THE INVESTIGATION OF NOCTILUCENT CLOUDS AND OTHER MESOSPHERIC PHENOMENA USING GROUND-BASED INSTRUMENTATION AND ROCKETS

SALLY ANNE EAST

DEPARTMENT OF PHYSICS AND ASTRONOMY

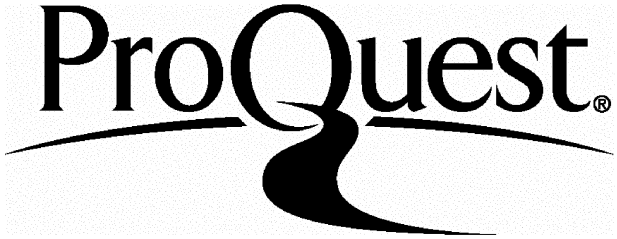
UNIVERSITY COLLEGE LONDON

THESIS PRESENTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

OF THE UNIVERSITY OF LONDON

May 1993



To My Beloved Mother,

Thank you for always

being there for me -

through all my struggles.

You're the greatest!

Abstract

The optical and dynamical properties of the summer phenomena known as Noctilucent Clouds (NLC) have been studied globally since the early 1960s. These clouds only occur naturally in the Earth's mesosphere, and are presently studied using remote sensing from rockets, and satellites in addition to ground-based observations.

Direct evidence of the topology, and structure of an aerosol layer, such as a NLC can be achieved using non-imaging photodiode/photometers housed on a rocket payload. The APL designed photometers utilize the spin of the rocket payload, therefore scanning the entire sky laterally, and producing a two-dimensional image of the aerosol layer traversed during the upleg and downleg flights. The APL photodiode/photometers were flown during the MAED (Middle Atmospheric ElectroDynamics) and NLC-91 (NoctiLucent Clouds 1991) summer rocket campaigns respectively. These multinational rocket campaigns were coordinated with ground-based and satellite observations. The resultant APL data provides a complementary data source to be compared and contrasted with other rocket experiments flown during the absence/presence of NLC and/or PMSE (Polar Mesospheric Summer Echoes).

A Bomem Michelson Interferometer (MI) was stationed in Sweden during the summer rocket campaign, NLC-91 provided measurements of the hydroxyl (3,1) band emission from a layer positioned ~87 km. The presented data were kindly provided by the Space Dynamics Laboratory, Utah State University; analyzed and interpreted by the author. The data gave a measure of the upper mesospheric conditions during the presence and absence of NLC during the rocket campaign. The interpretation of the raw data gave an indication of stratospheric filtering of upward propagating waves whose diminution could produce the upward forcing that may be involved in the NLC formation processes. Intensity and rotational temperature profiles deduced during the absence/ presence of NLC gave clear results of small- and large-scale waves, and possible correlation between the hydroxyl intensity, and mesopausal temperature.

A ground-based Imaging Fabry-Perot Interferometer (IFPI) stationed at the Bear Lake Observatory (BLO), Utah, (41.93°N, 111.42°W) has been operated, since 1989, to study the behaviour of the mesospheric winds at mid-latitudes deduced from the measurements of intensity, and wind speeds of the 8430 Å hydroxyl (6,2) band. The results presented have been studied during the summer periods from this site. This IFPI has provided an opportunity of observing this weak infra-red emission line, and provides a continuous monitoring of the mesopause region throughout the year. A comparison of characteristics on a night by night basis of hydroxyl intensity variations, and wind speed structure, and variations from the mid-latitude IFPI (BLO) with intensity fluctuations indicated by the Bomem MI as being associated with the absence/presence of NLC at summer high latitudes.

List of Acronyms

ALOHA-90	Airborne Lidar and Observations of the Hawaiian Airglow 1990
ALOMAR	Arctic Lidar Observatory for Middle Atmosphere Research
APL	Atmospheric Physics Laboratory
BLO	Bear Lake Observatory
BST	British Summer Time
CAMP	Cold Arctic Mesopause Project
CASS	Centre for Atmospheric and Space Science
CUPRI	Cornell University Portable Radar Interferometer
EISCAT	European Incoherent Scatter
FPI	Fabry-Perot Interferometer
GSFC	Goddard Space Flight Center
HRDI	High Resolution Doppler Imager
IGY	International Geophysical Year
IPD	Imaging Photon Detector
IRFWI	Infra-Red Field Widened Interferometer
LIDAR	Light Detection And Ranging
LT	Local Time
MA	Middle Atmosphere
MAED	Middle Atmosphere ElectroDynamics
MALT	Middle Atmosphere and Lower Thermosphere
MELTER	Mesosphere and Lower Thermosphere Explorer
MISU	Meteorology Institute of Stockholm University
MLTI	Mesosphere and the Lower Thermosphere and Ionosphere
NASA	National Aeronautics and Space Administration
NLC	Noctilucent Clouds
NLC-91	Noctilucent Clouds 1991

PC	Personal Computer
PCM	Pulse Code Modulation
PMC	Polar Mesospheric Clouds
PMSE	Polar Mesospheric Summer Echoes
RDES	ReaD ESrange
SDL	Space Dynamics Laboratory
SERC	Science and Engineering Research Council
SLIPS	Scattered Light Intensity Profile Sensor
SME	Solar Mesosphere Explorer
SOAP	Selective Optical Atmospheric Probe
SSC	Swedish Space Corporation
SPU	Signal Processing Unit
TAD	Thermal and Atmospheric Dynamics
TIMED	Thermosphere, Ionosphere and Mesosphere Energetics and Dynamics
TPU	Tape Processing Utility
UARS	Upper Atmosphere Research Satellite
UCL	University College London
UK	United Kingdom
USSR	Union of Soviet Socialist Republic
USU	Utah State University
UT	Universal Time
UVS	Ultraviolet Spectrometer Experiment
UWOMI	University of Western Ontario Michelson Interferometer
WFF	Wallops Flight Facility
WINDII	Wind Imaging Interferometer
WMO	World Meteorological Organization

List of Figures

Chapter 1

1.8 CUPRI data during the salvo A launch sequence of the NLC-91 Rocket Campaign. The top panel displays the post-processing signal-to-noise ratio versus altitude and time. The middle panel shows the radial velocity (positive is upwards and negative is downwards). The Doppler spectral width is displayed in the bottom panel.

Chapter 2

2.1A The final set-up of the instrumentation flown on the rocket payloads.

2.1B Schematic diagram of the APL photometer.

2.2 Silicon (Si) p-i-n photodiode with guard ring.

2.3 Typical spectral response plot.
Responsivity:0.2 A/W at 450 nm;0.35 A/W at 633 nm;
0.5 A/W at 900 nm; 0.15 A/W at 1064 nm; 7.9 mA/Im (2850 K source).

2.4 Absorption process; (a) intrinsic; (b) extrinsic; (c) free carrier.

2.5 Energy band structure of (a) GaAs; (b) Si; (c) Ge. Si is the indirect semiconductor material ie the absolute conduction band minimum is offset from the valence band maximum by a momentum. E_c = conduction band; E_v = valence band.

2.6 Light absorption in a semiconductor material.

2.7 Current-voltage characteristic of a photodiode.

2.8 Dependence of light intensity incident on a photodiode on penetration depth.

2.9 Principal response of a p-i-n photodiode to a very short optical pulse.

2.10 Pulse response for non-uniform carrier generation in the drift region, $\nu_n = 0$ for illumination from (____) n side; (_ _ _) p side; (_._) uniform generation.

2.11 (a) Photodetection process of a photodiode.

(b) Equivalent circuit of a p-i-n photodiode.

2.12 Variation of NEP of p-i-n photodiode with load resistance R_{eq} for 150 pA dark current, and various background currents.

2.13 An example of Pulse Code Modulation (PCM).

Chapter 3

3.1 Plot illustrates the height profile of the Stokes parameters I, and Q, obtained from the 450 nm photometer.

3.2 Plot illustrates the degree of polarization of the signal resulting from the Noctilucent Clouds for two cycles per revolution.

3.3 Same as 3.2, except the plot illustrates the degree of polarization for one cycle per revolution.

3.4 Spherical geometry showing the scattering plane as defined by the Sun vector, and the optic axis of the photometer. The rocket spin rate is denoted by ω;
ϵ = the elevation,
α_0 = the azimuth of the optic axis,
δ = the depression of the Sun,
α_s = the azimuth of the Sun,
θ = the scattering angle.

3.5 Components of the electric vector of the incident, and scattered light components, vibrating parallel, and perpendicularly to the scattering plane.

3.6 Dynamical coordinates of a spin-stabilized rocket payload.

3.7 N-S magnetic field derivation recorded at Esrange during salvo A of the CAMP Rocket Campaign, 3/4 August 1982. The vertical bars indicate the different payload launch times.

3.8 Esrange 27 MHz riometer absorption recorded at Esrange during salvo A of the CAMP Rocket Campaign, 3/4 August 1982.

3.9 Plot illustrates the height profile of the relative intensity data obtained by the MISU photometers at Esrange during salvo A of the CAMP Rocket Campaign, 3/4 August 1982. The Stoke parameters have been plotted for two cycles per revolution.

3.10 Same as 3.9, except that the Stoke parameters have been plotted for one cycle per revolution.

3.11 Positive ion density profiles measured on ascent of CAMP-P payload in salvo A. The absolute densities were normalized to the electron density measured by propagation experiments on the same payload at 90-95 km. At the NLC height, the proton hydrate concentrations were largest, and also the highest order proton hydrate density was measured, $H^+(H_2O)_7$.

3.12 Temperature profile obtained during the Noctilucent Cloud display on 3/4 August 1982 in salvo A.

3.13 Comparison of four temperature profiles obtained during the Noctilucent Cloud display on 3/4 August 1982 in salvo A.

3.14 Profiles of potential temperature and zonal wind velocity measured with a triaxial piezoelectric accelerometer on the TAD payload in salvo A. Five regions with negative potential temperature gradients indicate the presence of turbulent layers.

3.15 Nike-Orion 31.057 GE payload launch configuration.

3.16 Plot illustrates the height profile of the APL laterally scanning photometer data; (a) upleg flight; (b) downleg flight during the MAED Rocket Campaign during a Noctilucent Cloud display on 26 July 1986 in salvo A. Results from photometers 1 and 2 are similar, except for the π difference spin rate.

3.17 Same as 3.16b except the results presented were obtained by the MISU.

3.18 Temperature profile obtained on 26 July 1986 during a Noctilucent Cloud display in salvo A.

3.19 Positioning of the APL photometers housed in the Mother payloads on both Black Brants flown during the NLC-91 Rocket Campaign on 1st and 9th of August 1991 in salvoes B and A respectively.

3.20 Same as 3.16, except the rocket was launched into a Noctilucent Cloud display on 9th of August 1991 in salvo A.

Chapter 4

4.1 A simple two-beam interferometer, first invented by Michelson.
S = the source
M_1 = the fixed mirror
M_2 = the movable mirror
D = the detector
M_1' = the image of M_1 and M_2 as seen from D

4.2 The spectrum, and the interferogram of a monochromatic source.

4.3 (a) He-Ne Laser Light Path within the Bomem Michelson Interferometer; (b) 3-D view of the Bomem MB Series Scan Arm Movement.

4.4 Laser signal intensities as measured for the two opposite directions of motion of the scan arm.

4.5 The intensity distribution $S(\sigma) = \text{sinc } 2\pi(\sigma-\sigma_0)L$ for a monochromatic line at σ_0. The first zeros are at +/- 1/2L, where L is the maximum path difference.

4.6 The spectrum of a monochromatic source, when a phase shift of $\pi/2$ exists.

4.7 The observing geometry from Gulkana Airport, Alaska, (62.2°N,145.5°W).

4.8 The plot illustrates the relative intensity of the OH (3,1) band obtained by the UWOMI-1.

4.9 The plot illustrates the relative temperature of the OH (3,1) band determined by the UWOMI-1.

4.10 The plot illustrates the considerable night-to-night variability in the mean hydroxyl radiance (Taylor, personal communication).

4.11 The plot illustrates the distribution of the OH (3,1) band temperatures below ~154 K, on 3rd, 7th and the 11th of August 1986 (Taylor, personal communication).

4.12 The plot illustrates the distribution of the temperatures for the UWOMI-1 data set (Taylor, personal communication).

4.13 The plot illustrates the rotational temperatures determined on the 20th of August 1986 were unusually high, starting around 175 K, and not falling below 154 K during the night (Taylor, personal communication).

4.14 The photographic sequence illustrates the movement of the NLC during the night, the 9th of August 1991, Kiruna, Sweden.

4.15 The plot illustrates the nightly averages of the integrated OH (3,1) band radiance. The triangles illustrate the NLC occurrence.

4.16 The plot illustrates the thirty minute average of the integrated OH (3,1) band radiance. The triangles illustrate the NLC occurrence.

4.17 Same as 4.16, except the results presented are the rotational temperatures.

4.18 The plot illustrates the hourly averages of the integrated OH (3,1) band radiance about the local midnight.

4.19 Same as 4.18, except the results presented are the rotational temperatures.

4.20 Vertical profiles of high-latitude summertime temperature. The smooth curve is an average of 30 falling sphere measurements carried out at Kiruna, Sweden [von Zahn and Meyer, 1989]. The dashed extension to this curve connects with a point at 95 km [Groves, 1987], and indicates a lack of data. The wavy profile is from an active-falling sphere measurement by Philbrick **et al.**, [1984]. The heavy solid curve is the calculated frost point temperature for a 2-ppmv water vapour mixing ratio.

Chapter 5

5.1a The Gallium Arsenide (GaAs) detector head.

5.1b Schematic diagram of the GaAs detector head.

5.2 Schematic diagram of the Imaging Fabry-Perot Interferometer (IFPI).

5.3 Etalons E_1E_2 set-up of the IFPI illustrate the formation of circular interference fringes from multiple reflections.

5.4 Narrow concentric Fabry-Perot fringes corresponding to the multiple beam transmission pattern [Born and Wolf, 1987].

5.5 Equation (2.8), the Airy function [Born and Wolf, 1987].

5.6 IFPI reduced to radius-squared space data. Period of observation at 04:57:43 UT on the 2nd of June 1992, in the South-East mirror direction, BLO, Utah.

5.7 Schematic diagram illustrating of the viewing geometry of the IFPI, located at BLO, Utah.

5.8 The logarithmic plot illustrates one sigma error (in ms^{-1}) versus the peak intensity I(0) (number of photons).

5.9a The plot illustrates the relationship between the peak position and day number (1992) for a cutoff = 15, in the zenith direction.

5.9b Same as 5.9a, except the plot graphed illustrates the calibration lamp.

5.9c Same as 5.9a, except the plot graphed illustrates the difference = zenith - calibration.

5.10a IFPI reduced to radius-squared space data, with a cutoff value = 12.

5.10b Same as 5.10a, except the cutoff value = 15.

5.10c Same as 5.10a, except the cutoff value = 18.

5.11 The plot illustrates a typical cloudy night, 18th August 1992. The moon was present, 81% at 4 UT.

5.12 The plot illustrates a typical clear (cloudless) night, 28th August 1992. There was no moon present.

5.13 The plot illustrates the wind speed calculated from the peak intensity assuming a Gaussian distribution with a zero mean wind speed with the sigma calculated from the photon noise.

Chapter 6

6.1 Plot of $\sqrt{|\gamma|} \equiv \sqrt{(g|h|)}/2\Omega a$ versus frequency σ for zonal wave number s = 1. Positive equivalent depths ($\gamma > 0$): a) eastward phase speed ($\sigma > 0$); b) westward phase speed ($\sigma < 0$). In b) some selected periods in days are given, and a horizontal line is drawn at the diurnal period ($\sigma = -1/2$). Also in b), the cross indicates the main propagating diurnal tidal mode, and the circle indicates the "5-day wave". The indices n_E represent the equatorially trapped modes [adapted from Longuet-Higgins, 1968].

6.2 Dispersion curves for gravity waves. The full lines correspond to m = 0, the dashed line is for wave having the velocity of sound c.

6.3 Schematic diagram illustrating the breaking of vertically propagating inter-gravity waves in the mesosphere. The curves labelled (a) and (b), the linear non dissipative theory is approximately valid. At the level of (c) non linear effects are important, with irreversible deformation previously wavy material surfaces, and turbulence near the wave crests, presumably followed by small-scale mixing and dissipation.

6.4 Schematic diagram of the typical daily variation of the solar heating J_{solar} (heavy curve) at a point in the mid-latitudes. The horizontal dashed line represents the diurnal or zonal average of J_{solar}, which is assumed to be balanced by azonal-mean infra-red cooling, so that the net $\bar{J} = 0$. The term J′ is represented by the departure of the heavy curve from the dashed curve.

6.5 Hough functions for the two anti-symmetric and symmetric modes of the lowest order for the semi-diurnal solar tide [Hargreaves, 1979].

6.6a Plot illustrates separately, the individual OH intensity profiles for each viewing direction versus Universal Time (hours), for the 5th of June 1992.

6.6b Plot illustrates all the individual OH intensity profiles versus Universal Time (hours) mapped onto one set of coordinate axes, for the 5th of June 1992.

6.6c Same as 6.6a, except for individual wind speed (ms^{-1}) profiles.

6.6d Same as 6.6b, except for the wind speed (ms^{-1}) profiles.

6.7a Same as 6.6a, except for the 24th of August 1992.

6.7b Same as 6.6b, except for the 24th of August 1992.

6.7c Same as 6.6c, except for the 24th of August 1992.

6.7d Same as 6.6d, except for the 24th of August 1992.

6.8a Same as 6.6a, except for the 25th of August 1992.

6.8b Same as 6.6b, except for the 25th of August 1992.

6.8c Same as 6.6c, except for the 25th of August 1992.

6.8d Same as 6.6d, except for the 25th of August 1992.

6.9a Same as 6.6a, except for the 26th of August 1992.

6.9b Same as 6.6b, except for the 26th of August 1992.

6.9c Same as 6.6c, except for the 26th of August 1992.

6.9d Same as 6.6d, except for the 26th of August 1992.

6.10a Same as 6.6a, except for the 24th of August 1992.

6.10b Same as 6.6b, except for the 24th of August 1992.

6.10c Same as 6.6c, except for the 24th of August 1992.

6.10d Same as 6.6d, except for the 24th of August 1992.

6.11a Same as 6.6a, except for the 25th of August 1992.

6.11b Same as 6.6b, except for the 25th of August 1992.

6.11c Same as 6.6c, except for the 25th of August 1992.

6.11d Same as 6.6d, except for the 25th of August 1992.

6.12a Same as 6.6a, except for the 26th of August 1992.

6.12b Same as 6.6b, except for the 26th of August 1992.

6.12c Same as 6.6c, except for the 26th of August 1992.

6.12d Same as 6.6d, except for the 26th of August 1992.

6.13 Plot illustrates the averaged data from the various viewing angles over the six nights, 24th to 29th of August 1992.

6.14 Harmonic dial for the semi-diurnal drift components.

6.15a Plot illustrates the mesospheric wind speeds versus BST Time derived from the Sheffield meteor wind radar. The wind direction (SW) has been graphed during 24-29th of August 1992.

6.15b Same as 6.14a, except the wind direction is NW.

List of Tables

Chapter 5

5.1 Summary of the etalon parameters.

5.2 The errors tabulated are for the photon noise added to the Airy function. They do not take into account any background noise or any instrumental effects, and so represent the maximum obtainable accuracy in the presence of photon noise alone.

5.3 The overall averages tabulated are for the zenith, calibration peak 1, and the difference = calibration1 - zenith.

5.4 The overall averages tabulated are for the zenith, calibration peak 2, and the difference = calibration2 - zenith.

5.5 The overall averages tabulated are for the North, East, South, and West (NESW) peak positions, calibration peak 1, and the difference = calibration1 - NESW.

5.6 The overall averages tabulated are for the North, East, South, and West (NESW) peak positions, calibration peak 2, and the difference = calibration2 - NESW.

5.7 The dates, respective day numbers, and calculated offsets are tabulated for cloudy days encountered during the three month period, June - August 1992.

5.8 The table summarizes the data peak height for the range 500 to 2500, the data background, the calibration peak height, the calibration background, and the calculated error (ms^{-1}) for each corresponding data peak height.

5.9 Same as Table 5.8, except the integration time has been doubled.

Chapter 6

6.1 Summary of the various characteristics of the semi-diurnal modes influencing the Middle Atmosphere.

Appendix 3

A3.1 Summary of the complete number of days, June -> August 1992, inclusive, available for analysis, and interpretation. Lunar, and meteorological conditions have been stated for each day.

Contents

Chapter 1

1 Introduction to Noctilucent Clouds and Other Mesospheric Phenomena

1.1 Introduction

This chapter is an introduction to the summer phenomena known as Noctilucent Clouds, and their relationship with other mesospheric phenomena.

NLC occur near to the summer mesopause at a height of approximately 85 km, and are the highest natural cloud formation known to exist in the Earth's atmosphere.

The vertical structure of the Earth's atmosphere is described in Section 1.2 primarily to illustrate the location of the occurrence of NLC to the atmosphere as a whole, and a brief introduction to NLC is presented in Section 1.3.

An account of the history of the progress of NLC studies from the first recorded observation in 1885 up to the present day is presented in Section 1.4. Methods of observation are discussed from the earliest techniques to the state-of-the-art methods employed today.

The general characteristics of NLC are described in Section 1.5,

and a discussion of the relationship between NLC, and other naturally occurring manifestations is presented in Section 1.6.

1.2 Vertical Structure of the Atmosphere

The atmosphere of the Earth, although in movement in both the horizontal and vertical directions, is in fact always close to being in hydrostatic equilibrium. The most basic feature of the atmosphere is the decrease of pressure, and density with increasing height. The height variation is best described by the hydrostatic equation derived as follows.

Let a gas contain n gas molecules, each of a mass m, per unit volume, a cylinder of unit cross-section will contain a total mass of nm.dh, where dh is the height of the cylinder (see Figure 1.1). Under gravity, g, the cylinder will experience a downward force of nmg.dh, which in static equilibrium, is balanced by the net upward force due to the difference of pressure between the top and bottom faces:-

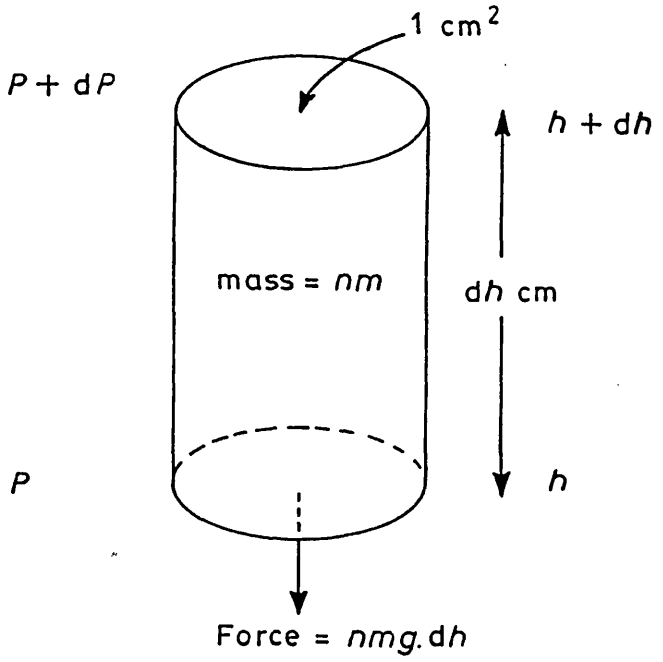
$$(P + dP) - P = -nmg.dh \qquad \textbf{(1.1)}$$

where P = the atmospheric pressure

therefore, Equation (1.1) becomes:-

$$\frac{dP}{dh} = -nmg \qquad (1.2)$$

Figure 1.1



But, by the gas law, given by:-

$$P = n\,k_B\,T \qquad (1.3)$$

where k_B = the Boltzmann constant = 1.38 x 10^{-23} JK^{-1}

T = the atmospheric temperature

Hence, Equation (1.2) can be re-written as:-

$$\frac{1}{P}.\frac{dP}{dh} = -\frac{m\,g}{k_B\,T} = -\frac{1}{H} \qquad \textbf{(1.4)}$$

where $H = k_BT/mg$ = the scale height

If H is constant, then Equation (1.4) can be written as:-

$$P = P_o \exp \frac{(-h)}{H} \qquad \textbf{(1.5)}$$

where P_o = the atmospheric pressure at a height h = 0

The scale height is a vertical distance in which P changes by a factor e. Even if g, T, and m are not independent of the height, H can be defined by Equation (1.4) in terms of the relative rate of change of concentration with height. As Chapman [1931][1] indicated in his classic paper on the Upper Atmosphere, this approximation is good over layers which are extended in height. Equation (1.5) can be presented as:-

$$P = P_0 \exp \frac{-(h - h_0)}{H} = \exp -z \qquad \textbf{(1.6)}$$

where $P = P_0$ at $h = h_0$

z = the reduced height = $(h - h_0)/H$

The Earth's atmosphere is commonly described as a series of layers defined by their thermal characteristics (see Figures 1.2). Specifically, each layer is a region where the change in temperature with respect to altitude has a constant sign. The layers are called "spheres", and the boundary between connecting layers is the "pause".

The troposphere

The lowest layer, called the troposphere, exhibits generally decreasing temperatures with increasing altitudes, at a rate of 10 Kkm^{-1}, up to a minimum known as the tropopause. The temperature and location of the tropopause vary with latitude and season (see Figures 1.2). At the equator, its mean altitude is located near 18 km, and the corresponding temperature is about 190 K, while in the polar regions its elevation is only about 8 km, and the temperature approximately 220 K.

The majority of the atmospheric gases occur in the troposphere. The troposphere is a turbulent layer, with the weather located at its base, dominantly driven by above air flow patterns.

The stratosphere

Above the tropopause, the stratosphere begins, exhibiting increasing temperatures with altitude up to a maximum of about 270 K at the level of the stratopause located near 50 km. This maximum is the result of the absorption of the Sun's ultra-violet rays by the ozone. At a still higher altitude, the temperature again decreases up to 85 km, where another temperature minimum is found.

The positive temperature gradient in the stratosphere prevents air in this region becoming turbulent. This inhibits air convection, so the tropopause acts as a "lid" confining the air turbulence to the troposphere. A characteristic trend in the stratosphere is that there is a tendency for the equatorial, and low latitude winds to change direction every twenty-six months (biennial oscillation). This phenomenon is poorly understood, and will not be discussed further, but it should be noted that it may have an important influence on the circulation in the troposphere, as well as the stratosphere.

The highest concentrations of ozone are found within the stratosphere. Ozone is formed by the photolysis of atomic oxygen (O_2) by the ultra-violet component of sunlight. Ozone is vital to the survival of the life on Earth, since it filters out much of the Sun's incoming ultra-violet radiation, which would otherwise cause damaging effects on most biological organisms, and eco-systems.

The mesosphere

The mesosphere (or Middle Atmosphere) is a region located between the stratopause (situated at 50-55 km), and the mesopause (situated at 80-85 km). The global mean temperature is 180 K at the mesopause, making it the coldest region in the atmosphere. At these altitudes, 80-85 km, the atmosphere is relatively transparent to ultra-violet, visible, and near infra-red solar radiation, and the local air temperature is set principally by mixing, due to radiation from the air itself, from the Earth's surface, from the Lower Atmosphere, and by bulk movement in the vertical direction, and turbulent dispersion of winds, gravity waves, and tidal components.

The global circulation of the Middle Atmosphere is largely the result of solar heating in the stratospheric ozone layer, imposing an upward (cooling) movement of air over the summer polar regions, with downward (warming) movement of the air over the winter polar regions.

Consequently, the Middle Atmosphere shows the paradoxical behaviour of being colder in the summer than in winter. This is the result of solar heating being less important at these high levels than the temperature changes caused as a result of expansion or compression when the air rises or falls.

In the troposphere, stratosphere, and mesosphere, the major constituents, N_2, and O_2, contribute about 80%, and 20%

respectively, of the number density, so that the mean molecular mass of the air varies minimally with altitude. Because of this common feature, the three layers are collectively known as the homosphere.

The thermosphere

The region located above the mesopause is called the thermosphere. The temperatures there increase very rapidly with altitude, and can range from as low as 500 K to temperatures in excess of 2000 K, depending on the level of solar geomagnetic activity. The composition at these altitudes is very different from that of the lower regions due to an increasing proportion of atomic oxygen, whose density becomes comparable to and even greater than those of O_2 and N_2 above the 130 km. The abundances of O_2 and N_2 decrease, primarily as a result of rapid photodissociation. In contrast to the homosphere, the mean molecular weight of air in this region therefore varies with altitude; for this reason the region above 100 km is also called the heterosphere.

Finally, at greater heights, above the range illustrated in Figures 1.2, the atmospheric temperature rises to several or a few thousand degrees Kelvin. This is the result of a continuous stream of charged particles, called the solar wind, impinging on the Earth's magnetosphere.

1.3 Introduction to Noctilucent Clouds

1.3.1 What are the phenomena known as Noctilucent Clouds?

Noctilucent Clouds or Night Luminous Clouds (from hereon, abbreviated to NLC) are the highest natural cloud formation known to exist in the Earth's Upper Atmosphere [Goldberg and Witt, 1977][2]. NLC are observed from the ground against the background of a twilight sky. To the observer, NLC appear to be a silvery or blue colour [Gadsden, 1982][3], however, their colour may be reddened at low elevation angles when the light passes through an increased path length of the troposphere [Fogle and Haurwitz, 1966][4].

NLC are only observed when the Sun has set below the horizon, between 6° and 16° [Fogle and Haurwitz, 1966][4], and this indicates that the cloud must be located at a considerable altitude. Stereophotogrammetric measurements [Witt, 1962][5] confirmed that the altitude of the clouds typically varied from about 80 km to 85 km.

In the Northern Hemisphere, NLC are most frequently observed at high latitudes, usually between 50°N and 60°N, during the boreal summer, June to August. In the Southern Hemisphere, NLC are observed at high latitudes, 50°S to 60°S, during the austral summer, December to February [Fogle, 1964; 1966][6,7]. This is largely due to the distribution of observers, and viewing conditions.

NLC are tenuous [Ludlam, 1957][8], and often resemble cirrus or cirrostratus, frequently displaying a delicate filigree pattern with marked wave structure. For clear classification of NLC types [Grishin, 1955][9], see Section 1.5.2.

1.4 Historical Approach

The first recorded observation of NLC was made by Backhouse [1885][10] in Kissingen, Germany, on June 8 of that year. NLC had been noted prior to the eruption of Krakatoa (1883), but the published accounts were insufficiently certain for a clear identification to be made [Arago, 1854; Scultetus, 1949, who discussed an observation made by Lavoisier][11,12]. Many of the pre-1885 observations had been made by the following observers:- Hartwig (in Tartu); Kiessling and Tseraskii (in Moscow), and Jesse (in Berlin).

NLC were first systematically studied by Jesse and Tseraskii in 1885, and in later years. Jesse [1896][13], made the first accurate height (triangulation techniques) and velocity measurements of these clouds, and also gave them their present day name.

The study of NLC waned during the First World War. However, Astapovich [1925][14] (in Russia), and a few years later, Stormer [1933][15] (in Norway), observed and measured the height of NLC. Stormer used pre-existing auroral observation instruments in order to measure and observe the height of the NLC. Vestine [1934][16], wrote a paper published in two consecutive parts

containing a useful catalogue of observations made during the period 1885-1933. Again, interest in NLC decreased during the Second World War. Post-war, NLC observations increased, Paton [1949][17] (in Scotland), and Khvostikov and Grishin (in Russia), began their work on NLC. Grishin [1955][9] published the first detailed classification of NLC structures and again, Grishin [1956][18] published the first spectrum of these clouds. Interest in NLC intensified during the International Geophysical Year (IGY) (1957-1958). Ludlam initiated the study of NLC in Sweden, which was taken up with enthusiasm by Professor Georg Witt from the Arrhenius Laboratory, Stockholm University. Since the IGY, the number of scientists actively involved with the study of NLC increased considerably. Systematic NLC observations were made from meteorological stations in both the Northern and Southern Hemispheres, between the latitudes 45° to 90°. Currie [1962][19], and later Fogle [1966][7] encouraged North American and Canadian observers. Prior to the summer of 1962, only a few sightings of NLC had been reported from the North American continent, whilst hundreds had been made in Europe and Russia. The question was raised as to why NLC do not occur over North America as often as elsewhere. To improve the general knowledge of NLC occurrence, a study of NLC over North America was initiated, involving the meteorological stations situated in Alaska and Canada. Using this network of observation stations, the resulting data showed that NLC occurred just as frequently over North America as over Europe and Russia. It was concluded that the scarcity of observations was due to the lack of interested and informed observers.

However, in 1960, under the direction of R.K.Soberman, the Meteor Physics Branch at the Air Force Cambridge Research Laboratories considered the possibility of direct rocket investigations of NLC. In June 1961, the first successful flight of the "Venus Flytrap", collected samples of meteoric dust in the Upper Atmosphere. With direct sampling available, it was concluded that the same principle could be applied to NLC samples. During the summer of 1962, at Kronogard, Sweden, four NASA Nike-Cajun solid fuel rockets were launched during a NLC display. Rocket flights have continued during the summer:- CAMP (Cold Arctic Mesopause Project) in 1982; MAED (Middle Atmosphere ElectroDynamics) in 1986; NLC-91 (Noctilucent Clouds 1991) in 1991. Rockets are the most direct way of analyzing the NLC environment in-situ.

LIDAR (LIght Detection and Ranging) measurements [Fiocco and Grams, 1969][20] were made at Kjeller (60°N) on twenty-one nights during July and August, 1966. The resulting measurements indicated that the 60 km to 70 km altitude region contained an appreciable amount of particulate material. Fiocco and Grams observed that the NLC layer existed at an altitude close to 74 km. Although changes were observed in height from 75 km to 73 km, in a period of 90 seconds, it should be noted that the height given for this cloud was still far outside the range of modern altitude measurements of NLC either by ground-based photography [Grahn and Witt, 1971][21] or rocket-borne photometers [Gadsden, 1982][3]. Hansen **et al**., [1989][22] have claimed to have obtained the first measurements of NLC by lidar.

Theon **et al.**, [1967][23] published the results from eight rocket soundings conducted from Kronogard, Sweden (66°N), and Barrow, Alaska (71°N), during the summers of 1963-1965 providing measured profiles of atmospheric temperature, pressure, density and wind to an altitude of 95 km. From these rocket-borne acoustic grenade techniques, a determination between the mesospheric temperature structure and the occurrence of NLC was made. Soundings were carried out in the presence and absence of NLC displays. At Barrow, the minimum temperature of the profile made in the presence of NLC was 139 K, which was 3 K warmer than the mesopause temperature of the control sounding. At Kronogard, however, the soundings conducted in the presence of NLC recorded mesopause temperatures varying from 130 K to 147 K.

Correspondingly, for soundings conducted during the absence of NLC, recorded mesopause temperatures varied from 129 K to 149 K. Since no significant temperature difference was noted between the soundings conducted in the presence and absence of NLC, the authors concluded that a mesopause temperature of less than 150 K was necessary, but not a sufficient condition for the existence of NLC.

Even today there are three methods of measuring temperature profiles in the mesopause region. These include the rocket-borne acoustic grenade technique [Nordberg and Smith, 1964; Nordberg **et al.**, 1965][24,25] and sodium (Na) lidar technique.

1.5 General Characteristics of NLC

1.5.1 Geometry of NLC

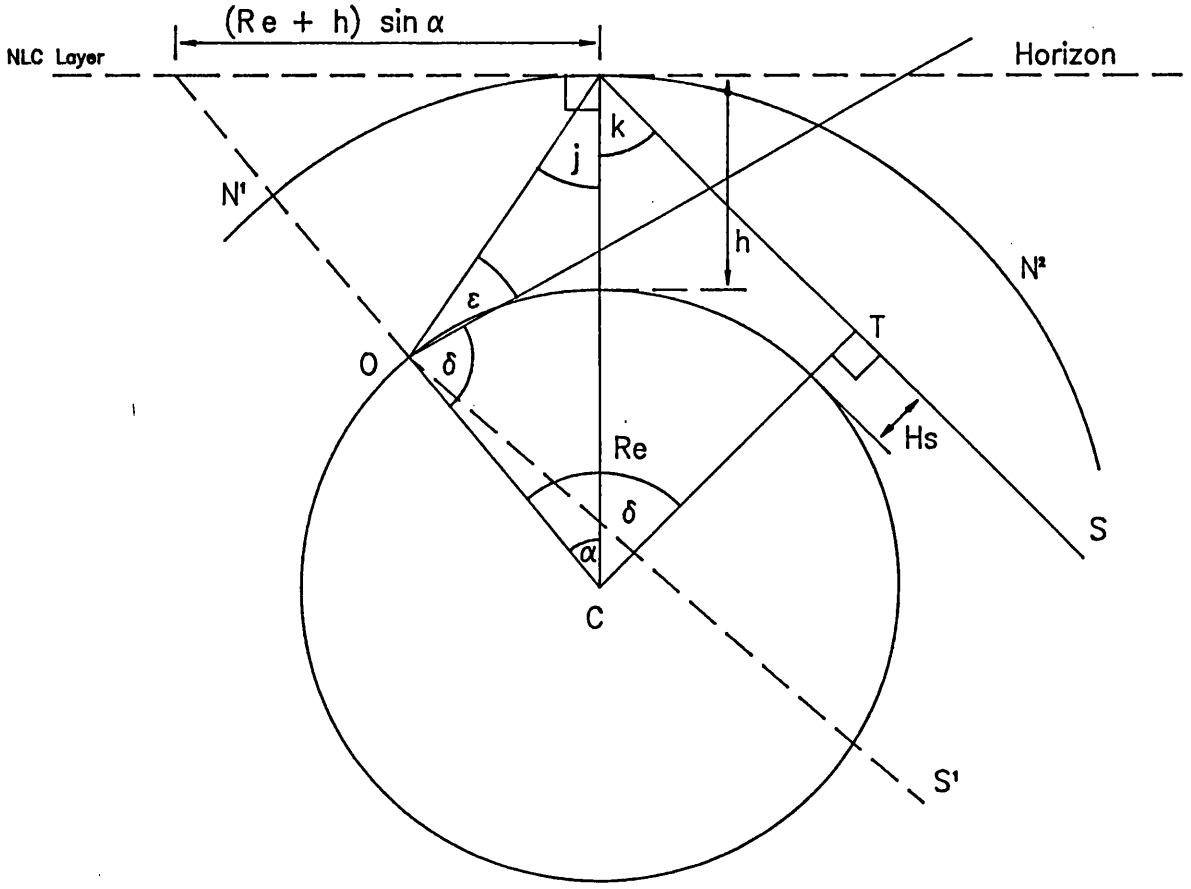
For NLC to be observed, the following conditions must be satisfied:-

1) there must be no obstructions in the line of sight of the observer, for example tropospheric clouds,
2) the 82 km region must be sunlit - this condition is fulfilled when the Sun is not more than 16° below the observer's horizon,
3) the sky background must be dark enough for the clouds to stand out - this is usually requires that the Sun is at least 6° below the horizon,
4) the observer must be in an appropriate latitude region to observe NLC.

The observing geometry is illustrated in Figure 1.3. Absorption and refraction of the sunlight in the Lower Atmosphere have been ignored.

An observer at O sees the NLC extend from the horizon up to an elevation ϵ given by the intersection of the solar grazing ray SN and the NLC layer N_1N_2 at the height h above the Earth's surface. Because of the effects of the atmospheric screening, only those rays incident upon the atmosphere at altitudes greater than h_s will be able to illuminate the NLC.

Figure 1.3



where h_s = the atmospheric screening height

h = the height of the NLC layer

δ = the solar depression angle

ϵ = the elevation of the NLC layer

R_e = the mean radius of the Earth = 6367 km

The maximum elevation ϵ, at which NLC can be seen, if they are present, is given by the equations:-

$$j = 90^\circ - \epsilon - \alpha \qquad (1.7)$$

$$k = 90° - \delta + \alpha \qquad (1.8)$$

$$\cos(\delta - \alpha) = \frac{(R_e + h_s)}{(R_e + h)} \qquad (1.9)$$

$$ON^2 = R_e^2 + (R_e + h)^2 - 2R_e(R_e + h)\cos\alpha \qquad (1.10)$$

$$\cos\epsilon = \frac{(R_e + h)\sin\alpha}{ON} \qquad (1.11)$$

Assuming that the atmospheric screening height is 10 km [Taylor **et al.**, 1984][26] and h = 83 km, the values for ϵ for various values of δ are shown in Table 1.1.

For sufficiently small solar depression angles, NLC may be seen to extend beyond the zenith. It is generally agreed that the sky background is too bright for NLC to be seen when the Sun is less than 5° below the horizon. Table 1.1 suggests that NLC cannot be seen above the horizon when the solar depression angle is more than 17°.

Requirement 2) depends upon the amount of light scattered from a region varies with the size of the scattering angle, and (in

the case of non-spherical or anisotropic scatterers) with the azimuth, of the direction of the scattered light. These angles, and the orientation of the scatterer are measured relative to the scattering plane, which is defined as the plane containing the direction of the incident light and a point at which the scattered light is detected. The scattering angle is zero in the direction of travel of the incident light (forward scattering), and 180° (back scattering) if the light travels anti-parallel to the incident light.

The extent and the frequency of tropospheric cloud and similar meteorological phenomena determines whether requirement 1) is satisfied or not. The tenuity of NLC means that even very thin or patchy cirrus will cause confusion or hide a display altogether. Since these clouds and the bulk of the aerosol constituting atmospheric haze exists within the troposphere, an effective screening layer of some several kilometres thickness may be expected. Intuitively, this layer would extend from the ground to the upper regions of the troposphere or lower regions of the stratosphere and would so attenuate the Sun's rays that they no longer illuminated the NLC field. Bright NLC displays could be seen through it, but detail will be lost.

Table 1.1

Solar Depression Angle, °	Maximum NLC Elevation, °
17	1
16	2
15	3
14	5
13	7
12	11
11	16
10	28
9	63
8	131
7	156
6	166
5	170

1.5.2 Brightness, Colour and Morphology

NLC are usually tenuous, and of low brightness. However, it has been noted that occasionally the clouds are bright enough to cast shadows and permit the reading of printed matter outdoors. Khvostikov [1964][27] and Sharonov [1960][28] have measured the albedo of NLC. NLC photography is mainly governed by brightness of the sky background, but also by the lens and speed of the film employed.

Characteristically, NLC have a silvery or blue colour [Gadsden, 1982][3]. It is now known [Gadsden, 1975; Avaste **et al.**, 1988][29,30] that the colour is due mainly to the preferential removal of red light by stratospheric ozone along the optical paths traversed

by the Sun's rays striking the clouds. Polarization measurements [Witt, 1960a; 1960b; Willmann, 1962; Tarasova, 1962; Vasilyev, 1962][31,32,33,34,35] showed that the light from the NLC is strongly polarized. Gadsden [1982][3] also notes the presence of circular polarization in scattered light.

Usually NLC resemble the form of thin cirrus-like streaks seen against a diffuse bluish-white structureless background. Sometimes only one or two isolated bands are visible, while at other times they tend to occur in extensive sheets in an almost continuous mass, resembling cirrocumulus or altocumulus undulations. The forms of NLC have been described by various authors [Stormer, 1933; 1935; Vestine, 1934; Grishin, 1955; Ludlam, 1957; Paton, 1964; Fogle and Gotaas, 1963; Fogle, 1965c][15,36,16,9,8,37,38].

Grishin [1955][9] gave the most clear classification of NLC forms. The following classifications are distinguished :-

I) Nebulae (or Flares, or Veils)
II) Bands
III) Billows (or Waves)
IV) Eddies (or Whirls)

Several photographs (see Figures 1.4 and 1.5) of NLC illustrating the forms described below are illustrated at the end of this chapter.

I) Nebulae (more commonly known as Veils) are very tenuous and lack a well defined structure. Veils are often present as a background for other forms. They resemble cirrus clouds, occasionally exhibiting a faintly visible fibrous structure, and a flickering luminosity. Veils are the simplest morphological form of NLC and often precede (by the order of half an hour) the appearance of NLC with a well defined structure.

II) Bands are long streaks with either diffuse edges (Type IIa) or sharply defined edges (Type IIb), sometimes extending hundreds of kilometres. Often Bands occur in groups arranged roughly parallel to each other or interwoven at small angles. Sometimes an isolated Band is observed. Little, if any change with time is observed. The predominant structures found in the NLC field are blurred Bands of low mobility, particularly when the brightness is low. Bands are observed throughout a display (for several hours). They move in a direction and with a speed that is different from that of the display as a whole. This form is perhaps the most striking feature of an NLC display. Smaller streaks with twists, or bends, may lie across the Bands, or branch out from them, giving the appearance of a feather. The distance separating pairs of Bands ranges from tens of kilometres to over a hundred kilometres.

Closely spaced serrations, roughly parallel with thin streaks are occasionally seen in the Veil background. The distance

separating serrations is of the order of a few kilometres.

III) Billows (or Waves) are groups of closely spaced short Bands. Occasionally, Billows may consist of straight, narrow Bands which are well defined, they tend to lie parallel to one another and are generally short (Type IIIa). At other times, they exhibit a wave-like structure with undulations in the short Bands (Type IIIb). The distance separating pairs of Billows is of the order of ten kilometres. Billows tend to lie across the direction of the Bands, but their alignment may differ noticeably in neighbouring parts of the sky. Sometimes they appear alone against the Veil background. In contrast to the Bands, the Billows may change their form and arrangement, or appear and disappear within several minutes.

IV) Eddies (or Whirls) which usually exhibit varying degrees of curvature, are sometimes observed in Band serrations and occasionally in Veils. The Eddies extend over variable arcs, and on rare occasions, complete rings or vortices with dark centres. This structural form indicates the presence of turbulence at the mesopause. Eddies of small radii of curvature (0.1° to 0.5°) are classified as Type IVa. Eddies have the form of a simple bend of one or several bends with a radius of 3° to 5° are classified as Type IVb. Large scale Eddies are classified as Type IVc.

Sometimes intersecting groups of long Bands are observed. At the

points of intersection bright knots will appear. These knots will move at greater speeds and in different directions from the Band systems. When determinations of the drift motion of NLC are made, knots and/or wave crests should not be used, as they do not represent the true motion of the air at NLC altitudes.

The occurrence of these groups of intersecting Bands led to the speculation that they may be a double mesopause, and two closely spaced layers of NLC. Other evidence pointing towards the possibility of a double mesopause has been discussed by Schilling [1964][39]. The question of whether or not there are sometimes two layers of NLC, was settled by precision height measurements of individual Bands in the two intersecting groups under investigation [Fogle and Haurwitz, 1966][4].

In complex displays, two or more (sometimes all) of the previously described forms are observed simultaneously. A powerful method for examining the structural changes, growth and decay of NLC is the use of time-lapse motion picture photography. One can obtain information on the evolution of the structural details of the NLC with the help of time-lapse motion pictures otherwise almost imperceptible to the naked eye. These films are of considerable use in determination of NLC velocities. See Thayaparan [1990][40] for photogrammetry determinations.

1.5.3 Height

The height of NLC has been measured directly, using photographic

triangulation, on a number of occasions [Jesse, 1896; Stormer, 1935; Paton, 1949; Khvostikov, 1952; Burov, 1959; 1966; Witt, 1962; Dirikis **et al.**, 1966][13,36,17,41,42,43,5,44]. The values ranged from 74 km to 92 km, the average height being 82 km [Gadsden and Schroder, 1989][45]. Today, after many rocket campaigns flown during NLC displays, it is now known that NLC occur at altitudes near to the summer mesopause (~85 km).

The general method that was used in determining the height of a NLC was that of parallactic photography [Witt, 1962][5]. In this method, Witt had the use of precise photo-theodolites placed at each end of a geodetically-surveyed baseline of length 51.5 km. The parallax between identical points in corresponding photographs was measured, and from this the height was obtained.

1.5.4 Variation of NLC Observations with Latitude

As a general guide, it has been accepted that NLC will be seen only when the solar depression is between 6° and 16°, and as mentioned before, these clouds are almost exclusively a high-latitude phenomenon.

It has been reported over North America that the highest and lowest latitudes from which NLC have been observed are 45.5°N and 71.3°N respectively, with maximum number of sightings in the latitude range 52.5°N to 66.5°N (see Figure 1.6). These results agreed with those made in the USSR [Bessonova, 1963; Sharonov, 1965][46,47].

Obviously the optimum latitude for NLC lies between 52.5° to 66.5°. Few displays are observed in latitudes North of 60°, because the sky background is far too bright for NLC to be observed even if they were present. It should be noted here, that a small number of artificial NLC have been observed at latitudes below 45°. These clouds have been associated with rocket launchings, from which exhaust products are rich in water vapour and solid particles [Meinel **et al**., 1963; WMO, 1966; Avaste **et al**., 1988][48,49,30].

1.5.5 Variation of NLC Observations with Longitude

Variation with longitude of the frequency of NLC occurrence was proposed by Sharonov [1965][47], who had analyzed the NLC observations reported from the USSR during 1958 and 1959. The observed NLC frequency was found to increase from East to West. To investigate this longitudinal variation further, a similar analysis was carried out in North America, during the years 1964 and 1965. For this analysis, the observing stations positioned in between 51°N and 61°N latitude range, were grouped into 40° sectors of longitude ie

55°W to 95°W
95°W to 135°W
135°W to 175°W

The results were normalized to take into account the nights when all stations in a sector were overcast. The results did not

indicate an increase in NLC activity from East to West as Sharonov had postulated. However, the greatest number of NLC sightings were reported from the middle sector (95°W to 135°W) which, in addition had the most observation stations and had the larger percentage of clear station nights.

But, there may be a longitudinal effect particularly at longitudes 0°E and 140°W, where the auroral zone crosses the regime of maximum NLC occurrence [Fogle, 1966][7]. With the evidence that a low mesopause temperature is required for NLC formation, one might expect the years of high sunspot activity to produce fewer NLC displays at latitudes 0°E and 140°W than at other longitudes. This is because the mesopause in the auroral zone frequently experiences considerable heating, associated with high sunspot activity (see Section 1.6.3).

