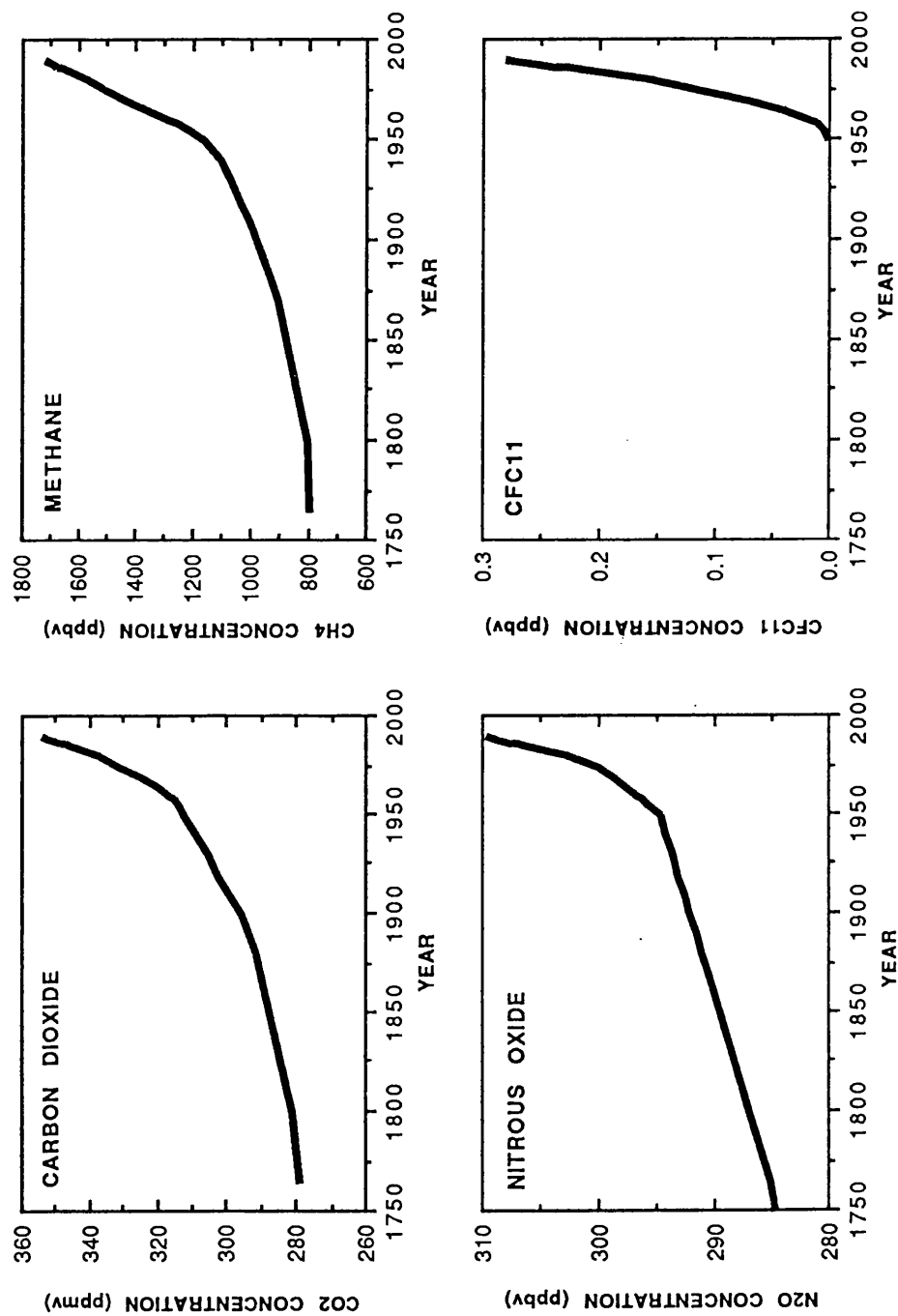


Figure 1.12 Plots illustrating the sharp increases in concentrations of several important greenhouse gases in recent times. (IPCC, 1990).

the rain forest (IPCC, 1990).

The exact evolution of sources and sinks of CO<sub>2</sub> and other greenhouse gases is not clear and represents a significant source of uncertainty in the prediction of climate change (WCRP, 1990).

### **1.3.2 Possible effects of climatic change**

It is generally agreed that the most likely climatic change over the next century or so is a rise in the average temperature of the surface of the Earth and the lower atmosphere due to an increased greenhouse effect resulting in a minimum rise of ~1-2°C (WCRP, 1990). Such a change in the climate would manifest itself in the following ways:

- 1) The thermal expansion of the oceans for a one degree rise in temperature, combined with the melting of inland glaciers could lead to a 2-3 feet rise in sea level (IPCC, 1990).
- 2) Changes in the weather patterns around the globe will occur. Modelling is not yet accurate enough to predict these in detail, although a major effect on the grain producing regions of the USA and USSR as predicted by some simulations (eg. Manabe and Stouffer, 1979) could be disastrous.
- 3) Changes in the extremes of weather might occur eg. more frequent droughts, storms and flooding (IPCC, 1990).

There are, however, some more serious scenarios which could result from a global temperature increase. These include :

- The possible collapse of the West Antarctic ice sheet. The West Antarctic ice sheet is known to be grounded below sea level and could potentially undergo a catastrophic collapse and surge off the continental shelf (eg. Thomas *et al*, 1979). This would lead to a rise in sea level of ~6 metres. At the moment it is thought that the ice shelf is buttressed in by the Ross and Filchner-Ronne ice shelves. Present ice sheet dynamic models indicate that a time scale of 100's of years is more likely for this to happen (Warrick and Oerlemans, 1990).
- The melting of the Greenland ice sheet could lead to a sea level rise of ~7 metres,

although a surge event is unlikely, placing this event on the 1000's of years timescale.

- Melting of the East Antarctic ice sheet would lead to a 60 metre rise in sea level, but again, the ice sheet is grounded above sea level and is unlikely to undergo rapid (100's of years) disintegration.

Any combination of the last three events would have lasting effects; the collapse of the West Antarctic ice sheet and/or the occurrence of the others on a shorter timescale than currently predicted would have disastrous social and economic consequences for both developing and developed nations. Such events must be considered a potentially serious risk since the important ice-albedo feedback mechanism leads to greater temperature rises at the poles (see Figure 1.13).

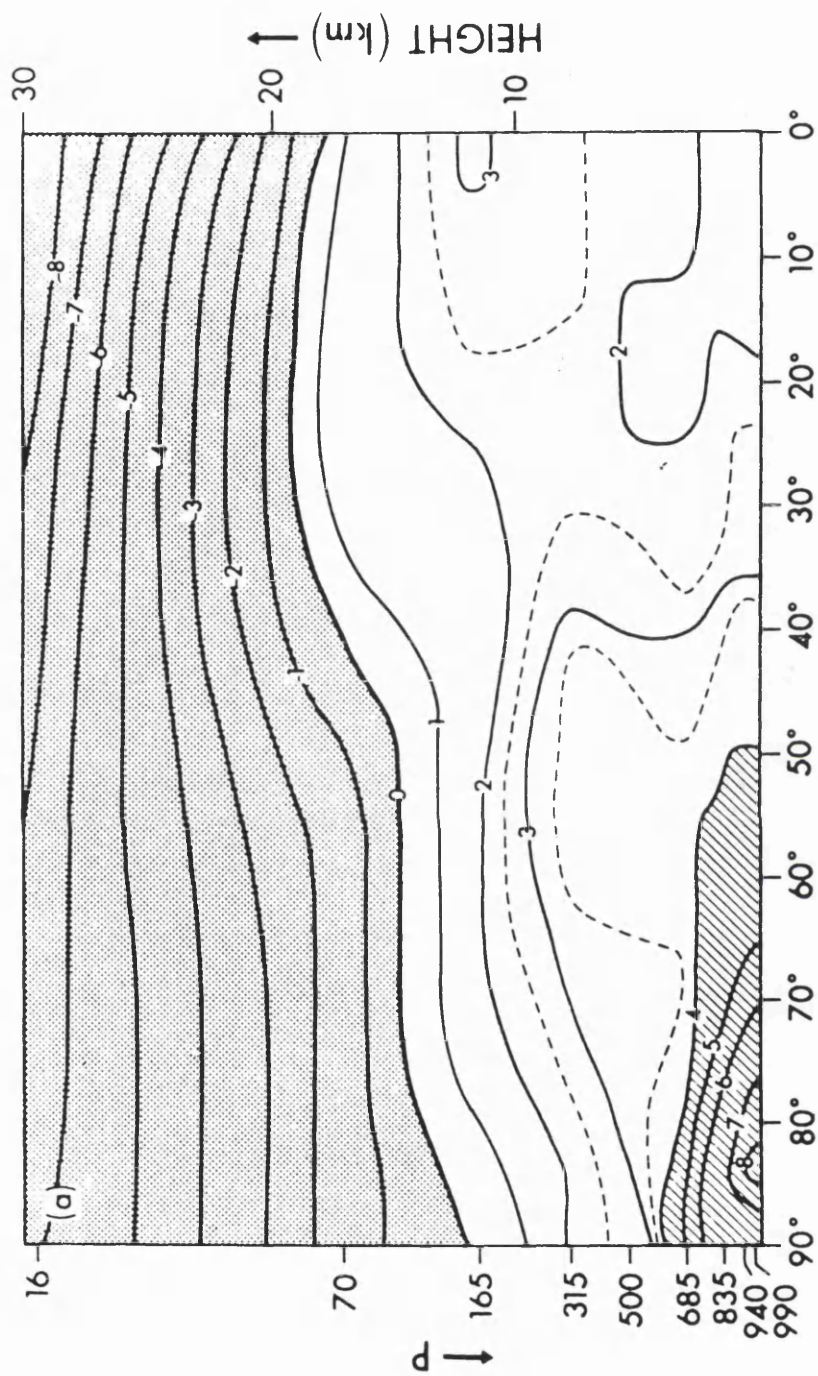
### **1.3.3 The World Climate Research Programme**

Concern for the prospect of global climate change and its possible outcomes resulted in the instigation of the World Climate Research Programme (WCRP). The WCRP developed from the Global Atmospheric Research Programme (GARP) and like GARP is jointly sponsored by the International Council of Scientific Unions (ICSU) and the World Meteorological Organisation (WMO). The objectives of the WCRP are to determine (a) to what extent climate can be predicted and (b) the extent of man's influence on climate. The programme is primarily concerned with changes which occur on timescales from several weeks to several decades. This timescale is consistent with the feasibility of obtaining comprehensive data sets, the practicalities of numerical modelling, and the major concerns of planners and decision makers.

Three objectives or streams of research in the WCRP have been identified, namely (1) long-range weather predictions over periods of several weeks, (2) interannual variability of the global atmosphere and the tropical oceans over periods of several years, (3) longer term variations. Two major experiments, the TOGA (Tropical Oceans and Global Atmosphere) and the WOCE (World Ocean Circulation Experiment) have been identified as the foci of the second and third streams. For these, new modelling efforts (especially coupled atmosphere/ocean models) and global observations, especially from satellites, of all components (in particular the ocean) of the climate system are required.

### **1.3.4 Climate Modelling**

It is clearly desirable to be able to predict the behaviour of the climate system under



**Figure 1.13** Changes in temperature with respect to altitude and latitude resulting from a doubling in the atmospheric  $\text{CO}_2$  content. Note that the greatest temperature rises will occur at the poles. (Houghton and Morel, 1984, after Manabe and Wetherald, 1980).

the changes in the environment that have been and are still are occurring. Although alternatives exist (eg. paleoclimate reconstruction, Cusbasch and Cess, 1990), the study of the behaviour of the climate system can best be performed by using numerical models because of its complexity.

- Numerical models are an ideal tool for studying the evolution of the climate system, since external influences can readily be changed within them.
- These numerical models have as a basis for their calculations the so-called set of 'primitive equations' which describe heat transfer, momentum exchange, moisture conservation, mass conservation and hydrostatic equilibrium.
- The common input for all climate models is the incident solar flux. Above this, models have the ability to change various factors such as CO<sub>2</sub> content. Atmospheric General Circulation Models (GCM's) regard the Sea Surface Temperature (SST) as a specific external influence. Such models are used for short- and medium- range weather forecasting and are applicable to stream one studies of the WCRP. Models which keep SST as internal to the model are known as Coupled Circulation Models (CCM's) or 'Full-Climate' models. These models are best suited for studying climate variations on the timescale of a few years such as El Niño, where only the tropical ocean is modelled (ie. TOGA - Tropical Ocean Global Atmosphere), and are used in stream two studies (Houghton and Morel, 1984). More complete modelling of the ocean circulation enables these models to be used for stream three studies such as the World Ocean Circulation Experiment (WOCE).
- The final goal of the modellers is to produce a fully coupled atmosphere/ocean/cryosphere/land model which will be more deserving of the title of a 'Full Climate Model'. Such a model is suited to stream three studies where the climate is modelled over a period of tens to hundreds of years.

The models explicitly calculate the evolution of weather patterns over the globe. The calculations are performed over as fine a resolution as today's supercomputers will permit such that the result is obtained within the desired time. For the purposes of weather forecasting, the data that is input to the models is updated every day, thus the calculations must be performed within this timescale. Since GCMs allow the computation to be performed faster than the models which interactively couple the

atmosphere with the ocean, and because the effect of ocean-atmosphere coupling is minimal over short timescales (see Section 1.2.3), these are the models used for weather forecasting.

There are various input parameters for the climate models. Those required for the General Circulation models include :

- The Sea Surface Temperature, obtained from ships, floating bouys and, increasingly, from satellite-borne radiometers.
- The measurements of surface humidity, air temperature and wind velocity from land-based meteorological stations and some ships.
- Vertical water vapour and temperature profiles of the atmosphere measured by meteorological balloon ascents (known as radiosondes, see Chapter 3 for a more complete description), and similar information derived from satellite-borne microwave sounders.

Such parameters are predicted by the Coupled Circulation models, but these measurements can be used to update such models and provide a check on the predictions made.

#### **1.4 The use of Remote Sensing data in Climate Research**

In addition to ground-based sources of data, remote sensing can provide valuable data for the climate research effort. Input data for models can be provided to determine initial and boundary conditions. An example of such data is the Sea Surface Temperature that can be obtained from satellite-borne radiometers such as the Advanced Very High Resolution Radiometer (AVHRR) carried on board the National Oceanic and Atmospheric Administration's (NOAA) series of satellites and the recently launched Along-Track Scanning Radiometer (ATSR) carried on board the European Space Agency's European Remote Sensing-satellite 1 (ERS-1). The advantages of remote sensing are that it provides frequent and regular sampling, it comes from a single source and the data can be processed to ensure that it is self-consistent and, hopefully, accurate. It is possible for satellite-borne instruments to provide both a direct measure of climate (eg. the air temperature obtained from a microwave sounder) and a proxy measure, that is, measurement of some other parameter which is affected by climate.

#### 1.4.1 The use of proxy indicators as monitors of climatic change

There are various phenomena that are affected by climatic change which can be measured by satellite:

- The sea ice extent is very sensitive to the mean winter air temperature due to the positive feedback mechanisms described earlier. Satellite-borne instruments such as passive microwave, Radar Altimeters (RA's) carried on Seasat, Geosat and the recently launched ERS-1 satellite and visible and near-infrared imagers can all be used to monitor the sea ice extent (Laxon, 1989).
- The mass and extent of the ice sheets is a measure of the climate. The equilibrium (or non-equilibrium) of mass balance can in principle be monitored by instruments such as RA's (Partington, 1988). Information on the ice extent can also be provided by imagers such as Multi-Spectral Scanner (MSS), AVHRR and Synthetic Aperture Radar (SAR).
- Lakes respond in various ways to climatic variations. For closed lakes (those without surface outlet), the volume fluctuations represent a measure of the aridity (evaporation/precipitation) of the region (Mason *et al*, 1985). Also, work done in Canada (Tramoni *et al*, 1985) and elsewhere shows a relationship between the freezing/thawing dates of lakes and the mean air temperature. It should be possible to monitor the icing of large lakes from space. Monitoring of the whole seasonal temperature cycle of a lake can provide information as to the mean annual air temperature (Shuter *et al*, 1983). Again, it should be possible to monitor lake temperatures from space (Brown *et al*, 1991).

The work described in the following chapters is related to finding techniques for monitoring lake area, lake temperature and for improving the retrieval of SST (see Chapters 5 and 6), particularly in tropical regions where the atmospheric correction is high (see Chapter 2, also Barton, 1983) and is also where the majority of solar heat input occurs (see Section 1.2.1). The lake areas work described in Chapters 3 and 4 can be combined with lake level measurement (Rapley *et al*, 1985) to provide the volume changes required for climate monitoring. The data used for measuring lake area and temperatures comes from the AVHRR instrument on the NOAA series of satellites. This data is similar to that provided by the ATSR which is carried on board the recently launched ERS-1 satellite. Both these instruments are described more fully in Chapter 2. Since ERS-1 also carries a radar altimeter, this will allow the synergistic use of both instruments over lakes, thus providing measurements of

volume change.

In summary, the goals of the work presented in this thesis are to develop new techniques for the remote sensing of climate related variables and to demonstrate improvements over existing techniques, where such techniques exist. In this way, it is hoped to provide new tools for the monitoring of the global climate using remotely-sensed data, thus further enabling the exploitation of data provided by the increasing number of new and improved satellite-borne instruments that are being launched, together with improved processing of existing data archives, thus providing the longer time series of measurements required for global climate studies.



## **Chapter 2**

# **An Introduction to Infrared Radiometry from Space**

### **2.1 Overview**

There is considerable potential offered by remote sensing for the global monitoring of the effects of climate and climate change, as outlined in the previous chapter. However, before we can begin to interpret remotely-sensed infrared data of the planet's surface, it is important to consider some physical mechanisms that have contributed to the signal that reaches the instrument. It is also instructive to investigate the operation of the instruments themselves, in terms of the way that the signal is both detected and calibrated. There are therefore several areas to consider and these are outlined below and a fuller discussion of some of these topics follows in the rest of this chapter.

Remote sensing of surface parameters from space-borne instruments involves the measurement of electromagnetic radiation which has undergone interaction with the atmosphere:

- Radiation can be absorbed by the clear atmosphere through interaction with various atmospheric constituents.
- The atmosphere can also emit radiation, particularly at long wavelengths in the thermal infrared region of the spectrum.

In addition, radiation can be scattered by aerosol particles and molecules, and also interacts with clouds, which reflect short wavelengths and absorb long wavelength fluxes.

It is clearly important to understand the nature of electromagnetic radiation itself, particularly black body radiation and the role of emissivity. It is also necessary to discuss the essential elements of the radiometers themselves, including:

- Optical system for generating the image, collecting the flux and concentrating it at the detector.

- The operational characteristics and requirements of the detectors themselves.
- The absolute calibration of the signal produced by the detector in response to the flux level.

It is also instructive to review current techniques that are employed to compensate for the effects that the atmosphere has had on the radiation emitted from the surface, for example, the methods used to correct for the change in perceived radiance introduced by the atmosphere. The identification and elimination of radiances which have been contaminated by clouds is also of interest, but further discussion on this topic is held over until chapter 6, where a cloud clearing scheme (Saunders and Kriebel, 1988) is employed in the analysis.

In addition, there are other effects that have some impact when attempting to extract climatically useful physical quantities, in particular the skin effect (eg. Robinson *et al*, 1984, Hepplewhite, 1989, Schluessel *et al*, 1990) causes problems for retrieval of bulk water temperatures; further discussion of this particular phenomenon is delayed until Chapter 6.

## 2.2 The nature of electromagnetic radiation

All bodies with a temperature greater than zero Kelvin emit electromagnetic radiation. For a perfect, or black body, radiator, the spectral distribution of the intensity of this radiation can be determined by the Planck distribution function:

$$B_{\lambda}(T) = \frac{2hc^2}{\lambda^5} \frac{1}{\exp(hc/\lambda kT) - 1} \quad (2.1)$$

where  $\lambda$  is the wavelength,  $T$  is the temperature of the body,  $h$  is Planck's constant,  $c$  is the speed of light and  $k$  is the Boltzmann constant. In general, bodies are not perfect radiators of electromagnetic radiation, and will emit somewhat less radiation than this for a given temperature. The actual value of the radiance can be calculated by multiplying the Planck exitance as calculated from the above relationship by the emissivity of the body. A black body therefore has an emissivity of 1 and, by comparison, water has an emissivity of ~0.99 over the wavelength range 3 - 14  $\mu\text{m}$ .

We now compare the Planck distribution curves for the temperature of the Sun (~6000 K) and the Earth (~285 K) as shown in Figure 2.1. Note that the radiation

from the Sun has been reduced by the inverse square law over the distance from the Earth to the Sun but the areas under the two curves are not the same unlike the case for the original discussions of effective temperature in Chapter 1. However, the total amounts of energy received and emitted must be equal to ensure that the Earth remains at a stable temperature. The reason for the areas being substantially different is that we are considering the flux intensity per square metre at the surface. Thus the Earth is only receiving this intensity over an area of  $\pi R^2$  whereas it is emitting from an area of  $4\pi R^2$ . In addition, about 33% of the incoming solar radiation is reflected (Goody & Walker, 1972), and the figure of  $70 \text{ Wm}^{-2}$  is the outgoing longwave flux after absorption by the atmosphere and does not include emission from the atmosphere. In addition, there is no consideration of convection and evaporation heat transfer processes in the atmosphere. It is also important to note that there is very little overlap between the curves and we can consider them independently.

The effects of absorption and scattering on incoming short-wave solar radiation are shown in Figure 2.2. Thus we can see that only a fraction of incoming short wave radiation reaches the ground and is absorbed. The Meteosat visible image shown in Figure 2.3 demonstrates the distribution of clouds and other high albedo sources at the solar radiation wavelengths ( $0.4 - 1.1 \mu\text{m}$ ).

The total power emitted per square metre for a perfect radiator is the integral of the Planck function which results in the Stefan-Boltzmann equation:

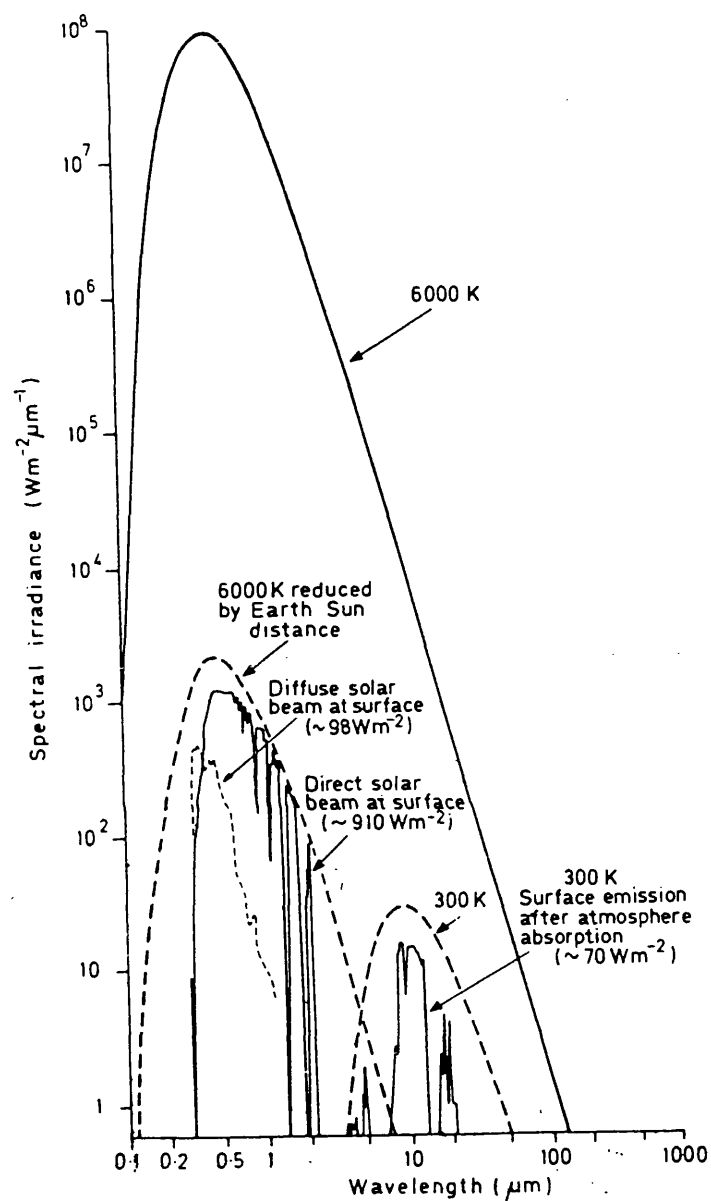
$$E = \sigma T^4 \quad (2.2)$$

The peak of emission for a given temperature can be found by setting the derivative of the Planck function to zero, resulting in the Wien displacement law :

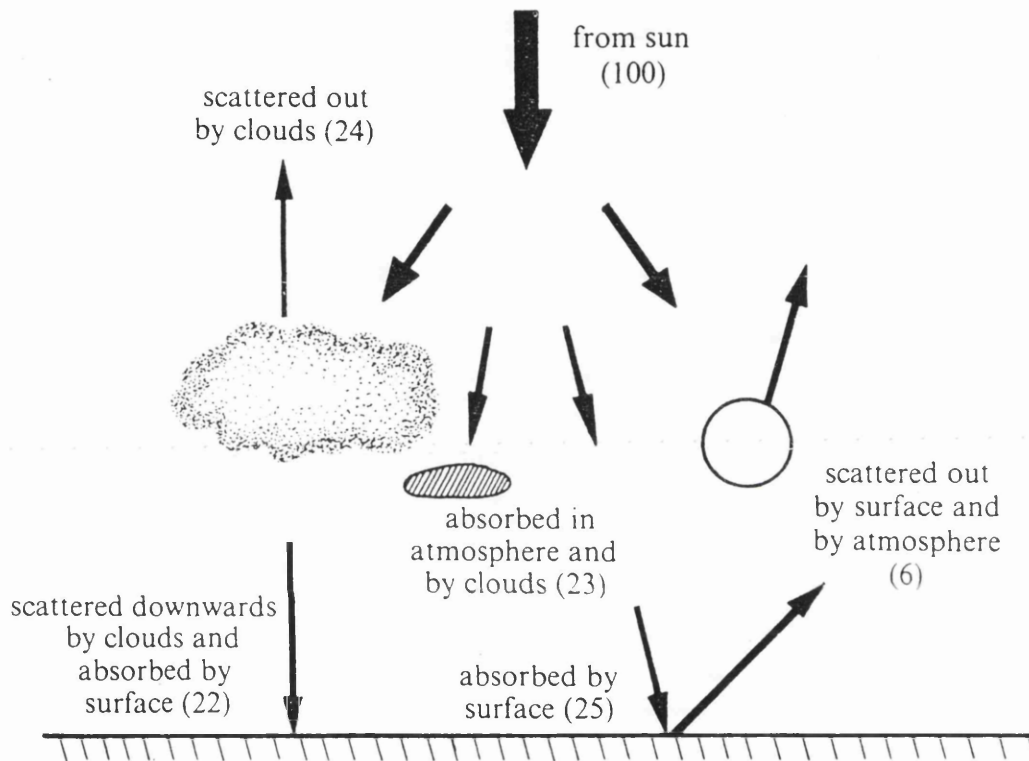
$$\frac{dI}{d\lambda} = 0 \quad \text{for } I_{\lambda \max} \quad (2.3)$$

$$\Rightarrow I_{\lambda \max} = A/T \quad (2.4)$$

where A is a constant of value  $2898 \mu\text{m K}$ . From this we can see that the peak of solar radiation is  $\sim 0.5 \mu\text{m}$ , and the peak of terrestrial radiation is  $\sim 10 \mu\text{m}$ . We know that the atmosphere is relatively transparent to short wave radiation, but it also allows



**Figure 2.1** Spectral irradiance curves for the Sun (6000 K), the solar flux at the mean Earth - Sun distance and the Earth. The effect of atmospheric absorption on both solar and terrestrial fluxes is also shown. The diffuse curve illustrates the additional scattering and absorption by clouds. (From Schott and Henderson-Sellers, 1984).



**Figure 2.2** The interactions of incident solar radiation with the Earth and its atmosphere. (From Houghton, 1987).



Figure 2.3 Full disc Meteosat band 1 (0.4 - 1.1  $\mu\text{m}$ ) image, illustrating various albedo sources on the surface and in the atmosphere. Data courtesy of Imperial College.

a fair quantity of long wavelength radiation in the 8 - 14  $\mu\text{m}$  region to escape, thus there is a reasonable quantity of flux available to be measured by space-borne sensors in both the short and long wavelength regions. An interesting question remains, however, as to the nature of the physical processes that cause these 'windows' in the electromagnetic spectrum.

### **2.3 Absorption processes in the atmosphere**

The absorption process involves the interception of photons by atoms and molecules which are capable of changing their energy state to a higher one by the addition of exactly the amount of energy contained in the photon. The main atmospheric constituents that hinder the passage of solar radiation are oxygen and ozone (particularly in the ultraviolet region), whereas outgoing terrestrial radiation is absorbed mainly by carbon dioxide and water vapour.

Ozone is formed in the stratosphere and mesosphere by photochemical processes, peaking in concentration at  $\sim 25\text{km}$  (Houghton, 1987). Ultraviolet radiation is absorbed by ozone in the Hartley band (200 - 300 nm); below  $\sim 70\text{ km}$  most of the energy of the incoming photons is used to increase the kinetic energy of the molecules. The absorption spectrum is effectively a continuum over this wavelength range, and this is the main heating mechanism for the stratosphere. The actual absorption process is one of photo-dissociation, where a single oxygen atom is liberated, later to recombine via one of several different mechanisms (Goody and Walker, 1972). The stratopause represents the peak of this absorption and resultant heat input at about  $\sim 50\text{km}$ .

The atmospheric constituents are more or less effective as absorbers of photons of various frequencies and energies across the spectrum. For a particular frequency, the effect of a certain constituent can be calculated by using the equation for simple exponential absorption, assuming that the absorption cross-section does not change along the path length. In practice, however, we need to treat the individual pressure and temperature levels separately to establish the absorption cross-sections before calculating the total column absorption because:

- The mixing ratios will not be constant, particularly in the case of water vapour
- Line shapes will be collision broadened by different amounts

- Doppler broadening can occur at higher levels in the atmosphere
- Lines of absorption overlap, particularly at the lower levels.

In the infrared, the absorption coefficient in molecular bands varies rapidly with frequency. There are many different discrete energy states for each molecule, resulting from the many different rotational and vibrational modes that are possible. The many different lines are also subject to various broadening mechanisms, such as Doppler broadening at the high mesospheric temperatures and collisional broadening at high pressures. These effects are dependent on temperature and pressure, and consequently vary throughout the atmosphere. In order to calculate the average transmission over a spectral region (eg. an atmospheric window), we need to integrate over frequency and pressure, with varying temperatures and concentrations of constituents at each pressure.

It is relatively easy to model a molecular band if we assume that the lines do not overlap and that only collision broadening processes are operating. Such assumptions are valid, for example, for carbon dioxide and water vapour between 20 and 60 km altitude. Below 20 km, the increased concentrations and pressure broadening result in line overlap. It is helpful to note the simple equation for the absorption coefficient at frequency  $\nu$  of a collision-broadened line centred at frequency  $\nu_0$  (Houghton, 1987):

$$k_{\nu} = \frac{s\gamma}{\pi\{(\nu - \nu_0)^2 + \gamma^2\}} \quad (2.5)$$

where  $s$  is the line strength (ie.  $s = \int k_{\nu} d\nu$ ) and the half-width,  $\gamma$ , is  $(2\pi t c)^{-1}$  where  $t$  is the mean time between collisions and consequently varies with pressure. The variation of half-width with pressure can be written as  $\gamma = \gamma_0(p/p_0)$  where  $p_0 = 1$  atmosphere and  $\gamma_0$  is typically 0.1 wavenumbers. Thus for a gas, density  $\rho$ , path length  $l$ , we have a transmission  $\tau_{\nu}$  of :

$$\tau_{\nu} = \exp(-k_{\nu}\rho l) \quad (2.6)$$

and the equivalent width or integrated absorbance,  $W$ , of the line is the integral of  $\tau_{\nu}$  over all  $\nu$  :



$$W = \int_{-\infty}^{\infty} dv(1 - \tau_v) = \int_{-\infty}^{\infty} dv\{1 - \exp(-k_v \rho l)\} \quad (2.7)$$

We can make two useful approximations for  $W$ :

- The weak approximation, where second order and higher terms are eliminated from the exponential, giving:

$$W = s\rho l \quad (2.8)$$

- The strong approximation, which renders  $\gamma^2$  insignificant in the broadening term, thus:

$$W = 2(s\gamma\rho l)^{1/2} \quad (2.9)$$

Taking the logarithm of both sides, we see that for (2.8) :

$$\log W \propto \log(s\rho l) \quad (2.10)$$

and for (2.9) :

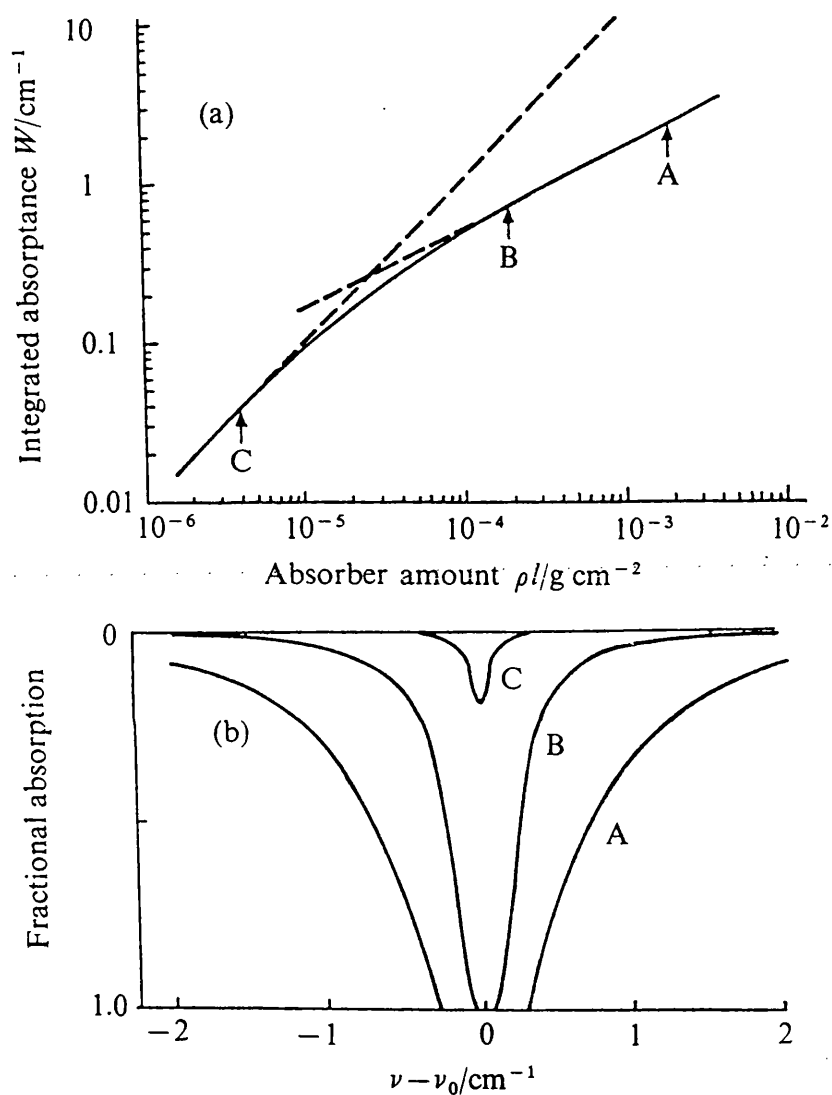
$$\log W \propto \{\log(\gamma s\rho l)\}/2 \quad (2.11)$$

Thus there are two distinct linear relationships, with a transition region in between. This is known as the curve of growth of a line (Figure 2.4). Note that the relationship does not account for changes in the half-width due to density changes.

For spectral bands containing a number of non-overlapping lines, we can deduce an average transmission over the interval  $\Delta\nu$ :

$$\tau = 1 - \frac{\sum W_i}{\Delta\nu} \quad (2.12)$$

The situation for a vertical path through the atmosphere is more complicated because



**Figure 2.4** (a) Curve of growth of a typical spectral line showing the linear and square root regions of growth. (b) Actual shapes of the line profile for different absorber path lengths corresponding to those marked in (a). (From Houghton, 1987).

of the density, pressure and temperature changes along the path. Between pressures  $p_1$  and  $p_2$  we have:

$$W = s \int c \rho dz \quad (2.13)$$

for a weak absorber, with  $c$  the fractional content by mass of the absorber responsible for the line and  $\rho$  the density of the atmosphere. Now  $\rho dz = -dp/g$ , and assuming a constant mixing ratio:

$$W = sc(p_1 - p_2)/g \quad (2.14)$$

If the absorption is stronger, the variation of the half-width with pressure becomes important. We then define a mean pressure  $\bar{p}$  for the path:

$$\bar{p} = \frac{\int p c \rho dz}{\int c \rho dz} \quad (2.15)$$

This is known as the Curtis-Godson approximation. This approximation is an exact one for the case of the strong approximation (Houghton, 1987), and is also exact for the weak approximation where  $\gamma$  is independent of pressure. For a constant mixing ratio, we can assume that  $\rho_{abs} = c\rho_{tot}$  and  $\bar{p} = (p_1 + p_2)/2$ . Thus integrating  $\rho$  over height gives  $c(p_1 - p_2)/g$  and bringing in the relationship for the half-width dependence on pressure (ie.  $\gamma = \gamma_0(p/p_0)$ ) and the strong approximation (2.9), we have:

$$W = 2 \left[ \frac{sc\gamma_0}{2p_0g} (p_1^2 - p_2^2) \right]^{\frac{1}{2}} \quad (2.16)$$

The average transmission over a spectral interval  $\Delta\nu$  for one absorber of concentration  $c$  is then:

$$\tau_v = 1 - \left[ \frac{2c}{p_0 g} (p_1^2 - p_2^2) \right] \frac{\sum (s_i \gamma_i)^{\frac{1}{2}}}{\Delta v} \quad (2.17)$$

As long as the transmission is high, this expression can be used calculate the transmission for any given atmospheric path.

The above derivation for transmittance over any particular frequency band is only valid for the case of non-overlapping lines. For the atmosphere it can only be applied to the path lengths and pressures above about 30 km altitude (Houghton, 1987). Various band models have been developed which allow for the overlap of spectral lines and it is worth describing two of them briefly.

The regular model was first proposed by Elsasser, and consists of an absorption band which is simulated by an array of equally spaced lines of equal strength and shape. The random model, proposed by Goody, consists of lines which are spaced randomly and whose line strengths follow some statistical law. The latter model results in a simple and useful expression for the average transmission  $\tau$  of a spectral interval of width  $\Delta v$  containing a large quantity of lines:

$$\tau = \exp(-\sum W_i / \Delta v) \quad (2.18)$$

where  $\sum W_i$  is the sum of the equivalent widths of all the lines in the interval  $\Delta v$ .

Another important consideration, particularly in the 8 - 14 $\mu$ m region of the infrared where the peak of terrestrial radiation lies, is that of continuum absorption. Although the atmosphere is reasonably transparent in this region, there are still some water vapour lines present, and also the presence of water vapour dimers. The many rotational and vibrational states of such molecular pairs, together with a low energy required to dissociate them, results in an effective continuum absorption. This absorption coefficient is dependent on the partial pressure of the water vapour since the dimer associations are more likely to occur when the partial pressure is high. Thus continuum absorption is of particular importance in humid atmospheres eg. in the tropics, where transmission through a vertical atmospheric path may only be 30% in this wavelength region.

## 2.4 Radiative transfer

Having considered the various atmospheric absorption processes that operate in the thermal infrared part of the spectrum, we might assume that it is a fairly simple matter to calculate the effect that the atmosphere will have had on the radiance perceived by the satellite sensor. If the atmosphere was much colder than the surface this would be a reasonable assumption. However, since the atmosphere is warm, in fact at a similar temperature to the surface, it is also radiating in the infrared at similar wavelengths to the ground-emitted radiance. So, not only does the atmosphere absorb some of the surface radiance, it also contributes to the signal observed from space. There is normally a net deficit since the mean brightness temperature of the atmosphere is lower than that of the surface.

Calculation of the effect of the intervening atmosphere on the perceived radiance can be simplified by the assumption of a plane-parallel atmosphere. Here, the atmosphere is assumed to be uniform in the horizontal dimensions in terms of pressure, constituency and temperature. The atmosphere is also assumed to be in thermodynamic equilibrium such that the vertical structure also remains constant with time and heat transfer is sustained only by the absorption and emission of infrared radiation.

These assumptions allow us to consider transfer of radiation only in the vertical dimension, since the net horizontal component of emitted atmospheric radiation must be zero. If we first consider the law of absorption (Houghton, 1987) which states that the absorption which occurs when radiation of radiance  $I$  ( $\text{Wm}^{-2}\text{sr}^{-1}$ ) traverses an elemental slab of atmosphere of thickness  $dz$  is proportional to the mass of the absorber  $\rho dz$  in unit cross-section of the slab where  $\rho$  is the density of the absorber and to the incident radiation itself:

$$dI = -I k \rho dz \quad (2.19)$$

where  $k$  is the absorption coefficient. Integrating (2.19) gives :

$$I = I_0 \exp(-\int k \rho dz) \quad (2.20)$$

for ground radiance  $I_0$ . The quantity  $\exp(-\int k \rho dz)$  is known as the fractional transmission,  $\tau$ , of the path, and  $\int k \rho dz$  is known as the optical path  $\chi$ .

The elemental layer of atmosphere will also emit radiation in an amount depending on

its temperature and effective emissivity. Under conditions of thermodynamic equilibrium, application of Kirchoff's law enables the amount emitted per unit area to be written as  $k\rho dzB(T)$  where  $B$  is the Planck function.

Although the Earth's atmosphere is not precisely in thermodynamic equilibrium, conditions of local thermodynamic equilibrium, or LTE, exist in the lower atmosphere such that the approximation to equilibrium allows black body emission to be utilised in the equation. The equation for radiative transfer through the slab, including both absorption and emission, is sometimes known as Schwarzschild's equation:

$$dI = -Ik\rho dz + B(T(z))k\rho dz \quad (2.21)$$

where  $T(z)$  is the temperature of the atmosphere at height  $z$ . This can be considered in terms of the change in optical path as identified in (2.20):

$$dI/d\chi = I - B(T(z)) \quad (2.22)$$

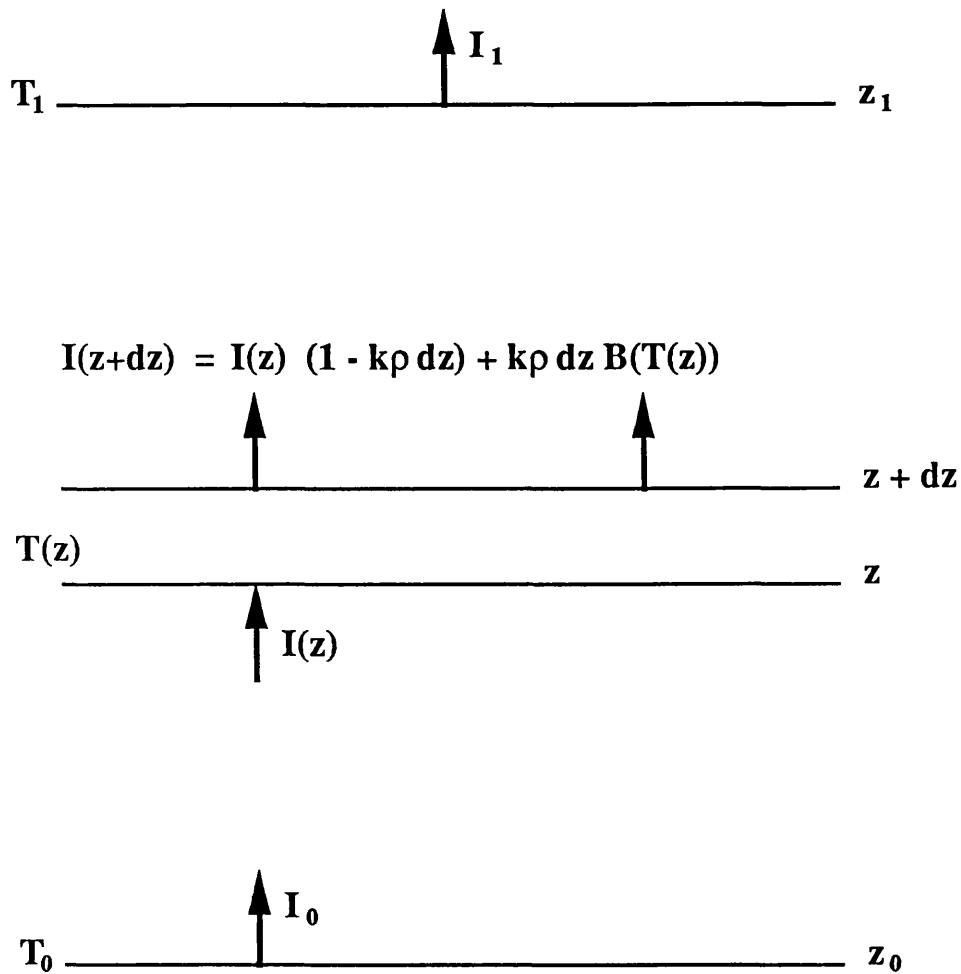
Integrating this equation yields:

$$\begin{aligned} I &= \int I \exp(-\chi) d\chi - \int B(T(z)) \exp(\chi) d\chi \\ &= I_0 \tau(0, \chi_0) + \int B(T(z)) d\tau(z, z_1) \end{aligned} \quad (2.23)$$

It is perhaps more instructive to derive equation (2.23) directly. We consider a thick slab between levels  $z_0$  and  $z_1$  with radiant intensity  $I_0$  incident vertically upwards at the base (due to the lower boundary at temperature  $T_0$ ) as shown in Figure 2.5. To calculate  $I_1$  (the intensity leaving the top of the slab), we consider the elemental layer of thickness  $dz$  at level  $z$  ( $z_0 < z < z_1$ ) and temperature  $T(z)$  emitting radiation in the upward direction  $k\rho dzB(T(z))$  as before. This emitted radiation will have to traverse a thickness  $z_1 - z$  before emerging from the thick slab at level  $z_1$ , the transmission being :

$$\tau(z, z_1) = \exp\left(-\int_z^{z_1} k\rho dz'\right) \quad (2.24)$$

Thus the contribution to  $I_1$  from elemental layer of thickness  $dz$  will be:



**Figure 2.5** Schematic illustrating contribution of elemental slab at height  $z$ , thickness  $dz$ , to the upwelling radiance observed at  $z = z_1$ .

$$\begin{aligned}
dI_1 &= k\rho dz B(T(z)) \exp\left(-\int_z^{z_1} k\rho dz'\right) \\
&= B(T(z)) d\tau(z, z_1)
\end{aligned} \tag{2.25}$$

ie.  $d\tau(z, z_1)$  is the fractional change in atmospheric transmittance between levels ( $z$  &  $z_1$ ) and ( $z + dz$  &  $z_1$ ). Thus, combining the absorption term from equation (2.20) with the emission term (2.25) integrated over the range of transmission :

$$I_1 = I_0 \tau(z_0, z_1) + \int_{\tau(z_0, z_1)}^1 B(T(z)) d\tau(z, z_1) \tag{2.26}$$

It is sometimes useful to express (2.26) in terms of the variable pressure, since measurement of atmospheric structure is often performed by balloon ascents which record the water vapour content, temperature etc. at various pressure levels. If we consider the limits to be  $p_0$ , the atmospheric pressure at ground level, and 0 (ie. the pressure at the top of the atmosphere):

$$I_\lambda = I_{0\lambda} t_\lambda(0, p_0) - \int_0^{p_0} B_\lambda(T(p)) dt_\lambda(0, p) \tag{2.27}$$

In this expression, the limits have been reversed, resulting in a minus sign before the integral. Also, note that the radiance at the top of the atmosphere is now considered to be wavelength dependent; previously this fact had only been implied in the definition of  $\tau$  (now called  $t_\lambda$ ).