1.5.6 Drift Motion of NLC

The drift motion of NLC provides information on the wind velocity at the mesopause. In principle, the observations are quite simple. To determine the true wind characteristics at NLC heights, velocity measurements of the NLC system as a whole rather than the velocity of individual bands or bright knots must be determined. Observations of the bands or knots generally yield values that are improbably high (200-300 ms^{-1}) and in a direction opposite to that of the true wind. Motion pictures/cinematography provide the most reliable velocity determinations [Taylor, 1986][50]. Cinematography clearly distinguished the various types

of motion. They also clearly showed the motion of Bands relative to the cloud system. NLC velocity measurements made in the Northern Hemisphere revealed that the motion is generally towards the south-west. Theon, Smith and McGovern [1969][51] reported on the results obtained from eight grenade soundings conducted from high latitude sites during the summers 1963-1965. Theon **et al.**, [1967][23] discussed the effect of temperature structure on the production of NLC. Wind profiles were measured simultaneously as part of those same grenade soundings. The wind speeds observed at the mesopause ranged from 24-84 ms^{-1} during NLC displays. Theon, Smith and McGovern [1969][51] concluded that the wind shears at the mesopause were not related to the occurrence of NLC, because similar shears had been observed at other latitudes and seasons. However, they did believe that the lower wind speeds at the mesopause, which accompanied the NLC displays, were significant. Today wind, temperature and hence, density profile measurements made in the NLC environment, are obtained from falling spheres [Schmidlin **et al.**, 1991][52] and radar techniques, for example EISCAT (European Incoherent Scatter).

1.5.7 Variation in Occurrence with Season

In the Northern Hemisphere, NLC have been reported as early as April [Gavine, 1987][53] and as late as October; however, the "season" for NLC in the Northern Hemisphere, when the greatest number of NLC sightings are made, is during the summer months June to August.

The sparse observations available from the Southern Hemisphere confirm the season to be truly summer; in the Southern Hemisphere, the clouds are seen during December to February [Fogle, 1965c][38]. Up-to-date references cannot be made, because of the lack of evidence, as there are still very few observing stations in the Southern Hemisphere, and no future plans are being considered (Dyson, personal communication).

1.5.8 Spatial Extent of NLC

At an altitude of 82 km, a circular region of about 1000 km in radius, and an area of about 3 million km^2 is theoretically visible from the ground, but only a part of this can be illuminated by sunlight against a dark sky. In the case of small NLC displays covering an area of a few tens of thousands km^2, the whole boundary of the NLC display can be seen from a station. However, at times extensive NLC displays have occurred which are too large for it all to be visible from any one place. The spatial extent of the widespread displays have been estimated to be sometimes in excess of 3 million km^2. Maps of widespread displays during 1963 and 1964 have been given by Fogle [1964a][54], and Fogle, Chapman and Echols [1965][55]. Avakyan **et al.**, [1986][56] reported on the results obtained from the fifth expedition of Orbital Station Salyut-6.

1.5.9 Duration

The observed duration of NLC varies from a few tens of minutes

to more than five hours. Occasionally, the extensive NLC have been seen on two or more successive nights, but whether they are continuous throughout the intervening daytimes remains uncertain.

1.5.10 NLC Formation

Despite many years of research, questions of the origin and composition of the NLC still have not been definitely answered. Naturally, because the NLC were observed in association with an increased awareness of twilight sky phenomena following the Krakatoa eruption in 1883, the earliest suggestions which were made related the cloud activity to the occurrence of volcanic eruptions. The dust from Krakatoa was hypothesized to appear after a delay of a couple of years in the Upper Atmosphere. These hypotheses later gave way to more complex explanations (see Section 1.6), to account for the clouds appearing in very dry regions of the atmosphere, and at low pressures [Hesstvedt, 1961; 1962; 1964; 1969b; Turco **et al**., 1982; Vestine and Deirmendijan, 1961; Witt, 1969][57,58,59,60,61,62,63].

Other suggestions were proposed relating the cloud activity to the influx of meteoric and cometary material, and solar activity as evidenced by the sunspot cycle. The correlation or lack thereof with the suggestions mentioned earlier should provide the necessary clues.

The theories for NLC formation have taken three general forms:-

1) vulcanism
2) extra-terrestrial dust
3) water ice nucleation

A satisfactory theory of NLC formation has still to be found which explains:-

i) their proximity to the mesopause,
ii) their appearance during only the summer months,
iii) their presence at latitudes greater than 45°,
iv) their seemingly random occurrence,
v) the reason they exist for only a few hours,
vi) their tenuous structure.

The aspects of the physics of NLC formation have been reviewed [Ludlam, 1957; Witt, 1962; Fogle and Haurwitz, 1966; Gadsden, 1982][8,5,4,3] with particular emphasis on the nucleation of ice particles and the types of ice structure nucleated.

There is a general agreement that the condensation nuclei of NLC particles are coated in a volatile substance which is almost certainly water ice [Hemenway, Soberman and Witt, 1964][64].

To allow NLC to be seen from the surface of the Earth and satellite [Avakyan **et al.**, 1986; Thomas, 1991][56,65] particles must form and grow to sizes that scatter sufficient sunlight. The temperature must be low enough and the mixing ratio of water vapour must exceed the ice saturation pressure at the temperature

involved to enable the nucleation process to proceed [Theon **et al**., 1967; Witt, 1969][23,63].

It is highly likely that nucleation is heterogeneous rather than homogeneous [Roddy, 1986][66]. Since there is evidence for the presence in the mesopause region of both cluster-ions and particles of meteoric origin [Goldberg and Witt, 1977; Bjorn **et al**., 1985][2,67] it seems highly likely that the nucleation is heterogeneous. Both experimental, and theoretical studies indicate that homogeneous nucleation is improbable. A sufficient supply of suitable nucleating particles must be available.

Hesstvedt [1961; 1962][57,58] proposed that ice particles growth was the result of water condensation on pre-existing particles. He assumed that the particles were of meteoric origin. However, Witt [1969][63] proposed that instead of meteoric dust particles, large ionospheric ions would act as condensation nuclei. Witt suggested that proton hydrates ($H^+(H_2O)_n$) were the most probable family of ions to grow to large sizes. He further speculated that large metal hydrate ions might be present in the ionosphere and that these might act as the condensation nuclei. If the latter hypothesis is true, it could be a clue to the problem of a possible link between meteoric activity and NLC formation (see Section 1.5.2).

Arnold [1980][68] gave further support to the idea of ion-induced nucleation as a possible NLC formation mechanism. Bjorn and Arnold [1981][69] made in-situ measurements leading to the evidence

of heavy ions with masses up to 360 amu which were interpreted as proton hydrates with up to 20 water molecules.

A temperature minimum value of 130 K, was measured during a NLC display, and recorded [Theon **et al.**, 1967][23]. However, Gadsden [1982][3] stressed that comparable very low temperatures (130 K to 147 K) had been measured at similar times of the year, and latitudes when NLC were not observed. A low mesopause temperature is a necessary, but not a sufficient condition alone for the formation of NLC.

Of course, large particles may be present, not least because of the meteoric particle production, but the number of such particles is small, and they are not confined to the Arctic latitudes.

Regarding the formation of NLC, several models, and some serious theories have been proposed. The models are highly simplified, and in some cases, inadequate attention has been paid to the crucial questions by the "model engineers". Firstly, models are Eulerian, but the process of NLC formation occurs in a Lagrangian frame of reference. The air parcel in which the particles grow, moves in a three-dimensional space, its temperature changing because of the wave motion, and velocity changes with altitude.

As to the key question itself, namely the starting point of particle growth-nucleation, it is not yet known whether the process is impulsive (a large number of nuclei produced at once

within a short period of time), or if there is a virtually continuous genesis. It is clear that the phenomenon needs nuclei, whether or not these are produced by hydration of oxonium ions (H_3O^+) (sufficient, but not necessary) or deposition of water onto Hunten's postulated "dust" of meteoric origin. Once nuclei begin to grow (whilst sedimenting and diffusing through the mesopause region), they will reach a maximum particle size depending on the amount of water that they encounter. However, it is important to remember that at 140 K, evaporation is very slow; even if there was no water at all, the particles would survive for a long time compared to their transport time scale. Only when the temperature exceeds 150 K-160 K will the particles evaporate within a very shallow layer that defines the cloud base. How this water is recycled is a question of modelling the vertical turbulent diffusion.

The climatology of the mesospheric temperature profiles indicates that the temperature profile is well behaved up to 82 km. Above 82 km, however, there are large variations in the temperature, and as was seen from past rocket flights, the NLC layer was not always uniform.

1.6 The Relationship Between NLC and Other Naturally Occurring Manifestations

1.6.1 NLC and Volcanic Eruptions

The idea of a connection between NLC and volcanic eruptions was

first expressed by Kohlrausch [1887][70] and Foerster [1906][71]. Bright mesospheric clouds were observed between 1885 and 1894, and had not been recorded previously. This suggested that their appearance was connected with some powerful catastrophic process. Such a phenomenon was the violent (Plinian) eruption of the volcano Krakatoa. The eruption involved the emission of 35 million tons of fine dust and yielded a Dust Veil Index of 1000. The optical anomalies noticed after Krakatoa were:- bright sunrise and brilliant sunsets, Bishop's ring, polarization anomalies, reduced transparency of the atmosphere [Vasilyev, Zhuravlev **et al**., 1965][72]. These anomalies lasted for about two years, then gradually weakened. However, NLC appeared only towards the end of this period. Kohlrausch and later, Wegner [1926][73] suggested that NLC particles were ice crystals resulting from the volcanic injection of water vapour into the Upper Atmosphere. However, the year-to-year variations of NLC did not seem to be connected with the eruptions, for example Mount Katmai, Alaska, 1912, was not followed by any increase in the frequency of reports of NLC. On the other hand, Spangenberg [1949][74] had tried to associate more frequent reports in the years 1932-1935 with an eruption in the Cordilleras in 1932, which threw up a cloud to 25 km. Fogle and Haurwitz [1966][4] also argued that increases in NLC activity followed the Mount Agung eruption in 1963.

With major volcanic eruptions occurring sufficiently often for some to lie within the three month observation period, and with NLC displays occurring in the absence of major volcanic

eruptions, it is hardly surprising that no close correlation can be made between volcanic activity and the occurrence of NLC. Further, in order for the fine volcanic dust to transverse through the stratosphere, it would take 2-3 years (Scott, personal communication). Again, no close correlation between volcanic eruptions and cloud activity can be drawn.

1.6.2 NLC and Influx of Extra-Terrestrial Material

Vestine [1934][16] pointed out that the period 1880-1887 was one of outstanding meteoric and comet activity. On June 30, 1908, a small comet fell in Siberia. During the same night an unusually brilliant display of NLC were reported by observers in Russia, Sweden, Denmark, Germany and England. Witnesses of the aforementioned comet reported seeing a light bluish trail residual in the atmosphere. Smaller meteors, large enough to be seen in daylight were observed, leaving residual smaller pale-silvery streaks, indicative of submicron particles. This conclusion was supported by the work of Fessenkov [1949][75].

However, no significant increase in NLC activity has been reported on the nights when the peak of Arietids (~June 8th), ζ - Perseids (~June 9th), Aquarids (~July 28th), and Perseids (~August 12th) meteor showers have occurred. Furthermore, the peak of NLC activity has a maximum around mid-July when there is no major meteor activity. This cannot be taken as evidence of strong dependence of NLC activity on an increased flux of meteors since NLC are observed in the Southern Hemisphere during the

months of December-February when the total meteor activity is relatively low.

On basis of the available evidence, it can be concluded that there is no clear correlation between NLC and meteor showers. For confirmation, a satellite continuously observing the occurrence of NLC and meteor activity, for a long duration, would be necessary.

1.6.3 NLC and the Sunspot Cycle

Vestine [1934][16] again postulated the possibility that increased NLC activity might correlate with reduced solar activity. When Vestine's paper was published, it appeared that the years 1887, 1899, 1911 and 1932 were times of the greatest reported frequency of NLC occurrence. These years were close to the solar minimum. As Vestine pointed out, this correlation with low sunspot activity was not entirely satisfactory, since the number of observers and interest over the years had fluctuated. Fogle [1965b][76] using a complete list of reported NLC sightings between 1885-1964, re-examined this question, and found no obvious correlation. Simmons and McIntosh [1983][77] have carried out an analysis of NLC over Western Europe during the years 1966-1982. They found two distinct minima in the years 1970 and 1980, and point out that Paton [1967][78] had noted a previous minimum extending over the years 1957 and 1958. They also found two maxima for NLC occurrence in the period considered, from 1964-1967, and 1974-1978. From their analysis, Simmons and McIntosh

supported an inverse relationship between annual sunspot number and frequency of NLC occurrence. The tentative conclusion is that the mesopause temperature is relatively higher at a maximum than at a solar minimum.

Thomas **et al.**, [1991][79] reported that Polar Mesospheric Clouds (PMC) occurrence rates are inversely dependent upon the solar UV flux, in particular solar Lyman-alpha (121.6 nm). The same general anti-correlation with solar activity is also present in NLC occurrence frequency [Gadsden, 1990][80]. Complicating the latter idea is the possibility that the mesopause temperature should vary directly with the solar cycle [Garcia **et al.**, 1984][81]. However, the evidence for solar cycle related temperature response is far from conclusive. The following questions still remain unanswered:-

i) does the occurrence of NLC vary with the solar activity?
ii) does the mean mesopause temperature vary with the solar activity?
iii) does the occurrence of NLC vary with mean mesopause temperature?
iv) with respect to ii), is the relationship direct or inverse?

1.6.4 NLC and Methane Concentration

Increasing concentrations of the greenhouse gas methane (CH_4), coupled with decreasing mesospheric temperatures, may also be related to the increased formation of NLC at 82 km over the cold

summertime polar caps. It is now believed that methane may facilitate the formation of NLC. Once methane reaches the mesosphere, photochemical reactions in the presence of atomic oxygen produces water vapour that is essential for the NLC formation.

Recently, Thomas **et al.**, [1989][82] discovered a historical record containing no reports of NLC displays before 1885 even though there were skilled observers monitoring the aurora and twilight phenomena at high latitudes, for many decades before then. Thomas **et al.**, [1989][82] suggested that the eruption of Krakatoa in 1883 may have influenced the discovery of the NLC two years later.

Since the discovery of NLC more than a century ago, the brightness, and frequency of these cloud displays have been increasing. Again, Thomas **et al.**, [1989][82] attributed the increasing cloudiness to the increasing methane, and hence, the increasing water vapour levels, while Gadsden [1990][80] attributes the cloudiness to the decreasing temperatures at the mesopause.

If the water vapour mixing ratio has been increased over the past century by 40% as Thomas **et al.**, [1989][82] suggested, the frost point will have increased by less than 2 K. Using a lapse rate in the vicinity of 82 km of 5 K km^{-1} [Lubken and von Zahn, 1991][83] the change in the frost point will correspond to a change in the mean height of the NLC base by ~0.4 km.

Such a change is undetectable when current methods of height

determination are utilized (by triangulation), but would presumably be measurable with laser investigation continued on a long term basis. Similarly, a reduction in the mesopause temperature of 7 K may cause a slight change in the lapse rate in the vicinity of the NLC, but the effect on the cloud height will be negligible.

Additional evidence of decadal changes in the mesopause density comes from the work of Clemesha **et al.**, [1992][84], who reported a statistically-significant lowering of the atmospheric Sodium layer between 1972, and 1987. The change in height by 0.7 km is consistent with a decrease of the temperature over Brazil of 5 K, although this is model dependent. Other studies suggest significant mesospheric cooling at non-polar latitudes.

1.6.5 NLC and Carbon Dioxide

Roddy [1986][66] raised the possibility that the increase in the "greenhouse effect" might alter the future behaviour of NLC. The effect on globally-averaged temperature, and composition has been discussed by Fels **et al.**, [1980][85], and more recently by Roble and Dickinson [1989][86]. Over most of the Middle Atmosphere, and the thermosphere, the effect of doubling carbon dioxide (CO_2) levels is to lower the future temperature. The change at the mesopause height is ~8 K. This is consistent with the results obtained from the general circulation model calculated heights (up to 80 km) by Rind **et al.**, [1990][87]. However, the latter results indicated a slight heating in the summertime mesopause

region from the CO_2 increases. This effect could be explained in terms of a changed circulation, which would counteract the direct radiative (cooling) effect. Moreover, this result is in doubt due to the local thermodynamic equilibrium (LTE) radiative heating algorithm [1]. This code becomes more accurate above the level of the vibrational levels located in the Upper Atmosphere.

Thomas (personal communication) examined the effect of a non-LTE radiative code, which is accurate throughout the mesosphere, and lower thermosphere. He adapted the "fixed-radiative heating" assumption [Fels **et al**., 1980][85]. This calculation assumes no dynamical response for a chemically-induced change in the radiative cooling/heating from the CO_2. The summertime polar regions warm up in response to an increase in the CO_2. The effect arises from the fact that the increased CO_2 opacity near the mesopause causes a greater net radiative exchange from the warm stratopause. This increased exchange occurs despite the lower upward flux leaving the cooler stratopause region. The average effect is to cool the upper mesosphere, in quantitative agreement with Roble and Dickinson [1989][86] result. However, this calculation indicates that local changes may be quite different from globally-averaged results.

In summary, the NLC height, which has not changed significantly from its canonical value of 83 +/- 1 km, since the nineteenth century [Gadsden and Taylor, 1992][88] indicates that neither the above mentioned changes in water vapour concentration nor in temperature would have altered the lower boundary of the

saturated mesopause region by more than a fraction of a kilometre. Therefore, the increase in the frequency of occurrence of NLC may arise from a change in the frequency with which structures appear (see Section 1.5.2). In general, the NLC are most readily detected when they show spatial structures, such as Bands, and Billows [Hines, 1991][89]. An apparent change in the frequency of occurrence of NLC would therefore come about through an increase in the frequency of occurrence of wave structures at the 82 km level. Many of the wave motions, particularly the larger scale Bands, seen in NLC, are thought to originate in the troposphere (see Chapter 6), so this increase in frequency could arise from the change in the prevailing winds in the stratosphere, and lower mesosphere acting as a filter to limit wave propagation.

Alternatively, there could be an increase in the number of potential wave sources. Either of these suggestions could explain the apparent increase in the number of NLC sightings without the necessity of there being a change in the greenhouse gases, and hence, an increase in the water vapour in the upper mesosphere.

1.6.6 NLC and Hydroxyl Layer

Measurements of the hydroxyl nightglow (altitudes ~ 87 km), provided an important method for remote sensing of the Middle Atmospheric dynamics in the vicinity of the mesopause. One such study, which has utilized hydroxyl measurements to investigate this region is the correlation of hydroxyl intensity and

rotational temperature with NLC occurrence (see Chapter 4). NLC often exhibit regular wave patterns similar to those seen in the hydroxyl emission [Taylor, 1986][50], (see Figure 1.7, and Table 1.2). Because of the relatively close proximity of the hydroxyl emission layer (halfwidth ~ 8 km), and mean height to the NLC layer, measurements of its temperature should therefore provide useful data on the atmospheric conditions during NLC displays. Previously, only two measurements of this type had been reported [Shefov, 1968; Harrison, 1973][90,91].

Harrison took spectrometric measurements of low elevation twilight sky from Calgary, Canada, during summer 1972. Although Harrison obtained spectra on several "cloud-free" nights, only one NLC display was observed. However, the average temperature and intensity for this night was similar to that of a "cloud-free" night, thus Harrison concluded that there was no significant change in the hydroxyl emission during the NLC. Shefov, on the other hand, reported a definite correlation between hydroxyl behaviour and NLC occurrence. These results remain in contention.

From the results discussed by Taylor **et al.**, [1989][92] there was no significant distinction between the mean of the hydroxyl temperature or its spread on NLC and other "cloud-free" nights. In particular the temperature was not observed to be significantly lower during NLC displays.

There seems to be no perceptible differentiation between the

intensity of the hydroxyl emission and its rotational temperature for nights containing NLC structures and all of the "cloud-free" nights.

Table 1.2

	NLC		OH		
	Bands	Billows	Broad Bands	Narrow Stripes	Ripples
height /km	82	82	85	85	85
number of waves	2-10	3-100	2-5	2-15	3-20+
average λ/km	20-50	5-10	100-?	20-70	5-15
speed ms^{-1}	10-60	50-100	_____	10-70	70-90
lateral extent /km	1-6 x 10^2	10-40	600	600	10-50
lifetime	1-3 hours	5-25 mins.	several hours	0.3-2 hours	45 mins.
amplitude	1.5-4	0.5-1	?	?	?

1.6.7 NLC and Aurora

Simultaneous occurrences of NLC and aurora were reported prior to 1965 [Smyth, 1886; Backhouse, 1886; Paton, 1951; Bryne, 1964; Schroder, 1965][93,94,95,96,97], and are still rarely reported events (see Figure 1.5C). However, during the summer of 1965, in a special observing programme, thirteen simultaneous displays of NLC and aurora were observed by the Geophysical Institute of the University of Alaska personnel [Fogle, 1966][7]. On seven out of thirteen occasions, the aurora occurred in the same region of the

sky as the NLC.

The alterations in brightness and form of the NLC after the onset of the aurora pointed towards a possible auroral-induced heating at or below the mesopause. Murcray [1957][98] gave experimental evidence pointing towards auroral-induced heating below the mesopause. The observed lower border of aurorae is generally found some 20-25 km above the clouds, ~100 km, and pulsating aurorae often extend down to 90 km. Moreover, it seems unlikely that heating at the mesopause could be caused by particle bombardment in-situ. Thus, the possibility of hydromagnetic wave heating at the mesopause could be ruled out, since this heating is confined to higher altitudes ~200 km [Chamberlain, 1961][99]. Maeda and Watanabe [1964][100] had suggested that pulsating aurorae were sources of infrasonic waves, and that these infrasonic waves may have provided some heating at and/or below the mesopause.

Today, however, it is clear that a NLC occurrence does not cause an affect on the aurora. They are initially independent phenomena [Schroder, 1965; Garcia et al., 1984; Gadsden and Schroder, 1989][97,81,45]. However, questions still remain unanswered:-

i) are NLC displays dispersed by auroral heating?
ii) do the auroral electrons accumulate for Polar Mesospheric Summer Echoes events?

1.6.8 NLC and Polar Mesospheric Summer Echoes

NLC and strong Polar Mesospheric Summer Echoes (PMSE) appear at similar heights, similar latitudes and at the same time of year. Both phenomena were observed several times during past campaigns [Taylor **et al**., 1989][91] using both VHF radars ie EISCAT, Tromso, Norway and rocket launches. However, few simultaneous occurrences have been observed. These observations indicate that NLC and PMSE are not closely correlated, at least during late summer (U.-P.Hoppe, personal communication).

As radar echoes are much too strong to be attributed to incoherent scatter, they must be semi-coherent, and therefore require the existence of structure with scale sizes as small as 0.67 m (half the EISCAT VHF λ). This is normally considered to be within the viscous subrange at mesospheric altitudes. One suggested scattering mechanism depends on the presence of large positive ions that can form in the very cold environment at the polar summer mesopause. These ions are probably cluster-ions.

Hydrated cluster-ions may also play a significant role in the generation of NLC [Witt, 1969; Arnold, 1980; Turco **et al**., 1982][63,68,61]. A significant fraction of the NLC particles may carry a negative charge, thus constituting another class of heavy, low mobility ions. As a consequence of electron attachment to aerosol particles, local depletions in the electron population may be expected in the vicinity of NLC.

Since PMSE and NLC are both manifestations of dynamics and chemistry of the summertime mesopause region and might then be related, a simultaneous investigation of the two phenomena is appropriate (see Figure 1.8). Visual and photographic observations were therefore made from Northern Sweden during the EISCAT PMSE campaign, in 1988 and the NLC-91 campaign, in 1991.

Ground-based remote sensing techniques are being further developed and hold enormous promise for the future. Radar studies are already revealing fascinating dynamical and turbulent regimes in the form of PMSE. It is hoped that the question of the relationship of PMSE and mesospheric clouds will be resolved in the near future. Lidar probing of the region is quite promising. The technology has matured to the point where bright mesospheric ice clouds are detectable [Hansen **et al**., 1989; Meriwether, 1991][22,101]. More powerful lidars are being currently developed which will have additional capability (see ALOMAR, Chapter 7).

1.6.9 NLC and Polar Mesospheric Clouds

The Ultraviolet Spectrometer Experiment (UVS) on board the Solar Mesosphere Explorer (SME) spacecraft has been recording scattered light from thin layers in the high latitude summertime mesosphere since its launch on October 6, 1981. The phenomenon known as Polar Mesospheric Clouds (PMC) were first observed by Donahue **et al**., [1972][102] from a satellite (OGO-6) photometer measurement at visible wavelengths. PMC were believed to be a brighter extension of the more familiar NLC phenomenon into the daylit polar cap,

approximately 65°-90° latitude [Jensen and Thomas, 1988][103]. The Southern Hemisphere clouds are brighter, because they are viewed at smaller scattering angles than the Northern Hemisphere PMC. Also, Northern Hemisphere clouds are somewhat higher (85 km) than the Southern Hemisphere clouds (83 km).

The ultimate goal for PMC studies is the measurement of both clouds and the underlying Rayleigh-scattered background under both daytime and night-time conditions. This achievement would probably change many of our long-held ideas about PMC/NLC and would usher in a powerful new way of studying the upper mesospheric region, not only in polar regions, but all latitudes.

New satellite studies of PMC/NLC are needed. There is a growing body of evidence that Northern and Southern PMC are inherently different [Olivero and Thomas, 1986; Thomas and Olivero, 1986; Thomas **et al.**, 1991][104,105,106]. Moreover, PMC appear to be brighter and are found at a greater height in the North than in the South. Jensen and Thomas [1988][103] speculated that this may be due to the presence of more water vapour in the Northern Hemisphere summer polar region. It is known that the Northern Hemisphere methane content is higher by approximately 8% than the Southern Hemisphere. The above statements lead to serious implications for NLC displays both in the Northern Hemisphere and the Southern Hemisphere.

With new measurements now becoming available from UARS (Upper Atmosphere Research Satellite) and the proposed NASA Space

Physics Division TIMED (Thermosphere, Ionosphere and Mesosphere Energetics and Dynamics) satellite (see Chapter 7, Future Work), the properties of NLC and PMC and their environment will be better understood in the near future.

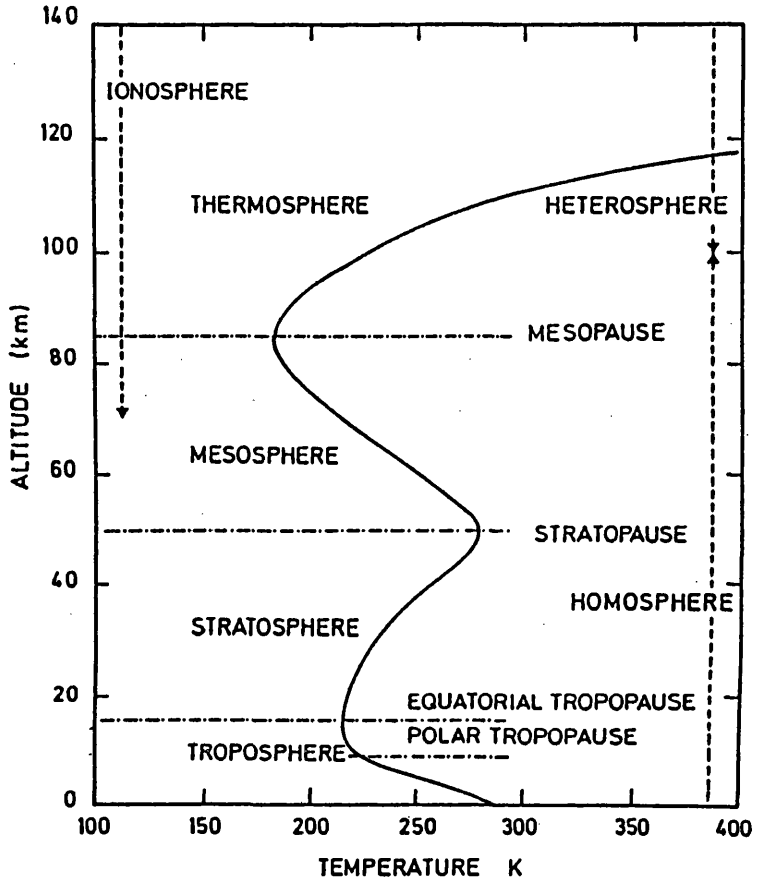


Figure 1.2a

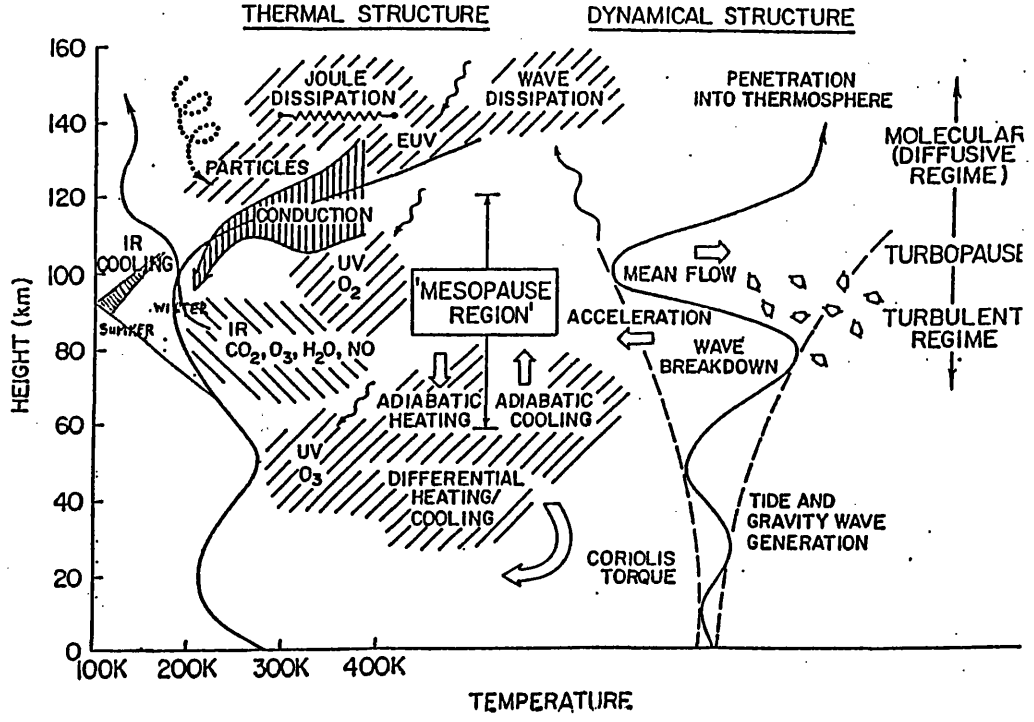
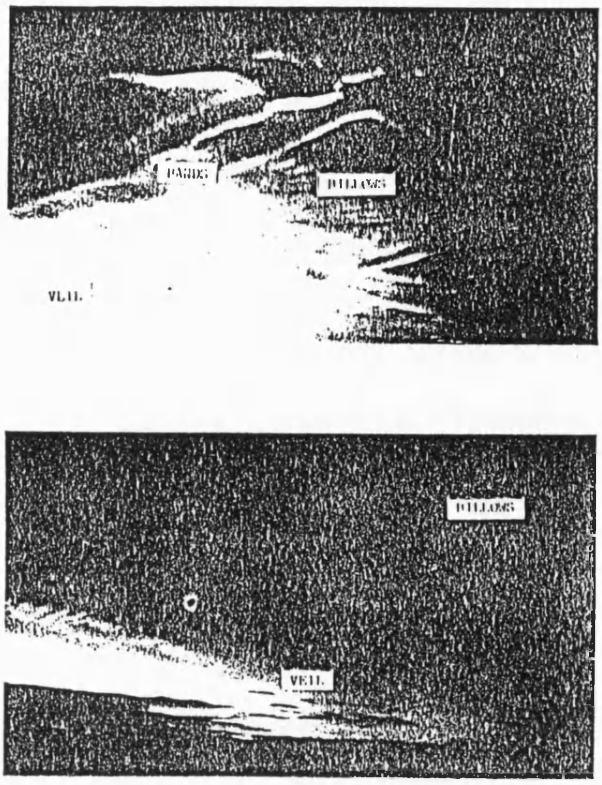


Figure 1.2b



Figures 1.4a and b

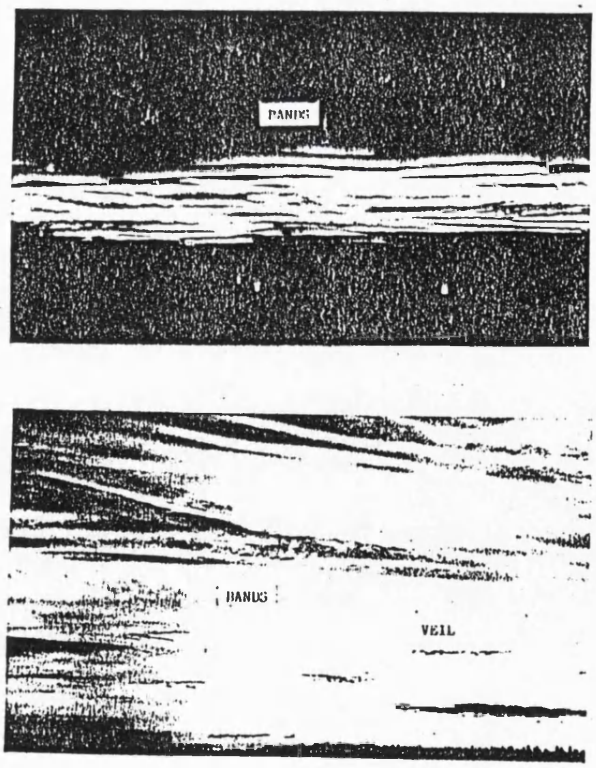
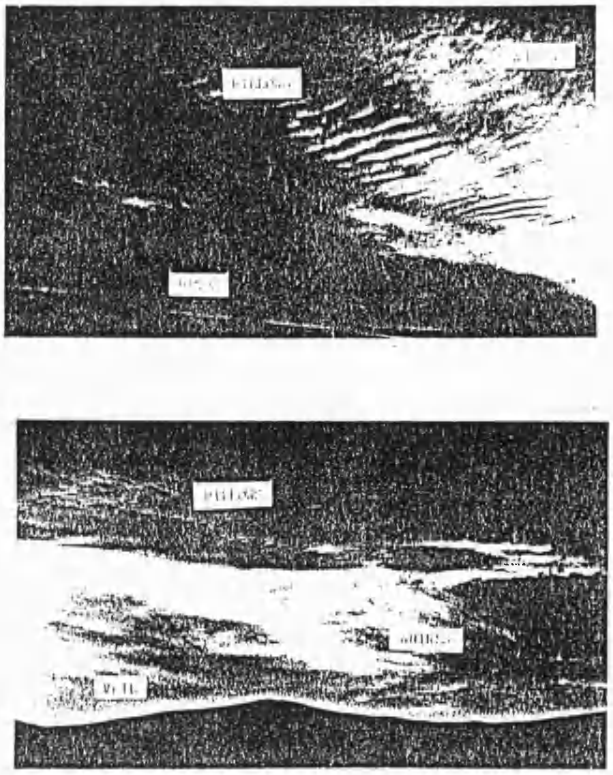


Figure 1.4c and d



Figures 1.4e and f

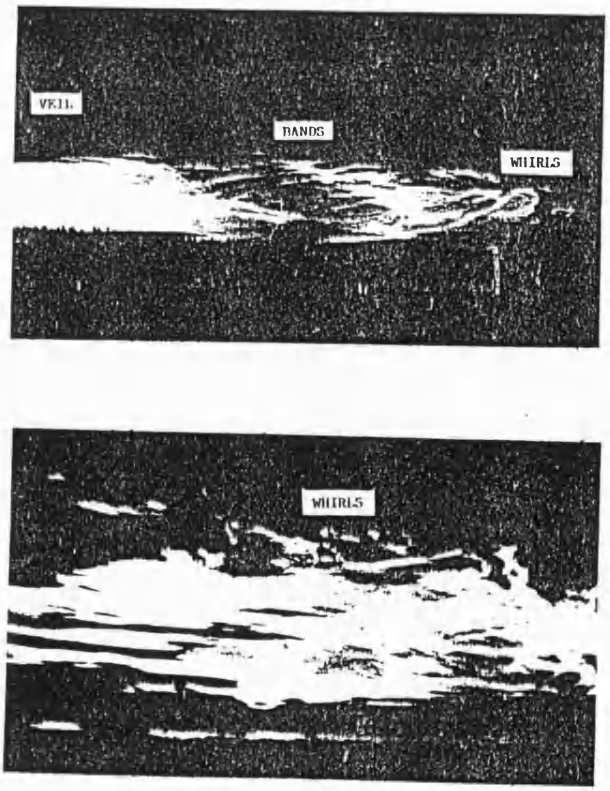


Figure 1.4g and h

(A) 15/16 JULY 1965 DISPLAY
Grande Prairie, Alberta

Veil (Type I)
Bands (Type II.b)
Intensity 4.

(B) 26/27 JULY 1965 DISPLAY
Watson Lake, Yukon Territory

Veil (Type I)
Bands (Types II.a,II.b)
Billows (Type III.a)
Intensity 4.

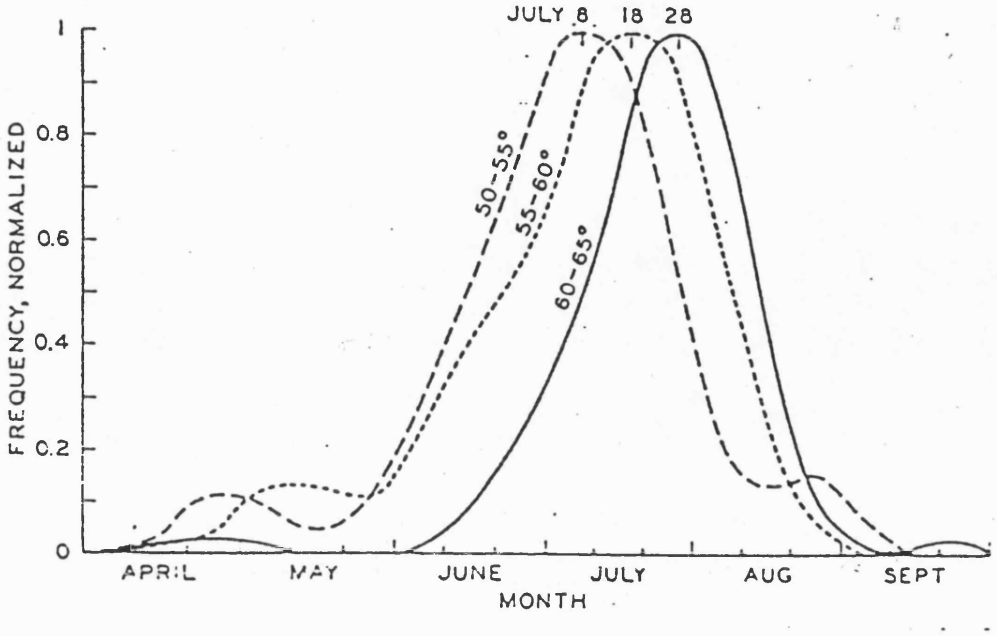
(C) 27/28 JULY 1965 DISPLAY (WITH AURORA)
Watson Lake, Yukon Territory

Veil (Type I)
Amorphous (Type V)
Intensity 2.

(D) 3/4 JANUARY 1966 DISPLAY
Punta Arenas, Chile

Veil (Type I)
Bands (Type II.a)
Intensity 2.

Figure 1.5 Unusual noctilucent cloud displays



Figures 1.6

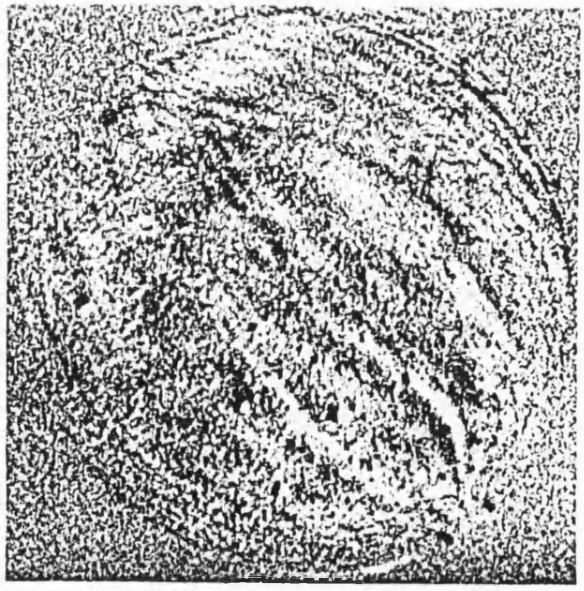


Figure 1.7

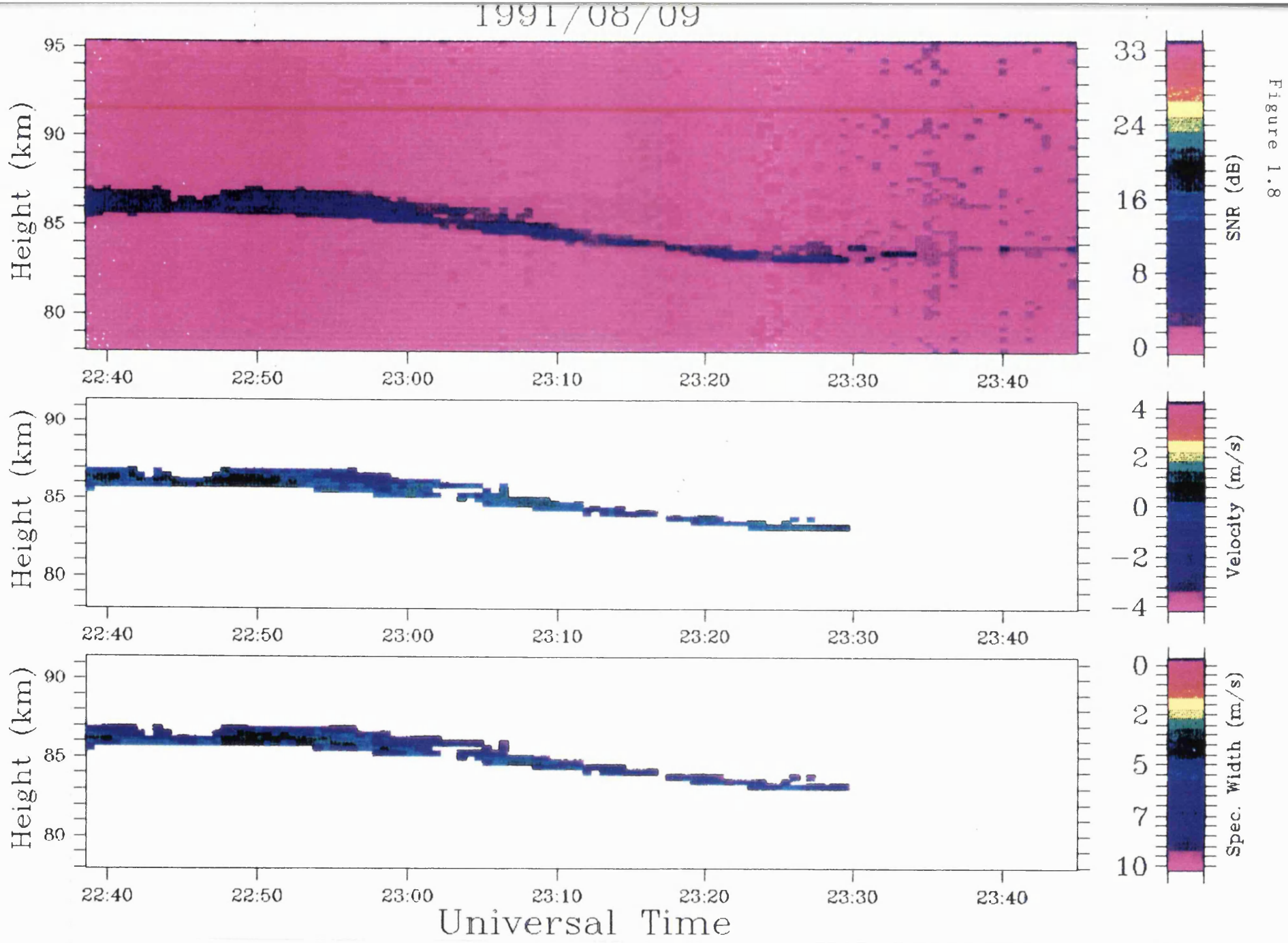

1991/08/09
Height (km)
95
90
85
80
22:40
22:50
23:00
23:10
23:20
23:30
23:40
33
24
16
8
0
SNR (dB)
Height (km)
90
85
80
4
2
0
−2
−4
Velocity (m/s)
Height (km)
90
85
80
0
2
5
7
10
Spec. Width (m/s)
Universal Time

Figure 1.8

Chapter 2

2 Rocket Instrumentation

2.1 Introduction

This chapter describes the APL rocket-borne instrumentation flown by the APL during the absence and presence of NLC. The same photodetectors were flown during the Middle Atmosphere ElectroDynamics (MAED) Campaign in 1986 and the recent Noctilucent Cloud 1991 (NLC-91) Campaign, launched from Esrange, Kiruna, Sweden, respectively.

Measurement of an aerosol layer structure is achieved using a non-imaging diode photometer consisting of a large area Silicon (Si) p-i-n photodiode. This device is compact, designed for low photo-conductive operation, and it is discussed in Section 2.2.

The principles of operation of a p-i-n photodiode are discussed in Section 2.3. With the success of the high sensitivity over a wide spectral range, fast response, low noise and excellent long-term stability of the Atmospheric Physics Laboratory, University College London photometers, the data collected yields optimum results.

Data collection and analyzing routines are discussed in Section 2.4. These techniques yield spatial morphologies observable by the photometers.

Finally, Section 2.5 summarizes the problems encountered during data collection and data analysis.

2.2 The p-i-n Photodiode

In general, the photometer consists of a detector with a filter whose spectral transmittance modifies the detector spectral responsivity to give an overall responsivity close to the $V(\lambda)$ function. The $V(\lambda)$ function is the internationally agreed function representing the spectral response of the average human eye.

Figure 2.1 illustrates the final set up of the instrumentation flown on the rocket payloads during the 1986, and 1991 NLC rocket campaigns. It should be noted that the lenses included are purely for focusing the incoming light onto the illumination window. The filter was present only in one of the three photometers at each launch during the latter rocket campaign. The intention was to measure the atomic oxygen emission intensity. In general, the APL photometers are narrow scanning instruments which scan the entire sky laterally from the rocket payload utilizing the payload spin. As a result of the spin, and coning motions of the payload, these instruments provide a two-dimensional imaging of the NLC layer/s from below when between the altitude regime of 60-80 km, and from above when the payload is over 90 km.

2.2.1 Silicon Photodiode

Due to its abundance, and relatively simple preparation of extremely pure, high-quality single crystals, such as Silicon (Si), it is thus an inexpensive material with a quasi-perfect lattice. Additionally, the fabrication of Si is most advanced, and so is passivation (a coating to protect the circuit from contamination), and surface stability. The electrical, and thermal properties are excellent ie fast response time and small temperature coefficient respectively.

The Si photodiode is one of the simplest and cheapest forms of photodetectors. Figure 2.2. illustrates a Si p-i-n photodiode most commonly used commercially. The diode is fabricated by diffusion of a highly doped top contact into a low doped epitaxial drift region based on a thick substrate material, if a short to medium drift region is the objective ($\leq 50\ \mu m$). Diodes with drift regions in excess of 100 μm are usually made from highly resistive material, into which both contact regions are diffused. As a diffusion mask for the top contact, which defines the diode area, is thermally grown Silicon Dioxide (SiO_2) or Silicon Nitride (SiN) is used, which also provides the surface passivation. SiO_2 is used as an anti-reflection coating $n_{ac} \approx 1.45$ on the illumination window. A contact ring overlapping the SiO_2 mask reduces the fringing fields and surface currents. It generates a depletion at the periphery of the p-n junction. The p-n junction is the region near the junction which becomes clear of holes and electrons (see Figure 2.2) as a result of the

initial migration. This is called the depletion region, and is typically less than 1 μm wide. The diffused contacts are usually a few microns thick to avoid carrier injection from the contact, and to keep the electric field strength at the curved periphery sufficiently low to prevent pre-breakdown there. The thickness of the top contact layer is conveniently small compared to the light penetration depth, and the diffusion current effects are still tolerable.

The p-i-n time response usually exceeds 1 ns both for pulse rise, and fall time. Peak quantum efficiencies of 60-85% are typical. Commercial diodes frequently have large surface areas compared to that of a fiber core. As mentioned in Section 2.3, the dark current is partly due to the surface currents, and edge break down. As shown in Figure 2.2, the guard rings or channel stoppers may be introduced in the same process, which generates the n^+ (heavily doped n-type region) contact layer. If an appropriate bias is applied, as indicated in Figure 2.2, these currents are shunted around the load resistor.

The time response may be improved without sacrificing the quantum efficiency. There has been three approaches to achieve this:-

a) metallic reflection of the light at the bottom contact,
b) side illumination,
c) multiple reflections incorporating both metallic and total reflections.

The author will only be discussing option b). However, further details of the reflective-type diodes and multiple reflective-

type diodes can be found in the literature [Lee and Sze, 1970; Muller, 1978][107,108].

2.2.1.1 Side Illumination

The quantum efficiency may be further increased if a large area diode is side -illuminated. As shown in Figure 2.2, a parallel light beam is incident directly on the drift region of the diode, or it enters the active region via multiple reflections. The first approach requires a very careful alignment of a parallel beam. If it is not parallel in itself, and with respect to the junction, or misaligned in the vertical direction, the optical power will enter the highly doped regions, and diffusion current effects may be significant. The alignment of such a diode requires a system of focusing lenses, and a drift region width exceeding the diode diameter ~100 μm. Hence, the diode response time will be of the order of 1 ns. As shown in Figure 2.3, the spectral response of such a diode may be flat up to a wavelength of 1.1 μm. Easier alignment is possible in the reflection-type diode, however, fabrication of this diode structure is more complicated, as it requires a steep, polished side wall. The latter problem gives rise to yet a further complication - the deposition of an anti-reflection coating. Thus, the diode performance will deteriorate, and surface leakage currents will be generated.

2.3 Principles of Operation

This section describes the operational mode of the Silicon (Si) p-i-n photodiodes, employed by the APL, and briefly considers the general characteristics of quantum efficiency, response speed, time response, and noise.