Another consideration which has importance for the global climate system and also some significance for the calculated flux reaching space is the contribution of the downwelling radiance from the atmosphere. It is this part that is primarily responsible for the "greenhouse effect" warming of the planet's surface and lower layers of the atmosphere. In addition, if we consider the lower boundary to be an imperfect radiator, that is, with emissivity less than one, there will be some radiation reflected from the surface that will contribute to the signal observed from space. If we consider the elemental layer at pressure level  $p$  and temperature  $T(p)$  (Figure



2.6), the contribution to the downwelling atmospheric radiance at  $p = p_0$  is:

$$dI_{0\lambda}' = B_{\lambda}(T(p))dt_{\lambda}(p_0, p) \quad (2.28)$$

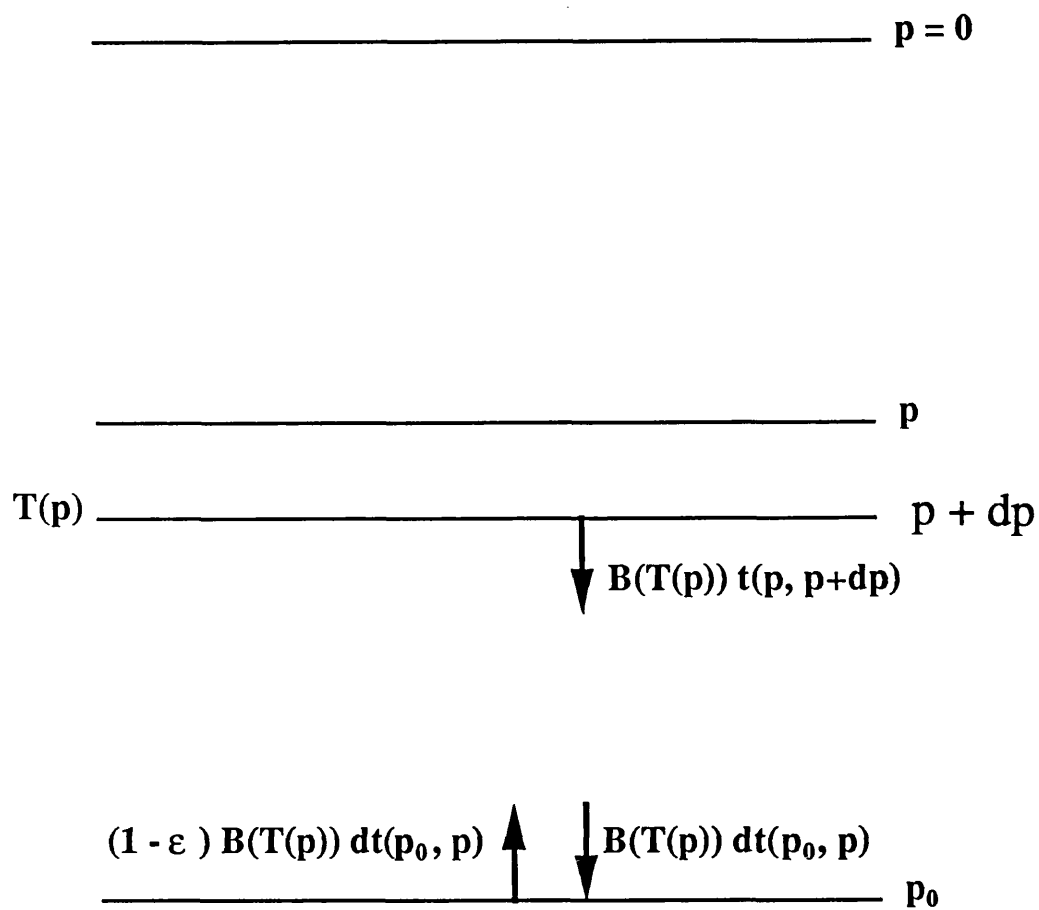
where  $I_{0\lambda}'$  is the downwelling atmospheric radiance at the surface and  $t_{\lambda}(p_0, p)$  is the transmission between pressure level  $p$  and the ground. Integrating from top to bottom of the atmosphere and including the effect of emissivity and ground temperature  $T_0$ , we can calculate the upwelling radiance at the ground,  $I_{0\lambda}$  which is used in equation (2.27):

$$I_{0\lambda} = B_{\lambda}(T_0)\epsilon_{\lambda} + (1 - \epsilon_{\lambda}) \int_0^{p_0} B_{\lambda}(T(p))dt_{\lambda}(p_0, p) \quad (2.29)$$

## 2.5 Atmospheric correction

Surface temperature measurements by infrared radiometry using meteorological satellites (eg. the AVHRR instrument on NOAA polar orbiting satellites) requires some form of correction to the brightness temperatures perceived by the satellite sensor in order to obtain the related ground temperature. The measurement is usually made in an atmospheric 'window' (see section 2.3) where the molecular absorption is at a minimum, generally between 10 and 12 $\mu$ m. Nevertheless, the influence of the atmosphere on the signal is not negligible in these regions. For example, Barton (1983) reports deficits of up to ~10K in the tropics for a vertical path through the atmosphere, and the effect is increased if the surface temperature is sensed at an angle through the atmosphere (Llewellyn-Jones *et al*, 1984). The primary absorber in this window is atmospheric water vapour (section 2.3) which is highly variable in content. If the water vapour content as a function of pressure is known, together with the temperature profile (eg. as measured by radiosonde balloon ascent) then the effect of the atmosphere can be calculated by performing numerical integration of the radiative transfer equation described in section 2.4. However, the existence of such data coincident with a satellite overpass is the exception rather than the rule, and the usefulness of the derived correction diminishes as distance and time increase from the site and occasion of the profile measurement.

A correction method based on two or more radiometric measurements of surface temperatures is clearly a more attractive proposition. Anding and Kauth (1970)



**Figure 2.6** The contribution of elemental slab at pressure  $p$  to the upwelling radiance at  $p = p_0$  via its reflected downwelling emission.

proposed the use of measurements at different wavelengths, and McMillin (1975) also proposed the use of measurements at different angles. This latter option will be considered in conjunction with the former in Section 2.6. The first validation of the multi-wavelength technique was performed in 1973 when Prabhakara *et al* applied it to Nimbus 3 and 4 IRIS (Infrared Interferometer Spectrometer) data. Since then, the technique has increasingly been used for the retrieval of surface temperatures, mainly of the sea (eg. McClain *et al*, 1981, 1989, Barton, 1983). The principle of multispectral correction by differential absorption can be derived from consideration of the radiative transfer equation (2.27) which can also be written in terms of the radiance deficit,  $\Delta I_\lambda$ , resulting from absorption and subsequent re-emission by the atmosphere (this version after Deschamps and Phulpin, 1980):

$$\begin{aligned}\Delta I_\lambda &= B_\lambda(T_0) - I_\lambda \\ &= - \int_0^{p_0} [B_\lambda(T_0) - B_\lambda(T(p))] dt_\lambda(0,p)\end{aligned}\quad (2.30)$$

We can also express this in terms of the temperature deficit,  $\Delta T_\lambda$ , between the radiometric temperature,  $T_\lambda$ , measured by the satellite at wavelength  $\lambda$ , and the surface temperature,  $T_0$ :

$$\Delta T_\lambda = T_0 - T_\lambda \quad (2.31)$$

Where  $T_\lambda$  is defined by the expression :

$$B_\lambda(T_\lambda) = I_\lambda \quad (2.32)$$

Figure 2.7 shows the near-linear behaviour of the Planck function between 273K and 310K for wavelengths at 10.8 $\mu$ m and 12.0 $\mu$ m (AVHRR channels 4 and 5 respectively). We therefore assume that, for small values,  $\Delta T_\lambda$  can be related to  $\Delta I_\lambda$  by :

$$\Delta T_\lambda = \frac{\Delta I_\lambda}{\left(\frac{dB_\lambda}{dT}\right)_{T_0}} \quad (2.33)$$

### Black Body radiance vs temp for 10.8 $\mu\text{m}$ and 12.0 $\mu\text{m}$

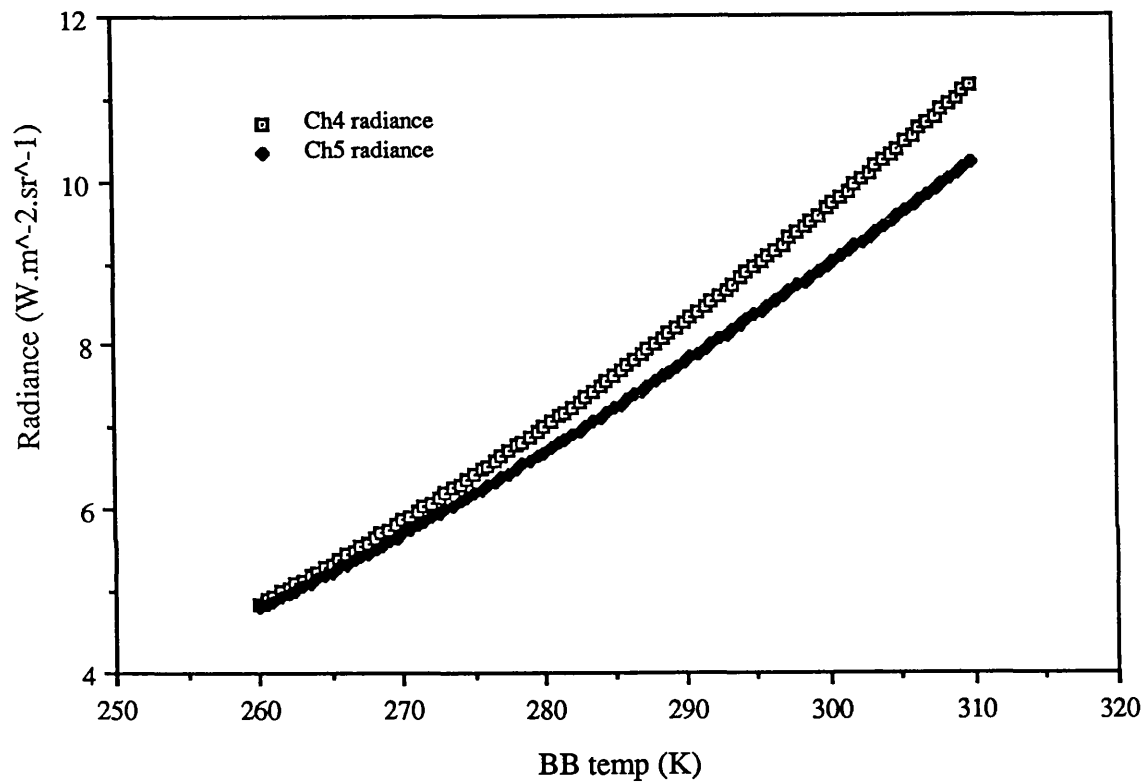


Figure 2.7 The plot of radiance versus temperature for the Planck function at the wavelengths of AVHRR channels 4 & 5, with temperatures ranging between 260 and 310 K. The main point is to illustrate the validity of the use of the first order Taylor expansion for a small range of values about any likely target temperature.

thus assuming that  $(\frac{dB_\lambda}{dT})_{T_0}$  is approximately constant. Indeed, for a temperature of 285K and  $\Delta T_\lambda$  of 1 Kelvin, this assumption leads to an error of 0.5% in the radiance estimate and hence a negligible error in  $\Delta T_\lambda$  since  $B_\lambda(T) \propto T^5$  at 11 $\mu$ m.

In order to separate the wavelength dependency from the integral in equation (2.30) we make the following approximations :

- (a) Absorption by atmospheric gases is low enough to allow the relationship between differential change in transmittance and absorbing gas amount to be:

$$dt_\lambda(0,p) = -k_\lambda dU(p) \quad (2.34)$$

where  $k_\lambda$  is the absorption coefficient and  $U(p)$  is the content of the absorbing gas integrated between pressure levels 0 and  $p$ . Thus we are considering the absorption to be proportional to the total content of the gas in the column, since the optical thickness is much less than one and each elemental layer in the column absorbs little of the total signal, consequently successive layers are effectively attenuating the same signal.

- (b)  $B_\lambda(T(p))$  can be expanded about the surface temperature  $T_0$  by using the first-order Taylor approximation :

$$B_\lambda(T(p)) = B_\lambda(T_0) + \left(\frac{dB_\lambda}{dT}\right)_{T_0}(T(p) - T_0) \quad (2.35)$$

From equations (2.30) and 2.33) we have:

$$\Delta T_\lambda \left(\frac{dB_\lambda}{dT}\right)_{T_0} = - \int_0^{P_0} [B_\lambda(T_0) - B_\lambda(T(p))] dt_\lambda(0,p)$$

and substituting for  $B_\lambda(T(p))$  gives:

$$\Delta T_{\lambda} = - \int_0^{p_0} (T_0 - T(p)) dt_{\lambda}(0, p)$$

Finally, substituting for  $dt_{\lambda}(0, p)$  (approximation (a)) gives the following simple expression for  $\Delta T_{\lambda}$ :

$$\Delta T_{\lambda} = -k_{\lambda} \int_0^{p_0} (T_0 - T(p)) dU(p) \quad (2.36)$$

Note that the integral depends only on the vertical structure of temperature and absorber content in the atmosphere, conditions which are the same for all wavelengths. Thus the wavelength dependency is expressed solely in terms of the absorption coefficient and we can write:

$$\Delta T_{\lambda} = k_{\lambda} f(T(p)U(p)) \quad (2.37)$$

For the temperature differences observed at two different wavelengths, we have :

$$\Delta T_{\lambda_1} k_{\lambda_2} = k_{\lambda_1} \Delta T_{\lambda_2}$$

and since  $\Delta T_{\lambda} = T_0 - T_{\lambda}$  :

$$(T_0 - T_{\lambda_1}) k_{\lambda_2} = (T_0 - T_{\lambda_2}) k_{\lambda_1}$$

$$\therefore T_0 = T_{\lambda_1} \frac{k_{\lambda_2}}{(k_{\lambda_2} - k_{\lambda_1})} + T_{\lambda_2} \frac{k_{\lambda_1}}{(k_{\lambda_2} - k_{\lambda_1})} \quad (2.38)$$

Thus it is possible to obtain the surface temperature,  $T_0$ , from a linear combination of radiometric temperatures,  $T_i$ , at wavelengths  $\lambda_i$  :

$$T_0 = a_0 + \sum_{i=1}^n a_i T_i \quad (2.39)$$

where coefficients  $a_i$  depend upon the absorption coefficients  $k_i$  at wavelengths  $\lambda_i$ . The constant term  $a_0$  is added to take account of the nearly constant contributions of terms such as CO<sub>2</sub> emission and surface reflection. This linear relationship theoretically provides an exact solution for sea-surface temperature as long as the approximations made remain valid. There are several effects which introduce non-linearity into the equation :

- Equation (2.34) should be written:

$$dt_\lambda(0,p) = -k_\lambda t_\lambda(0,p)dU(p) \quad (2.40)$$

so that the influence of the lowest layers of the troposphere are appropriately weighted by the wavelength-dependent factor  $t_\lambda(0,p)$ , ie. there is some attenuation of the signal before it reaches the upper layers.

- $k_\lambda(p,T)$  is dependent on temperature and pressure and is variable with wavelength. This is particularly important in the 10-12  $\mu\text{m}$  region where the continuum absorption of water vapour (the dominant attenuating process in this window) is dependent on the vapour partial pressure.
- The first-order approximation of the Planck function is not valid at 3.7  $\mu\text{m}$  since it is at the toe of the distribution and the dependence of emitted energy on temperature is decidedly non-linear in this region; the dependence of emitted energy is  $\sim T^{13}$ .

In practice it has been found that these non-linearities have been ignored to date (eg. McClain, 1989), thus we can consider that the temperature deficits in two channels are related in a linear way, such that the difference between the radiant temperatures is proportional to the total temperature deficit for either channel. It is therefore sometimes more helpful to consider the equation for retrieval of surface temperature as:

$$T_0 = a_0 + T_1 + a_2(T_1 - T_2) \quad (2.41)$$

since, from consideration of the values of the constants in equation 2.38, we can see that the constant multiplying  $T_1$  is positive and one larger in magnitude than the negative one which multiplies  $T_2$ , which allows factorisation of the difference in temperatures.

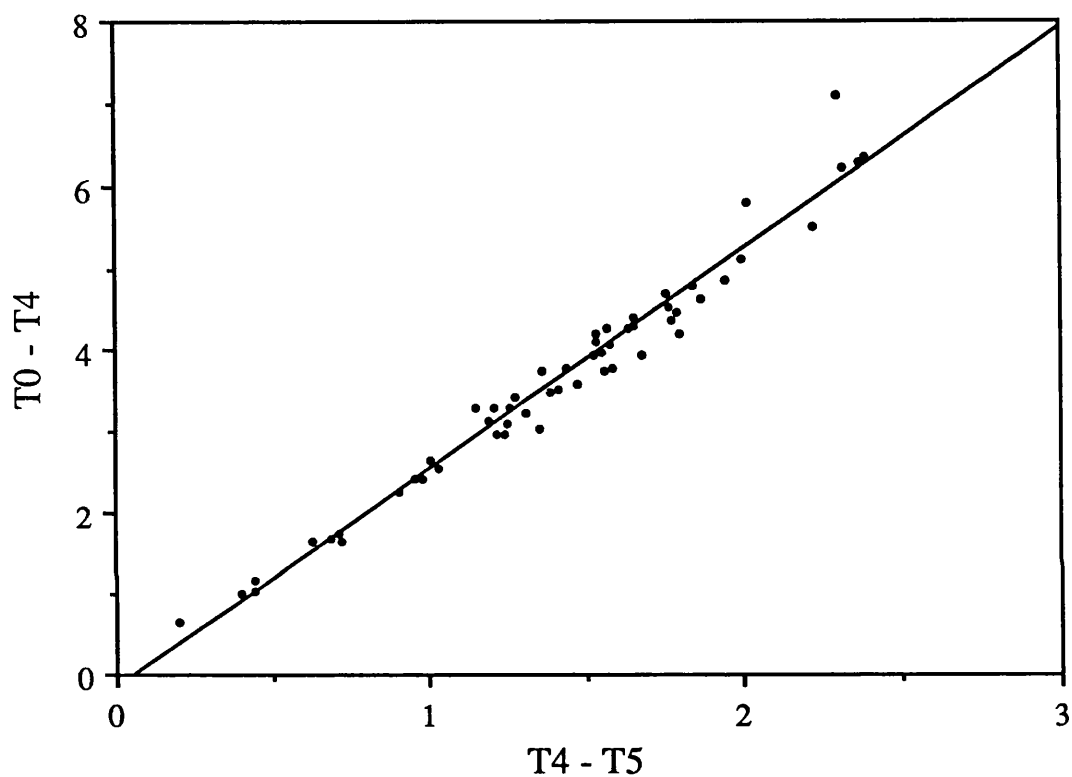
Thus the surface temperature is equal to the radiant temperature in one channel plus a correction, the correction being proportional to the difference in the observed radiant temperatures of the two channels. This relationship is shown in Figure 2.8, which shows the corrections required for given temperature differences as measured in AVHRR channels 4 and 5, which are centred at 10.8 and 12.0  $\mu\text{m}$  respectively.

The brightness temperatures were generated using the Atmospheric Transmission Model developed at RAL by A. Zavody. This program uses data from radiosonde ascents to calculate the line by line absorption and emission of the sample atmospheres. The output is the perceived radiance in channels 4 and 5 for a given ground temperature and intervening atmosphere, with the radiosonde data providing the concentrations of various trace gases plus water vapour and the temperature for pressure intervals in the atmosphere. The relationship between the absorption profiles in the two channels is demonstrated for 3 different atmospheres in Figure 2.9. It is also possible to calculate the atmospheric effect for various slant paths through the atmosphere in order to obtain a correction for the increased atmospheric path as the satellite views off-nadir. Such scan angle dependent coefficients were demonstrated by Llewellyn-Jones *et al* (1984).

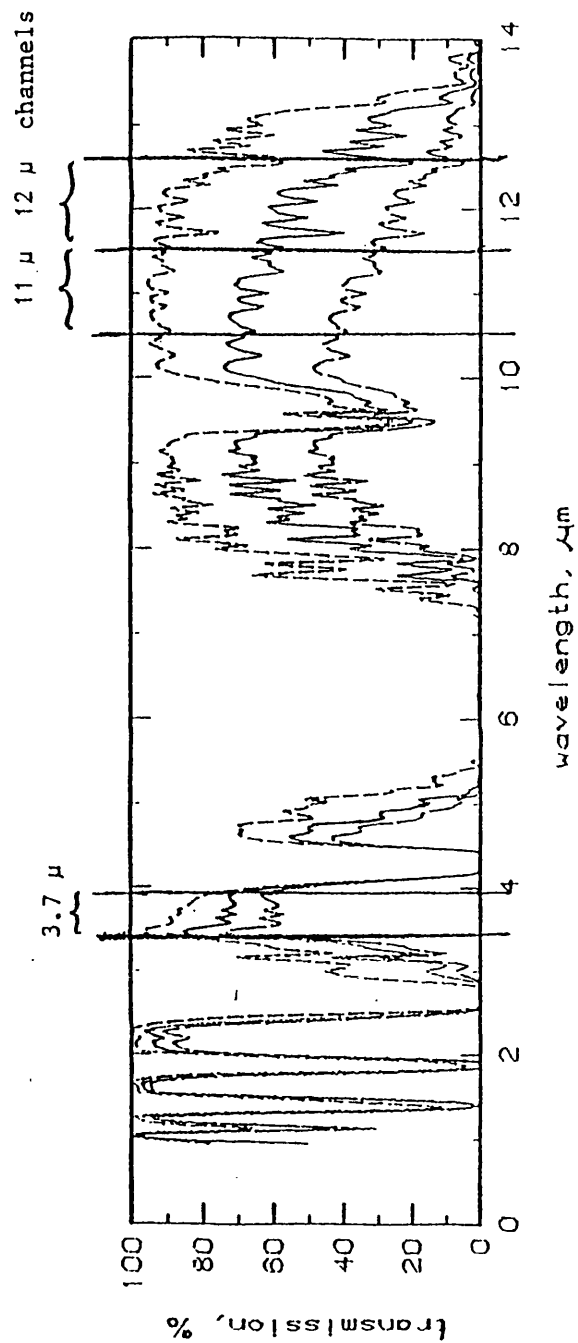
In practice the coefficients in the equation for sea-surface temperature from AVHRR are often empirically derived from correlations with ground truth data provided by buoys and ships (McClain *et al*, 1983), and the resultant equation is described as "tuned". This method permits the elimination of instrumental systematic errors and other factors such as the "skin effect" which is the strong thermal gradient which occurs in the last few millimetres of the sea-surface due to the evaporation of water from the surface layer and other heat exchange processes between the water and atmosphere. Since the skin depth of the IR radiation is much smaller than this depth, the radiation from the sea-surface is not emitted from a layer at the bulk temperature, which is measured by buoys and required by climate modelers, but a layer at some lower temperature. Previous research shows that the magnitude of this effect ranges typically from  $\sim 0$  degrees to  $\sim 1$  degree K over the ocean (Robinson *et al*, 1984, Hepplewhite, 1989, Schluessel *et al* 1990).

In addition, coefficients generated from atmospheric transmission programs are often tuned to a particular geographical area (eg. Llewellyn-Jones *et al*, 1984) by selecting a series of radiosondes over the region of interest. This is of particular interest where





**Figure 2.8** The relationship between channel 4 atmospheric correction and the difference between brightness temperatures in channels 4 and 5 is shown for a wide range of different atmospheres. The brightness temperatures have been calculated using a line-by-line transmission model and 56 radiosonde profiles, and the straight line represents a conventional 'split-window' algorithm.



**Figure 2.9** This figure shows three calculated atmospheric transmission spectra between 1 and 14  $\mu\text{m}$  wavelength for 7 mm, 29 mm and 54 mm of precipitable water. These were generated by a line-by-line transmission model developed by A. Zavody at RAL, and show increased absorption for the wetter atmospheres. The approximate positions of the 3 AVHRR thermal channels are also shown.

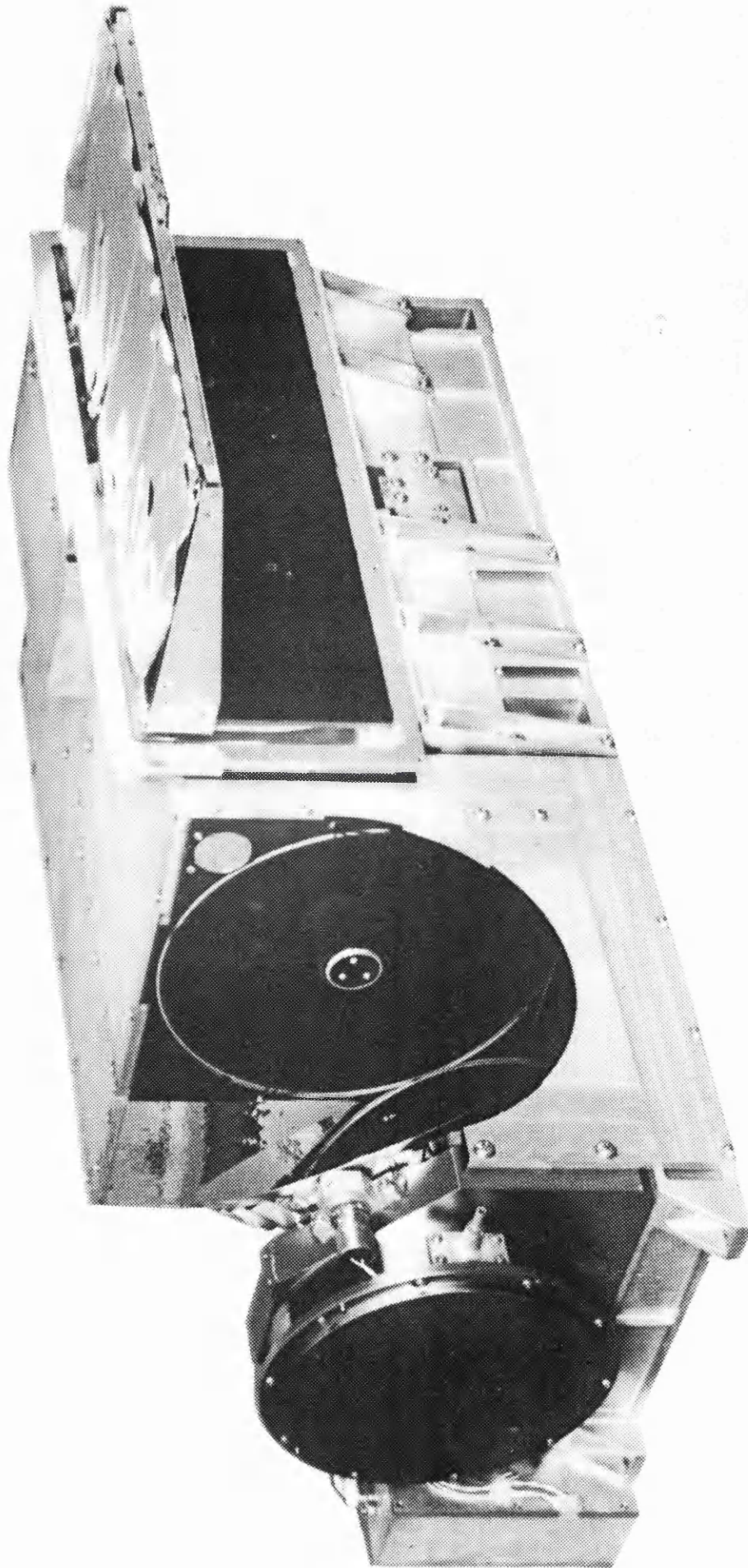
the atmospheric transmission is fairly high, since the inclusion of low transmittance tropical atmospheres in the analysis causes a marked change in the coefficients produced.

## **2.6 Radiometer Design with AVHRR as an example**

We now turn our attention to the mechanism by which the radiance emerging from the top of the atmosphere is detected and measured by the satellite. The AVHRR instrument, whose data is used in the majority of work for the rest of this thesis, is shown in Figure 2.10. The basic method of radiance collection is by a telescope optical system that focuses the radiant energy onto a detector. A scene is imaged by scanning the field of view of the optical system to produce a large area of coverage. Although both electronic and mechanical systems are used for remote sensing instruments, thermal infrared instruments usually employ the latter. In mechanical (also known as 'whiskbroom') scanning, the field of view is changed by angular rotation of the sensor system (eg. Meteosat) or angular movement of a scan mirror, usually in the cross-track direction (eg. AVHRR). In the case of AVHRR and other Low Earth Orbit satellites, along-track motion of the satellite provides coverage in the other dimension. The principal advantage of the whiskbroom type of sensor is that only a few (one per spectral channel in the case of AVHRR) detectors are needed and calibration is therefore relatively uncomplicated. The main disadvantage is that the sensor only captures a tiny fraction of the radiation directed towards it since the scan mirror/optical system is only aligned with any particular ground element for  $\sim 25 \mu\text{sec}$  in the case of AVHRR. The result of this is that the detector must be well cooled in order to limit the noise on the measurement of the signal.

### **2.6.1 Telescope and receiver systems**

Many telescopes for space sensors make use of mirrors to form the primary image. There are particular advantages for multispectral systems since mirrors are free from chromatic aberrations. A spherical mirror, whilst being easy to manufacture, has the disadvantage of spherical aberration, where the focal length of circular elements decreases from centre to edge of the mirror. The primary mirror is usually a parabolic one, which has the property of producing a perfect focus for rays which are incident parallel to the axis of the mirror. The disadvantage of the parabolic mirror is coma, such that rays incident at an angle to the axis result in distorted images. Fortunately, the field of view (determined by field stops) of whiskbroom scanning radiometers is



**Figure 2.10** Features of the AVHRR instrument can be seen, including the passive radiant cooler and its sun shield. The diagonal scan mirror with its circular motor housing can be seen to the left of the main aperture

small (eg. 1.4 milliradians for AVHRR) and all the radiant energy is focussed onto the detector array.

The primary mirror concentrates the incoming rays into a cone, but before the radiation is focussed to a point it is intercepted by a convex hyperbolic (in the case of the Cassegrain telescope design, eg. AVHRR) mirror which lengthens the cone and reflects it back through a small hole in the centre of the primary mirror. The main advantage of these compound reflector telescope designs such as the Cassegrain is that they offer long focal length in a short physical length, thus reducing the size and weight of supporting structures which is important for space-based applications. The resolution of a mirror of diameter  $D$  is given by the Rayleigh criterion, thus the minimum angular separation that can be resolved in the case of AVHRR at  $12\mu\text{m}$  and aperture 20cm is:

$$\alpha = 1.22\lambda/D$$

$$\approx 0.07 \text{ milliradians}$$

although this resolution is lost since the sensor integrates the signal over the entire field of view.

Mirrors are usually made of low expansion materials such as Zerodur and ULE titanium silicate which are chosen for their thermal properties and high stiffness-to-density ratios; the latter property helping to reduce the weight of the optical system. The reflecting surfaces of such figured mirrors are coated with evaporation-deposited metallic films; aluminium being particularly useful since its reflectance is above 90% from  $0.2\mu\text{m}$  to  $20\mu\text{m}$ , although gold is employed for the ATSR primary mirror. The metal film is usually coated with silicon monoxide or magnesium fluoride to protect it from the space environment.

Since the positional relationship between the various elements of the optical system must be maintained, the structure of the instrument must have a high degree of rigidity, good resistance to thermal deformation and be light in weight. The structural materials most commonly in use are aluminium (eg. AVHRR) and graphite/epoxy (eg. ATSR) with beryllium in occasional use in components requiring high specific stiffness. Although graphite/epoxy has better stiffness-to-density than aluminium, it is particularly prone to absorption and subsequent outgassing of water vapour (Chen,

1985), resulting in the potential for sensor contamination. Metal matrix composites are under consideration for future instruments.

Once the radiation has been focussed onto the detector assembly, beam-splitters divide up the radiant energy and pass it through filters in front of each detector.

### 2.6.2 Detector types

For ultraviolet and visible radiation, a silicon photodiode will generate a current proportional to the flux incident on the active surface area. This photocurrent produces a voltage across the load resistor which is the signal that is amplified and recorded. For the measurement of low level *dc* radiation signals, the unbiased photovoltaic type has distinct advantages since it has no dark current. The biased photoconductive type has a much faster response time (nanoseconds instead of microseconds) and is therefore advantageously employed in the measurement of rapidly varying *ac* fluxes. Such silicon-based semiconductor devices are employed as the detectors for AVHRR channels 1 and 2 (0.63 $\mu\text{m}$  and 0.86 $\mu\text{m}$ ) for example.

Terrestrial infrared radiation can be measured by measuring the photocurrents liberated inside materials such as HgCdTe in a similar manner to the silicon photodiodes described above. HgCdTe detectors are employed for AVHRR channels 4 and 5 and channels 3 and 4 of the forthcoming ATSR (all at 10.8 $\mu\text{m}$  and 12.0 $\mu\text{m}$ ). In order to operate at these wavelengths it is necessary to cool the detectors to reduce the thermally liberated electron currents in the lattice.

### 2.6.3 Detector cooling systems

There are two types of cooling systems employed for most infrared detectors. The first is the passive radiant cooler which can cool systems to cryogenic temperatures by radiating thermal energy to the low temperature space environment. The second type are the closed cycle mechanical refrigerator systems which provide cooling at lower temperatures by rejecting heat at higher temperatures via a heat engine cycle.

Passive radiant coolers have been the most widely used to date (eg. for AVHRR); the AVHRR coolers maintain the detector assembly at a temperature of  $\sim 105\text{K}$  (ITT, 1989). The main advantages are:

- Reliability and simplicity
- No moving parts
- No power consumption

and the principle disadvantage is that they require shielding from sources of radiation such as the Sun (eg. see Figure 2.10). Active Stirling Cycle coolers are now being used (eg. for ATSR) as early problems such as the lifetime of dry-lubricated bearings and seal wear have been overcome. New designs have magnetic bearings and clearance seals, linear drive motors and employ dynamic balancing. Their chief advantage is that they are capable of cooling to lower temperatures than passive coolers, resulting in, for example, a temperature of 80K for the ATSR detector assembly. The noise reduction of a factor 4 for ATSR (which has a smaller primary mirror) over the AVHRR detectors, is a combination of this improved cooling and increased dwell-time per pixel.

## **2.7 Calibration of the radiometer signal**

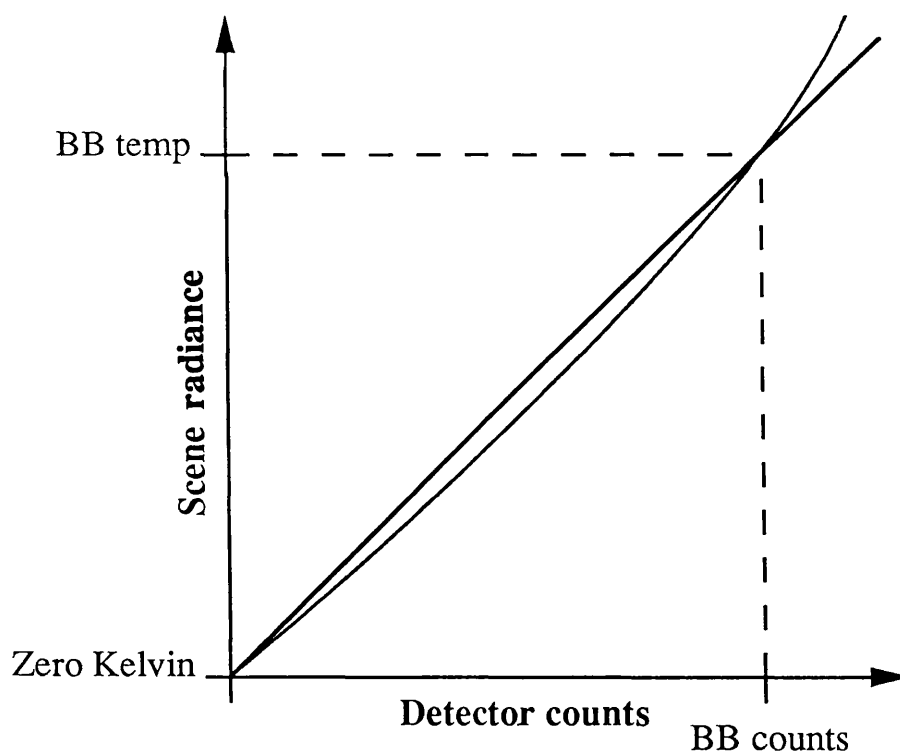
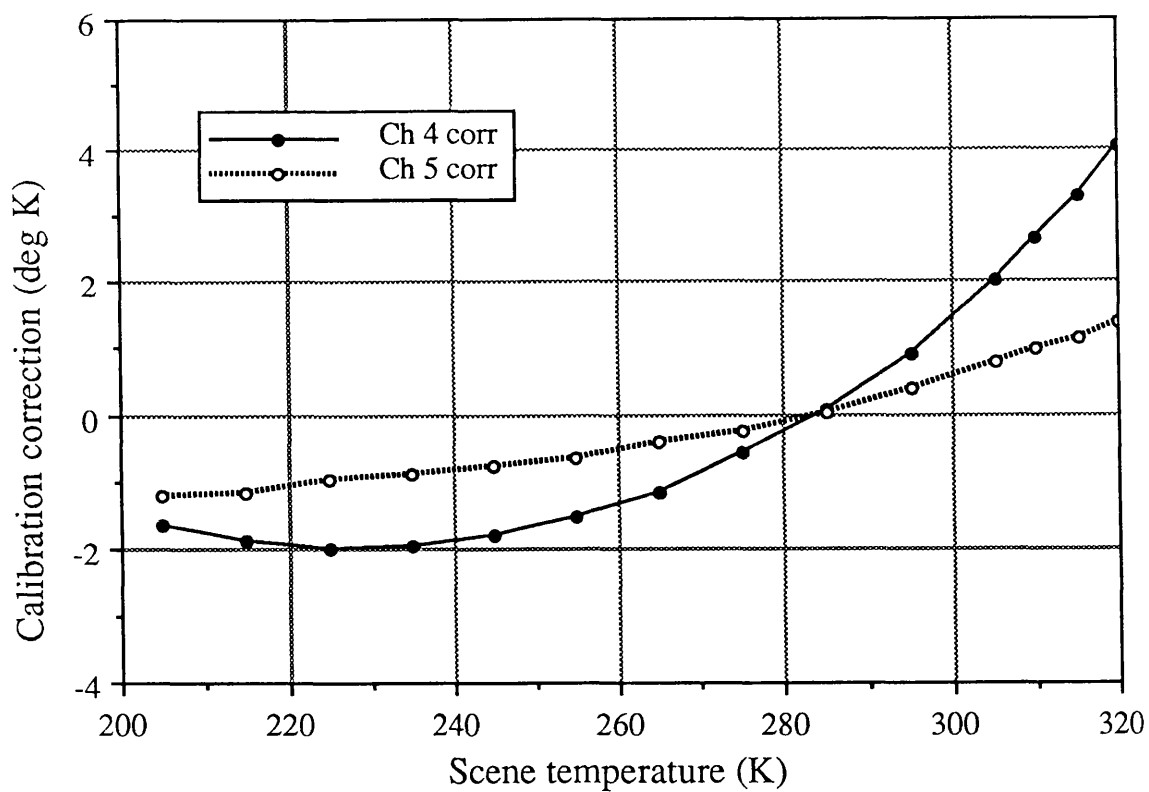
### **2.7.1 Need for on-board calibration of the system**

The detector/optical system must be calibrated on board the satellite. The reasons for this on-board calibration are : -

- The zero offset and gain of the detectors will drift with time (for example due to  $1/f$  noise, to which HgCdTe detectors are prone (Chen,1985)).
- A decrease in the effectiveness of the cooler system will result in increased noise and may also introduce a bias into the output of the detectors.
- There may be a degradation of the optical surfaces, resulting in less radiant flux reaching the detectors for a given brightness temperature.
- The instrument will be heating up and cooling down during each orbit, resulting in a variation in the background thermal emission from the body of the instrument which can be seen by the detectors due to aberrations of the optics.

### **2.7.2 Calibration schemes for thermal channels**

The AVHRR instrument employs a two point calibration scheme to correct for the effects mentioned in the previous section. The two sources that are nominated to have known radiance are an on-board calibration target and the view to deep space, at temperatures of ~285K and effectively absolute zero respectively. One disadvantage with this calibration scheme is the non-linear response of the detector with respect to radiance, and hence brightness temperature, as illustrated in Figure 2.11. This results



**Figure 2.11** Top shows non-linear correction values for AVHRR channels 4 and 5 that must be applied to brightness temperatures calculated from the simple two-point calibration scheme. The lower sketch illustrates the origin of the shape of the non-linear correction curves. Note that the straight line and curve agree at zero kelvin and the black body temperature (typically ~285 K) but not for other temperatures.



in the radiance calculated from the simple relationship  $R = I \times G + C$  ( $R$  = radiance,  $I$  = detector counts,  $G$  = gain of detector calculated from the two calibration points and  $C$  is the calculated intercept) being incorrect for most incoming radiances. In practice, a non-linear correction is applied to obtain the true radiance. The value of the non-linear correction is obtained from the Non-Linear Correction (NLC) table compiled by Lauritson before the launch of the instrument. Since no values for the correction have been derived since launch, the current derived radiances are probably in error. It is the large difference between the radiances for space and the internal calibration target that makes the values for non-linear correction in the table so critical. It is clear that the interpolation between these two radiances is taking place over a large point separation. A further problem with AVHRR is the uncertainty in the radiance from the calibration target, which is known to suffer from variable thermal gradients of up to  $\sim 1^\circ\text{C}$ .

The in-flight calibration scheme of ATSR makes use of two simulated black body radiation sources<sup>1</sup> at different known temperatures. The two temperatures are chosen to be roughly at each end of the range of brightness temperatures expected to be observed from the sea surface ( $\sim 0$  to  $30^\circ\text{C}$ ). This two-point calibration scheme is an improvement over the current AVHRR calibration which, as mentioned above, makes use of the radiance of space together with a single on-board calibration target.

For ATSR, the interpolation is being carried out over a range of  $\sim 30\text{K}$  instead of  $\sim 285\text{K}$ , so the non-linear correction values are fixed much more precisely. In addition, the black body sources are designed to be much more accurate.

### ***Accuracy requirement for the ATSR Black Body sources***

The performance requirement is that any error in retrieved sea surface temperature due to the inaccuracy of the Black Body sources should be less than  $0.1\text{K}$ . This requirement stems from the overall system requirement of  $0.35\text{K}$  for the retrieval of Sea Surface Temperatures. This  $0.1\text{K}$  accuracy requirement translates to certain requirements in three areas (Mason *et al*, 1990).

- 1) A high and constant emissivity for the Black Bodies. By considering the effect within the SST retrieval algorithms the aim has been established for an emissivity

<sup>1</sup> From here on, when describing the ATSR calibration sources, the phrase 'Black Body' is used for convenience instead of 'simulated black body radiation source'.

of  $> 0.998$  in all channels.

- 2) Good thermal uniformity within the Black Body cavity. The design aim is for any gradients on the viewed surface to be  $< 30$  mK.
- 3) Accurate and sufficient temperature monitoring of the Black Bodies throughout the mission life. The aim is for absolute temperature monitoring to within 30 mK.

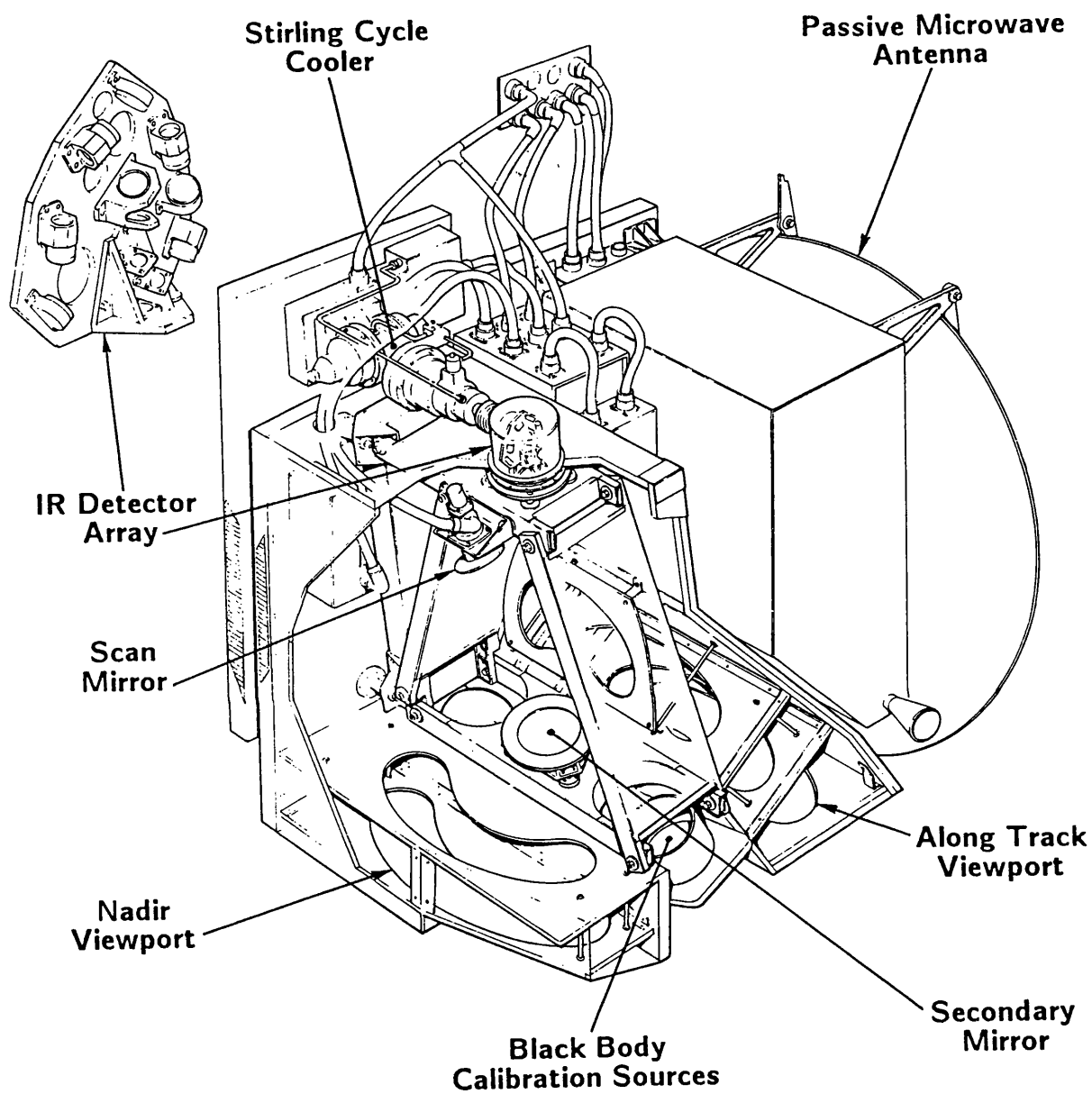
The thermal uniformity and temperature monitoring and calibration tests had already been carried out. The emissivity design had been performed partly by calculation and partly by measurement on a separate sample surface, thus it was still necessary to measure the emissivity of the black bodies by experiment in order to complete their calibration.

## **2.8 The Operation of the Along-Track Scanning Radiometer (ATSR)**

The techniques described in the rest of this thesis, although having been tested using data from the current series of AVHRR instruments, have been developed with ATSR in mind. It is therefore of interest to consider the main features of this instrument, primarily to demonstrate the improvements it offers over AVHRR, some of which have already been mentioned.

The ATSR is an instrument designed to measure thermal infrared radiation from space with the primary purpose of retrieving SST to unprecedented accuracy for climate research compared with retrievals from current generation weather satellites such as AVHRR and is currently flying on the recently-launched ERS-1 satellite. There are some novel features of this instrument compared with the NOAA AVHRR, the most important of which are those of along-track scanning and internal calibration, the latter having been covered in the previous section. These features are designed to improve the correction for atmospheric attenuation together with an improved absolute accuracy of the measured radiances respectively.

We now consider the instrument as shown schematically in Figure 2.12. The along-track scan direction provides a different path length through the atmosphere, thus providing, for each wavelength channel, two radiances for the same region of sea. The radiance obtained from the sea surface for the along-track direction will coincide with that at nadir some two minutes later when the satellite passes over the same area of sea. It is unlikely that the atmosphere above the sea will have changed significantly in this short period of time, so the radiance for the along-track and nadir



**Figure 2.12** Cut-away diagram of the ATSR instrument showing the main advanced features of the instrument. The microwave sounder antenna is shown folded in its pre-launch position. (Diagram courtesy of RAL).

directions can be considered to have been attenuated by the same atmosphere, giving a two-fold increase in the number of parameters for the atmospheric correction algorithm used to obtain Sea Surface Temperature (SST). This is expected to improve the atmospheric attenuation correction from  $\sim 0.5\text{K}$  (currently achieved by AVHRR) to  $\sim 0.3\text{K}$  (Zavody, 1982). The three thermal IR channels used are the same as those for AVHRR, centred at 3.7, 11 and 12  $\mu\text{m}$  in the atmospheric transmission 'windows' (see Figure 2.9).

A near-infrared channel is also carried with the primary purpose of detecting cloud contaminated pixels over the ocean during the daytime which might affect the sea surface temperature retrieval and also provides good distinction between land and water. The channel is centred at 1.6  $\mu\text{m}$  and thus can also discriminate between clouds and ice or snow, which will be useful for determination of sea-ice boundaries and concentrations. Operation of this channel is shared with the one at 3.7  $\mu\text{m}$  for several reasons, the primary one being that there is a restriction on the amount of space available for ATSR data in the ERS-1 telemetry stream. This requirement led to the selection of 1.6  $\mu\text{m}$  data being transmitted during the daytime and 3.7  $\mu\text{m}$  data being transmitted at night since there is no detectable short wavelength reflected radiation to be received at this time and, conversely, the 3.7  $\mu\text{m}$  radiances are contaminated by reflected solar radiation during the day (eg. Llewellyn-Jones *et al*, 1984, and see Figure 1.8) and thus their inclusion can lead to errors in SST retrieval.

For the purposes of global monitoring, one important difference between the ATSR and AVHRR is that, for the former instrument, all data is recorded at maximum resolution of 1 km whereas the default for the latter is a resampled and sub-sampled format which has an effective resolution of  $\sim 4$  km at nadir and only short passes can be recorded per orbit at full resolution upon user request. Although the coverage of the ATSR instrument is significantly less (500 km swath as opposed to 3000km), it still provides repeat coverage at least every week and more frequently than this at high latitudes, except above  $\sim 84^\circ$  where no data is available due to the orbital inclination of the satellite.

## 2.9 Summary

We have discussed several topics that will be useful when considering the extraction of physical quantities from remotely-sensed data for use in climate studies. In particular, an understanding of the processes involved in radiative transfer through the atmosphere together with a review of techniques used to correct for atmospheric

effects on surface-emitted radiance will be invaluable when developing new techniques (see Chapter 5). In addition, knowledge of the design and operation of space-borne radiometers provides useful insight into the effects of instrumentation that will have to be taken into account. In this regard, calibration techniques required for the conversion of detector counts to brightness temperature together with an awareness of their possible accuracy limitations is required when using thermal infrared radiometer data to validate new atmospheric correction techniques (see Chapter 6). Additionally, knowledge of the imaging mechanism allows conversion of pixel measurements to physical distances (Chapters 3 and 4) and aids the geolocation procedures required for satellite - *in situ* comparisons (Chapter 6). Finally, an understanding of the relative merits of the AVHRR and ATSR instruments gives us confidence that techniques applied to AVHRR data will not only be suitable for use with that provided by ATSR but also that, particularly in the case of temperature measurements, the results will be improved with the latter instrument.

## Chapter 3

### Remote Sensing of Lake Areas:

### Technique Development, Validation and Comparison

#### 3.1 Review of Lake Area Remote Sensing

##### 3.1.1 The Importance of Lake Volume Monitoring for Climate Research

As discussed in Chapter 1, the precise mechanisms of climate have received much attention in recent years and many climate models have been developed. It is necessary to devise techniques which will be capable of testing these models by providing a monitor of current climate fluctuations for comparison with the predicted trends. The remote sensing of various phenomena that are influenced by climate allows us to detect the results of any variability and infer the nature of change that takes place. Examples of such proxy indicators include the annual cycle of growth and decay of sea ice (Laxon, 1990) and the volume change cycles of closed lakes.

The importance of lake volume monitoring for climate research is discussed by Mason *et al* (1985). It is well known that lake volumes fluctuate in response to changes in the evaporation and precipitation rates over their catchment basins. Measurements of lake volumes are not only important for economic purposes but also provide a climate record. In particular, closed lakes (those without surface outlet) are known to fluctuate considerably in response to variations in the climatic conditions (ie. changes in the precipitation and evaporation rates) in their catchment areas. Past lake volume fluctuations can be inferred from the morphology of the lake basin and from sediment cores. Paeleoclimate models (Street-Perrott and Harrison, 1985) based on reconstruction of past lake volume changes can then be tested against current lake volume fluctuations.