As aforementioned, the p-i-n photodiode consists of an undoped (intrinsic) semiconductor region sandwiched between two highly doped p, and n contact junctions. The incident light penetrates this diode through a window (see Figure 2.2) in the top metal contact. It is absorbed when passing the semiconductor region, and electron-hole pairs are generated. These carriers are separated in the electric field supplied by the reverse bias to the diode, and thus induce a current in the electrical circuit.

2.3.1 The Absorption Process

Of the three prime ways of light absorption in semiconductors - intrinsic, extrinsic, and free-carrier absorption (see Figure 2.4) - the photovoltaic process usually employs the intrinsic, as both types of generated charge carriers have to be mobile, so that they can separate, and induce an electric current, or voltage. Intrinsic absorption involves the transition of an electron from the valence band to a free state in the conduction band, it requires the optical frequency υ of the incoming light to be:-

$$h\upsilon \geq E_g \qquad (2.1)$$

or

$$\frac{h\,c}{\lambda} \geq E_g \qquad (2.2)$$

where E_g = the energy gap

h = the Planck constant = 6.63×10^{-34} Js

c = the velocity of light

λ = the wavelength of the light

The cutoff (or longest) wavelength λ_0 subject to intrinsic absorption at the absorption edge of the semiconductor is obtained for the equality in Equation (2.2):-

$$\lambda_0 = \frac{h\,c}{E_g} \qquad (2.3)$$

For wavelengths shorter than λ_0 the absorption, characterized by the absorption coefficient α, rises more or less sharply depending on the band structure of the semiconductor material. Silicon is an indirect semiconductor material (see Figure 2.5). The absolute conduction band minimum is offset from the valence band maximum by a momentum. Therefore, a transition from the

valence to the conduction band in addition to a photon of energy exceeding E_g involves a phonon to supply the electron with the necessary momentum. As this coincidence of the photon, and phonon absorption is much less probable, the rise in the absorption coefficient with increasing photon energy is accordingly smoother. If this energy finally exceeds the energy gap of the direct transition, Silicon turns into a direct material also, with a steep increase in α.

Depending on α, the incoming light intensity Ξ_0, which penetrates the illumination window, decays along its path through the semiconductor material. This decay can be written as:-

$$\Xi(x) = \Xi_0 \exp(-\alpha x) \qquad \textbf{(2.4)}$$

Instead of the absorption coefficients, its inverse, $1/\alpha$, the light penetration depth is often used. According to Equation (2.4), at this depth, the intensity has fallen to a fraction 1/e of its initial value. From Equation (2.4), the fraction of an optical power, absorbed along a distance w, is expressed as:-

$$\frac{\Xi_{abs}}{\Xi_0} = 1 - \exp(-\alpha w) \qquad \textbf{(2.5)}$$

which approaches 1 for absorption widths considerably larger than

the light penetration depth (see Figure 2.6).

2.3.2 The Photovoltaic Effect

If light is absorbed in a semiconductor device, which contains a p-n junction like the p-i-n diode of Figure 2.2, the optically generated carriers are separated in the junction region because of the electric field there. The carriers traversing the junction creates a current, the short-circuit photocurrent i_{ph}, which shifts the current-voltage characteristic of the diode, as indicated in Figure 2.7, is given by:-

$$i = i_0 \left[\exp\left(\frac{q\,v}{n\,k_B\,T}\right) - 1\right] - i_{ph} \qquad \textbf{(2.6)}$$

where i_0 = the reverse dark current of the diode

v = the applied voltage

n = the diode ideality factor ($1 \leq n \leq 2$)

k_B = the Boltzmann constant = 1.38×10^{-23} JK^{-1}

The diode is open-circuited, thus a photogenerated voltage v_{ph} develops:-

$$v_{ph} = \frac{n\,k_B\,T}{q} \ln\left(\frac{i_{ph}}{i_0} + 1\right) \qquad \textbf{(2.7)}$$

Photodiodes, as used by the APL, for optical communication are used neither in the short- nor the open-circuit mode, because they are to respond to weak, and rapidly varying optical signals. In fact they are reverse-biased, and thus function as a current source i:-

$$-i = i_0 - i_{ph} \qquad (2.8)$$

with a source conductance G_d given by:-

$$G_d = \frac{d\,i}{d\,v_B} \qquad (2.9)$$

where v_B = the reverse bias voltage

In order to guarantee a small source conductance with a correspondingly linear and sensitive response also for weak photocurrents, the dark current of the diode ought to be small and independent of the voltage.

Back-biasing the diode is advantageous for the following three reasons:-

i) The photocurrent remains a linear function of the generated carriers up to relatively high optical powers and associated currents even at finite load resistances.

ii) The capacitance and the resistance of the diode are minimized, because the depletion region within the diode extends.

iii) The bias establishes a high electric field throughout the i layer of the p-i-n diode or at least, if a p-n diode with not so lightly doped i region is used, within an appreciable part of the diode. Thus, the carriers travel at higher velocities and are gathered from a larger part of the diode than without bias, which improves both the time response and increases the number of carriers contributing to the current.

2.3.3 Quantum Efficiency and Responsivity

When optical power P = ΞA, where A is the illuminated area, hits the illumination window of the diode, it will generally not be completely absorbed in the semiconductor, and generate electron-hole pairs. In addition, not all of the generated carriers will contribute to the photocurrent i_{ph}. The ratio of the number of charge carriers contributing to the current, and the number of photons hitting the diode surface per second is denoted as quantum efficiency η:-

$$\eta = \frac{\frac{i_{ph}}{q}}{\frac{P}{h\nu}} = \frac{hc}{q\lambda}\,\frac{i_{ph}}{P} \qquad (2.10)$$

where q = the electronic charge

η = the effective or external quantum efficiency ≤1

Another expression commonly used is the responsivity, R:-

$$R = \frac{i_{ph}}{P} \qquad \textbf{(2.11)}$$

by which the photocurrent to be expected from a certain power can be computed. Substituting Equation (2.10) into Equation (2.11) results in the following expression:-

$$\eta = \frac{hc}{q\lambda} R \qquad \textbf{(2.12)}$$

Both η, and R are functions of the diode material, geometry, wavelength, and temperature. The quantum efficiency would approximate 1 very closely if all the optical power were absorbed in the lightly doped i layer of width w_i. The carriers are then separated immediately after their generation by the applied electric field and swept out of the depletion region. To this end, w_i must be large as compared to the light penetration depth $1/\alpha$, and the drift time of the carriers through this layer must be short compared to their lifetime. The latter is nearly always true even at relatively low electric field strengths.

However, there are a variety of physical effects which contribute to a reduction of quantum efficiency or responsivity, respectively. The author will evaluate these effects, as the path of penetrating light is followed through the diode (see Figure 2.8):-

i) Because of the high refractive index n_s of the semiconductor, an amount of light, Ξ_r, is reflected at the air-semiconductor transition. The reflectivity R_n is the magnitude of the reflection coefficient r squared:-

$$R_n = |r|^2 = \frac{(n_s - n_0)^2 + k_s^2}{(n_s + n_0)^2 + k_s^2} \qquad \textbf{(2.13)}$$

where

$$k_s = \frac{\lambda}{4\ \pi\ \alpha} \qquad \textbf{(2.14)}$$

is the imaginary part of the refractive index, and n_0 is the refractive index of air.

Since n_s is between 3 and 4 for most semiconductors, about one-third of the incident light is reflected. These reflection losses can be avoided by choosing an appropriate anti-reflection coating, usually a quarter-wavelength layer of a medium with the

required refractive index n_{ac} and transparency at the wavelength region of interest:-

$$n_{ac} \approx \sqrt{(n_s\, n_0)} \qquad \textbf{(2.15)}$$

where n_0 = the refractive index of air .

SiO_2 (Silicon Dioxide) is used for an anti-reflection coating with a refractive index n_{ac} in the range of 1.5-2.2.

ii) When the light of intensity Ξ at a wavelength in the usable range of the semiconductor strikes the diode, firstly, it is absorbed in the top, highly doped contact layer. Since there is no electric field in this layer, the generated minority carriers - electrons in the p^+ layer, then holes in the n^+ layer - diffuse toward the depletion region of the diode. Depending on the thickness of this top layer w_T, and the absorption coefficient α, a fraction of intensity:-

$$\frac{\Xi_T}{\Xi_0} = T_n\, [1 - \exp\, (-\alpha\, w_T)] \qquad \textbf{(2.16)}$$

with T_n = (1 - R) for no anti-reflection coating (solid curve in Figure 5.8), and T_n = 1 with an ideal anti-reflection coating (dashed curve in Figure 2.8), respectively, are absorbed there.

iii) The most desired situation exists when all the incident light is absorbed within the depletion region. The fraction of power absorbed is then:-

$$\frac{\Xi_I}{\Xi_0} = T_n \exp(-\alpha\, w_T)\ [1 - \exp\ (-\alpha\, w_I)] \qquad \textbf{(2.17)}$$

iv) If the sum of w_T and w_I is not large compared to $1/\alpha$, the incident light will penetrate into the substrate and again, will generate minority carriers, which will diffuse to the drift region. The absorbed intensity Ξ_s takes the following form:-

$$\frac{\Xi_s}{\Xi_0} = T_n \exp\ [-\alpha\ (w_T + w_I)]\ [1 - \exp\ (-\alpha\, w_s)] \qquad \textbf{(2.18)}$$

If the substrate thickness is large as compared to $1/\alpha$, all of the remaining light is absorbed there (last term in the brackets is zero). The absorption of optical power in the substrate, and the associated recombination loss might seem to be tolerable for three reasons:-

1) If $(w_T + w_I)$ is larger than $1/\alpha$, the light intensity entering the substrate is decreased considerably.
2) As the substrate doping usually is not as high in the diffused contacts, both the minority carrier lifetime, and diffusion

coefficient are relatively large. Therefore, the diffusion length governing the distance, wherein the carriers will reach the drift region, will be relatively long (≥10 μm in Silicon).

3) Because of the relatively thick substrate, surface recombination is insignificant.

2.3.4 Time Response

Due to the distributed generation of the carriers within the diode, and to the different mechanisms of their transport, the time response of the photodiode depends on the wavelength, semiconductor material (Si, in this instance), diode geometry, and load impedance of the electrical circuit. The response can be estimated from the time constants involved, and the fraction of carriers, the response of which they govern. A typical response to a short optical pulse is illustrated in Figure 2.9. Besides, the form of the optical pulse, three phenomena shape the electrical response, namely 1) the carrier drift in the depletion region, 2) the carrier diffusion in the non-depleted region, 3) the RC constant of the circuit impedance, and diode capacitance.

2.3.4.1 Carrier Drift Time Constant

Whenever an electron or hole of charge q drifts with constant velocity through the depletion layer of the photodiode, it gives rise to a conduction current i_{dr} in the electrical circuit during its drift time t_d:-

$$i_{dr} = \frac{q}{t_d} \qquad (2.19)$$

The distance travelled by the carrier, and the drift velocity v_d determine t_d. The velocity v_d depends on the applied electric field strength F as expressed by:-

$$v_d = \mu F \qquad (2.20)$$

where μ = the carrier mobility, until a drift saturation velocity v_s at a drift saturation field strength F_s. However, if the carriers are generated at the edge of the i layer or are injected from the highly doped contact regions, they have to transverse the whole of the depletion layer w_I. The drift time then becomes:-

$$t_{dpn} = \frac{w_I}{\mu_{pn} F} \qquad (2.21)$$

for $F \leq F_{spn}$

and

$$t_{dpn} = \frac{w_I}{v_{spn}} \qquad (2.22)$$

for $F \geq F_{spn}$

where p = the holes

n = the electrons

From Equations (2.19), (2.21), and (2.22), it is clear that a high electric field strength results in a short drift time and a high conduction current.

If the incident optical power is sinusoidally modulated as a frequency ω, then so will the photocurrent. The carrier drift time will delay this current and at higher frequencies also weaken it. The frequency response (FR), is given by:-

$$FR = \frac{1 - \exp(-j\,\omega\, t_{dpn})}{j\,\omega\, t_{dpn}} V \qquad (2.23)$$

If the electrons and holes are generated within the i layer, they are separated and, depending on the applied field, will drift in opposite directions at unequal velocities. The time response of the diode will be faster in this case, as the average distance between carriers have to travel is less than w_I. The frequency

response takes the form:-

$$FR = \frac{1 - \exp(-j\,\omega\,t_{dn})}{(\omega\,t_{dn})^2} + \frac{1 - \exp(-j\,\omega\,t_{dp})}{(\omega\,t_{dp})^2} + \frac{1}{j\,\omega\,t_{dn}} + \frac{1}{j\,\omega\,t_{dp}} \qquad \textbf{(2.24)}$$

Due to a finite absorption constant α, the carrier distribution in the i layer decays exponentially. Whereas this decay changes the electrical response very little for equal carrier velocities, it effects it considerably if they are different. Figure 2.10 schematically depicts the pulse responses for illumination from the direction in which either the faster (solid line), or slower (dashed line) carriers drift to. The latter is preferred for a fast time response, as a larger fraction of slow carriers has to travel only a short distance. As holes need a higher electric field strength to saturate than electrons do, this is of importance for the design of low-voltage diodes, which evidently should have the p^+ contact on top.

2.3.4.2 Diffusion Time Constant

Carriers which are generated in the highly doped contact regions or in other parts of the diode without electric field, but close enough to the drift region, have to diffuse to the drift region before they contribute to the photocurrent. If an appreciable amount of the optical power is absorbed there, this carrier diffusion will cause a time spread of the carriers reaching the

drift zone. Therefore, the associated diffusion time constant adds to the carrier drift time, and may dominate the frequency response.

2.3.4.3 The RC Time Constant

In addition to the delay mechanisms inside the diode, the imbedding network, and diode housing in combination with the diode impedance may limit the time response of the diode. Figure 2.10 represents the simplified equivalent circuit of the p-i-n photodiode imbedded in the detection circuit. It includes the current source i_{ph} with its conductance G_d, the depletion layer capacitance C_d in parallel, followed by the diode series resistance R_s, and the load impedance R_L. the package capacitance C_p, and inductance L_p are indicated by dashed lines.

The main elements governing the time response are the depletion layer capacitance of the diode, and the series and load resistances R_s and R_L. The resulting RC time constant t_{RC} takes the form:-

$$t_{RC} = (R_s + R_L)\ C_d \qquad \textbf{(2.25)}$$

For this response to be fast, R_s is small, usually below 10 Ω, and R_L should be low in high speed circuits ≈ 50 Ω. Moreover, the depletion layer capacitance must be reduced to a minimum.

If the diffusion time constants govern the time response, then the optimum design of the p-i-n photodiode with respect to the time response depends strongly on the light penetration depth in the semiconductor material at the received wavelength, but depends also on its electrical properties, and on the diode geometry.

If the light penetration depth is relatively large, a reduction of the i layer width below the light penetration depth in order to speed up the carrier drift time constant will increase the amount of light absorbed in the substrate, and result in a slowly decaying diffusion current.

For frequencies considerably lower than $1/2\pi t_d$, where t_d is the time constant, this diffusion current will contribute to the quantum efficiency. However, at higher frequencies, it will give rise to an underlying dc component, and contribute negligibly to the rf photocurrent, and hence, reduce the quantum efficiency. A process to speed up this diffusion process in the substrate is to apply a relatively high built-in field strength. This can be achieved by out diffusion of the substrate doping into the epitaxial drift region. Table 2.1 summarizes the design criteria for a single material, fast, and sensitive p-i-n photodiode. The trade-off between time response and quantum efficiency with respect to the depletion layer width, becomes obvious by combining Equations (2.17), (2.21), and (2.22), and ignoring reflection, and absorption in the top contact layer results in the following expression:-

$$t_d = \frac{1}{\alpha \, v_s \ln(1 - \eta)} \qquad (2.26)$$

Table 2.1

Property	Short Time Response	High Quantum Efficiency
drift layer	short (drift time) long (capacitance)	long ($w_I \gg 1/\alpha$)
drift layer doping	low (F_s throughout)	low (recombination)
top contact layer	thin	thin
surface recombination	high	low (low frequency)
substrate recombination	high	low (low frequency)
area	small (capacitance)	large (beam diameter)

2.3.5 Noise in p-i-n Photodiodes

The minimum optical power a photodiode can detect is limited by noise. The noise originates in a variety of sources, and their individual contribution depends on the wavelength, diode design, and the semiconductor material, and the electrical circuit. In the wavelength region of interest for optical communication, the detection is limited either by thermal or shot noise. Signal- or background-limited detection hardly applies, as the modulation bandwidth is neither so narrow nor so extremely wide and the background illumination can readily be kept small.

Figure 2.11 illustrates the generalized photodetection process

of a photodiode. Since the diode resistance has little effect, only the shot noise is generated in the photodiode. This noise is due to the photogenerated currents of the signal P_s ($\sqrt{i_s^2}$), and background illumination P_b ($\sqrt{i_b^2}$), and the diode reverse dark current ($\sqrt{i_d^2}$). Applying Equations (2.10), and (2.11), the photogenerated noise currents assuming white noise takes the form:-

$$\langle i_s^2 \rangle = 2\, q\, i_s\, B = \left(\frac{2\, q^2\, \eta}{h\, \nu}\right) P_s\, B = 2\, q\, R\, P_s\, B \qquad \textbf{(2.27)}$$

$$\langle i_b^2 \rangle = 2\, q\, i_b\, B = \left(\frac{2\, q^2\, \eta}{h\, \nu}\right) P_b\, B = 2\, q\, R\, P_b\, B \qquad \textbf{(2.28)}$$

where B = the bandwidth of the detection circuit

Accordingly, the dark current noise source is:-

$$\langle i_d^2 \rangle = 2\, q\, i_d\, B \qquad \textbf{(2.29)}$$

The dark current depends strongly on the diode material, geometry, passivation, and fabrication procedure. In general, it comprises of three components, the individual height of which may differ by orders of magnitude. Theses components are:-

1) The carrier generation-recombination (G-R) current in the drift zone:-

$$i_{dgr} = \frac{A\, n_i\, w_I}{\tau_I} \qquad (2.30)$$

where n_i = the intrinsic carrier concentration

τ_I = the minority life-time in the i layer

2) The minority carrier diffusion current of the highly doped contact regions, which form the potential barriers in the p-i-n diode:-

$$i_{di} = A\, q\, n_i^2 \left(\frac{1}{N_D}\sqrt{\frac{D_p}{\tau_p}} + \frac{1}{N_A}\sqrt{\frac{D_n}{\tau_n}}\right) \qquad (2.31)$$

where N_D = the donor concentration in the highly doped regions

N_A = the acceptor concentration in the highly doped regions

3) The reverse current flowing across the surface, and periphery of the diode i_{ds}. This current is commonly termed as the surface leakage or leakage current. As it depends strongly on the diode preparation, material and geometry, no definite expression can be given.

However, in order to obtain the total dark current:-

$$i_d = i_{dgr} + i_{di} + i_{ds} \qquad (2.32)$$

the minority carrier lifetimes should be as long as possible, which contradicts the claim for a short time response. However, since the lifetimes and diffusion coefficients decrease with increased doping, a high doping concentration in the contact layers is favoured for a low diffusion current. This demand is prevents thin contact layers being possible, which according to Equation (2.18) guarantees a short time response and a high quantum efficiency. The doping of the drift region, on the other hand, should be as low as possible for a small G-R current.

A low dark current requires the smallest diode geometry possible with respect to quantum efficiency. A small leakage current demands good passivation, low surface recombination velocity, and restricted peripheral regions. As i_{ds} flows across the periphery of the diode, it varies approximately as the square root of the area.

In general, the most significant noise contribution in the broad-band detection circuit incorporating the p-i-n photodiodes is the thermal (or Johnson) noise i_L^2 of the load input resistance R_L:-

$$\langle i_L^2 \rangle = \frac{4\ k_B\ T\ B}{R_L} \qquad \textbf{(2.33)}$$

Another, significantly smaller, noise contribution $\langle i_B^2 \rangle$ arises from the bias resistors R_B, as:-

$$R_L \ll R_B \qquad \textbf{(2.34)}$$

usually holds:-

$$\langle i_B^2 \rangle = \frac{4\ k_B\ T\ B}{R_B} \qquad \textbf{(2.35)}$$

Introducing the effective resistance to be:-

$$R_e = \frac{R_L\ R_B}{(R_L + R_B)} \qquad \textbf{(2.36)}$$

Then, combining Equations (2.33), (2.35), and (2.36) yields the thermal noise source expression:-

$$\langle i_{th}^2 \rangle = \frac{4\ k_B\ T\ B}{R_e} \qquad \textbf{(2.37)}$$

Only if the signal is strong enough to compete with the noise, can it be detected. Measures for weighing the effect of noise on signal transmission are the signal-to-noise ratio (SNR) for analog signals, and the bit error rate (BER) for digital signals. Whereas the SNR can be derived in a straightforward manner, the BER depends strongly on the pulse shape, and pulse length compared to the repetition rate, on extinction rate, code, and equalization in a rather complicated manner. However, the author will restrict the discussion to the SNR, which is the signal power at the output of the detection circuit divided by the average noise power:-

$$\frac{S}{N} = \frac{\frac{1}{2}\, \acute{\imath}_s^2}{\langle i_n^2 \rangle} \qquad \textbf{(2.38)}$$

where $\langle i_n^2 \rangle$ = the sum of the square of the currents of all the noise sources

$$\langle i_n^2 \rangle = \langle i_s^2 \rangle + \langle i_b^2 \rangle + \langle i_d^2 \rangle + \langle i_{th}^2 \rangle \qquad \textbf{(2.39)}$$

where $\acute{\imath}_s$ = the signal current amplitude obtained from a sinusoidally modulated optical signal:-

$$P_s(t) = P_{s0}\,(1 + m \sin \omega\, t) \qquad \textbf{(2.40)}$$

where P_{s0} = the time average

m = the modulation index

ω = the frequency of the modulating sine function

From Equations (2.10), and (2.40), the expression for the signal current amplitude follows:-

$$i_s = \frac{\eta\, q\, \lambda}{h\, c}\, m P_{s0} \qquad \textbf{(2.41)}$$

The signal noise current is determined by the average optical signal power P_{s0}. Combining Equations (2.27), (2.28), (2.29), (2.37), (2.38), (2.39), and (2.41):-

$$\frac{S}{N} = \left(\frac{\eta\, q\, \lambda}{h\, c}\right)^2 \frac{(m\, P_{s0})^2}{4B\left[q\,(i_s + i_b + i_d) + \frac{2\, k_B\, T}{R_e}\right]} \qquad \textbf{(2.42)}$$

Consider m = 1 ie 100% modulation, and solving Equation (2.42) for P_{s0}, incorporating Equation (2.27), results in the optical signal power at the detector input necessary for a SNR as required by the system criteria:-

$$P_{s0} = \frac{2\, h\, c\, B}{\lambda\, \eta}\left(\frac{S}{N} + \sqrt{\left(\frac{S}{N}\right)^2 + \frac{S}{N\, B\, q}\left(i_b + i_d + \frac{2\, k_B\, T}{q\, R_e}\right)}\right) \qquad \textbf{(2.43)}$$

The optical power P_{s0} which generates a signal amplitude just equal to the noise amplitude (S/N = 1) at the output is termed the minimum detectable power P_{smin}. From Equation (2.43), the following expression results:-

$$P_{smin} = \frac{2\ h\ c}{\lambda\ \eta}\ \sqrt{B}\ (\sqrt{B} + \sqrt{B + \frac{i_b + i_d}{q} + \frac{2\ k_B\ T}{q^2 R_e}}) \qquad \textbf{(2.44)}$$

Dividing Equation (2.44) by $\sqrt{B}$ yields the noise equivalent power (NEP):-

$$NEP = \frac{2\ h\ c}{\lambda\ \eta}\ (\sqrt{B} + \sqrt{B + \frac{i_b + i_d}{q} + \frac{2\ k_B\ T}{q^2 R_E}}) \qquad \textbf{(2.45)}$$

Figure 2.12 depicts the NEP as a function R_e with dark current as a parameter. Since the background illumination can be made vanishingly small, the associated noise can be neglected.

2.3.6 Calibration of the p-i-n Photodiode

The following calibration procedure is necessary:-

I) Spectral variation measurements.

2.3.6.1 Spectral Variation Calibration

The p-i-n photodiode is a stable detector, and the spectral response will depend on the wavelength and intensity of the incoming light. The measurement requires uniform illumination over the photodiode window. This is achieved by using a source of white light, such as a shielded light bulb, a selection of filters and a focusing lens mounted in front of the p-i-n photodiode on an optical bench. The apparatus set-up was housed in a blackened box, awaiting the completion of the dark room. The voltage and current outputs are read from an oscilloscope, and a digitized ammeter respectively. These laboratory measurements of the spectral calibration were used only as a confidence check. Figure 2.3 illustrates the responsivity (AW^{-1}) versus filter wavelength (nm) of standard calibrations, which are more accurate than the recorded APL measurements. To enhance the maximum uniform illumination, the photodiode used was decreased to a 10 mm x 8 mm square, using an opaque material (black sticky tape), therefore becoming an overfilled than under-filled detector.

Ideally, the calibration laboratory should be temperature controlled. Humidity control is not necessary unless there is the possibility of condensation on the components.

2.4 Data Analysis

2.4.1 Collecting Data

During the MAED (Middle Atmosphere ElectroDynamics) Campaign, and the NLC-91 (NoctiLucent Cloud 1991) Campaign, the APL payloads

housing the photodiodes were flown on the Mother Section of the Nike-Orion 31.057 GE, and the Black Brant VB 21.103 GE and the VB 21.104 GE, respectively.

The Mother Section on the Nike-Orion 31.057 GE contained a Vector 200 k-bit PCM (Pulse Code Modulation) system, C-band radar transponder, accelerometer, 2-axis magnetometer, and related components. The PCM system was configured of 64 analogue inputs. In Pulse Code Modulation, only certain discrete values are allowed for the modulating signal. The modulating signal is sampled, as in other forms of pulse modulation, and any sample falling within a specified range of values is assigned a discrete value. Each value is assigned a pattern of pulses (see Figure 2.13). The filtered PCM output modulates a 2 Watt transmitter operating at 2269.5 MHz. The telemetry system was powered by nickel cadmium cells, which had a discharge-rate of one hour.

The data from the photodiodes 1, and 2 (channels A1, and A2, respectively) (see Table 2.2) were sampled at a rate of 1250 samples per second (sps) with two words (wd) every frame (fr) every interval (int) ie 2-1-1.

During the MAED Campaign, the following ground-based telemetry was available:-

I) Magnetic Tape Recorders - NASA's ground stations at Esrange provided support for the launches. Redundant direct record tape

recorders were the prime data record.

II) Paper Tape - Recorder formats that were used for real time flight data and prior pre-launch payload check.

The raw data were collected in the field on magnetic tapes of density 1600 bpi (bytes per inch). On returning to the UK, the complete data records stored on the magnetic tapes were transferred using TPU (Tape Processing Utility) onto the APL DEC microvax for subsequent reduction, analysis, and permanent storage.

The Mother Section of the EFIELD payload on each of the Black Brant VB 21.103 GE, and the VB 21.104 GE respectively, were two freefall payloads in a Mother-Daughter configuration. The Mother Section was envisioned to house photometers to measure the aerosol layer passage and structure, a second vector electric field instrument, an X-ray detector, and a solid state detector to measure ionizing night-time radiations, and one or more plasma probes to measure the ion, and electron density, and conductivity.

The Mother Section contained a 800 k-bit PCM link. A radar transponder provided trajectory data. Both 3-axis gyro, and a 3-axis magnetometer were flown for attitude determination. The filtered PCM output modulates a 5 Watt transmitter operating at 2246.5 MHz. The telemetry system was powered by nickel cadmium

cells which had a discharge-rate of thirty-five minutes.

Table 2.2 illustrates the signal sampling routine pattern for a select number of frames. Those frames omitted, 5 -> 11, follow the same format as frames 2 -> 4, and 12.

Table 2.2

F SF	1	2	3	4	12	13	14	15	16
1	SFID	A1	A2	A3	A11	A12	A28	A14	SYNC
2						A13	A29	A15	
3							A30	A14	
4							A31	A15	
5						A16	A32	A14	
6						A17	A33	A15	
7						A18	A34	A14	
8						A19	A35	A15	
9						A20	A36	A14	
10						A21	A37	A15	
11						A22		A14	
12						A23		A15	
13						A24		A14	
14						A25		A15	
15						A26		A14	
16	SFID	A1	A2	A3	A11	A27		A15	SYNC

The data from the photodiodes 1, 2, and 3 (channels A48, A49, and A52, respectively), were sampled at a rate of 800 samples per second with the following signal sampling routine:-

Channel	Photodiode	WD	INT	FR	INT
A48	1	36	50	1	2
A49	2	36	50	2	2
A52	3	27	50	3	2

During the NLC-91 Campaign, the same ground-based telemetry was made available as in the MAED Campaign. Again, the raw data were collected in the field on magnetic tapes of density 1600 bpi. The complete data records stored on the magnetic tapes were transferred from the respective flight tapes to the APL DEC microvax using the program STRP1991, specifically written for this raw data.

STRP1991 provided a fast efficient data record transfer procedure taking thirty minutes per magnetic tape compared to several hours per magnetic tape required by the TPU.

2.4.2 Data Transfer from Magnetic Tapes to the APL DEC Microvax

As aforementioned, on returning to the UK, the respective stored data files/records were transferred from the magnetic tapes to the APL DEC microvax using the specific TPU facilities available on the Vax 11/730.

The TPU was a viable transfer procedure for the data obtained from the MAED Campaign, since it had the same stored signal format as that of a previous rocket flight. Due to the lack of sufficient storage space on the Vax 11/730, the output data file was stored on the microvax. The unprocessed output file is still

in an unreadable format and requires further processing. On running the program RDES (ReaD Esrange) on the TPU output file,the sequence of data bytes were swapped around, so that the least significant became the most significant, and vice versa. The resultant RDES output file had the following layout:-

Time Relative to Launch = I4 seconds
Record Number I4, Format Number I2
IDCTR PHO1 PHO2 DIOD1 DIOD2 SUN1X SUN2X SUN1Y SUN2Y MGROL MGY MGP

where I2, and I4 indicate an integer of two and four digits long respectively.

IDCTR	identity counter
PHO1/2	photodiode 1, 2 respectively (APL)
DIOD1/2	diode 1,2 respectively (MISU)
SUN1X/2X	sun sensors 1X, 2X respectively
SUN1Y/2Y	sun sensors 1Y, 2Y respectively
MGROL	magnetometer roll
MGY	magnetometer yaw
MGP	magnetometer pitch

The aforementioned headers re-occurred every sixteen points. It should be noted that the time rate recorded did not increment at a regular interval. However, there were sufficient number of data points to average over every sixteen points without losing structural detail.

Finally, the resultant RDES output file was "tidied up" using the program TIDY specifically written to strip all headers from the data file, bar the first row of column headers - helpful indication for future reference. The program then averaged the data for every sixteen points, and replaced the identity counter column with the respective stripped time header.

The TIDY raw data output file was now ready to be plotted out and studied, so that promising data could be selected for further analysis (see Section 2.4.3). The graphs were plotted on a suitable local PC graphics terminal, and plotted on a Tektronix 4691 ink jet plotter, or a HP Paint Jet plotter.

As aforementioned, STRP1991 was the main program used to transfer the NLC91 data from the magnetic tapes to the microvax. This program was written specifically for this raw data, since:-

1) the signal code format has altered since 1986 ie MAED, and thus, the TPU procedure was no longer viable, and
2) to save time running subsidiary programs.

Therefore, STRP1991 incorporates similar processes such as RDES and TIDY.

Hence, the data was written out in the following layout:-

TIME PHO3 S1X PHO1 MGX MGY PHO2 SIY TEMP

where

TIME	time relative to launch
PHO1/2/3	photodiode 1, 2, 3 respectively (APL)
S1X/Y	sun sensors 1X, 1Y respectively
MGX/Y	magnetometer X, Y respectively
TEMP	temperature of the atmosphere

STRP1991 has averaged the data over every tenth of a second, and removed all default values ie -100.00, an occasion when the respective instrument measured an undefined quantity. The number of data points measured per tenth of a second were random, varying from four to six, so STRP1991 was written in order to detect the change in the tenths of a second, and average over the appropriate number of data points. Again, the resultant raw data file was readily available for the on-line graphic facilities. However, unlike the MAED data, this raw data could not be further reduced, and ultimately analyzed, since on plotting the raw data, no promising data sets were found (see Section 2.5). In brief, Section 2.5 will deal with the respective inflight problems encountered during the 21.103, and the 21.104 rocket flights. The data return were poor, and further data reduction, and analysis could lead to erroneous conclusions (see Chapter 3).

The derivation of the look direction of the photodiode on a rocket payload is discussed in Appendix 1.

2.5 Problems Encountered During Data Analysis

During the MAED and NLC-91 rocket campaigns respectively, the APL narrow scanning photodiodes were employed to scan the entire sky laterally from the rocket payloads by utilizing the rocket payload spin. As a result of the spin, and coning motions of the payload, the photodiodes provide a limited two-dimensional imaging of the aerosol layer from below at an altitude of 60-80 km, and from above at an altitude over 90 km.

The MAED plots shown in Chapter 3, indicative of a traversed aerosol layer illustrate the expected gradient change (negative/positive, see Chapter 3) in the photometer reading. During upleg/downleg flight, the aerosol layer will act as a filter, thereby decreasing number of photons reaching the photodiode window. Hence, on exiting the layer, the number of photons reaching the photodiode window increases. The MAED plots indicate the presence of a strong NLC layer, however a conclusion cannot be deduced from the NLC-91 Campaign.

On the 1st of August 1991, the Black Brant VB 21.103 GE was launched into a PMSE (Polar Mesospheric Summer Echoes), and not a NLC layer. However, there is some speculation as to whether the rocket payload passed through a weak NLC layer or not (Goldberg, personal communication). The results from this flight are highly speculative due to the collision between the Mother and Daughter payload (see Chapter 3).

The 21.103 developed attitude problems when the Daughter payload collided with the Mother payload shortly after separation. This caused the Mother to cone severely, and tumble. The magnetometer data were of poor return, since it is believed that for a short period during tumbling, the Mother and Daughter were parallel, and thus resulted in the magnetometer booms being snapped off. Therefore, without the magnetometer data, the author was unable to make a detailed analysis of the APL photodiode data. This payload did not meet all the minimum criteria, and thus, was termed a failure.

On the 9th of August 1991, the Black Brant VB 21.104 GE was launched into a weak NLC. This flight also experienced problems. Due to the collision between the Daughter and Mother during the 21.103 flight, the despin, and the payload separation times on the 21.104 were altered to overcome another collision situation (see Table 2.3).

Table 2.3

Time /secs	21.103	21.104
despin	69	66
nosecone eject	72	69
daughter eject	75	72
payload separation	78	75

During the 21.104 payload separation, the Mother experienced a slight coning. Once the nosecone had ejected, the coning became larger. When the Daughter was ejected, the Mother went into a 60° half angle cone. Fortunately, the gimbal did not lock the gyro,

thus minimum success was met.

Although, the program was conducted as planned, the campaign was hampered by natural manifestations. Often winds negated launch for several of the US payloads, or thunderstorms that flooded the launch area wiped out power for two nights. The lack of NLCs prolonged the launch date. Thus, with these obstacles, the launch into NLCs became later and later, therefore approaching the end of the NLC season.

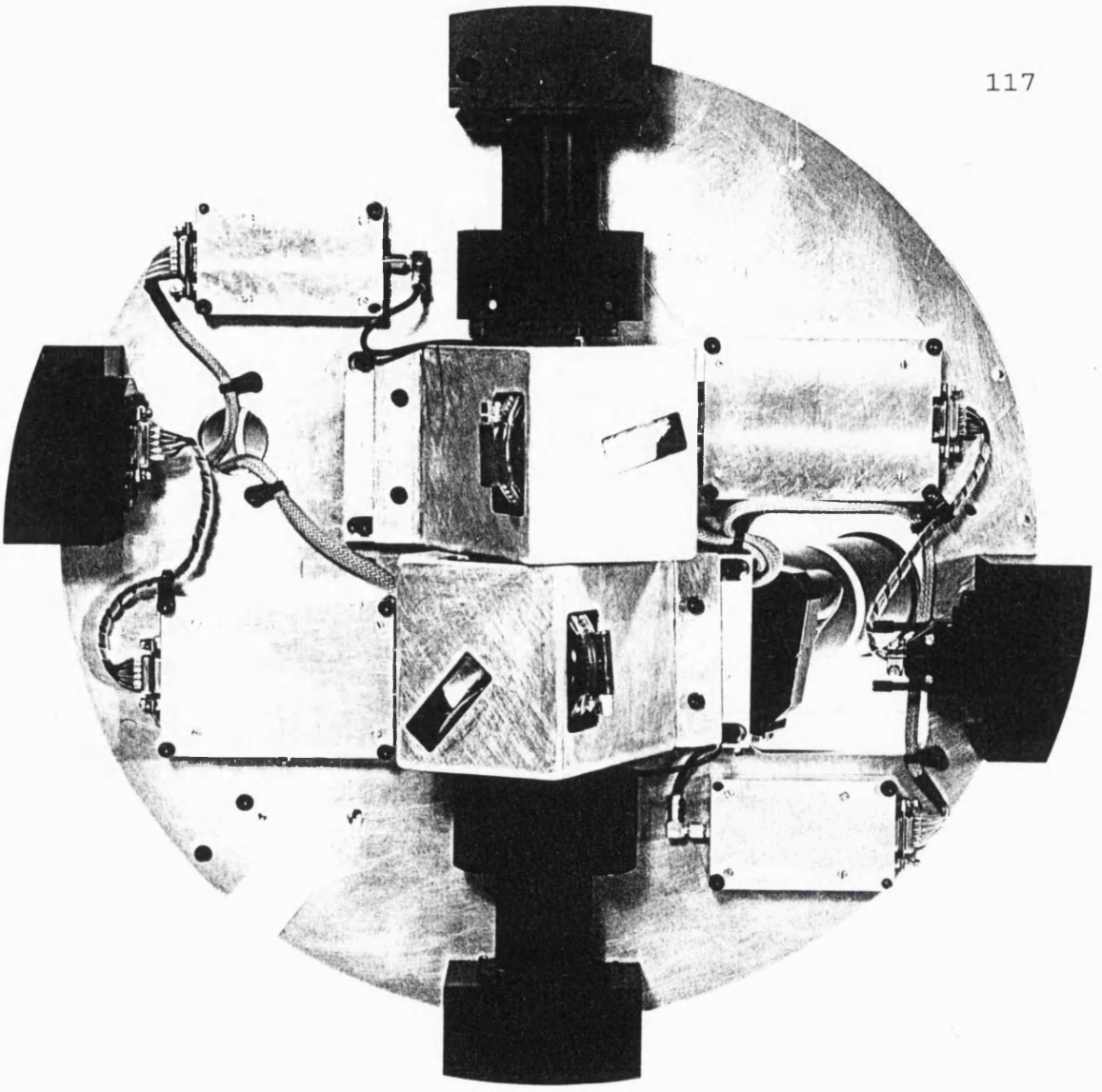


Figure 2.1A

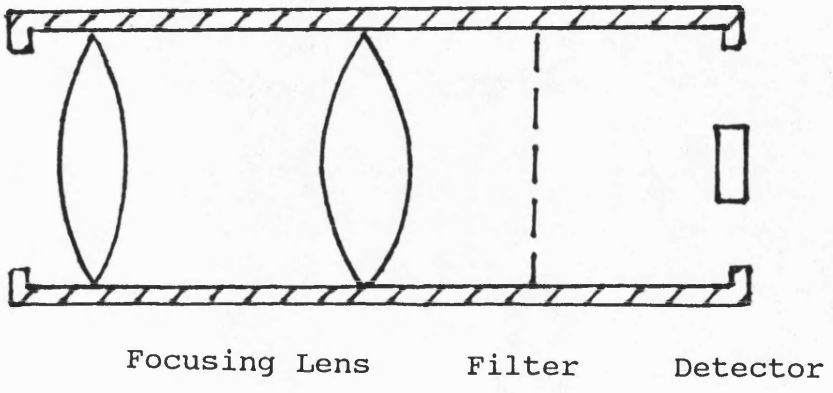


Figure 2.1B

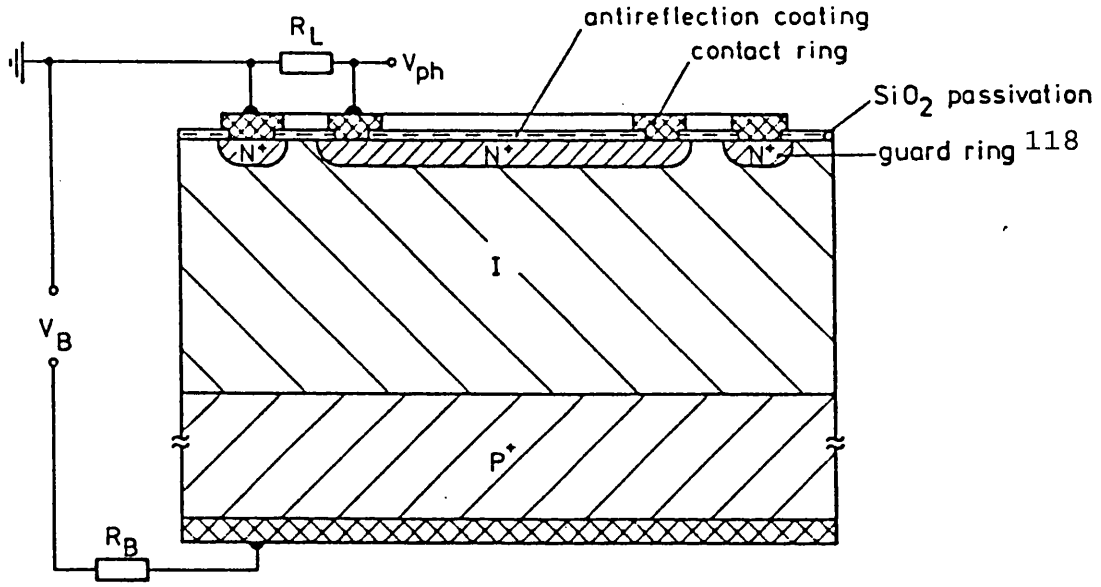


Figure 2.2

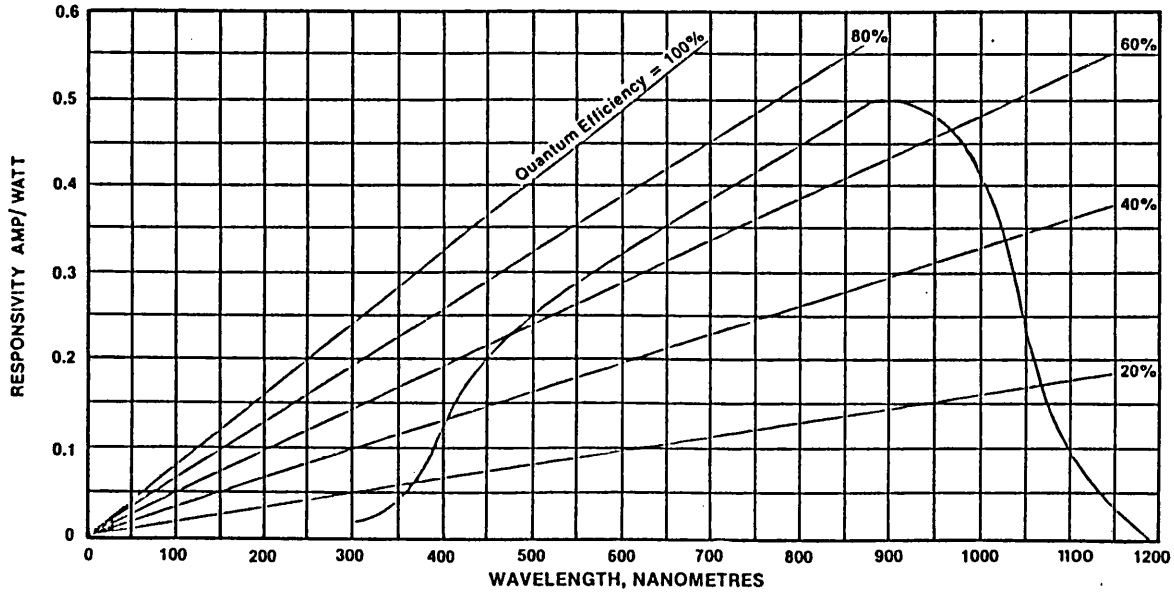


Figure 2.3

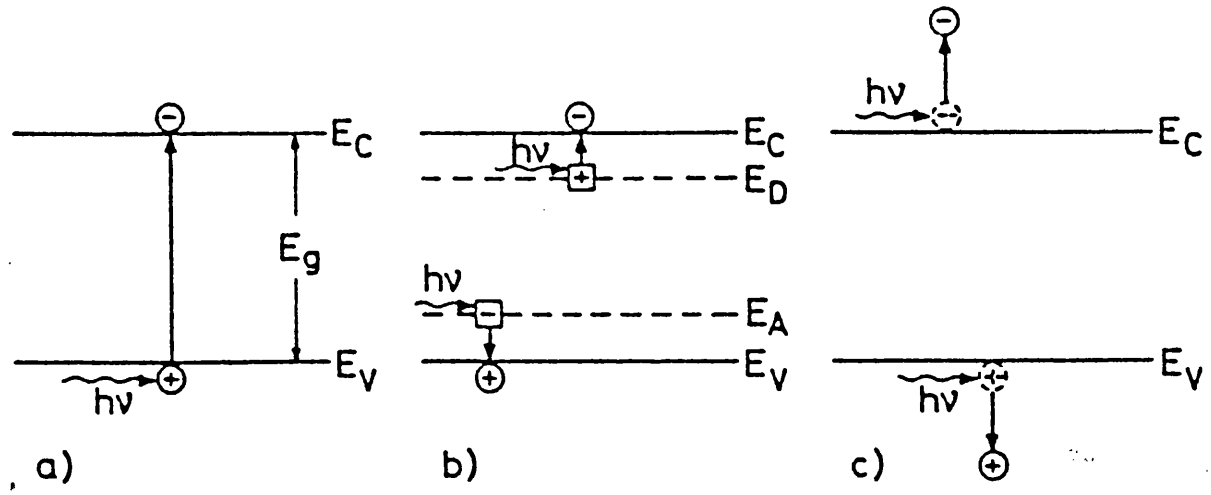


Figure 2.4

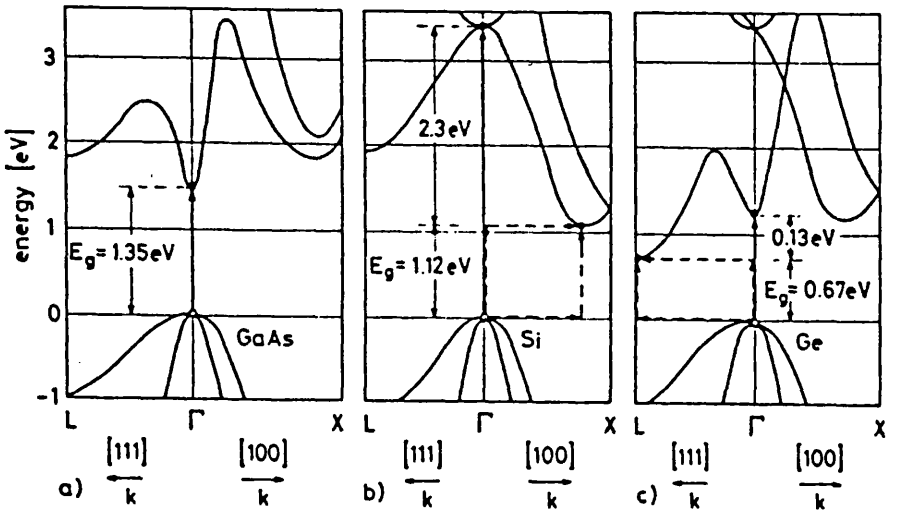


Figure 2.5

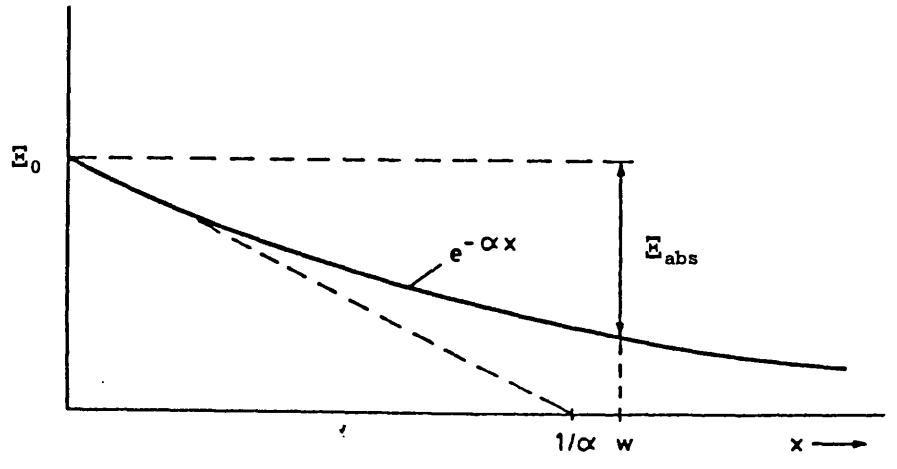


Figure 2.6

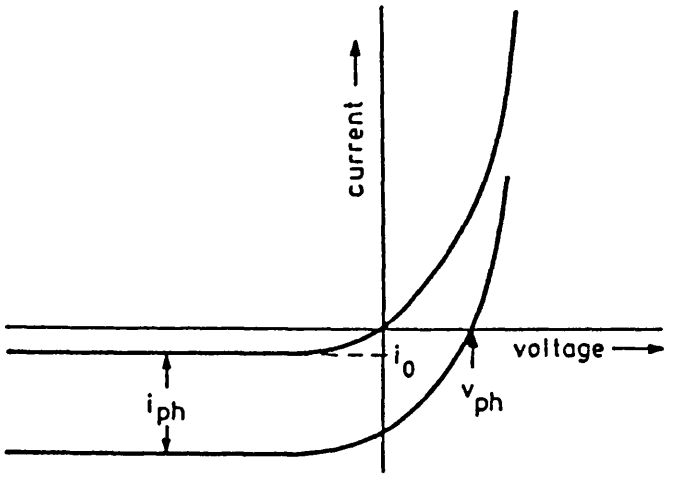


Figure 2.7

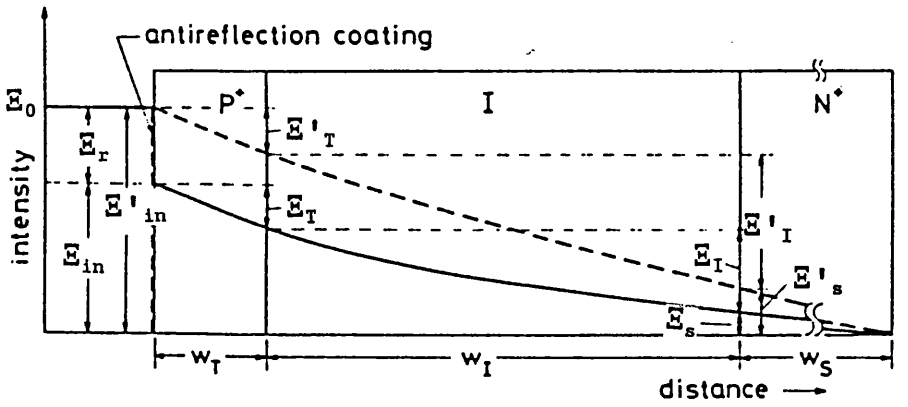


Figure 2.8

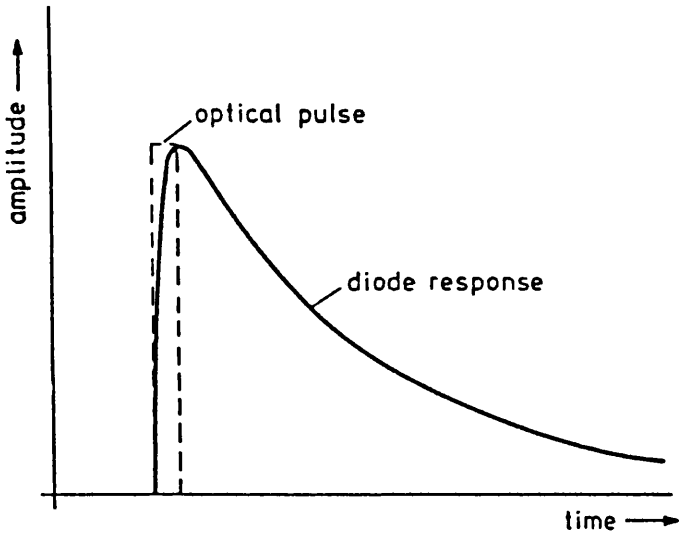


Figure 2.9

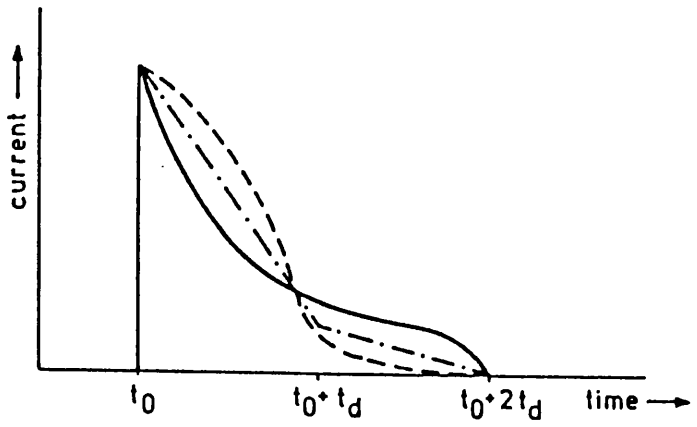
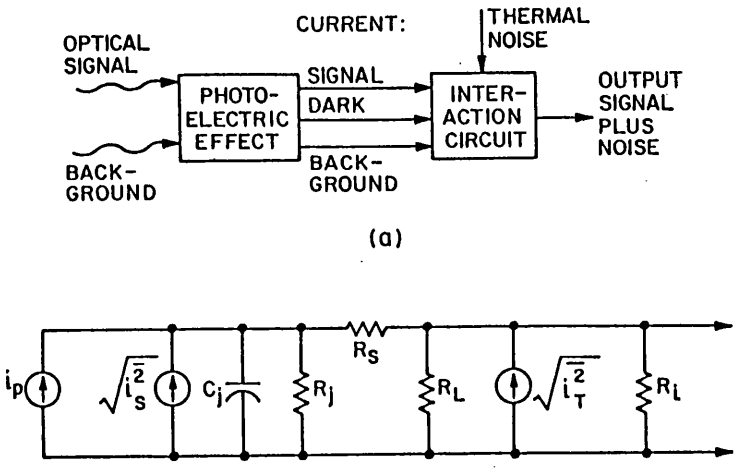


Figure 2.10



Figures 2.11a and b

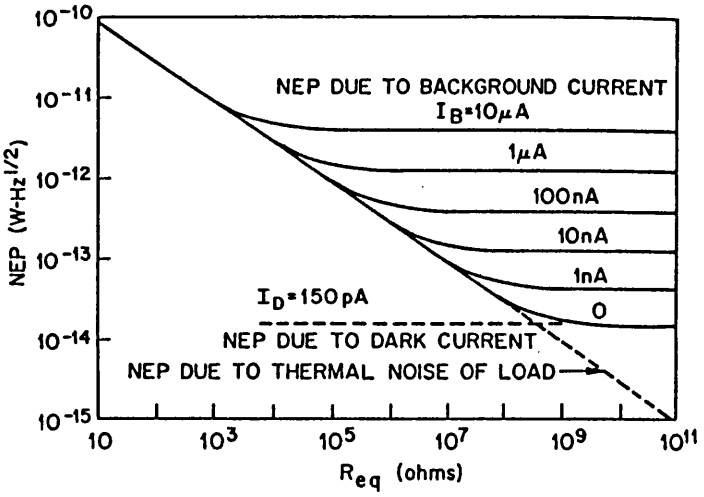


Figure 2.12

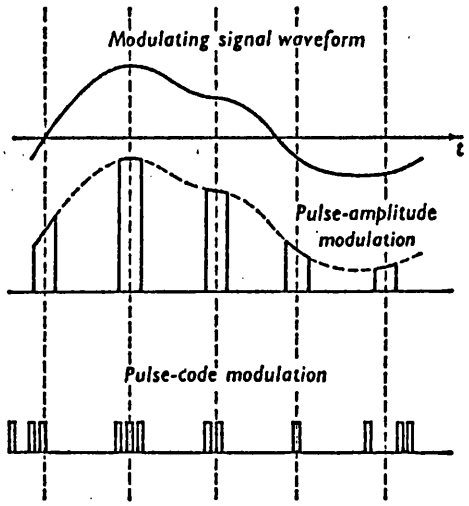


Figure 2.13

Chapter 3

3 Rocket Instrumentation Results

3.1 Introduction

This chapter will discuss in detail the results obtained from three rocket campaigns; CAMP (Cold Arctic Mesopause Project) in 1982, MAED (Middle Atmosphere ElectroDynamics) in 1986, and the NLC-91 (NoctiLucent Clouds 1991) in 1991 conducted during the summer months located at Esrange, Kiruna, Sweden.

The author spent a month during 1992 working with Prof. Georg Witt, of the Arrhenius Laboratory, (Stockholm, Sweden) studying the Meteorological Institute of Stockholm University (MISU), rocket data collected from the CAMP, and MAED rocket campaigns. The CAMP results will be discussed in detail in Section 3.2.

The MAED rocket campaign was instrumented primarily to study the electrical environment in the vicinity of a Noctilucent Cloud (NLC). The results obtained from the aerosol detector experiment, developed by the MISU, and the complementary data acquired by the APL photometers will be presented and discussed in Section 3.3.

The multinational rocket campaign, NLC-91, was organized to determine, with in-situ experiments, the dynamical, electrodynamical, physical and chemical parameters of the NLC layer, combined with ground-based experiments. The objectives and

results obtained will be discussed in Section 3.4.

3.2 CAMP 1982 Rocket Campaign

3.2.1 Introduction

The CAMP campaign was conducted at Esrange (67.9°N, 21.4°E), Kiruna, Sweden. The rocket measurements in CAMP were supported by ground-based, aeroplane, and satellite remote sensing experiments. Importantly for CAMP, there was the UV-backscatter experiment on the SME satellite [Thomas, 1984][109] which, during the summer 1982, gave a continuous survey of the polar mesospheric clouds (PMC) at NLC heights over the summer polar cap region. The rocket-borne techniques in CAMP included optical photometers to detect NLC particles, mass spectrometers for positive and negative ion species, high resolution electron, and total positive ion probes, accelerometer measurements of density, temperature, and wind, chemical release wind measurements, ionization rates from solar, and particle sources, and minor species such as atomic oxygen and argon.

This section covers a selection of rocket and ground-based measurements at the time of the first salvo (3/4 August). Table 3.1 summarizes a select number of rocket payloads flown, and their flight objectives, during the CAMP Campaign.

Table 3.1

Rocket Payload	Instrument	Parameter	Date	Time /(UT)	PI
SOAP1	Photometer Lyman-α	NLC altitude, Temperature	3-Aug-82	23:49	Witt
SOAP2	Resonance Fluoresce.	Atomic Oxygen	3-Aug-82	23:49	Witt Dick-inson
TAD	Active Falling Sphere	Temperature, Density, Wind Velo., Turbul. Waves	4-Aug-82	00:16	Phil-brick

3.2.2 Experimental Technique

The SOAP (Selective Optical Atmospheric Probe) payload carried the two MISU photometers positioned at 180° apart each equipped with identical filters, centred at 450 nm, with a bandwidth of 9 nm. The radiance observed by the rocket-borne filter photometer in a sunlit atmosphere contained contributions from molecular light scattering, light scattered by particulates, selective emission, the sky background, and spurious light scattered by parts of the rocket instrumentation. Since each of the above sources of signal has in general a different degree of polarization, the signals were separated with a minimum number of assumption when measurements were made with the polarization sensitive photometers. By means of digital data analysis technique it was possible to derive from the measured signal smoothed values of both Stoke parameters I, and Q, and hence, the polarization. The background, and the light scattered from the instrument sunshade was masked out from the data by subtracting the apogee data. Both of these components vary with the rocket

attitude, and corrections including this variation have been made during the data processing. Figure 3.1 illustrates the height profile of the Stokes parameters I, and Q obtained from the 450 nm photometer. This filter channel was selected for the comparison as there were no significant emission features near this wavelength. The abrupt change in the observed signal near 82 km for both I, and Q was due to the presence of a well developed NLC display. In the absence of the excess scattering, the signal would be expected to decrease with the atmospheric pressure. The observations from the atmospheric profile can be explained by the spatial variations of the overhead radiance scanned by the precession of the payload.

The observation from the 450 nm photometer showed that the degree of polarization (see Figures 3.2, and 3.3) illustrates the degree of polarization of the signal from the clouds, and the atmosphere, within a few per cent of the theoretically expected Rayleigh scattering degree of polarization. Above the clouds, the measured degree of polarization decreases sharply, then further decreases between the clouds, as the atmospheric scattering contribution diminishes.

3.2.3 Data Analysis

The nature of the scattering medium is characterized by the degree of polarization, and the colour index of the scattered radiation [Witt, 1968][110]. These parameters are discussed in detail in Appendix 2. Figures 3.4, 3.5, and 3.6 referred to in

Appendix 2 can be found at the end of this chapter.

3.2.4 Measurements

The measurements made in CAMP were obtained from a total of ten rocket payloads launched in two salvoes. The main salvo (salvo A) with seven rockets launched at 23:02 UT on the 3rd of August, and the smaller salvo (salvo B) with three rockets began launching at 23:30 UT on the 11th of August, 1982. Further details of the different payloads, and instruments launched have been published by Bjorn [1984][111]. Salvo A was launched into a well developed, and stable NLC display observed from a spotter aeroplane 500 km South of Esrange over a period of seven rockets launches. The altitude of the NLC observed with rocket-borne photometers, and positive probes was 83 +/- 1 km. See Figure 3.7. This salvo was launched a few days after a major solar proton event, which had caused a much larger electron and ion densities than originally anticipated (Witt, personal communication). Figures 3.7, and 3.8, indicate the measurements of the North-South magnetic field, and the riometer absorption at 27 MHz in Kiruna during salvo A on 3/4 August. High electron densities, 5 x 10^3 cm^{-3} at 80 km, and 3 x 10^5 cm^{-3} at 100 km were measured as a consequence of the high ionization rate. The advantage of a ionization level provided an excellent signal-to-noise ratio obtained with the incoherent scatter radar echoes, measured by EISCAT (European Incoherent Scatter Radar at Tromso, Kiruna, and Sodankyla), which made vertical wind velocity measurements possible.

Six optional photometers were used on the payloads:- CAMP-N, CAMP-P, and the SOAP payloads in salvo A for the detection of the NLC layer. The photodiode has a narrow field-of-view to the side of the payload. The two curves indicated in both Figures 3.9 and 3.10, illustrate the different rocket spin angles. The enhancement above the normal Rayleigh scattered light of the background atmosphere at 81-83 km originated from the presence of visible NLC particles (Witt, personal communication). The other photometers on CAMP-P, and CAMP-N also detected the scattered light from the NLC layer at similar heights. On CAMP-P payload, positive charge density was measured with a spherical probe consisting of an inner sphere 1.7 cm, and an outer grid with a diameter of 4 cm, respectively. The grid was biased with -5 V relative to the rocket ground, and the accelerating potential between the grid, and the sphere was 10 V. The derived positive charge density was (see Figure 3.11) calibrated against the electron density measurement at 95 km from the propagation experiment on the same payload. From Figure 3.11, the unguarded MS-tip was flown on the same payload as the boom probe (CAMP-P). It revealed a distinct positive charged layer at the NLC height during the rocket descent. The boom probe being of a smaller cross-sectional area than the MS-tip probe, 50 cm^2, could explain the reason why the boom probe failed to observe the positively charged aerosol layer.

The cross-section presented to the positive charges entering the grid on the boom probe was mass and mobility dependent. The cross-section for heavy ions, and charged aerosols and heavy ions

were much larger on the MS-tip probe. A contribution of photoemission at both probes by Lyman-α radiation was possible. However, the observed increase of positive charge on the descent of CAMP-P was not explicable by the enhancement of the photoemission from the scattered Lyman-α at the NLC particles. The only explanation that can describe why positively charged aerosols existed was due to the faster photo-detachment process compared to the time for electron loss by attachment to these particles in the sunlit mesosphere. The photo-detachment depends on the size and nature of the particles in question.

In-situ measurements of the temperature were made in salvo A, using the measurements of density by a falling accelerometer sphere, bead thermistor measurements of temperature (Schmidlin, personal communication), and scale height temperature measurements from the absorption of the Lyman-α sunlight and the EISCAT measurements [Kofman **et al.**, 1984][112]. See Figure 3.12. The temperature profile between 55 km and 120 km was derived from the high resolution density measurement of the triaxial accelerometer on the TAD (Thermal and Atmospheric Dynamics) payload.

The wave structure between 80-100 km has three minima, 139 K at 83.5 km, 199 K at 89.4 km, and 110 K at 93.6 km respectively. The first minimum was located at the NLC region. The temperatures of the two latter minimum were significantly lower than any which have been previously measured in the atmosphere [Schmidlin, 1992][113]. Prior to 1991, the lowest recorded temperatures, 130-140 K, were measured at Pt. Barrow (71°N) with grenades [Theon **et**

al., 1967][23]. However, the grenade techniques were not of sufficient altitude resolution to resolve temperature structures of 5-7 km vertical wavelength as observed in CAMP at the mesopause. A temperature profile was determined from integrating the profile of the molecular absorption UV-measurement (Lyman_α) on the SOAP payload. See Figure 3.13. This derived temperature was obtained assuming hydrostatic equilibrium. Again, insufficient height resolution to resolve the wave structure observed with the triaxial accelerometer. The Lyman_α measurement gave a temperature measurement of about 120 K in the altitude regime of 93 km. This mean temperature obtained from EISCAT agreed with that obtained from the triaxial accelerometer in the altitude region of 68-90 km.

The wave structure found in the temperature profile was much more pronounced in the zonal wind velocity profile, see Figure 3.14. The mean zonal wind velocity was determined from the triaxial accelerometer on the TAD payload, is westward below 94 km, and eastward above 94 km, in good agreement with Garcia and Soloman [1985][114], and radar measurements [Nastrom **et al**., 1982][115]. A well developed wave with five vertical wavelengths were measured in the zonal wind at 65-98 km with a mean vertical wavelength of 7.5 km. The exponential growth of the amplitude with increasing height was evident only at heights below 70 km. The wave above was affected by dissipation by five turbulent layers measured at 75.5 km, 79.5 km, 86.5 km, 89 km, and 93 km.

Vertical wind amplitude measurements were available from EISCAT

measurement on the 2nd of August around 23:00 UT, the evening before salvo A was launched. The measurements by EISCAT were made in the biostatic mode. By using the envelope around the measured wind velocities, an amplitude range of +/- 5 ms^{-1}, wind amplitude and vertical wavelength of approximately 7.5 km (Rottger, personal communication).

The following Table 3.2, summarizes the features observed at the NLC height.

Table 3.2

Rocket Payload	Parameter	Altitude/ km
CAMP-P	Positively Charged Aerosol Layer Highest Order and Maximum Density of Proton Hydrates	83.7 up 83 down 82 +/- 0.5
SOAP 1	NLC - Photometers O - Density Depression	82.4 low NLC 83.0 upp NLC 83.3
TAD	First Temperature Minimum	83.5

In-situ measurements of temperature, and wind indicate that there were five thin layers of turbulence in the mesosphere in the altitude region 75-93 km, and a wave of wavelength 7.5 km. These observed turbulent layers dissipated this wave.

3.3 MAED 1986 Rocket Campaign

3.3.1 Introduction

The Middle Atmosphere ElectroDynamics (MAED) campaign was

conducted at Esrange, Sweden during July 1986. Three Nike-Orion rockets, and six Super-Loki Sphere vehicles were launched on the 26th of July, during a period of an NLC sighting over Esrange. The presence of the NLC was confirmed by several rocket-borne photometer, and plasma probe profiles. The NLC height was 83 +/- 2 km. A 132 K temperature minimum was located at the height of the observed NLC.

The rocket campaign was instrumented primarily to study the electrical environment in the vicinity of the NLC. The rocket measurements in the MAED were supported by ground-based, aeroplane, and satellite remote sensing experiments. The SME satellite provided leading limb coverage during the launch window period, 23rd of July to the 8th of August, reporting daily on the activity within the field of interest.

The rocket-borne techniques in MAED included an aerosol detector experiment (developed by the MISU), an optical aerosol detector (developed by the APL, UCL), mass spectrometers for positive and negative ion species, Maynard electric field probes, accelerometer measurements of density, temperature, and wind, ionization rates from solar and particle sources.

In this section, emphasis will be made on the results obtained from the aerosol detector experiment (MISU), and the optical aerosol detector (the p-i-n photodiode, as discussed in Chapter 2).

3.3.2 Measurements

The measurements made in MAED were obtained from a total of nine rockets launched in two salvoes. The main salvo started at 23:37:19 UT on the 26th of July. The first salvo was launched into a well developed and stable NLC display observed by a spotter aeroplane 500 km South of Esrange over the period of the launches. Further details of the different payloads, instruments, and flight requirements are available, at request, from NASA/GSFC. The altitude of the NLC observed with rocket-borne APL photometers was 83 +/- 2 km.

The Mother payload, of the Nike-Orion 31.057 GE Mother-Daughter configuration (see Figure 3.15), housed the photometric aerosol detector experiment, and the optical aerosol detector experiment. The doors associated with the former experiment were located 180° apart exposing the two photometers which were themselves angled at 30° to the rocket centreline. As a part of the optical aerosol detector, two photodiodes and two sun sensors were located aft of the doors, and were exposed throughout the flight.

The photometer aerosol detector experiment was similar to that discussed and illustrated in Section 3.2. The APL photometers were positioned 180° apart aft of the doors. These detectors were unfiltered, and were primarily flown to confirm whether or not an aerosol layer was traversed during the rocket flight. Although these photometric detectors provide a limited two-dimensional imaging of the aerosol layer, it is sufficient as a confirmation

method as to the presence or absence of an aerosol layer, such as NLC.

The measured electrical properties of this campaign still remain unpublished (Pfaff, personal communication). However, from the APL, and the MISU plotted results (see Figures 3.16, and 3.17), it is clear that on the upleg flight, a strong, stable NLC was traversed at a height of 83 +/- 2 km. The MISU results confirm this observation, at a height of 82.2 km. The gradient change clearly illustrates an aerosol layer (as discussed in Chapter 2). On the downleg flight, the NLC appears to have drifted downwards to a height of 83 +/- 3 km. Witt confirms this observation, the upper NLC layer being at 83.2 km. However, the NLC profile was no longer smooth, glitches appeared in the profile. The cloud became "patchy", and tenuous. Although, it was believed to be stratified as observed during the CAMP campaign (Rees, personal communication), this is improbable, since neither the photometer nor the temperature profiles illustrate a step structure nor more than one minima, respectively. See Figure 3.18. As only one minimum can be seen, it is feasible that the NLC layer was breaking up under the influence of a prevailing north-eastward wind. Another approach is to plot the radiosonde density data logarithmically. Linearly, all detail of the cloud structure is lost. However, a logarithmic curve illustrates the structural detail of the aerosol cloud, clearly indicating cloud stratification. Although these techniques are not conclusive, the results would strongly suggest that only one layer occurred being strong and stable, but by the downleg flight, the NLC was

dispersing.

In-situ temperature measurements were made throughout the campaign, before and after each rocket launch during each salvo. Temperature, density, pressure, and wind parameters were measured employing a radiosonde. As aforementioned, during the 31.057 GE launch, between 80-100 km only one temperature minimum was measured. The temperature minimum occurred at 132 K, at an altitude of 83 +/- 2 km. There were no significant turbulent layers indicative of dynamically unstable layers.

The following table, Table 3.3, summarizes the features observed at the NLC height.

Table 3.3

Rocket Payload	Parameter	Altitude/km	PI
31.057 GE	NLC - photometer	83 +/- 2 ↑ 83 +/- 3 ↓	Rees
31.057 GE	NLC - photometer	82.2 low.NLC 83.2 upp.NLC	Witt
Radiosonde	Temperature Min.	132 K @ 83 +/- 2	Schmidlin

The brief summary indicates that a single aerosol layer was observed on the night of the 26th of July 1986 in which a layer became tenuous and drifted by 3 km over a period of approximately 100 seconds.

3.4 NLC-91 Rocket Campaign

A multi-national rocket and radar program took place during July/August 1991 in Northern Scandinavia and the Barents Sea. The main objectives of this campaign were to determine with in-situ experiments the dynamical, electrodynamical, physical, and chemical parameters of a NLC layer combined with ground-based visible, radar, lidar, and micro-wave experiments. The altitude resolution for this campaign was between 50-100 m.

3.4.1 Introduction

The project involved several rocket launchings mainly from Esrange, Kiruna to improve the present knowledge on physical. chemical, dynamical, and electrodynamical conditions, and processes in the neighbourhood of a NLC layer. The use of the ground-based radars, mainly EISCAT's VHF radar at 224 MHz, and the CUPRI radar at approximately 50 MHz, for the first time provided an opportunity to investigate together with in-situ measurements, PMSE (Polar Mesospheric Summer Echoes), and its relationship to NLC. Previous rocket measurements of the electron density, ion composition, turbulence, atomic oxygen, water vapour, and other parameters in the vicinity of NLC indicated that such layers were accompanied by significant variabilities in a vertical scale of the order of 100 m. New instrumental rocket techniques with the necessary height resolution were used in the NLC-91 program.

Tables 3.4a and b summarize the payloads flown, flight details, and their experimental objectives.

Table 3.4a

Salvo	Rocket Payload	Launch/UT	Date
A	Viper 6	22:53:00	9-Aug-91
	TURBO A	23:15:00	9-Aug-91
	DECIMALS A	23:15:15	9-Aug-91
	MISTI A †	23:30:00	9-Aug-91
	EFIELD A * †	23:30:20	9-Aug-91
	PEP A †	23:40:00	9-Aug-91
	Viper 7	00:06:00	10-Aug-91
	Viper Chaff 2 †	00:24:00	10-Aug-91
	DECIMALS B	01:37:00	10-Aug-91
	Viper 8	01:42:00	10-Aug-91
B	Viper 1	01:03:00	1-Aug-91
	MISTI B †	01:82:00	1-Aug-91
	EFIELD B * †	01:28:20	1-Aug-91
	PEP B †	01:39:00	1-Aug-91
	TURBO B	01:40:00	1-Aug-91
	Viper 2	01:54:00	1-Aug-91
	Viper Chaff 1 †	02:24:00	1-Aug-91
	Viper 3	02:39:00	1-Aug-91
C	MISTI C †	23:21:00	5-Aug-91
	Viper 4	23:32:00	5-Aug-91

* APL flown instrumentation

† NASA contribution

The photometric detectors were developed by the APL, to detect the presence and structure of a NLC encountered. These data sets provide complementary information with that obtained by the MISU regarding the vertical structure. SLIPS (Scattered Light Intensity Profile Sensor) photometers developed by the University of Stockholm were used to measure the height profile of the scattered light by the NLC particles.

Table 3.4b

Rocket Payload	Objectives
TURBO	Provided instrumentation to study neutral and small scale plasma waves and turbulence
DECIMALS	Dynamical Electrodynamical, and Chemical Interactions of the Mesopause at Arctic Latitudes in Summer Designed for the study of heterogeneous interactions between the neutral and ionized components of the NLC particles in the summer high latitude mesosphere.
MISTI	Mesospheric Ionization Structure and Turbulence Investigation Equipped with electron density and a Faraday experiment.
EFIELD	Mother-Daughter configuration system concentrated on measurements in the field of electrodynamics, in particular AC/DC E-fields, conductivity,signatures of atmospheric turbulence, and the effect of the presence of NLC particles in electron and positive ion density.
PEP	From Penn State. Gerdien condenser measured conductivity, mobility, and ion density. It was able to distinguish between their mobility, and ion density of low and high mass ions in the NLC region.
Falling Spheres CHAFF	Meteorological rockets provided additional information on atmospheric temperature, horizontal wind velocities, neutral air density, and turbulence on the attitude range of 30-90 km.

The two radar stations, EISCAT (Tromso), and CUPRI (Cornell University Portable Radar Interferometer) (Esrange) were available to look at PMSE during the NLC rocket launch period, 8th of July -14th of August. A PMSE may possibly originate from the thin layers of approximately 100 m thickness with special properties, which do affect the electron diffusivity as a result of the presence of high mass ions and/or special turbulence

within the layer. The PMSE as seen by these radars cannot be directly related to the occurrence of an NLC, but may be a signature of a structure which is relevant for transport and build-up of NLC particles.

The main salvo at Esrange, Kiruna required an aeroplane to observe NLC displays over Esrange.

The contribution made by the USSR was the launching of ten M-100B rocket payloads from Heiss Island. The instruments employed were coordinated with salvo A at Esrange, Kiruna. Similar experiments were flown from Heiss Island including:- E-fields, SLIPS, scattered radiation photometers and chaff foils. The USSR results obtained will not be discussed in this thesis.

3.4.2 Measurements

Measurements in NLC-91 were obtained from a total of nine rocket payloads launched in three salvoes (see Table 3.4a).

Figure 3.19 illustrates the set-up on each payload flown on the Mother Section of each EFIELD payload. On each payload, one of the three APL photometers were filtered. The filtered photometers were flown with the intention of measuring the O_2 emission intensity. A filter centred at 7630 Å with a bandwidth of 120 Å was employed on one of the photometers flown in salvo B, and a filter centred at 7625 Å with a bandwidth of 125 Å was employed on one of the photometers flown in salvo A.

3.4.2.1 1st August 1991 (Salvo B)

At the time of the salvo B launch (PMSE, 1st of August 1991), it was noted that the riometer, and the magnetometer records (Pfaff, personal communication) indicated that salvo B was launched during the recovery phase of a small geomagnetic substorm. These findings were verified by the energy deposition measurements on each EFIELD B payload, which showed a weak particle precipitation during the PMSE night.

As aforementioned in Chapter 2, the Black Brant flights suffered the most difficulties. Salvo B developed attitude problems when the Daughter Section collided with the Mother Section payload shortly after separation. As a result of this collision, modifications were made in the timing sequence for various activities during the upleg flight (see Table 2.3) to avoid a further collision in Salvo A.

The APL preliminary results for both flights have been reported [East, 1992][116], however, no NLC characteristics could be discerned during salvo B, because of the severe coning by the payload. The unfiltered photometers flown during this salvo appeared to indicate that the payload may have passed through an aerosol layer, possibly a cloud, but little else could be differentiated. From the hard copies of the magnetometer, and riometer data, it was highlighted in a post flight plan meeting (Goldberg, personal communication) that the Black Brant may have passed through an extremely weak, and tenuous aerosol layer.

However, no other data sets obtained from this payload confirmed this observation. The rocket attitude problems compromised the data quality, thus inhibiting the primary objectives from being met. The results are still in their preliminary state - a more sophisticated analysis is required ensuring that noise is not mistaken as spatial morphologies.

3.4.2.2 9th August 1991 (Salvo A)

During salvo A, the magnetometer, and riometer indicated that the salvo was launched into a relatively quiet period, very little disturbance was recorded (Pfaff, personal communication).

Salvo A was launched into a NLC display, that had been observed some distance away from Esrange. At 22:30:00 UT, the aircraft observer, 500 km South of Esrange, had confirmed that the cloud cover had moved in above the rocket range. At this time, the CUPRI radar was also detecting the PMSE, the campaign scientists made the decision to proceed with salvo A. By 23:45:00 UT, the rockets were launched, the PMSE layer began to fade. Visual confirmation of the NLC above Esrange continued until 0:30:00 UT, (the 10th of August), when the contrast of the clouds against the brightening sky became too low. A PMSE layer resurfaced briefly between 01:44:00, and 02:05:00 UT.

During the salvo A flights, EFIELD A payload suffered similar attitude problems as that experienced by EFIELD B payload. The laterally scanning APL photometers did detect an aerosol layer

on both the upleg, and downleg flights. Again, with the severe rocket attitude problems sustained, the data presented here (see Figures 3.20a, and b) are still in their preliminary state. However, it may be possible (pending further analysis) to extract the two-dimensional structural information from salvo A Mother Section photometers during both the upleg and downleg flights. Figure 3.20a indicates a feature at a height of 74 km. There was some concern with such a result especially when the height of a NLC is usually at ~82 km. On re-plotting, and seeking advice, the feature still occurred at 74 km. It is proposed that the upleg feature may be spurious. The downleg results indicate an aerosol layer around 82 km. This result was confirmed by the conductivity data obtained from both the Mother and Daughter payloads. A "ledge" was observed at ~85 km. During salvo A, Pfaff noted an approximate 30% decrease in the positive conductivity, and an approximate 50% decrease in the negative conductivity at about 84 km, which is directly above the NLC layer altitude. Again, these results were consistent with the energy deposition results [Hale, 1992][117]. However, Pfaff remains uncertain as to whether an AC field was in this data set.

The results obtained from the EFIELD A payload were confirmed further with the results obtained from the ion mass spectrometer (SIDAMS) flown on the DECIMALS A payload. Very heavy water clusters with a high concentration were observed within the NLC layer height. Furthermore, much layering of the various water cluster species appeared to occur within the cloud region. TURBO A showed an enhancement in electron density. These measurements

were made nearly simultaneously with DECIMALS A, illustrating a spike around the correct height, but this result is still under investigation.

Further, although the APL photometers were only flown on the first two salvos, it should be noted that during the final salvo, C, the lowest temperature of 100 K was recorded [Schmidlin, 1992][117]. This is the lowest temperature on record, and as yet no satisfactory theory has been proposed. The resultant low temperature could be due to a tropospheric storm, but until the meteorological data records can be studied, this outcome will have to remain unanswered.

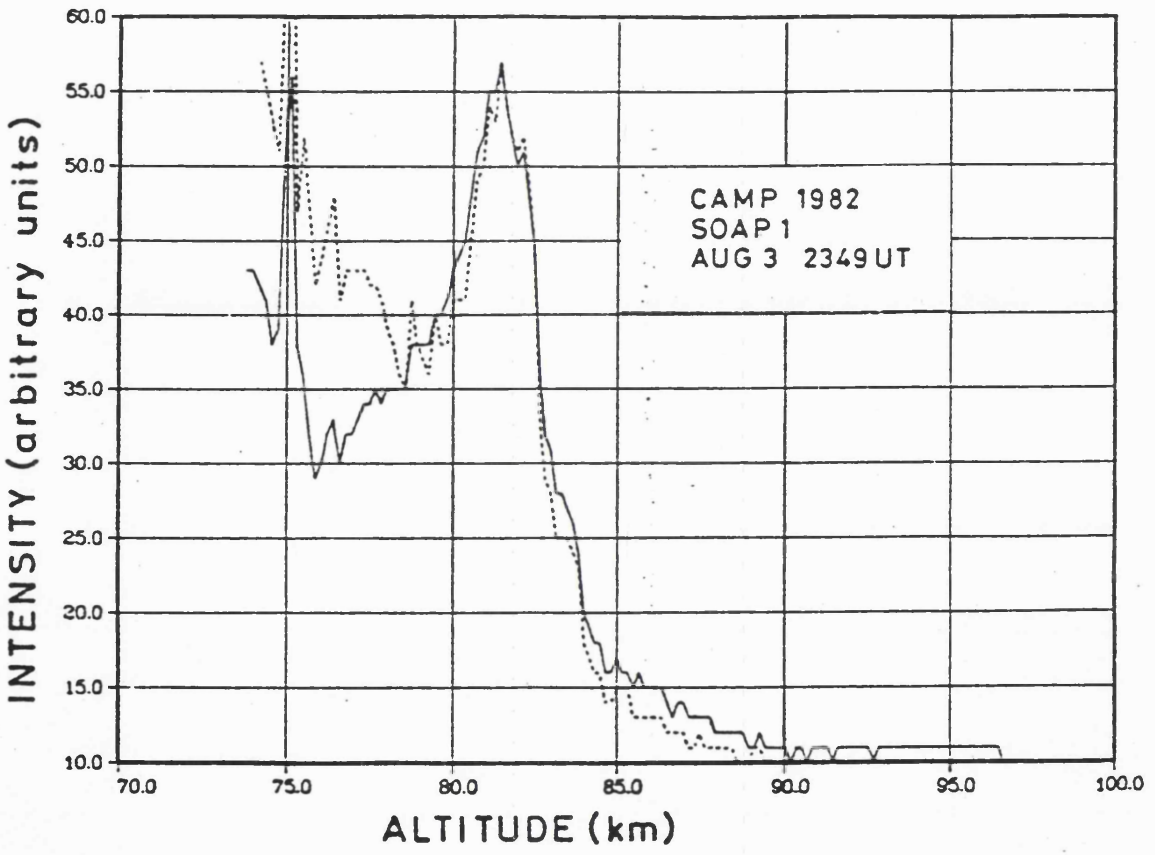

CAMP 1982
SOAP 1
AUG 3 2349 UT
INTENSITY (arbitrary units)
ALTITUDE (km)
60.0
55.0
50.0
45.0
40.0
35.0
30.0
25.0
20.0
15.0
10.0
70.0
75.0
80.0
85.0
90.0
95.0
100.0

Figure 3.1

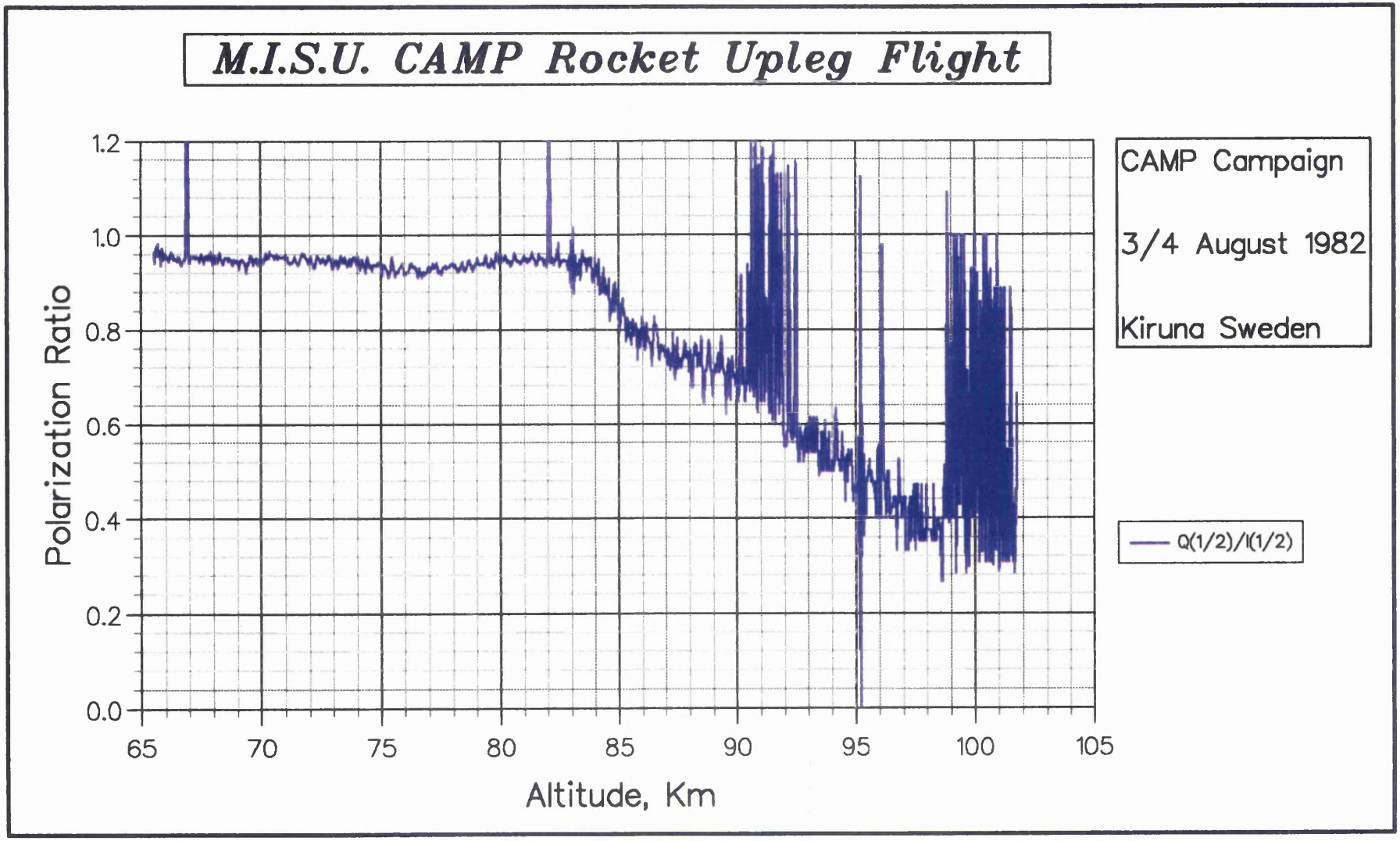

M.I.S.U. CAMP Rocket Upleg Flight
CAMP Campaign
3/4 August 1982
Kiruna Sweden
Q(1/2)/I(1/2)
Polarization Ratio
1.2
1.0
0.8
0.6
0.4
0.2
0.0
65
70
75
80
85
90
95
100
105
Altitude, Km

Figure 3.2

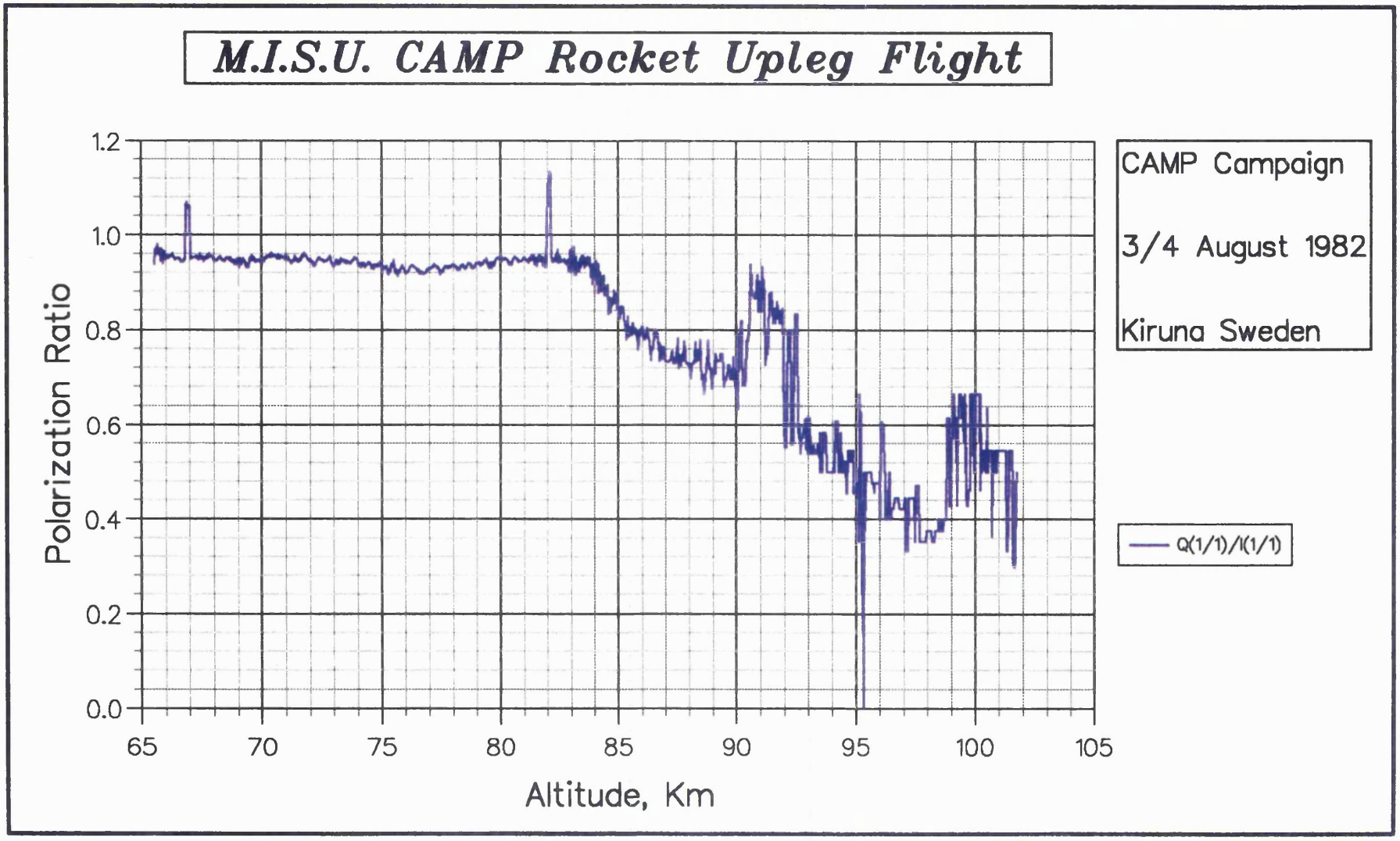

M.I.S.U. CAMP Rocket Upleg Flight
Polarization Ratio
1.2
1.0
0.8
0.6
0.4
0.2
0.0
65
70
75
80
85
90
95
100
105
Altitude, Km
CAMP Campaign
3/4 August 1982
Kiruna Sweden
Q(1/1)/I(1/1)

Figure 3.3

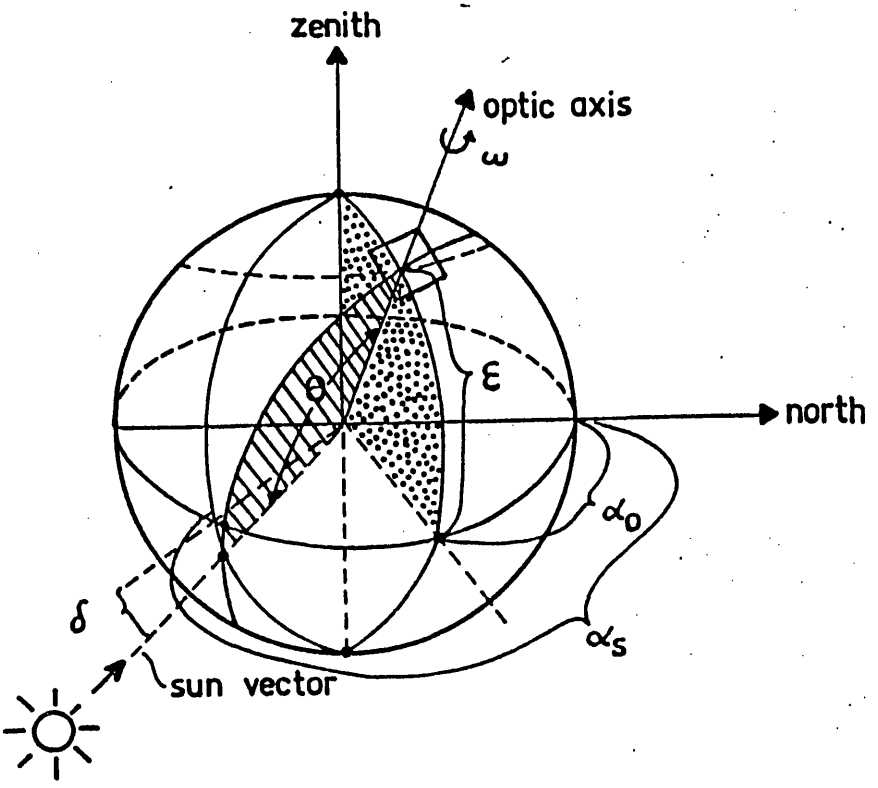


Figure 3.4

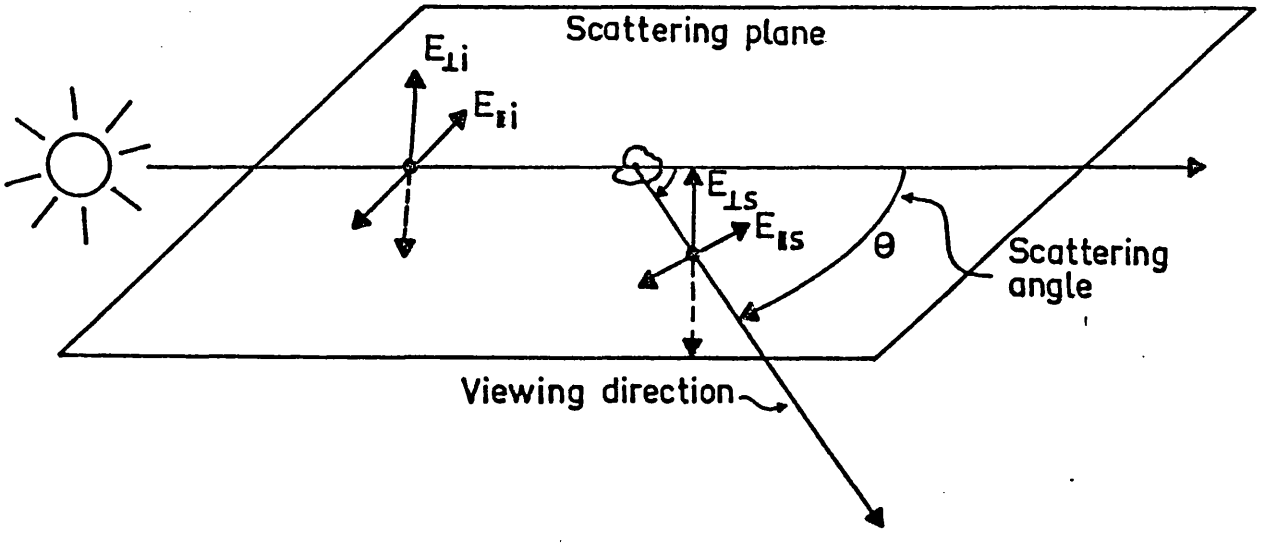


Figure 3.5

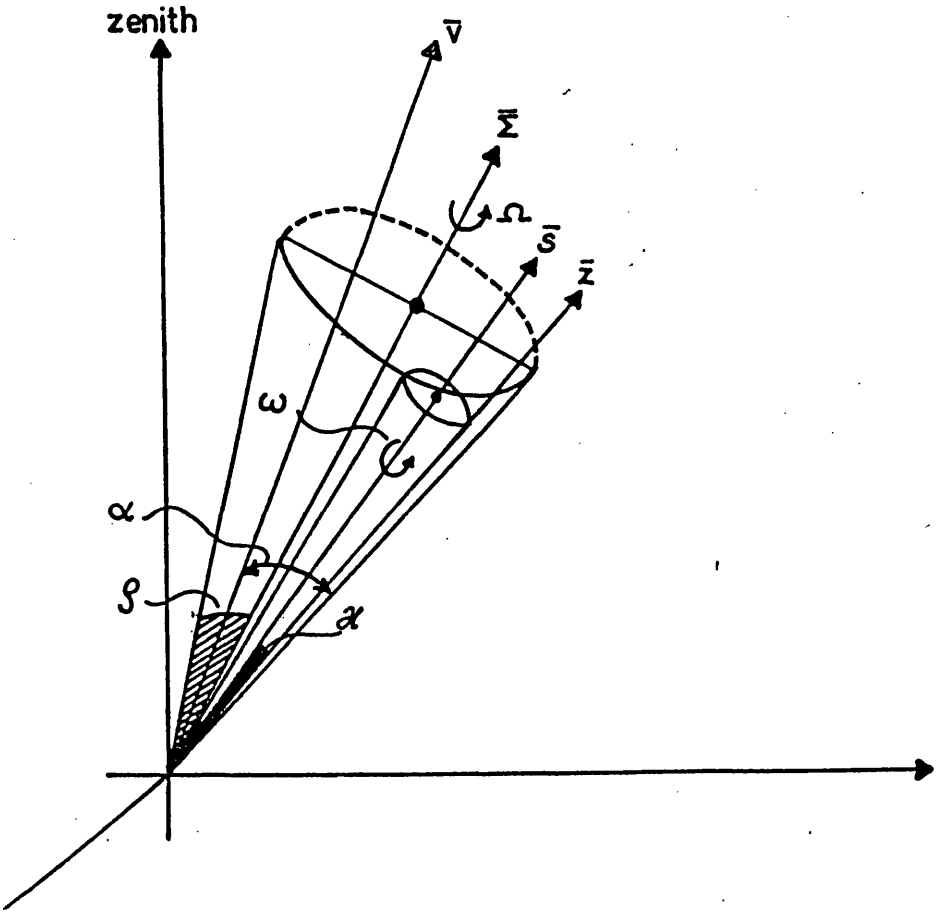


Figure 3.6

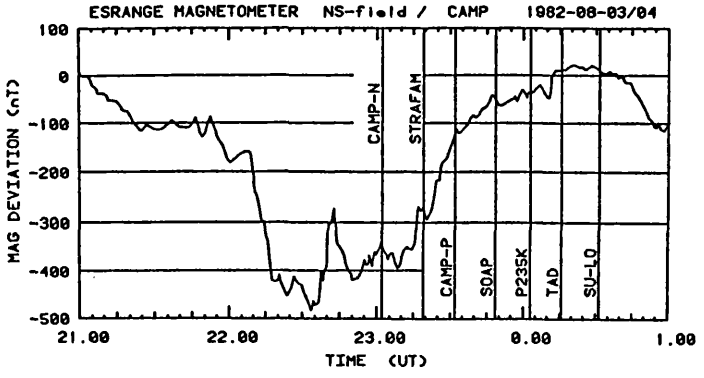


Figure 3.7

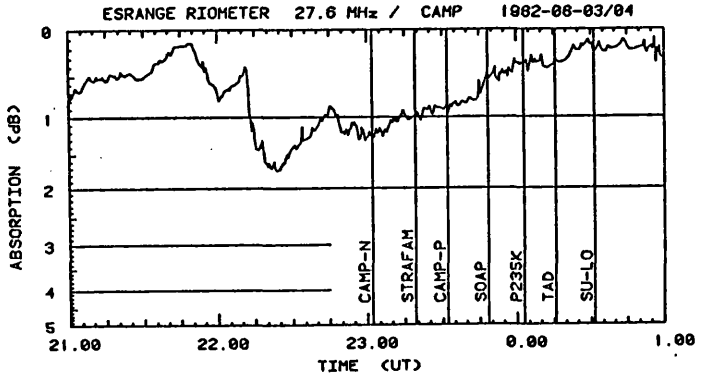


Figure 3.8

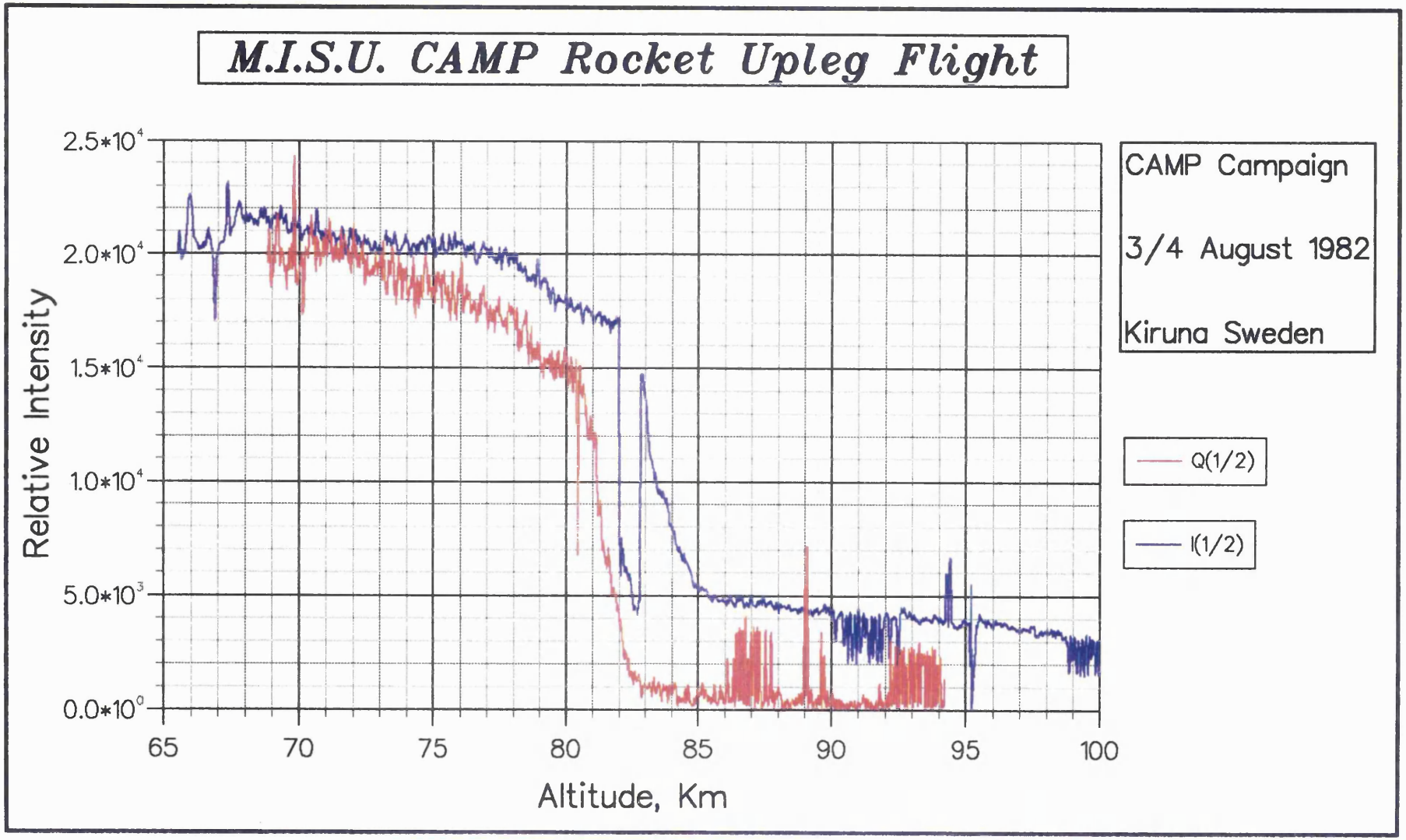

M.I.S.U. CAMP Rocket Upleg Flight
Relative Intensity
2.5*10^4
2.0*10^4
1.5*10^4
1.0*10^4
5.0*10^3
0.0*10^0
65
70
75
80
85
90
95
100
Altitude, Km
CAMP Campaign
3/4 August 1982
Kiruna Sweden
Q(1/2)
I(1/2)