Currently, measurements of lake volume fluctuations have been made on relatively few lakes using ground based measurements of the levels and areas of the lakes. These methods rely on survey data to obtain the relation between lake level and lake area, which is then used in conjunction with one parameter (usually lake level) to obtain the lake volume change. It has been recently pointed out, however, (Mason *et al*, 1985) that remote sensing of levels and areas from satellite altimeters and imagers

offer an alternative method for measuring lake volume changes.

This alternative approach can supplement the current ground-based methods, and also be tested against them. In addition, however, remote sensing offers the potential to monitor large numbers of lake volumes on a global scale. Many of the most climatically sensitive lakes occur in semi-arid regions (see Figure 3.1) for which there is little or no *in situ* data. The other main advantage of remote sensing is that it offers frequent and regular sampling, and the data provided can be processed to ensure that it is accurate and self-consistent.

### ***Lake water balance***

Lakes which have no surface outlet or groundwater seepage are the most easily modelled. Mason *et al* (1985) describe the water balance of such a closed, sealed lake as follows :

$$dV/dt = R - A_L(E_L - P_L) \quad (3.1)$$

where  $V$  = lake volume,  $R$  = run-off rate of water from the catchment basin,  $E_L$  = evaporation rate over the lake per unit area,  $P_L$  = precipitation rate per unit area over the lake and  $A_L$  = the area of the lake. Since  $dV/dt = A_L dL/dt$ , the equation can be rewritten in terms of lake level,  $L$ ,

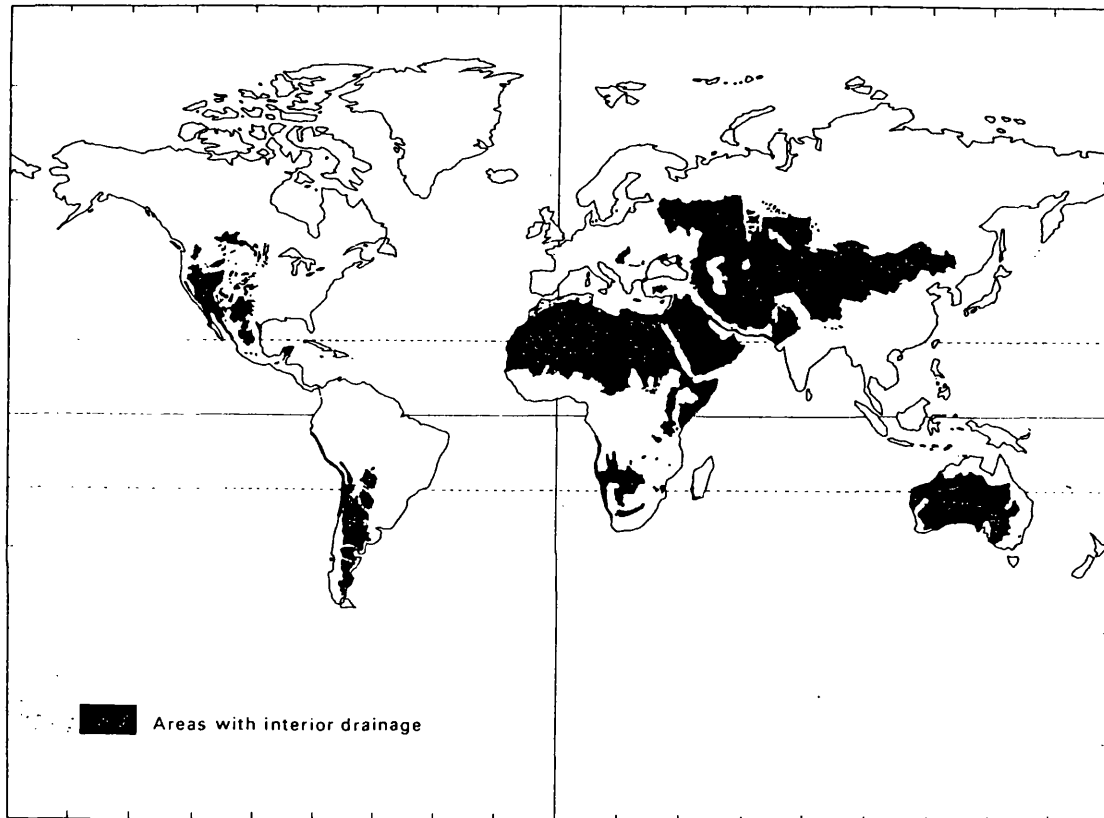
$$dL/dt = R/A_L - (E_L - P_L) \quad (3.2)$$

or, in terms of  $A_L$ ,

$$dA_L/dt = dA_L/dL [R/A_L - E_L - P_L] \quad (3.3)$$

Thus in general, to monitor a lake's response to changes in climate, ie. in precipitation or evaporation rates, it is necessary to monitor either  $L(t)$  or  $A(t)$  with an independent knowledge of  $A_L(L)$ . For many lakes ground survey data is not available, ie.  $A_L(L)$  is not known, but this problem can be solved by simultaneous remote sensing of both  $A_L(t)$  and  $L(t)$ .

The European Space Agency's satellite ERS-1 (European Remote Sensing satellite-1), which was recently launched, is well equipped to carry out these monitoring and measuring tasks. Both an imaging radiometer (the ATSR, see chapter 2) and a pulse-limited radar altimeter (RA) are carried on board the satellite. Using the data, it will



**Figure 3.1** The areas of the world where inland drainage can occur, ie. where closed lakes can exist.  
(From Street-Perrott and Harrison, 1985).



be possible to build up curves of  $A_L(L)$  for each lake. If the range of each curve is great enough it will be possible to calculate lake volume change using past and present satellite data from just one instrument (eg. AVHRR or the Seasat and Geosat RAs). Thus such remotely-sensed data should provide a valuable proxy indicator of climatic change from the late 1970's onwards. Rapley *et al*, 1987 estimated that, globally, there are ~1500 lakes of area  $> 100 \text{ km}^2$  of which ~200 are closed, and up to half of these should be able to be monitored regularly (at least once per month) by the radar altimeter on ERS-1. The figure for the number of closed lakes  $> 100 \text{ km}^2$  has been recently validated by work at MSSL (Mason, private communication, 1991).

An accuracy of a few percent in the measurement of lake area variations will allow changes in aridity as small as a few percent for lake areas of  $\sim 100 \text{ km}^2$  and over (Mason *et al*, 1985). An important question, then, is whether such high measurement accuracy can be achieved using a coarse resolution instrument such as AVHRR or ATSR. This chapter describes a study to investigate this question, developing and validating a simple technique (Harris and Mason, 1989) and comparing it with an alternative method.

### 3.1.2 Current algorithms for measuring lake areas

Since the advent of the Earth Resources Technology Satellite (ERTS), or Landsat, series, which carry the Multi Spectral Scanner (MSS) with a resolution of  $\sim 80$  metres, there has been little interest in using coarse resolution data (such as is provided by, for example, NOAA satellites) for measuring lake areas. Later versions of the Landsat satellite carry the Thematic Mapper (TM) instrument which has a resolution of  $\sim 30$  metres. For large lakes (above  $100 \text{ km}^2$ ) the resolution of these instruments is such that measurement of the areas to within a few percent should present little problem. However, such an instrument is impractical for global monitoring because coverage is too infrequent and data volumes would be too high, thus it is necessary to use instruments which have coarser spatial resolution. The main problem caused by coarse resolution is that of edge pixels (pixels which fall across the boundary between water and land and thus have an intensity between that of the land and the lake), which may account for a significant proportion of the area of the lake.

The problem of edge pixel measurement has been considered in Landsat images containing smaller bodies of water. Malila and Nalepka (1975) show that it is possible to detect up to 97% of the total area of water in an image, but their results consistently underestimate the area of water, thus indicating that a systematic error is

inherent in the method. Their method involves setting thresholds for the intensity levels of the edge pixels such that any pixel with an intensity that is further away from the mean water pixel intensity than the threshold is considered to contain no water. Pixels which are within the threshold are considered to contain a fraction of water, the fraction being the difference between the intensity of the pixel and the mean intensity of the pixels containing 100% water divided by the difference between the mean intensities of the 100% land pixels and 100% water pixels; such a method is best considered in conjunction with the idealised image histogram shown later in Figure 3.3. The threshold is required in order to avoid the inclusion of land pixels in the calculation of the lake area. The drawback with this method is that some pixels to the land side of the threshold set within the histogram still contain a small fraction of water but are not included in the analysis; this is probably the cause of the systematic bias in the results. It will be seen in the next section that thresholds for water and land should be set symmetrically with respect to the brightnesses of the water and land in order to be valid, and further that it is not actually necessary to employ the fractional brightness method described above.

Malila and Nalepka mention that there is a trade-off between the number of pixels containing some water that their technique fails to detect and the number of land pixels that are wrongly classified as containing some water. Such a technique applied to an area for which there was no *in situ* data would result in errors which would compensate for each other to a greater or lesser extent, depending on the parameters selected. Since the selection of the parameters is dependent on *a priori* knowledge of the composition of the image (eg. already knowing in advance which pixels contained some water), complete cancellation of the errors is unlikely. In any event this situation of self-cancelling errors must be regarded as unsatisfactory.

Barker (1975) also considered the problem of measuring lake areas with MSS data and found empirically that, for small lakes, the most successful method involved determining the intensity mid-way between the intensities of the water and land and counting all the pixels with an intensity closer to that of the water than the mid-point. To my knowledge, however, no studies of the validity of this method have been performed, and no work has been done with AVHRR. The work described in this chapter largely consists of a theoretical study and development of this simple technique, and an evaluation of the method.

After this study had been completed, however, another technique for retrieving sub-pixel areas was brought to my attention. Palmer and Forte (1990) describe a method

for image segmentation by assessing and plotting local isoluminance contours (LIC). The method is compared with the simple histogram technique later and so merits some description here. The method of segmentation has four basic stages (see Figure 3.2):

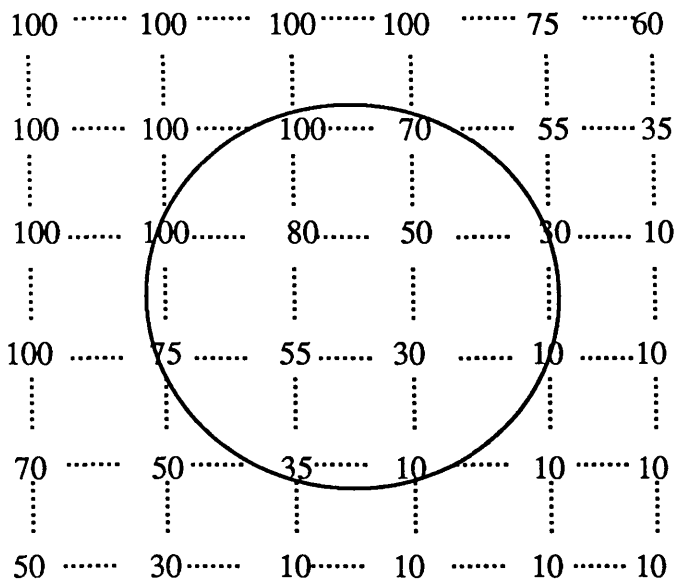
- Construction of a high resolution model of the intensity function  $I(x,y)$  to allow sub-pixel accuracy segmentation.
- Look in a circular neighbourhood of diameter  $4 \times 4$  pixels centred on an image grid point (ie. the point of intersection of  $2 \times 2$  pixels) to determine whether there is a grey level contrast greater than a predefined threshold between a bright and dark region (with values  $B$  and  $D$  respectively) within the neighbourhood.
- If such a contrast exists, check whether the isoluminance contour with value  $(B+D)/2$  passes through the central "unit cell" (ie. an area one pixel square in the centre of the neighbourhood) and if so plot it in the unit cell.
- If the value of the local isoluminance contour changes from one unit cell to another, join the end points of the contours in each unit cell to preserve connectivity.

In order to calculate the area of a closed outline such as would be represented by the coastline of a lake, the cross products of vectors joining successive end points of the contours to a chosen start point were summed and the result divided by two. The use of the cross product allows the calculation of area of any complex closed shape since the angle between the vectors was taken to be positive in one sense and negative in the other.

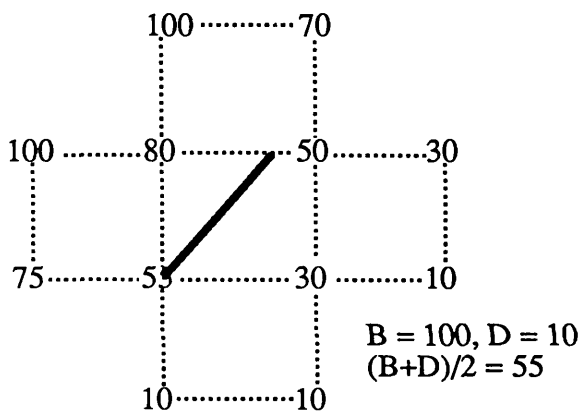
### **3.1.3 Lough Neagh as a case study for investigating and validating the methods**

The main objective is therefore to develop techniques for measuring lake areas with remote sensing instruments (AVHRR), and to assess their accuracy with a European case study. Any technique considered should be generalised so that it can be applied to lakes of various sizes down to  $\sim 100 \text{ km}^2$ , and also for lakes that occur in different locations, since the most climatically sensitive lakes occur in semi-arid regions.

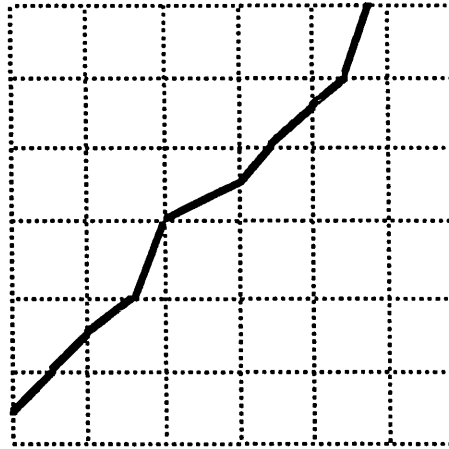
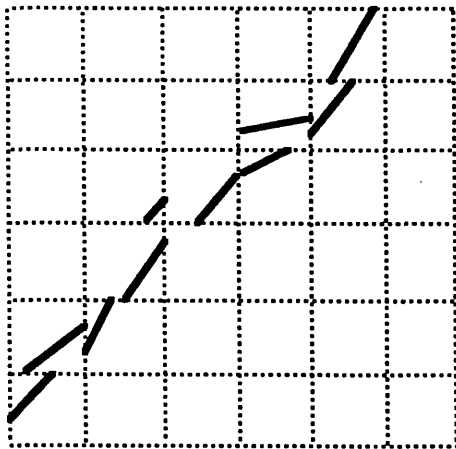
The study involves analysing digital AVHRR images of a climatically insensitive open lake (one which possesses a surface outlet), that has been well surveyed by ground-based methods. The lake selected for the case study must ideally be within range of the Dundee Satellite Station for ease of data acquisition. In addition, a high



Section of a digitised image with an edge passing through it from the bottom left corner to the top right corner. Edge detection is done locally within neighbourhoods of pixels shown by the circle.



For each neighbourhood determine the brightest (B) and darkest (D) pixel. If  $(B - D)$  is greater than a set threshold then an edge with luminance  $(B+D)/2$  passes through the neighbourhood. The edge can then be plotted depending on the high resolution model of the image used.



Operation is repeated for every neighbourhood within the image. The result is a disjointed edge plot. A line-joining algorithm is used to join the individual edges into a continuous line.

Figure 3.2 Stages in the local isoluminance contour (LIC) method of edge detection.

resolution TM (Landsat) image of the lake will be helpful in assessing the effects of coarse resolution and for comparison of techniques.

From consideration of the above, it was decided to select Lough Neagh, in Northern Ireland, as the case study. The choice of Lough Neagh has the following advantages :

- It is one of the largest lakes in Northern Europe, and has an area of  $\sim 390 \text{ km}^2$  which is a good size for the purposes of this project.
- It is stable in size because it is an open lake.
- Since it is within range of Dundee Satellite Station, AVHRR coverage is readily available.
- Landsat data for the UK is available from the Royal Aerospace Establishment (RAE) at Farnborough.
- The region has been well surveyed by the Ordnance Survey.

### 3.2 The development of a simple area algorithm

#### 3.2.1 Introduction to the main problems

There are five main problem areas to be considered:

- i) Edge pixels
- ii) Atmosphere
- iii) Geometrical distortion
- iv) Pixel effects
- v) Automation

##### *i) Edge pixels*

Edge pixels have a brightness intermediate between the water and the land. This fractional brightness represents the fractional area covered. Thus a pixel which contains a fraction  $\rho$  of land at brightness  $l$  and  $(1-\rho)$  of water at brightness  $w$  will have a brightness  $i$  such that:

$$i = \rho l + (1-\rho).w \quad (3.4)$$

An AVHRR image of Lough Neagh (total area ~390 km<sup>2</sup>) has some ~140 edge pixels, so it can be seen that a substantial fraction of the total area of Lough Neagh is contained within the edge pixels. Since the ratio of lake perimeter to lake area increases with decreasing lake size, the fraction of total area contained within the edge pixels increases for these smaller lakes. For lakes with a more developed coastline the fraction of edge pixels will be greater still. These points can be summarised in the following equation (Hutchinson, 1975) :

$$P = D_L 2(\pi A)^{1/2} \quad (3.5)$$

where  $P$  is the lake perimeter,  $A$  is the lake area and  $D_L$  is the development of the coastline.

$D_L$  can never be less than 1, which represents a circular coastline.

Other factors to consider are :

- The brightness of the lake and/or shore may not be uniform due to variations in brightness temperature (for infra-red channels) or reflectivity (for the visible/near infra-red channels).
- The contrast between the lake and shore may be very small under certain conditions or in certain wavebands.
- These effects may be seasonal (eg. lake vegetation in summer, lake ice in winter, land cover variations) or diurnal (eg. there will be no data for visual wavebands at night).
- There may be some instrumental effects dependent on the waveband of the detector used. For example AVHRR channel 3 (near infra-red) has severe noise problems and it is likely that this data will be unusable. There will also be less pronounced detector noise on the other channels.

## *ii) Atmosphere*

There are two possible problems that are related to the effects of the atmosphere :

- Absorption and scattering caused by clouds, fog, aerosols or air molecules

- Distortion due to refraction of the image; this is unlikely to have any great effect since the lake is contained within only a small angular part of the total image, and is rarely displaced by more than about  $30^\circ$  from the nadir of the satellite and atmospheric refraction only becomes significant at very low angles of elevation.

### *iii) Geometrical distortion*

The factors to be taken into account are :

- There will be a  $\cos\theta$  projection of the lake area being imaged by the satellite due to the angle  $\theta$  between the viewing direction of the satellite and the normal to the lake surface.
- There are two components to this angular projection. Firstly, there is an angle between the viewing direction and the nadir direction of the satellite. Secondly, earth curvature will result in the plane of the lake being turned by a further angle relative to the satellite nadir direction which must also be taken into account when applying the total angular correction to the projected area.
- The distance between the sensor and the lake will determine the relationship between the angle that the mirror sweeps through per pixel and the corresponding across-track dimension the pixel represents on the ground.
- Since the attitude control of the NOAA satellite ensures that pitch is continuously maintained around the orbit (Barbière, personal communication, 1991), successive scan lines can be considered to be arcs of great circles perpendicular to the sub-satellite track and thus get closer together on the Earth's surface further away from nadir. This situation is analogous to longitude lines being closer together further from the equator.
- The across-track component of earth rotation will have no effect on the measured area since the result will simply be to skew the image. The along-track component will, however, result in an increase in the effective area since the distance between successive scans will be larger (for a retrograde orbit) than would be the case for a non-rotating earth.

### *iv) Pixel effects*

These are :

- The Instantaneous Field of View (IFOV) of the optical system which will determine the total area being integrated into one pixel. There may be an overlap of the IFOVs of successive pixels resulting in the same piece of land or water contributing to two adjacent pixels, although the symmetrical nature of this effect means that it is unlikely to bias the area measurement.
- The actual dimensions of the pixel, both in the across-track direction (determined electronically) and along track direction. The latter is determined by the scan rate of the instrument and the velocity of the satellite (a function of altitude). The pixel dimensions will therefore vary as the altitude of the satellite varies. This effect can be considered as part of the problem of geometric correction of pixel areas.

#### *v) Automation*

Any technique for measuring large numbers of lake areas on a regular basis will have to be devised with automation in mind. To this end the guidelines for the technique should be :

- The technique devised should be fast, ie. complicated algorithms should be avoided (if possible) and the number of stages of computation should be minimised.
- The technique should be applicable to a large variety of lakes without requiring special modifications for each type.

As can be seen from previous discussion, the major problem to be overcome is how to extract the area of the lake contained within the edge pixels which on its own represents a potential error of ~20% , nearly an order of magnitude greater than the aim stated earlier. Once the basic pixel area has been extracted from the data, the other main effects of geometrical distortion and the pixel dimensions can then be corrected for using the necessary information which can be obtained from the images and from Dundee Satellite Station.

The rest of this section is thus devoted to devising a method for extracting the basic pixel area of the lake from the data.



### 3.2.2 Approaches for extracting the basic pixel area

There are two main approaches to this problem :

- i) Utilisation of the spatial image data to count pixel areas. For example, the area contained in each edge pixel could be obtained by utilising the fractional brightness relation using the values of the brightnesses of adjacent pixels which are totally land or water.
- ii) Reduction of the image to a one-dimensional frequency histogram of number of pixels of brightness  $i$  against brightness  $i$  and identifying the component of the brightness distribution that represents the lake. Selection of thresholds for land and water as proposed by Malila and Nalepka (1975) would enable the fractional brightness relation to be applied to the intermediate edge pixels. Since such approach involves the minimum of spatial information and requires no extensive calculations it is to be preferred to i).

The other main consideration that favours the second approach is that it allows a fuller investigation into the edge pixels and promotes a clearer understanding of the problems involved.

### 3.2.3 Idealised one-dimensional histogram

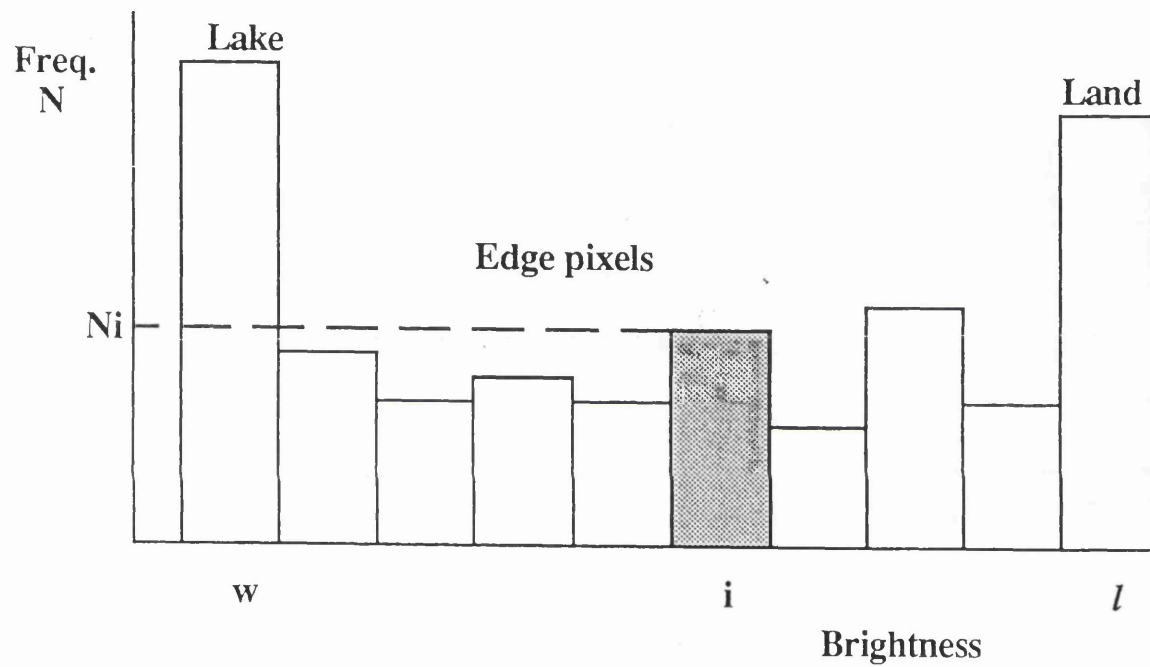
Consider an idealised histogram, making the following assumptions :

- i) The water is uniformly bright.
- ii) The land is uniformly bright.
- iii) There is no cloud cover or other atmospheric effect.
- iv) The pixels do not overlap.

Assumptions (i) and (ii) mean that all of the edge pixels will have intensities intermediate between the intensities of the lake and land.

A window (eg. rectangular) can be drawn around this ideal lake and a histogram plotted of frequency against brightness for all pixels within the window. For water at one brightness,  $w$ , and land at another brightness,  $l$ , and edge pixels in between, the resultant histogram is shown schematically in Figure 3.3.

The fractional area of water (in pixels) contained within  $N_i$ , the number of pixels of brightness  $i$ , is given by :



**Figure 3.3** Idealised histogram with uniform land and water illustrating the fractional brightness relation for intermediate pixel intensities.

$$A_i = [(l - i)/(l - w)]N_i \quad (3.5)$$

The total lake area  $A$  (in pixels) is given by :

$$A = \sum_{i=w}^{i=l} A_i \quad (3.6)$$

$$\therefore A = \sum_{i=w}^{i=l} [(l - i)/(l - w)]N_i \quad (3.7)$$

It is now necessary to study whether this simple approach can be applied in the case of a real histogram from a Lough Neagh image.

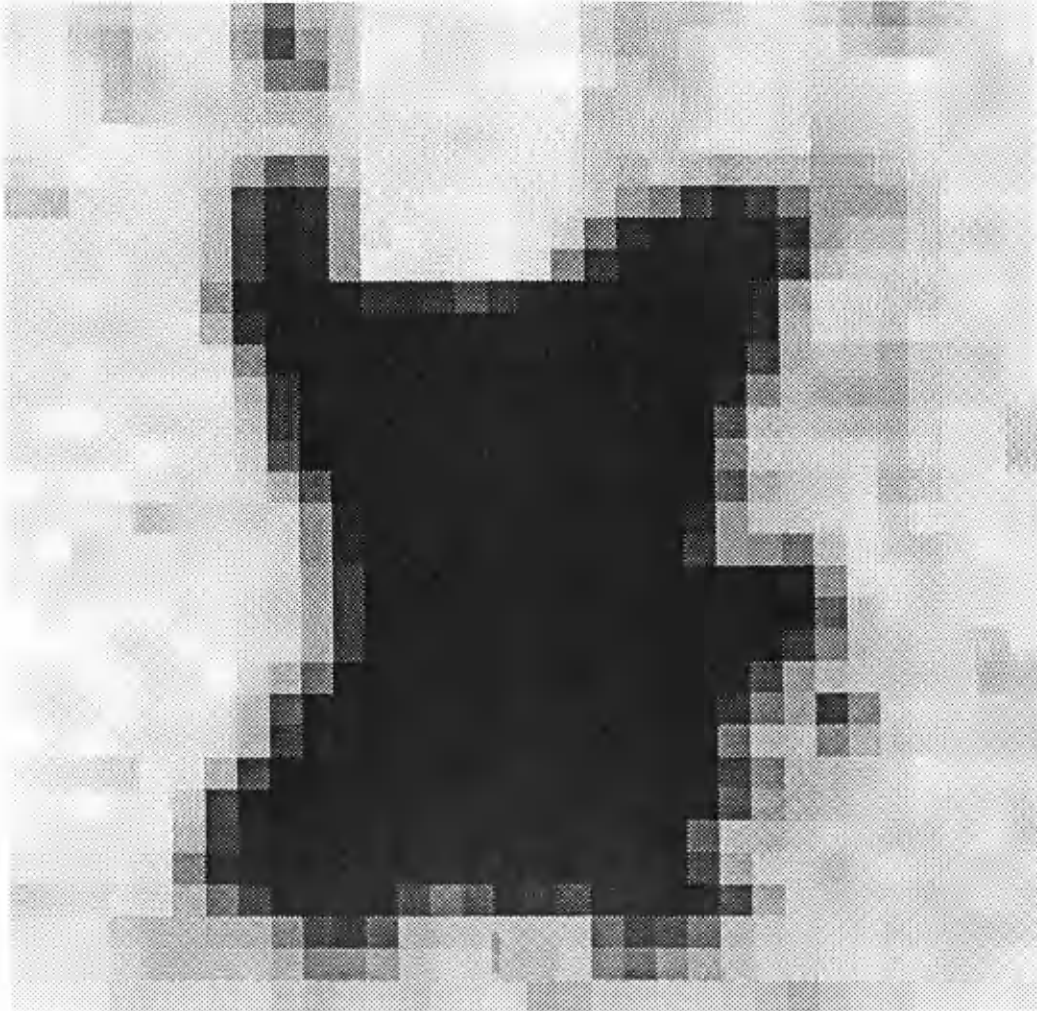
#### 3.2.4 Real one-dimensional histograms

If an appropriate area of the image is selected, enclosing the lake, and a histogram output of the number of pixels of brightness  $i$  against brightness  $i$  is obtained (see Figure 3.4), the following features are apparent :

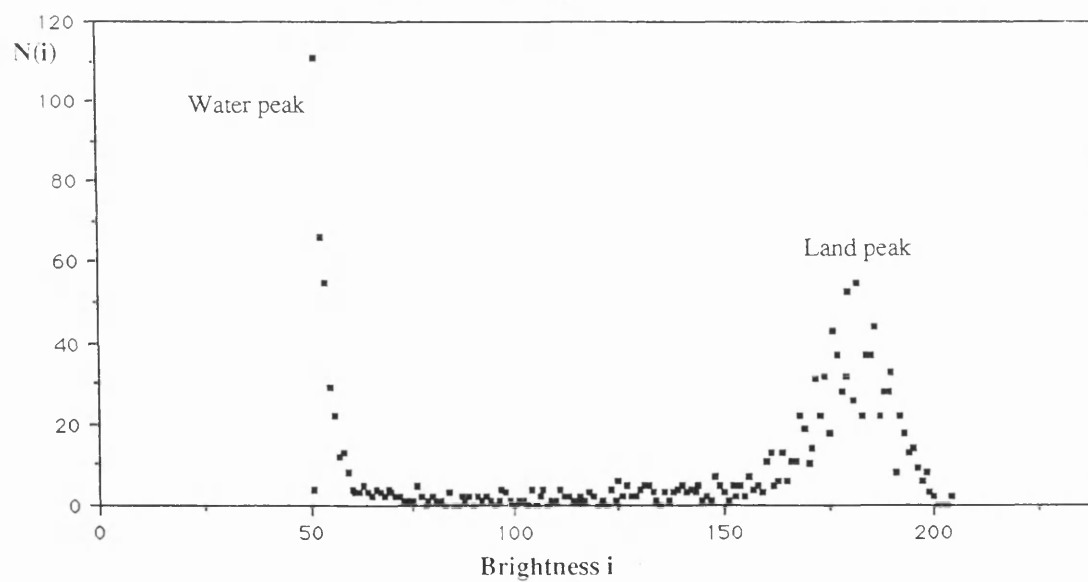
- i) The lake shows up as a sharp peak in the distribution (a good approximation to the ideal case) which is to be expected since the water is a homogeneous body with a uniform reflectivity.
- ii) The land distribution is more widely spread, presumably due to the different ground cover types with varying reflectivities.
- iii) The edge pixel distribution stretches between the two main peaks, and merges with the land distribution towards the land end.

#### 3.2.5 Problems caused by the broad land distribution

- i) The edge pixel distribution extends into the land distribution, with the pixels containing a large proportion of land merging with the land distribution. The fractional areas of these pixels are not directly obtainable. These edge pixels merged with the land distribution contain a low (up to ~30%) proportion of water,



Channel 2 histogram



**Figure 3.4** A typical channel 2 AVHRR image of Lough Neagh (top); the effects of poor resolution can clearly be seen. The image data can be represented as a histogram, as shown.

thus indicating that the total amount of water contained in these pixels as a percentage of the total is only a few percent. Since this is of the same order as the accuracy to which we wish to measure the area of the lake, it is clear that any technique devised will have to be capable of obtaining the area of water contained within these pixels. Failure to extract the area from these pixels probably led to the bias observed by Malila and Nalepka.

***Possible solutions :***

- It may be possible to devise curve-fitting techniques to subtract the land distribution. This would require knowledge of the land distribution function which will vary from lake to lake, seasonally and sometimes with the time of day (eg. different materials will heat up at different rates during the day and thus have changing brightness temperatures).
  - Thresholds could be defined for water and land, regarding all pixels outside these thresholds to be wholly land or lake, and apply the fractional brightness relation defined in equation (3.7) to the pixels intermediate between the thresholds (Malila & Nalepka 1975). Knowledge of the edge pixel distribution would be desirable in order to know how much of the area of the lake would be excluded by this method.
- ii) The broad land distribution modifies the shape of the edge pixel distribution. Since the land distribution is broad, pixels containing the same proportion of land may not have the same brightness (because the land they contain can have a broad range of brightness values). This broadening effect is greater at the land end of the distribution because the pixels there contain more land.

***Possible solution :***

- It may be possible to de-broaden the edge pixel distribution. This would require knowledge of the functional form of the land distribution (which is variable for reasons mentioned above) and the pre-broadened edge pixel distribution function.

The above solutions require a knowledge of the edge pixel distribution, and this was investigated. As will be seen below, this in turn led to simpler solutions to the problems of the broad land distribution and the area measurement itself.

### 3.2.6 The nature of the edge pixel distribution

In order to determine the edge pixel distribution, firstly in an idealised case, one possible numerical method would be to draw a coastline randomly on a pixel and calculate the fraction of the pixel covered by water. If the process is repeated enough times a frequency distribution of the fractions can be obtained.

In order to draw such a coastline it is necessary to define both the position and the shape of the curve in some random way. Since a very large number of repeats would be required to give a reasonably smooth distribution, a computer program, with algorithms to define the random coastline, would be required. However, the method for defining the random nature of the position and shape of the coastline is not obvious. Therefore only a simple numerical model was undertaken and is described in the next section. However, even without determining the shape of the distribution, it is possible to make some deductions about it.

Consider the coastline drawn on the pixel (as before) as a line defining the boundary between land and water. Assuming that there are no inherent geomorphological preferences which affect the random nature of a coastline, then the water and land should be interchangeable (ie. an island would have the same edge pixel distribution as a lake). Thus if the process mentioned above is repeated an infinite number of times, then for every pixel which has land one side of a given boundary and water the other there will be a pixel with a similar boundary which has the water and land interchanged. Thus the distribution will be symmetrical about the mid-point between the land and water. This implies that the average area of water contained in an edge pixel is 50%. This assumption leads to a very simple algorithm for lake area, and turns out to be the same as the method proposed by Barker that was mentioned in section 3.1.2.

### 3.2.7 Simple lake area method assuming symmetrical edge pixel distribution

Consider Figure 3.5, which shows an idealised symmetrical histogram. Since there are the same number of edge pixels to the left of the distribution as there are to the right,

$$\frac{\sum_{w+1}^{l-1} [(l-i)/(l-w)] N_i}{\sum_{w+1}^{l-1} N_i} = \frac{1}{2} \quad (3.8)$$

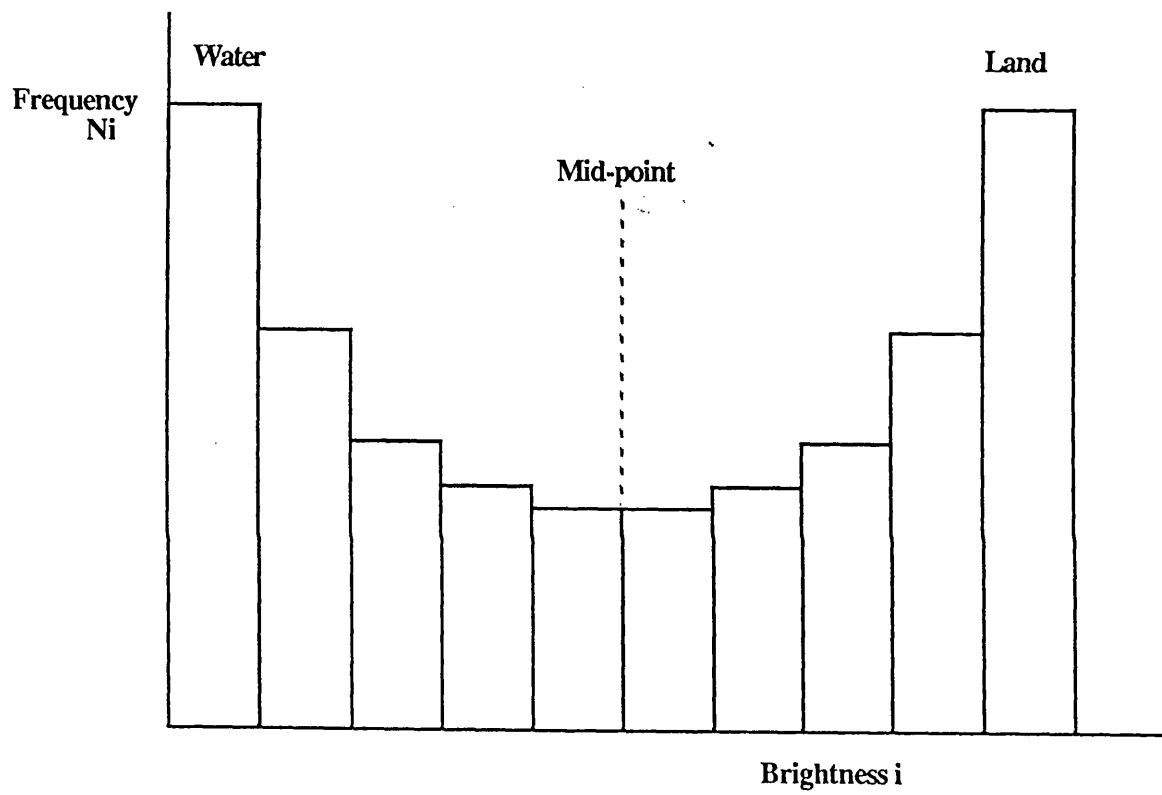


Figure 3.5 Idealised histogram with a symmetrical edge pixel distribution.

ie. average coverage is 50%, assuming a symmetrical edge pixel distribution.

$$\therefore A_E = A_p \sum_{w+1}^{l-1} [(l-i)/(l-w)] N_i = \frac{1}{2} \sum_{w+1}^{l-1} N_i \quad (3.9)$$

where  $A_E$  is the area of lake contained in edge pixels and  $A_p$  is the area of one pixel

Also,

$$\frac{1}{2} \sum_{w+1}^{l-1} N_i = \sum_{w+1}^{(l+w)/2} N_i \quad (3.10)$$

$$\therefore A = A_p N_w + A_E = A_p \sum_{i=w}^{(l+w)/2} N_i \quad (3.11)$$

Hence by simply placing a threshold halfway between the two peaks in the histogram, and counting up all the pixels to the lake side of the threshold, we obtain the total area of the lake in pixels.

This method should be applicable to real histograms, provided that the following assumptions are valid :

- i) The edge pixel distribution is symmetrical about the mid-point between land and water.
- ii) The half-width of the land distribution (at minimum) is not greater than half the distance between the mean values for land and water, ie. the land pixels do not spread beyond the mid-point between the land and lake peaks.
- iii) The broadening of the edge pixels due to the broad land distribution has a negligible net effect on those edge pixels which are capable of being broadened into the other half of the distribution. This implies that the number of edge pixels originally in the lake half of the distribution which are broadened into the land half of the distribution approximates closely to the number of edge pixels broadened into the lake half of the distribution from the land half, thus rendering



the net exchange of edge pixels negligible.

- iv) The land and water *within the edge pixels* have the same distributions as the *overall* land and water distributions.

It was decided to make a detailed study of this symmetrical edge-pixel distribution technique, with two main objectives. Firstly, it is important to investigate whether assumption (i) is valid. Secondly, it is important to determine the minimum statistical error on the lake area measurement using this technique for the real case of a non-infinite coastline.

As a first stage in any future, more fundamental, study of the problem of edge pixels, it was also decided to investigate the precise shape of the edge pixel distribution, as discussed above, by deriving a simple numerical model of an infinite coastline .

These studies are described in the following section.

### **3.3 Simulated Results**

#### **3.3.1 Study of the edge pixel distribution for a non-infinite coastline**

Since the case study is to be performed on AVHRR images of Lough Neagh, it is desirable to study the edge pixel distribution associated with Lough Neagh's coastline. This real coastline differs from an infinite one in two main aspects :

- i) The real coastline is not infinite.
- ii) The real coastline is closed.

Point i) means that there will be statistical fluctuations on the areas obtained by using the simple histogram technique, which will determine the minimum statistical error of this method. Point ii) means that there is a slight positive curvature of the coastline, which may result in some asymmetry of the edge pixel distribution due to there being a slight preference for the land to be on the outside of the curve. The study will help to decide if this is significant in determining the area of the lake.

#### ***Precision of the method***

Given the assumptions described in section 3.2.7, the precision of the method is given by :

$$\Delta A = \sqrt{(\Delta B)^2 + (\Delta M F_E(M))^2} \quad (3.12)$$

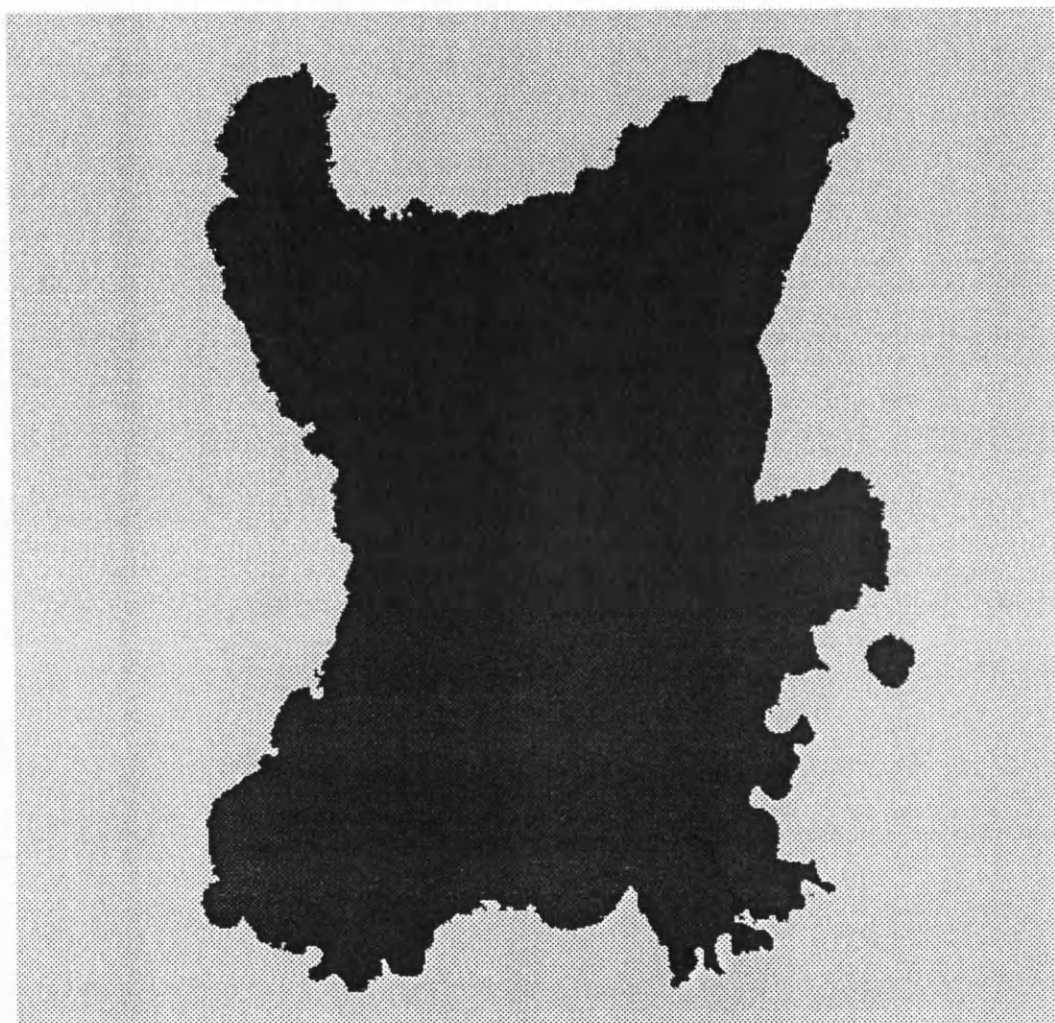
where  $\Delta A$  is the error on the area of the lake,  $\Delta M$  is the precision with which the mid-point of the edge pixel distribution can be determined,  $\Delta B$  is the error due to the random fluctuations in the edge pixel distribution and  $F_E(M)$  is the edge pixel distribution at the mid point.

$\Delta B$  was investigated (this being the best possible precision for the method) using a Landsat TM image of Lough Neagh (pixel size 30 metres). The image was density sliced in order to have the land and water as two discrete intensities as shown in Figure 3.6 (ie.  $\Delta M = 0$ ). This black and white image was then operated on, so as to average the intensities in each  $32 \times 32$  pixel square, thus simulating AVHRR resolution (see Figure 3.7). This operation was repeated for 64 evenly sampled phases across the pixel grid (simulating a number of different AVHRR passes), and the area of the lake was calculated for each image using the simple method described earlier. The results from this statistical study show that, for Lough Neagh, the method has a best possible precision of  $\pm 0.5\%$ . We now compare this with the result expected from binomial statistics. If we assume that any edge pixel has a probability  $p = 1/2$  of being to the lake side of the mid-point, and there are about 140 edge pixels in any given image :

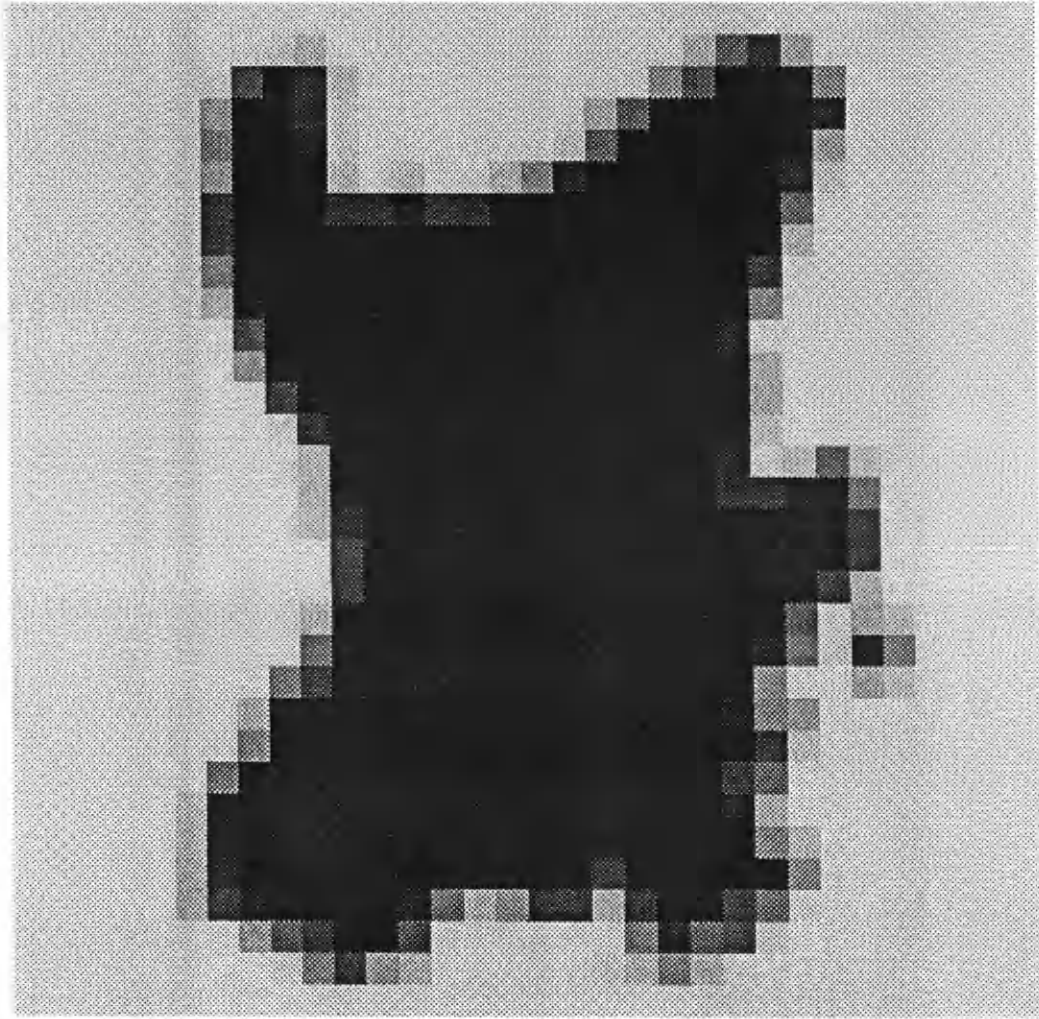
$$\begin{aligned} \sigma_{binomial} &= \sqrt{N p (1-p)} \\ &= [\sqrt{140}]_{1/2} \\ &\approx \pm 6 \text{ pixels } (\pm 1.5\% \text{ of total area}) \end{aligned}$$

The reason why the results obtained from the analysis of the TM image of Lough Neagh are better than those expected from binomial statistics is that the fraction of water contained in one pixel is not independent of the fractions of water contained in the others. As the pixel grid is moved across the lake, the pixels on the left edge are becoming more filled with land, whereas those on the right side of the lake are becoming more filled with water, and are moving to the lake side of the mid-point, thus the net result is only a small change in the measured area of the lake.

It was found that the shape of the edge pixel distribution could be modelled by assuming the coastline across a pixel could be represented by a straight line. A detailed description of the modelling of the edge pixel distribution is given in



**Figure 3.6** A high resolution Thematic Mapper image of Lough Neagh which has been density-sliced in order to render the land and water at two discrete intensities. Rivers and small islands have been removed.



**Figure 3.7** Here, the image in Figure 3.6 has had its resolution degraded to approximate that of AVHRR by binning up the pixel values.

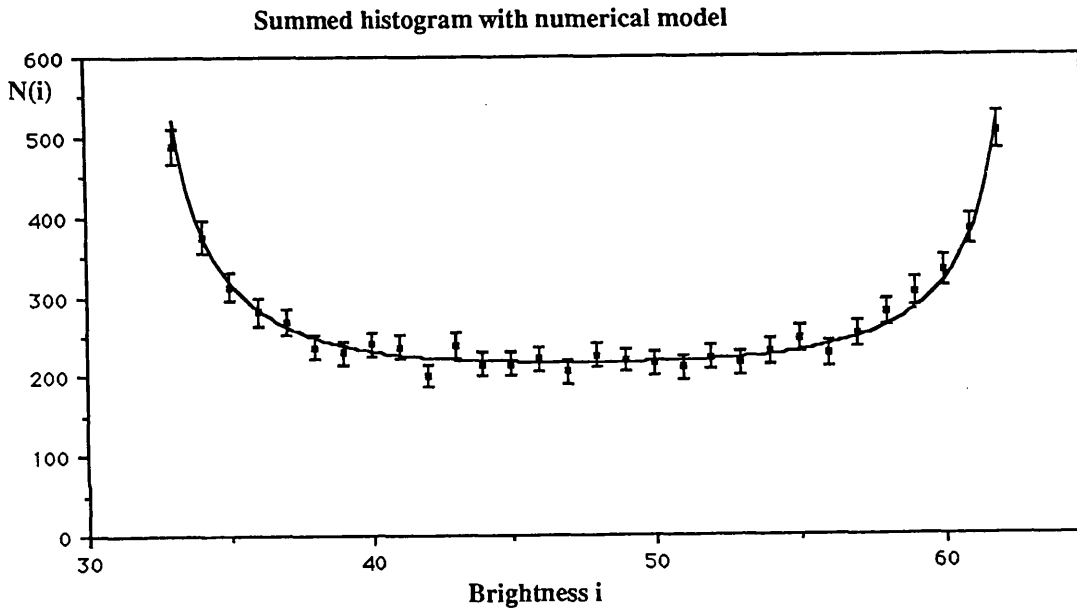
Appendix A and a summary is presented here. If any orientation of the straight line is equally likely, an analysis of the probabilities involved gives a  $P(A)$  against  $A$  (probability of area  $A$  against  $A$ ) relationship shown in Figure 3.8. This simple model shows a good fit to the averaged data, thus it will be useful for future detailed investigation of the effects of asymmetry due to land broadening and mid-point location error on the final precision.

### **3.3.2 Investigation of the precision of the LIC edge detection method**

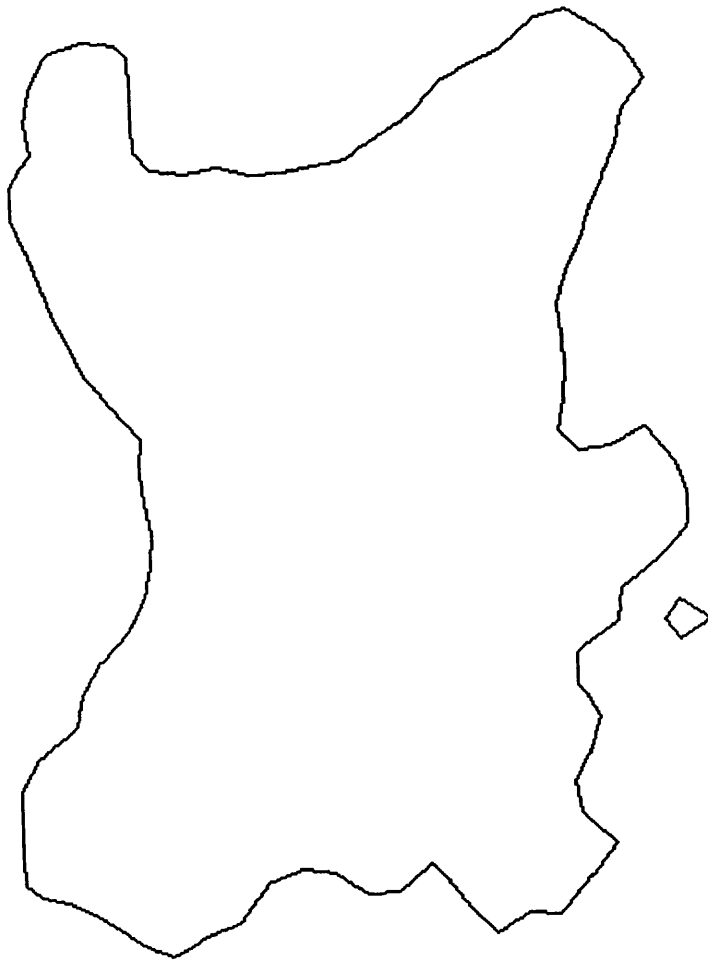
For comparison, it was decided to apply the LIC edge detection method to the same resampled images. This would allow a determination of the best possible precision attainable by the local isoluminance contour technique when applied to data of similar resolution to that of the AVHRR and ATSR. In principle, the LIC method should be capable of determining the exact position of the isoluminance contour within the edge pixel for the situation of uniformly bright land and water. If this were the case, a perfect assessment of the area should occur for the image produced by each phase of the resampling grid. In practice, at certain points along the coastline, sub-pixel features result in there being no pure water or land pixels present in the local intensity sampling regions, thus the position of the contour in the central cell is not determined w.r.t. the true intensity levels of land and water. The extent to which this occurs for a given lake outline determines the accuracy of the technique. The precision for the LIC method when applied to the density-sliced and resampled images of Lough Neagh (see Figure 3.9) was a remarkably low 0.2%, which is particularly encouraging when considered in the light of its potential for measuring the areas of closed lakes in semi-arid regions which may have non-uniform coastlines due to peripheral vegetation, salt flats and other effects. A second simulation was performed using MSS imagery of a closed lake and the results are described in the next chapter.

### **3.4 Application of the histogram technique to AVHRR images.**

The histogram technique described in section 3.2.7. was applied to cloud-free AVHRR images of Lough Neagh. The area of Lough Neagh (and the surrounding small lakes) was obtained from the UK Ordnance Survey at Belfast (Sharkey, personal communication, 1988), and found to be 390.2 km<sup>2</sup>. The peaks for the water and land were found by excluding the wings of the distributions outside the full-width half-maximum (to eliminate most of the edge pixels and thus minimise their biasing effect) and finding the mean of each. Twenty-nine AVHRR images of Lough Neagh were analysed, and the results are shown in Table 3.1. The areas have been geometrically corrected for the various distortions outlined in Section 3.2.1, the



**Figure 3.8** The sum of all histograms obtained from the binup histogram analysis is shown, together with the edge pixel distribution calculated using the numerical model described in Appendix A. The error bars shown are  $\pm\sqrt{N(i)}$  since the value for each bin is much less than the greatest possible number (ie. Poisson statistics apply).



**Figure 3.9** Result of local isoluminance contour (LIC) edge detection method applied to typical simulated AVHRR image (Figure 3.7). Comparison with the original density-sliced TM image (Figure 3.6) demonstrates the relative success of the technique.

**Table 3.1** Results for area of Lough Neagh obtained from AVHRR data. The areas (km<sup>2</sup>) have been corrected for satellite altitude and viewing angle as described in Appendix B.

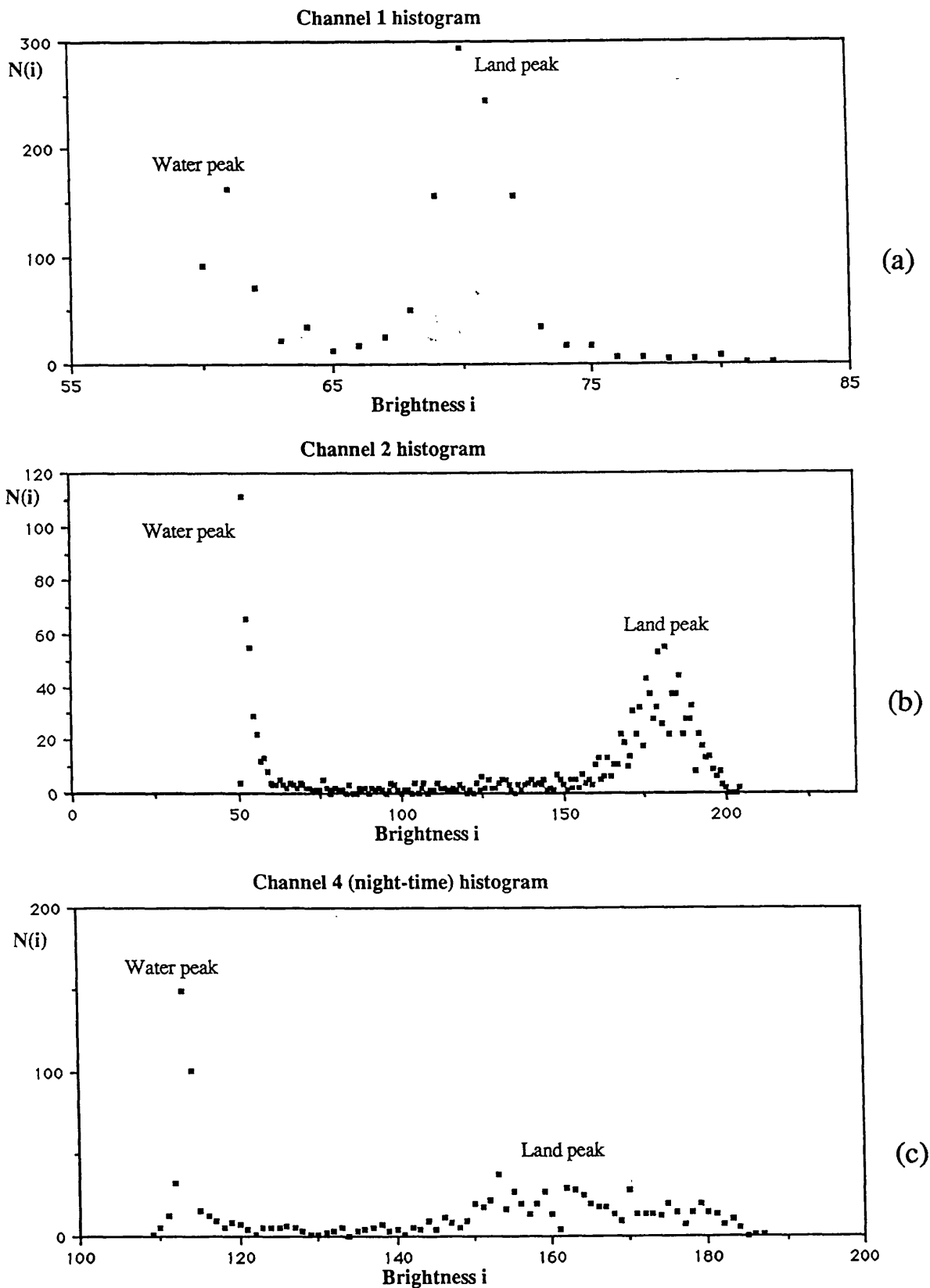
	Day-time				Night-time	
	Channel-1	Channel-2	Channel-4	Channel-5	Channel-4	Channel-5
	378.6	393.3	365.3	364.6	386.4	392.9
	—	391.8	371.1	373.8	381.2	377.6
	371.1	386.6	356.3	365.2	383.9	383.9
	396.1	395.1	391.4	393.7		
	367.5	385.1	378.4	373.7		
	381.6	393.8	371.5	374.9		
Mean	379.0	391.0	372.3	374.3	383.8	384.8
$\sigma$	$\pm 9.9$	$\pm 3.8$	$\pm 10.9$	$\pm 9.9$	$\pm 2.6$	$\pm 7.7$



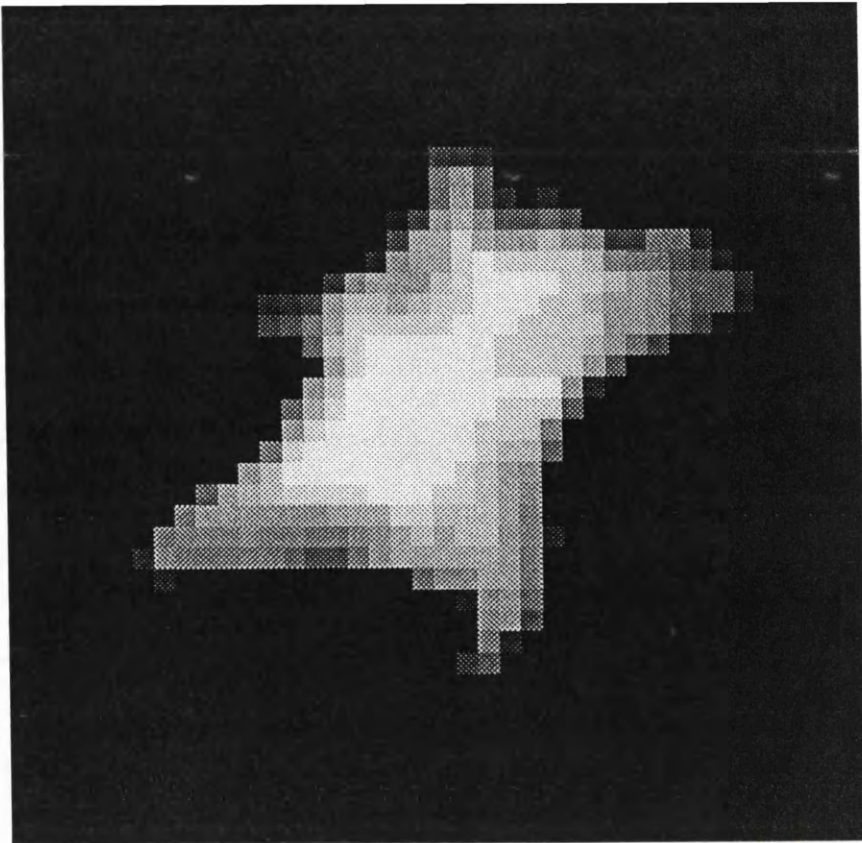
mathematics of which are more fully described in Appendix B.

Channel 3 was not used since the data was noisy, making it impossible to distinguish the water and land peaks in the histogram, and channels 1 and 2 clearly give no information at night. The following points can be noted :

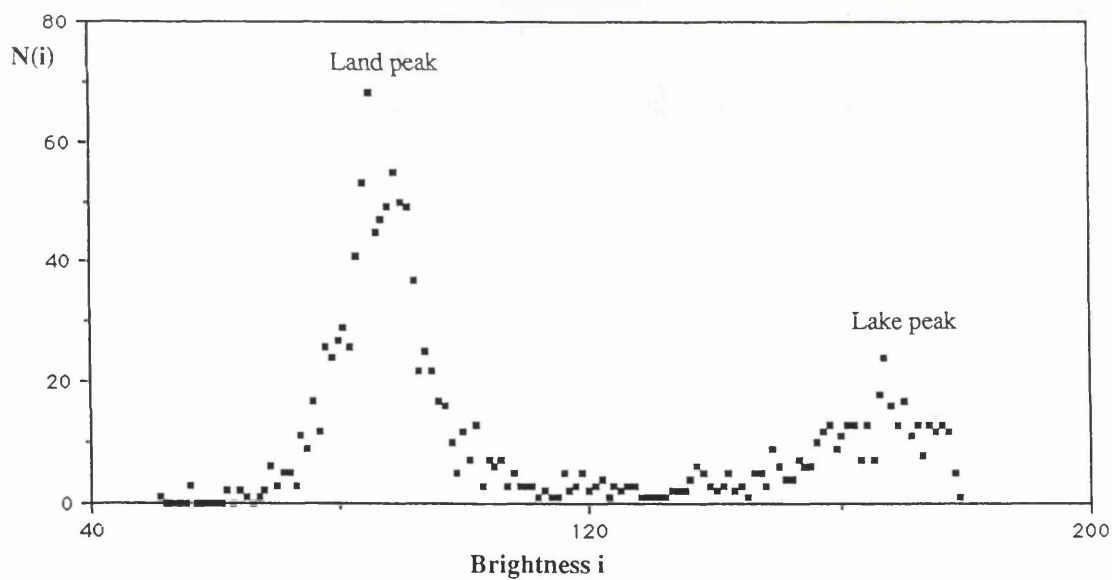
- i) Channel 1 gives a mean area which is 3% lower than the ground value, and the areas have a relatively large spread of  $\pm 10 \text{ km}^2$  ( $\pm 3\%$ ). This is due to the very poor contrast between water and land in the visible band resulting in a smaller separation between the peaks relative to their widths (see Figure 3.10).
- ii) Channel 2 gives the best results, with a spread of 1% and a mean area that is very close to the ground survey value. This is to be expected since there is good contrast between the land and water in the image (see Figure 3.4).
- iii) Daytime mean areas from channels 4 and 5 are 4% to 5% lower than the ground survey value, and have relatively large spreads. The reasons for these results are as follows:-
  - During the day, the land and lake are warmed by the sun.
  - In the early morning (~8 am) the water and land have similar brightness temperatures and the contrast between them is poor, hence there is a large scatter in areas calculated for images obtained at this time of day.
  - Later in the day, the contrast improves as the land heats up faster than the lake (because the heat is only stored in the surface layer of the land). However, the edge of the lake, where the water is shallow, heats up faster than the middle (see Figure 3.11), thus the brightness temperature of the water contained in the edge pixels is higher than the mean brightness temperature of the whole lake. This results in the true mid-point of the edge pixel distribution being shifted towards the land peak relative to the calculated mid-point; thus a low area is obtained.
- iv) The results for night-time channels 4 and 5 images are ~1.5% lower than the ground value, but the spread is low ( $\pm 1\%$ ), indicating that there may be some constant bias effect. It can be seen from the histogram in Figure 3.10 that the land



**Figure 3.10** Typical channel-1, -2 and -4 histograms ((a), (b), (c) respectively). The channel-4 histogram is similar to the channel-5 histogram obtained from the same pass, but the peaks are more widely separated.



Channel 5 (afternoon) histogram



**Figure 3.11** A typical afternoon thermal infrared image which has been contrast-stretched to show the edge warming effect. The different orientation of the image is due to the pass being in the north-bound direction. The histogram output of the image shows the spread of the water peak, thus giving an indication of the difference in brightness between the middle and edge of the lake.

(which is now cooler than the lake) is represented by a very flat, multiple-peaked distribution, but the water is now a very sharp peak, which is why the result is much improved. It was also observed that the contrast between land and water was better for channel 4 than channel 5. This is due to there being more absorption in channel 5 than channel 4; the differential absorption forming the basis for the 'split-window' atmospheric correction as described in Chapter 2.

For the channel 2 (day) and channels 4/5 (night) histograms, the estimate of the error  $\Delta M.F_E(M)$  is between 0.5% and 1%, which, with  $\Delta B = 0.5\%$  provides a total error  $\sim 1\%$ , similar to the precision observed. This fact, together with the low residual biases ( $\sim 0\%$  and  $1.5\%$  respectively) implies that the three assumptions for the method hold reasonably well. Therefore the simple method can provide lake areas that are accurate enough to detect aridity changes as small as  $\sim 1\%$ , as described earlier.

### 3.5 Conclusion

The results for this case study show that it is possible to measure the area of Lough Neagh with AVHRR data to an accuracy of a few percent; this is expected to be well-matched to the MSSL climate research application. The best results are obtained by using channel 2 during the day, when the thermal IR channels produce poor results, and by using the thermal IR channels at night, when channel 2 provides no information. The channel 2 results have a very high accuracy, and both these cases have a precision that is very close to the best possible precision (obtained by simulation with the Landsat image). This precision itself is surprisingly good, being due to the correlations in the fluctuations of the edge pixel water fractions, as described in section 3.2.7. In addition, the LIC method has been shown to have even better intrinsic precision than the simple histogram technique and, although there was only time for a preliminary evaluation in this study, may hold significant advantages when applied to the measurement of lakes in semi-arid regions.

The results show promise for the proposed global monitoring of closed lakes by the RA and the coarse resolution ATSR instrument on ERS-1, although the effects of the conical scan of ATSR have yet to be investigated. The potential problems of measuring lake areas in semi-arid regions are investigated in the next chapter, together with the application of an area measurement technique on a time series of AVHRR images of a climatically-sensitive closed lake in Ethiopia.

## Chapter 4

# Time Series Remote Sensing of a Climatically Sensitive Lake

### 4.1 Introduction

The studies performed in Chapter 3 demonstrated the feasibility of measuring lake areas with coarse resolution instruments such as AVHRR and ATSR to the order of a few percent. However, there are still some important questions to be answered, particularly with regard to the suitability of the various techniques for the measurement of areas of closed lakes. In considering the nature of closed lakes, there are some features which may potentially influence measurements:

- The lack of outlet, high evaporation rates and varying shorelines lead to the deposition of soluble mineral salts, indeed, such lakes are often referred to as salt lakes (eg. Great Salt Lake, Utah). Such deposition preferentially occurs in the coastline regions with shallow slopes since the margins are exposed at low water and, in addition, the water temperatures and evaporation rates are highest (eg. see Chapter 3). In such areas, large salt flats are common, and, as will be seen later, these have an albedo that is substantially different from that of the surrounding land.
- Since the majority of closed lakes occur in semi-arid regions of the world, it is possible that they may have a significant influence on the local hydrology of the area. One potential result of this might be the growth of vegetation around the periphery of the lake which may again have significantly different albedo from the surrounding land.

Both of these features are likely to have the same type of impact on the measurement of the areas of closed lake. In particular, the histogram technique relies on several assumption, and the validity of one of these is called into question by the occurrence of either of the situations described above. The main potential problem is that the histogram technique requires the mean brightness value of land *within the edge pixels* to be the same as that for the overall land brightness distribution within the histogram. Clearly, any effect which can alter the mean brightness of the land close to the shore, where the edge pixels exist by definition, relative to the overall land

brightness distribution will cause the area determined by the histogram technique to be in error. For example, in the case of the salt flats, where the albedo is likely to be higher than the surrounding land, a pixel which contains 50% water and 50% salt flat will have a higher brightness value than would a similar pixel which contained water and normal soil. As a result, the pixel containing an area of salt flat will be to the land side of the mid-point of the edge pixel distribution, and an underestimation of the lake area will result. This situation is analogous to the edge-warming effect encountered for daytime thermal infrared scenes of Lough Neagh in the previous chapter.

The LIC edge detection method potentially offers a means of overcoming this problem since its determination of the position of the coastline is based on the detection of *local* isoluminance contours. However, the local sampling area extends for a diameter of four pixels and can still be affected unless at least one pixel within the sampling region is completely filled by salt flat and another completely filled with water. In particular, this may present a problem where the extent of the salt flat is less than a pixel in size.

There is clearly a need for a validation study to determine which of the two methods is most suitable for the measurement of the areas of lakes in semi-arid regions. In addition, once the most suitable technique has been determined, it will be of considerable interest to apply it to a time series of AVHRR images of a climatically sensitive lake to see if we can detect the changes in area that might be expected.

#### 4.2 Lake Abiyata as a Case Study

The lake selected for the purposes of this study is Lake Abiyata in Ethiopia. The lake is in the highlands about 100 kilometres south of Addis Ababa and is in a valley that is part of the Rift system. There are several reasons why this lake is particularly suitable for both the technique assessment and the time series measurements:

- The intention for global lake volume change analysis is to use remotely-sensed measurements of both area (obtained from imaging radiometer data) and level (using data from Radar Altimeters). Despite the lack of usable altimeter data over this region due to loss of lock caused by the mountainous terrain, volume changes may be inferred since previous work, conducted between 1969 and 1973, has been done on the lake using *in situ* measurements. Indeed, volume changes can be more easily calculated if required, although the volume change analysis is beyond the scope of this thesis and it is the accurate measurement of lake area that

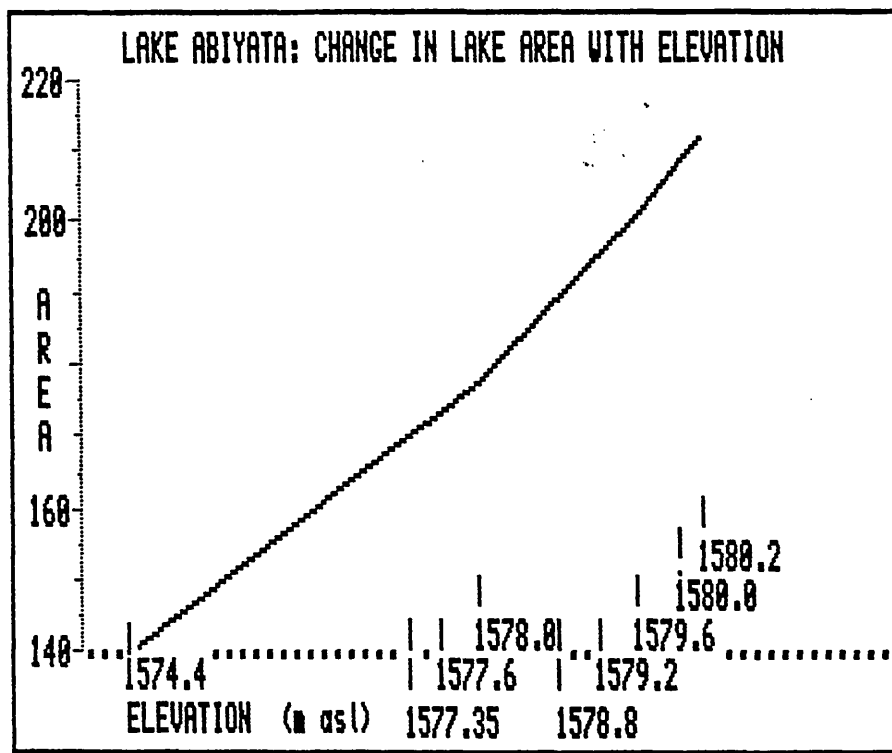
concerns us here.

- In this previous work (Makin *et al*, 1976), a combination of ground survey and bathymetry measurements was used to obtain the relationship between level and area (see Figure 4.1(a)). This information was used in combination with a time series of level measurements to calculate the changes in area (and hence volume) of the lake. The lake displayed considerable variation in area over the four year period (see Figure 4.1(b)), ranging from 170 km<sup>2</sup> to 210 km<sup>2</sup> during that time, thus we might expect to see similar variation in lake area over a relatively short time series of images spanning only a few years
- The lake is considerably smaller than Lough Neagh and approaches the lower limit of the size of lakes which we wish to measure. In addition, the lake has the typical features of a salt lake that may cause problems for the area measurement algorithms; these can be seen in Figure 4.2.
- The lack of a groundstation in the region to receive the 1.1 km High Resolution Picture Transmission (HRPT) coverage of the area is offset by the interest that various customers of AVHRR data have in the area of the Rift valley. Although there is insufficient storage capacity on board the spacecraft to enable storage of a complete orbit, NOAA will record segments of an orbit at full resolution; this data is known as Local Area Coverage (LAC) and is similar to direct broadcast HRPT data. NOAA have been recording data with increasing frequency over the region for customers since 1985.
- The main interest in the region is the study of vegetation cover which requires data from the short wavelength channels in the visible and near-infrared, thus only the daytime passes have been archived. It was found in the previous chapter that the results for lake area measurement obtained from channel 2 are superior to those from any other AVHRR channel due to the large contrast between water and land, thus the availability of this data is of importance in this study.

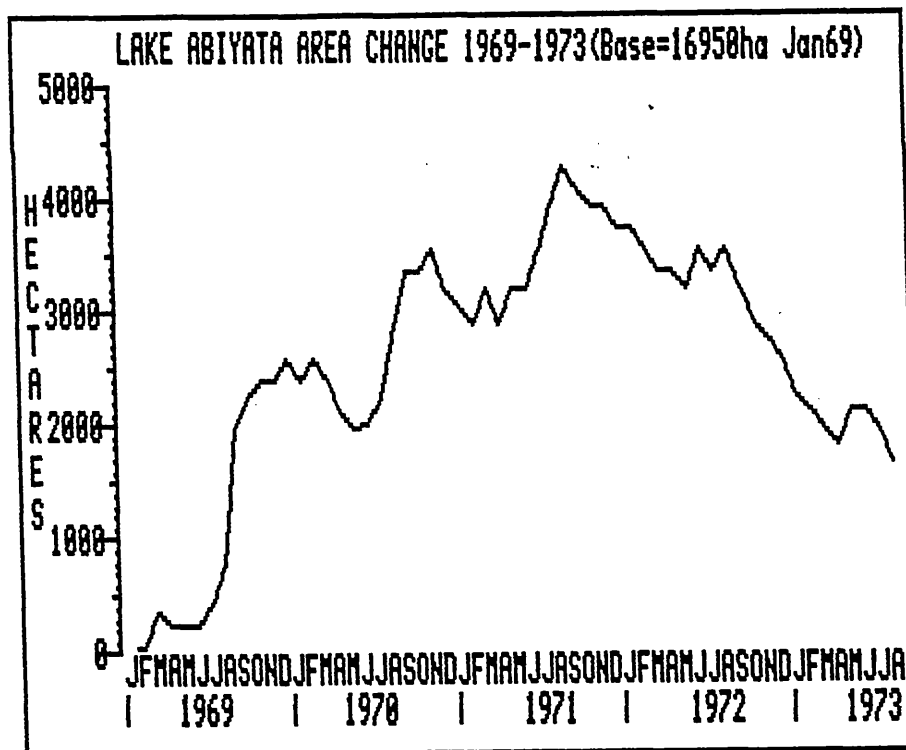
Before we can make use of the archived AVHRR data to obtain a time series of area measurements of the lake, it is first necessary to determine which of the two techniques is most suited to the task.

#### 4.3 A Comparison of Area Measurement Techniques for a Closed Lake

There are no *in situ* measurements of the lake area that are contemporaneous with the



(a)



(b)

Figure 4.1 The level - area relationship for Lake Abiyata derived by ground survey is shown in (a). The area cycle derived from four years of monthly level measurements and the level - area relationship is shown in (b), illustrating variable nature of the lake. (From Makin *et al*, 1976).



AVHRR data archive over the region, so an alternative means of evaluating the suitability of the area measurement techniques is required. The method chosen is similar to that adopted for the study of the precision of the two techniques in chapter 3, namely, the use of high resolution Landsat imagery which is then degraded to approximate AVHRR data. A Landsat MSS image of the region was obtained for 2<sup>nd</sup> November 1988 and is shown in Figure 4.2. The three band false colour composite image shows interesting features, some of which have already been mentioned above, which may have an impact on the results obtained by the area measurement techniques. Firstly, a large area of salt flats is visible at the north end of the lake, where the lake shore slope is at its most gentle. The marks of previous shorelines can also be distinguished on the land surrounding the lake and these are furthest apart at the northern end, thus confirming the shallow nature of the slope in this part of the lake. Note also that there are some much smaller bright areas of exposed salt flats along other parts of the shoreline; these may cause problems to both methods as outlined above.

The other important feature to be seen in the image is the close proximity of Lake Abiyata to the other lakes in the area. In fact, closer inspection of the marks left by ancient shorelines indicates that all three lakes were joined as one lake in previous times. The close proximity of the lakes will have an impact on the application of the histogram area measurement technique, since it will be necessary to exclude any pixels that originate from the other lakes from the analysis.

The problems mentioned above meant that a modification to the histogram technique was required to facilitate its successful operation in this area. The change implemented was to restrict the calculation of the brightness histogram to a closely-fitting user-defined polygon region around the lake. This modification had the following advantages:

- Pixels from the neighbouring lakes could be excluded from the analysis; this would not have been possible if a rectangular window of the type employed in the Lough Neagh case study had been used.
- The restriction of land pixels for inclusion in the histogram to those within the immediate vicinity of the shore meant that the mean value of the overall land peak in the histogram was more likely to be representative of the brightness value of the land contained within the edge pixels.



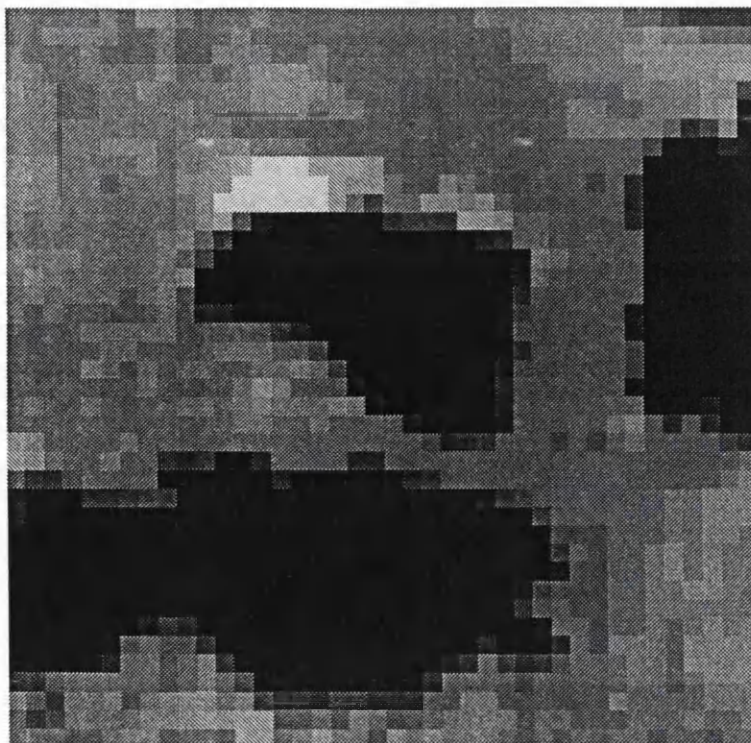
**Figure 4.2** Uncorrected false colour composite image (MSS bands 1, 2 and 4) of Lake Abiyata for 2<sup>nd</sup> November 1988. Note the large area of salt flats on the northern shore of the lake and smaller areas of salt around the perimeter. In addition, the marks of previous shorelines can be discerned, some of which encompass the other lakes.

The one disadvantage of restricting the number of pixels was that determination of the land peak of the histogram for such a polygon region would be more uncertain than would be the case for a rectangular window. This problem, however, was found to be more than offset by the second advantage mentioned above.

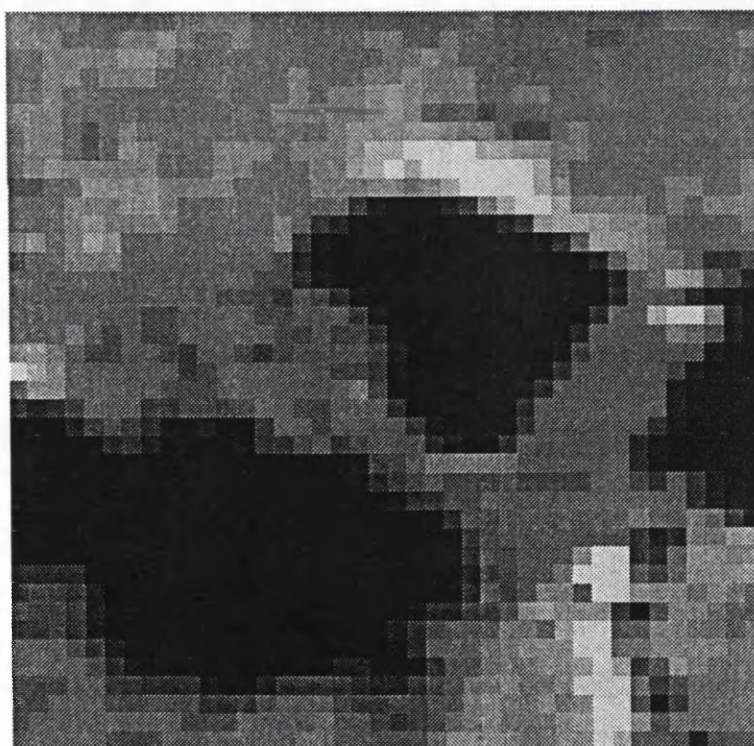
The MSS channel 4 image was selected for the resampling analysis because it covered a similar spectral range to AVHRR channel 2. The nominal resolution of MSS data is 56 metres across-track and 79 metres along-track, although the IFOV is 79 metres square, thus the data is oversampled in the across-track dimension in a similar ratio to that of AVHRR at nadir. It was decided to resample the data by averaging the pixels in every 14×14 pixel square in order to give an effective resolution of ~0.8 km across-track by ~1.1 km along-track. The main difference between this analysis and the one performed in chapter 3 is that the pixel values remained unaltered since the image was not density sliced. The resampling procedure was repeated for all possible phases across the pixel grid, resulting in 196 different resampled images and the two methods were then applied to each of these in turn.

An example of a resampled MSS band 4 image of the lake is shown in Figure 4.3(a) together with an AVHRR channel 2 image (Figure 4.3(b)) for comparison. Note that the salt flats at the north end of the lake (see Figure 4.2) are still clearly evident in both images. The close-fitting polygon employed in the histogram technique is shown in Figure 4.4 together with the resultant brightness histogram of the region. The main feature to note in the histogram is the indistinct nature of the land peak due to the restricted number of pixels included in the histogram, and the biasing contribution of the brighter pixels over the salt flats. The isoluminance contour for the same resampled image determined by the LIC method is shown in Figure 4.5. Note that the technique has had some success in detecting the small inlet on the northern shore of the lake.

The r.m.s. spread for the histogram technique was  $\pm 2.9$  pixels, representing ~2% of the total area of the lake. The most likely reason for the relatively poor performance of the algorithm is the small quantity of pixel values available for the determination of the land peak, which, as outlined above, is itself likely to be unrepresentative of the mean brightness of the land contained within the edge pixels. The LIC method, although still potentially troubled by sub-pixel features, returned an r.m.s. spread of  $\pm 0.9$  pixels, representing less than 1% of the total area of the lake. The mean values for the areas retrieved indicated that both methods were underestimating the area of



(a)



(b)

**Figure 4.3** A resampled MSS band 4 image is shown in (a), with the intention of simulating an AVHRR pass at nadir. A comparable AVHRR channel-2 image is shown in (b). The different orientation of the lake in the two images is due to the passes being south-bound in the case of MSS and north-bound in the case of AVHRR.



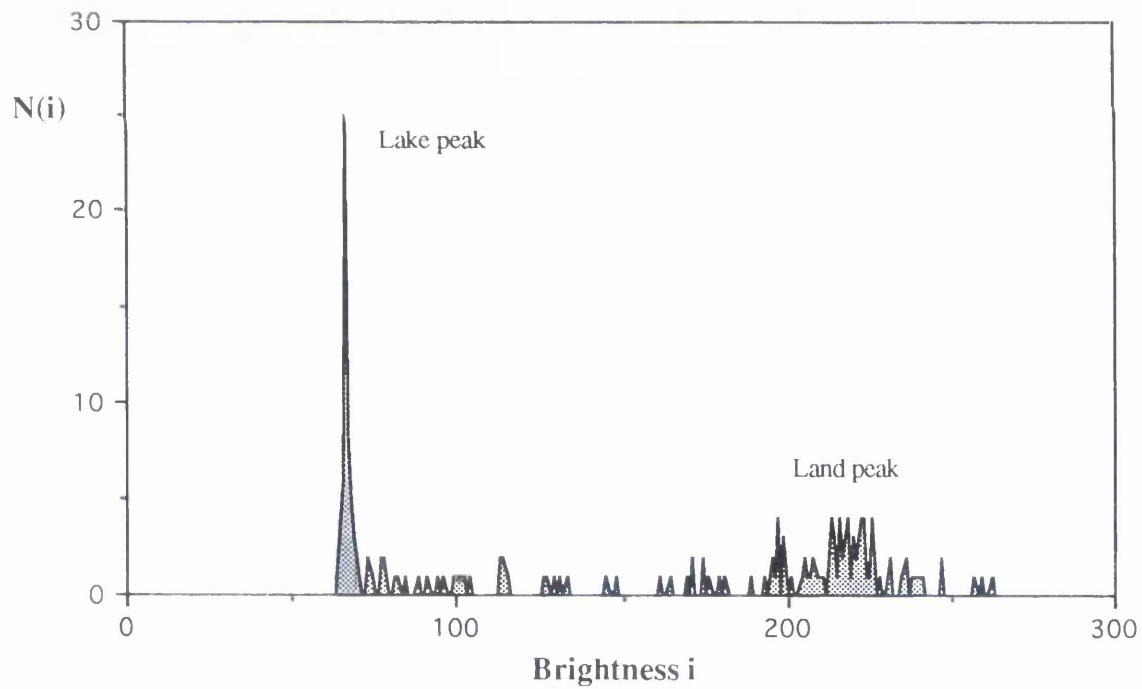
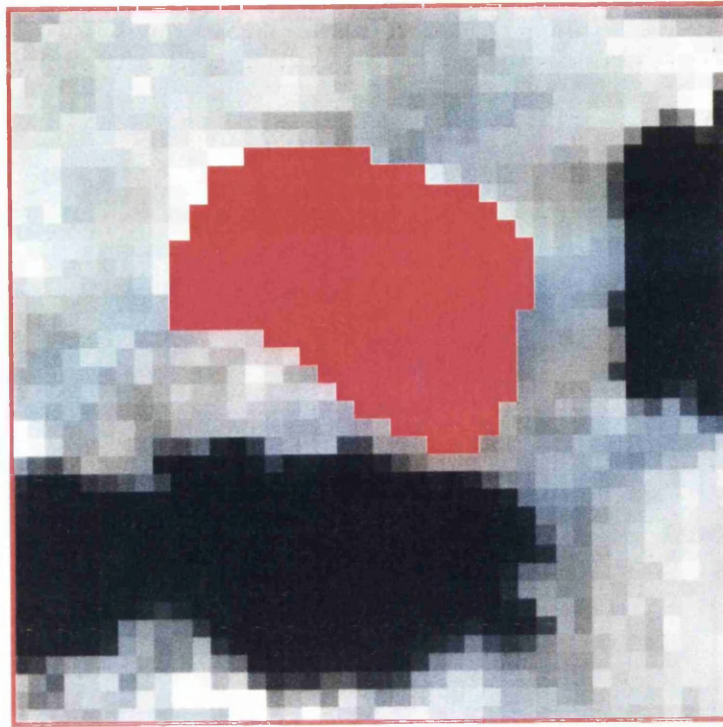
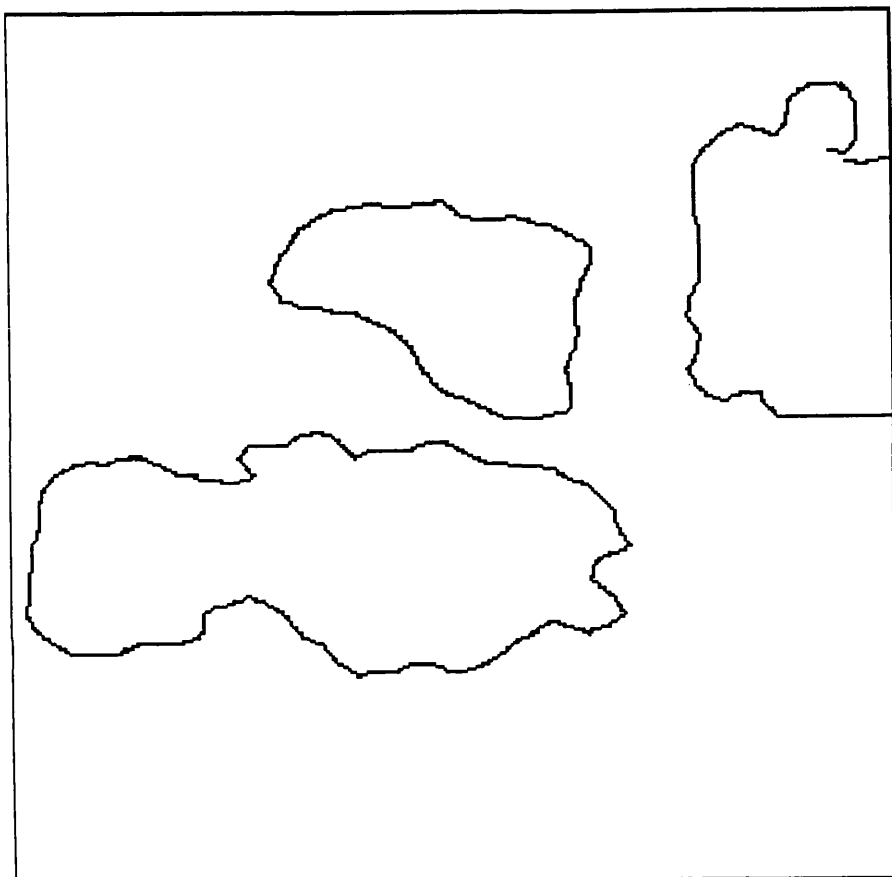


Figure 4.4 The user-defined close-fitting polygon employed in the histogram technique is shown together with the resulting brightness histogram. Note the rather indistinct nature of the land peak, together with brighter areas well to the right of the main peak.



**Figure 4.5** The result of the application of the LIC edge detection method to the resampled MSS band 4 image. Comparison with the original MSS image (Figure 4.2) reveals the relative success of the technique.

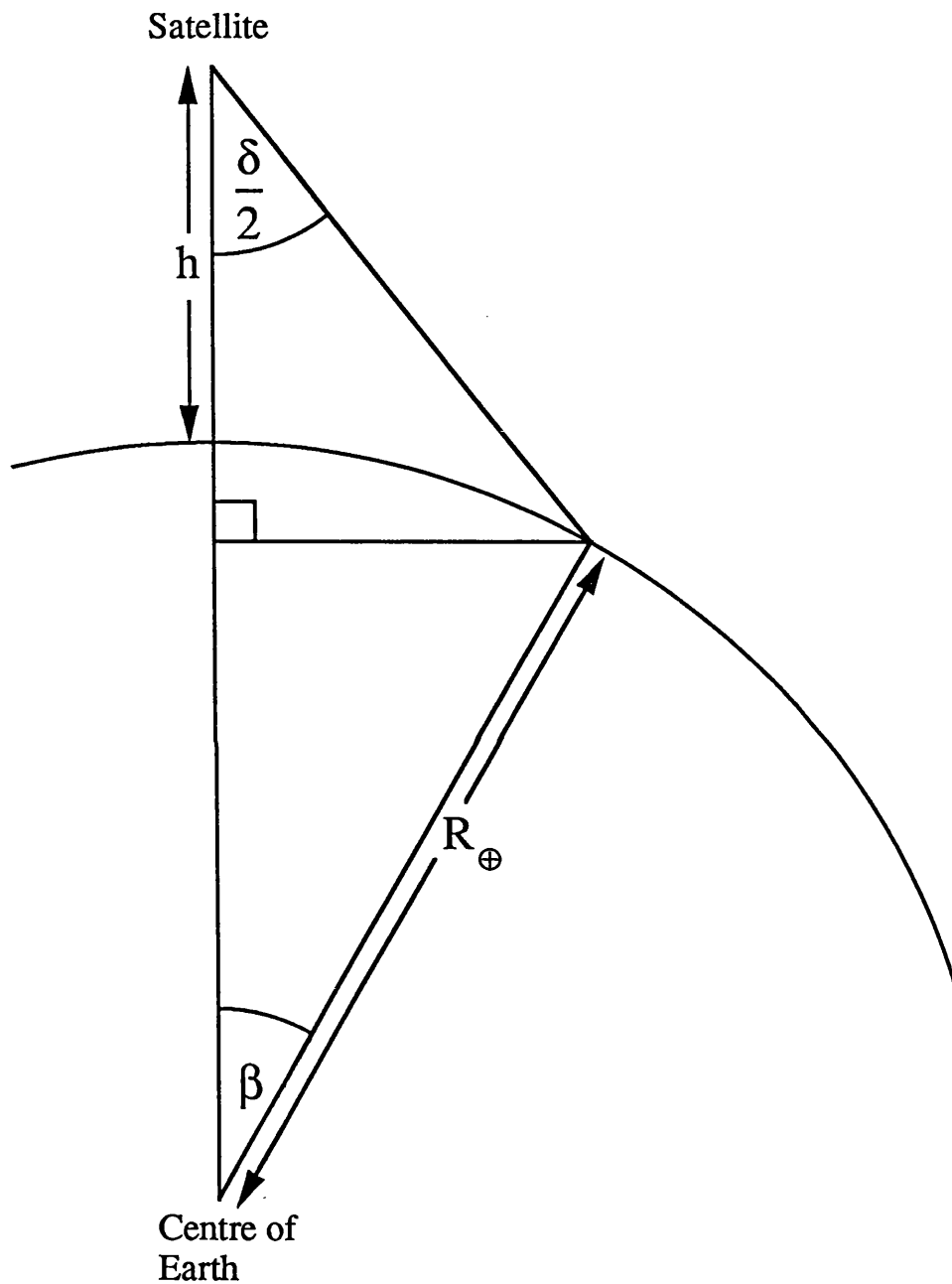
the lake. The histogram technique underestimated the area of the lake by an average of ~2%, the most likely reason again being that, because of the salt flats, the mean value of the land peak in the histogram was lower than the brightness value of the land within the edge pixels, resulting in the calculated mid-point of the edge pixel distribution being closer to the lake side of the histogram than the true mid point. The average area returned by the LIC method was ~1% low, presumably caused by salt flats that were less than 1 pixel in size around the border of the lake. Such sub-pixel features would result in the edge pixel that contained them, although not the brightest in the local sampling region due to the presence of water in the pixel, being brighter than would be the case if the land it contained was not salt flat, thus the fractional cover of water would be underestimated.