Figure 3.9

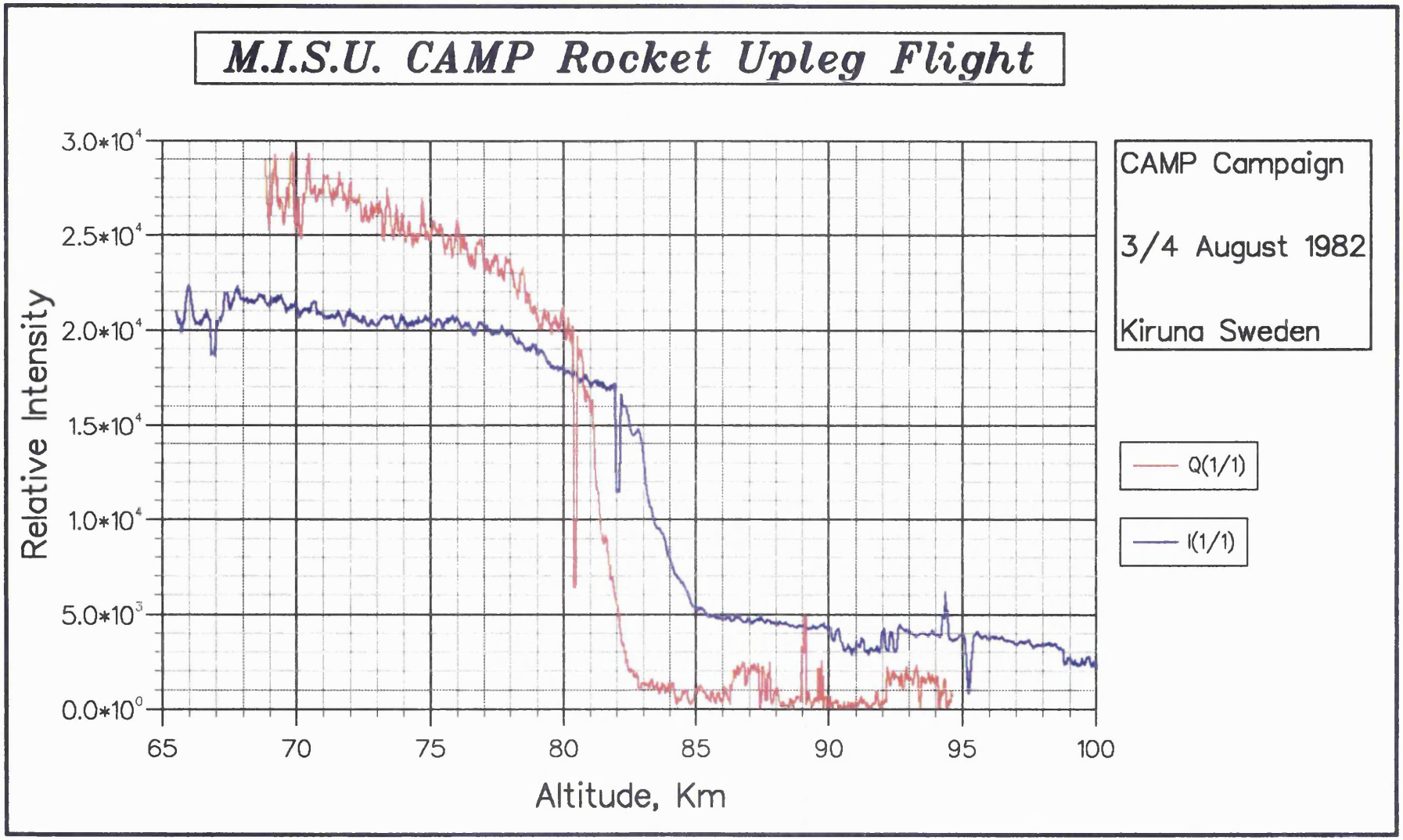

M.I.S.U. CAMP Rocket Upleg Flight
CAMP Campaign
3/4 August 1982
Kiruna Sweden
Q(1/1)
I(1/1)
Relative Intensity
3.0*10^4
2.5*10^4
2.0*10^4
1.5*10^4
1.0*10^4
5.0*10^3
0.0*10^0
65
70
75
80
85
90
95
100
Altitude, Km

Figure 3.10

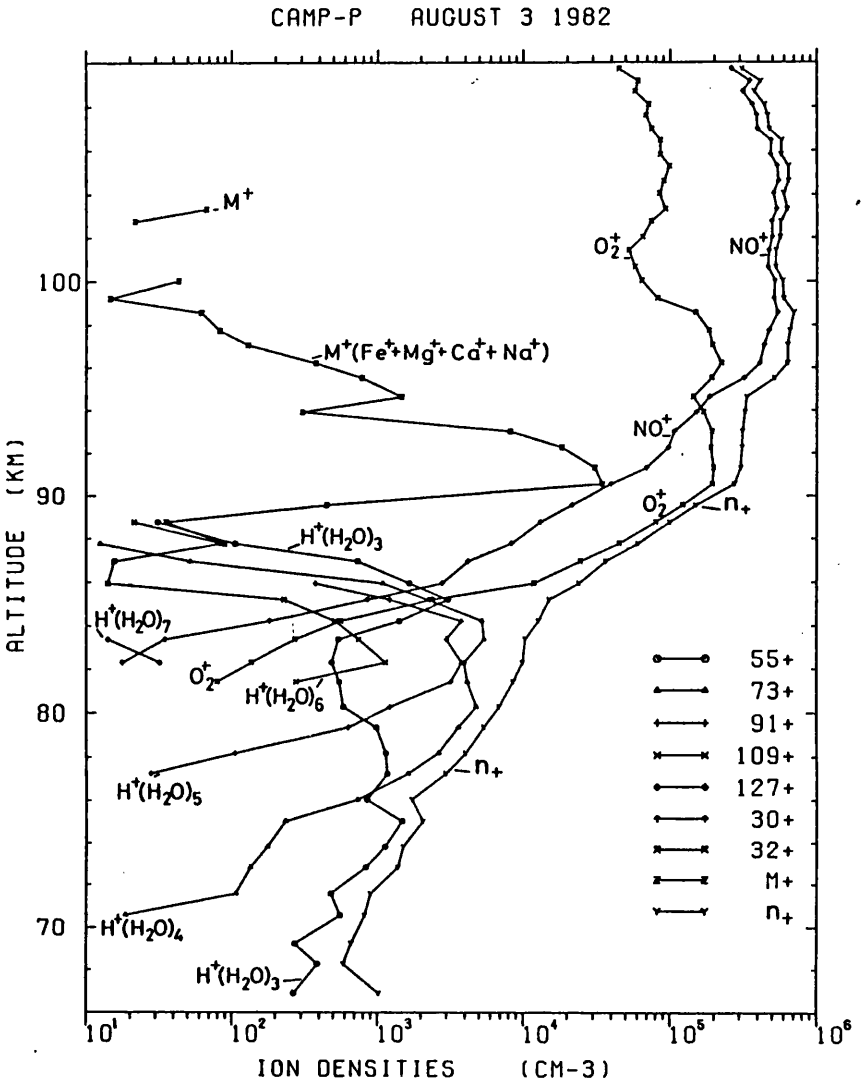

CAMP-P AUGUST 3 1982
ALTITUDE (KM)
100
90
80
70
10^1
10^2
10^3
10^4
10^5
10^6
ION DENSITIES (CM-3)
M^+
$M^+(Fe^+ + Mg^+ + Ca^+ + Na^+)$
O_2^+
NO^+
NO^+
O_2^+
n_+
$H^+(H_2O)_3$
$H^+(H_2O)_7$
O_2^+
$H^+(H_2O)_6$
n_+
$H^+(H_2O)_5$
$H^+(H_2O)_4$
$H^+(H_2O)_3$
55+
73+
91+
109+
127+
30+
32+
M+
n_+

Figure 3.11

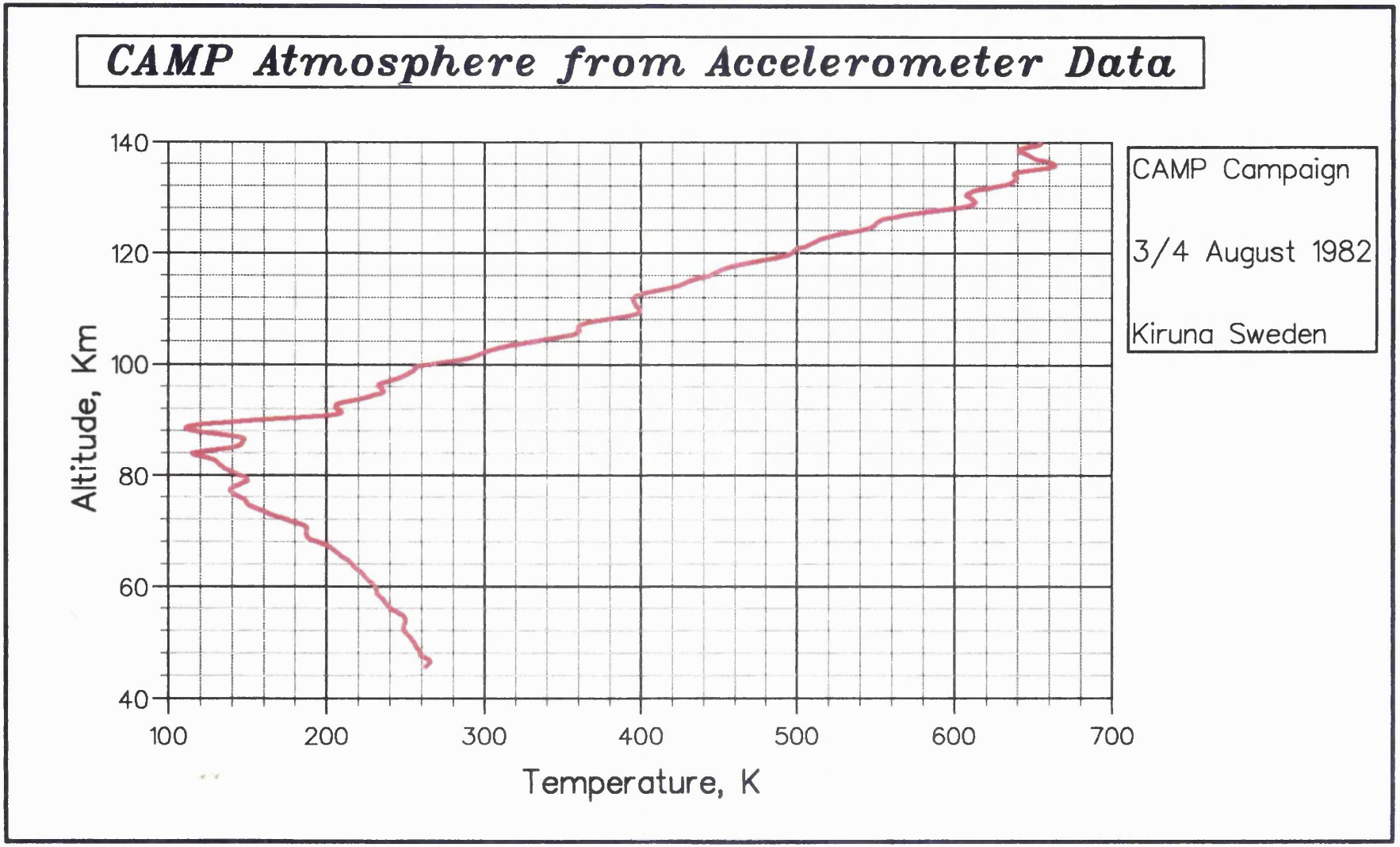

CAMP Atmosphere from Accelerometer Data
CAMP Campaign
3/4 August 1982
Kiruna Sweden
Altitude, Km
140
120
100
80
60
40
100
200
300
400
500
600
700
Temperature, K

Figure 3.12

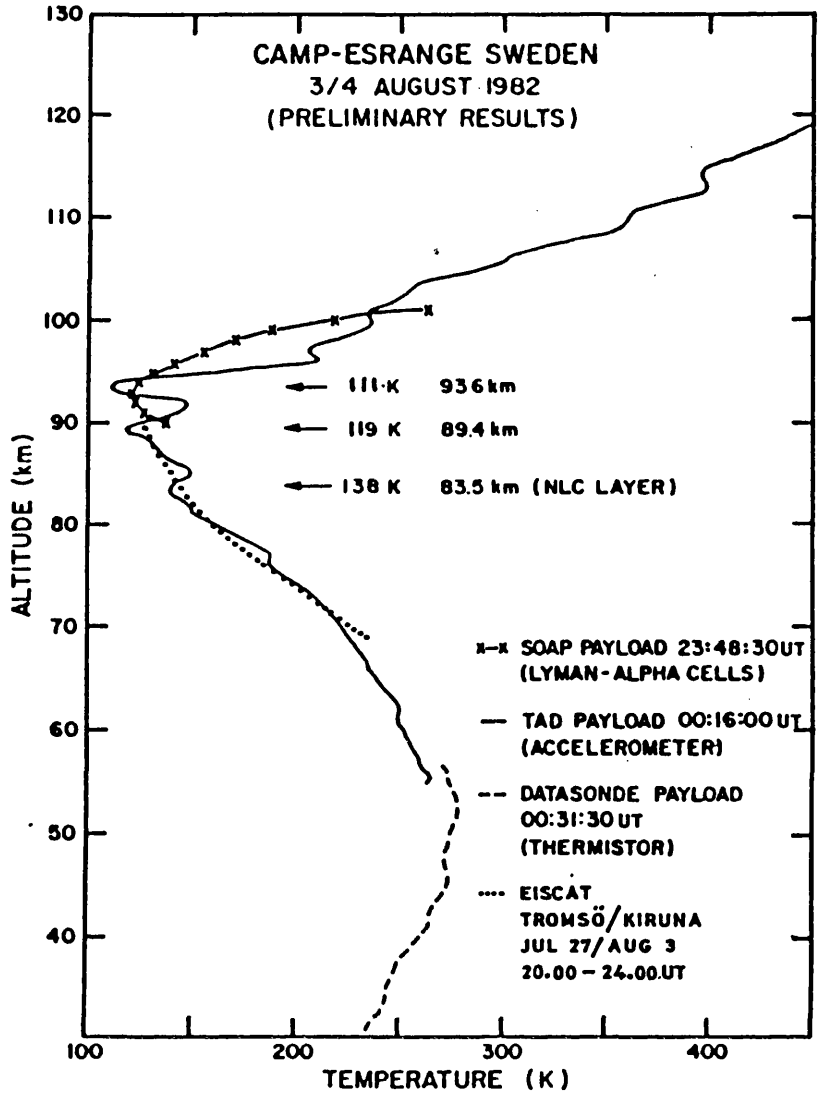


Figure 3.13

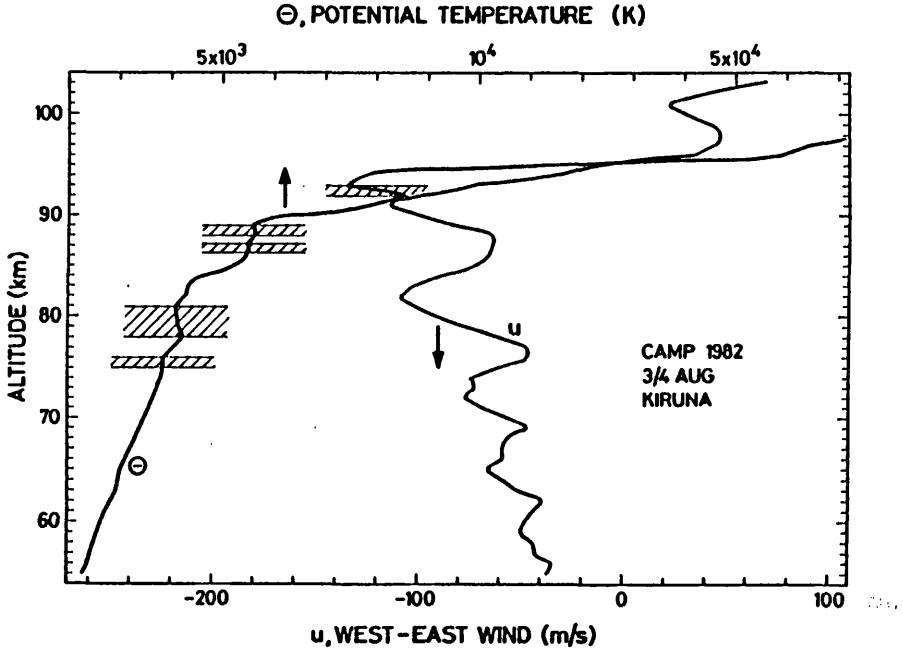


Figure 3.14

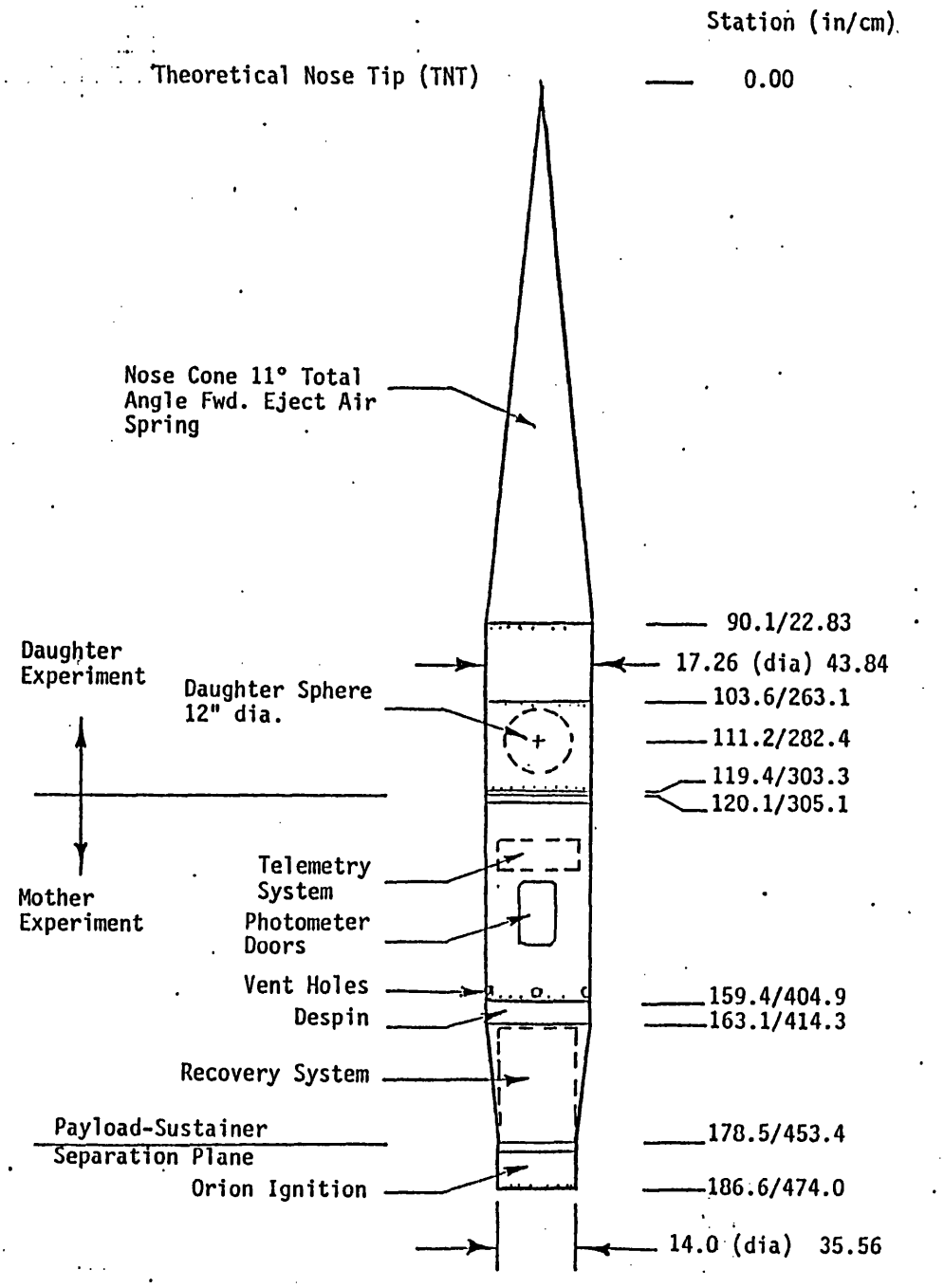

Station (in/cm)
Theoretical Nose Tip (TNT)
0.00
Nose Cone 11° Total Angle Fwd. Eject Air Spring
90.1/22.83
Daughter Experiment
17.26 (dia) 43.84
Daughter Sphere 12" dia.
103.6/263.1
111.2/282.4
119.4/303.3
120.1/305.1
Telemetry System
Mother Experiment
Photometer Doors
Vent Holes
159.4/404.9
Despin
163.1/414.3
Recovery System
Payload-Sustainer Separation Plane
178.5/453.4
Orion Ignition
186.6/474.0
14.0 (dia) 35.56

Figure 3.15

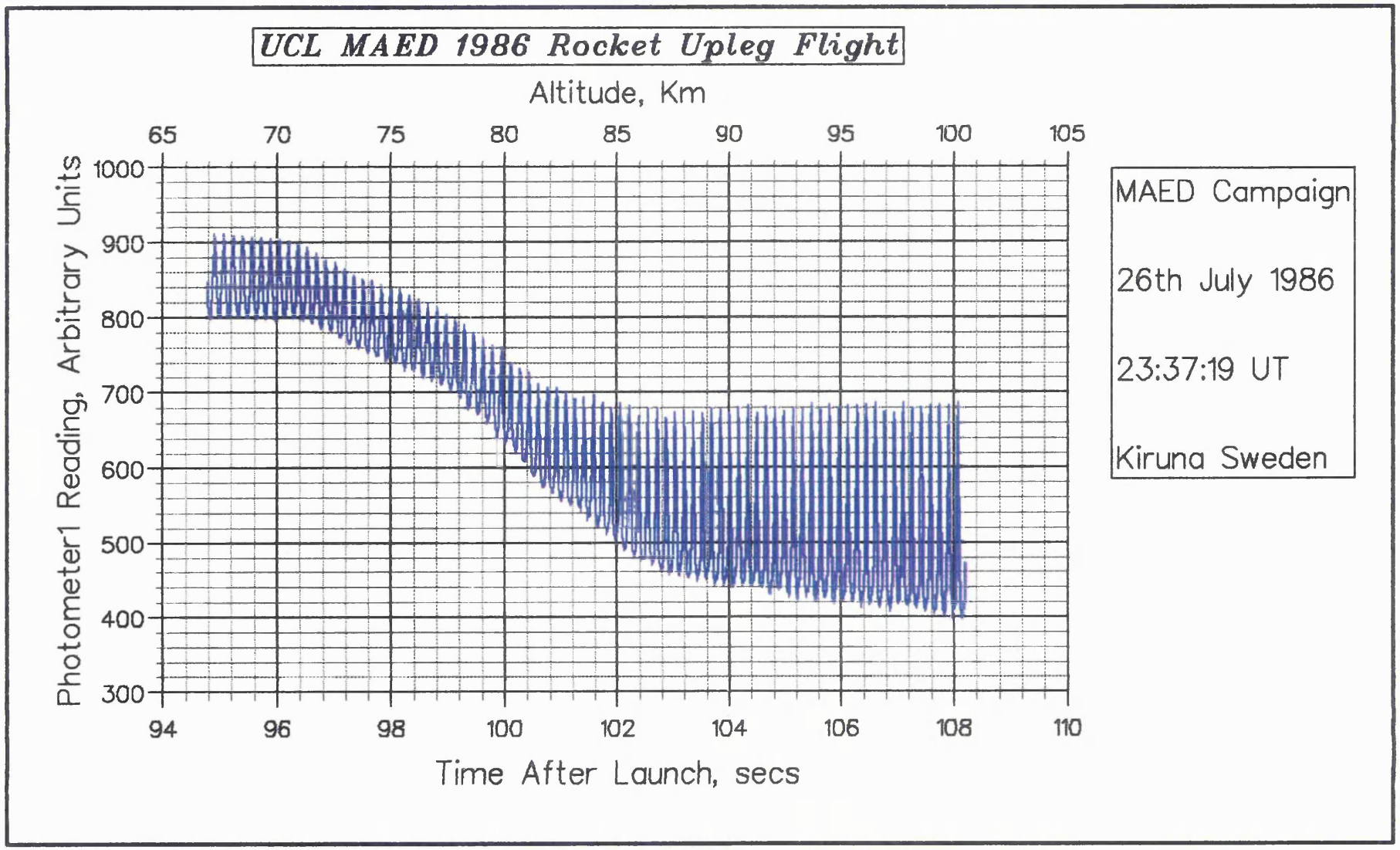

UCL MAED 1986 Rocket Upleg Flight
Altitude, Km
65
70
75
80
85
90
95
100
105
Photometer1 Reading, Arbitrary Units
1000
900
800
700
600
500
400
300
94
96
98
100
102
104
106
108
110
Time After Launch, secs
MAED Campaign
26th July 1986
23:37:19 UT
Kiruna Sweden

Figure 3.16a

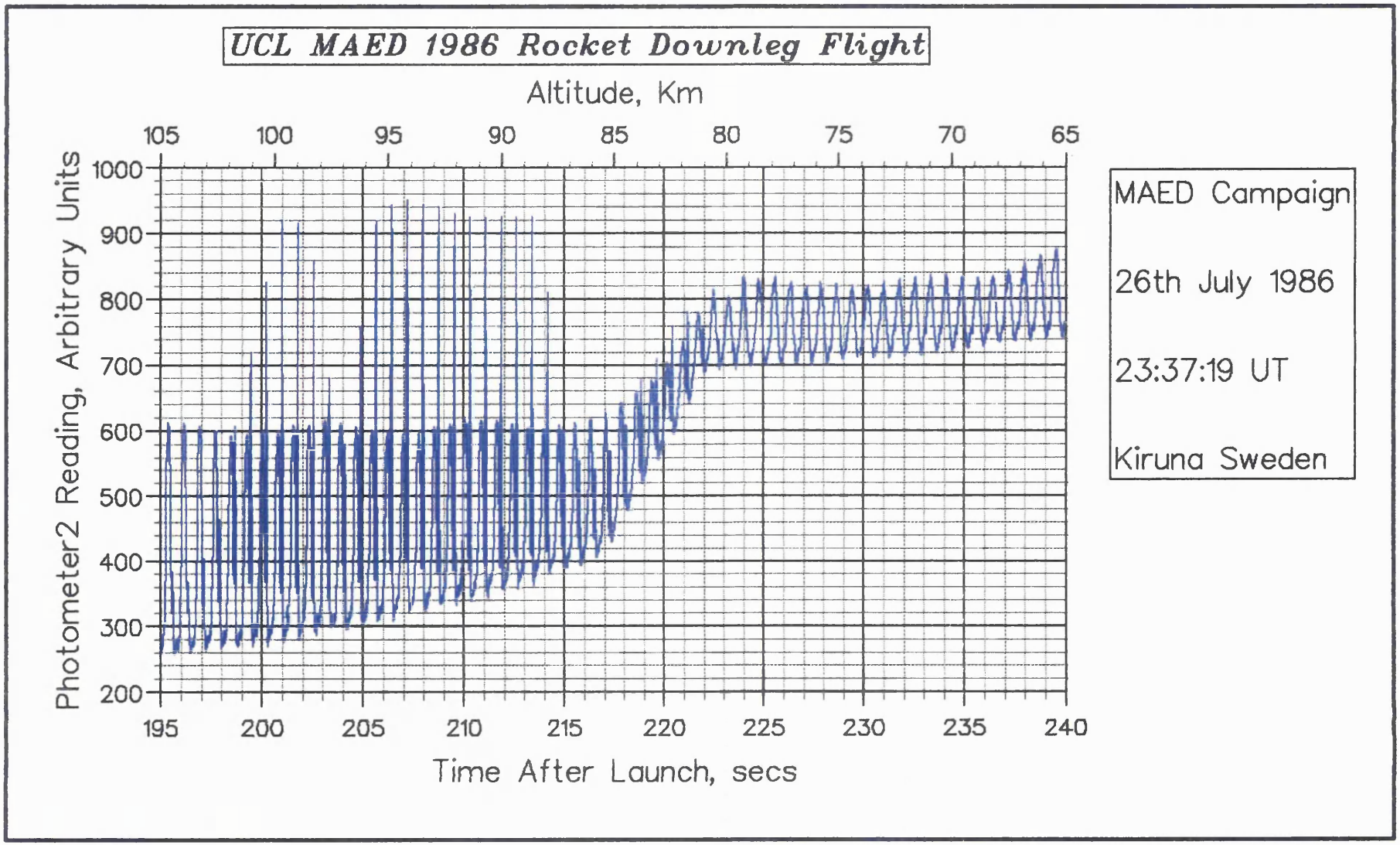

UCL MAED 1986 Rocket Downleg Flight
Altitude, Km
105
100
95
90
85
80
75
70
65
Photometer2 Reading, Arbitrary Units
1000
900
800
700
600
500
400
300
200
195
200
205
210
215
220
225
230
235
240
Time After Launch, secs
MAED Campaign
26th July 1986
23:37:19 UT
Kiruna Sweden

Figure 3.16b

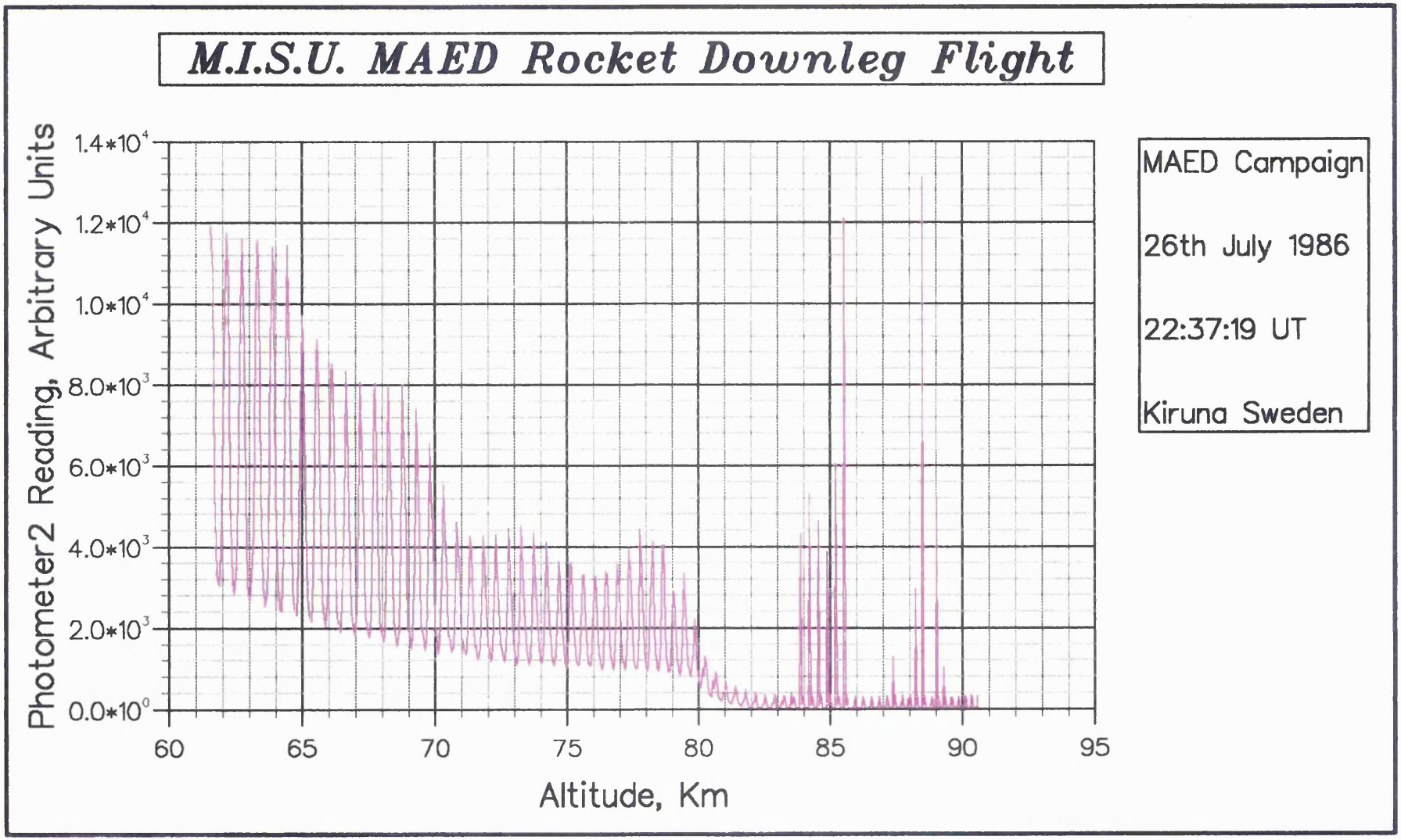

M.I.S.U. MAED Rocket Downleg Flight
Photometer2 Reading, Arbitrary Units
1.4*10^4
1.2*10^4
1.0*10^4
8.0*10^3
6.0*10^3
4.0*10^3
2.0*10^3
0.0*10^0
60
65
70
75
80
85
90
95
Altitude, Km
MAED Campaign
26th July 1986
22:37:19 UT
Kiruna Sweden

Figure 3.17

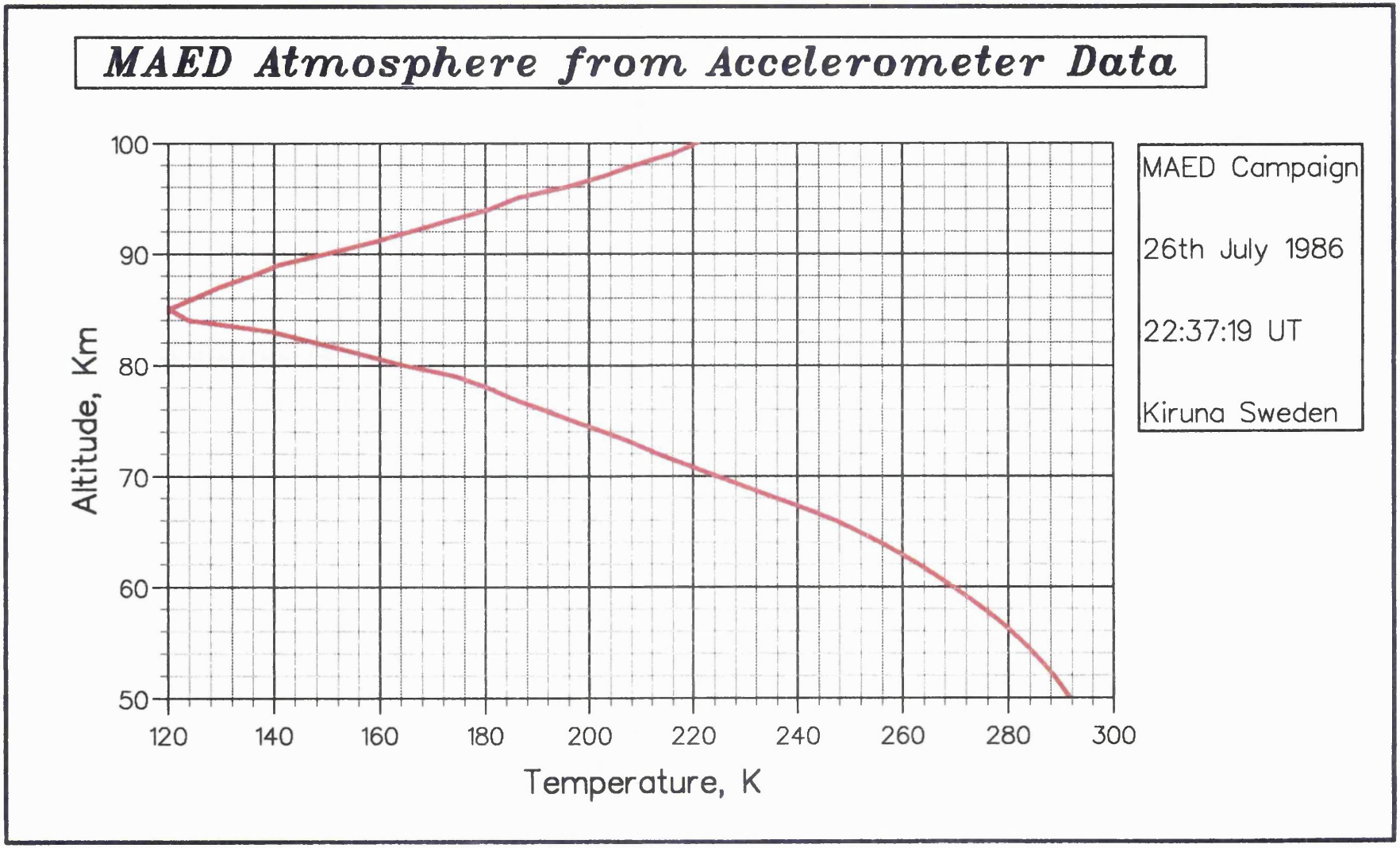

MAED Atmosphere from Accelerometer Data
100
90
80
70
60
50
Altitude, Km
120
140
160
180
200
220
240
260
280
300
Temperature, K
MAED Campaign
26th July 1986
22:37:19 UT
Kiruna Sweden

Figure 3.18

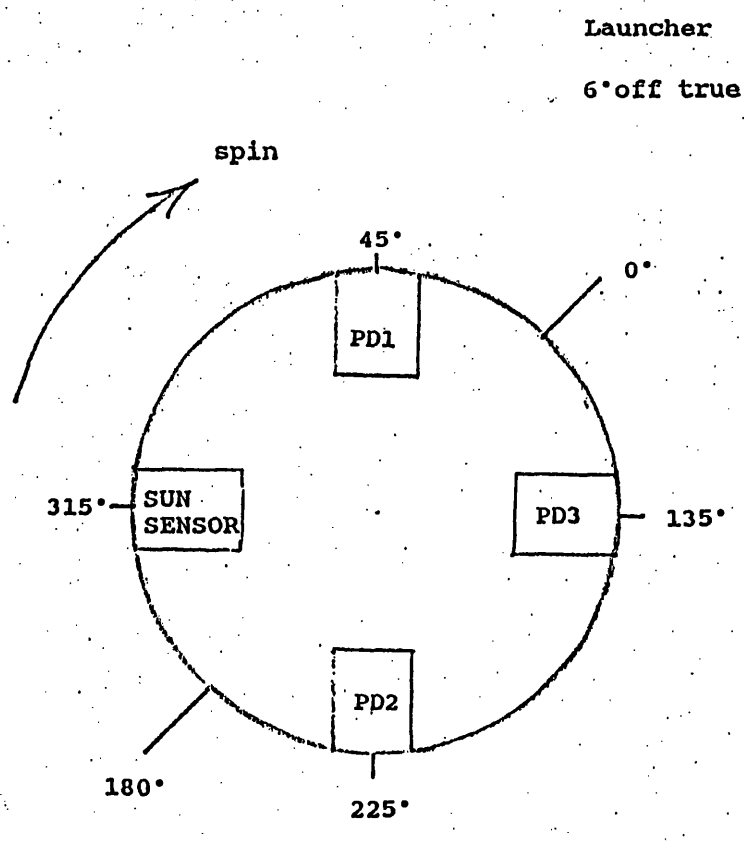

Launcher
6°off true North
spin
45°
0°
PD1
315°
SUN
SENSOR
PD3
135°
PD2
180°
225°
London instruments looking forward

Figure 3.19

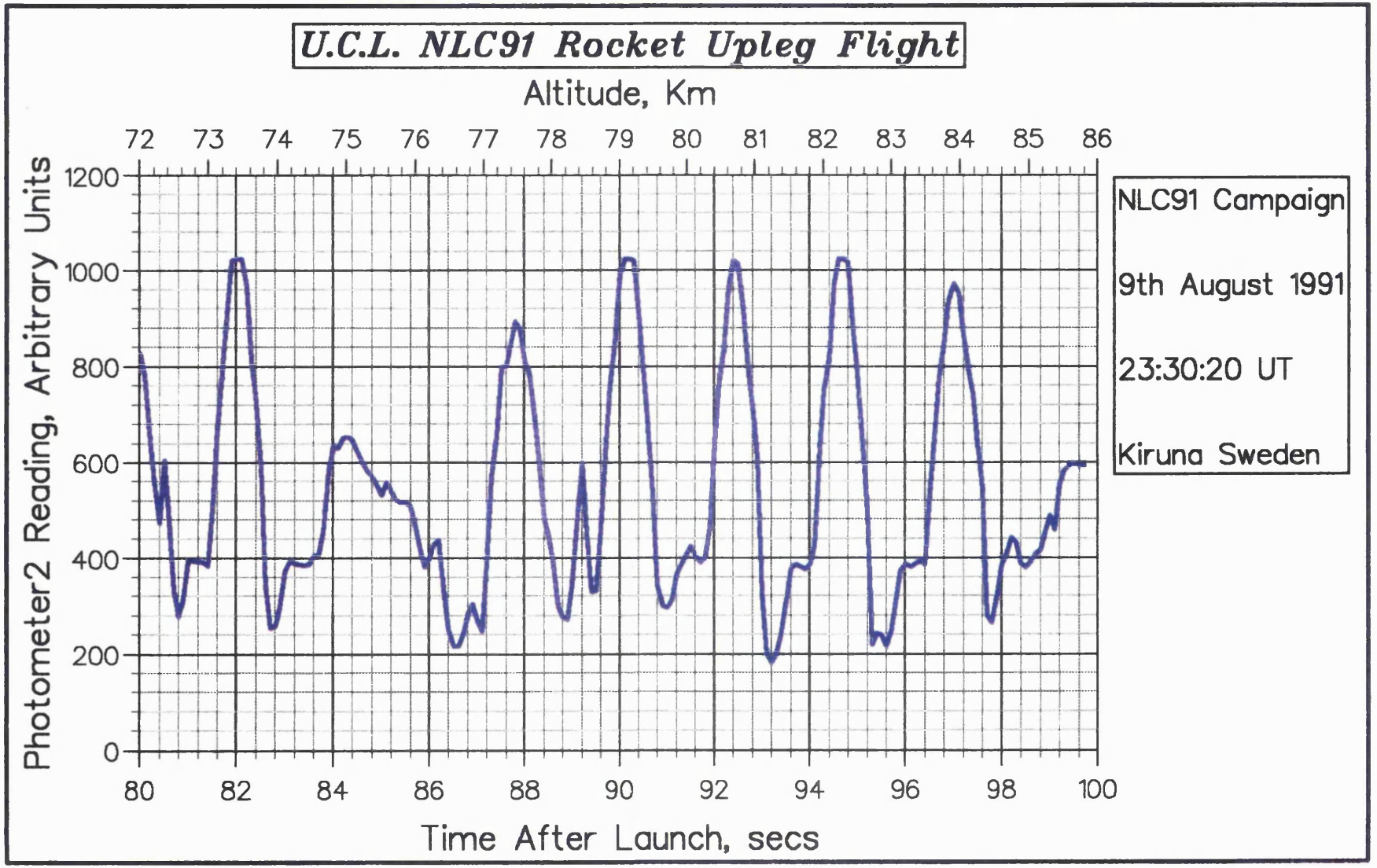

U.C.L. NLC91 Rocket Upleg Flight
Altitude, Km
72 73 74 75 76 77 78 79 80 81 82 83 84 85 86
Photometer2 Reading, Arbitrary Units
0 200 400 600 800 1000 1200
80 82 84 86 88 90 92 94 96 98 100
Time After Launch, secs
NLC91 Campaign
9th August 1991
23:30:20 UT
Kiruna Sweden

Figure 3.20a

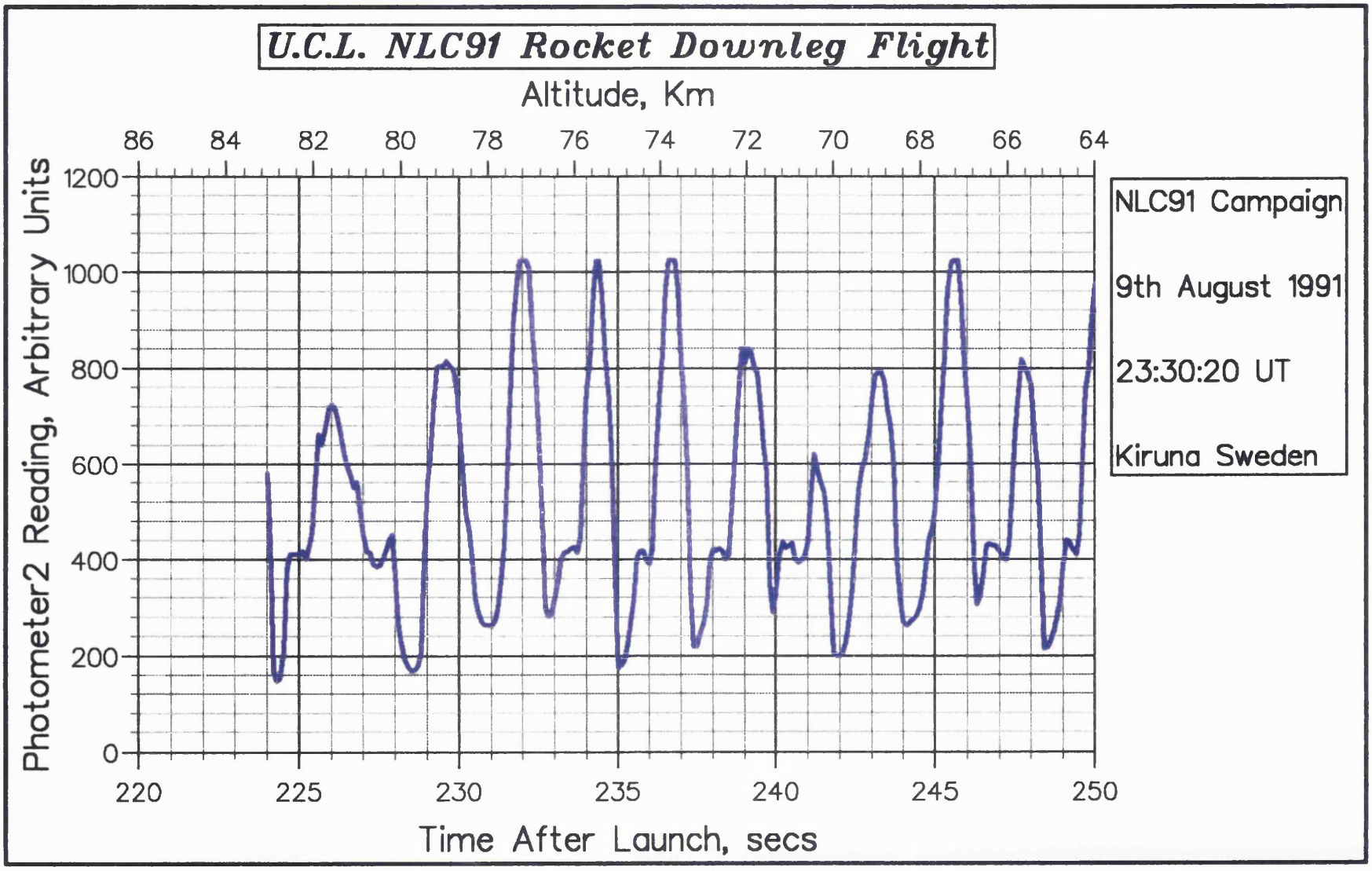

U.C.L. NLC91 Rocket Downleg Flight
Altitude, Km
86
84
82
80
78
76
74
72
70
68
66
64
Photometer2 Reading, Arbitrary Units
1200
1000
800
600
400
200
0
220
225
230
235
240
245
250
Time After Launch, secs
NLC91 Campaign
9th August 1991
23:30:20 UT
Kiruna Sweden

Figure 3.20b

Chapter 4

4 The Application of the Bomem Michelson Interferometer to the Study of the Hydroxyl (3,1) Band Radiance and Temperature

4.1 Introduction

This chapter describes the ground-based instrumentation deployed in the observations of the near infra-red (NIR) hydroxyl (OH) emission from the Meinel (3,1) band at 1.0 to 1.6 μm.