As a final check on the suitability of the two techniques, a cloud-free AVHRR LAC scene of the lake for the next day (3<sup>rd</sup> November 1988) was obtained from NOAA and the two techniques were applied to the channel 2 image. The areas obtained from the AVHRR data for the two methods were geometrically corrected using the procedures described in Appendix B. The pixel area of the lake on the original full resolution MSS band 4 scene was obtained by using the histogram technique. Since the fraction of the pixel area of the lake in the MSS scene contained within the edge pixels is only ~4%, and the error estimate from binomial statistics (likely to be conservative, see chapter 3) in determining this fraction is dependent on the root of the number of edge pixels ( $\approx 1/2 \times \sqrt{1100}$ ), the total error on the measurement using this technique is assessed to be less than 0.1%. Conversion of the pixel area to square kilometres requires knowledge of the scan angle moved through per pixel and the satellite altitude. The former parameter is not stated explicitly in the Landsat Data User's Handbook but can be calculated using the knowledge that the instrument is designed to have an active scan angle which produces a swath width of 185.0 km for a nominal altitude of 705.3 km. The MSS data used in the analysis were raw scanlines that were not corrected for variations in mirror velocity between successive scans, thus the angle per pixel was calculated by obtaining the average scan line length in the region of the lake and employing the following formulae:

From Figure 4.6 we have:

$$\tan\left(\frac{\delta}{2}\right) = \frac{R_{\oplus} \sin\beta}{R_{\oplus}(1 - \cos\beta) + h} \quad (4.1)$$

where  $\delta$  is the active scan angle,  $R_{\oplus}$  is the radius of the earth and  $h$  is the nominal



**Figure 4.6** The trigonometry for the calculation of the active scan angle,  $\delta$ , of the MSS instrument. The angle in radians at the centre of the earth,  $\beta$ , is obtained by dividing the half swath width (92.5 km) by the earth radius;  $\delta$  can then be obtained from application of the Sine rule.



altitude of the satellite. Having obtained the angle  $\delta$ , the average angle per pixel,  $\eta$ , is simply obtained:

$$\eta = \frac{\delta}{\text{average } n^{\circ} \text{ of pixels per scan}} \quad (4.2)$$

This value, together with the altitude of the spacecraft calculated from the satellite state vector in the header information and the the height of the lake above sea level, was then used to geometrically correct the area obtained from the MSS scene in a similar way to that described in Appendix B for the AVHRR data. The geometrically corrected area obtained from the MSS image was 133.1 km<sup>2</sup>, representing a large decrease from the baseline area surveyed in 1969 of 170 km<sup>2</sup> (see Figure 4.1). By comparison, the areas obtained from the histogram and LIC methods from AVHRR were 130.0 km<sup>2</sup> and 131.8 km<sup>2</sup> respectively, indicating both that the LIC method was the most accurate and that the simulated results had produced realistic values for the errors associated with the two area measurement techniques.

Based on the results from the above analysis, it was decided to employ the LIC edge detection method for measuring the area of Lake Abiyata in a time series of AVHRR images.

#### **4.4 Area cycle of Lake Abiyata obtained from a time series of AVHRR images**

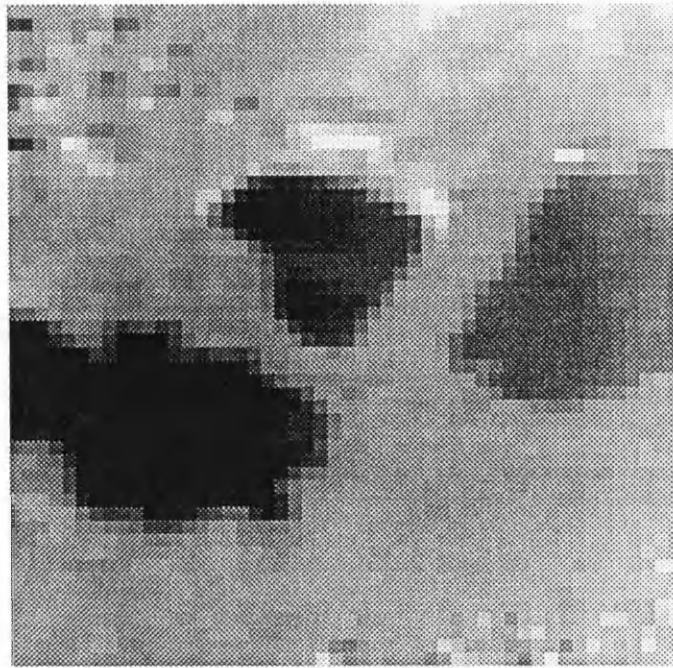
The temporal spacing desired for the time series analysis of lake area was ~1 cloud-free image per month over the period of the study. It was felt that this represented a good compromise between data cost and temporal resolution, matching that of the previous dataset, and significantly better results were unlikely to be obtained from more frequent observations since lake areas act as low pass filters to changes in climate (Mason *et al*, 1991). As mentioned previously, the NOAA archive of LAC data of the region commences in mid-1985, but regular coverage was only established from the end of 1987. In addition, the lake is situated in a valley in the Ethiopian highlands and the high relief of the nearby mountains promotes considerable cloud cover as moist air is forced to rise over them. As a result, the availability of cloud-free scenes until the end of 1987 is somewhat limited. As an example, for one six month period, only six scenes were archived for the region, only two were cloud free over the lake and these were fifteen days apart. Thus the provision of more than one cloud free scene per month on a regular basis was unlikely to have been possible.

The total number of cloud free scenes obtained for the lake from mid 1985 to early 1991 was fifty two. Since the scenes were all daytime passes, the channel 2 image was selected for the area measurement as this had produced the best results in the previous studies due to the strong contrast between water and land. The LIC technique was applied to each image in the following manner:

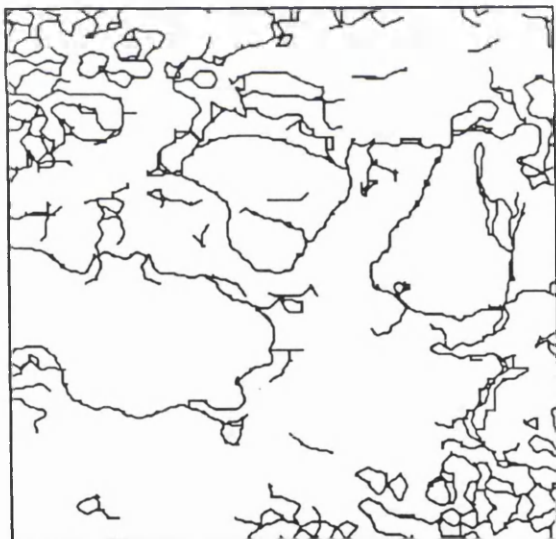
- A starting value for the contrast threshold that would determine the presence of an edge within the local sampling region was input and the area enclosed by the resulting contour obtained.
- The process was repeated with higher threshold values until the area enclosed by the contour did not change. The resulting contour was assumed to be the true edge of the lake based on the premise that this would have the edge of highest contrast, and its area was adopted as the area of the lake.

The feature of the LIC method allowing alteration by the user of the minimum contrast level at which the algorithm will determine the existence of an edge has an additional advantage. Some unwanted features such as thin cloud could be excluded from the analysis by increasing the threshold until their less distinct edges were no longer detected by the algorithm. An example of this process is shown in Figure 4.7; application of the histogram technique to such an image would be difficult and the result obtained for the area of dubious value.

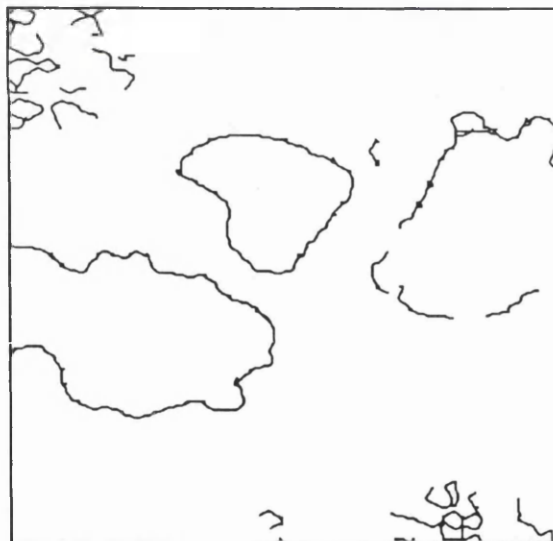
The pixel areas obtained for the series of images were corrected for geometric distortions as outlined in Appendix B. The satellite altitudes for the various passes that were required for geometric correction had to be calculated from Brower mean orbital elements supplied by NOAA since the NOAA LAC product has the TIP data which contains satellite altitude and attitude information stripped out. The results for the lake areas obtained from the AVHRR scenes are shown in graph form in Figure 4.8. The areas are shown without individual error bars and are joined together to illustrate the higher frequency trends, although this is not necessarily meaningful during certain times when the measurements are sparse. A sample point with error bars is also illustrated to give an idea of the reliability of the trends seen. Since the viewing angle of the lake varied considerably for different scenes up to a maximum of  $\sim 60^\circ$ , the average pixel area of the lake was  $\sim 100$  rather than the 150 obtained from the simulation analysis which represented AVHRR data for the nadir view. As a result, it was decided to adopt a typical error figure of 2% for the measurements which might be regarded as a conservative estimate, except when considering the bias



(a)

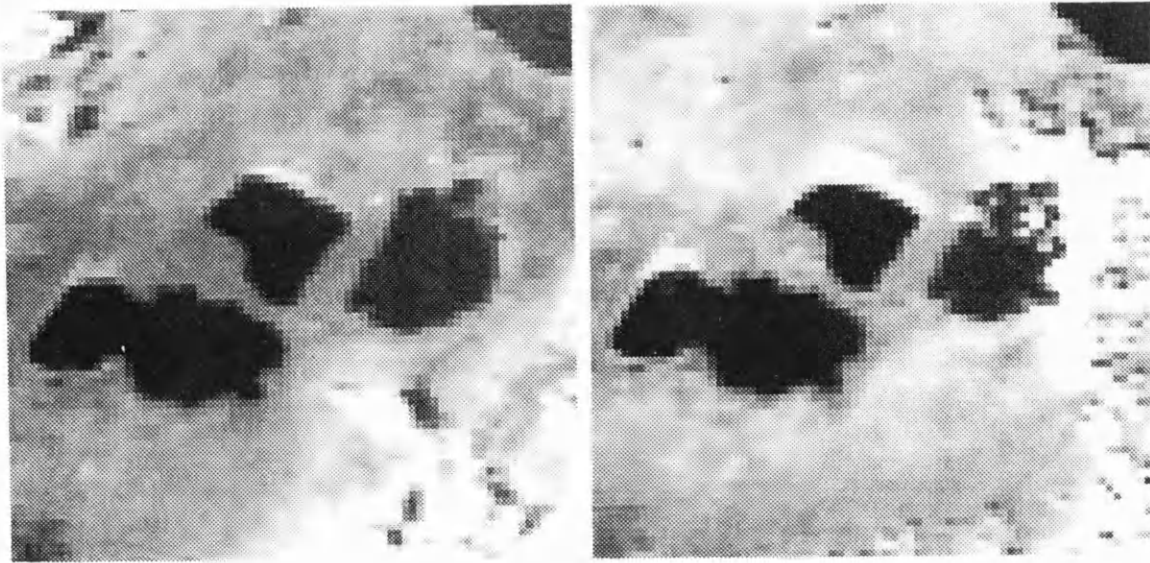


(b)

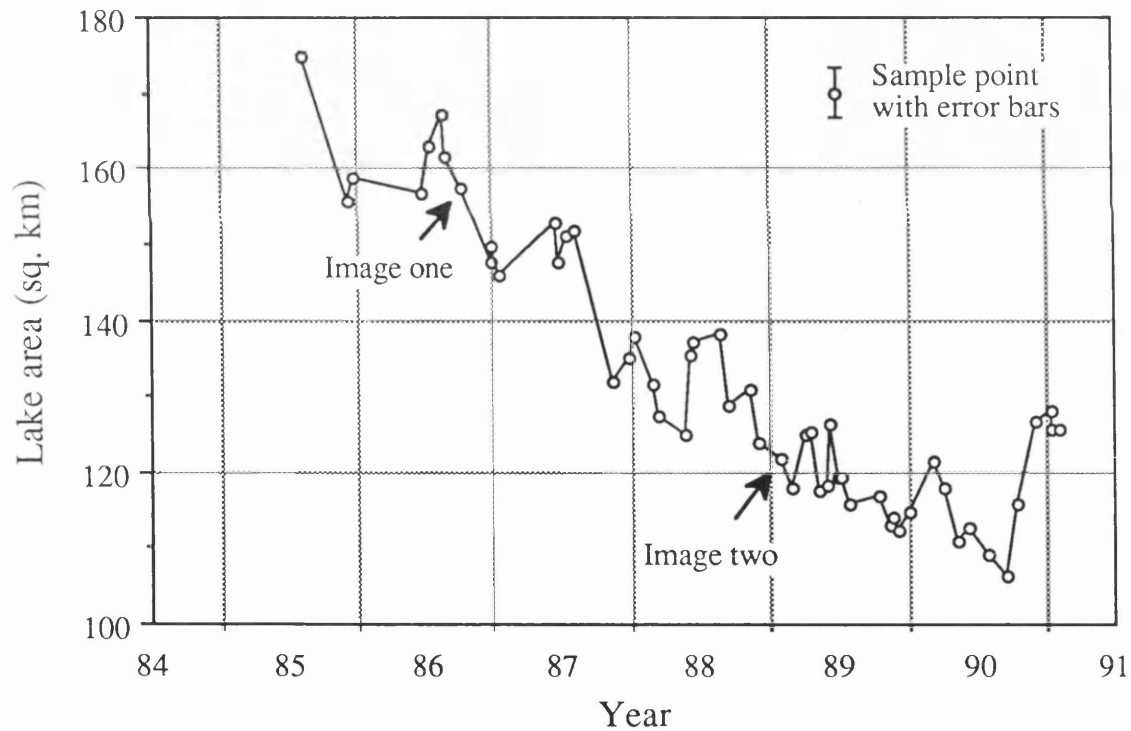


(c)

**Figure 4.7** An AVHRR channel-2 image of Lake Abiyata is shown in (a). Note the presence of some thin cloud in the image which is also detected by setting a low threshold for the LIC method (b). The area of the lake can still be retrieved by choosing an appropriate threshold (c).



### Lake Abiyata Area Cycle



**Figure 4.8** The area cycle obtained from the time series of AVHRR image is shown, together with 2 sample images to illustrate the appearance of the lake at different times. Note large area of salt flats which have been exposed in the second image.

in area estimation obtained during the simulation.

In addition, two sample channel 2 scenes are shown to illustrate the the appearance of the lake for different retrieved areas. In the first image, which represents a lake area of 157 km<sup>2</sup>, the salt flats at the shallow northern end of the lake are not strongly evident. By contrast, in the second image, taken less than two and a half years later, the lake has retreated considerably, revealing a large extent of salt flats at the northern end as the area has decreased to 122 km<sup>2</sup>.

#### 4.5 Discussion of Results

The first result of note is the considerable variation seen in lake area from mid-1985, when the lake measured more than 170 km<sup>2</sup> in area, to mid-1990, when it had decreased to less than 110 km<sup>2</sup>. The area of the lake in mid-1985 is equivalent to the baseline area measured in 1969, when the lake proceeded to increase by 40 km<sup>2</sup> in area in less than four years (see Figure 4.1(b)). The change in area seen in this present study is even greater in magnitude but is still of the same order and is thus a credible result in terms of expected variability. The overall trend in lake area for the period under study is that of steady decline for the most part which is sharply reversed at the end of 1990 and the beginning of 1991.

Closer inspection of the data shows that an annual cycle in lake area can be distinguished which is superimposed on the general downward trend. The supply of runoff water to the lake is not evenly distributed about the year as the rains in this region are the monsoonally driven ones that also supply the Blue Nile. The summer wet season runs from July to September, while the winter rains come between November and March, thus sharp increases in lake area can be detected during these periods for most years. Note the large increase in area for the middle of 1988, but that little increase in area is observed for the summer of 1989. Also, the onset of the summer monsoon for 1990 seems to have been delayed, and has combined with the winter rains to produce a very large increase in area of ~20 km<sup>2</sup> at the end of the year.

Climate records for the adjoining Sahel region (World Climate Data Programme (WCDP), 1990) show below normal rainfall for eighteen years preceding 1988, with an intensification of the drought between 1986 and 1987, although not as severe as the well-known extremes of 1983 and 1984. During 1988 the rainfall returned to near normal values in the Sahel, with severe flooding in and around Khartoum in the Sudan occurring in August of that year. This latter event has been partly attributed to



run off from heavy falls in the Ethiopian Highlands during the summer months (WCDP, 1990), and the occurrence of such precipitation events is evident in the sharp increase in lake area over this period.

The sudden change from a moderate warm phase of the Southern Oscillation to a pronounced cold phase in 1988 resulted in many precipitation anomalies around the globe, and the flooding in Khartoum may have resulted from this (WCDP, 1990). However, the absence of significant increase in lake area during 1989 indicates that the change in precipitation pattern during 1988 may only have been temporary. Recent work (WCDP, 1990, after Ropelewski and Halpert, 1989) indicates that there is a strong correlation between the phase of the Southern Oscillation and precipitation patterns in monsoonally driven regions, and, in particular, drier than normal conditions for the region are linked to high index cold phases, which would be in agreement with the area cycle results seen for Lake Abiyata.

#### 4.6 Conclusions

The consistency of results for the area of Lake Abiyata bears out the high precision obtained from the simulation study for the LIC technique. Thus the method has been shown to be capable of dealing with the major potential problems that the various features of closed lakes are capable of producing. Although the technique is relatively computer intensive, the amount of CPU time has been substantially reduced by the introduction of ever faster processors and now requires about fifteen seconds per contour measurement. Since Lake Abiyata is at the bottom end of the size of lakes that we wish to measure with sensors such as ATSR, this study shows that we can expect accurate lake area cycles to be determined for most, if not all, of the World's closed lakes  $\geq 100 \text{ km}^2$  in area.

However, the interpretation of lake area cycles with respect to climate requires some care. For example, it is possible that the continued decline in lake area during 1989 and the majority of 1990 could still be observed if the precipitation rate had increased since the equilibrium time constant for Lake Abiyata is  $\sim 19$  years (Mason *et al*, 1991) and if the period of the climate signal is substantially less than this, there will be a phase lag between any change in precipitation and the corresponding change in lake area.

In conclusion, the results from this study are encouraging for the prospect of the monitoring of closed lakes around the globe. The generation of reliable time series of area measurements with coarse resolution 1 km data has been demonstrated, and the

major trends identified have been seen qualitatively to have relevance to climate signals. The lack of global coverage with 1 km resolution data from AVHRR will be overcome by the provision of ATSR data from ERS-1 for which all the data are recorded at full resolution, albeit with a reduced swath of ~500 km resulting in a daytime pass for each lake of at least once week. Ultimately, the available archived LAC data will prove invaluable in extending the climate record of the various lakes backward as the data from ERS-1, ERS-2 and Polar Platforms are used to monitor the current and future trends.

## Chapter 5

# A New Technique for Atmospheric Correction and Total Water Vapour Retrieval

### 5.1 Introduction

Sea surface temperature (SST) is an important parameter for climate research, and, in principle, satellite remote sensing offers the synoptic global datasets which are required for climate modelling. The accuracy requirement, however, is high; an r.m.s. accuracy of a few tenths of a degree K, for example, would be necessary to accurately monitor the growth of SST anomalies (typically of only a few degrees Kelvin in magnitude). Indeed, the WCRP have stated that their requirement for SST monitoring is  $\pm 0.5^{\circ}\text{C}$  (WCRP, 1986). The main problems of retrieving surface temperatures from space are described in chapter 2, but it is useful to recap on some points here.

For current thermal infrared measurements from space, an important source of error is the estimate of the 'atmospheric correction' (the observed deficit in brightness temperature at the top of the atmosphere), which can be up to  $\sim 10$  K in the tropics (Barton, 1983). This deficit is due to absorption of the signal by the atmosphere which is partially compensated for by atmospheric emission. During the last decade, techniques have been devised to account for the atmospheric correction, and the resulting algorithms applied to data from the NOAA Advanced Very High Resolution Radiometer (AVHRR). These 'split-window' or 'dual-window' algorithms are based on correlations between the SST and the brightness temperatures measured in different spectral bands, and their derivation is more completely described in chapter 2. These relations are not exact, however, since they are based on approximations to the radiative transfer equation. The result is that the algorithms are probably responsible for the majority of the error in SST measurements. The intrinsic residual r.m.s. error of such algorithms can be up to about half a degree in regions where the atmospheric correction is high (eg. the tropics), as will be seen later, and yet this is where, for example, SST anomalies such as that associated with El Niño occur; such events have major global climatic effects. It is in order to improve the atmospheric correction that the Along Track Scanning Radiometer (ATSR) is currently flying on the European Space Agency's ERS-1 satellite. Its correction technique utilises dual-angle views of the ocean through different atmospheric paths, with the aim of providing SST's globally to  $\sim 0.3$  K accuracy. The reader is referred to Chapter 2 for a



more complete description of the novel features of the ATSR. Simulations (eg. see later) indicate that residual atmospheric correction errors as low as  $0.15 \text{ K}^1$  should be feasible in ideal conditions for a two channel implementation of this technique using a linear combination of the split-window channels viewing the sea surface at nadir and  $55^\circ$ . The attainment of similar accuracies with conventional instruments such as AVHRR would be a valuable goal, particularly when applied with respect to the archived datasets from such instruments.

A further point is that current algorithms are optimised for maritime atmospheres, and yet the retrieval of surface temperatures over inland water bodies, particularly lakes, is of considerable interest for monitoring climatic change. Lake surface temperatures (lake ST) reflect long term changes in air temperature and George (1989) has shown that time profiles of lake ST can show features (eg. periodicity) which cannot be detected in the equivalent air temperature records. Also, continental atmospheres are often not in equilibrium with the lake surface temperature and conventional algorithms can experience problems (Irbe, personal communication, 1990), for example over-correction can result when the lake is much warmer than the atmosphere, as may happen at night. The remote sensing of the surface temperatures of lakes using AVHRR could benefit from a more robust atmospheric correction algorithm, and the monitoring of large numbers of lakes globally should prove to be a valuable dataset for climate research. A further parameter of importance for climate modelling is the total atmospheric water vapour content (Woods, 1984). This parameter is also important for the tropospheric correction for radar altimeters (Rapley *et al*, 1987). Indeed, a microwave radiometer dedicated to this latter purpose is being flown on ERS-1 and has a predicted accuracy of  $0.4 \text{ g/cm}^2$  independent of value (Bernard, personal communication, 1989). Algorithms using AVHRR data alone can achieve similar accuracy for this parameter (Schluessel, 1989). Microwave instruments such as the Special Sensor Microwave Imager (SSM/I) which carry a 22GHz water vapour channel can produce accurate images of global atmospheric water vapour content (Schluessel and Emery, 1990) but only over the ocean. Conventional 'split-window' water vapour retrievals also experience large errors over land due to the difference of the temperatures of the land and atmosphere eg. hot land during the day. Knowledge of the daytime water atmospheric vapour content would be invaluable for correction of NDVI measurements in tropical regions (Sharmann, 1991, personal communication).

---

<sup>1</sup> Since the ATSR has a swath of only 500 km, the viewing angle for the nadir viewport doesn't extend beyond  $\sim 25^\circ$ , so the intrinsic accuracy quoted is valid for the whole swath.

In this chapter a novel extension to the standard 'split-window' atmospheric correction technique is proposed which promises significant improvements in SST and lake ST retrieval for instruments such as AVHRR, as well as providing an accurate measure of atmospheric transmittance and total atmospheric water vapour (Harris and Mason, 1990). An explanation of the proposed technique is presented, together with a preliminary semi-empirical evaluation based on results derived from a standard set of 56 radiosonde profiles and an atmospheric transmission model.

A radiosonde profile consist of a series of measurements of atmospheric temperature and water vapour concentration made at various pressure levels by a balloon ascent; the results are transmitted back by radio as the balloon is usually not recoverable. The atmospheric transmission model functions in the following way:

- Radiosonde profile data is read into the model and is interpolated to fill 128 equal pressure intervals in the atmosphere. Equal pressure intervals are used since they represent equal masses of absorbing atmosphere.
- A database of line strengths and positions for the various atmospheric constituents at 8 different pressures (1000, 800, 600, 400, 200, 100, 50 & 25 mb) produced by the Air Force Geophysical Laboratory is used and the values for intermediate levels are interpolated. The need for measurements at different pressures can be understood in the light of the discussion of line broadening mechanisms in Chapter 2.
- The model calculates the absorption for each of the 128 layers and then integrates the absorption from each layer to the top of the atmosphere (TOA) and also integrates the absorption from the layer to the bottom of the atmosphere (BOA). The emissivity of the layer is equal to its absorptivity, so the emitted radiance from each layer is calculated and the model assesses the radiance contribution at TOA and BOA, having accounted for the intervening attenuation. The individual contributions to upwelling and downwelling atmospheric emission are then summed for all 128 layers, and the total transmittance is also calculated.
- The atmospheric transmission and upwelling and downwelling fluxes are calculated for 0.04 wavenumber intervals across the 10.3 - 11.3 and 11.5 - 12.5  $\mu\text{m}$  bands and these are then convolved with the AVHRR filter functions for Channels 4 & 5 respectively.

- It is now possible to input the temperature and emissivity of the surface, thus defining both the radiance emitted from the surface and the reflected downwelling atmospheric radiance. These are summed, attenuated by the atmospheric transmission and added to the upwelling atmospheric radiance at TOA to give the radiance as seen by the satellite.
- Finally, the model has the facility to permit different angles of view to be input, and the change in path length is incorporated into the above calculations. The variation of emissivity of water with angle is catered for by a lookup table where the emissivity has been calculated using the Fresnel equations with the refractive index of water at the frequencies of interest.

The atmospheric transmission model used in these studies was provided by A. Zavody of Rutherford Appleton Laboratory and is the same one that is used for calculating the pre-launch atmospheric correction coefficients for the ATSR. For the purposes of the simulations described in this chapter, all atmospheres are considered to be cloud-free and containing no aerosols. The accuracy and feasibility of the technique are discussed, together with a first evaluation of the effects of radiance measurement errors. A variety of potential applications are noted and the need for a validation study is discussed.

## 5.2 Accurate Transmittance and Total Water Vapour

It is known that sea surface temperature fields may exhibit significant spatial variability (typically ~1K in 50 km, Minnett, personal communication, 1989) and the spatial gradient may be even greater than this in coastal regions and on the surface of large lakes. In contrast, the atmosphere is, in general, expected to exhibit less variability over a similar distance. Indeed, the validation study presented in chapter 6 demonstrates that this is a reasonable assumption for the purposes of SST measurement. Here a new technique is described which makes use of the concept of spatially varying water surface temperature beneath a constant atmosphere to obtain information about the atmospheric absorption and hence the total atmospheric water vapour. In the next section this new information is incorporated in a new atmospheric correction algorithm. The technique is also applicable to any surfaces for which the emissivity is constant.

The radiative transfer equation leads to the classical expression for the radiance,  $I_\lambda$ , measured from space through an atmosphere for a surface of emissivity  $\epsilon_\lambda$  (eg. combining equations 2.27 and 2.29 we have) :

$$I_{\lambda} = \varepsilon_{\lambda} B_{\lambda}(T_0) t_{\lambda}(0, p_0) \int_0^{p_0} B_{\lambda}(T(p)) dt_{\lambda}(0, p) + t_{\lambda}(0, p_0) (1 - \varepsilon_{\lambda}) \int_{p_0}^0 B_{\lambda}(T(p)) dt_{\lambda}(p_0, p) \quad (5.1)$$

where  $T_0$  is the surface temperature;  $B_{\lambda}(T)$  is the Planck function;  $t_{\lambda}(0, p)$  is the atmospheric transmittance between pressure levels 0 and  $p$ ;  $t_{\lambda}(p_0, p)$  is the atmospheric transmittance between pressure levels  $p_0$  and  $p$ ; and  $T(p)$  is the air temperature at pressure level  $p$ . This represents a combination of equations 2.27 and 2.29 in chapter 2. For the purposes of deriving the new method, it is more convenient to simplify the expression to :

$$I_{\lambda} = I_{o\lambda} t_{\lambda}(0, p_0) + A_{D\lambda} + A_{R\lambda} \quad (5.2)$$

where  $I_{o\lambda} = \varepsilon_{\lambda} B_{\lambda}(T_0)$  is the surface *emitted*<sup>2</sup> radiance; and the terms  $A_{D\lambda}$  and  $A_{R\lambda}$  refer to the radiance components due to atmospheric emission observed directly and after reflection from the surface respectively (see Figure 5.1).

We now consider the atmosphere to be unchanged, but introduce a deviation in the surface temperature of  $\Delta T_0$  resulting in a change of ground radiance  $\Delta I_{o\lambda}$  and space-measured radiance  $\Delta I_{\lambda}$ . The expression for the space-measured radiance, as shown in Figure 5.1, is therefore :

$$I_{\lambda} + \Delta I_{\lambda} = (I_{o\lambda} + \Delta I_{o\lambda}) t_{\lambda}(0, p_0) + A_{D\lambda} + A_{R\lambda} \quad (5.3)$$

Subtracting equation (2) we have :

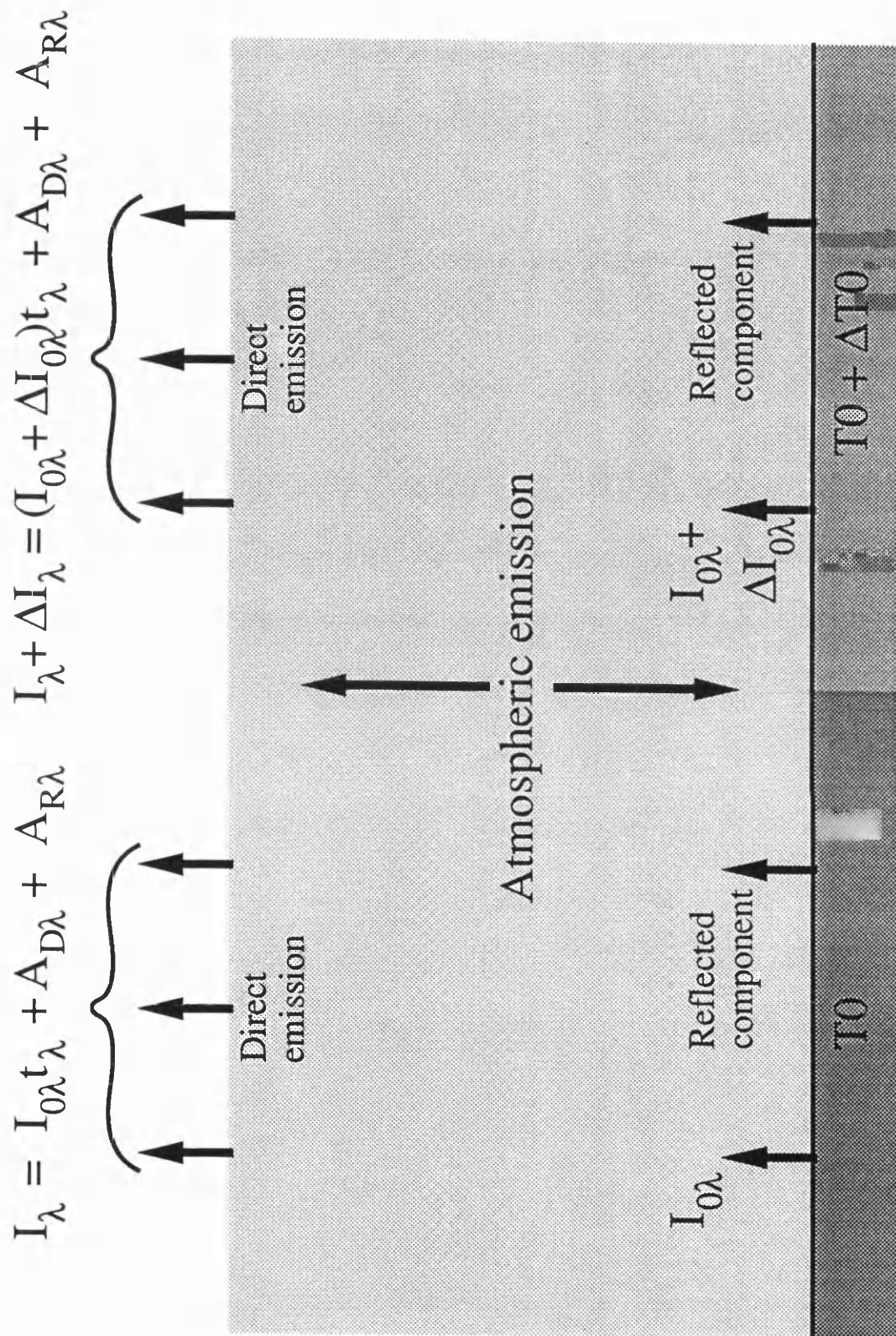
$$\Delta I_{\lambda} = \Delta I_{o\lambda} t_{\lambda}(0, p_0) \quad (5.4)$$

Thus the change in observed radiance is independent of the atmospheric emission terms. We now consider the ratio of the corresponding changes in radiance for two different wavelengths  $\lambda_1$  and  $\lambda_2$  :

$$\frac{\Delta I_2}{\Delta I_1} = \frac{\Delta I_{o2} t_2(0, p_0)}{\Delta I_{o1} t_1(0, p_0)} \quad (5.5)$$

---

<sup>2</sup> We are distinguishing between thermal radiance emitted from the surface and reflected atmospheric emission



**Figure 5.1** This figure illustrates the different radiance components that sum to make the radiance observed by the satellite. Note that the increase in radiance observed due to a change in surface temperature is independent of atmospheric emission.



For small values,

$$\Delta T_{\lambda} = \frac{\Delta I_{\lambda}}{\left(\frac{dB_{\lambda}}{dT}\right)_{T_0}} \quad \text{and} \quad \Delta T_0 = \frac{\Delta I_{0\lambda}}{\epsilon_{\lambda} \left(\frac{dB_{\lambda}}{dT}\right)_{T_0}} \quad (5.6)$$

This type of relationship has been used previously in Chapter 2 (equation 2.33); its validity for AVHRR channels 4 and 5 is best illustrated by considering Figure 5.2. From equations (5.5) and (5.6) we have :

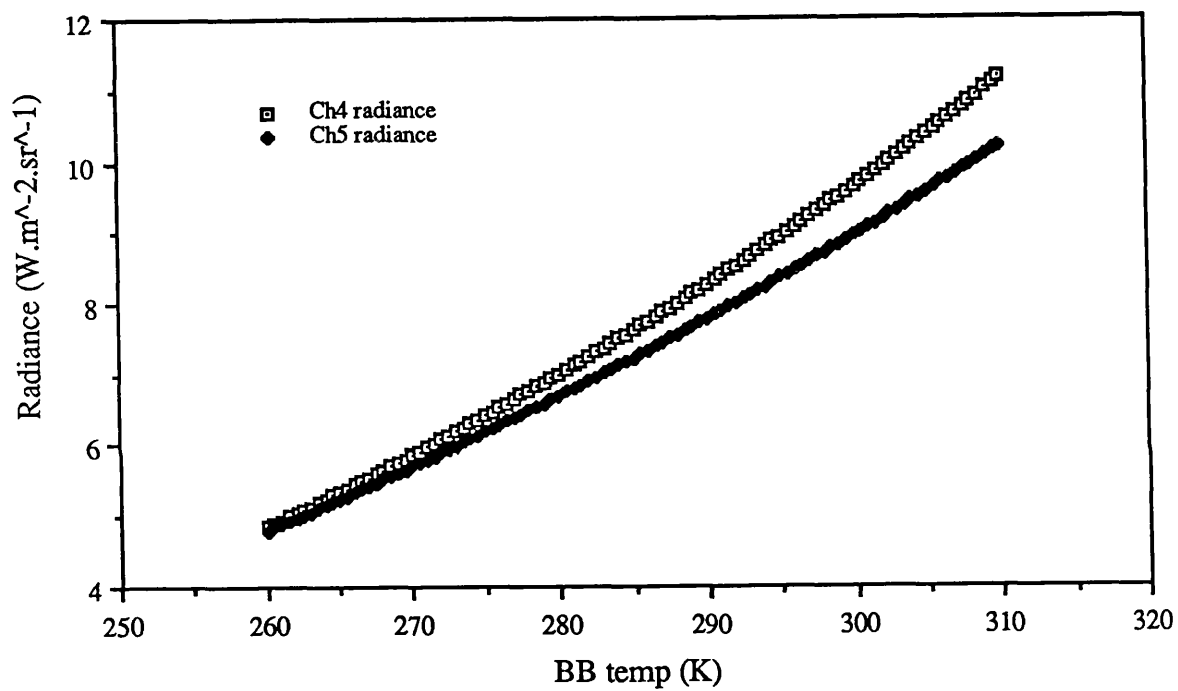
$$\frac{\Delta T_2}{\Delta T_1} = \frac{\epsilon_2}{\epsilon_1} \frac{t_2(0, p_0)}{t_1(0, p_0)} = R_{2,1} \quad (5.7)$$

Thus for a given change in surface temperature we expect a simple linear relationship between the ratio,  $R_{2,1}$ , of the changes in brightness temperatures in the two wavebands for a given change in surface temperature and the ratio of the transmittances in the two wavebands. Over the range of ocean temperatures likely to be observed ( $\sim 273\text{K}$  to  $\sim 310\text{K}$ ) and a change  $\Delta T_0$  of 1 K, the use of the approximation in equation (5.6) results in an error of less than  $\sim 0.5\%$  in the estimation of  $\Delta T$  and hence a negligible error in  $R_{2,1}$ .

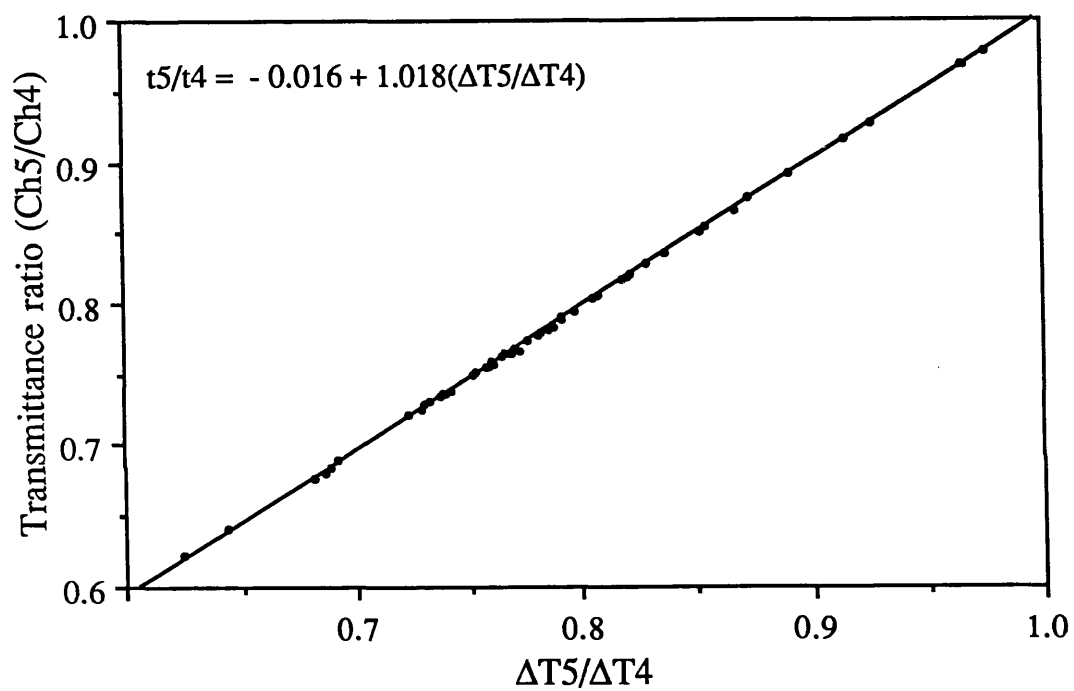
The validity of this relationship has been tested using the line-by-line radiative transfer model described earlier and a set of 56 ocean radiosonde profiles evaluated for the wavebands corresponding to AVHRR channels 4 and 5 viewed at nadir (see Figure 5.3). The radiosondes cover seasonally and geographically diverse conditions and were originally supplied by NOAA. Since in this case (sea water) the band-averaged emissivities for the  $\sim 11\mu\text{m}$  and  $\sim 12\mu\text{m}$  channels are 0.990 and 0.986 respectively (at 288K), it can be seen that the coefficient derived for the slope is within  $\sim 1\%$  of the theoretical prediction. The relationship of  $R_{2,1}$  to the total atmospheric transmittance for each band is shown in Figure 5.4. We can see from Figure 5.2 that the differential of the Planck function wrt temperature is slightly greater for Channel 4 than Channel 5 and therefore expect the gradient of the line shown in Figure 5.3 to be slightly higher.

Since the majority of absorption in the 10 - 12  $\mu\text{m}$  band is due to water vapour, we can now try to extract an expression for the total column content from the transmittance measurement. Let us make the commonly-used approximation that the absorption can effectively be considered as a continuum absorption within each band, and that the resulting band-averaged absorption coefficients are independent of gas temperature

### Black Body radiance vs temp for 10.8 $\mu\text{m}$ and 12.0 $\mu\text{m}$

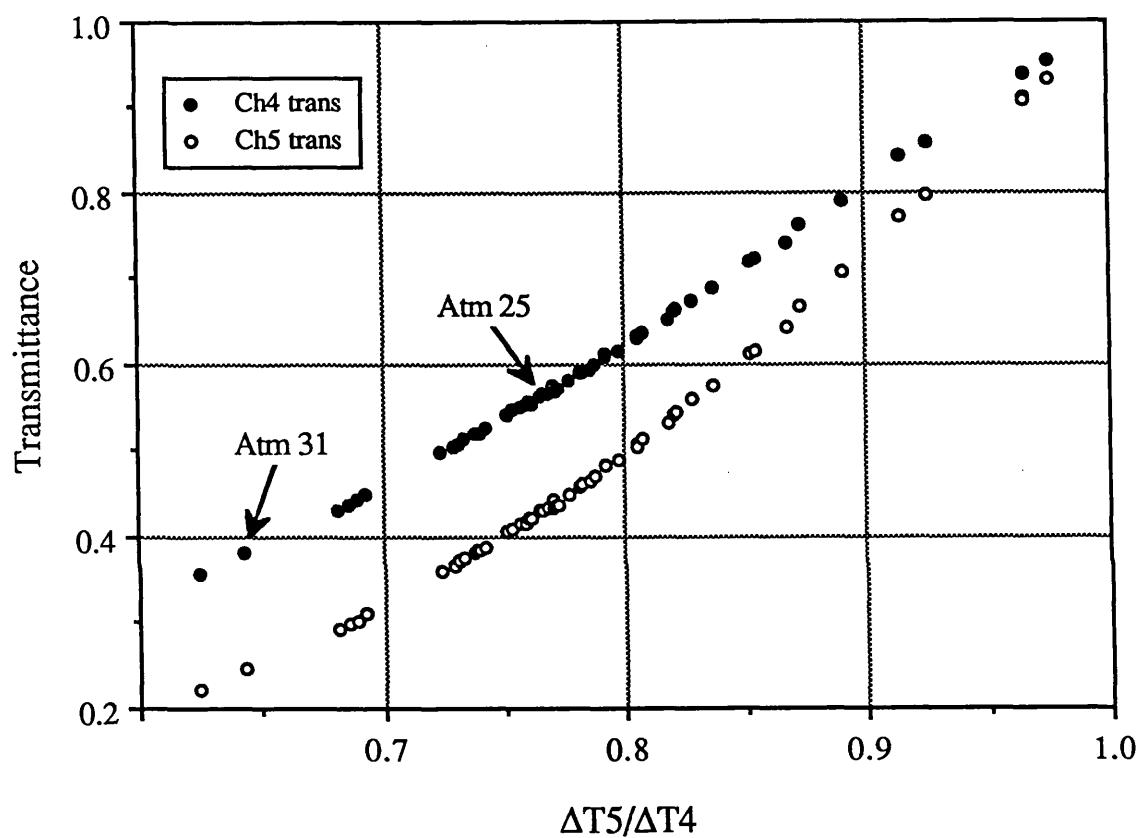


**Figure 5.2** The plot of radiance versus temperature for the Planck function at the wavelengths of AVHRR channels 4 & 5, with temperatures ranging between 260 and 310 K. The main point is to illustrate the validity of the use of the first order Taylor expansion for a small range of values about any likely target temperature.



**Figure 5.3** The linear relationship between the ratio of transmittances in AVHRR channels 5 and 4 and  $\Delta T5/\Delta T4$  is demonstrated.





**Figure 5.4** The relationship between the atmospheric transmittances in channels 4 and 5 and  $\Delta T5/\Delta T4$  is shown. Note how the plot of channel 5 transmittance is more strongly curved than that for channel 4. This point is more clearly illustrated in Figure 5.7.

(eg. Deschamps and Phulpin, 1980). We then have :

$$t_{\lambda}(0, p_0) \approx \exp(-k_{\lambda}U(0, p_0)) \quad (5.8)$$

where  $k_{\lambda}$  is the absorption coefficient for the waveband  $\lambda$  and  $U(0, p_0)$  is the total content of absorbing gas. Using the same approximations, we can then further divide the absorbing gases into water vapour and other gases :

$$t_{\lambda}(0, p_0) = \exp(-k_{O\lambda}U_O(0, p_0)) \exp(-k_{W\lambda}U_W(0, p_0)) \quad (5.9)$$

where  $k_{W\lambda}$  and  $k_{O\lambda}$  are the band-averaged absorption coefficients for water vapour and other gases respectively;  $U_W(0, p_0)$  and  $U_O(0, p_0)$  are the total column contents of water vapour and other gases respectively<sup>3</sup>. The exponents multiply their effect since the lines overlap, this is to be expected since the dominant absorption process is the water vapour continuum which overlaps all the lines in the window. Applying this to equation (5.7), we have :

$$\frac{\Delta T_2}{\Delta T_1} \approx \frac{\varepsilon_2}{\varepsilon_1} \exp((k_{O1} - k_{O2}) U_O(0, p_0)) \exp((k_{W1} - k_{W2}) U_W(0, p_0)) \quad (5.10)$$

which reduces to :

$$\log_e \left( \frac{\Delta T_2}{\Delta T_1} \right) \approx (k_{W1} - k_{W2}) U_W(0, p_0) + \text{const.} \quad (5.11)$$

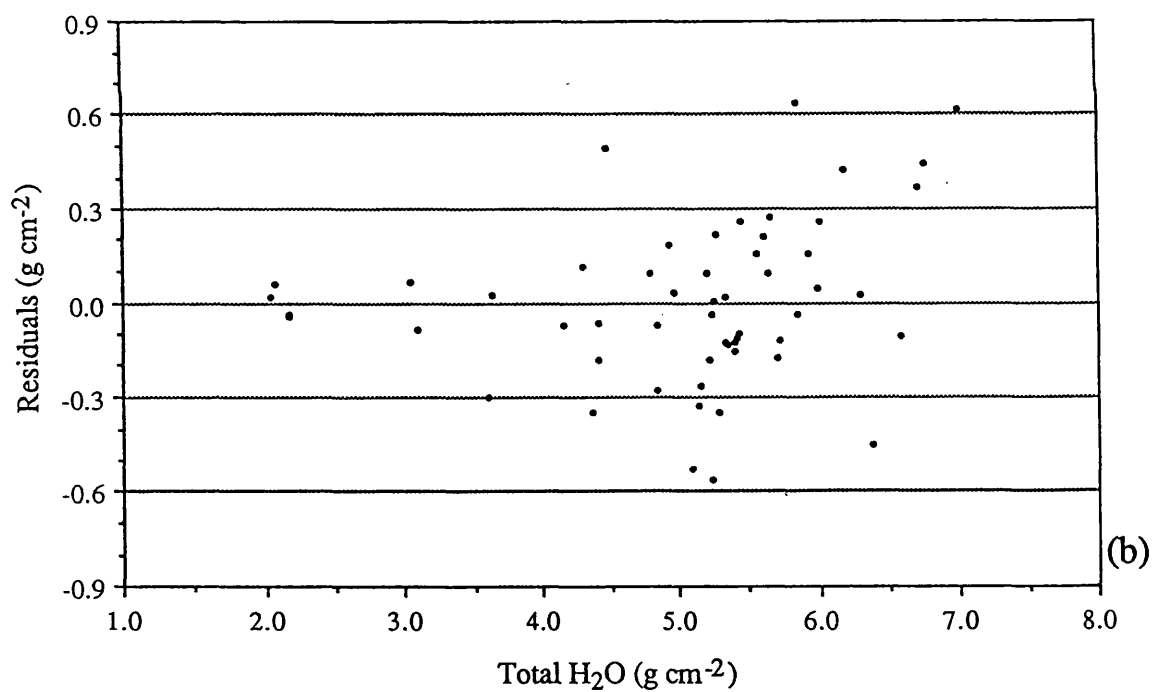
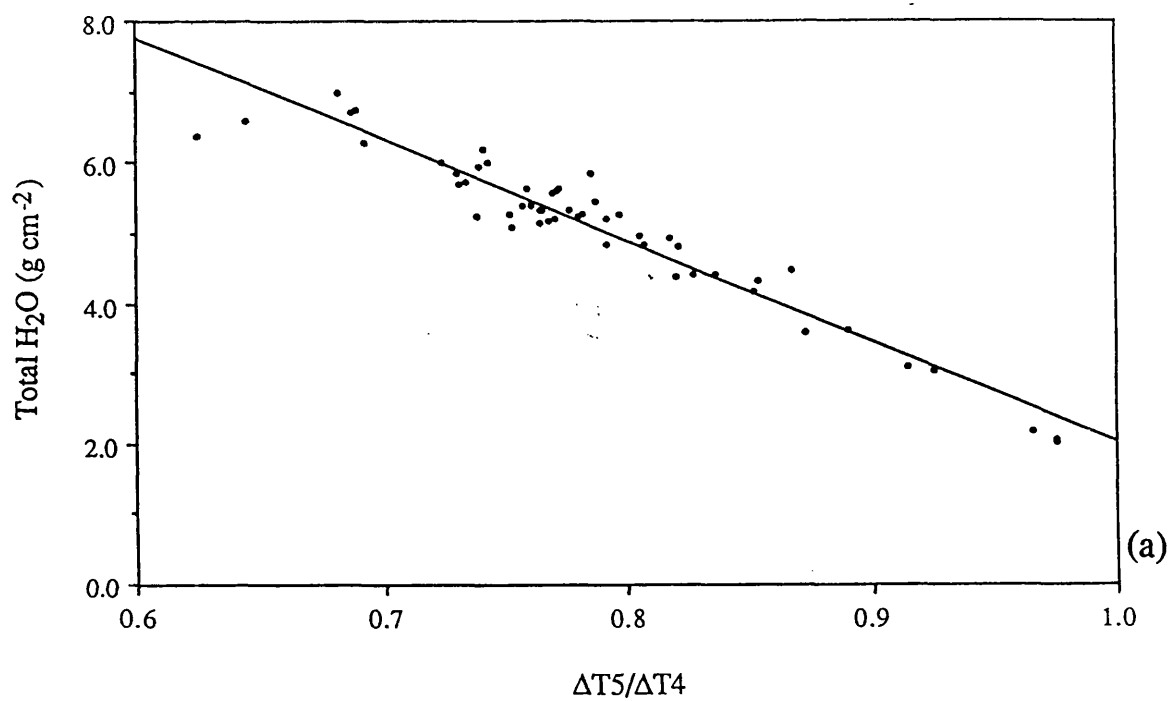
making the approximation that the effect of other atmospheric gases remains constant. In practice, for the 56 radiosondes used, it is found that the magnitude of  $(k_{W1} - k_{W2})U_W(0, p_0)$  is small and it is reasonable to take the first order expansion of the exponent in equation (5.10) :

$$\frac{\Delta T_2}{\Delta T_1} = K U_W(0, p_0) + \text{const.}' = R_{2,1} \quad (5.12)$$

where  $K$  is a constant. For the case of the 56 radiosondes and AVHRR channels 4 and 5, we plot the total column content of water vapour against  $R_{2,1}$  in Figure 5.5(a). It can be seen that the data are indeed fitted fairly well by a linear expression, the overall r.m.s. spread of 0.31 g/cm<sup>2</sup> being due to the above approximations not being

---

<sup>3</sup> The other assumption required here is that all gases have the same mixing ratio. Although this is somewhat unrealistic, we expect equation (5.9) still to be reasonably valid because water vapour is the principal absorber.



**Figure 5.5** The relationship between total column content of water vapour and  $\Delta T5/\Delta T4$  is shown in (a), together with a straight line fit to the data; coefficients are  $H_2O = 16.36 - 14.34(\Delta T5/\Delta T4)$ . The residuals to a second order polynomial fit are shown in (b), demonstrating that there is a greater spread when measuring larger values of water vapour.

completely valid, in particular the absorption coefficients presenting a dependence on temperature and pressure, especially for the continuum absorption by water vapour dimers (see Chapter 2). The residuals to the linear fit actually display a slight curvature; when a second-order polynomial is fitted, the residuals display no curvature (Figure 5.5(b)), and the r.m.s. deviation is reduced from 0.31 g/cm<sup>2</sup> to 0.26 g/cm<sup>2</sup>. The error is not independent of value; the fit is very good for low to medium water vapour content but somewhat worse at larger values. In percentage terms, the r.m.s. accuracy is ~5% independent of value. For comparison, the standard 'split-window' retrieval algorithm with AVHRR (a multiple linear correlation of  $(T_4 - T_5)$  and  $T_4$  with value) has an absolute accuracy of 0.45 g/cm<sup>2</sup> independent of value when applied to our dataset. This is comparable to the accuracy of 0.4 g/cm<sup>2</sup> predicted for the 2 channel microwave sounder on ERS-1 (Bernard, personal communication, 1989). An improvement to ~0.3 g/cm<sup>2</sup> is possible with the synergistic use of AVHRR with the nadir-viewing High resolution Infrared Radiation Sounder (Schluessel, 1989). The algorithm presented here is thus superior to current methods for low to middle values of water vapour, and comparable at high values.

### 5.3 Improved Atmospheric Correction

The concept of a varying surface temperature under a constant atmosphere can also be used to provide a new improved atmospheric correction algorithm. Let us first consider again the standard methods employed to determine where improvements might be made. Current techniques for the correction of space-measured brightness temperatures that have been developed are based on the correlation of the surface temperature with brightness temperatures retrieved from the different wavelength channels. A complete description of the theory behind the standard 'split-window' method is given in Chapter 2 so only a summary follows. A linear relationship between the atmospheric corrections for each channel can be derived from the radiative transfer equation by making a number of approximations (eg. Deschamps & Phulpin, 1980) leading to the following expression :

$$T_0 - T_4 = a_0 + a_1(T_4 - T_5) \quad (5.13)$$

where  $T_4$  and  $T_5$  are the brightness temperatures measured in AVHRR channels 4 and 5 respectively, and  $T_0$  is the ground temperature. Some authors (eg. Llewellyn-Jones *et al*, 1984) allow the coefficients for each channel to vary independently :

$$T_0 = a'_0 + a'_1 T_4 + a'_2 T_5 \quad (5.14)$$

however, the sum of the coefficients  $a'_1$  and  $a'_2$  is still  $\approx 1$ . It has been shown that coefficients derived for the nadir view are not appropriate for larger airmasses (Llewellyn-Jones *et al*, 1984) and that it is necessary to calculate new coefficients for different angles of view. Table 5.1 shows the coefficients derived for equation (5.14) for five airmasses (1.0 to 2.0 in steps of 0.25) representing view angles from  $0^\circ$  to  $60^\circ$ . The various factors that can affect the signal due to change in viewing angle are:

- Change in emissivity of the surface, resulting in a change in the weight of the  $T_4$  term and a change in the reflected downwelling atmospheric radiance.
- A change in the column content of the absorbing gases; this results in a decrease in the transmission thus the atmospheres become less thin and the multiplier of the  $(T_4 - T_5)$  term becomes larger. Also, there is a change in the constant term since the constituents such as  $\text{CO}_2$  which have a more or less equal effect in both channels increase their contribution to the total absorption.

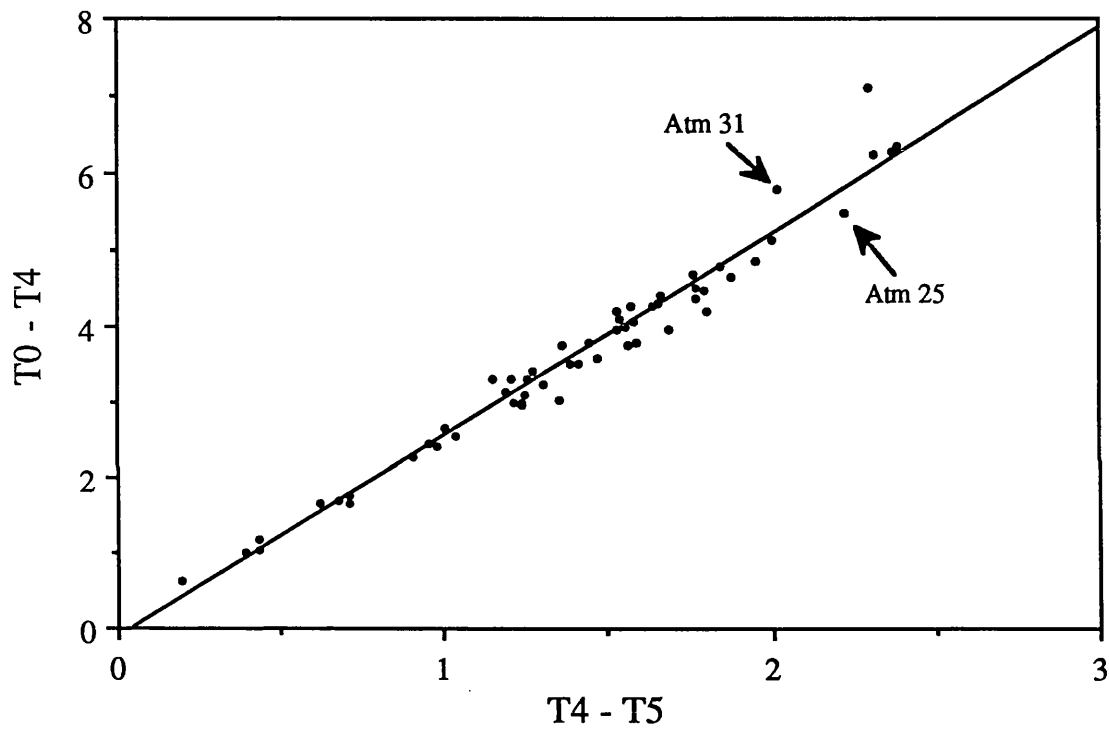
The above algorithms represented by equations 5.13 and 5.14 have been used extensively for SST retrieval (eg. Barton, 1983, McClain, 1989). Our result for equation (5.13) using the radiosonde data is shown in Figure 5.6, the algorithms in equations (5.13) and (5.14) have similar residual scatters ( $\sim 0.25$  K r.m.s.) for the nadir view. This is comparable to results obtained from other simulation studies (eg. Barton, 1983, obtains 0.22 K for 39 radiosondes), although realised SST accuracies so far are usually worse than this (eg. see Chapter 6) due to instrumental and other systematic errors such as the oceanic skin effect. The residual scatter, even in the simulations, is due to the approximations used to derive equations (5.13) and (5.14) not being completely valid for all atmospheres. The main approximations used are that the absorption coefficients are independent of temperature, and that the atmospheric transmittance is high, resulting in the simplification :

$$dt(0,p) \approx -k dU(0,p) \quad (5.15)$$

where  $dt(0,p)$  is the change in transmittance and  $dU(0,p)$  is the change in the content of absorbing gas between pressure levels  $p$  and  $p + \Delta p$ , and  $k$  is the absorption coefficient. This latter approximation is not very realistic for many atmospheres. For example, Figure 5.4 shows, for the 56 atmospheres described earlier, the total transmittances for each channel (obtained from the model) plotted as a function of  $R_{2,1}$ . It can be seen that the channel 4 transmittance is as low as 0.36, where the use

**Table 5.1 Split-window coefficients for 1 to 2 airmasses (NOAA-9 AVHRR)**

<b>Airmass</b>	<b>Constant</b>	<b>T<sub>4</sub> term</b>	<b>T<sub>5</sub> term</b>	<b>R.M.S. error (K)</b>
1.0	-1.61	3.653	-2.648	0.251
1.25	-3.11	3.861	-2.851	0.326
1.5	-5.68	4.055	-3.036	0.390
1.75	-8.05	4.180	-3.152	0.437
2.0	-12.38	4.335	-3.291	0.488



**Figure 5.6** The relationship between channel 4 atmospheric correction and the difference between brightness temperatures in channels 4 and 5 is shown for a wide range of different atmospheres. The straight line represents the 'split-window' relation described in equation 5.13.

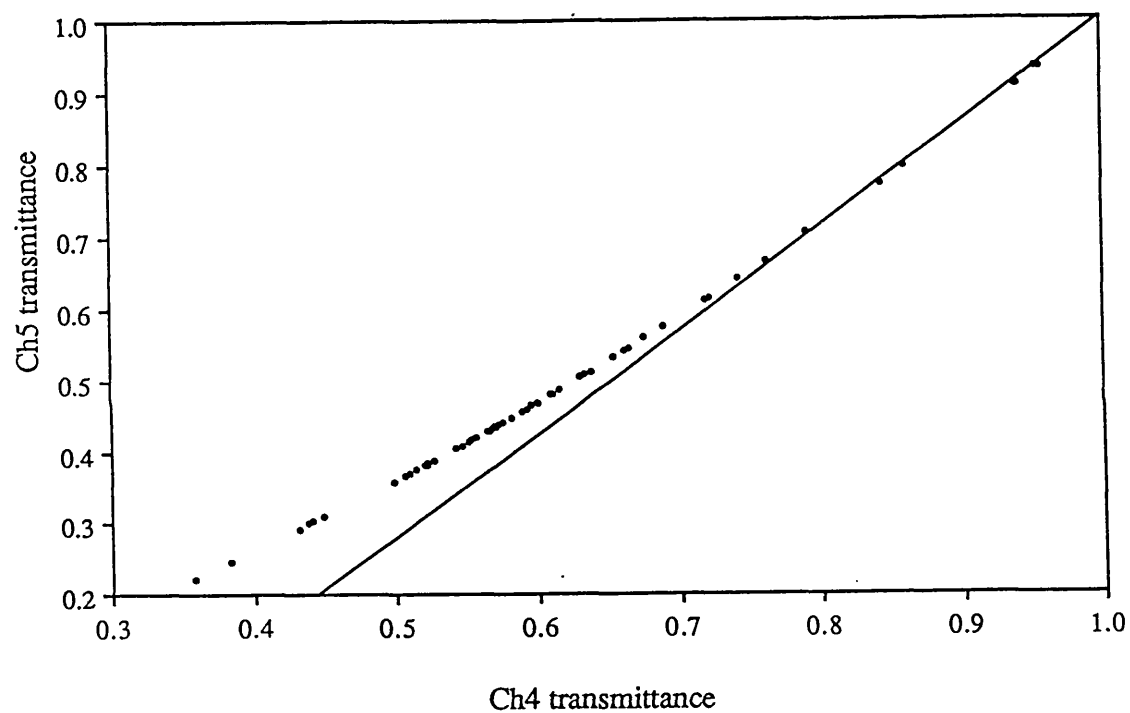
of the approximation of equation (5.15) would lead to a factor of up to 2.77 error in the estimation of  $dt(0,p)$ .

We would expect no effect if the linear relation between the atmospheric corrections in the two channels remained unchanged, but this is in fact not the case. By consideration of Figure 5.6, we can see that the spread of points increases as the atmospheric correction increases. This is primarily due to the (band-averaged) absorption coefficient for water vapour in the 11.5 - 12.5  $\mu\text{m}$  window being higher than that in the 10.3 - 11.3  $\mu\text{m}$  window, resulting in the above approximation (equation (5.15)) being less valid for channel 5 than channel 4 for any optically thick atmosphere. This is the reason for the relation between the transmittances in different channels becoming non-linear (see Figure 5.7). Thus, as we move from a thin to a thick atmosphere, the increase in the atmospheric correction ( $T_0 - T_5$ ) will no longer be proportional to the increase in ( $T_0 - T_4$ ) and hence the ( $T_4 - T_5$ ) value will lie off the predicted straight line (equation 5.13). It is expected, therefore, that the incorporation of extra information on the transmittance of each atmosphere, as derived in section 5.2, could lead to an improved algorithm for the atmospheric correction.

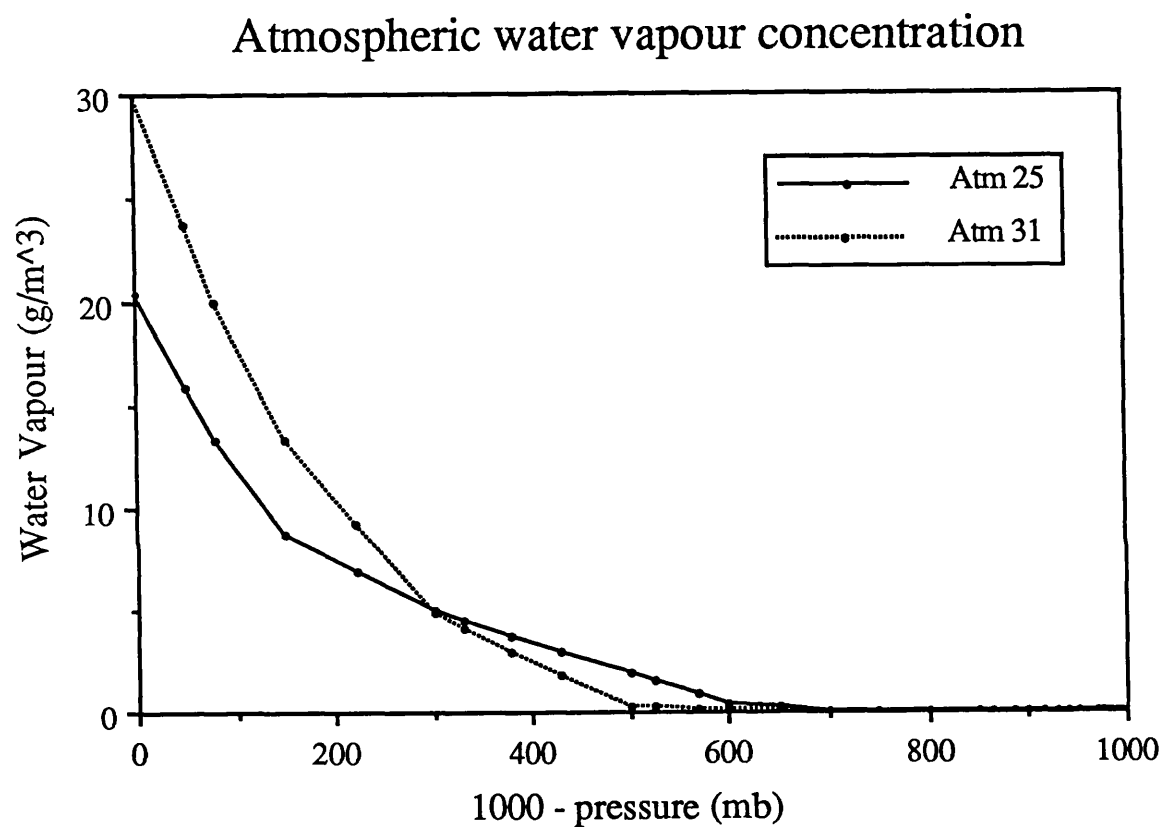
To illustrate these points, and to show how a new algorithm might best be achieved, let us consider the example of atmospheres 25 and 31 as marked in Figure 5.6. These have similar atmospheric corrections (5.5 K and 5.8 K respectively), but their different ( $T_4 - T_5$ ) values lead to differences in the prediction of the correction using the standard split window algorithm (equation 5.13). Figure 5.4 confirms that their total transmittances are also substantially different. Atmosphere 31 has a low channel 4 transmission, 0.38, where the approximation of equation 5.13 is poor and would lead to an error in the estimate of  $dt(0,p)$  of up to a factor  $\sim 2.8$ . Atmosphere 25, however, has a considerably higher channel 4 transmission (0.57, where the approximation would lead to an error in the estimate of  $dt(0,p)$  of a maximum of only  $\sim 1.8$ ) thus the approximation (equation 5.15) is more valid. The reason that atmosphere 25 has a rather high atmospheric correction with respect to its transmission is because the water vapour content is biased to higher (and therefore colder) levels, as shown in Figure 5.8, where the emission does not compensate so completely for the absorption.

Thus if we were to weight the atmospheres' ( $T_4 - T_5$ ) values by their transmittances, such that the ( $T_4 - T_5$ ) values are increased by a factor which is greater for the lower





**Figure 5.7** This plot demonstrates the relationship between the atmospheric transmittances in the two 'split-window' channels. The deviation from a straight line at low values demonstrates that atmospheres are optically thicker in channel 5 than channel 4.



**Figure 5.8** This plot demonstrates that the distribution of water vapour for atmosphere 25 is biased to higher levels than for atmosphere 31. This results in the atmospheric corrections being the same for both atmospheres despite the larger quantity of water vapour in atmosphere 31.

transmittance case, we might expect the points to move closer together on the plot. The parameter which we shall use for this weighting is  $R_{2,I}$ , which, as described earlier, is an almost linear function of the transmittance. It has a value of 1 at 100% transmittance and tends to zero as the absorbance approaches 100%, and so the simplest weighting of  $(T_4 - T_5)$  is  $[(T_4 - T_5)/R_{2,I}]$ . The atmospheric correction is plotted against this weighted parameter in Figure 5.9, and a new straight line fitted. It can be seen that not only do atmospheres 25 and 31 now lie very close on the line, but the overall spread of all the atmospheres is reduced. The r.m.s. spread on the atmospheric correction is reduced by a factor of 2.2 from 0.25 to 0.12 K.

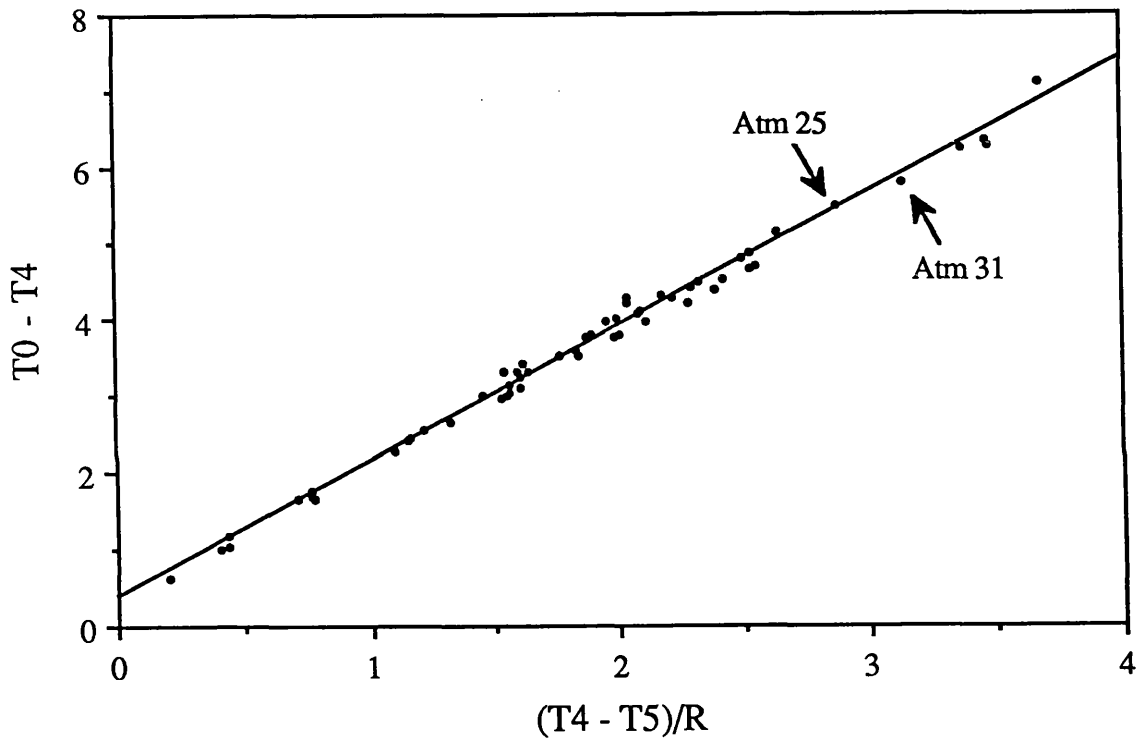
Furthermore, it is reasonable to allow the coefficients to vary independently, since the transmittance is a different function of the  $R_{2,I}$  for each channel, and so the final algorithm for SST retrieval has the form :

$$T_0 = a_0 + a_1(T_4/R_{2,I}) + a_2(T_5/R_{2,I}) + a_3T_4 \quad (5.16)$$

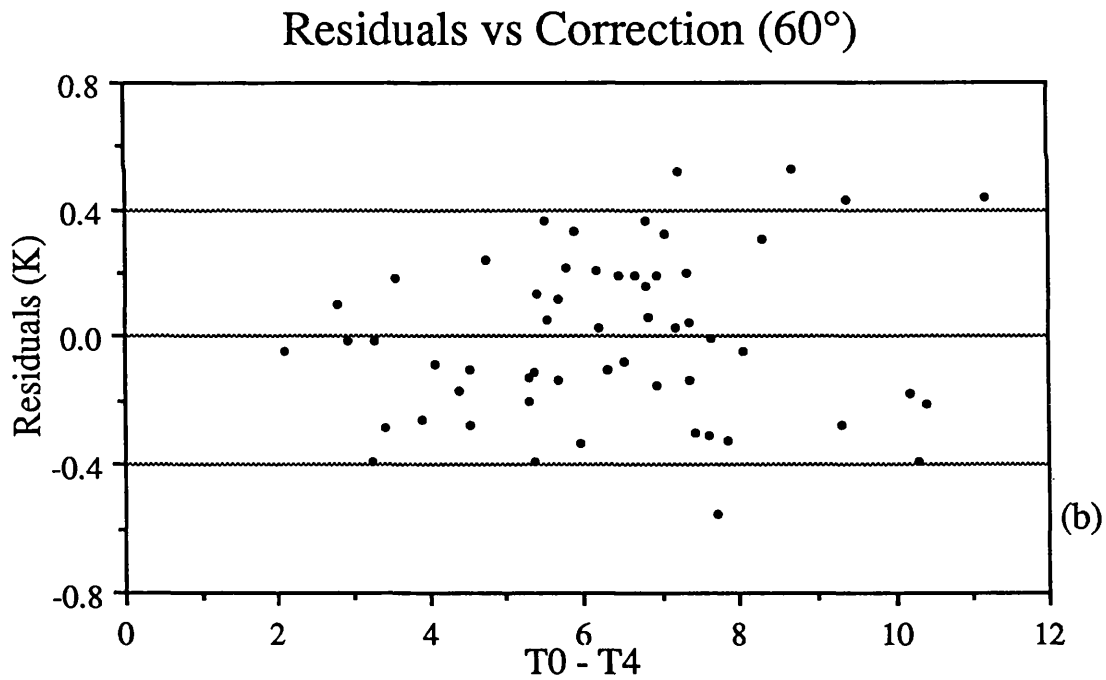
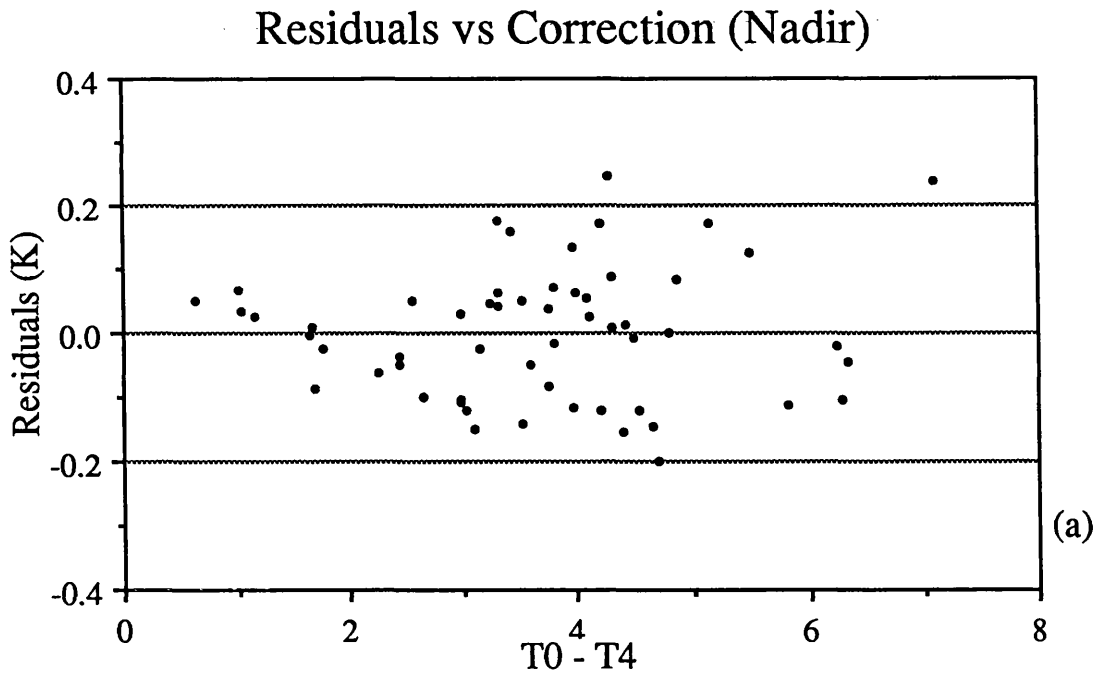
The r.m.s. deviation of the points for this algorithm is 0.106 K (see Figure 5.10(a)), representing a substantial improvement of a factor 2.4 over the conventional algorithms (equations (5.13) and (14)) and the actual reduction in r.m.s. is ~0.15 K. Figure 5.10(b) shows the residuals for the best fit regression for the new algorithm with a path length of 2.0 airmasses. The spread is now ~0.3K compared with ~0.5K for the conventional algorithm. Table 5.2 shows the new algorithm coefficients for 5 different airmasses together with their r.m.s. spread and the reduction factor in r.m.s. from the corresponding conventional split-window algorithm. Note that the coefficients multiplying  $T_4/R$  and  $T_5/R$  change very little in comparison with the standard algorithm coefficients shown in Table 5.1 because the weighting by  $R$  increases to compensate for the larger airmasses, although the constant and  $T_4$  multiplier still change due to the reasons discussed earlier. For comparison, the r.m.s. error at nadir attained for these radiosondes by an algorithm using a linear combination of the brightness temperatures for two different angles ( $0^\circ$  and  $55^\circ$  respectively) and two (split-window) channels (similar to that to be used with ATSR during the daytime) is found to be 0.147 K so this new algorithm promises improvements comparable with the addition of dual angle viewing. The practical limitations on the achievable error will be discussed in the next section.

#### 5.4 Limits to accuracy

These algorithms, based on the measurement of  $R_{2,I}$ , will have two limitations on



**Figure 5.9** Atmospheric correction for channel 4 is plotted against  $(T_4 - T_5)/(\Delta T_5/\Delta T_4)$ , together with a straight line representing a trial atmospheric correction algorithm. Note that atmospheres 25 and 31 have moved together on the plot.



**Figure 5.10** The residuals for the final algorithm at nadir (a) and 60° (b) are plotted with respect to channel 4 atmospheric correction. Note that the spread does not increase significantly at the higher values of atmospheric correction.

**Table 5.2 New algorithm coefficients for 1 to 2 airmasses (NOAA-9 AVHRR)**

Airmass	Constant	$T_4$ term	$T_4/R$ term	$T_5/R$ term	R.M.S.	Improvement Factor
1.0	-2.64	1.009	1.669	-1.668	0.106	2.4
1.25	-4.74	1.019	1.681	-1.681	0.139	2.3
1.5	-7.48	1.031	1.679	-1.680	0.177	2.2
1.75	-10.10	1.043	1.668	-1.671	0.224	2.0
2.0	-12.93	-1.056	1.634	-1.638	0.265	1.8

accuracy. The first is the validity of the assumption that the atmosphere is constant over the area of sea under consideration, and the second is the accuracy to which  $R_{2,I}$  can be determined. The two are not mutually exclusive since it will be easier to determine  $R_{2,I}$  over a large area of sea but the atmosphere is less likely to be constant. We present an initial examination of the probable errors in determining  $R_{2,I}$ . The assumption of constant atmospheric transmission is not expected to cause large errors, and results for SST retrieval using different sized areas for calculating  $R_{2,I}$  presented in Chapter 6 seem to bear this out.

In order to assess the error in determining  $R_{2,I}$  it is necessary to make some working assumptions. The first assumption is that the noise on the brightness temperatures retrieved from the channels is 0.1 K (c.f. NE $\Delta$ T for AVHRR channels 4 & 5 = 0.12 K). We are not concerned here with the calibration and other absolute errors since these will be the same for each channel. Although errors in the non-linear correction table (see Chapter 2) may be different for each channel and the absolute values of  $T_4$  and  $T_5$  may be wrong, over a small range of temperature the values of  $\Delta T_4$  and  $\Delta T_5$  will be unaffected. The second assumption is that the spread of brightness temperatures observed from space is 1 K in channel 4 over the area observed and that there is an equal number of pixels in each brightness interval. The digitisation noise has been ignored since it is less than the NE $\Delta$ T and can be effectively removed by averaging. If we consider the case of just two pixels that are 0.5 K apart<sup>4</sup> in both channels 4 and 5 :

$$\Delta T_5 / \Delta T_4 = (T_5 - T_5') / (T_4 - T_4') = R_{2,I} \quad (5.17)$$

The error in this expression will be :

$$\begin{aligned} \Delta / R_{2,I} &= \sqrt{[\Delta / T_5 - T_5' / (T_4 - T_4')]^2 + [\Delta / T_4 - T_4' / (T_5 - T_5')]^2} \\ &= 0.40 \end{aligned} \quad (5.18)$$

If we regard the errors on subsequent pairs of pixels (see footnote 4) as independent, the error will reduce as the root of the number of pairs used. For an area of 50x50 pixels (a typical averaging area for global SST maps, Llewellyn-Jones *et al*, 1984) the error is reduced to ~0.01. The impact of such a small error (corresponding to ~1% error in transmittance) on the determination of total column content of water vapour

<sup>4</sup> The increment of 0.5 K is chosen in order to maintain the same separation for subsequent pairs over the full range of 1 Kelvin.

by equation (5.12) will be minimal, but there are still some implications for the determination of SST by equation (5.16). If we consider the errors on  $T_4$  and  $T_5$  to be negligible through averaging, the error in SST,  $\Delta T_0$ , may be approximated to :

$$\{\Delta T_0\}^2 = \{[a_1(T_4 - T_5)/(R_{2,1})^2] \Delta R_{2,1}\}^2 + 0.106^2 \quad (5.19)$$

since the coefficients  $a_1$  and  $a_2$  are almost identical in magnitude but opposite in sign. This expression gives a value for  $\Delta T_0$  of 0.15 K for a worst case of  $(R_{2,1})^2 = 0.36$  and  $(T_4 - T_5) = 2.3$  K. Thus the accuracy of the algorithm is degraded by only ~40% in the worst case.

## 5.5 Applications and Conclusions

We now consider the possible applications to which this technique should be suitable. In the first instance, the algorithm provides a very accurate means of calculating the atmospheric transmittance. This may be used not only for extracting an accurate (~5% error) value for the total column water vapour content but may also be used as an input for radiation budget studies. The water vapour content may be used for meteorological studies and to provide the 'wet' correction for radar altimeter derived heights. The latter will be invaluable, for example, in extracting ocean topography, and for accurate lake levels which can be used as proxy indicators of climatic change (Mason *et al*, 1985 and Chapters 3 and 4).

There is considerable potential for providing an improved correction for SST, particularly in the tropical regions where current algorithms are experiencing problems due to the low atmospheric transmittances. The above simulation indicates precisions that are comparable to, or better than, the along track scanning technique of ATSR. The predicted r.m.s. error for the algorithm is expected to be somewhat degraded in practice, in the same way as the standard algorithms, since the errors due to sensor calibration and skin effect have yet to be included, although in the ERS-1 and Polar Platform era the former, at least, is expected to be characterised to an accuracy of  $\leq 0.1$  K. Retrieval of the surface temperatures of large lakes also benefit from the use of this algorithm, (see Chapter 6) since although atmospheres over land may be substantially different from ocean ones, the algorithm is expected to be more independent of atmospheric characteristics such as the distribution of absorbing gases.

It should also be possible to derive an equivalent algorithm for use over land if the emissivities of the cover types are known (eg. uniform regions such as deserts).



There is also the potential for making use of temperature variations of the same piece of land over a short time during which the atmosphere can be expected to remain reasonably constant (eg. future 'split-window' geostationary sensors). In addition, the differing emissivities of different cover types have the potential to provide the required spread in brightness temperatures, although the following points will then arise:

- 1) A change in emissivity will alter the reflected component of atmospheric emission as well, thus the derived value of  $R_{2,I}$  will be different.
- 2) There is a requirement for emissivity to be the same in both channels (or in a fixed ratio) for all cover types.
- 3) The problem of distinguishing between changes in brightness temperature due to actual temperature differences on the ground and emissivity changes may result in an ambiguity in the expected relation between  $R_{2,I}$  and total atmospheric transmittance.

It is possible, however, that these effects may display stable mean characteristics over individual regions, thus there would be the potential to empirically determine  $R_{2,I}$ /transmittance relationships for them. As regards the extraction of atmospheric water vapour content for atmospheric correction of the NDVI (and the 'wet' tropospheric correction for altimeter-derived heights), it is useful to note that vegetated surfaces have high emissivities (eg. up to ~0.98, Becker, 1987) and that the requirements on the accuracy for determination of  $R_{2,I}$  are much less stringent. Also, we expect the majority of observed brightness temperature variability to be genuine temperature variations due to the albedo and heat capacity differences of different cover types (Prata, 1991).

To summarise, the novel technique presented here has several potential uses in remote sensing over water for SST, LST and atmospheric water vapour, leading to improved global datasets for climate research. The next chapter demonstrates the application of the new temperature correction algorithms to satellite data and validation of the measurements.

## Chapter 6

# A First Validation Study of the New Technique for Atmospheric Correction

### 6.1 Introduction

The new technique proposed in Chapter 5 shows great promise for reducing the error associated with atmospheric correction of thermal infrared brightness temperatures measured from space. The simulation study indicates that there is a potential reduction of a factor 2.4 in the r.m.s. spread for the nadir view and similar levels of improvement for other angles, although these results are based on accurate knowledge of both the space-measured brightness temperatures and the ratio  $\Delta T_5/\Delta T_4$ . Although a preliminary assessment of the errors in SST retrieval that might be associated with an error in the measurement of  $\Delta T_5/\Delta T_4$  was performed, there is clearly a need to investigate the accuracy of atmospheric correction attainable with real AVHRR data. The need for validation arises from several important questions:

- What tolerance can be permitted, for the purposes of atmospheric correction, in the assumption of a constant atmosphere for the region over which the ratio is being assessed?
- How accurately can the ratio be measured with the AVHRR instrument within the region? The ability to average out the effects of detector noise and digitisation is most important, and it is unlikely that the two are well-matched in practice (Dudhia, 1989). This point is closely linked with the above, since reduction in error can be achieved by calculating the result for a larger area which provides a greater range of temperatures and more points to average over but is less likely to satisfy the first requirement.
- Moreover, what effect will errors associated with the above two points have on the precision of the new atmospheric correction algorithm?
- The most important question, however, will be whether the new algorithm offers any real improvement in atmospheric correction in comparison with the conventional 'split-window' algorithm.

It was decided that a comparison with the results for a conventional 'split-window' algorithm would be the most useful indication as to the impact of the potential sources of error outlined above. In addition, two variations of the new algorithm were tried, the variation being the area over which  $\Delta T_5/\Delta T_4$  was being assessed.

## 6.2 The Validation Dataset

The validation dataset consisted of a series of buoy temperature measurements in the Great Lakes together with the AVHRR/2 data from corresponding NOAA satellite overpasses of the region, and was supplied by G. Irbe of CCAH (Canadian Climate Center). There are a few points to be noted about both the *in situ* and satellite data. The Great Lakes buoy data has the following characteristics:

- The buoys are anchored and their positions known to within 1 minute of arc in both latitude and longitude for the Canadian side of the lakes. The latitude and longitude positions of the buoys on the American side are known to 10 seconds of arc.
- The temperatures are measured by a thermistor mounted on the hull of the buoy at a depth of about 1 metre. The accuracy of the measurement is  $\pm 0.1^\circ\text{C}$  and the measurement itself is quoted to one decimal place.
- The temperature measurements are made every hour, so the temporal coincidence/difference between satellite and *in situ* measurements is  $\pm 30$  minutes at worst.

The satellite used to provide the AVHRR data was NOAA-11. There are a few points to note about this particular satellite and the data it provides:

- This is the first satellite for which NOAA have published Non-Linear Correction (NLC) values (see Chapter 2) for three different instrument temperatures. The calibration accuracy that can be attained using this type of NLC is estimated to be  $\pm 0.2^\circ\text{C}$  (Brown *et al*, 1985). The improvement results because it helps to overcome the problem resulting from the changing radiance contribution reflected off the reference target as the instrument temperature varies around the orbit.
- Noise on channel 3 is substantially reduced in comparison with previous NOAA satellites. This information will aid the declouding procedure, but it will not be used in a 'triple-window' algorithm for the purposes of this study.

- The satellite overpasses are at about 07:45 GMT (~01:30 local time) for the early morning scenes, and 18:45 GMT (13:30 local time) for the one afternoon scene. Although the lakes are not particularly turbid, the single afternoon pass may be affected by the diurnal thermocline.

The atmospheres which are to be expected in the region will be continental in nature, and will not, in general, be in thermodynamic equilibrium with the lakes. The lower layers of the atmosphere will have been influenced by the temperature of the land over which they have been passing; this land may display considerable variation in temperature throughout the day and from day to day due to its low thermal diffusivity. As a result, we might expect considerable heat exchange between the lower atmosphere and lake surface, thus the atmospheres over the Great Lakes must be at least as variable in nature in a spatial sense as those over the ocean. We may therefore expect that this dataset will prove to be a tough test of the robustness of the algorithm.

### **6.3 Analysis procedure for the AVHRR data**

The data were supplied as full swath scenes of 2048×1536 pixels in a proprietary CCAH format on 6250 bpi tape. Initially, a full size raw counts channel 4 image was generated for each overpass and a number of 512×512 subscenes were selected which covered the lakes which had buoy locations that were clear of cloud on first inspection. For each 512×512 subscene, the raw counts were converted to brightness temperatures in the manner outlined in Chapter 2 (or NOAA Technical Memorandum 107, 1979) and using the values supplied in Appendix B (1988 revision).

#### **6.3.1 Declouding procedure**

Each subscene was then declouded using the APOLLO algorithms described by Saunders and Kriebel (1988). These consist of a series of hierarchical tests which are different for day and night; a diagram outlining the operation of the scheme is shown in Figure 6.1. There are certain points to note about the various tests:

- i) The gross cloud check which consisted of selecting a temperature in channel 5 below which all pixels would be excluded often served as an effective land mask for the nighttime scenes since the land had cooled down below the water temperature as the heat is stored only in the surface layer.
- ii) The spatial coherence test was modified so that every pixel was placed in the

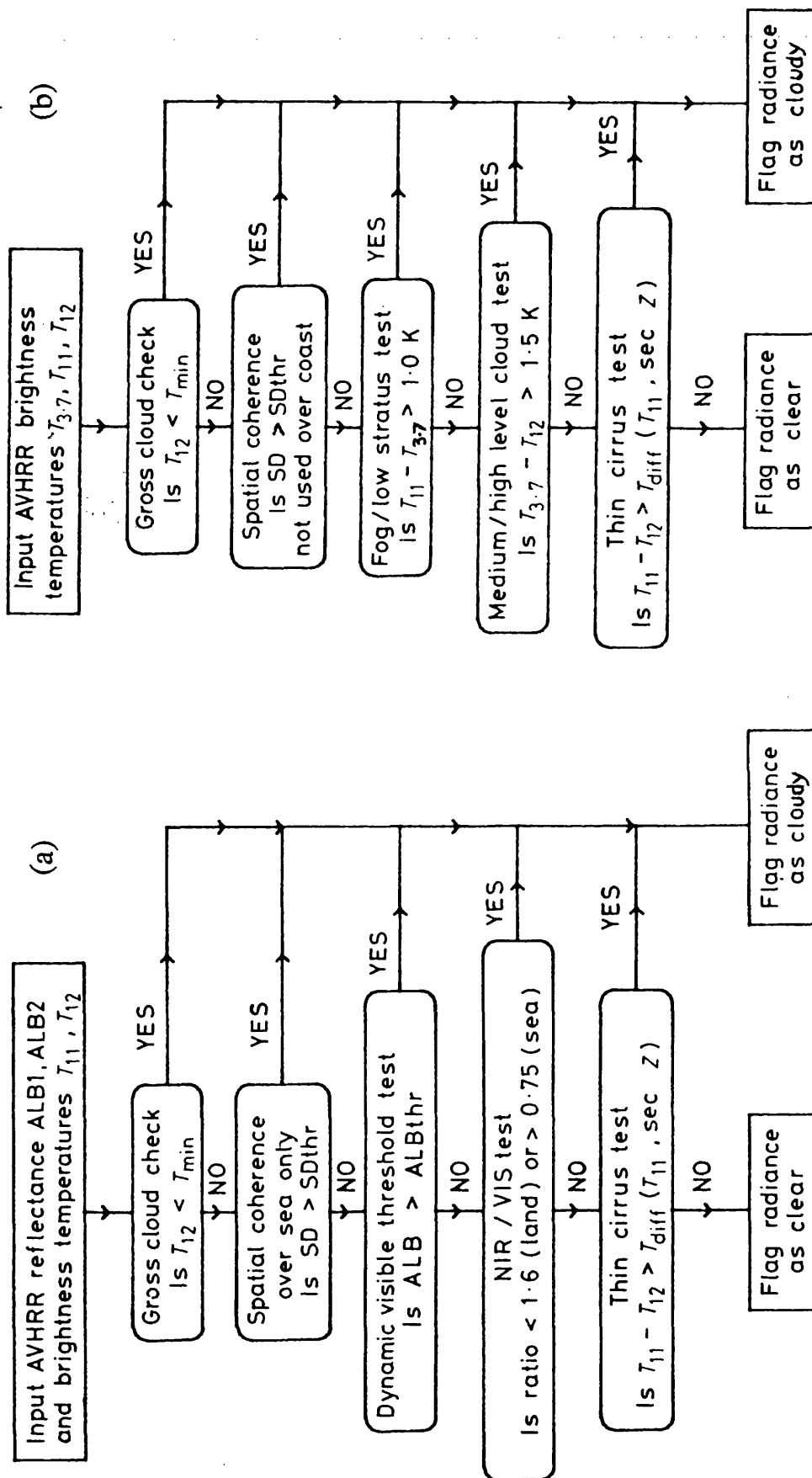


Figure 6.1 Schematic illustration of APOLLO decoupling tests for (a) day and (b) night AVHRR scenes. (From Saunders and Kriebel, 1988).

centre of a 3×3 square and rejected if its temperature was more than 0.3 degrees different from the mean of all the pixels within the box which had not been flagged as cloudy by previous tests. Saunders and Kriebel recommend that the threshold should be set to 1.0 degree for coastal regions due to the large surface gradients, but it was found that only a few pixels were being misidentified as cloudy with a threshold of 0.3 degrees, and these were in the regions of strongest gradient. Since it was considered most important to ensure the removal of all cloudy pixels the smaller threshold was used.

- iii) The test for medium level sub-pixel cloud was based on the difference between brightness temperatures in channels 3 and 5. The threshold was modified to 2.5 degrees since some apparently cloud-free areas were being flagged as cloudy using the default threshold suggested by Saunders and Kriebel of 1.5 degrees. The areas which were consistently being removed were the shallowest and therefore warmest parts of the lakes. Since these areas were likely to be much warmer than the atmosphere, we would expect less compensation for atmospheric absorption and therefore the differences between these channels to be greater since the absorption in channel 5 is greater than that in channel 3.
- iv) Lake Superior is extremely deep (up to ~300 metres) towards its Northern end and the inhibition of circulation results in a thermal bar (Hutchinson, 1975) at a temperature of 4°C which persists throughout the summer. The result of this is that a large area of the lake is at an almost constant temperature of 4°C, which is likely to be considerably cooler than the atmosphere above it. Thus there is the potential for warm air which comes off the land at night to cool over the lake and for fog to form. Thus it is frequently the case that much of Lake Superior fails the fog test which is based on the emissivity difference between channel 3 and 4; the test being positive if  $T_4 - T_3 > 1.0$  degrees.
- v) An alternative cause of higher temperatures in channel 4 than channel 3 might be the greater atmospheric emission in the optically thicker 10.3 - 11.3  $\mu\text{m}$  band than in the 3.45 - 3.95  $\mu\text{m}$  band. This situation would be unlikely over the ocean since maritime atmospheres' upward emission does not usually compensate fully for their absorption, but may be possible in the situation of Lake Superior where the water temperature is only 4°C. The stronger dependence of the Planck function on temperature for channel 3 for temperatures around 280 K ( $\sim T^{13}$  at 3.7  $\mu\text{m}$  as opposed to  $\sim T^5$  at 10.8  $\mu\text{m}$ ) means that the extra radiance contribution from atmospheric emission at the

longer wavelength will have more effect on the brightness temperature calculated via the inverse Planck function from the radiance sensed by the instrument. Whether this effect could produce brightness temperatures that are a degree or more higher in channel 4 than channel 3, resulting in the removal of potentially cloud-free pixels, is not clear, and so it was decided to perform the test with the default threshold since there was reason to believe that fog could form over much of Lake Superior by the mechanism outlined above.

- vi) The test for thin cirrus is based on the opacity difference in the 'split-window' channels for this type of cloud. The test is positive if  $T_4 - T_5 > \text{threshold}$ ; the threshold is calculated by interpolation from a table of values covering a range of channel 4 temperatures and air masses. The table values themselves were generated from the output of a radiative transfer model and a global set of maritime clear sky radiosondes (Saunders and Kriebel, 1988), and thus were probably too generous for the region in which they were being applied. In addition, this test was the last in a series of hierarchical tests and few pixels which had escaped the other tests were flagged as cloudy. The test did, however, help to remove some drop-outs which occurred in channel 5 but not channel 4.

Line and pixel drop-outs in channel 4 were usually removed by the spatial coherence test, but, particularly in the case of individual pixels, some drop-outs remained, mostly those which existed in channel 5 but not in channel 4. Two additional tests were added to ensure the removal of all remaining drop-outs. Drop-outs remaining in channel 5 were removed by a spatial coherence test similar to that applied to channel 4 in the declouding tests, and a final gross check threshold (the same one which had been applied to channel 5 during the declouding tests) was applied to the channel 4 brightness temperatures.

- vii) In addition, the single daytime scene had the channel 2 dynamic threshold set generously to allow an apparent aerosol contamination to pass the test.

### 6.3.2 Calculation of gradient and retrieved temperature

Having removed all pixels in the subscene that were subject to cloud contamination as determined by the above tests and additionally the pixel and line drop-outs in the data, the algorithms were applied to the brightness temperature images. The gradient  $\Delta T_5 / \Delta T_4$  was calculated for the image in the following way:

- A window of 32×32 or 64×64 pixels was defined, and a scattergram of  $T_5$  against

$T_4$  was generated for the remaining cloud-free pixels within the window if the number of cloud-free pixels was at least 200.

- The mean and standard error of all  $T_5$  values within each  $T_4$  bin of the scattergram were calculated and these values were passed to a weighted least squares fitting routine, using the standard error on the mean value of  $T_5$  to determine the weighting.
- The routine returned the straight line coefficients together with an error on the fit. The maximum allowable error on the fit was 0.02 for the 32×32 pixel window and 0.01 for the 64×64 pixel window. The reason for the different error tolerance thresholds was that the 64×64 pixel window potentially had four times as many pixels available for the gradient calculation and it was expected that the error would at least reduce as the root of the number of pairs of pixels used. The additional advantage of the larger window area was that it could cover larger regions with larger spread of temperatures, but this might be offset by the risk of greater atmospheric variability.
- The windows were moved across the subscenes in steps of 8 pixels for the smaller window and 16 pixels for the larger window. The gradient calculation was assigned to the central 8×8 or 16×16 pixel box within each window position; thus the results were a smoothed average for the whole window. This approach implies that it was expected that any change in the atmosphere within the window could be fitted well by a first order expression.

The algorithms for surface temperature retrieval could then be applied to the data. The values of  $T_4$ ,  $T_5$ ,  $\Delta T_5/\Delta T_4$  and scan angle for each cloud free pixel within the image were used to calculate the corrected surface temperature image. The procedure was as follows:

- The airmass for each pixel was calculated as the cosecant of the viewing angle, itself obtained from simple trigonometry using the scan angle and satellite orbital height.
- The coefficients for the tabulated airmasses immediately above and below the actual airmass were used in conjunction with the values of  $T_4$ ,  $T_5$  and  $\Delta T_5/\Delta T_4$  for each pixel to calculate corrected surface temperatures. The results using the conventional 'split-window' coefficients were obtained in a similar fashion.



- The surface temperature for the pixel was calculated from a simple linear interpolation by airmass between the results for the two tabulated airmasses. Results beyond 2.0 airmasses were calculated by linear extrapolation of the values obtained from the coefficients for 1.75 and 2.0 airmasses.