Measurement of the OH (3,1) band is achieved using a Bomem Michelson Interferometer (MI) operating in the NIR region of the spectrum at ~0.0025 μm resolution. Using the Bomem MI has the advantage that the reduced light scattered in this spectral range, provides the possibility of measuring the hydroxyl temperature and radiance from the ground with the Sun only 3.5° below the horizon. The principles of operation, and basic design of the Bomem MI are discussed in Section 4.2.

The author had the opportunity to spend two months during 1992 working closely with Dr Patrick Espy, of the Space Dynamics Laboratory (Utah State University, USA), studying the USU Bomem MI data collected during the summer rocket campaign, NLC-91. The author re-analyzed, and interpreted the resulting data giving a measure of the mesospheric conditions during the absence/presence of NLC during the rocket campaign. The data analysis and routines are discussed in Section 4.3.

The interpretation of the raw data gives a strong indication of stratospheric filtering whose diminution could produce the upward forcing that may be involved in the NLC formation processes. Intensity, and rotational temperature profiles, from other campaigns, deduced during the absence/presence of NLC give clear indications of small- and large-scale waves. These results are discussed in Section 4.4.

4.2 The Michelson Interferometer

4.2.1 Basic Design

Figure 4.1 illustrates the arrangement of a typical Michelson Interferometer (MI). Light rays from a source S are incident at 45° on a glass plate B, which is partially silvered. At B, the incident light beam is split into two beams, one of which is reflected onto mirror M_2, and the other is transmitted through B onto a mirror M_1. Mirrors M_1, and M_2 are mutually perpendicular, and are both perpendicular to their respective incident beams. Both beams are reflected from the mirrors back onto plate B, where they recombine, and are reflected onto the detector where an interference pattern may be measured. A virtual image of M_1 exists, parallel to M_2 as shown in the diagram. These planes M_1', and M_2 constitute a virtual plane-parallel layer of optical thickness equal to twice the optical path difference between plate B and M_1, and plate B and M_2.

Mirror M_2 is capable of moving parallel to the incident beam

thereby increasing or decreasing the optical path difference (OPD) between the beams. If this difference is equal to a full wavelength or an integral number of wavelengths of a particular radiation then the combining beams will interfere constructively, and the detector will measure maximum intensity, while if it is a half wavelength or an odd number of half wavelengths of this radiation, then the beams will interfere destructively giving a minimum intensity reading at the detector. If the instrument is used to observe a monochromatic source, a sinusoidal interference pattern will be measured by the detector as the mirror is moved. Figure 4.2 illustrates a resultant interferogram of a monochromatic source.

A Fourier Transform (FT) of the incident spectrum will be discussed in Section 4.3. The centre of this pattern corresponds to the zero path difference (ZPD) position which occurs when the distance travelled by each of the two interfering beams exactly equal. This is important when recording interferograms since all frequencies yield constructive interference at this point.

The coherence length of an optical source is the maximum optical path difference between two beams for which they form an interference pattern. The coherence length varies with the degree of spectral purity of the source.

4.2.2 Theory of the Michelson Interferometer

The complex amplitude from the superposition of two coherent

beams of amplitudes A_1, and A_2, and wavenumber σ with an OPD of χ is given by:-

$$A = A_1 + A_2 \exp(i\delta) \qquad (4.1)$$

where $\delta = 2\ \pi\ \sigma\ \chi$, is the phase difference between the two beams. This results in an intensity given by:-

$$I_\sigma = A\ A^* = A_1^2 + A_2^2 + 2\ A_1\ A_2 \cos\delta \qquad (4.2)$$

When A1, and A2 are equal, then Equation (4.2) becomes:-

$$I_\sigma(\chi) = 2\ A^2\ (1 + \cos\delta) = I_{av}\ (1 + \cos(2\ \pi\ \sigma\ \chi)) \qquad (4.3)$$

If B is varied, then the signal at the detector, I_σ, will vary from a maximum of $2I_{av}$ to a minimum of 0. If the source has a spectral radiance $S(\sigma)$, the average flux reaching the detector within the range $\sigma + d\sigma$ is proportional to $S(\sigma)d\sigma$, and to the instrument response, a function of σ which is the result of transmission and reflection coefficients variation within the instrument as a function of wavenumber. If this source function is designated $B(\sigma)$, Equation (4.3) becomes:-

$$I_\sigma = \int_0^\infty B(\sigma)\ (1 + \cos\ (2\ \pi\ \sigma\ \chi))\ d\ \sigma$$

$$= \int_0^\infty B(\sigma)\ d\ \sigma + \int_0^\infty B(\sigma)\ \cos\ (2\ \pi\ \sigma\ \chi)\ d\ \sigma$$

$$= I_{av} + \int_0^\infty B(\sigma)\ \cos\ (2\ \pi\ \sigma\ \chi)\ d\ \sigma \qquad \textbf{(4.4)}$$

For negative values of σ, $B(\sigma)$ equals 0 which means that the lower integration limit of this term can be taken as $-\infty$ replacing zero. The second term in Equation (4.4) contains all the information on the incident spectrum. Re-defining $I(\chi)$ as the modulated section of the interferogram, replacing $I(\chi)$ with $I(\chi) - I_{av}$, results in the following expression:-

$$I_\sigma\ (\chi) = \int_{-\infty}^\infty B(\sigma)\ \cos\ (2\ \pi\ \sigma\ \chi)\ d\ \sigma \qquad \textbf{(4.5)}$$

Therefore, $I(\chi)$ is the cosine transform of $B(\sigma)$, and hence, the spectrum can be calculated from the interferogram, by taking the inverse cosine transform.

$$B(\sigma) = \int_{-\infty}^{\infty} I_{\sigma}(\sigma) \cos(2 \pi \sigma \chi) \, d\chi \qquad (4.6)$$

4.2.3 The Bomem Michelson Interferometer

4.2.3.1 The Mechanical and Optical Design

The mechanical design of an interferometric spectrometer is crucial to its operation. The sensitivity response and stability can be affected by the mechanical design of the system. The mirrors in modern day interferometers are cube corner reflectors. See Figure 4.3a. This form of mirror system has the ability of maintaining an output beam parallel to the input beam independent of their orientation. This ensures that the recombinant infra-red radiation from both arms intersects the beamsplitter at exactly the correct position, and angle for interference to occur.

Figure 4.3b illustrates the rigid arm clasping both mirrors in the Bomem instrument. The scan arm movement is rotary, in order that the distance between the mirrors does not change, but the orientation of the scan arm to the incident beam, and the beamsplitter does, thereby altering the path difference between the two beams. The "centre of gravity" of the scan arm arrangement lies at the pivot position, which allows the mechanism to be insensitive to the instrument orientation, and to vibrations. The advantage is that the Bomem can be operated even when on board an aeroplane as during the ALOHA-90 Campaign

(Airborne Lidar and Observations of the Hawaiian Airglow) (Espy, personal communication). A linear frictionless motor controls the scan arm movement.

The pivot is built of two steel blades aligned perpendicular to one another at the top, and bottom of the scan assembly. The dimensions of the pivot blades have been optimized to obtain a very sturdy pivot in the perpendicular direction but allowing the scan arm to have a free horizontal rotation. The flexibility of the steel blades is such that the movement required for maximum rotation never exceeds the elastic deformation range of the steel ensuring a long lifetime for the complete assembly.

The instrument houses an internal Helium-Neon (He-Ne) laser which is used to maintain a constant wavelength sampling interval. The laser beam is aligned along the same optical path as the incident atmospheric radiation as shown in Figure 4.3. The small area of the beamsplitter, and other optical components used for the radiation from the laser are optimized for this wavelength, as opposed to the remainder is optimized for the IR radiation. Figure 4.3 illustrates that the laser beam is split by the same beamsplitter arrangement as the IR light. A sinusoidal interference pattern is detected by a separate Indium Gallium Arsenide (InGaAs) detector as the path difference is altered. This corresponds to the case of an essentially monochromatic source, which was described in Section 4.2.1. The conversion of the IR signal from analog to digital is triggered at the zero crossings of the laser fringe signal. This gives an accurate, and

constant sampling interval of $\lambda_{He\text{-}Ne}/2$ nm.

The two detectors measuring the internal laser light are illustrated in Figure 4.3. The outputs of these detectors determine the direction of the scan arm, using quadrature detection, in order to correctly monitor, and control its speed and direction of motion. The laser beam is initially randomly polarized, and can therefore be observed as two individual but collinear beams, one horizontally polarized, and the other vertically polarized. The beams are separated, and recombined as previously described in Section 4.2.1, but in one of the arms, the beams are transmitted through a birefringent optical material, which has a different refractive index of the horizontally and vertically polarized beams, such that there is an optical retardation relationship between the beams. At the output, these two beams are isolated, and their interference patterns are detected separately. A phase shift will exist between these patterns. The direction of the scan arm determines which pattern lags/leads the other. See Figure 4.4. The waveforms A and B represent the intensities measured at the two laser signal detectors as a function of the OPD. If the optical movement has the direction of arrow 1, waveform A will appear to lead waveform B as shown in Figure 4.4. If this movement is in direction 2, waveform B will lead that of A as shown in Figure 4.4. The digital electronics recognize the relative phase relationship, and therefore, can deduce the direction of the scan arm from the two signals.

The ZPD position is located in the Bomem MI by means of an internal white light source, and a corresponding detector. Each time the interferometer is switched on, a small light bulb is powered, radiation from which propagates through the interferometer using the same portion of the beamsplitter that the laser beam utilizes. The scan arm is rotated, and an interferogram recorded. At the location of ZPD position, the interferogram intensity increases dramatically allowing its exact position to be identified; the white light source is then extinguished. The complete procedure takes about thirty seconds, and the instrument has a visible indicator LED which is illuminated during the search for the ZPD position.

4.3 Data Analysis

4.3.1 Resolution and Apodization

The maximum frequency which may be determined by the interferometer is referred to as the Nyquist frequency, N_f.

The Nyquist frequency is expressed as:-

$$N_f = \frac{1}{2\,\Delta\,\chi} \qquad (4.7)$$

where $\Delta\chi$ = the sampling interval = the change in the path difference between successive detector signal

measurements

The number of points in a recovered spectrum, the output from the transform program, TRNSFORM, will span the interval $-N_f$ to N_f. This number is identical to the number of sampling points recorded in the original interferogram. Therefore, the resolution element in the final spectrum is given as:-

$$df = \frac{2\ N_f}{number\ of\ sampling\ points} \qquad \textbf{(4.8)}$$

The resolution of the MI is determined by this minimum frequency interval, and hence, it follows that for a constant sampling interval, higher resolution may be obtained by increasing the maximum path difference of the interfering beams. The computational routines employed to process the cosine transform or a full FT utilize the Cooley-Tukey algorithm, which depends on the number of data points of the input data set to be an integral power of 2. The maximum number of sampling points used by the Bomem MI, is 32,768 (2^{15}).

Ideally, for the perfect spectrum, an infinitely small sampling interval, and an infinitely long maximum path difference would be preferred. However, the spectrum produced under real conditions will represent a convolution of what is known as the instrumental line shape, and the true spectrum. The instrumental line shape (ILS) for a MI is a sinc function, and is defined as:-

$$f(\sigma) = \frac{\sin(2\pi\sigma L)}{2\pi\sigma L} \qquad (4.9)$$

where σ = the frequency

L = $N\Delta\chi = \lambda_{max}/2$ = the maximum path difference between the ray paths

N = the maximum number of sampling points recorded by the instrument for a fixed $\Delta\chi$

$\Delta\chi$ = the sampling interval

This is the result of a cosine transform of a truncated pure cosine function, and is equivalent to the result obtained from transforming an interferogram of a monochromatic source. It results in a spectrum similar to that illustrated in Figure 4.5, where both negative and positive side-lobes are in evidence. The first negative side-lobe reaches a minimum of ~22% of the maximum intensity at σ_0. If a second, but weaker, spectral line was present in the spectrum at this frequency then it might not be seen in the computed spectrum. The solution to this problem is to perform what is known as apodization. This has the effect of increasing the width of the central "Airy disc", and reducing the intensity of the ring patterns making it possible to resolve the neighbouring spectral line. A compromise must be met, that is between the reduction of the side-lobes, and the loss in spectral resolution caused by the subsequent widening of the spectral lines. The maximum path difference, given in Equation (4.9) as 32,768/2 x λ/2, which implies an unapodized resolution of 1 cm^{-1}

using the Rayleigh criterion.

There are various apodization functions that can be employed. One simple and commonly used function, known as the triangulation apodization, has the form:-

$$A_1(\delta) = 1 - \left|\frac{\delta}{\Delta}\right| \qquad (4.10)$$

for $-\Delta \leq \delta \leq \Delta$

$$A_1(\delta) = 0 \qquad (4.11)$$

for $\delta > |-\Delta|$

where Δ = the maximum path difference = $2\ \pi\ \sigma\ L$

A FT of this function has the form:-

$$f(\sigma) = \Delta\ sinc^2(\pi\ \sigma\ \Delta) \qquad (4.12)$$

When $\sigma = 0$, then any function will be unity, and decreases with increasing path difference which will serve as an apodization function. A cosine apodization function:-

$$F(D) = \frac{1 + \cos(\pi D)}{2} \qquad \textbf{(4.13)}$$

where D = the OPD divided by the maximum OPD

This function is employed, providing a good compromise between the minimization of the side-lobes, and the deterioration of the spectral resolution.

Each interferogram is apodized, therefore, before transformation to a spectrum. This is achieved by multiplying each point on the interferogram by its corresponding point on the apodization function. Thus, the central point corresponding to the ZPD position is given a weighting of unity, while points on either side of this are given a reduced weighting extending all the way to the two end points which are given a weighting of zero.

4.3.2 Phase Correction

Equation (4.5) has been assumed up to now to be an accurate representation of the interferogram:-

$$I_\sigma(\chi) = \int_{-\infty}^{\infty} B(\sigma) \cos(2\pi\sigma\chi)\, d\sigma \qquad \textbf{(4.5)}$$

In practice, however, there is an additional term, the phase

angle, δ, to describe the actual measured interferogram. The corrections to the phase angle, may arise due to optical, electronic, or sampling effects. The proceeding two equations illustrate two important effects in the cosine term:-

I) If the interferogram is not symmetrical about $\delta = 0$, that is if the central sampling point does not correspond directly with the zero path difference, but corresponds to a phase angle of say $\delta = \epsilon$ then the interferogram will take the following form:-

$$I(\delta) = \int_0^\infty B(\sigma) \cos [2 \pi \sigma (\delta - \epsilon)] \, d\sigma \qquad \textbf{(4.14)}$$

II) Electronic filters which are designed to remove the high frequency noise from an interferogram have the effect of putting a wavenumber dependent phase lag, θ_σ, on each cosinusoidal component of the interferogram, and the resultant is given by:-

$$I(\delta) = \int_0^\infty B(\sigma) \cos[2 \pi \sigma \delta - \Theta_\sigma] \, d\sigma \qquad \textbf{(4.15)}$$

Recall,

$$\cos (A - B) = \cos A \cos B + \sin A \sin B \qquad \textbf{(4.16)}$$

the second term in the phase angle has the effect of adding sine components to the cosine wave interferogram. If $\theta_\sigma = \pi/2$ in Equation (4.15), the interferogram of a monochromatic source would look like that illustrated in Figure 4.6. In general $-\pi/2 < \theta_\sigma < \pi/2$, and the cosine transform will be intermediate between the sinc function.

The process of removing these sine component effects from a spectrum is called the phase correction. This is achieved by calculating the complex FT of the interferogram, that is two spectra are calculated, the real, and imaginary spectrum, from the cosine transform, and sine transform respectively. The phase corrected spectrum is then calculated by taking the modulus of this complex spectrum:-

$$|B(\sigma)| = \sqrt{[Re^2 + Im^2]} \qquad \textbf{(4.17)}$$

4.3.3 Instrumental Response

The transmission and reflection coefficients of the various optical components in the interferometer are not constant at all wavelengths. Additionally, it should be noted that the response of the detector varies significantly over the wavelength range of the instrument. The spectrum deduced by obtaining a FT of the detected interferogram, therefore, is not the true representation of the actual incident spectrum, and all recorded spectra must be corrected for such instrumental effects.

The instrument response for the interferometer is derived using the method similar to that described by Espy **et al.**, [1988][118], and Lowe **et al.**, [1991][119]. A low power (1 W) calibration lamp was supplied by Bomem Inc. for this purpose. The "true" instrumental response is obtained by normalizing the spectrum recorded by the interferometer, using the calibration lamp as a source, to the spectral radiance calibration curve.

A regular check must be kept of the instrument response to account for the any variations in the instrumental response with time, for example changing absorption coefficients in the beamsplitters caused by aging or changes in the detector response. Each spectrum resulting from the interferometer is then adjusted for instrumental response. This is achieved by normalizing the recorded spectrum of the instrument response. Once a spectrum is known accurately, calculations such as temperature determination can be performed with confidence.

4.3.4 Temperature and Band Intensity Calculation

From the literature [Mies, 1974][120], the mean time between collisions of neutral atmospheric molecules is approximated to:-

$$t = 4.4 \times 10^{-12} \exp(0.19\, h) \qquad \textbf{(4.18)}$$

where h = the altitude

OH (3,1) band molecules have a mean altitude of ~87 km, will consequently have collision times 6.6 x 10^{-5} seconds or collision frequencies of ~15 per μs. The excited hydroxyl radical, when produced, will have a rotational distribution which is significantly non-Boltzmann. It is generally accepted that the excited hydroxyl radical is produced by the exothermic reaction of atomic hydrogen and ozone (see Equation (4.19):-

$$O_3 + H \text{ --------} \rightarrow OH^* + O_2 \qquad \textbf{(4.19)}$$

where H = atomic hydrogen

O_3 = atomic ozone

OH^* = excited hydroxyl radical

O_2 = atomic oxygen

The 3.34 eV of energy which is released by this reaction is just enough to excite the OH radical to the maximum observed vibrational level, v = 9 (3.23 eV), but not enough to excite it to the unobserved level, v = 10.

However, given that the radiative lifetime for OH molecules is 4-6 ms, a molecule before it decays to a lower excited state, will have an average of 75 collisions which is sufficient to thermalize its rotational distribution. This has been confirmed by many measurements [Turnbull and Lowe, 1983; Espy **et al.**, 1988; Pendleton **et al.**, 1989][121,118,122]. When a vibrationally excited hydroxyl molecule decays to a lower vibrational state, its

rotational distribution will conform to a Boltzmann distribution. This thermalization of the rotational states is fundamental to the calculation of temperatures from hydroxyl emission spectra.

The rate at which vibrationally excited states decay to lower states, for rotational states obeying Boltzmann statistics, depends on the population in the excited state, the temperature of the excited molecules and on the Einstein spontaneous emission coefficients. The actual equation for the photon emission intensity is given as:-

$$I_{(J',v'\rightarrow J'',v'')} = N_{v'}\, A_{(J',v'\rightarrow J'',v'')}\, 2(2J' + 1)/Q_{v'}(T_{ROT}) \exp X/Y \qquad \textbf{(4.20)}$$

where $X = -E_{v'}(J')$

$Y = K\, T_{ROT}$

J' = the upper state total angular momentum

I = the intensity of the vibration-rotation transition from v',J' to the state v'',J''

k = the Boltzmann's constant

The total concentration of molecules in the upper vibrational state is given by $N_{v'}$, while $Q_{v'}$ is the partition function or state sum, and is defined as:-

$$Q_{v'}(T) = \Sigma\, (2J' + 1) \exp -E_{v'}/kT \qquad \textbf{(4.21)}$$

where $2J' + 1$ = the degeneracy of the rotational state J' which, for a particular temperature, is constant

$E_{v'}$ (J') = the upper state energy

The energy of a state or a term value is denoted by the expression $F_{v'}$ (J'), where:-

$$E_{v'} (J') = F_{v'} (J') * 100\ hc \qquad \textbf{(4.22)}$$

where h = the Planck's constant

c = the speed of light

From Equation (4.20), A, the Einstein spontaneous emission coefficient for the vibration-rotation transition. It is the reciprocal of the lifetime expectation value of the particular state.

Re-arranging Equation (4.20), and taking logs, the following relationship is obtained:-

$$\ln \{I_{(J',v'\rightarrow J'',v'')}/(2(2J' + 1)A_{(J',v'\rightarrow J'',v'')})\}$$
$$= \ln \{N_{v'}/Q_{v'}(T)\} - F_{v'} (J')100\ hc/kT \qquad \textbf{(4.23)}$$

Equation (4.23) represents the straight line graph, which on being plotted yields a gradient of 100hc/kT, from which the temperature can be derived. The intercept, ln ($N_{v'}/Q_{v'}$) can be related to the overall band intensity.

4.4 Discussion and Interpretation of Data

4.4.1 Alaska, 1986

The interferometer-spectrometer observations were made of the zenith sky from the Optical Observatory of the Poker Flat Research Range (65.13°N, 147.48°W) during August 1986. Measurements using an Infra-Red Field Widened Interferometer (IRFWI) were obtained from the 2nd to the 20th of August, and from the University of Western Ontario Michelson Interferometer (UWOMI-1) from the 7th to the 20th of August UT. Data were collected for solar depression angles greater than 6°. Since the hydroxyl rotational temperature can be derived from the spectra obtained even in the presence of relatively dense meteorological cloud, observations were continued provided that the signal level was adequate. Because of the diffusing effect of the clouds, rotational temperatures obtained in their presence were averaged over the entire sky, the corresponding hydroxyl intensities were not valid.

Photographic observations of the northern twilight sky were made from the Gulkana Airport, Alaska (62.2°N, 145.5°W), approximately 320 km South of Poker Flat. The observing geometry is shown in Figure 4.7. Any NLC over the interferometer site were observed at an elevation of ~12°, and at an azimuth of 344°N in the photographic data, assuming that the NLC height was 83 km. The prevailing twilight conditions limited the observing period to typically 4-6 hours centred on the local midnight (8:00 UT).

NLC were sighted on eight occasions during the eighteen day campaign. Due to the presence of meteorological cloud several of these displays were only seen from the Poker Flat site. However, on three occasions (the 3rd, 7th, and 11th August) NLC were photographed from Gulkana. Although, the observations were restricted by the changeable weather conditions, sufficient photographic and visual records have shown that on three other occasions (4th, 9th and 14th August), the twilight arch was clear with no discernable NLC.

All of the observations discussed here (visual, photographic, and spectrometric), were made when the solar depression angle at the observing sites was greater than 6°. Although, the detection of NLC requires that the mesopause region is illuminated by sunlight, it is important to note that the transition from day to night conditions for the hydroxyl layer occurs when solar radiation capable of photodissociating ozone no longer reaches the upper mesosphere due to strong screening by stratospheric ozone located at ~40 km. This transition occurs for solar depression angles ranging from about 3° to 5°. Thus, for these measurements, there is the apparent paradox that the NLC are sunlit, but the hydroxyl emission layer is under night-time conditions.

4.4.2 Measurements During NLC Displays

4.4.2.1 7th August 1986

The sky was mostly cloudy over the Gulkana Airport, but by 11:15 UT, it had cleared considerably, and the NLC were detected at the nominal azimuth extending above a low elevation cloud up to 30°. The NLC display was characterized by a set of bright Billows with some fainter Bands and patchy structure. By 12:00 UT several faint North-South aligned Bands were also photographed at 60° elevation. At the Poker Flat site, the sky remained clear throughout the night, and data were obtained by both spectrometers from 08:30 to 12:00 UT. Several NLC Bands were sighted to the North around 09:15 UT parallel to the horizon and extending up to 40° elevation. By 10:30 UT a second set of Bands emanating from the northern horizon were also observed. These were later seen to extend over the interferometer site by the Gulkana observers.

Figure 4.8 illustrates the relative intensity of the OH (3,1) band obtained by UWOMI-1. The data shows a near sinusoidal wave-like variation which appears to increase significantly in amplitude towards dawn. The narrower field IRFWI instrument observed a similar shaped, but more structured, variation in the intensity. The rotational temperature determined from the UWOMI-1 data set is shown in Figure 4.9. The plot exhibits a steady downward trend from an initial temperature of 165 K to 155 K during the period 11:15 UT to dawn, when NLC were confirmed

overhead. Superimposed on this trend, there is a small, but definite wave-like oscillation in the temperature in phase with the OH intensity variation. A detailed analysis can be found by Lowe **et al.**, [1991][119].

4.4.3 Discussion

During the course of the campaign significant changes in the average rotational temperature (~15 K) and the relative intensity, as much as 100% were observed on both NLC, and cloud-free nights. The data were often characterized by the presence of short period (<1 hour) wave-like variations, particularly in the intensity records, which temporarily modulated the hydroxyl brightness by typically 10 to 25%. The corresponding rotational temperature perturbations were less apparent, around 3-8 K, and are comparable in amplitude with the data obtained at mid- and low-latitudes using the same instrumentation. In general, the short period variations appeared superimposed upon much longer time scale changes, possibly associated with tidal or planetary waves.

The distinction between the Gulkana, and the Poker Flat NLC observations are emphasized: the Gulkana measurements contain definite evidence of NLC over the interferometer site, whereas the Poker Flat sightings refer mostly to the NLC displays observed in the northern twilight sky, not necessarily including the zenith. From the data there is a clear tendency for the mean temperature on the nights when NLC were confirmed (over Poker

Flat) to be lower than those on nights when no NLC were detected. Temperature data for two consecutive nights, the 3rd, and 4th of August were collected from the IRFWI. On 3rd of August, the hydroxyl temperature was ~150 K, for most of the night including the period when the NLC were detected overhead by the Gulkana observers (10:30-11:10 UT). However, on the following cloud-free night, the rotational temperature was significantly higher, around 160 K, indicating an ~10 K temperature difference on these two occasions for majority of the time. Only during the latter part of the night did the temperatures approach similar values.

In contrast, to this finding the intensity of the hydroxyl emission during the period when the NLC were identified overhead on the 3rd of August, was similar to that measured on the following cloud-free night for the same period. This result differs from the observations of Shefov [1961; 1968][123,89] who reported an enhancement in the hydroxyl intensity by a factor of 1.5 to 2 during the NLC displays followed by a sharp decrease in intensity by 2 to 3 times lower than average on the next day when there were no NLC. The Figure 4.10, indicates a considerable night-to-night variability, as was observed by Shefov, but there is no obvious difference in the mean hydroxyl radiance; and its diurnal variation was found on NLC nights to be comparable with cloud-free nights [Taylor **et al**., 1989][91].

To investigate the tendency for lower average hydroxyl temperatures during NLC nights, the interferometer results show the distribution of the OH (3,1) band temperatures obtained by

each instrument (see Figure 4.11). There does appear to be distributions present in the data sets. In the IRFWI data (see Figure 4.11) only three nights, the 3rd, 7th, and 11th of August, show distributions with a substantial fraction of the temperatures below about 154 K. These three nights coincide with the positive identification of the NLC over the Poker Flat site, at least for part of each night. In comparison, on the 4th and 14th of August very few of the spectra indicate temperatures below 154 K. Both of these dates coincide with confirmed cloud-free sightings.

The distribution of temperatures for the UWOMI-1 data set are illustrated in Figure 4.12. Once again a significant number of the spectra obtained on the NLC nights fell below 154 K. There were no UWOMI-1 data for the 3rd of August, but on the 10th of August a large number of spectra with temperatures well below 154 K were recorded. As aforementioned, meteorological cloud prevented detailed observations to be made from Gulkana on this occasion, but reports from the Poker Flat site indicated that an intense NLC display was visible over much of the northern sky including the zenith around local midnight. On the 9th of August, the interferometer data indicated higher temperatures. Although the data were relatively sparse, due to cloud, the sky was clear at Gulkana for much of the night, and no NLC were detected. This is in good agreement with the IRFWI measurements of the 4th, and 14th of August.

In a detailed analysis of the UWOMI-1 data recorded during this

campaign, Lowe **et al.**, [1991][119] have investigated the variation of the rotational temperature with the vibrational level. They found that on average, the (4,2) band temperature was about 3 K less than the (3,1) band temperature. Other bands from higher vibrational levels indicated temperatures higher than (3,1) temperature, but the results of Pendleton **et al.**, [1989][124] suggest that the higher vibrational levels may not be fully thermalized as a result of insufficient number of collisions. The (3,1) band was chosen in this analysis, because it has a high signal-to-noise ratio in both data sets, and because the spectral response of each instrument is relatively flat in its vicinity.

Shefov [1968][89] did not observe a correlation between the hydroxyl rotational temperature, and the occurrence of NLC. However, this was not unexpected when it is considered that the average temperature difference appears to be quite small (<10 K), and that natural variations, induced for example by the passage of short period gravity waves, can act to mask its detection. The correlation reported here was found using spectra capable of yielding temperatures with a high precision, better than 2 K, and a high temporal resolution (<1 min); by making a critical assessment of the common volume data to determine the presence/absence of NLC rather than a general appraisal of the sky conditions. Moreover, based on the spread in temperatures observed on both NLC and cloud-free nights [Taylor **et al.**, 1989][91], it was concluded in an initial investigation that there was no obvious relationship between the hydroxyl rotational temperature, and the occurrence of NLC.

Coordinated observations of NLC are simple in concept, but are naturally restricted by the twilight conditions [Fogle and Haurwitz, 1966][4], and by the occurrence of meteorological cloud, both of which can prevent or confuse the detection of NLC. As NLC displays may last for several hours, then the author cannot preclude the data taken prior to and/or after the periods when the clouds were confirmed overhead at Poker Flat site did not contain NLC. Although the (3,1) band data indicate a very high percentage of spectra below 154 K during the confirmed NLC periods, the actual proportion of "cold spectra" as a percentage of an entire nights data may be somewhat lower, depending upon the geographical extent, and duration of the NLC region.

Data obtained on two nights nearing the end of the campaign, the 17th, and 20th of August, seem to conflict with the above results. On each of these occasions NLC were seen only from Poker Flat, and no information was available from the Gulkana site. On the 17th of August, several East-West closely aligned bands were observed in the twilight sky extending to high elevations. The spectral measurements were limited by intermittent cloud, and indicated temperatures around 160 to 170 K throughout the night. On the night of the 20th of August, the visual records were supported by all-sky photographs taken from Poker Flat illustrating several near North-South aligned bands stretching from the northern horizon to close to the zenith. Compared with the interferometer data recorded earlier in the campaign, the rotational temperatures determined on this occasion (see Figure 4.13) were unusually high starting around 175 K, when the all-sky

photographs were taken, and at no time during the night did they drop as low as 154 K. As there were no data available from Gulkana to corroborate these sightings some questions inevitably remain concerning the nature of these two displays. However, it is interesting to note that Shefov [1968][89], also reported that the hydroxyl temperature on some NLC nights was significantly higher than the average.

To investigate this apparent dilemma, the conditions under which these measurements were made need to considered in more detail. Model calculations on NLC formation by Gadsden [1981][125] and Turco **et al**., [1982][61], and PMC by Jensen **et al**., [1988][126] indicate that ice nucleation will occur most effectively at low temperatures below 130-140 K. Rocket measurements at high latitudes (69°N) have shown that the summertime mesopause is remarkably stable exhibiting a mean value of 130 K at an altitude of ~90 km [von Zahn and Meyer, 1989][127]. Considerable small-scale structure has also been observed in the thermal profiles with temperatures occasionally reaching as low as 110-120 K [Philbrick **et al**., 1984; Fritts **et al**., 1988][128,129]. At these low temperatures particle growth occurs in a saturated region immediately below the nucleation level. If, the conditions remain favourable for growth, a visible NLC layer may form several kilometres below the mesopause, at an altitude of typically 83 km [Gadsden, 1981][125].

Using a water vapour mixing ratio of a few parts per million, thought to be typical of the mesopause [Arnold and Krankowsky, 1977][130], each model shows that particle growth will take several

hours before they reach a sufficient size (> 50 nm) to scatter light above the twilight foreground. During this period, the particles are swept away from the nucleating region by the prevailing horizontal mesospheric winds which can be as large as 50-100 m/s. Thus, although the ice grains only fall a few kilometres in altitude they may have travelled several hundred kilometres downstream by the time they are observed as NLC.

Baker and Stair, [1988][131] found that, on average, the hydroxyl emission peaks at an altitude of ~87 km and has a typical night-time halfwidth of ~8 km. High latitude observations of the high emission are rare, and this result was based mainly on rocket soundings made at lower latitudes. To a first approximation, the rotational temperature derived from the interferometer data is the brightness-weighted temperature of the entire emitting region. Thus, under nominal conditions the hydroxyl observations should provide a measure of the average temperature of a region of atmosphere whose lower boundary typically coincides with the base of the NLC layer, and which extends upwards somewhat more than one scale-height.

The majority of the data presented here indicates that the mean temperature of the region immediately above the cloud base was slightly cooler (<150 K) on nights when NLC were observed than on other cloud-free nights. Due to the averaging effect of the emissive layer, the actual temperature differences may have been somewhat larger. This suggests that for much of the time the prevailing conditions in the vicinity of the overhead mesopause

at Poker Flat were not dissimilar from those at the latitude of cloud nucleation. On the 17th and 20th of August, the average hydroxyl temperatures were substantially higher than on previous NLC displays, and nucleation may have occurred at a distance from the observation site. In this case, the overhead mesopause may have exhibited little or no semblance to that of the NLC formation and growth region. Alternatively, the observation of higher hydroxyl temperatures during the latter part of the campaign may reflect a change in the mean altitude of the hydroxyl layer with respect to the mesopause region possibly associated with the onset of a seasonal transformation.

In summary, from this limited data set a very tentative conclusion can be drawn. NLC are more frequently observed on nights when the OH (3,1) band rotational temperature fell below 154 +/- 2 K for a substantial fraction ie >15% of the time. This is consistent with the current ice particle growth models. However, as particle nucleation usually takes place at some distance from the region at which the clouds are observed the detection of "cold" hydroxyl spectra cannot be used as a likely indicator for the existence of NLC. There is a clear need for further observations of this type to investigate the details of this temperature relationship which has until now proved difficult to prove.

4.4.4 Sweden, 1991

All hydroxyl observations were made in the zenith and were

accompanied by observations documenting the presence or absence of NLC over the interferometer.

The MI-spectrometer used during this period of studies was operated in the near infra-red region of the spectrum, 1.0 to 1.6 μm, at about 0.0025 μm resolution. The Bomem has a 1.5° circular field-of-view, and a 1 cm diameter entrance aperture. However, the field widened device used at Poker Flat, Alaska during the 1986 campaign had a much higher throughput, and will acquire a spectra with a signal-to-noise ratio on the order 500:1 in 20 seconds. The individual rotational temperature fits in this case have a precision of less than 1%. The Bomem scans each interferogram rapidly, but integrates the individual interferograms together for about 5 minutes in order to achieve a signal-to-noise ratio of about 100:1. This translates to about a 3 to 5% error in the temperature fit. In all cases, the fit to the radiance is slightly better. Currently, all of the data have been converted from interferograms to spectra, a necessity in a Fourier Transform system. All of the data for 1986, 1991, and 1992, have been reduced to yield the OH M(3,1) band radiance and rotational temperatures.

At first, the primary objectives of the campaigns were to look at the gravity-waves in the hydroxyl, and to compare those with the structures in the NLC. During 1991, and 1992, a closer look was made of the climatology of the hydroxyl during the NLC season. The climatological data suggests the presence of large-scale planetary disturbances and presently, the correlation

between these and the occurrence of NLC are being examined.

4.4.5 Measurements During NLC Displays

4.4.5.1 1st August 1991 (salvo B)

From Table 3.4, the period of observation that is of most interest with respect to the multinational rocket and radar program, ie NLC-91 is the 1st to 13th of August 1991, during which the US payloads EFIELD-A/B were flown housing the APL photometers. On the night of salvo B, 1st of August, the sky at Lycksele, Sweden was clear at 20:30:00 UT, by 22:14:00 UT some patchy clouds had moved in slowly dissipating to a mostly clear sky with a few patches of cloud, by 23:30:00 UT; however, no NLC were observed. This observation suggests that the tentative conclusions drawn with respect to the photometric results collected from the APL on the aforementioned night were incorrect (see Chapter 3). However, on the night of salvo A, the 9th of August 1991, the results obtained from the Bomem interferometer seem to correlate with the results gained from the APL photometers flown, indeed a NLC display was observed. See Section 4.4.5.2.

4.4.5.2 9th August 1991 (salvo A)

As on the preceding night, there were high cirrus clouds, however at midnight, there was a NLC display. It was clear in the zenith, see Figure 4.14, for photographic sequence of the movement of the

NLC during the night. The stars were visible through the NLC display, for example Capella. This display was just below Capella with a good eastward extent. It had no coherent structure. Some Billows (East-West), and at least two North-South Bands were visible. One bright Band moved over Esrange. There were patchy clouds to the West of the bright Band, and to the North, as far as Andoya. Rockets were launched, and an aerosol layer was detected using the APL photometers (see Chapter 3). From the Bomem interferometer, a good series of spectra were obtained.

4.4.6 Discussion

During the course of the campaign, in particular looking at the band radiances, and rotational temperatures, no significant changes in either the temperatures or the radiances were observed. Figure 4.15 illustrates the nightly averages of the integrated band radiance. There is some indication that the integrated band radiance increases after a NLC display, but not during the display. However, it should also be noted that the integrated band radiance has increased without NLC being present. Now considering the corresponding nightly averages of the rotational temperature, again, there is some indication that the rotational temperature will decrease after a NLC display, but this is not always the case (Espy, personal communication). The rotational temperature can also be seen to decrease before the NLC display. Figures 4.16 and 4.17, illustrate thirty minute averages of both the integrated band radiance and rotational temperature. Both of these figures give a better indication of

the aforementioned trends. Figure 4.15 illustrates a strong trend, the integrated band radiance during a NLC was not as intense as that following a NLC display. Considering the corresponding rotational temperatures, it can be clearly seen that the temperatures on the nights of NLC displays were lower than those nights on which no NLC were observed. Finally, looking at the one hour averages of integrated band radiance and rotational temperatures, around the local midnight (see Figures 4.18 and 4.19), this time series reveals several long-period fluctuations in the hydroxyl radiance and temperature which appear to be correlated with the appearance of NLC. These fluctuations are interpreted as the signature of large, planetary-scale disturbances that are responsible for changing the local mesospheric air mass or wind field over the interferometer. Although estimated to be a mixture of a nominal 16-day wave and possibly a ~5-day wave [Espy and Witt, 1991][132], a more accurate assessment of the period of the disturbance was not possible, since the observations were carried out at different geographic locations. During the NLC-91 campaign, the interferometer was moved gradually northward, as the solar depression angle permitted, in order to coordinate with the EISCAT observations.

From the above investigation of the climatology of the high latitude, mesospheric hydroxyl during the summer NLC season, a new direction of investigation has become clear. The author believes that a closer look should be made as to the situation of low temperature and low radiance. Does this combination still

enable NLC to be observable? This postulate is examined further in Chapter 7.

As found in the 1986 data, the presence of short-period wave-like perturbations were also observed in 1991. However, it has been postulated that this region is shielded from gravity and planetary waves by the filtering effect of the generally easterly (westward) underlying stratospheric flows. Being cut off from this momentum source ,and exposed to the strong upwelling resulting from the general circulation pattern, the mesopause region can cool below 140 K. At these very low temperatures, small amounts of available mesospheric water vapour can condense onto suitable nucleation centres into high altitude clouds, or NLC.

The presence of NLC and PMC (Polar Mesospheric Clouds) within the summer mesosphere has been well documented by ground-based, and satellite observations, respectively. However, there are as yet few experimental measurements of the specific mesospheric conditions necessary for their formation. The critical parameter of mesospheric temperature has been measured in a limited number of cases using in-situ rockets (see Chapter 3) probes. However, due to the solar scattering conditions necessary for observing NLC it is difficult to employ traditional, non infra-red ($\lambda<1\mu$), spectroscopic measurements of hydroxyl nightglow emissions in order to provide widespread remote sensing of the temperature near the mesopause. The few measurements that have been made during the summer season have shown that the temperature of the mesosphere changes little on days when the clouds are present as

compared to days when they are absent. Indeed, many of the infra-red hydroxyl measurements have shown the temperature near the mesopause to be above the "frost" point during periods when NLC were present [Espy and Witt, 1991; Taylor **et al.**, 1989; Shefov, 1967; Harrison, 1973][132,91,133,90].

In an effort to reconcile the hydroxyl measurements, there has been much speculation that the cloud formation process requires the presence of waves in the mesosphere. These wave structure, produce temperature profiles with wave-like oscillations, and allow the clouds to form in the cold minima while keeping the average temperature high. Although, there is experimental evidence for gravity waves structuring the mesospheric temperature [Philbrick **et al.**, 1984][128], there are as yet no observations that unambiguously relate the appearance of the clouds to the level of gravity wave activity, or to the phase and amplitude of the tides and planetary waves present in the polar summer mesosphere.

A future proposal known as the Arctic Noctilucent Cloud Campaign-93, will remotely sense the residual gravity and planetary waves present in the Arctic summer mesosphere using infra-red optical techniques. This will be a unique opportunity of using both aircraft-borne and ground-based MI to measure the hydroxyl, and to determine both the spatial and temporal characteristics of the wave fields. This information will then be correlated with the stereo-photographic measurements of the cloud height, and structure in order to infer the relationship between the

formation of mesospheric clouds and the presence of small- and large-scale waves.

Using the USU MI, observing hydroxyl in the near infra-red (NIR) spectral region, has the advantage that the reduced light scattering in this spectral range, provides the possibility of measuring the hydroxyl temperature and radiance from the ground with the Sun only 3.5° below the horizon. Therefore, on the aircraft, due to the decrease of the scattering with altitude, the existing instrumentation should be capable of meeting these future developments of measuring hydroxyl temperature and radiance as far North as 65°-70°. Obviously to increase the range of these instruments, polarizing filters would be added to improve the rejection of solar scattered light.

The critical nature of the mesospheric temperatures during the high-latitude summer in determining the rates of cloud growth has engendered a number of measurements and techniques. A review and summary of some sixty in-situ measurements by von Zahn [1990][137] indicated that at 59°N (fifteen measurements), the average summer mesopause was located at 88 +/- 2 km, and had a temperature of 148 +/- 13 K. The altitude profile of the intensity-weighted mean temperature given by von Zahn and Meyer [1989][127] for 68°N showed that the temperature gradient below the mesopause is ~4 Kkm^{-1}. Thus, if one used this gradient, and the lower limits for the mesopause temperature and altitude, then it would be expected that a particle growth region would be at an altitude region of ~82 km, where the temperature would be ~159 K. Although, this

extreme case is in general agreement with the observed altitude range of NLC, the average temperature values would be too high for effective nucleation, and cloud growth according to the present models.

In-situ techniques for measuring temperature, reviewed by von Zahn [1990][134], are generally assumed to be accurate, as they are not integrated over large time periods nor altitude ranges. However, these in-situ measurements have found mesopause temperatures to be low. During the NLC-91 campaign carried out at Esrange, Kiruna, Sweden, Schmidlin [1992][113] measured a mesopause at 90 km with a temperature less than 100 K. But, neither the ground-based nor the aircraft observers South of the launch site detected any NLC during this extremely low temperature event. Again, this emphases the fact that a NLC display does not necessary occur when the temperature is sufficiently low. A temperature of 100 K would correspond to a saturation water mixing ratio of less than 1 part in 10^{12}.

Along with the in-situ measurements, the hydroxyl rotational temperature has also been used as a measure of mesospheric conditions in the presence of NLC [Espy, and Witt, 1991; Taylor **et al.**, 1989; Shefov, 1967; Harrison, 1973][132,91,133,90]. At night-time, the hydroxyl radical is produced naturally in the atmosphere in a relatively thin, ~8 km thick layer, at an altitude of ~87 km [Baker, and Stair, 1988][131]. The hydroxyl formation chemistry produces a strong, chemiluminescence of the NIR, vibrational-rotational bands, the Meinel bands, which may

be observed at night from the ground. And since hydroxyl can make several collisions before radiating, the rotational distribution of energy within these bands may be used to remotely deduce the natural temperature at the altitude of the hydroxyl layer. Thus, since the nominal hydroxyl emitting altitude is near the average summertime mesopause altitude as given by von Zahn [1990][134], the summer hydroxyl temperatures should reflect the mesopause temperature.

The typical values found for high latitude hydroxyl temperatures reflect the same behaviour as the in-situ measurements. For example, Shefov [1967][133] from 56°N typically illustrate temperatures in the vicinity of 160 K. Similarly, the USU NIR data taken from Stockholm during the NLC-91 campaign, yield an average temperature of 164 +/- 6 K. The data taken for 19th of June through the 22nd of August 1992, show a similar value of 163 +/- 8 K (Espy, personal communication). Once again, if the hydroxyl temperature is placed at either the lowest or highest limits of its observed altitudes, then the lower limits of these average temperatures would be consistent with the observed NLC altitudes. However, the average temperature and altitude values would preclude NLC formation.

It has been suggested that the difference between the in-situ, and the hydroxyl temperature measurements may be the result of the smoothing produced in the hydroxyl measurements which integrates over the non-negligible thickness of the hydroxyl layer. It may be shown that these differences are minor if the

temperature profile resembles the smooth, systematic profile of von Zahn, and Meyer [1989][127]. However, during the CAMP (1982) campaign, Philbrick **et al.**, [1984][128] measured a highly structured temperature profile using an active falling sphere technique (see Chapter 3). These oscillations were interpreted as the perturbation caused by an internal gravity wave moving through the mesopause region. As shown in Figure 4.20, these fluctuations drive the temperature well below the "frost" point over restricted altitude ranges, as well as up to 150 K in some places. Since it is assumed that gravity waves caused the Band structures seen in NLC, these structured temperature profiles have led to the speculation that gravity waves might be necessary for the formation of mesospheric clouds by providing growth regions in the cold troughs of the waves. This provides a mechanism whereby particle nucleation and growth could occur over even though the measured hydroxyl temperature, which is generally an average over time and altitude, is above the "frost" point.

Another feature of the gravity wave mechanism is that the upward component of acceleration associated with these waves could counteract the particle sedimentation rates and allow the particles to grow larger than the models [Turco **et al.**, 1982; Jensen and Thomas, 1988][61,103] presently allow [Rosenfeld, 1986][135]. However, the average particle thus levitated would be exposed to the downward phase propagation of the temperature wave and thus, would alternately be exposed to both hot and cold temperatures. Since the time scales for cloud growth are on the order of hours to days, while the sublimation rates rapidly approach seconds for

temperatures near 150 K, the wave may actually destroy the cloud. Indeed, when Turco **et al.**, [1982][61] examined the effects of a gravity wave propagating through a NLC, they found that the cloud brightness decreased substantially due to the hysteresis associated with the disparate growth, and sublimation rates. They suggested that the gravity waves could promote cloud growth by providing nucleation sites in the cold troughs. However, since the gravity wave periods were short compared to the growth times, the wave activity needed to cease once particles had nucleated or the gravity waves would erode the ice particles.

Since wave motions in the mesosphere could explain the variability of NLC, Jensen **et al.**, [1989][136] sought to include in their model the effects of waves whose periods were more commensurate with the particle growth times. Although it was believed that planetary waves would be severely attenuated by the underlying winds in summer [Andrews **et al.**, 1987][137], Vial [1989][138] predicted that the semi-diurnal tide would dominate at high latitudes. Indeed, Myabo and Harang [1988][139] used hydroxyl temperature data to show that springtime tides could create temperature variations near the mesopause between 5 and 10 K. Hence, Jensen **et al.**, [1989][136] included a tide-like 24 hour temperature wave of +/-7 K in their cloud growth model. They found that the cloud brightness could change by up to a factor of 7, and that these changes lagged behind the input wave by several hours. Thus, this long period waves seem to strike a balance and allow both enhanced nucleation and significant growth during the cold cycle before beginning to erode the particles

during the warm phase.

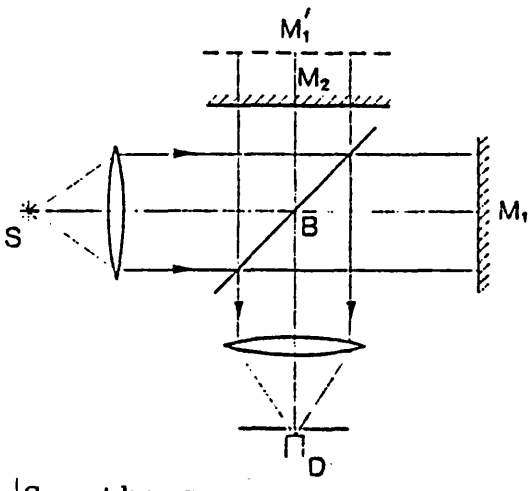
In addition to the aforementioned tidal variations atmospheric planetary waves can also create temperature fluctuations on the order of 6 to 10 K in the mesosphere [Vincent, 1990][140]. In the past, it had been thought that all planetary waves were filtered out before reaching the summer mesosphere by the underlying stratospheric easterlies [Andrews **et al.**, 1987; Houghton, 1986; Thomas, 1991][137,141,142]. However, Vincent [1990][140] maintains that it is only the forced planetary waves that are excluded from the mesosphere. Planetary-scale transient motions associated with the normal or migrating modes, as well as waves generated locally above the filter by baroclinic instabilities of the stratospheric jets, are often observed in the summertime mesosphere [Vincent, 1990; Carter and Balsley, 1982][140,143]. Spectral analysis of long-term radar data suggests that this transient wave energy tends to maximize in the summer near the mesopause with the most frequently reported periods falling into three intervals. These are referred to as the "16-day", "5-day", and "2-day" waves, respectively. A similar summary table of these transient Rossby-gravity normal modes that are observed in the summer mesosphere is reproduced after Vincent [1990][140] (see Table 4.1).

Table 4.1

Period /days	Nominal Period ("x" day wave)	Ampli-tude	Vertical Wavelength /km	Zonal Wave-number
10-20	16	5-10	> 100	1?
4-7	5	5-30	25 to > 100	1
1.9-2.2	2	10-50	50 to > 100	3

Presently, the effect of these planetary waves on the nucleation and growth of mesospheric clouds has not been explored in models. However, the long sequence of summer high latitude hydroxyl measurements of Espy and Witt [1991][132] revealed both 5- and 16-day waves which appeared to modulate the NLC. This result might have been anticipated from the tidal modulation of the NLC brightness demonstrated in the model of Jensen **et al.**, [1989][136] since these transient waves have amplitudes similar to the tidal modes, and like the tides, they have temporal scales on the same order as those required for particle growth. In addition to changing the mesopause temperature like the tides, these long-period planetary disturbances, which dominate the wind energy in the horizontal plane [Carter and Balsley, 1982][143], could create significant horizontal transport. Consequently, their strong meridional components [Vincent, 1990][140] could transport and mix water vapour rich in air from the South with nucleated particles from the Northern PMC regions to enhance the NLC growth. This could explain the factor of 4 variations in the mid-latitude mesospheric water vapour observed by Bevilacqua **et al.**, [1987][144] over a forty hour period. Moreover, as suggested by Thomas [1991][142], the formation of giant mesoscale cloud complexes

[Fogle, 1965][38] may require the presence of waves of planetary waves. An up-to-date summary and discussion of the recent, and future proposed work characterizing these long-period fluctuations is given in Chapter 7.



S = the source
M_1 = the fixed mirror
M_2 = the movable mirror
D = the detector
M_1' = the image of M_1 and M_2 as seen from D

Figure 4.1

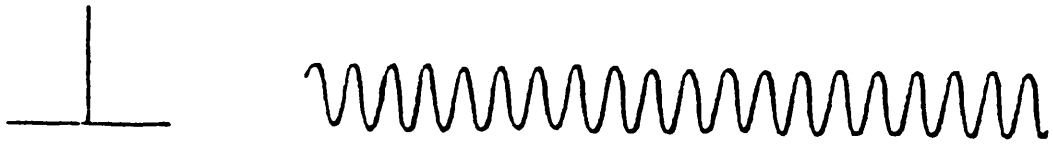
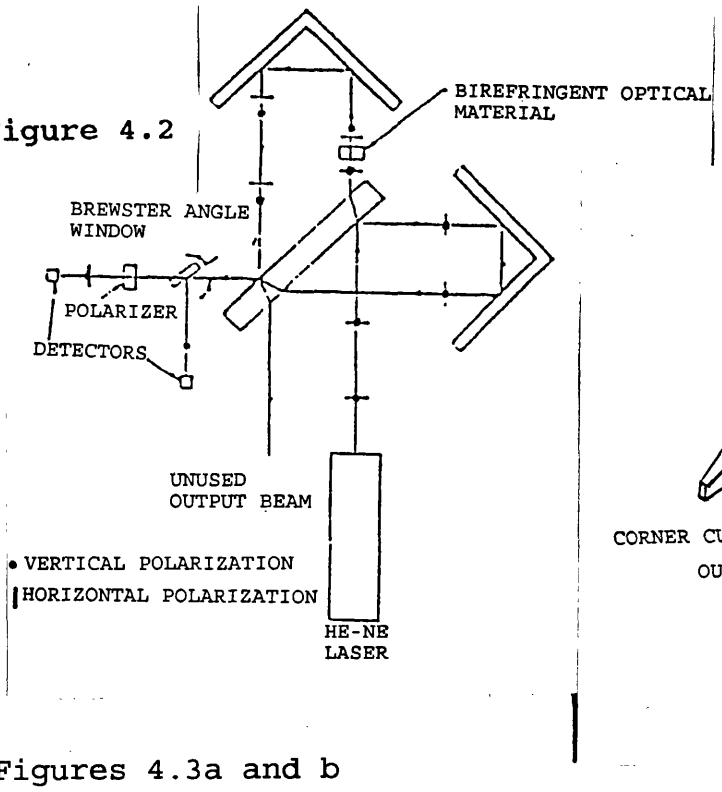
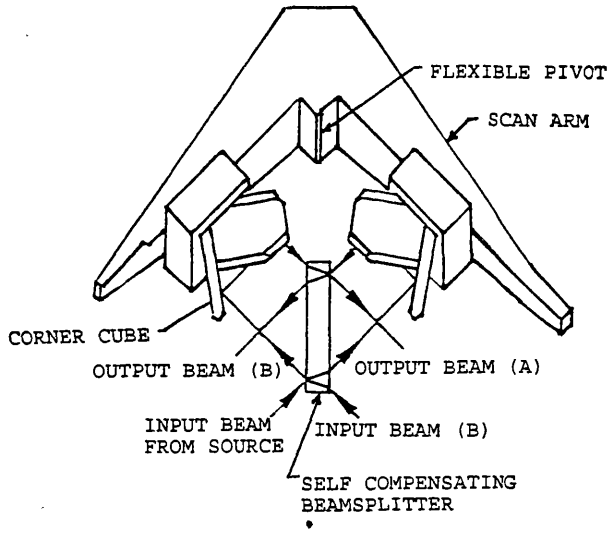


Figure 4.2





Figures 4.3a and b

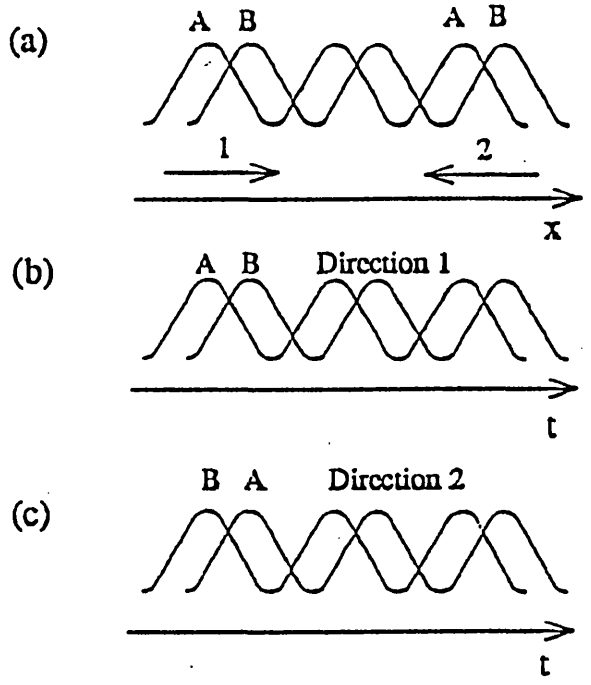


Figure 4.4

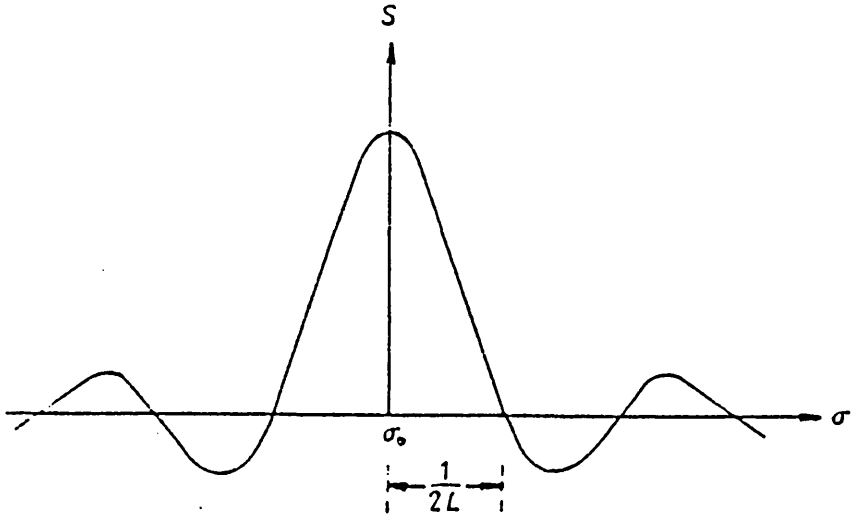


Figure 4.5

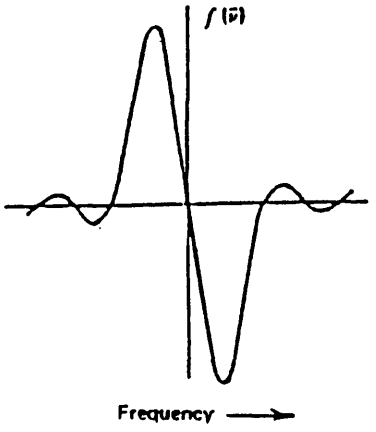


Figure 4.6

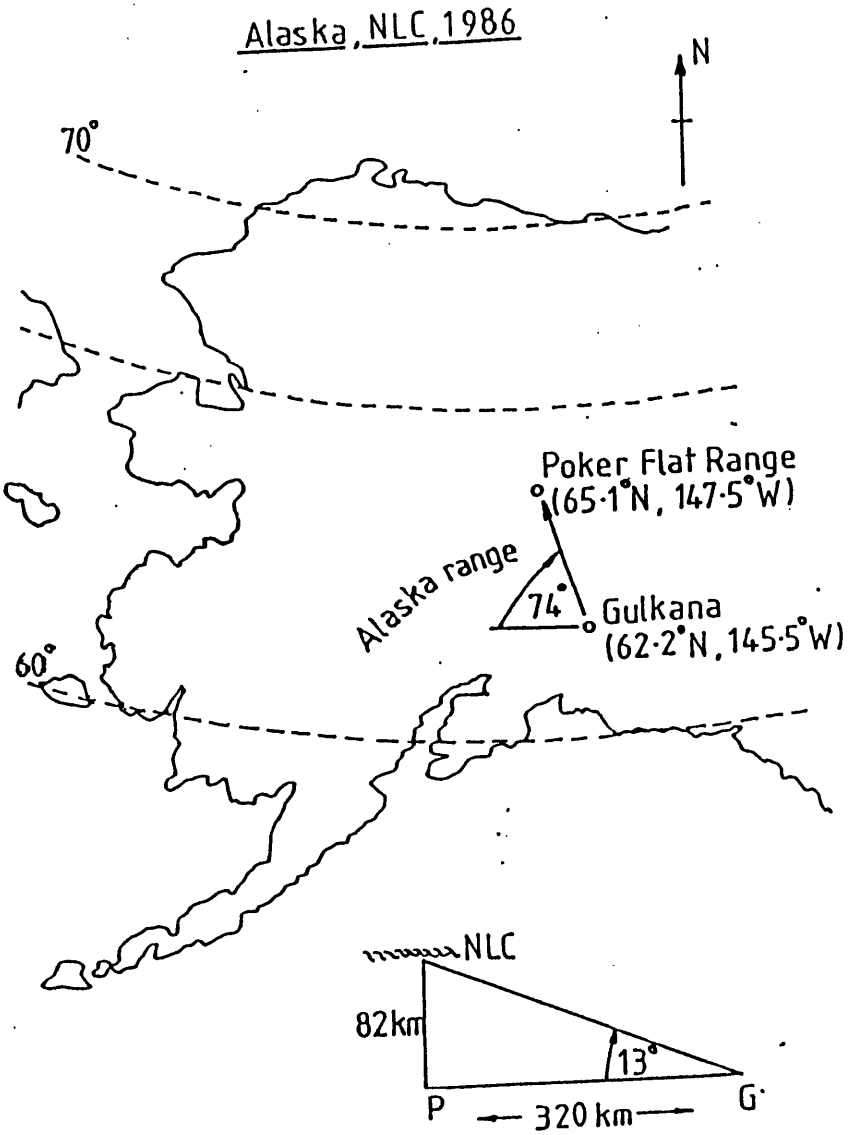

Alaska, NLC, 1986
N
70°
60°
Poker Flat Range
(65·1°N, 147·5°W)
Alaska range
74°
Gulkana
(62·2°N, 145·5°W)
NLC
82km
13°
P
320 km
G

Figure 4.7

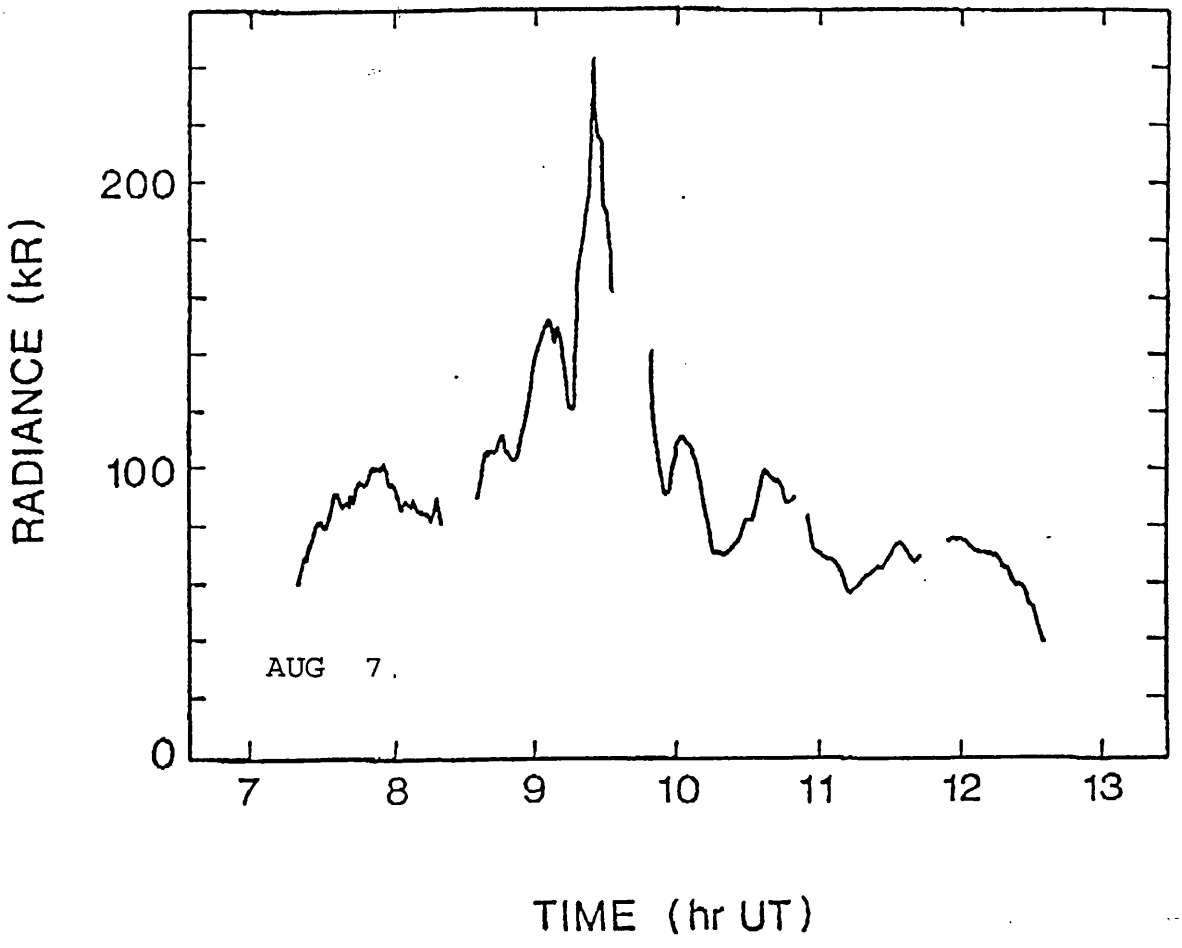


Figure 4.8

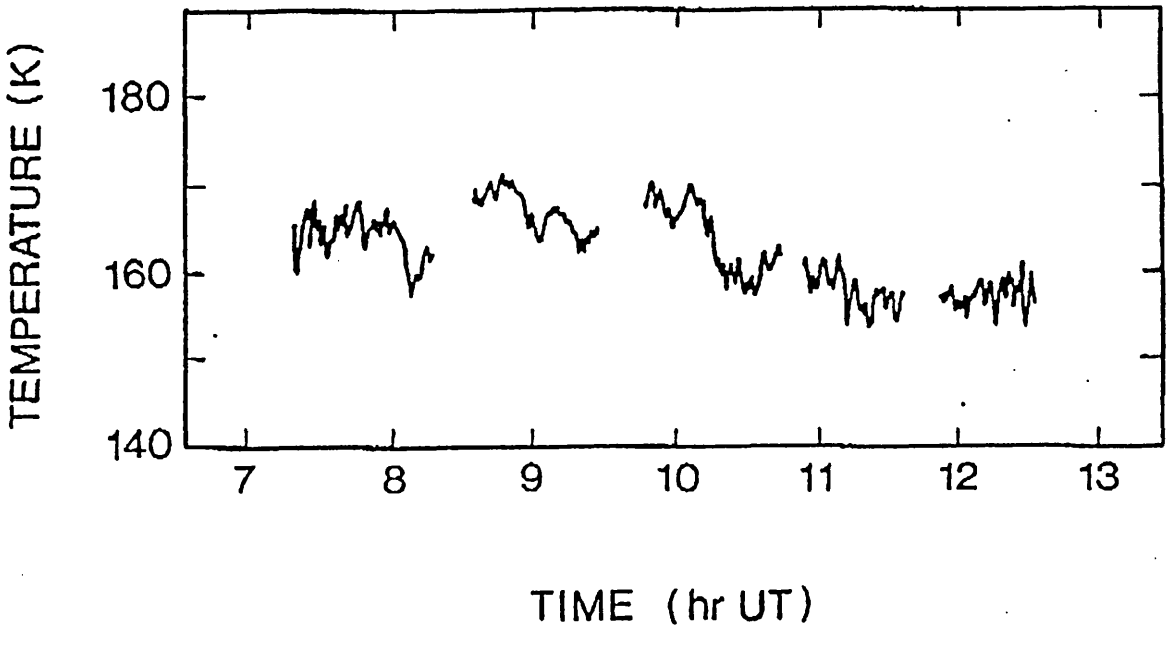


Figure 4.9

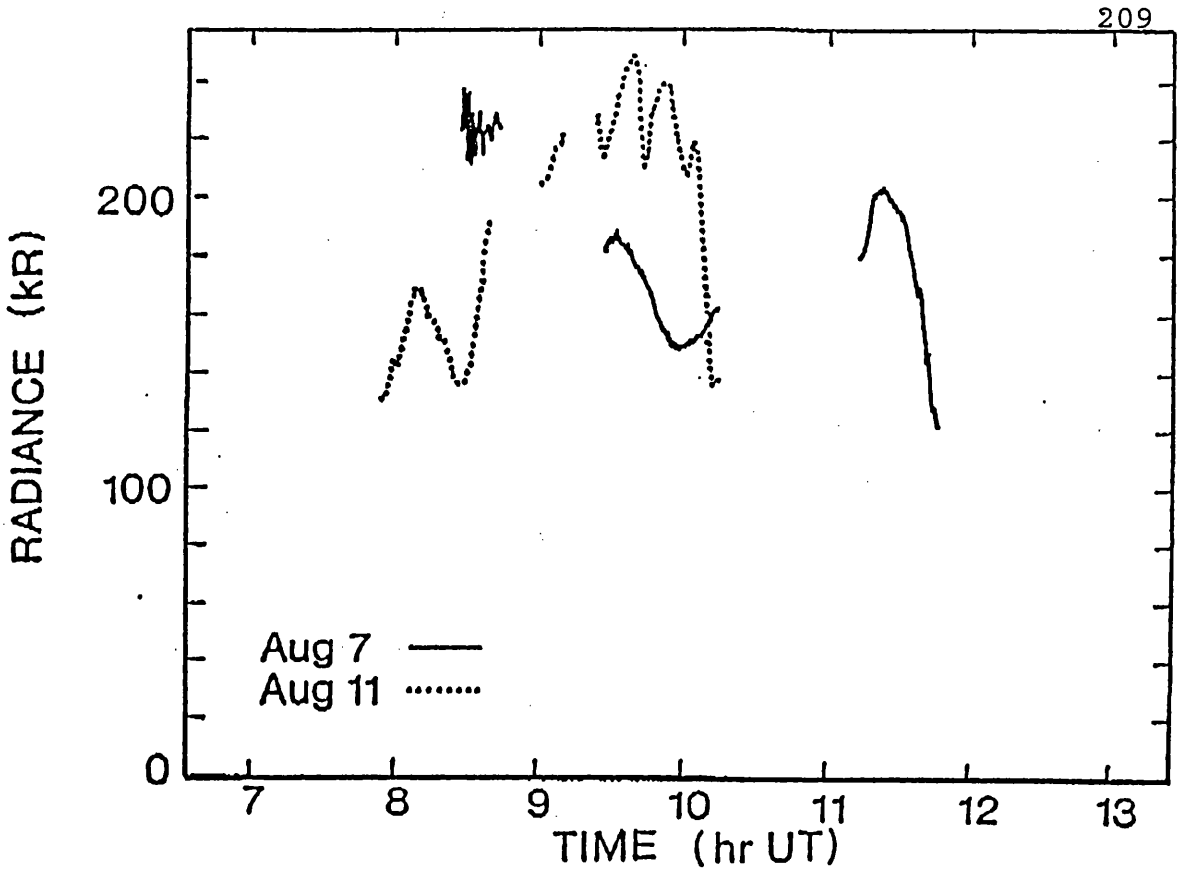


Figure 4.10

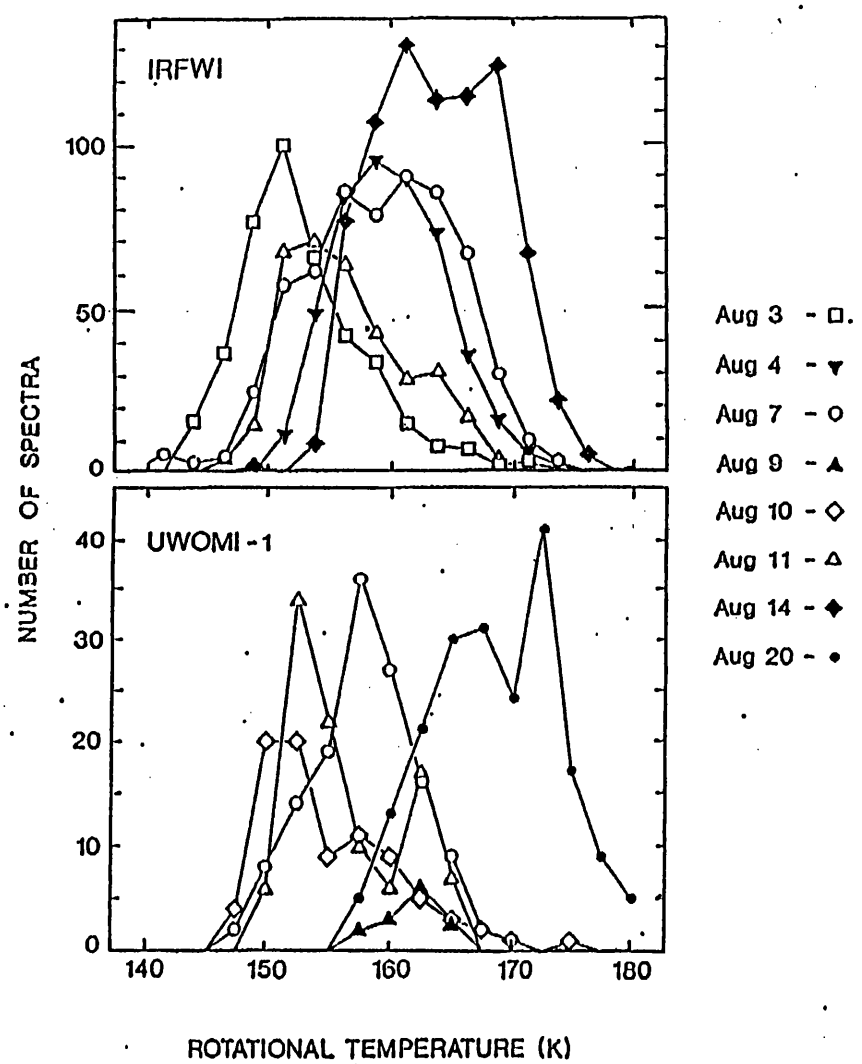


Figure 4.11

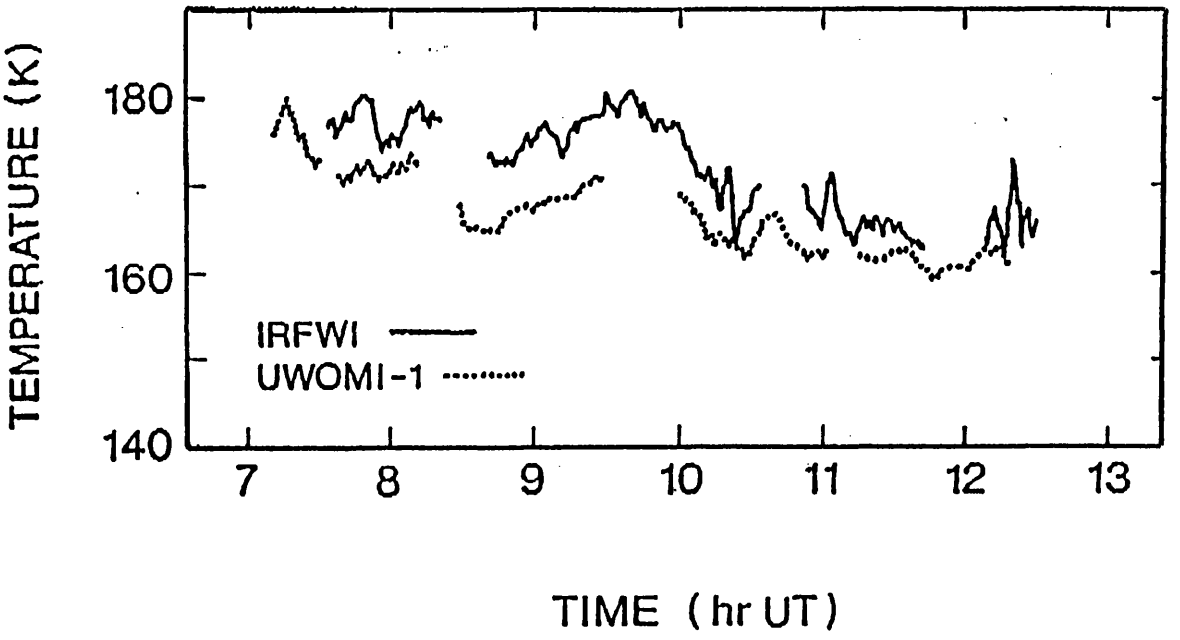


Figure 4.12

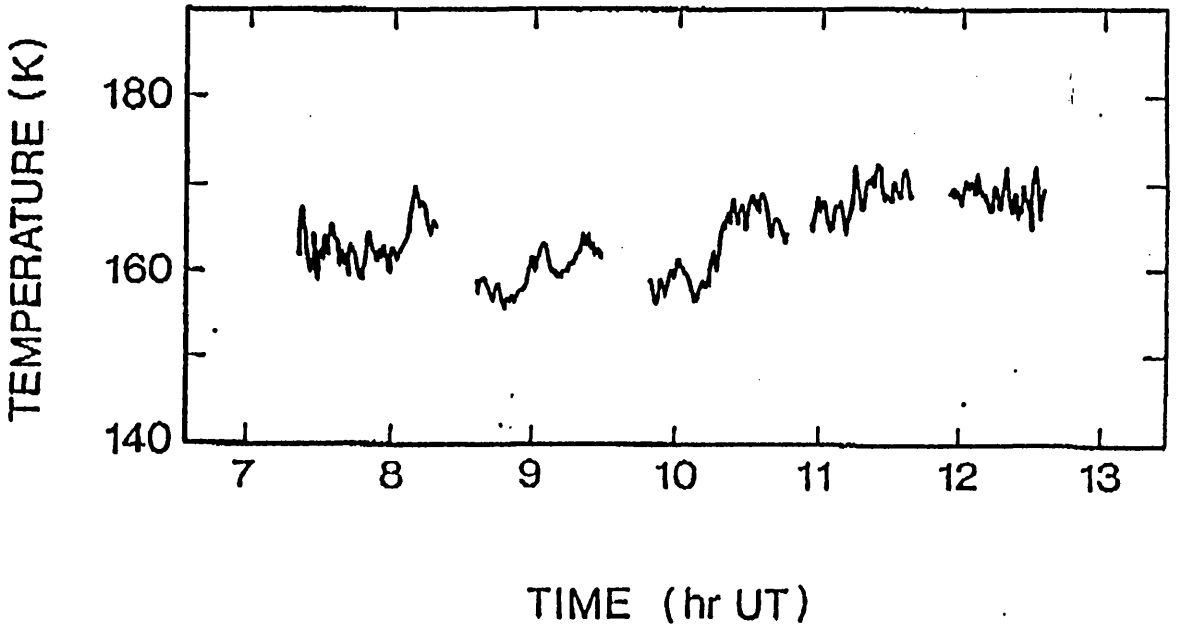


Figure 4.13

Figure 4.14

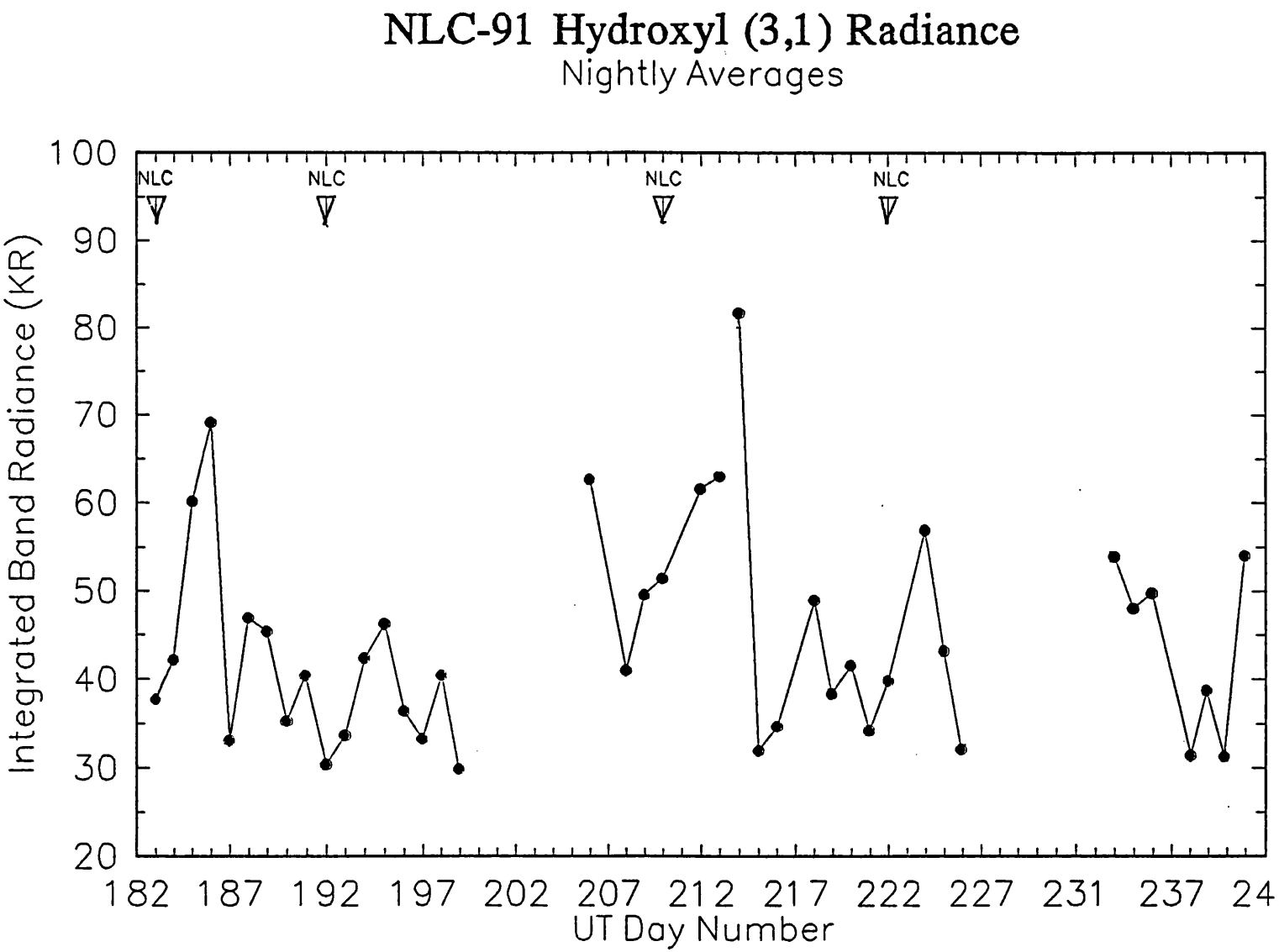

NLC-91 Hydroxyl (3,1) Radiance
Nightly Averages
Integrated Band Radiance (KR)
100
90
80
70
60
50
40
30
20
NLC
NLC
NLC
NLC
182
187
192
197
202
207
212
217
222
227
231
237
242
UT Day Number

Figure 4.15

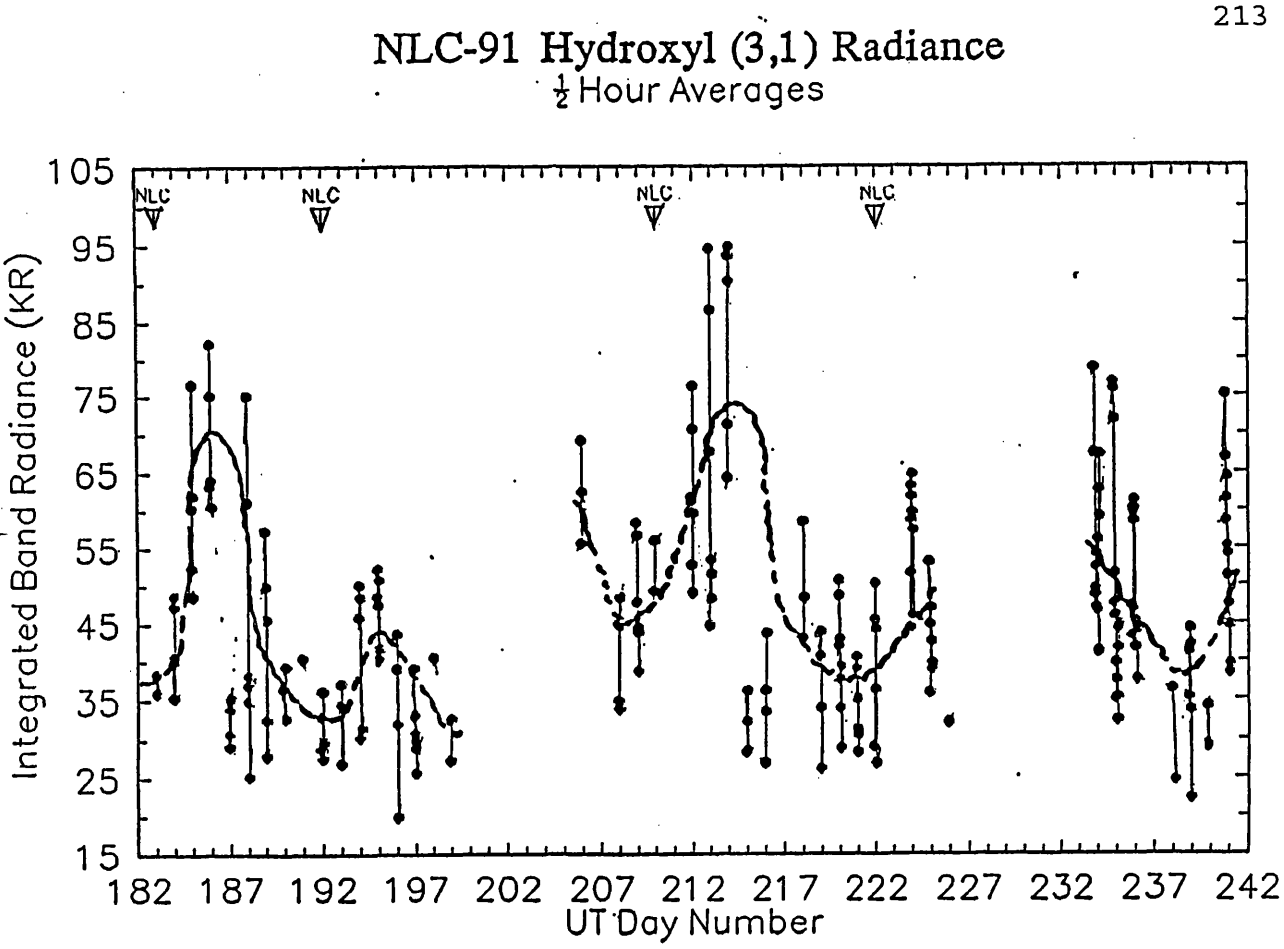


Figure 4.16

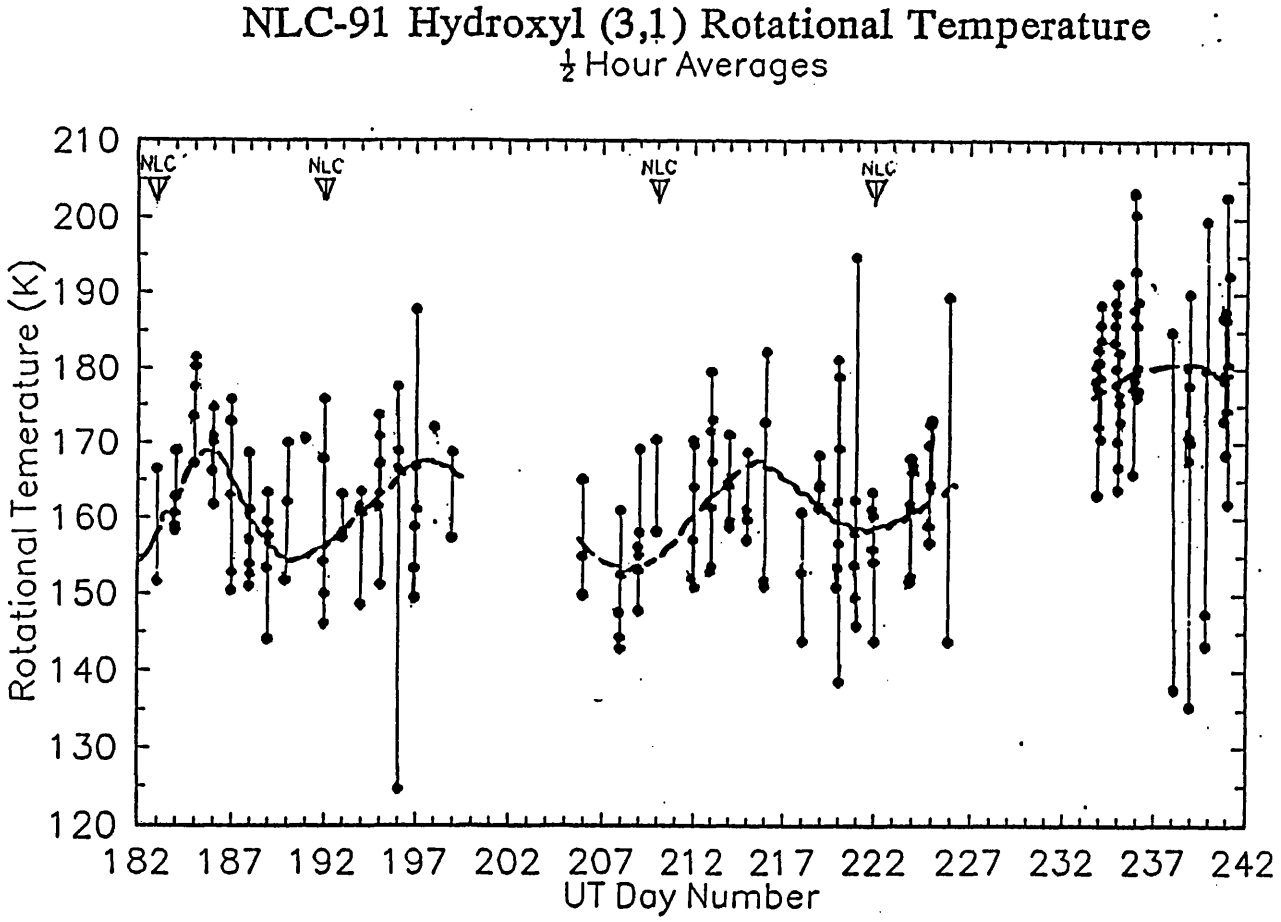


Figure 4.17

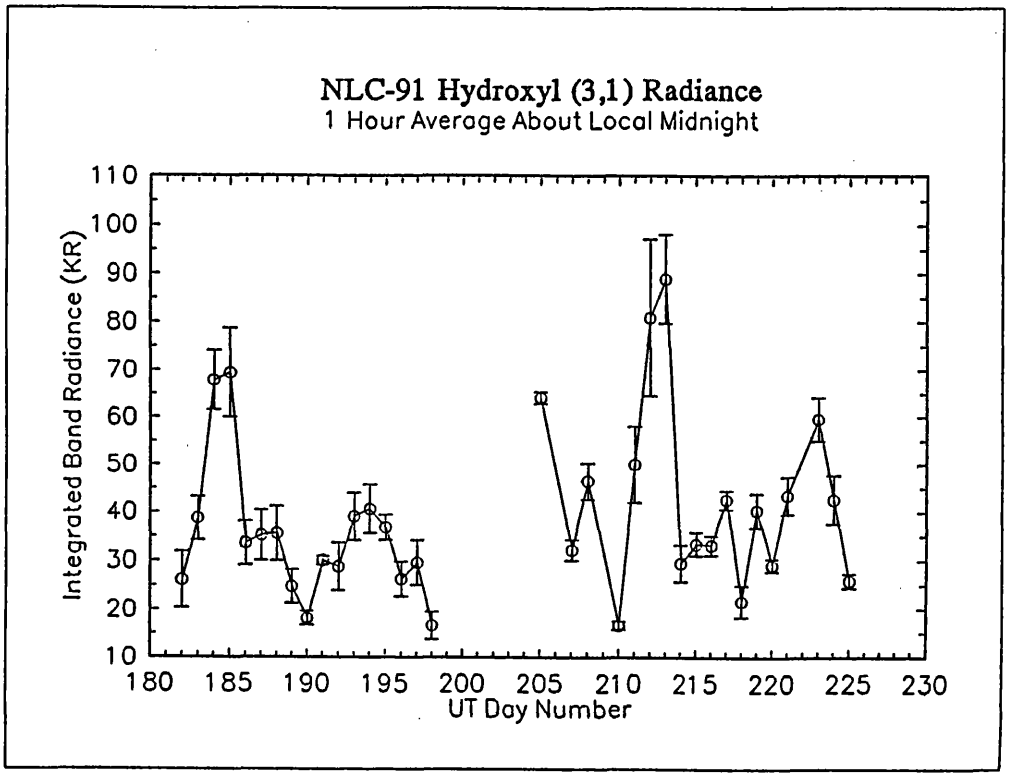


Figure 4.18

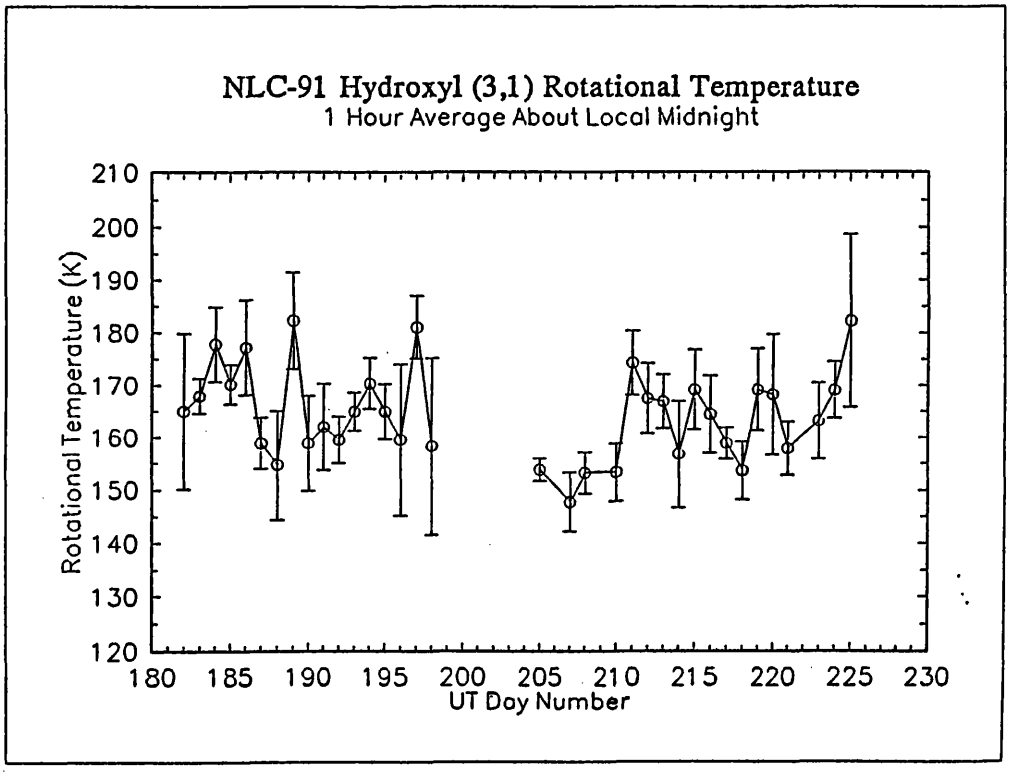


Figure 4.19

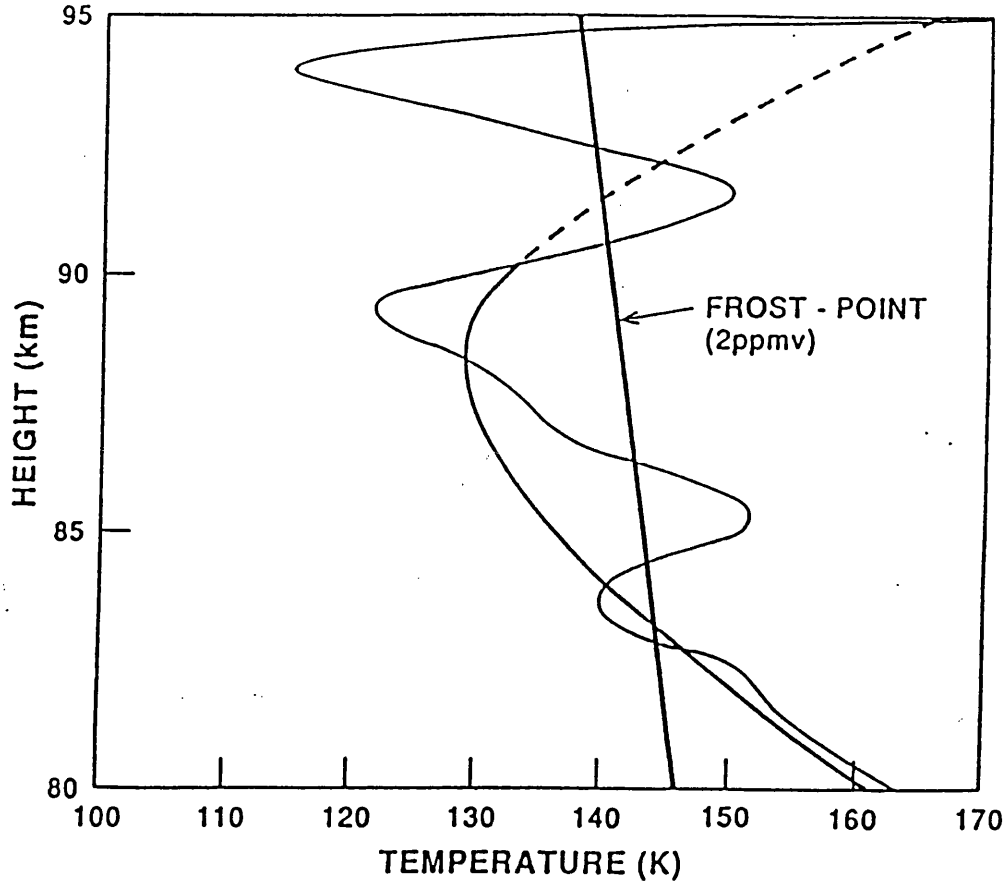

95
90
85
80
HEIGHT (km)
FROST - POINT
(2ppmv)
100
110
120
130
140
150
160
170
TEMPERATURE (K)

Figure 4.20

Chapter 5

5 The Fabry-Perot Interferometer Deployed to Measure Mesospheric Winds from Bear Lake Observatory

5.1 Introduction

This chapter describes the ground-based instrumentation deployed in the observations of the infra-red (IR) hydroxyl (OH) emission from the ($P_1(3)$) emission line in the Meinel (6,2) band at 8430 Å.