The results for all algorithms were averaged over 4×4 pixels in order to reduce the effects of detector noise and digitisation of the signal. The NEΔT quoted for AVHRR is 0.12°C per pixel (ITT, 1989) and with the digitisation interval being ~0.1°C at 285K, we would expect to be able to remove this latter effect by averaging over 16 pixels since, for a single space-observed brightness temperature, a significant number of these pixels should fall outside one digitisation bin. The effective NEΔT is expected to be reduced to 0.03°C by this averaging procedure. Llewellyn-Jones *et al* (1984) report that no significant improvement in retrieval error could be achieved by reducing the NEΔT below 0.02°C (averaging over 6×6 pixels) thus we expect the above procedure to be a good compromise between brightness temperature errors propagated through the algorithms and the contribution to error resulting from point comparisons of buoy and satellite data since the Great Lakes display significant spatial variability of surface temperature.

### 6.3.3 Image Rectification and Geolocation

Having generated corrected surface temperature images for the algorithms, it was necessary to geolocate the scenes in order to compare the temperatures at the buoy sites with the *in situ* measurements. It was decided to break this task down into two steps:

- 1) resampling of the channel 4 brightness temperature image and corrected temperature images for the various algorithms in order to remove all non-linear distortions, and
- 2) warping the resultant images to a map containing the shorelines of the Great Lakes obtained from the World Data Base II and the buoy positions.

The above approach has several advantages:

- Non-linear distortions in the image can be calculated analytically and removed in an automated manner.
- If these distortions, which have a known mathematical form, are not removed before warping the image, the warper is unlikely to be able to adapt its polynomial

coefficients to successfully mimic the function required to remove them over the whole image, although it may be successful in removing the distortions within a smaller region of the image.

- The removal of all non-linear distortions before warping the image to the map allows use of simple first order equations of the form:

$$X_{out} = a_0 + a_1X_{in} + a_2Y_{in} \quad (6.2)$$

and

$$Y_{out} = b_0 + b_1X_{in} + b_2Y_{in} \quad (6.3)$$

The geometrical transformations that such equations can perform are limited to resizing, rotating, translating and skewing the image, which is all that was required and desired.

- The generation of coefficients for equations 6.2 and 6.3 requires only three coordinate pairs of tiepoints, thus greatly reducing the amount of operator time required to warp each image. By comparison, a bi-quadratic warp would have required the identification of at least 10 well-spaced tiepoints in each scene, and the misidentification of any of the tiepoints would have been much harder to detect and correct.

The resampling procedure itself was broken down into two steps. Firstly, the effect of panoramic distortion in the across-track dimension was removed as follows:

- The along-track separation between the centres of successive scans was calculated using the satellite altitude scan rate of the instrument assuming the satellite velocity could be approximated to that obtained from a circular orbit. No correction was made for the effect of Earth rotation on the separation between the centres.
- The distances from the edge of the subscene for successive pixel centres across the scan line were calculated using the following equation:

$$D(i) = [\beta(i) - \beta(1)]R_{\oplus} \quad (6.4)$$

where  $D(i)$  is the distance on the surface of the Earth of pixel  $i$  from the first

pixel in the subscene,  $R_{\oplus}$  is the radius of the Earth and  $\beta(i)$  is as shown in Figure 6.2 and can be calculated using the following equation:

$$\beta(i) = \sin^{-1}[\sin(n(i) \times \eta)(R_{\oplus} + h)/R_{\oplus}] - n(i) \times \eta \quad (6.5)$$

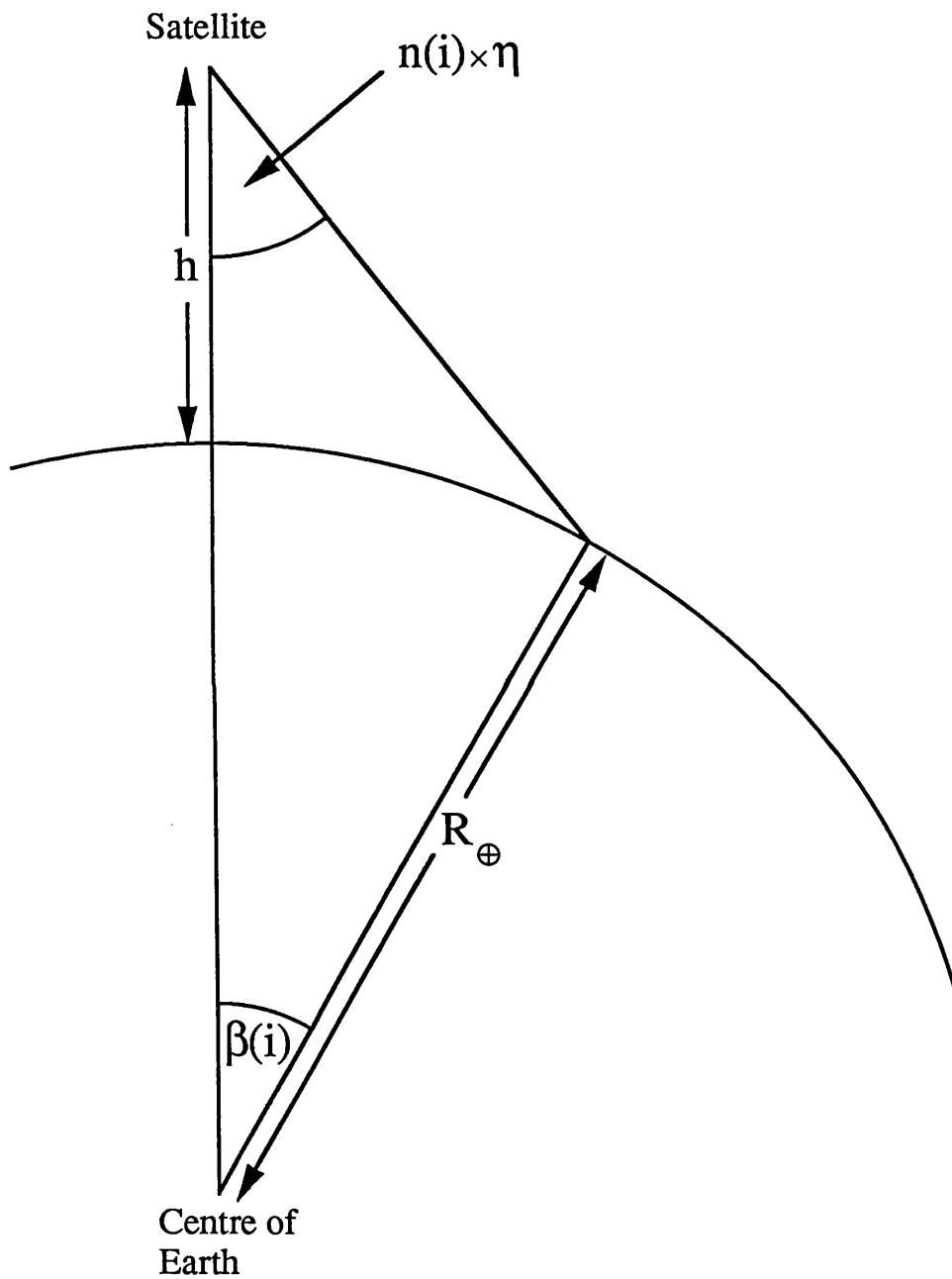
Here,  $h$  is the satellite altitude,  $n(i)$  is the number of pixels from nadir (image centre) for pixel  $i$ , and  $\eta$  is the angle in radians that the scan mirror of the instrument turns through between successive pixel centres. This latter parameter is taken to be a constant, its value being calculated from the knowledge that 39,936 samples are available for every second (six revolutions of the scan mirror).

- The pixel values along each scanline were then linearly resampled in increments of equal distance, the distance chosen being the along-track separation between successive scan lines which had been calculated in the first step.

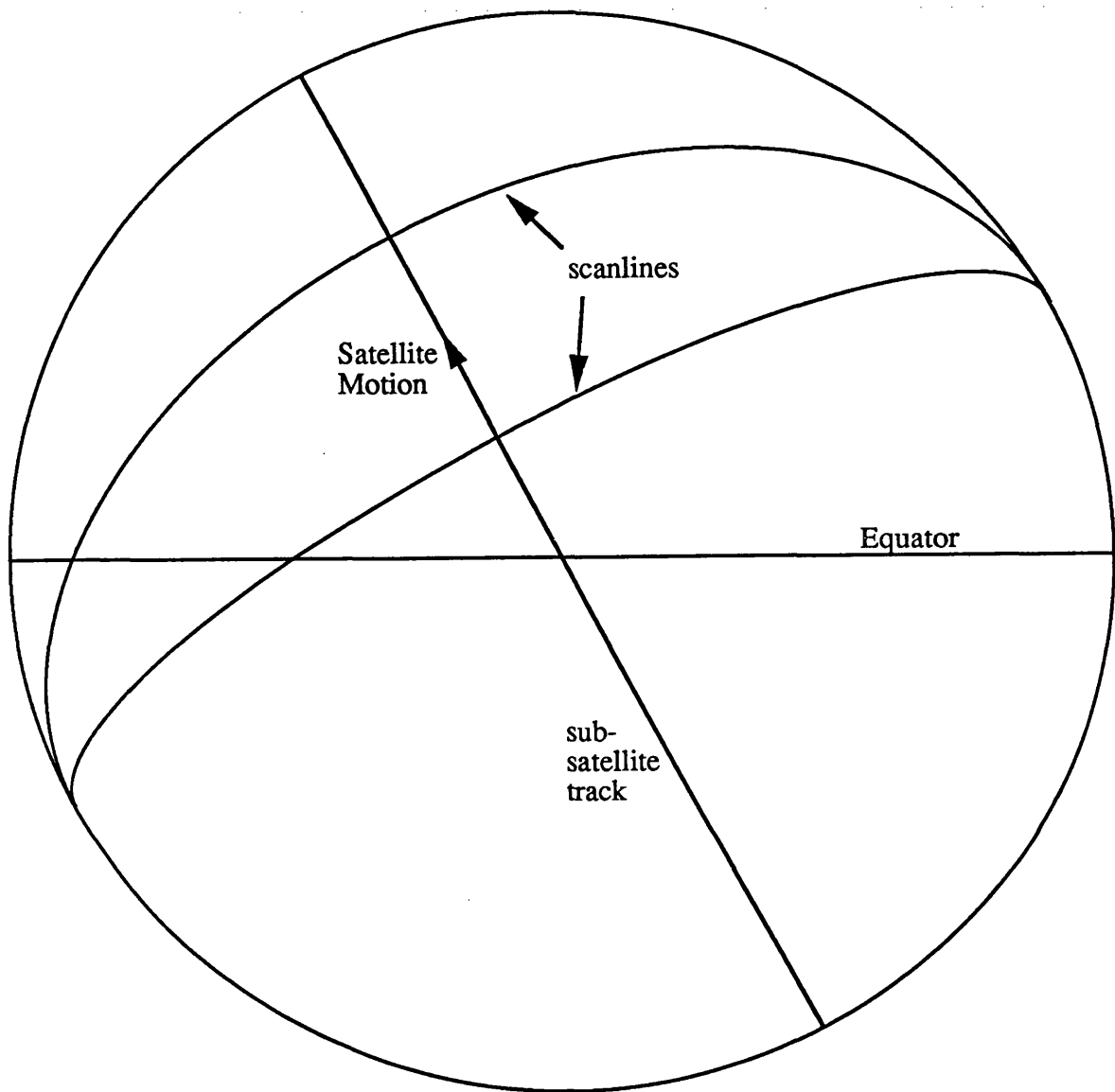
Secondly, the distortion due to the separation between scan lines decreasing away from nadir was removed, again by linear resampling but this time in the along-track direction. The effect is similar to that of longitude lines becoming closer together away from the equator; in this case the equator is represented by the sub-satellite track and the longitude lines are represented by scan lines. The geometry of the situation is illustrated in Figure 6.3. The actual steps to the removal of this distortion are as follows:

- The angle at Earth centre (the angle  $\beta$  in Figure 6.2) from the sub-satellite point is calculated for each across-track pixel position in the image. Each successive column reduces its angle by the same amount since the scan lines have already been resampled to represent equal across-track distance per pixel.
- Each along-track column of pixels is linearly resampled to a length that is  $\cosine(\beta)$  of its original length, thus pixel columns at nadir do not change their length. The use of cosine means that this procedure is unaffected by the angle  $\beta$  becoming negative past the nadir position. The resampled columns are shifted so that their centre positions remain unaltered.

The resampled subscenes were then warped to an image containing the shorelines of the Great Lakes generated from the World Data Base II and the buoy positions. The warping procedure was first carried out on the channel 4 brightness temperature images since it was easier to identify coastal features before declouding and pixel



**Figure 6.2** Trigonometry required for calculation of the angle subtended at the centre of the earth,  $\beta$ , from knowledge of the across-track pixel number in the image and application of the Sine rule.



**Figure 6.3** An illustration of the 'barrel distortion' effect on AVHRR images resulting from the extended swath of the instrument imaging points significantly closer to the pole of the orbit than the great circle represented by the sub-satellite track.

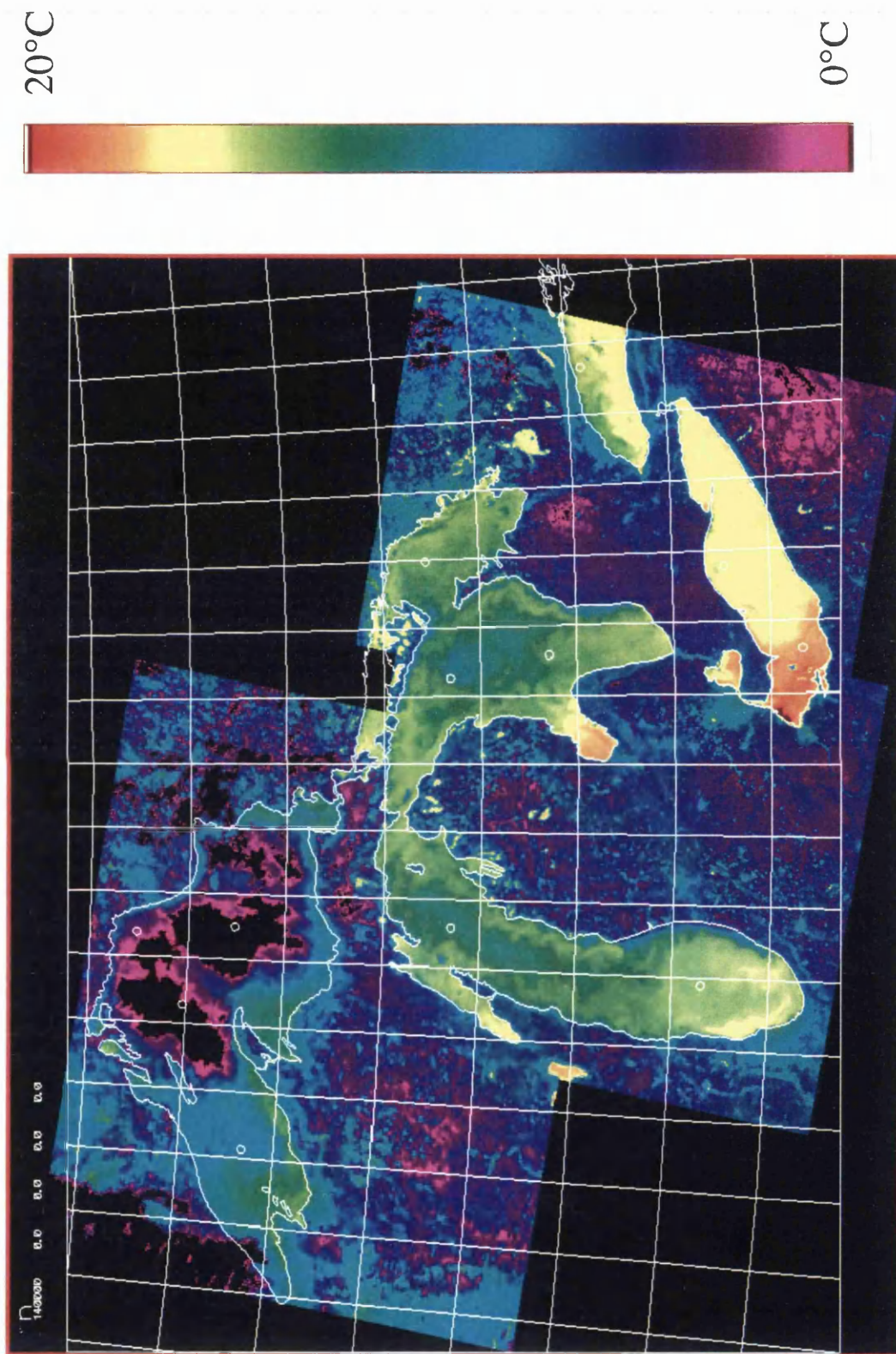
averaging. Three well-spaced tiepoints were selected and the warp applied. The fit of the coastline to the image around each tiepoint was closely examined, and the tiepoints adjusted in turn until a good fit was obtained for all three locations. A final inspection of the overall fit was used to confirm the suitability of the tiepoints selected; visual inspection indicated an error within  $\pm 1$  pixel r.m.s. for the scenes. Figure 6.4 shows an example of the geolocation procedure; three resampled channel 4 subscenes, the coastline and buoy positions marked as circles can all be seen. Other features that can be seen are the temperature gradients in the lakes and the range in brightness temperatures from Lake Superior ( $\sim 4^{\circ}\text{C}$ ) to Lake Erie ( $\sim 18^{\circ}\text{C}$ ).

The tiepoints identified in warping the channel 4 images were then reapplied to the corrected temperature images and the results obtained for the various algorithms at the marked buoy positions that had passed the cloud clearing tests.

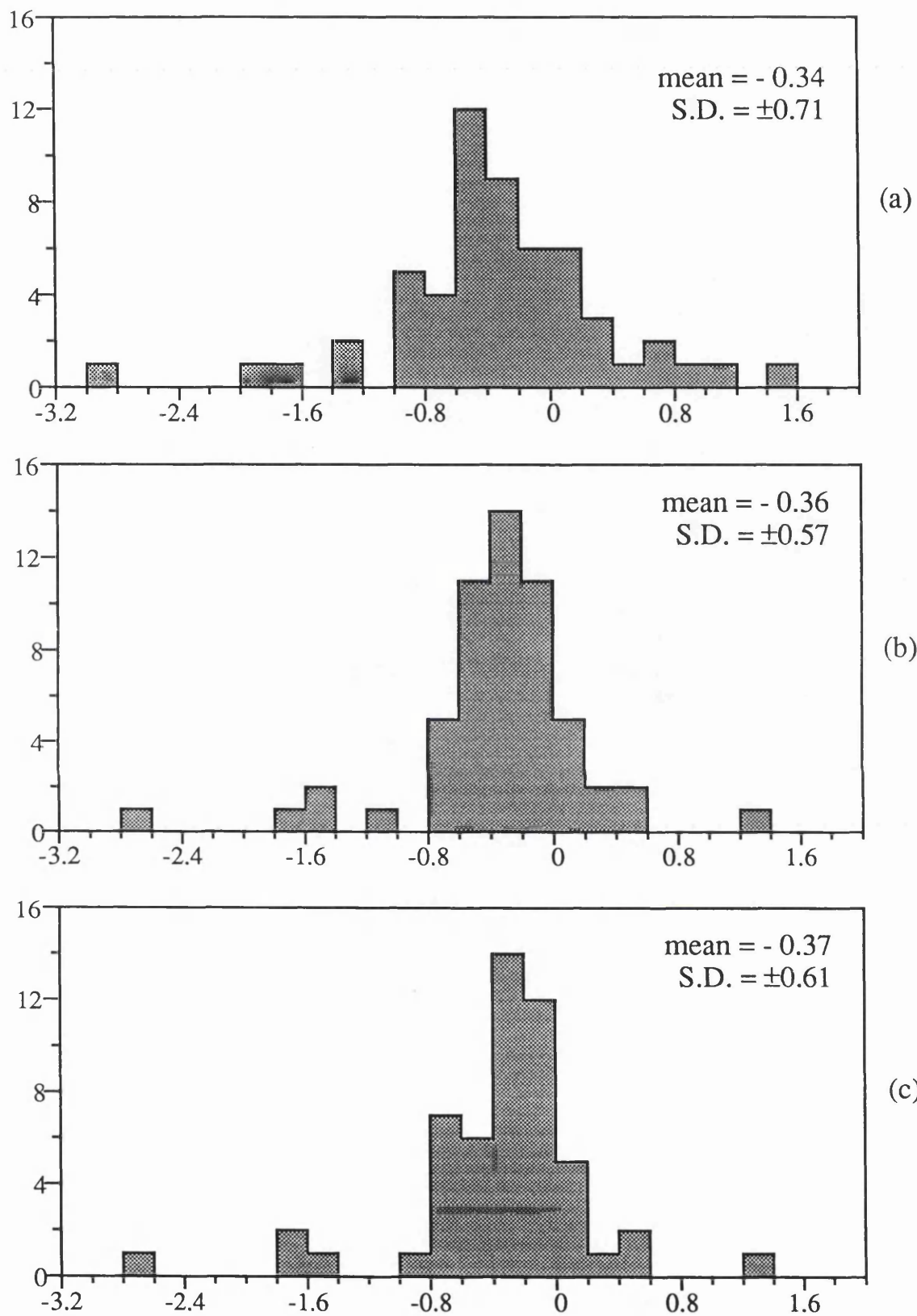
#### 6.4 Results obtained for the various algorithms

The analysis yielded 56 cloud-free satellite - *in situ* comparisons for the standard algorithm and the new algorithm using the gradient calculated from the  $64 \times 64$  pixel window. The variation of the new algorithm which used the gradient calculated from a window of  $32 \times 32$  pixels yielded 53 satellite - *in situ* comparisons since, on a very few occasions, the gradient error was above the threshold of 0.02. As mentioned previously, the lakes were often at significantly different temperatures from the surrounding land which will have influenced the lower layers of the atmosphere, thus we would not expect the atmosphere to be in thermal equilibrium with the lakes and therefore at its most thermally unstable. The fact that the gradient error calculated using the  $64 \times 64$  pixel window never exceeded the threshold of 0.01 at the comparison sites is an indication that the atmospheric regime did not change significantly on this scale.

The frequency histogram of *in situ* temperature minus retrieved temperature for the conventional 'split-window' algorithm is shown in Figure 6.5(a). Note the large central peak and also the large spread of outliers which have a range from -2.9 degrees, representing an overestimate of the temperature by the algorithm, to +1.6 degrees which represents an underestimate. Note also the large standard deviation of the distribution ( $\pm 0.7$  degrees) and the mean bias of -0.3 degrees. Figure 6.5(b) shows the same information for the new algorithm using the gradient calculated from  $64 \times 64$  pixel windows. Here, the central peak is particularly narrow and the presence of the outliers, ranging from -2.7 degrees to +1.3 degrees, is even more noticeable. The



**Figure 6.4** Result of the geolocation procedure applied to three contemporaneous channel 4 subscenes of the Great Lakes. Note the good fit of the coastline, the buoy positions and the range of temperatures in the lakes. The dark areas in the middle of Lake Superior are cloud.



**Figure 6.5** Frequency histograms for the conventional (a) and new atmospheric correction algorithms using the gradient calculated from 64x64 and 32x32 pixel windows ((b) and (c) respectively). Note the different shapes of the peaks for the different algorithms and the presence of outliers in all three plots.



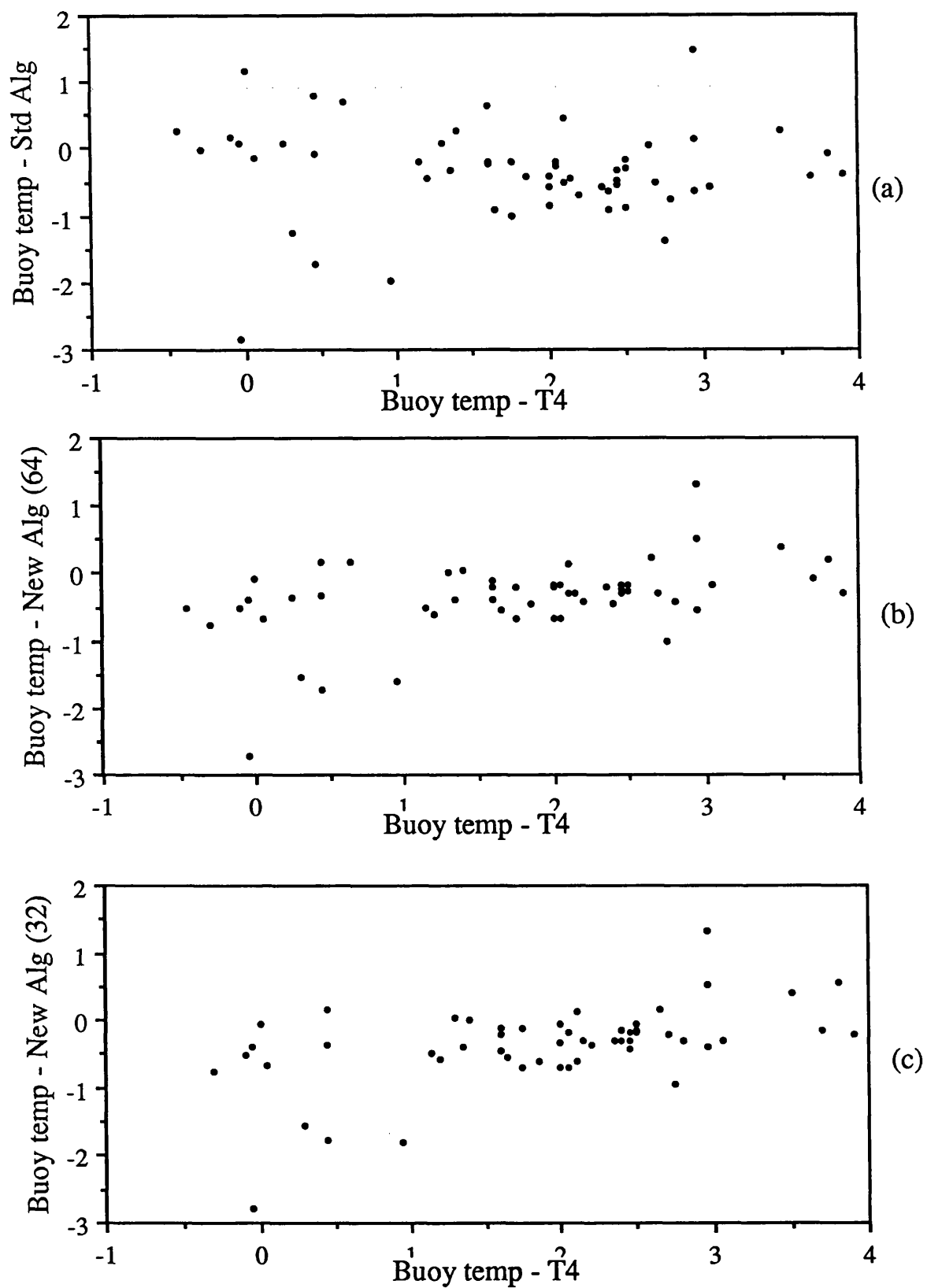
standard deviation is slightly reduced from  $\pm 0.7$  degrees to  $\pm 0.6$  degrees but the mean bias is virtually unchanged. The histogram for the new algorithm using the gradient obtained from the smaller window size is shown in Figure 6.5(c). The features that can be seen are very similar to those for the other variant of the new algorithm, with the mean and standard deviation being virtually identical.

The plots of *in situ* temperature minus retrieved temperature versus *in situ* temperature minus channel 4 brightness temperature for the various algorithms are shown in Figure 6.6. The X-axis parameter illustrates the range of atmospheric corrections for the dataset, although the true *in situ* skin temperature - satellite temperature difference is not known. The apparent atmospheric corrections range from  $\sim 0$  degrees to 4 degrees, and there are some trends in the spread of *in situ* - algorithm temperature difference that can be seen. The main feature, particularly noticeable for the new algorithm, is the increase in spread at low and high atmospheric corrections. In addition, for the new algorithm, there is a slight dependence of algorithm error on the atmospheric correction, with overcorrection being predominant at low values of atmospheric correction and undercorrection at high values. The trend for the conventional algorithm is less clear, but appears to be in the opposite direction.

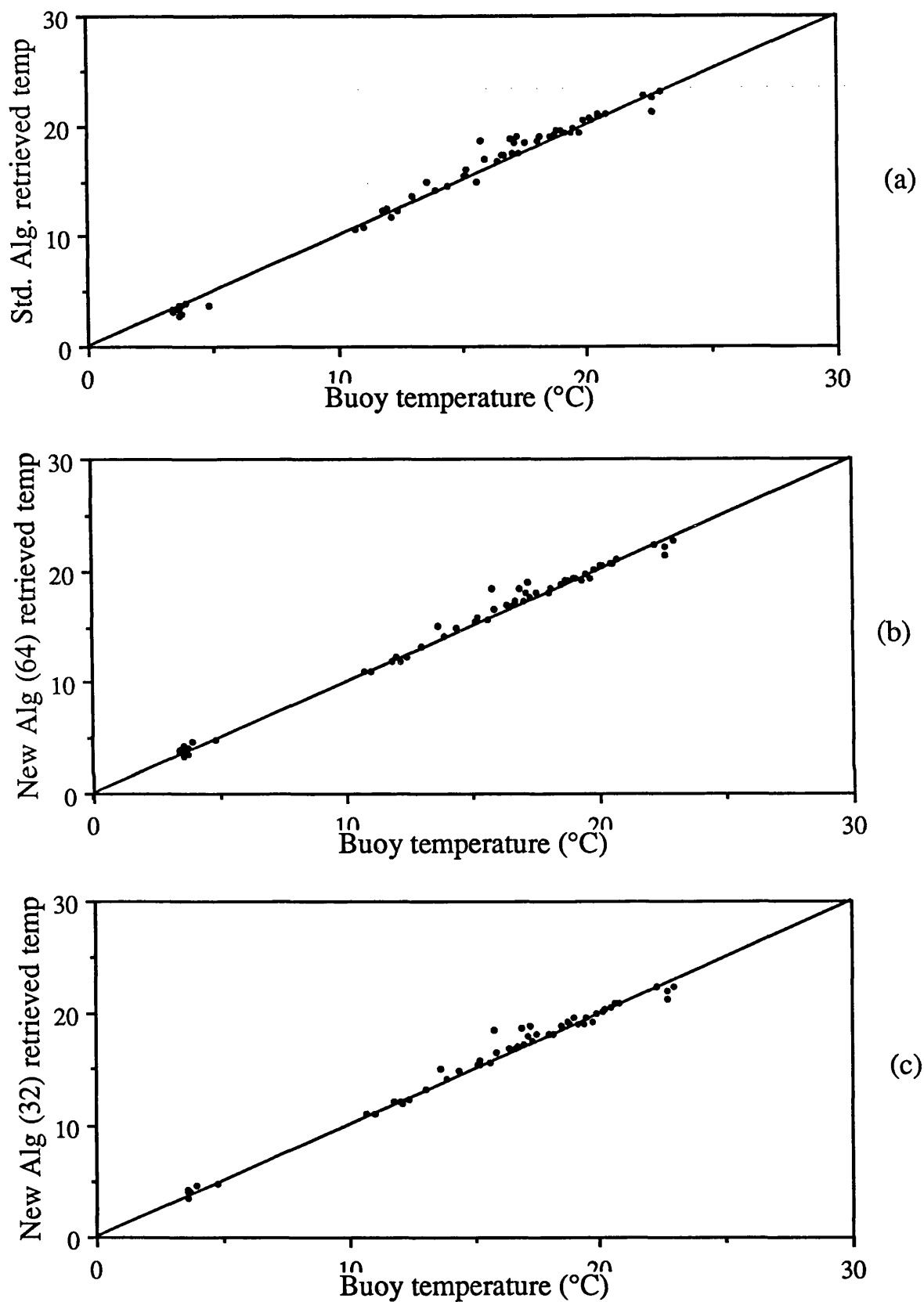
The plots of retrieved temperature versus *in situ* temperature for the conventional and new algorithms are shown in Figure 6.7. The main point illustrated by this figure is the range of temperatures within the dataset, although it is also possible to distinguish the trends of overcorrection and undercorrection described above, except that they occur for low and high temperatures rather than low and high corrections. Figure 6.8 shows the plot of *in situ* minus channel 4 temperature versus *in situ* temperature; a trend of low atmospheric correction for low temperatures and high atmospheric correction for high temperatures can be seen. Possible reasons for these and other results will be discussed in the next section together with any implications for the various algorithms.

## 6.5 Discussion of Results

The main purpose of this validation study is to assess the feasibility of the new atmospheric correction technique and to decide if it is capable of providing a significant improvement in atmospheric correction with respect to the conventional 'split-window' method when applied to real AVHRR data. To this end, a reasonable range of both atmospheric corrections and surface temperatures is required and these



**Figure 6.6** Plots of *in situ* minus retrieved temperature versus *in situ* minus channel 4 temperature are shown for the various algorithms. Note the dumbbell nature of the spread with outliers prevalent at low and high corrections, particularly for the new algorithms (64×64 (b) and 32×32 (c) respectively).



**Figure 6.7** Retrieved temperatures versus *in situ* temperatures for the conventional (a) and new algorithms (64×64 pixel window (b) and 32×32 pixel window (c)).

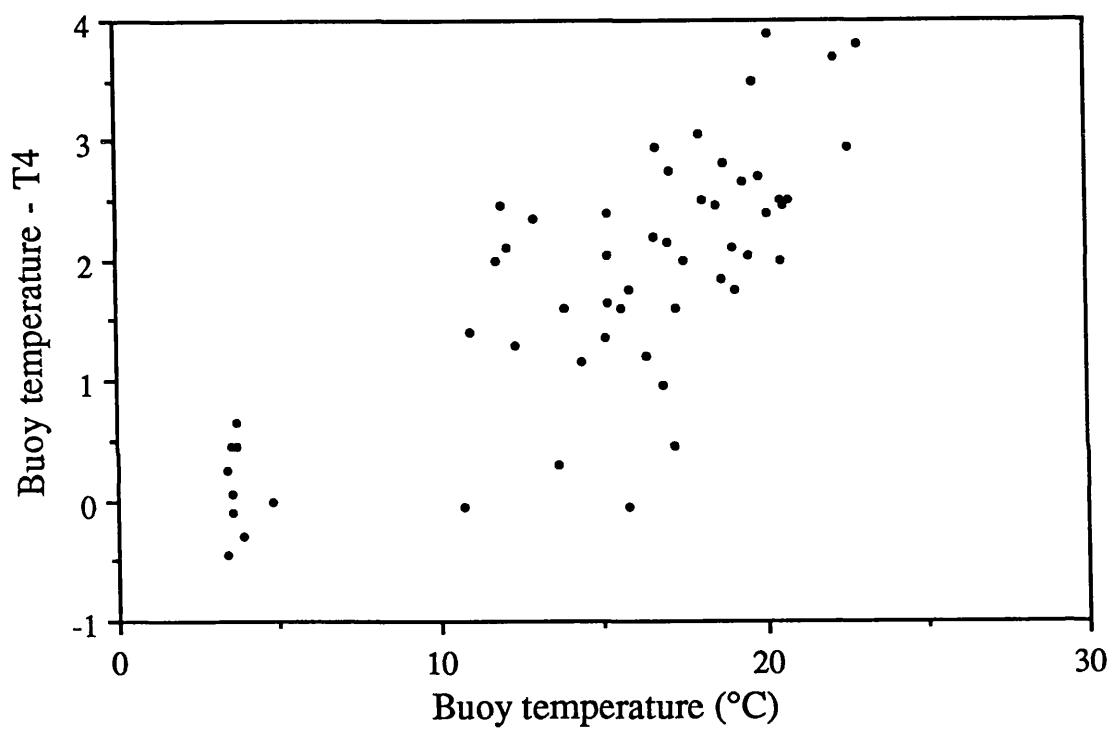
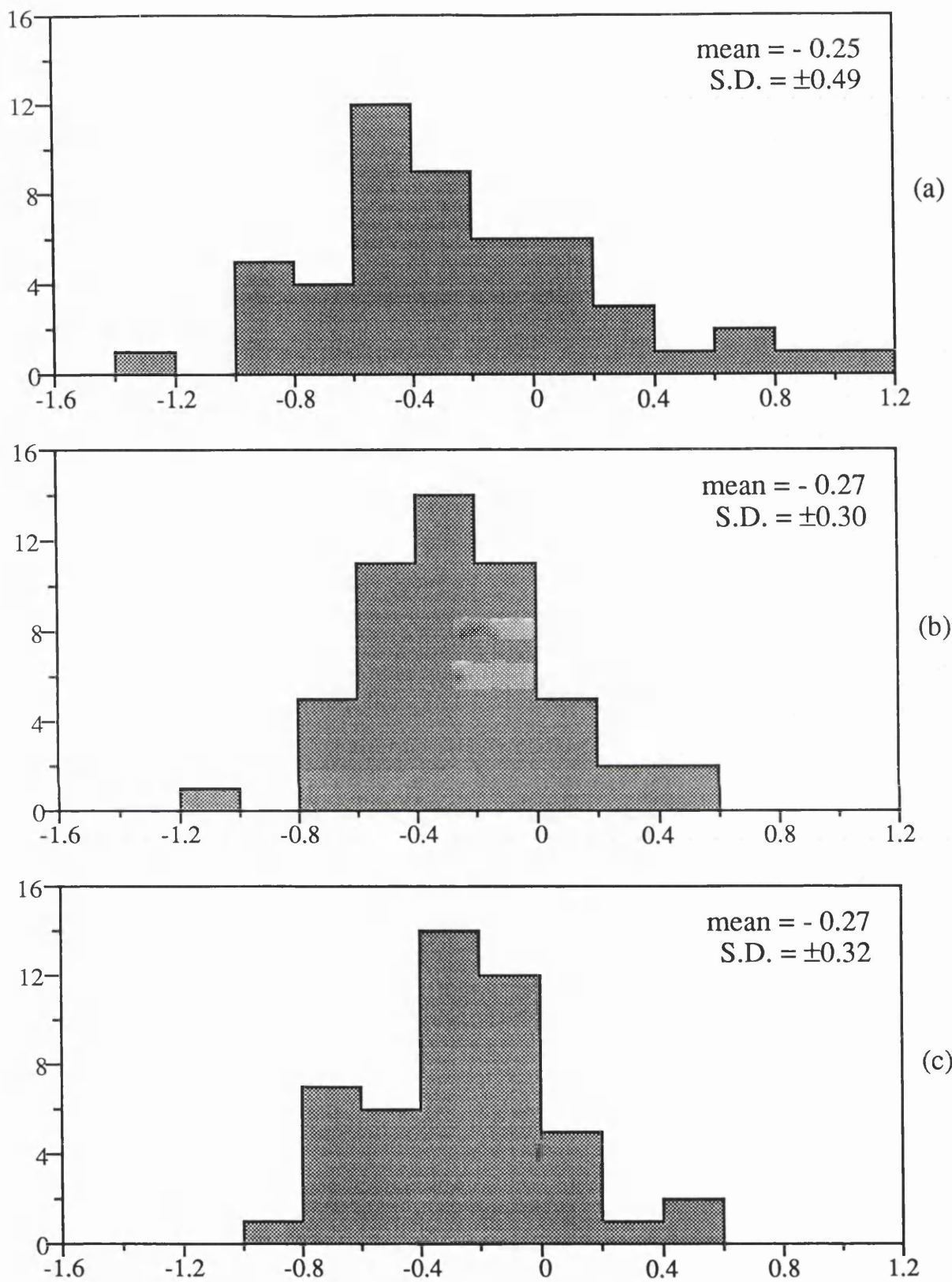


Figure 6.8 *In situ* minus channel 4 temperature versus *in situ* temperature is plotted to illustrate the relationship between surface temperature and atmospheric correction.

have been provided by the Great Lakes dataset. The plot of *in situ* minus channel 4 temperature versus *in situ* temperature shown in Figure 6.8 demonstrates a clear trend between atmospheric correction and surface temperature. For the oceanic case, a higher water temperature results in more water vapour being liberated into the atmosphere, thus the atmospheric absorption increases and atmospheric correction along with it, thereby producing a similar trend. Since the atmospheres over the different lakes must be reasonably similar due to their geographic proximity, we may conclude that, in the case of the Great Lakes, the relationship occurs for different reasons than it does over the ocean. In addition, the range of gradients for the dataset was from 0.85 to 0.93, implying similar atmospheres for the duration of the study as well as spatially.

For this validation study, it is possible that the difference in temperature between the atmosphere and lake surface is the main cause for both the large and small atmospheric corrections observed. If the lake is much hotter than the atmosphere, as may be the case for temperatures in Lake Erie, there will be little compensation for absorption of the surface-emitted radiance by atmospheric emission, and a relatively large temperature deficit will result. Conversely, when the temperature of the lake is much lower than that of the atmosphere, as may be the case for Lake Superior, there will be significant emission from the atmosphere, resulting in a very small brightness temperature deficit.

The outliers in the histogram results for both the new algorithm and the conventional 'split-window' method (shown in Figure 6.5) have a disproportionate effect on the magnitude of the standard deviation calculated for the *in situ* minus retrieved temperatures. If the five worst outliers for the new algorithm are excluded from the analysis, the histograms now appear as shown in Figure 6.9. Note that the outliers were common to all algorithm results, and all histograms now show a significant decrease in spread. The standard deviation for the conventional 'split-window' algorithm is reduced from  $\pm 0.7$  degrees to  $\pm 0.5$  degrees, but the mean bias is virtually unchanged. The results for the new algorithm show a more significant improvement; for both the  $64 \times 64$  and  $32 \times 32$  pixel windows the r.m.s. spread is reduced from  $\pm 0.6$  degrees to  $\pm 0.3$  degrees, but the mean biases remain unchanged at  $-0.3$  degrees. In addition, the histograms now appear more normal in nature upon visual inspection. Since the r.m.s. difference between the  $64 \times 64$  and  $32 \times 32$  window results is  $\pm 0.1$  degrees, with zero bias, further discussion of the new algorithm will be confined to the results obtained from the  $64 \times 64$  pixel window version of the algorithm.



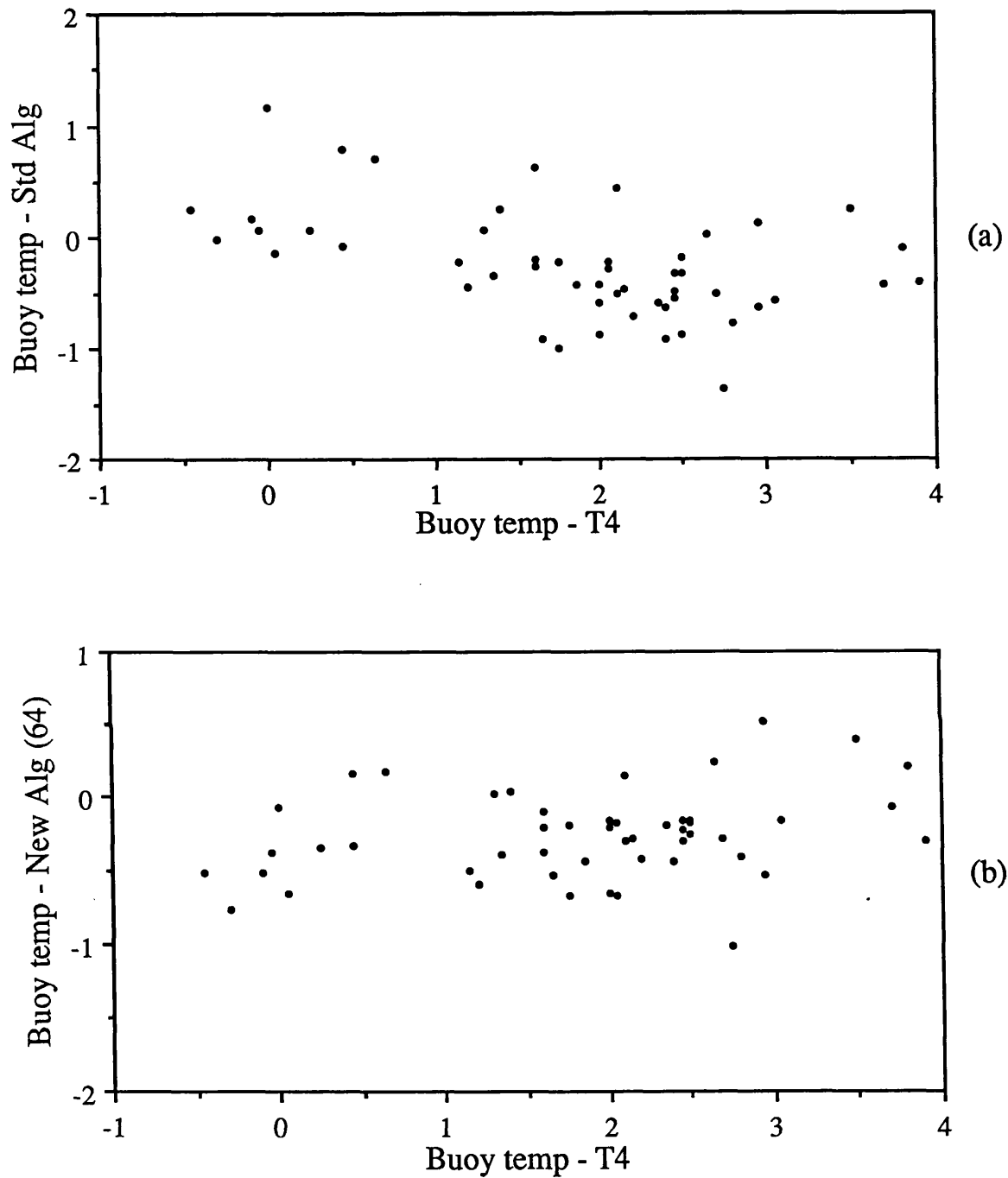
**Figure 6.9** Frequency histograms for the conventional (a) and new atmospheric correction algorithms using the gradient calculated from 64x64 and 32x32 pixel windows ((b) and (c) respectively). The five worst outliers have been removed in order to facilitate the study of the main distributions and their attendant statistics, which are much improved for the new algorithms.

One possible cause of the outliers is that of large skin effects on the surface of the lakes. As has been suggested previously, the lower layers of the atmosphere may be at significantly different temperatures to the lake surface, particularly at the extremes of the water temperature range, for example, in Lake Superior and Lake Erie. It is this situation which has been suggested to be a possible cause of the trends seen in the range of atmospheric corrections observed (see Figure 6.8). In such cases, the atmosphere will be trying to establish thermodynamic equilibrium with the lake and significant heat exchange may take place at the boundary between water and air. The magnitude of the difference between the surface skin temperature, which the radiometer sees, and the bulk temperature, which is measured by the buoy, reflects the amount of heat exchange taking place at the boundary. The three main heat exchange processes which can operate in the absence of downwelling shortwave radiation, which is the case for all but one of the images, are (eg. Robinson *et al*, 1984, Schluessel *et al*, 1990):

- Conduction (sensible heat transfer)
- Latent heat transfer, either evaporative cooling or heating due to condensation
- Radiative heat transfer

The final parameter which can have an effect on the magnitude of the skin - bulk temperature difference observed is the wind speed, since a strong wind can quickly break down any skin effect, although the difference quickly re-establish itself after the disturbance has ceased (eg. Robinson *et al*, 1984, Hepplewhite, 1989). Since the wind is likely to be the most variable parameter on short timescales, we might expect the observed skin effects to be most variable when the magnitudes are large. Indeed, Figure 6.6(b), for example, shows that the spreads in *in situ* minus retrieved temperatures are largest at high and low atmospheric corrections, where we expect the net heat transfers, and consequently the skin effects, to be at their greatest. This also provides some justification for removing the outliers from the analysis, namely that they may possibly be due to large differences between the skin and bulk temperatures at the *in situ* sites rather than errors in atmospheric correction.

If we now consider the plots of *in situ* minus retrieved temperature against *in situ* minus channel 4 brightness temperature for both the conventional and new algorithms with the outliers removed (as shown in Figure 6.10), there are still some important trends that can be discerned. The result for the new algorithm (Figure 6.10(b)) shows a trend of overestimation of surface temperature for low atmospheric corrections, and

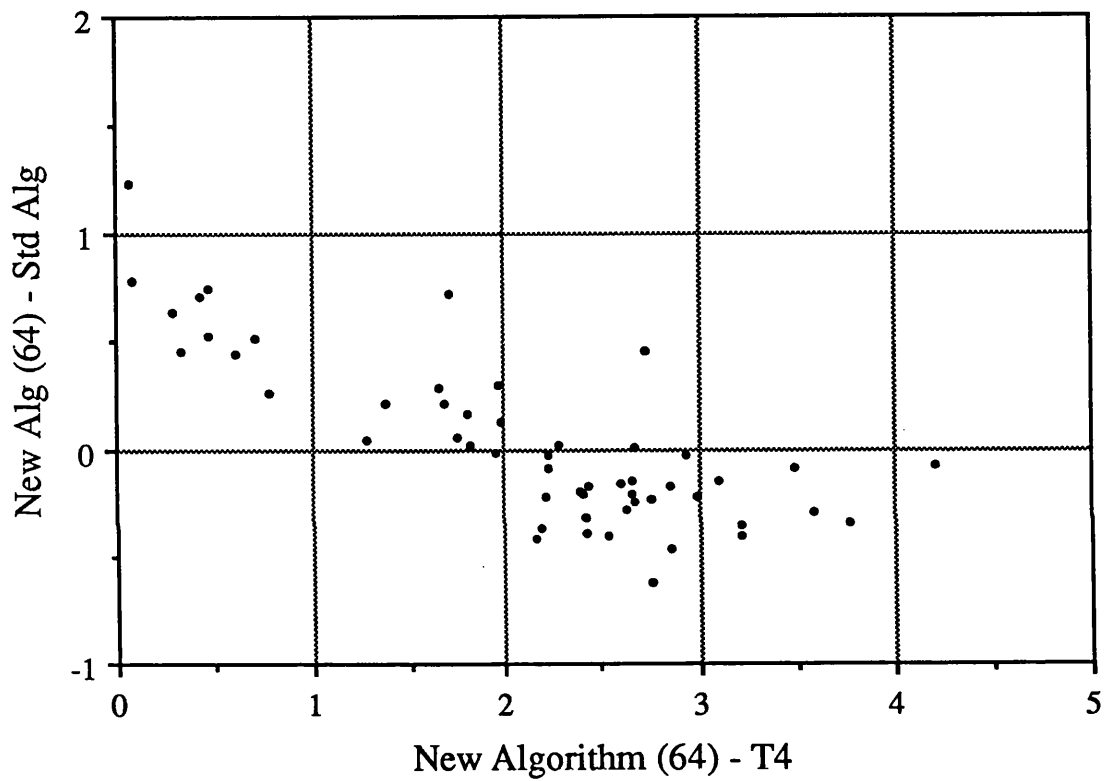


**Figure 6.10** *In situ* minus retrieved temperature versus *in situ* minus channel 4 temperature is plotted for the conventional 'split-window' algorithm and the 64x64 variant of the new algorithm. Note the different trends in atmospheric correction error for the two algorithms.

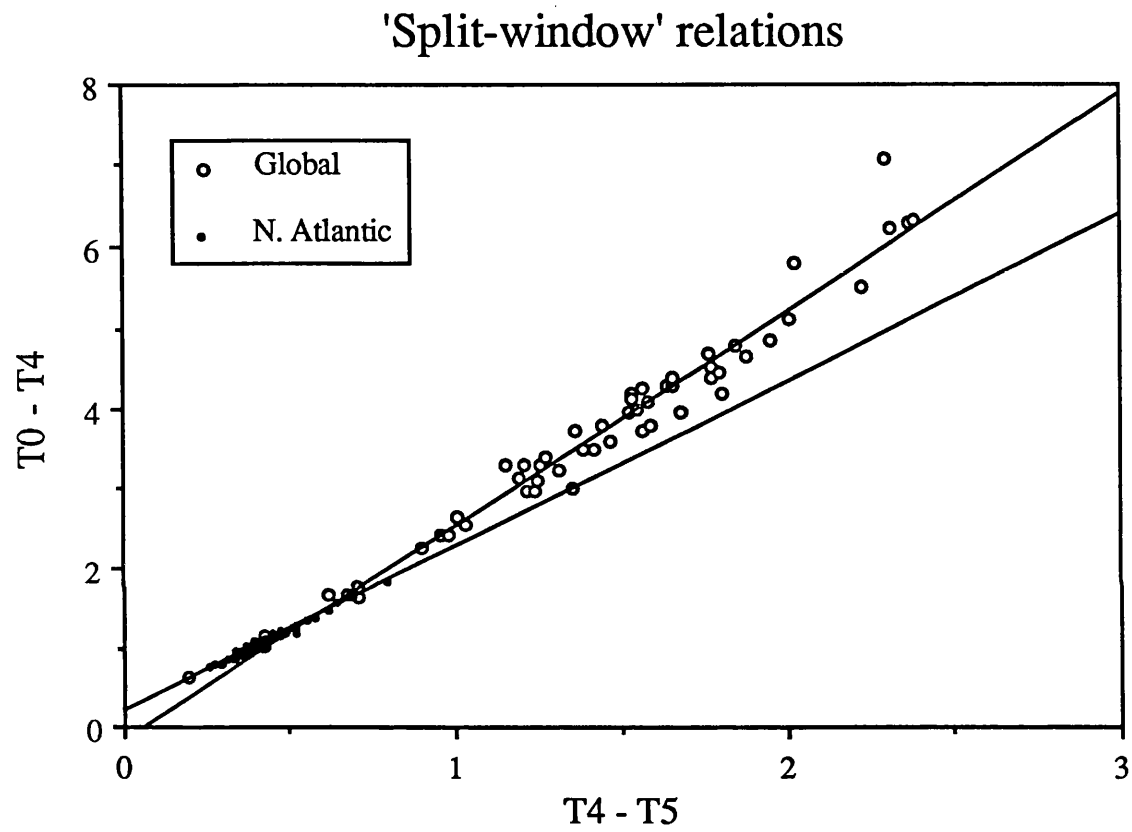


underestimation for higher atmospheric corrections. As can be seen from previous discussion, we would expect the net transfer of heat for low atmospheric corrections to be into the lake since the atmosphere is likely to be much warmer than the water. This would result in a positive skin effect, since the atmosphere would be trying to deposit its heat into the lake through the boundary, and may be the cause for the temperature retrievals for the new algorithm being higher than the *in situ* buoy measurements. For the case of large atmospheric corrections, where the lake is much warmer than the atmosphere, it is likely that the net heat flux will be out of the lake, resulting in a negative skin effect; this could be the reason for temperature retrievals being lower than the *in situ* buoy measurements.

The plot for the 'split-window' method (Figure 6.10(a)) shows the reverse of the trend observed for the new algorithm. Since the arguments for the former trend have been established, it is important to discuss the possible reasons for this further potential failing of the conventional algorithm. If we now consider the new algorithm to be measuring the 'true' skin temperature (still with some scatter) and plot the differences between the retrieved temperatures for the new algorithm and 'split-window' method against atmospheric correction (new algorithm retrieval minus channel 4 brightness temperature), a clear trend in atmospheric correction error can be seen (Figure 6.11). The difference between the retrievals for the new and conventional algorithms changes by about one degree as the atmospheric correction moves from ~0 degrees to 4 degrees. The most likely reason for this trend is that the 'split-window' coefficients were derived from a global set of ocean radiosondes, some of which had significant atmospheric absorption (see Chapter 5). Atmospheres over continental North America are likely to have a low water vapour content, and thus be of reasonably high transmittance, indeed, as mentioned earlier, the gradient values for the dataset ranged from about 0.85 to 0.93, thus indicating that no atmospheric transmittances were lower than 0.7 in channel 4. If we compare a 'split-window' algorithm generated from a set of radiosondes obtained in the North Atlantic, representing high atmospheric transmittances and consequently more likely to be representative of atmospheres over the Great Lakes, with the one used in this analysis (see Figure 6.12), a significant difference in atmospheric correction can be seen. As the atmospheric correction changes from zero to 4 degrees, the difference in atmospheric correction for the two algorithms changes by about one degree, similar to the change observed in Figure 6.11.



**Figure 6.11** New algorithm retrieval minus conventional algorithm retrieval versus new algorithm retrieval minus channel 4 temperature is plotted as an attempt to discern the trend in atmospheric correction error for the conventional algorithm. The retrieval for the new algorithm is regarded as the 'true' skin temperature.



**Figure 6.12** 'Split-window' relations derived from 56 global and 57 North Atlantic radiosondes are plotted to illustrate the difference in atmospheric correction predicted by the two algorithms over the range of 0 to 4 degrees of correction.

## 6.6 Conclusions and Recommendations

The first point to be made is that application of the new algorithm to real AVHRR data is feasible and relatively straightforward. More importantly, for this case study, the new algorithm demonstrates a significant reduction in atmospheric correction error over the conventional 'split-window' method from  $\pm 0.5$  degrees to  $\pm 0.3$  degrees, confirming the potential outlined in chapter 5. Indeed, we cannot expect to achieve better results than this using buoy measurements as ground truth since the variability of the skin effect is known to be of this order (Robinson *et al*, 1984). The bias in retrievals of  $+0.3$  degrees for all algorithms is not easily explained, but may be due to calibration error, (eg. Brown *et al*, 1991, report a similar bias for NOAA-11 satellite retrievals of lake temperature in Northern Ireland) or possibly an error in the atmospheric database of absorption coefficients since Zavody (personal communication, 1991) has suggested an increase of 30% in the value for the water vapour continuum.

The small r.m.s. difference between the retrievals for the two variants of the new algorithm is encouraging since we might expect atmospheres over the Great Lakes to display considerable spatial variability as they try to establish thermodynamic equilibrium with the water, and the assumption of a constant atmosphere is the key to the success of the technique.

It will be useful to carry out further validation work on the new algorithm using *in situ* radiometric skin temperature measurements as the ground truth for comparison with the algorithm retrievals in order to assess the true precision attainable with real AVHRR data, together with radiosonde data to clarify the water/atmospheric temperature difference question. For operational use, satellite retrievals of surface temperature using AVHRR data may now be limited primarily by the accuracy of skin effect models and calibration error. Additional validation using ATSR data will be of interest since the latter error is expected to be reduced below  $\pm 0.1$  degrees. Finally, we would expect the benefits of the new algorithm to be more noticeable in tropical conditions where atmospheric transmission is low and consequently the primary 'split-window' assumption is most called into question, thus the new algorithm shows great promise, particularly with respect to its application to the ten year archive of AVHRR/2 data for global change studies.

## Chapter 7

### Concluding Remarks and Recommendations

The advantages remote sensing offers with respect to the necessary task of global monitoring of climate have already been stated in the preceding chapters. Thus the aims of the work presented in this thesis, as described in Chapter 1, are to develop new and improved remote sensing techniques for monitoring the earth's climate, either directly or by use of proxy indicators.

#### 7.1 Summary of work presented

Before the research into new techniques was presented, a review of various aspects of the climate system and the physical processes that operate within it was undertaken, with the primary purpose of gaining an understanding of the nature of potential problems involved in climate monitoring. The merits of remote sensing for climate research were described, together with some potential areas of exploitation with respect to global monitoring. In addition, a review of infrared radiometry from space was undertaken to provide the necessary background of understanding required to enable the development of new techniques for the exploitation of remotely-sensed data. Some of the physical processes that contribute to the signal measured by the instrument were reviewed, including atmospheric absorption processes and radiative transfer. The nature of instrumentation was also studied in order to gain understanding of the relative merits of various systems and assist in proper exploitation of the data. Current techniques employed for overcoming the effects of the atmosphere on remotely-sensed brightness temperatures were also studied in order to provide the necessary background for further development.

The first of the research goals pursued in the thesis was the development of techniques for measurement of lake areas for climate research. Having identified lake volume cycles as potential proxy indicators of climatic change (Mason *et al*, 1985), the task was to develop sufficiently accurate measurement techniques for lakes of areas down to  $\sim 100 \text{ km}^2$  using the coarse resolution data provided by instruments with sufficiently frequent global coverage such as AVHRR. A simple histogram-based lake area measurement technique was developed and a case study of an open lake was performed as a validation exercise. A simulation study demonstrated that the technique has an intrinsic accuracy of 0.5% for a lake of area  $390 \text{ km}^2$ , and the

validation study using AVHRR data realised an accuracy of 1% which is within the requirement for climate research (Mason *et al*, 1991). After this work was completed, a new edge-finding technique (Palmer and Forte, 1990) was brought to my attention, and a repeat of the earlier simulation study indicated a significant improvement (intrinsic error 0.2%) on the histogram technique.

The lake area research continued with a case study of a closed lake in a semi-arid region. The purpose of this study was to assess the impact of potential problems with measuring the areas of lakes in semi-arid regions such as the existence of salt flats and peripheral vegetation around the shore, and to decide whether climate signals could be detected from a time series of area measurements. Lake Abiyata in Ethiopia was chosen because it is small ( $\sim 130 \text{ km}^2$ ) and consequently at the lower end of lake sizes which we wish to measure (Harris and Mason, 1989). In addition, the lake is known to exhibit variability on short timescales (Makin *et al*, 1976) and AVHRR LAC coverage from mid-1985 is available from the NOAA archive. Simulations of AVHRR passes, using degraded high resolution data, were undertaken to decide on the most appropriate area measurement technique to employ for the time series study; the LIC method was found to be the most accurate with less than 1% rms spread and a small bias of  $\sim 1\%$ . The technique was then used to measure the area of Lake Abiyata from 52 AVHRR scenes taken between mid-1985 to the beginning of 1991. The results show a clear trend, with the lake area falling from  $\sim 170 \text{ km}^2$  to  $\sim 110 \text{ km}^2$  in 5 years. This seems to demonstrate a qualitative link with the climate over the period, although precise interpretation of such data is difficult (Mason *et al*, 1991) and beyond the scope of this thesis.

Research continued with an investigation into the problem of satellite remote sensing of water surface temperatures. Accurate measurement of SST from satellites is considered to be an important goal for climate prediction (eg. Woods, 1984), particularly in the tropics where the major quantity of solar radiation is absorbed (see Chapter 1), but this is where the atmospheric correction is high and conventional algorithms experience problems (eg. Barton, 1983). Also, lake temperature cycles are known to exhibit features which cannot be distinguished in the equivalent air temperature records (George, 1989), but conventional remote sensing techniques often experience problems (Irbe, 1990). This resulted in the development of a new algorithm for atmospheric correction. Simulations, using a standard set of ocean radiosondes and an atmospheric transmission model, indicated that the new algorithm offered a factor  $\sim 2$  improvement in intrinsic accuracy over the standard 'split-window' technique, and thus is comparable to the along-track scanning technique of ATSR.

By-products of the algorithm are an accurate measure of atmospheric transmittance and improved total column water vapour retrieval.

Validation of the atmospheric correction algorithm was performed on a dataset of AVHRR images and corresponding *in situ* buoy measurements of the Great Lakes. The nature of the Great Lakes meant that the validation was likely to be a tough test of the algorithm (see Chapter 6). The results realised a significant reduction in spread for the new algorithm with respect to the conventional 'split-window' method, and hence confirmed the potential of the technique. Some biases and outliers were found in the results for all methods, and possible mechanisms for these were discussed.

## **7.2 Assessment of contributions with respect to original aims**

The results obtained from the analysis presented in this thesis are encouraging in themselves, but it is important to relate them back to the original goals outlined in Chapter 1. In summary:

- The development and application of techniques for lake area measurement with coarse resolution data can be seen as a valuable achievement with respect to the goal of the remote sensing of proxy indicators of climate and climate change. The results for the closed lake case study show, albeit qualitatively, a link between lake area change and climate.
- The new atmospheric correction technique developed demonstrates significant improvements over the conventional 'split-window' method in both the simulations and validation study. SST retrieval, particularly in the climatically-important tropical regions, is likely to benefit from use of this new algorithm. In addition, the validation study performed on the Great Lakes is an indication of the potential of the technique in the remote sensing of lake temperatures.
- Potentially useful by-products of the atmospheric correction algorithm are the accurate measurement of atmospheric transmittance in the 'split-window' channels and the improved measurement of total water vapour content at low to middle values. The latter parameter is already identified as being a useful one for climate research (Woods, 1984).

## **7.3 Recommendations for further study**

The results obtained in Chapter 4 show that the LIC edge detection method has been shown to be suitable for accurate area measurement of a small ( $\sim 130 \text{ km}^2$ ) closed lake

in a semi-arid region, coping well with the problems that this entails, and further, that the use of lake areas as proxy indicators of climate change is worthwhile. The simulation results in Chapter 5 show that the new atmospheric correction technique promises significant improvements for a wide range of ocean atmospheres, and the good results obtained in the validation study (Chapter 6), as well as demonstrating improvement over existing techniques, imply that its application could successfully be extended to the measurement of lake surface temperature. These promising results indicate that further exploitation of the techniques is required, with a view to the processing of both future and existing global datasets. The following work is therefore recommended:

- Use of the level - area relation for Lake Abiyata to obtain lake volume changes, and subsequent analysis to provide a more quantitative description of the response with respect to climate.
- Identification of the other closed lake regions (eg. the Tibetan Plateau) for which suitable archives of remotely-sensed data exist in order to provide a more global record of this useful proxy indicator of climate. In addition, for lakes where no level - area relation exists, attempt to combine the imagery with satellite altimeter height measurements to obtain volume changes.
- Investigation of the imaging characteristics of the ATSR and its suitability for monitoring lake areas with a view to exploiting the continuous global coverage provided by ERS-1, and the forthcoming ERS-2 and Polar Platforms.
- Further validation studies for the atmospheric correction algorithm, particularly over the ocean, and, if possible, in tropical regions where the atmospheric corrections are high and, as a consequence, the benefits of the algorithm are expected to be greater. More comprehensive *in situ* datasets would be preferred, including measurement of the ocean skin temperature and radiosonde data to answer questions about the atmospheric transmission and water vapour retrievals, and to check the results of radiative transfer models. Use of ATSR data for the nadir view could also assist by minimising calibration error.
- Validation studies using the algorithm over small lakes ( $\sim 100 \text{ km}^2$ ) will be needed to ascertain the suitability of the technique for lake ST retrieval. The main question will be whether  $\Delta T_5/\Delta T_4$  can be measured to sufficient accuracy with a



relatively small number of pixels. The potential for using the technique over land should also be investigated; the main question here being the impact of the emissivities of different cover types.

- Investigation into refinement of the atmospheric correction algorithm, together with an assessment of the additional information provided by the along-track view of ATSR. Improvements to the water vapour retrieval may also be possible by taking into account the temperature dependence of the absorption coefficients.
- Use of the measure of atmospheric transmission in radiation budget studies. NOAA GAC data could be used to calculate the global atmospheric transmission on a daily basis, although problems with data volume, and the effects of cloud and land would need investigation.

The developments outlined above, if successful, should provide new and better information on several aspects of the global climate system. The data provided by the increasing number of new and improved instruments that are being launched, together with the existing data archives, should benefit from improved analysis techniques such as those described above, thus providing a more complete picture of the Earth's climate.

## References

- Anding, D., and Kauth, R., 1970, Estimation of sea surface temperature from space. *Rem. Sens. Environ.*, **1**, 217 -220.
- Barker, J.L., 1975, Monitoring Water Quality from Landsat. *NASA Technical Memorandum* TM X-71006, NASA, GSFC, Maryland, USA.
- Barnola, J.M., Raynaud, D, Korotkevich, Y.S., and Lorius, C., 1987, Vostok ice core provides a 160,000 year record of atmospheric CO<sub>2</sub>, *Nature*, **329**, 408 - 414.
- Barton, I.J., 1983, Dual channel satellite measurements of sea surface temperature. *Quart. J. Roy. Met. Soc.*, **109**, 365 - 378.
- Becker, F., 1987, The impact of spectral emissivity on the measurement of land surface temperature from satellite. *Int. J. Rem. Sens.*, **8**, 1509 - 1522.
- Berger, A.L., 1980, Astronomical theory of paleoclimates, in 'Climate variations and variability: Facts and theories', (ed. A.Berger), 501 - 526. Reidel, Holland.
- Berger, A.L., 1980, Spectrum of Climatic Variations and Possible Causes, in 'Climate variations and variability: Facts and theories', (ed. A.Berger), 411 - 432. Reidel, Holland.
- Brown O.B., Brown, J.W., and Evans, R.H., 1985, Calibration of Advanced Very High Resolution Radiometer Infrared Observations. *J. Geophys. Res.*, **90**, C6, 11,667 - 11,677.
- Brown, S.J., Mason, I.M., and Harris, A.R., 1991, Lake Surface Temperatures from Space for Climate Research. Proceedings of 5th AVHRR data Users Meeting, Tromsø, 24<sup>th</sup> - 28<sup>th</sup> June 1991, in press.
- Budyko, M.I., Ronov, A.B., Yanshin, A.L., 1985, History of the Earth's atmosphere, Gidrometeoizdat, Leningrad, 208 p (R). (English translation: Springer Verlag, 1987)

- Budyko, M.I., Golitsyn, G.S, and Izrael, Y.A., 1988, Global Climate Catastrophes. Springer Verlag, New York.
- Chen, H.S., 1985, Space Remote Sensing Systems - An Introduction. Academic Press, Inc., Orlando, Florida.
- Cusbasch, U., and Cess, R.D., 1990, Processes and modelling, in 'Climate change - the IPCC scientific assessment', 69 - 91, (ed. Houghton, J.T., Jenkins, G.J., and Ephraums, J.J.), Cambridge University Press.
- Deschamps, P.Y., and Phulpin, T., 1980, Atmospheric correction of infrared measurements of sea surface temperature using channels at 3.7, 11, and 12  $\mu\text{m}$ . *Boundary Layer Met.*, **18**, 131 -143.
- Dudhia, A., 1989, Noise characteristics of the AVHRR channels. *Int. J. Rem. Sens.*, **10**, 637 - 644.
- Ford, J.P., Cimino, J.B., and Elachi, C., 1983, Space Shuttle Columbia Views the World with Imaging Radar: the SIR-A Experiment, JPL publication 82 - 95.
- George, G.D., 1989, The Thermal Characteristics of Lakes as a Measure of Climate Change. Proceedings of the conference on 'Climate and Water', Helsinki, Vol.1, p 402 - 412. The Academy of Finland.
- Goody, R.M., and Walker, J.C.G., 1972, Atmospheres, Prentice Hall (Eaglewood Cliffs, New Jersey).
- Harris, A.R., and Mason, I.M., 1989, Lake Area Measurement Using AVHRR - A Case Study. *Int. J. Rem. Sens.*, **10**, 885 - 895.
- Harris, A.R., and Mason, I.M., 1990, An Extension to the 'Split-Window' Technique giving Improved Atmospheric Correction and Total Water Vapour. *Int. J. Rem. Sens.* (in press).
- Hepplewhite, C.L., 1989, Remote observations of the sea surface and atmosphere. The oceanic skin effect. *Int. J. Rem. Sens.*, **10**, 801 - 810.
- Houghton, J.T., 1987, The Physics of Atmospheres, Cambridge University Press.

- Houghton, J.T., and Morel, P., 1984, The World Climate Research Programme, in 'The Global Climate', (ed. J.T.Houghton), 1 - 12. Cambridge University Press.
- Hutchinson, G., 1975, A Treatise on Limnology. Vol. 1 part 1. Wiley.
- Intergovernmental Panel on Climate Change (IPCC), 1990, Policymakers Summary, in 'Climate change - the IPCC scientific assessment', vii - xxxix, (ed. Houghton, J.T., Jenkins, G.J., and Ephraums, J.J.), Cambridge University Press.
- ITT, 1989, AVHRR/2. Publicity material.
- Kukla, G.J., 1980, Pleistocene Climate on Land, in 'Climate variations and variability: Facts and theories', (ed. A.Berger), 207 - 232. Reidel, Holland.
- Lamb, H.H., 1972, Climate: Present, Past and Future, Volume One. Methuen.
- Lamb, H.H., 1982, Climate History and the Modern World, Methuen.
- Lauritson, L., Nelson, G.J., and Porto, F.W., 1979, Data extraction and calibration of TIROS-N/NOAA radiometers. *NOAA tech memo* TM NESS 107.
- Laxon, S.W.C., 1990, Satellite Radar Altimetry over Sea Ice. PhD Thesis, Univ. London.
- Leith, C.E., 1984, Global Climate Research, in 'The Global Climate', (ed. J.T.Houghton), 13 - 24. Cambridge University Press.
- Llewellyn-Jones, D.T., Minnett, P.J., Saunders, R.W., and Zavody, A.M., 1984, Satellite Multichannel Infrared measurements of sea surface temperature of the N.E. Atlantic Ocean using AVHRR/2. *Quart. J. Roy. Met. Soc.* **110**, 613 - 631.
- Makin, M.J., Kingham, T.J., Waddams, A.E., Birchall, C.J., and Eavis, B.W., 1976, Prospect for irrigation development around Lake Zwai, Ethiopia. Land Resources Study, Land Resources Division, UK Ministry of Overseas Development. Tolworth 26: 270pp.

- Malila, W., and Nalepka, R., 1975, Advanced processing and information extraction techniques applied to ERTS-1 MSS data. pre-print, Environmental Research Institute Michigan. Ann Arbor, Michigan.
- Manabe, S., and Stouffer, R.J., 1979, A CO<sub>2</sub> Climate Sensitivity Study with a Mathematical Model of Global Climate, *Nature*, **202**, 491 - 493.
- Manschke, A., 1989, Remote Sensing of Cloud Base Temperatures in Convective Situations, pp 97 - 100, Proceedings of 4th AVHRR data Users Conference, EUM P06 ISSN 1015 - 9576, Eumetsat.
- Mason, I.M., Rapley, C.G., Street-Perrott, F.A., and Guskowska, M.A.J., 1985, ERS-1 Observations for Climate Research. ESA-SP-233 (Paris: European Space Agency), 235 - 241.
- Mason, I.M., Sheather, P., and Bowles, J., 1990, On-board calibration of infrared radiometers. Proceedings of Royal Aeronautical Society conference on 'Earth Observation Satellites - The Technology behind the Image'. RAS, January 1990.
- Mason, I.M., Guskowska, M., Rapley, C.G., and Street-Perrott, F.A., 1991, The Response of Lake Levels and Areas to Climatic Change. In preparation.
- McClain, E.P., 1981, Multiple atmospheric-window techniques for satellite-derived sea surface temperatures. In *Oceanography from Space*, edited by J.F.R.Gower (New York: Plenum Press) pp 73 - 85.
- McClain, E.P., Pichel, W., Walton, C., Ahmad, Z., and Sutton, J., 1983, Multichannel Improvements to Satellite-derived Global Sea Surface Temperatures. *Adv. Space Res.*, **2**, 43 - 47.
- McClain, E.P., 1989, Global Sea Surface Temperatures and Cloud Clearing for Aerosol Optical Depth Estimates. *Int. J. Rem. Sens.*, **10**, 763 - 770.
- McMillin, L.M., 1975, Estimation of sea surface temperature from two infrared window measurements with differential absorption. *J. Geophys. Res.*, **80**, 5,113 - 5,117.

- Palmer, D., and Forte, P., 1990, Image Segmentation using Local Isoluminance Contours. In *Communication, Control, and Signal Processing*, ed. E. Arıkan, Elsevier Science Publishers, B.V. 1386 - 1390.
- Partington, K.C., 1988, Studies of the Ice Sheets by Satellite Altimetry. PhD Thesis, Univ. London.
- Prabhakara, C., Dalu, G, and Kunde, V.G., 1974, Estimation of sea surface temperature from remote sensing in the 11 - 13  $\mu\text{m}$  window region. *J. Geophys. Res.*, **79**, 5,039 - 5,044.
- Prata, A.J., 1991, Land surface temperature measurements from the AVHRR. Proceedings of the 5th AVHRR data Users Meeting, Tromsø, 24<sup>th</sup> - 28<sup>th</sup> June, 1991. To be published.
- Rapley, C.G., Griffiths, H.D., Squire, V.A., Olliver, J.G., Birks, A.R., Cooper, A.P.R., Cowan, A.M., Drewry, D.J., Gorman, M.R., Guskowska, M., Laxon, S., Mason, I.M., McIntyre, N.F., Novotny, E., Paterson, R., Scott, R.F., Street-Perrott, F.A., 1985, Applications and Scientific Uses of ERS-1 Radar Altimeter Data. ESA report 5684/83/NL/BI.
- Rapley, C.G., Guskowska, M., Cudlip, W., and Mason, I.M., 1987, An Exploratory Study of Inland Water and Land Altimetry using Seasat data. ESA contract report 6483/85/NL/BI (European Space Agency: Paris).
- Robinson, I.S., Wells, N.C., and Charnock, H., 1984, The Sea Surface Thermal Boundary Layer and its Relevance to the Measurement of Sea Surface Temperature by air-borne and space-borne Radiometers. *Int. J. Rem. Sens.*, **5**, 19 - 45.
- Ropelewski, C.F., and Halpert, M.S., 1989, Precipitation Patterns associated with the High Index phase of the Southern Oscillation. *J.Climat.*, **2**, 268 - 284.
- Saunders, R.W., and Kriebel, K.T., 1988, An Improved Method for detecting Clear Sky and Cloudy radiances from AVHRR data. *Int. J. Rem. Sens.*, **9**, 123 -150.

- Schott, J.R., and Henderson-Sellers, A., 1984, Radiation, the Atmosphere and Satellite Sensors. In 'Satellite Sensing of a Cloudy Atmosphere' (ed. A. Henderson-Sellers), Taylor and Francis.
- Schuermans, C., 1980, Solar Activity and Climate, in 'Climate variations and variability: Facts and theories', (ed. A.Berger), 559 - 576. Reidel, Holland.
- Second World Climate Conference (SWCC), 1990, Conference Statement from the Scientific and Technical Sessions, S&T/SWCC/Nº 13.
- Shuter, B.J., Schlesinger, D.A., and Zimmerman, A.P., 1983, Empirical predictors of annual surface water temperature cycles in North American lakes. *Canadian J. Fish. Aquat. Sci.*, **40**, 1838 - 1845.
- Schluessel, P., 1989, Satellite-derived low level atmospheric water vapour content from synergy of AVHRR with HIRS. *Int. J. Rem. Sens.*, **10**, 705 - 722.
- Schluessel, P., Emery, W.J., Grassl, H., and Mammen, T., 1990, On the bulk - skin temperature different and its impact on satellite remote sensing of sea surface temperature. *J. Geophys. Res. (Oceans and Atm.)*, **95**, 13,341 - 13,356.
- Street-Perrott, F.A., and Harrison, S.P., 1985, Lake levels and Climate reconstruction. In *Paleoclimate Analysis and Modelling*, ed. A.D.Hecht (New York: Wiley) pp 291 - 340.
- Thomas, R.H., Sanderson, T.J.D., and Rose, K.E., 1979, Effects of a Climatic Warming on the West Antarctic Ice Sheet, *Nature*, **227**, 355 - 358.
- Tramoni, F., Barry, R.G., and Key, J., 1985, Lake ice cover as a temperature index for monitoring climatic perturbations. *Zeitschrift Für Gletscherkunde und Glazialgologie*, **21**, 43 - 49.
- Untersteiner, N., 1984, The cryosphere, in 'The Global Climate', (ed. J.T.Houghton), 121 - 140. Cambridge University Press.
- Warrick, R., and Oerlemans, J., 1990, Sea Level Rise, in 'Climate change - the IPCC scientific assessment', 257 - 281, (ed. Houghton, J.T., Jenkins, G.J., and Ephraums, J.J.), Cambridge University Press.

- Watson, R.T., Rodhe, H., Oeschger, H., and Siegenthaler, U., 1990, Greenhouse Gases and Aerosols, in 'Climate change - the IPCC scientific assessment', 1 - 40, (ed. Houghton, J.T., Jenkins, G.J., and Ephraums, J.J.), Cambridge University Press.
- Webster, P.J., and Stephens, G.L., 1984, Cloud - Radiation Interaction and the Climate Problem, in 'The Global Climate', (ed. J.T.Houghton), 63 - 78. Cambridge University Press.
- Woods, J.D., 1984, The upper ocean and air - sea interaction in global climate. In 'The Global Climate', (ed. J.T.Houghton), 141 - 178. Cambridge University Press.
- World Climate Data Programme (WCDP), 1990, The Global Climate System - Climate System Monitoring June 1986 - November 1988, WMO World Climate Data Programme, United Nations Environment Programme, CSM R 84/86.
- World Climate Research Programme (WCRP), 1986, Satellite-derived Sea Surface Temperatures for Global Climate Applications. WCRP Publications Service, WMO/TD N° 93.
- WCRP, 1990, 'Global Climate Change', a scientific review presented by the WCRP (text by P. Morel).
- Wunsch, C., 1984, The Ocean Circulation in Climate, in 'The Global Climate', (ed. J.T.Houghton), 189 - 203. Cambridge University Press.
- Zavody, A.M., 1982, Annex J of 'The Along-Track Scanning Radiometer with Microwave Sounder', a response to the ERS-1 Announcement of Opportunity. RAL unpublished manuscript.



## Appendix A

### A Simple Numerical Model of an Infinite Coastline

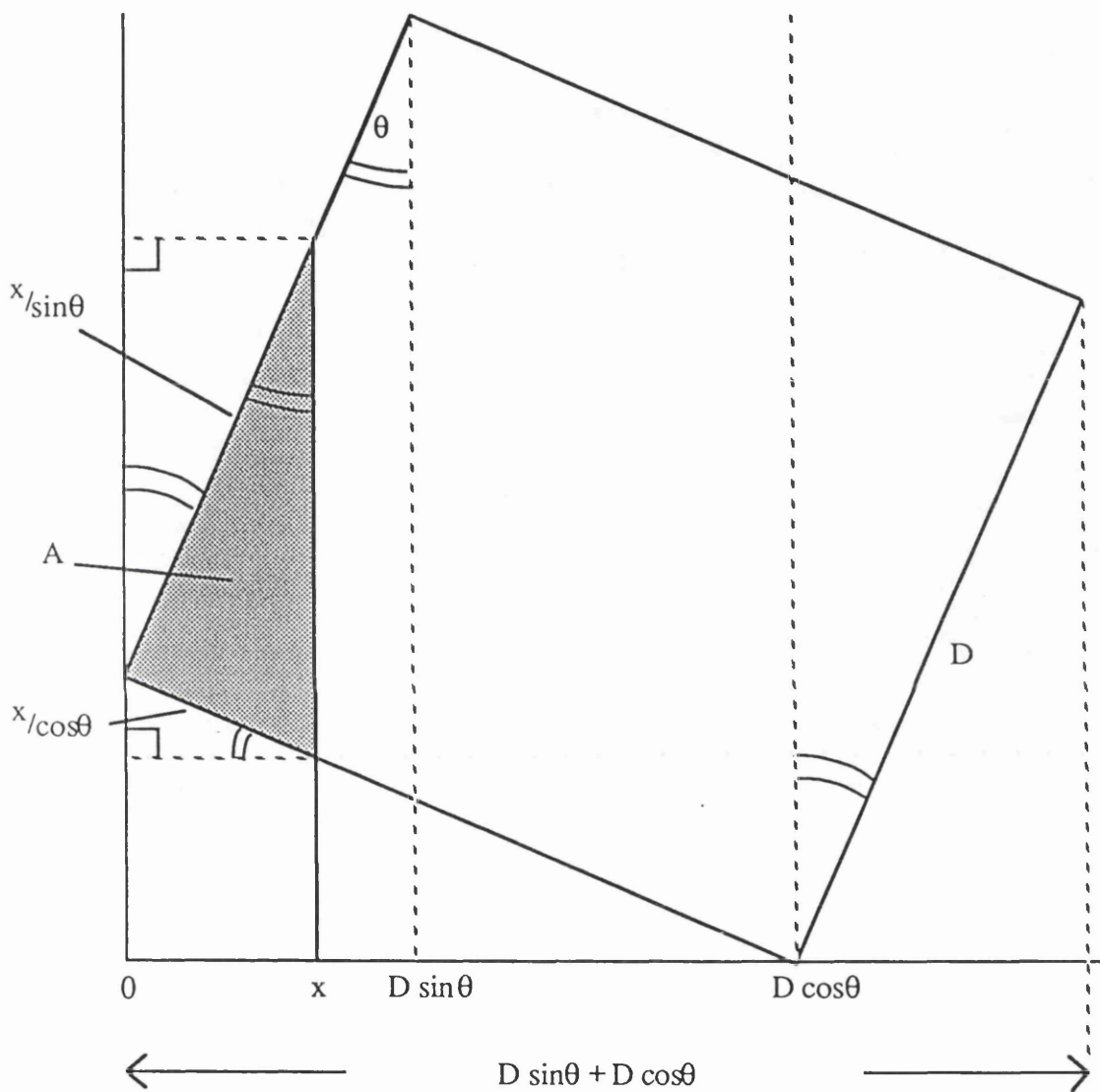
As a first stage in a more fundamental study of the problem of edge pixels, it was decided to produce a simple numerical model of an infinite coastline in order to obtain an approximation to the shape of the edge pixel distribution. The assumptions made for this model are :

- i) The pixel is regarded as a square with side of length  $D$ . This is the simplest case and is a reasonable approximation for most satellite images.
- ii) As a first approximation to the shape of the coastline, the simplest case has been chosen, namely a straight line across the pixel.
- iii) Any orientation of the pixel with respect to the coastline is regarded as equally likely.
- iv) Any position of the pixel with respect to the coastline is regarded as equally likely.
- v) The model makes no distinction between land and water, thus the model will be symmetrical about the mid point by its very nature.

Points (iii) and (iv) mean that some orientations will have a higher probability than others, and a weighting factor will have to be applied in order to obtain the final distribution.

We consider the general case, with the pixel turned at an arbitrary angle  $\theta$  to the coastline (see Figure A.1).

The range of values of  $x$  between  $D\cos\theta$  and  $D(\sin\theta + \cos\theta)$  is symmetrical with respect to the range  $0 \leq x \leq D\sin\theta$ .



**Figure A.1** This diagram shows the coastline drawn at an arbitrary angle  $\theta$  to the pixel sides. The shaded portion represents the fractional area  $A$  of the pixel covered by water when the coastline is at an arbitrary position  $x$  along the x-axis.

$$A = \frac{x^2}{2 \cdot \sin\theta \cos\theta} = \frac{x^2}{\sin 2\theta} \quad (\text{A.1})$$

When  $x = D \sin\theta$ ,  $A = A_{\text{crit}}$ , beyond this point  $A$  increases uniformly with  $x$  until  $x = D \cos\theta$  where the probability distribution reverts to a form that is symmetrical to the first part ( $0 \leq x \leq D \sin\theta$ ).

$$A_{\text{crit}} = \frac{D^2 \sin^2 \theta}{2 \cdot \sin\theta \cos\theta} = \frac{D^2 \tan\theta}{2} \quad (\text{A.2})$$

$$\text{Also, } \frac{\frac{D^2 \tan\theta}{2}}{\int_0^{\frac{D \sin\theta}{2}} P(A) dA} = \int_0^{\frac{D \sin\theta}{2}} P(x) dx = \frac{\sin\theta}{\sin\theta + \cos\theta} \quad (\text{A.3})$$

since only this fraction of the total range of  $x$  is being considered  
ie.  $0 \leq P(x) \leq (\sin\theta)/(\sin\theta + \cos\theta)$ .

$$\text{and } \int_0^A P(A) dA = \int_0^x P(x) dx \quad (\text{A.4})$$

and we assume that  $P(x) = C$ .

$$\therefore \int_0^{D \sin\theta} C dx = C \cdot D \sin\theta = \frac{\sin\theta}{\sin\theta + \cos\theta}$$

$$\therefore C = \frac{1}{D(\sin\theta + \cos\theta)} \quad (\text{A.5})$$

$$\therefore P(x) = \frac{1}{D(\sin\theta + \cos\theta)}$$

$$\therefore \int_0^x P(x) dx = \int_0^x \frac{1}{D(\sin\theta + \cos\theta)} dx = \frac{x}{D(\sin\theta + \cos\theta)} \quad (\text{A.6})$$

From (A.4) & (A.6) we have

$$\int_0^A P(A) dA = \frac{x}{D(\sin\theta + \cos\theta)}$$

From (A.1)  $x = \sqrt{A} \sqrt{\sin 2\theta}$

$$\therefore \int_0^A P(A) dA = \frac{\sqrt{A} \sqrt{\sin 2\theta}}{D(\sin\theta + \cos\theta)}$$

$$\therefore P(A) = \frac{\sqrt{\sin 2\theta}}{2D(\sin\theta + \cos\theta)} \cdot A^{-\frac{1}{2}} \quad (A.7)$$

Let us check for normalisation :

$$\begin{aligned} \frac{D^2 \tan\theta}{2} \int_0^2 P(A) dA &= \frac{\sqrt{(D^2 \tan\theta)/2} \sqrt{2 \sin\theta \cos\theta}}{D(\sin\theta + \cos\theta)} \\ &= \frac{D}{D} \cdot \frac{\sqrt{\sin\theta}}{\sqrt{\cos\theta}} \cdot \frac{\sqrt{2}}{\sqrt{2}} \cdot \frac{\sqrt{\sin\theta} \cdot \sqrt{\cos\theta}}{(\sin\theta + \cos\theta)} \\ &= \frac{\sin\theta}{\sin\theta + \cos\theta} \end{aligned}$$

which agrees with (A.3).

Since there are no discontinuities in the distribution, the (constant) probability associated with the middle part of the distribution, in the range  $(D^2 \tan\theta)/2 > A > D^2(1 - (\tan\theta)/2)$ , is the same as the probability for  $A = (D^2 \tan\theta)/2$ .

Therefore :

$$\begin{aligned}
 \text{For } 0 \leq A \leq \frac{D^2 \tan \theta}{2} \quad P(A) &= \frac{\sqrt{\sin 2\theta}}{\sin \theta + \cos \theta} \cdot \frac{A^{-1/2}}{2D} \\
 \text{For } \frac{D^2 \tan \theta}{2} < A < \frac{D^2 \cot \theta}{2} \quad P(A) &= \frac{\sqrt{\sin 2\theta}}{\sin \theta + \cos \theta} \cdot \frac{[(D^2 \tan \theta)/2]^{-1/2}}{2D} \\
 \text{For } \frac{D^2 \cot \theta}{2} \leq A \leq D^2 \quad P(A) &= \frac{\sqrt{\sin 2\theta}}{\sin \theta + \cos \theta} \cdot \frac{(D^2 - A)^{-1/2}}{2D}
 \end{aligned}$$

The next step is to select a finite set of angles evenly distributed over the range  $0 \leq \theta \leq \pi/4$  and input them into take these general solutions, summing the distributions over the range. When summing the distributions over the range  $0 \leq \theta \leq \pi/4$  it is necessary to apply a weighting factor, since the x-dimension of a pixel determines the relative probability of that pixel orientation being selected in the model. Therefore the total probability is given by

$$P_T(A) = \sum_{\theta=0}^{\pi/4} [P(A)] w$$

where  $w = \sin \theta + \cos \theta$ , the x-dimension of the pixel divided by D.  $P_T(A)$  is the total summed probability for any area A. This relationship was evaluated numerically, and the results are shown in Figure 3.8.

## Appendix B

### Geometric Correction of AVHRR Pixel Areas

The factors to be taken into account in order to convert AVHRR pixel areas to physical measurements are as follows:

- There will be a  $\cos\theta$  projection of the lake area being imaged by the satellite due to the angle  $\theta$  between the viewing direction of the satellite and the normal to the lake surface.
- There are two components to this angular projection. Firstly, there is an angle between the viewing direction and the nadir direction of the satellite. Secondly, earth curvature will result in the plane of the lake being turned by a further angle relative to the satellite nadir direction which must also be taken into account when applying the total angular correction to the projected area.
- The distance between the sensor and the lake will determine the relationship between the angle that the mirror sweeps through per pixel and the corresponding across-track dimension the pixel represents on the ground.
- Since the attitude control of the NOAA satellite ensures that pitch is continuously maintained around the orbit (Barbière, personal communication, 1991), successive scan lines can be considered to be arcs of great circles perpendicular to the sub-satellite track and thus get closer together on the Earth's surface further away from nadir. This situation is analogous to longitude lines being closer together further from the equator.
- The across-track component of earth rotation will have no effect on the measured area since the result will simply be to skew the image. The along-track component will, however, result in an increase in the effective area since the distance between successive scans will be larger (for a retrograde orbit) than would be the case for a non-rotating earth.

Some data on various instrument parameters, together with the satellite altitude (obtained either from Brouwer mean elements or the TIP header in the AVHRR

HRPT data stream), are required. The following data can be found in the NOAA Polar Orbiter Data Users Guide.

- i) The scan rate of the telescope mirror is 6 revolutions per second.
- ii) The pixel sample rate is 39,936 samples per second.
- iii) The number of samples allocated to earth view per scan is 2048, centred at nadir (assuming that the satellite has zero attitude roll).

From (i) and (ii) we can calculate the angular size of each pixel in the across-track direction. Each sample (pixel) represents an angle  $\eta$ , where

$$\eta = \frac{6 \times 2\pi}{39936}$$

$$= 9.43988 \times 10^{-4} \text{ radians.}$$

Next, the number of pixels,  $n$ , from the centre of the scan to the centre of the lake was counted for each pass. The angular distance between the nadir direction of the satellite and the centre of the target is simply  $n \times \eta$ .

The angular distance  $\beta$  (with respect to the centre of the earth) between the sub-satellite point (on the surface of the earth) and the target was calculated as follows :

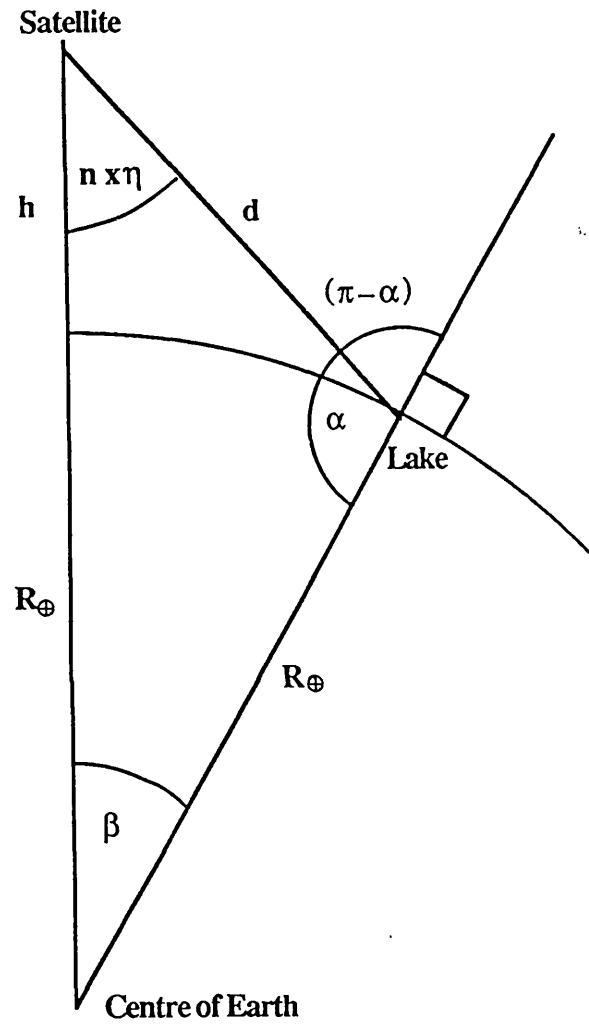
Referring to Figure B.1 and applying the sine rule, we have

$$\frac{R_{\oplus}}{\sin(n \times \eta)} = \frac{R_{\oplus} + h}{\sin \alpha}$$

$$\text{Hence} \quad \alpha = \sin^{-1} \left[ (R_{\oplus} + h) \sin(n \times \eta) / R_{\oplus} \right] \quad (\text{B.1})$$

Also, since the internal angles of a triangle sum to  $\pi$  radians,

$$\beta = \pi - \sin^{-1} \left[ (R_{\oplus} + h) \sin(n \times \eta) / R_{\oplus} \right] - (n \times \eta) \quad (\text{B.2})$$



**Figure B.1** Diagram showing the trigonometry involved in calculating the angle subtended at the centre of the earth.



But since  $\sin \alpha = \sin(\pi - \alpha)$ , (B.1) will give the first quadrant angle and (B.2) can be rewritten as

$$\beta = \sin^{-1}[(R_{\oplus} + h)\sin(n\pi)/R_{\oplus}] - (n\pi) \quad (B.3)$$

The across-track dimension of the pixel can now be obtained by repeating the process with  $n$  increased by one. The across-track dimension is simply obtained:

$$X = |\beta(n+1) - \beta(n)| R_{\oplus} \quad (B.4)$$

The along track dimension of the pixel is determined primarily by the scan rate of the AVHRR instrument and the velocity of the sub-satellite point on the surface of the earth. The orbital velocity of the satellite is determined by the orbital height. Since the eccentricity of NOAA satellites is extremely low (from the Brouwer mean elements), the orbit can be regarded as locally approximately circular. Newton's second law when applied to a circular orbit gives :

$$\frac{GMm}{(R_{\oplus} + h)^2} = \frac{m v^2}{(R_{\oplus} + h)} \quad (B.4)$$

On the earth's surface,  $g = (GM) / R_{\oplus}^2$ , ie.  $GM = gR_{\oplus}^2$ . Therefore equation (B.4) becomes

$$\begin{aligned} v^2 &= \frac{gR_{\oplus}^2}{(R_{\oplus} + h)} \\ \therefore v &= \left[ \frac{gR_{\oplus}^2}{(R_{\oplus} + h)} \right]^{1/2} \end{aligned} \quad (B.5)$$

Since we require the velocity of the sub-satellite point on the earth's surface, the velocity obtained in (B.5) must be multiplied by a factor of  $R_{\oplus}/(R_{\oplus} + h)$ ,

$$\therefore v' = \frac{R_{\oplus}}{(R_{\oplus} + h)} \left[ \frac{gR_{\oplus}^2}{(R_{\oplus} + h)} \right]^{1/2} \quad (B.6)$$

To get the along track dimension,  $Y$ , of the pixel in km we must divide the velocity of the sub-satellite point,  $v'$ , by the scan rate;

$$\therefore Y(\text{km}) = v'/6 \quad (B.7)$$

However, there are two further corrections required for the along track dimension of the pixel, namely those of 'barrel' distortion and earth rotation. The correction for 'barrel' distortion (see Chapter 6) is easily calculated from knowledge of the angle  $\beta$ :

$$Y' = Y \cos \beta \quad (\text{B.7})$$

and the correction for earth rotation can be obtained by first calculating the angle,  $F$ , between the sub-satellite track and the meridian that passes thro' the image centre when the lake is imaged. If  $i$  is the latitude of the sub-satellite point when the lake is imaged and  $I$  the (co)inclination of the orbit, then, from Figure B.2 and using spherical trigonometry:

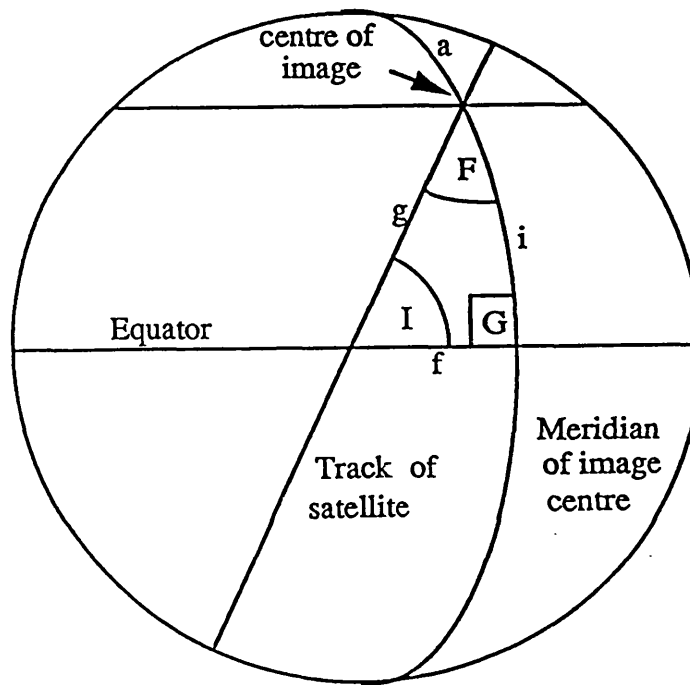
$$\text{sine rule} \quad g = \sin^{-1}[\sin i / \sin I] \quad \text{since } G = 90^\circ$$

$$\text{cosine rule} \quad f = \cos^{-1}[\cos g / \cos i] \quad "$$

$$\text{sine rule} \quad F = \sin^{-1}[\sin f / \sin g] \quad "$$

The along track component of earth rotation is then  $\sin F$ , with the magnitude of the rotation itself being reduced by a factor  $\cos i$  from that at the equator.

The area,  $A_p$ , of the pixel can be obtained by multiplying the along track and across track dimensions of the pixel.



**Figure B.2** Spherical triangle formed by the meridian passing through the sub-satellite point, the sub-satellite track and the equator.