Measurement of the 8430 Å emission band is achieved using a Fabry-Perot Interferometer (FPI) coupled to a high sensitivity Imaging Photon Detector (IPD). The IPD is a compact, two-dimensional image encoding system, designed for use in these low-light level applications, and it is discussed in Section 5.2.

The principles of operation, and construction of the Fabry-Perot etalons are discussed in Section 5.3.

The Imaging Fabry-Perot Interferometer (IFPI) located at the Bear Lake Observatory (BLO), Utah is introduced in Section 5.4

Data collection and analy zing routines are discussed in Section 5.5.

Finally, Section 5.6 summarizes the problems encountered during

data analysis and data collection.

5.2 The Imaging Photon Detector

The Imaging Photon Detector (IPD) is a small compact two-dimensional image encoding system, employed for low-light level applications where the available signal is photon limited. A detailed description of the earlier design and development of the IPDs utilized in the Atmospheric Physics Laboratory (APL) at University College London (UCL), may be found in the literature [Rees **et al.**, 1980; 1981; McWhirter **et al.**, 1982][145,146,147]. A similar device, using discrete anode elements, is described by Killeen **et al.**, [1983][148].

Since the 1980s, the IPD has been employed by the APL as an imaging device in a wide variety of applications, the most important of which are discussed by Rees **et al.**, [1981]; Rees **et al.**, [1982]; Smith **et al.**, [1986]; Winser **et al.**, [1988]; Rees **et al.**, [1984]; Batten **et al.**, [1988][149,150,151,152,153,154]. The theory of the operation of the IPD is discussed in Section 5.2.1. The photocathode used is examined in Section 5.2.2. The image collection, and storage are discussed in Section 5.2.3.

5.2.1 Theory of Operation of the Imaging Photon Detector

The IPD is a high-sensitivity, two-dimensional photon counting device. The "X", "Y" position coordinates of each photon event

are output in either digital, and/or analogue form, enabling the image to be viewed in real time on an oscilloscope and/or send the digital data to a micro-computer for collection, integration, storage, and subsequent analysis.

The IPD utilizes a proximity-focused transparent photocathode, three micro-channel plates (MCPs) in a "Z" configuration and a proximity-focused "linear" resistive anode as the image encoder, are contained in a sealed tube evacuated to typically 10^{-8} torr. The detector head is shown in Figure 5.1. A device with a 1.0 mm photocathode to MCP gap, and a potential difference of about 250 V across this gap, has a typical point spread function of 86 microns Full Width Half Maximum (FWHM) [McWhirter **et al.**, 1982][147].

The IPD consists of three MCPs held in a "Z" configuration staked one on top of each other. The "Z" configuration minimizes ion feedback, which is further reduced by a thin (about 7 nm) film of Aluminium Oxide on the first of the three MCPs [Csorba, 1979][155]. This layer reflects any electrons or ions back which can originate from the MCP, and thus protect the photocathode from irreparable damage and degradation, therefore maximizing the working life of the IPD. Within the first MCP all the electrons are contained within a single channel (diameter about 12 microns). However, on both the successive MCP interfaces a certain amount of spreading occurs so that an electron cloud of ~10^7 electrons emerges from 25 to 80 channels of the final MCP. This charge cloud is proximity focused on to a position sensitive

anode encoder.

The position sensitive encoder is in the form of a resistive anode sheet [Lampton and Carlson, 1979][156]. This consists of linear resistors located at the apexes of the four circular arcs, to produce the effect that the resistive anode is infinite in extent. The position of the charge cloud centroid is determined from the ratios of the pulse amplitudes, so that a continuous read-out of location in two orthogonal axes is obtained. This operation is performed in the Signal Processing Unit (SPU).

If the resistive anode has corners labelled A, B, C, and D, then the algorithm used to calculate the two-dimensional location of the charge cloud (X,Y) is given as:-

$$X_{position} = \frac{(A + B) - (C + D)}{(A + B + C + D)} \qquad Y_{position} = \frac{(B + C) - (A + D)}{(A + B + C + D)}$$

IPDs have a saturation limit set by the dead-time of the system, which is of the order 7 μs. The maximum observable intensity across the whole photocathode is 1/dead-time, and is thus of the order 150 000 counts per second.

The charge pulses collected from the IPD anode do not all have the same amplitude. The pulse height spectrum is a function of channel plate saturation, and is normally well defined. The upper and lower threshold (UT and LT) controls set a pulse height window which rejects noise and pulses which are too large or too small for the processing electronics to handle. The LT is set at around - 0.25 V, and the UT is normally set to 3.0 V. The position of the LT is fairly critical and is just above the noise level. MCP noise occasionally gives rise to low-level pulses which appear as "hot-spots" on the image. These can often be removed by adjusting the LT, though sometimes at the expense of genuine photon counts. The purpose of the UT is to block large pulses caused by a combination of pulse overlap, internal discharge and cosmic rays which do not represent real data.

The diameter of the photocathode of the IPDs used is 18 mm. The image is processed so that it is stored in a 256 x 256, integer x 2 array, with the first row being used as a header and storing information appropriate to the image. The spatial resolution achieved with 256 x 256 pixels is quite adequate for most

applications, and an image of this dimension is a convenient size for use with a micro-computer.

5.2.2 The Gallium Arsenide Photocathode

There are currently three types of photocathode used in the APL, namely the S20, the S25, [Winser **et al.**, 1988; Batten **et al.**, 1988; Meredith, 1990][152,154,157], and more recently, the Gallium Arsenide (GaAs) photocathode. The GaAs IPD is the most sensitive photocathode in the near infra-red parts of the spectrum. It has roughly a constant Detective Quantum Efficiency (DQE) of around 20% from 6500 Å to 8000 Å. On the short wavelength side, the DQE falls off to 10% at 6300 Å and to 5% at 6000 Å. On the long wavelength side, the DQE falls off to 10% at 8600 Å and to 5% at 8900 Å. The GaAs device has to be cooled to around -20°C to obtain an acceptable level of thermionic emission. Currently, one of the APL GaAs IPDs is located at the Center for Atmospheric and Space Sciences (CASS), Utah State University (USU), where it is used as the detector of one of the APL FPIs to derive mesospheric winds by measuring the Doppler shift of emissions from hydroxyl (OH) at 8430 Å.

5.2.3 Image Collection and Storage

The IPD is controlled by a microcomputer such as an IBM PC. The run program is used to collect, display, store and/or reduce images of the desired integration time.

The reduced file is named automatically according to the UT date. The method adopted for naming files is best described by referring to an example filename:-

Example of PC filename: V92704.TAB

The first label (V) represents the FPI location, and the second two digits represent the year the observation was made. The third digit represents the month, and follows the assigned notation 1 - 9, represent the months of January to September, and A - C, represent the months of October to December respectively. The last three characters after the decimal point refers to the type of data contained in file. Thus, the example filename was a file containing interpolated peak positions at fifteen minute intervals for data collected from the Utah IFPI observing 8430 Å emission on 4th July 1992.

5.3 Fabry-Perot Etalon

The principles of the Fabry-Perot etalon are outlined in Section 5.3.1, and the construction of the Fabry-Perot etalons are discussed in Section 5.3.2.

5.3.1 Theory of the Fabry-Perot Etalon

In principle, the Fabry-Perot etalon consists of two glass plates set so that the two surfaces facing each other are perfectly flat and parallel, separated by a distance d (see Figures 5.2 and

5.3). The two plates are slightly wedge-shaped, and are anti-reflectance coated to reduce interference from any reflections off the outer surfaces, while the two inner surfaces are coated with a dielectric film for high reflectance, R_f. Its operation is based on the multiple interference of a single beam between the two inner surfaces, but with each reflection a small fraction of light is transmitted out. These successive reflections between the first and second surfaces produce various path differences. From the geometry, the optical path difference between adjacent transmitted rays is given by:-

$$2d \cos \Theta \qquad (5.1)$$

The phase lag Φ of each beam is exacted by:-

$$\Phi = \frac{2 \pi \mu}{\lambda} d \cos \Theta \qquad (5.2)$$

where μ = the refractive index of the dielectric film

λ = the wavelength of light

When the path difference is an integer number of wavelengths, positive interference occurs resulting in a bright fringe, and when the path difference is an integer plus a half number of wavelengths, destructive interference occurs resulting in a minima in light throughput. Therefore, the condition for

constructive interference is shown by:-

$$n\lambda = 2d\cos\Theta \qquad (5.3)$$

Figure 5.4 illustrates the narrow concentric Fabry-Perot fringes corresponding to the multiple beam transmission pattern.

The particular wavelength difference where the nth order for one wavelength λ overlaps with the (n + 1)th order for another wavelength $\lambda + \Delta\lambda$ is known as the Free Spectral Range (FSR). This will occur when:-

$$n\lambda = 2d\cos\Theta \qquad (5.4)$$

and

$$(n + 1)\ (\lambda + \Delta\lambda) = 2d\cos\Theta \qquad (5.5)$$

Thus, $\Delta\lambda$, which is the change in the wavelength necessary for a fringe to move through one FSR is given by:-

$$\Delta\lambda = \frac{\lambda^2}{2d + \lambda} \qquad (5.6)$$

since $\lambda \ll d$

Equation (5.6) becomes:-

$$\Delta \lambda = \frac{\lambda^2}{2d} \qquad (5.7)$$

It can be shown [Born and Wolf, 1987][158] that by summing the amplitudes of light rays emerging from the etalon, the transmitted intensity I(t), is given by Equation (5.8):-

$$I(t) = \frac{I(0)}{\left(1 + F \sin^2 \left(\frac{\Phi}{2}\right)\right)} \qquad (5.8)$$

where

$$F = \frac{4R_f}{(1 - R_f)^2} \qquad (5.9)$$

where I(t) = the transmitted intensity

I(0) = the maximum intensity

Φ = the phase difference between adjacent transmitted rays

= ((4 * π * n * d)/ λ) * cos θ

R_f = the reflectance of the coated surface of the etalon plates

n = the refractive index of the material in the gap

d = the etalon gap

λ = the wavelength of the light

θ = the angle of emerging rays

F = the coefficient of finesse of the etalon

Equation (5.8) is illustrated by the Figure (5.5), intensity versus phase difference for several different values of reflectance.

The finesse of the instrument, f ($=\pi\sqrt{F}/2$), is an important parameter since it determines the ratio of the FSR to the Full Width Half Maximum (FWHM) of the Fabry-Perot fringe.

From Equation (5.4), the central fringe is given by:-

$$n = \frac{2d}{\lambda} \qquad \textbf{(5.10)}$$

Since all the fringes in the field-of-view are of a high order, and close to normal incidence, the small angle approximation can be made:-

$$\cos \Theta = 1 - \frac{\Theta^2}{2!} + \ldots \qquad \textbf{(5.11)}$$

Equation (5.4) becomes:-

$$n \lambda = 2d \left(1 - \frac{\Theta^2}{2!} + \ldots\right) \qquad \textbf{(5.12)}$$

For small angles, the radius of the observed fringe at the detector is linearly related to the angle of incidence. Thus, the order of the fringe is proportional to the square of the angle of incidence θ, and hence, the square of the fringe radius. The spacing of the fringes will be linear in radius-squared space.

5.3.2 Construction of the UCL Fabry-Perot Etalon

The design and fabrication of the Fabry-Perot etalons used by the APL have been discussed in the literature [Rees **et al.**, 1981; Killeen **et al.**, 1983][148,149]

The etalons are constructed from two plates of high quality fused silica separated by three spacers made from Zerodur. Zerodur is a polycrystalline glass ceramic of low thermal expansion coefficient. The two inner surfaces are polished flat to better than 3 nm rms ie λ/200. These faces are coated with a multi-layer dielectric coating, such as Zinc Oxide, to give a peak

reflectivity of over 90% over a specified wavelength range.

The two outer surfaces are flattened to within 300 nm, and are nominally set with a wedge angle to the inner etalon surfaces of 45 arc minutes. This offsets the fringes caused by interference between the inner and outer surfaces. The fringes are further decreased by anti-reflection coating on the outer surfaces resulting in a low reflectivity over the wavelength range at which the etalon is to be operated.

The etalon is housed in a mount made of Invar, (metal alloy with a low coefficient of thermal expansion) which is used to mechanically hold the etalon, and to allow adjustment of the etalon gap by three cantilever springs.

Once the etalon plates are correctly set up, they are sealed in an air-tight container, and evacuated to about 1-10 mm Hg. This prevents atmospheric pressure changes from altering the refractive index of air in the gap. It also minimizes the effects of any temperature instability.

5.4 The Imaging Fabry-Perot Interferometer

The Imaging Fabry-Perot Interferometer (IFPI) located at the Bear Lake Observatory (BLO), Utah (41.93°N, 111.42°W) has been optimized to measure the 8430 Å emission from the $P_1(3)$ emission line. This instrument has been utilized to observe this near infra-red emission. At this latitude, except for periods of very

high solar geomagnetic activity, only the extremely weak airglow emissions are normally observable.

The IFPI consists fundamentally of a scanning mirror, a Fabry-Perot etalon, a telephoto lens, an interference filter and a Gallium Arsenide (GaAs) Imagaing Photon Detector. A calibration lamp is also provided to monitor the instrument stability.

The optical components are mounted on a rigid four-bar optical bench which is extremely stable. This assembly ensures accurate alignment of the instrument. The temperature of the etalon is monitored using a thermal control unit, and controlled to 0.1°C.

The Fabry-Perot etalon has a working diameter of 150 mm, with a fixed etalon gap of 20.49 mm, designed to overlap the lambda-doubled (6,2) hydroxyl (OH) line and so, enhance the signal from the weak OH emission. The etalon parameters are summarized in Table 5.1.

Table 5.1

working diameter of etalon plates	150 mm
etalon gap	20.49 mm
Free Spectral Range (λ units)	17.34 pm
Free Spectral Range (velo. units)	6171 ms^{-1}
effective focal length of telephoto lens	~1100 mm
diameter of IPD	18 mm
Free Spectral Range (R^2 bins)	140.9
velocity shift per bin	43.80 ms^{-1}

The telephoto lens consists of a positive lens of focal length

750 mm and a negative lens of focal length -200 mm, set up to produce an effective focal length of approximately 1100 mm.

The image is recorded by an 18 mm diameter GaAs-photocathode IPD. This arrangement maps approximately 1.8 Free Spectral Ranges onto the detector window (see Figure 5.6).

The Free Spectral Range (FSR) in radius-squared bins is determined by examining a sequence of calibration images, and evaluating the average value of the separation of the peaks. A value of 140.9 was deduced for this parameter, yielding a velocity shift of 43.8 m/s per bin.

5.4.1 Mode of Operation of the IFPI

The scanning mirror of the IFPI rotates in a clockwise direction through eight positions of the sky at a 60° zenith angle. The standard mirror numbers are:-

Mirror Direction	Standard Number
North (N)	1
East (E)	2
South (S)	3
West (W)	4
North West (NW)	5
North East (NE)	6
Zenith (Z)	7
Calibration Lamp	8

South West (SW)	9
South East (SE)	10

The mirror scans the mirror positions, stopping to take an image at each position. The mirror is driven by a motor, and the opto-sensors sense the markers on the rim at the top of the IFPI which indicate the stopping positions. The mirror will then remain at each marker for an integration time of 240 s whilst the IPD collects photons. The integration time for the calibration lamp is 120 s. The field-of-view of the IFPI is 1°. The viewing geometry of the IFPI is clearly illustrated in Figure 5.7. The height of the OH emission layer is taken to be at 87 km. The scanning mirror rotates at a 60° zenith angle. Utilizing the right-angled triangle rule, the unknown length, O, is given by:-

$$O = 87 \tan 60^\circ$$

Therefore, the possible geographic extent covered is twice the distance O, ie 300 km.

5.5 Data Analysis

5.5.1 Collecting Data

The light rays impinging on the Fabry-Perot etalon enter at angles within the 1° field-of-view. This allows the small angle approximation to be used in calculating the positions of the

fringes (see Section 5.3.1). A direct proportionality between the peak at an angle θ, and the square of the radius of the peak, R, may be derived, resulting in the Airy function of Equation (5.8) translating simply into a function dependent on R^2. Once a two-dimensional image has been collected, it is reduced to a one-dimensional image in terms of the R^2 values, and stored as a one-dimensional array of data.

The centre of the image must be found with care during the calibration procedure. It is essential to obtain the best possible value for the centre, since all the data will be reduced to radius-squared using the radii determined from this centre position.

5.5.2 Calibration of the Imaging System

The following two calibrations improve the quality of the results. It should be noted that the calibration lamp discussed throughout this work is a RF (Radio Frequency) excited discharge tube of Ne gas, maintained at a pressured of a few Torr. Although absolute measurements are not obtained for both the calibration and sky-line observations, the RF calibration lamp emits a wavelength that is within 5 Å of the line that is being observed, 8430 Å in this case. A more accurate method of calibration determination is presently being investigated (Rees, personal communication).

5.5.2.1 Thermionic Emission Calibration

The thermionic emission is equivalent to measuring the dark current of the photocathode. Thermal electrons emitted from the photocathode cannot be discriminated from photoelectrons and are a source of noise. However, the rate of thermionic emission for a given IPD can be removed by performing a thermionic emission calibration. Since, the thermionic emission is a function of temperature, this calibration is performed at the temperature at

which the detector is to be operated. The photocathode is light-sealed, and a long integration (about 1 hour) performed. This is the thermionic emission data file. The long integration period is required to allow a useful signal-to-noise ratio to be obtained, since the thermionic emission rate is normally very low, of the order 10^{-2} - 10^{-3} counts per pixel per second.

5.5.2.2 Flat Field Calibration

The detector does not respond uniformly to an even source of light. However, the IPD is a photometrically and geometrically stable device, so that the flat field response depends mainly on the wavelength of the incoming light obtained by observing a diffuse white light source.

5.5.3 Fitting Peaks

Data analysis involves fitting a function to the peaks in order to obtain values of the radius, line intensity, line width, and background intensity.

The IFPI measures the frequency spectrum S(f), which is excited by a source E(f). Since this source has a finite spectral width, the resulting spectrum is represented by the convolute of S(f), and E(f). The convolution is represented by "*". This spectrum is then detected with the FPI whose transmission profile which is a function of frequency, is given by F(f). The ideal measured data Y(f) is an interferogram which has the following convolution

[Cooper, 1971][159]:-

$$Y(f) = S(f) * E(f) * F(f) \qquad \textbf{(5.13)}$$

The transmission profile F(f) is the convolute of several instrument qualities as follows [Cooper, 1971][159]:-

$$F(f) = A(f) * D(f) * M(f) * P(f) \qquad \textbf{(5.14)}$$

where A(f) = the ideal etalon Airy function
D(f) = the defect finesse due to curvature
M(f) = the microscopic imperfection effect
P(f) = the pixel aperture effects

The fitting routine adapted in this data analysis was developed by Batten [1989][160] which deals with the data in Fourier space, and therefore has the potential to be modified to trace out temperatures. A smooth fit may be made over the whole data profile, as opposed to just one peak, using a cosine bell (Hanning) filter [Koopmans, 1974][161] to cut out high frequency components.

5.5.4 Determination of Winds from Peak Positions

The main problem encountered in deriving the winds from the peak

positions, arises as a result of the absolute zero Doppler shift (or base-line) position determination. The base-line is deduced using the assumption that the average vertical wind component is normally negligible compared to the horizontal wind component [Aruliah, 1991; Aruliah and Rees, 1993][162,163].

For a passive emission line, the Doppler shift is simply given as:-

$$v = \frac{c\,\Delta\,\lambda}{\lambda} \qquad (5.15)$$

where v = the line-of-sight Doppler velocity
c = the speed of light
$\Delta\lambda$ = the Doppler shift in absolute wavelength
λ = the mean operating wavelength

The principles of error propagation, and the criteria strictured throughout the data analysis are outlined, and discussed in the following sections.

5.5.4.1 Errors in Peak Positions Anticipated on the Basis of Photon Counting

Ideal FPI fringes may be described by the Airy function. This function takes the form (recall Equation (5.8)):-

$$I(t) = \frac{I(0)}{\left(1 + F \sin^2 \left(\frac{\Phi}{2}\right)\right)}$$

The Airy function represents an idealized case which does not take into account the finite spectral width of the source, or certain secondary instrumental effects such as the defect finesse due to curvature, microscopic imperfection effects, or pixel aperture effects. However, the aim of this investigation is to prove the accuracy of the peak fitting routines, and consequently the Airy function may be used to represent the absolute maximum obtainable accuracy in the presence of photon noise alone.

The random fluctuation in the photon flux is described by a Poisson distribution. This distribution has the important property that the root mean square fluctuation in the average flux N is simply the square root of N.

A good test of the fitting accuracy may therefore be obtained by simulating data with the appropriate values of I(0). It was decided to perform the test on profiles calculated with values I(0) ranging logarithmically from 10 to 10000. The scale is sufficiently wide to cover the anticipated intensity range during an observation. For each value of I(0) tested twenty-four radial profiles with Poisson noise are generated, and the fit performed on each. The mean, and standard deviation of each peak position were determined. The results are presented in Table 5.2, in terms of a fraction of the FSR for two peaks [Meredith, 1991][164]. These

results were then converted into R^2 bins, and then ms^{-1}. These are specific values calculated for the BLO IFPI.

Table 5.2

Intensity	1σ error fraction of FSR	1σ error R^2 bins	1σ error ms^{-1}
10	0.0182	2.57	112.50
20	0.0154	2.17	95.00
50	0.00922	1.30	56.90
100	0.00603	0.85	37.20
200	0.00443	0.62	27.10
500	0.00296	0.42	18.40
1000	0.00174	0.24	10.50
2000	0.00108	0.15	6.56
5000	0.000894	0.126	5.51
10000	0.000704	0.099	4.33

Figure 5.8 plots one sigma (σ) error in ms^{-1} against the peak intensity, I(0) for the BLO IFPI. The following relationships are obtained:-

$$\sigma\ (m/s) = \frac{400.00}{\sqrt{I(0)}} \qquad \textbf{(5.16)}$$

$$\sigma\ (bins) = \frac{9.14}{\sqrt{I(0)}} \qquad \textbf{(5.17)}$$

The zenith peak intensity, normalized to 100 s, has

approximately 200 counts. The actual zenith peak intensity taken over an integration period of 240 s has approximately 500 counts. The one sigma error on the zenith peak position is approximately 0.41 bins ie 17.8 ms^{-1}.

The calibration peak 1 position, normalized to 100 s, has approximately 2200 counts. The actual calibration peak 1 intensity taken over an integration period of 120 s has approximately 2600 counts. The one sigma error on the calibration peak 1 position is approximately 0.18 bins ie 7.8 ms^{-1}.

The calibration peak 2 position, normalized to 100 s, has approximately 600 counts. The actual calibration peak 2 intensity taken over an integration period of 120 s has approximately 700 counts. The one sigma error on the calibration peak 2 position is approximately 0.35 bins ie 15.1 ms^{-1}.

5.5.4.2 Determination of the Base-line

The best method for determining the base-line involves evaluating the average value of the distance between the zenith peak position, and the calibration peak position (calibration - zenith) throughout the night, and using this as an offset from the calibration peak position.

It was therefore decided to take a detailed look at the zenith, and calibration peak positions, and the difference (calibration - zenith) as a function of time. Over the three month period under

examination the following results were obtained:-

	calibration peak	zenith peak	calibration - zenith
mean	22.30	107.70	-85.50
sigma	1.67	1.90	0.94

Examination of the results revealed a systematic drift of the instrument over this period, amounting to approximately 6 bins in three months. The standard deviation quoted on the calibration, and zenith peak positions is thus not representative of the true error to be experienced on any given night. The standard deviation on the difference, however, is likely to be representative of the errors present on the base-line determination (assuming the drift is the same for both the calibration, and zenith peak positions).

The fitting of the peak positions is very sensitive to the cutoff parameter, and so it was decided to investigate the effect of changing the value of this parameter on the base-line determination. The cutoff parameter is utilized to remove higher order harmonics from the data. There are in total 256 bins in radius-squared space, therefore, there are 256 associated different harmonics. The chosen cutoff parameter value only uses that number of harmonics from the total number of harmonics available. Mean values of the calibration peak position, the zenith peak position, and the difference (calibration - zenith), together with their associated errors were determined on a nightly basis, and plotted as a function of time for a variety of values of the cutoff parameter. Nights with less than four sets of neighbouring calibration, and zenith peak positions were ignored in this analysis. The standard deviation was calculated from the spread of the values on any one given night, and was not determined by the fitting program. Figures 5.9a, b, and c, illustrate the peak positions averaged throughout the night for the zenith (a), the calibration (b), and the zenith - calibration (c). Only nights with greater than three data points have been plotted. The cutoff is 15 for each of these plots.

The overall averages, calculated over all the available nights for a particular value of the cutoff parameter are tabulated in Table 5.3. The error on the zenith, and calibration peak positions is the average of the standard deviations for each night. This removes the systematic error associated with the instrumental drift, and is therefore indicative of the error on any given measurement. These are the values to be compared with those expected on the basis of photon counting alone.

The overall averages, calculated over all the available nights for a particular value of the cutoff parameter are tabulated in Table 5.3. The error on the zenith, and calibration peak positions is the average of the standard deviations for each night. This removes the systematic error associated with the instrumental drift, and is therefore indicative of the error on any given measurement. These are the values to be compared with those expected on the basis of photon counting alone.

Table 5.3

cut	zpeak	zsigma	calpeak	cal-sigma	diff	diff-sigma
12	106.34	0.62	23.47	0.37	-82.86	0.59
15	107.73	0.86	22.32	0.38	-85.41	0.75
16	108.15	0.93	22.29	0.40	-85.86	0.83
18	108.22	0.91	22.52	0.45	-85.70	0.89
20	108.26	0.88	22.71	0.45	-85.56	0.90
25	108.15	1.23	22.89	0.51	-85.26	1.22
30	108.11	1.90	23.20	0.46	-84.91	1.87

From Table 5.3, the errors indicate that using a lower value for the cutoff is advantageous. However, the fit to the IFPI data reduced to radius-squared space deteriorates below a cutoff value of 15, but sigma improves paradoxically. Figures 5.10a, b, and c, illustrate the quality of the fit when the cutoff value is

fixed at 12, 15, and 18, respectively.

At best the errors are approximately twice that expected on the basis of the photon counting alone. However, the intensity of the calibration peak is similar to the background, whilst the zenith peak is about a quarter of the background value, and this will certainly result in larger errors, due to the statistical fluctuation of the background.

An estimate of the errors expected due to this additional source may be determined as follows.

5.5.4.3 Error on the Peak Position in the Presence of the Background Noise

The ratio corresponding to an intensity I_n in the absence of any background intensity, is given from:-

$$\frac{\sqrt{I_n}}{I_n} = \frac{\sqrt{I(0) + B}}{I(0)} \qquad \textbf{(5.18)}$$

$$I_n = \frac{I(0)^2}{2\,(I(0) + B)} \qquad \textbf{(5.19)}$$

The error on the peak position in the presence of the background noise, B, is:-

$$\sigma\ (m/s) = \frac{400\ \sqrt{I(0) + B}}{I(0)} \qquad \textbf{(5.20)}$$

$$\sigma\ (bins) = \frac{9.14\ \sqrt{I(0) + B}}{I(0)} \qquad \textbf{(5.21)}$$

where peak intensity = I(0) counts

background intensity = B counts

maximum intensity (peak + background) = I(0) + B

error on peak intensity = $\sqrt{}$I(0)

error on the background intensity = $\sqrt{}$B

error on maximum count = $\sqrt{}$(I(0) + B)

5.5.4.4 Error on the Calibration Peak in the Presence of the Background Noise

peak intensity = 2500 counts

background intensity = 2600 counts

error $(ms)^{-1}$ = 11.42

error (bins) = 0.26

Hence, the expected error on the calibration peak position in the presence of this background noise is 0.26. The observed error is larger than this because the calibration peak position has a systematic drift associated with it.

5.5.4.5 Error on the Zenith Peak in the Presence of the Background Noise

peak intensity = 500 counts

background intensity = 2000 counts

error (ms^{-1}) = 40

error (bins) = 0.91

Hence, the expected error on the zenith peak position in the presence of the background noise is about 0.91 bins, which is indeed what is observed.

Since the average expected wind is of the order of 1 bin, the results presented above suggest that it would be difficult to utilize this particular method for a satisfactory base-line determination.

It was therefore decided to repeat the above analysis using the second calibration peak. The results of this analysis are tabulated in Table 5.4.

Table 5.4

cut	zpeak	zerr	calpeak	cal-sigma	diff	diff-sigma
12	106.30	0.60	161.79	0.48	55.67	0.64
15	107.73	0.86	163.62	0.42	55.89	0.81
16	108.12	0.94	164.17	0.42	56.05	0.80
18	108.17	0.89	164.12	0.49	55.94	0.95
20	108.30	0.86	164.27	0.53	55.96	0.94
25	108.22	1.20	164.60	0.77	56.37	1.39
30	108.29	1.93	164.60	1.10	56.31	2.14

These results are marginally worse than those using calibration peak 1, and so, it was decided to search for an alternative method of the base-line determination.

Another method of determining the base-line involves calculating the average of the North, East, South, and West (NESW) peak positions throughout the night. The above analysis was repeated using the average values of NESW, and the results obtained are tabulated in Table 5.5.

Table 5.5

cut	NESWpk	NESW-sigma	calpeak	cal-sigma	diff	diff-sigma
12	107.32	0.46	23.71	0.39	-83.61	0.43
15	108.20	0.47	22.51	0.43	-85.70	0.41
16	108.44	0.47	22.49	0.46	-85.95	0.41
18	108.44	0.44	22.70	0.50	-85.74	0.46
20	108.35	0.41	22.78	0.51	-85.57	0.47
25	108.28	0.46	23.02	0.53	-85.26	0.55
30	108.39	0.65	23.37	0.50	-85.02	0.73

The analysis was once again repeated using the second calibration peak. The results are tabulated in Table 5.6.

Table 5.6

cut	NESWpk	NESW-sigma	calpeak	cal-sigma	diff	diff-sigma
12	107.32	0.46	162.40	0.54	55.08	0.55
15	108.21	0.47	163.90	0.47	55.69	0.49
16	108.44	0.46	164.51	0.48	56.07	0.48
18	108.45	0.44	164.35	0.58	55.90	0.57
20	108.36	0.41	164.32	0.60	55.97	0.59
25	108.38	0.44	164.72	0.84	56.34	0.87
30	108.48	0.61	164.74	1.24	56.26	1.34

These results illustrate that a better method involves using the average of the NESW over the night, although the errors are still high, and are of the same order of magnitude as the wind to be determined.

5.5.4.6 Estimate of the Base-line

The base-line was estimated by calculating the average value of (North - calibration), (East - calibration), (South - calibration), and (West - calibration) on cloudy nights. The results are tabulated in Table 5.7.

Table 5.7

date of cloudy night	UT day number	average offset
6-June-92	157	-85.8069
7-June-92	158	-85.7529
8-June-92	159	-85.7856
20-June-92	171	-86.0592
25-June-92	176	-85.4533
26-June-92	177	-85.5552
4-July-92	185	-86.1659
8-July-92	189	-86.2166
22-Aug.-92	238	-86.2166
average		-85.8578

The presence of a cloud causes the emitted light to be scattered, so that the Doppler shifts from all directions are scattered, and tend to be mixed together with the Doppler shifts from all other directions. As a result, the light collected by the IFPI has lost its directional information, and is an average of all the Doppler shifts in the sky.

5.5.4.7 Determination of Meteorological Conditions

The determination of the meteorological conditions at BLO is not

a trivial task, as the Observatory is unmanned and currently no instrumentation exists at the site to monitor the weather conditions.

The weather conditions may be estimated by studying the plots of the background intensities. Figure 5.11, for example, illustrates the background intensities observed on the 18th of August 1992, in the presence of an 80% Moon. The interference from moonlight is clearly visible. In comparison, Figure 5.12 illustrates the background intensities observed on the 28th of August 1992, in the absence of the Moon. Here, the curves are seen to nearly overlap, or have a "bunched" appearance. This is indicative of a uniform background intensity across the entire sky. A good clear observing night will thus be characterized by such a pattern employing this procedure.

The results from this first estimate are compared with the observations made from Logan (D.Jensen, personal communication). However, Logan is 40 miles from BLO, and on the other side of the Wasatch Mountain Range, and so these readings will not always be indicative of the conditions at the Observatory.

In screening the data, the nights that exhibited any evidence of cloud and/or lunar contamination were eliminated (see Appendix 3). The resulting cloud-free data is discussed in Chapter 6.

5.5.4.8 Determination of the Line-of-Sight, and the Error on the Line-of-Sight Wind

The line-of-sight wind, is calculated from the following expression:-

V_{ls} = (data - calibration - offset) x velobin

where velobin = 43.8 = velocity per bin (see Table 5.1)

offset = -85.8578 (see Table 5.7)

The error on the data peak, σ_d, is given as:-

$$\sigma_d\ (m/s) = \frac{400\sqrt{I_d + B_d}}{I_d} \qquad \textbf{(5.22)}$$

The error on the calibration peak, σ_c, is given as:-

$$\sigma_c\ (m/s) = \frac{400\sqrt{I_c + B_c}}{I_c} \qquad \textbf{(5.23)}$$

Finally, the error on the line-of-sight wind, σ_w, is given by:-

$$\sigma_w = \sqrt{\sigma_d^2 + \sigma_c^2} \qquad (5.24)$$

$$\sigma_w = \frac{400}{I_d\, I_c} \sqrt{I_c^2\ (I_d + B_d) + I_d^2\ (I_c + B_c)} \qquad (5.25)$$

The detailed error analysis discussed in Section 5.5.4 initially was carried out on the derivation of the winds, and shows that the errors become large at the lowest observed intensities (see Table 5.8). The analysis indicates that at least 1000 counts peak intensity is required ie peak intensity above the background, to achieve a satisfactory wind determination. In the summer of 1992, the measured peak intensities varied from 500 to 2500, and on the basis of this analysis, it is recommended that the integration time is doubled (see Table 5.9), and eliminate the observations made in the NW, NE, SW, and SE to keep the cycle time the same.

Table 5.8

data peak height	data background	calib. peak height	calib. background	error ms^{-1}
500	2000	2500	2500	41.6
1000	2000	2500	2500	24.6
2000	2000	2500	2500	16.9
2500	2000	2500	2500	15.5

The analysis was repeated using figures appropriate to a doubling of the integration time. The results are tabulated in Table 5.9.

Table 5.9

data peak height	data background	calib. peak height	calib. background	error ms^{-1}
1000	4000	5000	5000	29.4
2000	4000	5000	5000	17.4
3000	4000	5000	5000	13.7
4000	4000	5000	5000	12.0
5000	4000	5000	5000	11.0

The results illustrate that an improved method utilizes twice the integration observation period over the night, although the errors are moderate, they are no longer of the same order of magnitude as the wind to be determined.

5.5.4.9 Statistical Analysis of Randomly Generated Wind Speeds

The resulting data indicated an apparent correlative trend between low peak OH intensities, with high wind speeds, and vice versa. It was suggested (Rees, personal communication) that where the peak OH intensities were falling below 500, the working hypothesis might no longer be valid, since, as the peak OH intensities decreased, the wind speed errors increased. To prove that a true correlation exists between the peak OH intensity, and the wind speeds, a random number generator was utilized. This simulation calculated the wind speed from the peak OH intensity assuming a Gaussian distribution with a zero mean wind speed, and a one sigma error (ms^{-1}) calculation from the photon noise. The calibration lamp error was assumed to be zero. The simulation was operated for a number of runs, each run included a different

maximum peak OH intensity input, and a fixed number of data points. The graphic output from each run clearly indicated the spread of the data points about the zero mean wind speed. The test simulation established that the BLO data were real and not shrouded by noise contamination as first speculated. Figure 5.13 illustrates a plot obtained from a simulation run where the maximum peak OH intensity was set at 10000, and two thousand data points were implemented. Figure 5.13 clearly illustrates that data associated with an intensity below 500 gives unreliable winds.

5.6 Problems Encountered During Data Analysis

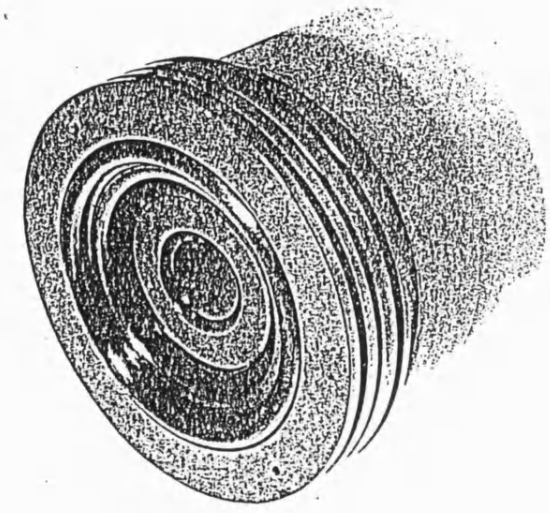
The following section outlines the problems encountered during the 8430 Å data analysis for the summer months of June, July, and August 1992.

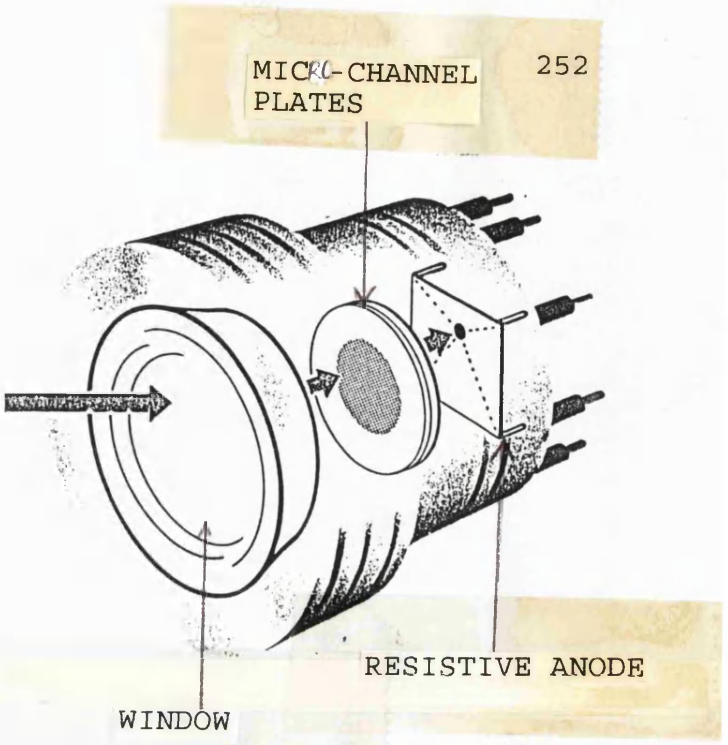
5.6.1 Flat Field Determination

A flat field file was not used in this work. The only flat field available dates from the 26th of January 1992. The next flat field recorded was taken in early November 1992 when the IFPI setup was altered. A trial run was made applying the January flat field, however, the results were implausible. The flat field file was itself of poor quality ie the peak-like structure lies in between the bin number range of 200, and 256. Flat field calibrations need to be monitored on a regular basis.

5.6.2 Instrumental Drift

The IFPI is mounted in a building (a converted mobile trailer) that is thermally controlled, where the temperature is maintained constant to within 3-5°C. The etalon of the IFPI itself has its own thermal control, so it is stable to better than 0.1°C [Rees **et al**., 1982a][150]. Normally, the system works well, and any small temperature drifts are easily compensated for with the use of the calibration lamp data. Occasionally, problems do arise. For example, the data collected between the 31st July, and 15th August 1992 has an unusually high photon count, greater than 12k in some cases. This was simply the result of a very high dark count rate. The thermo-electric cooler of the IPD had failed, so that the IPD was operating at room temperature with a very high thermionic emission. Unless there is an operator present at the observatory site at all times, these problems can go unnoticed for several days or in the worse cases some weeks.





Figures 5.1a and b

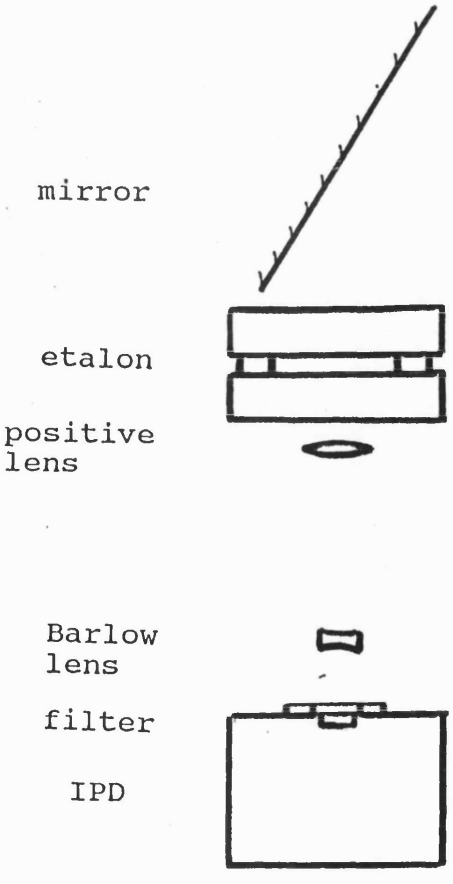
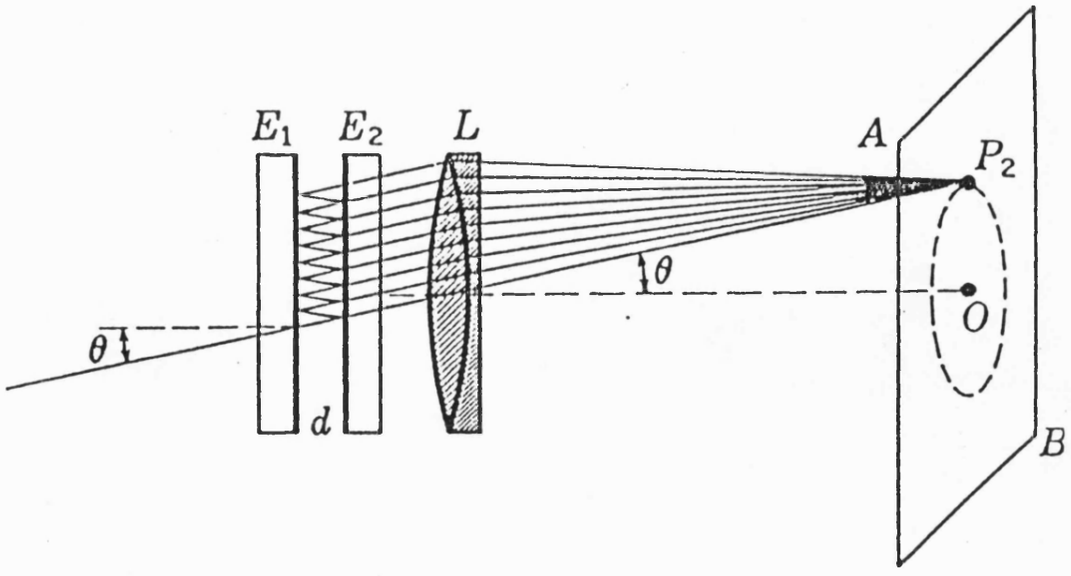


Figure 5.2



Figures 5.3

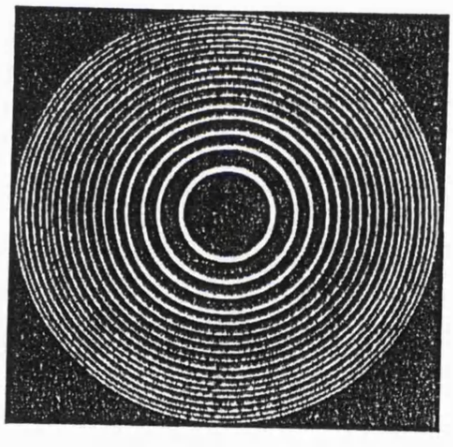


Figure 5.4

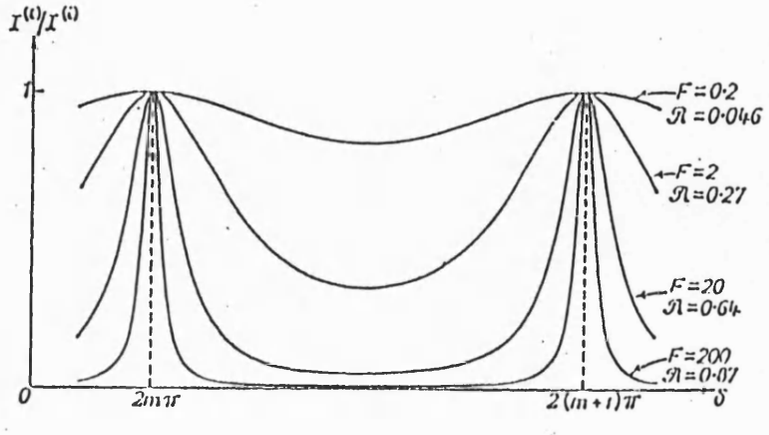


Figure 5.5

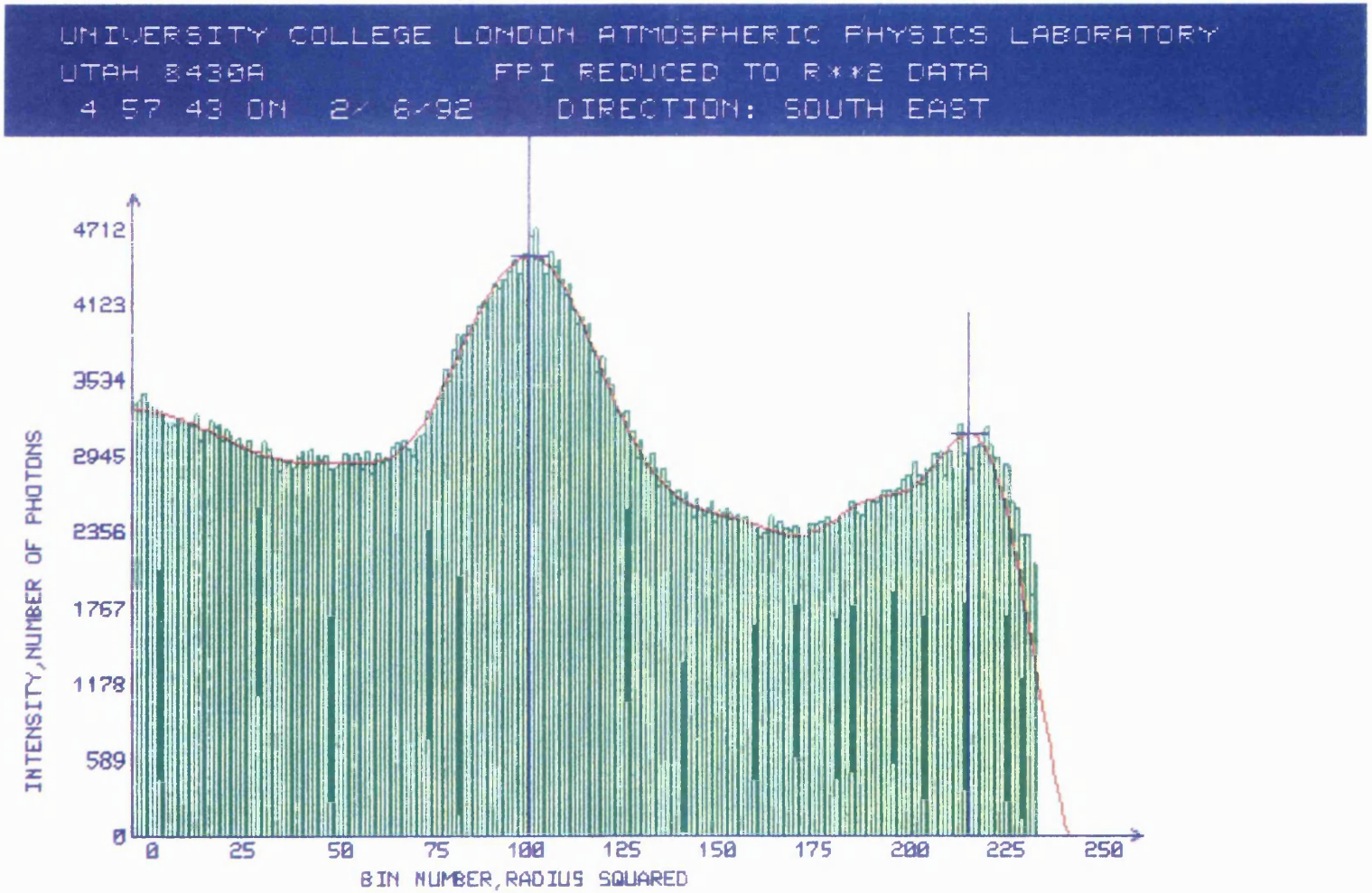
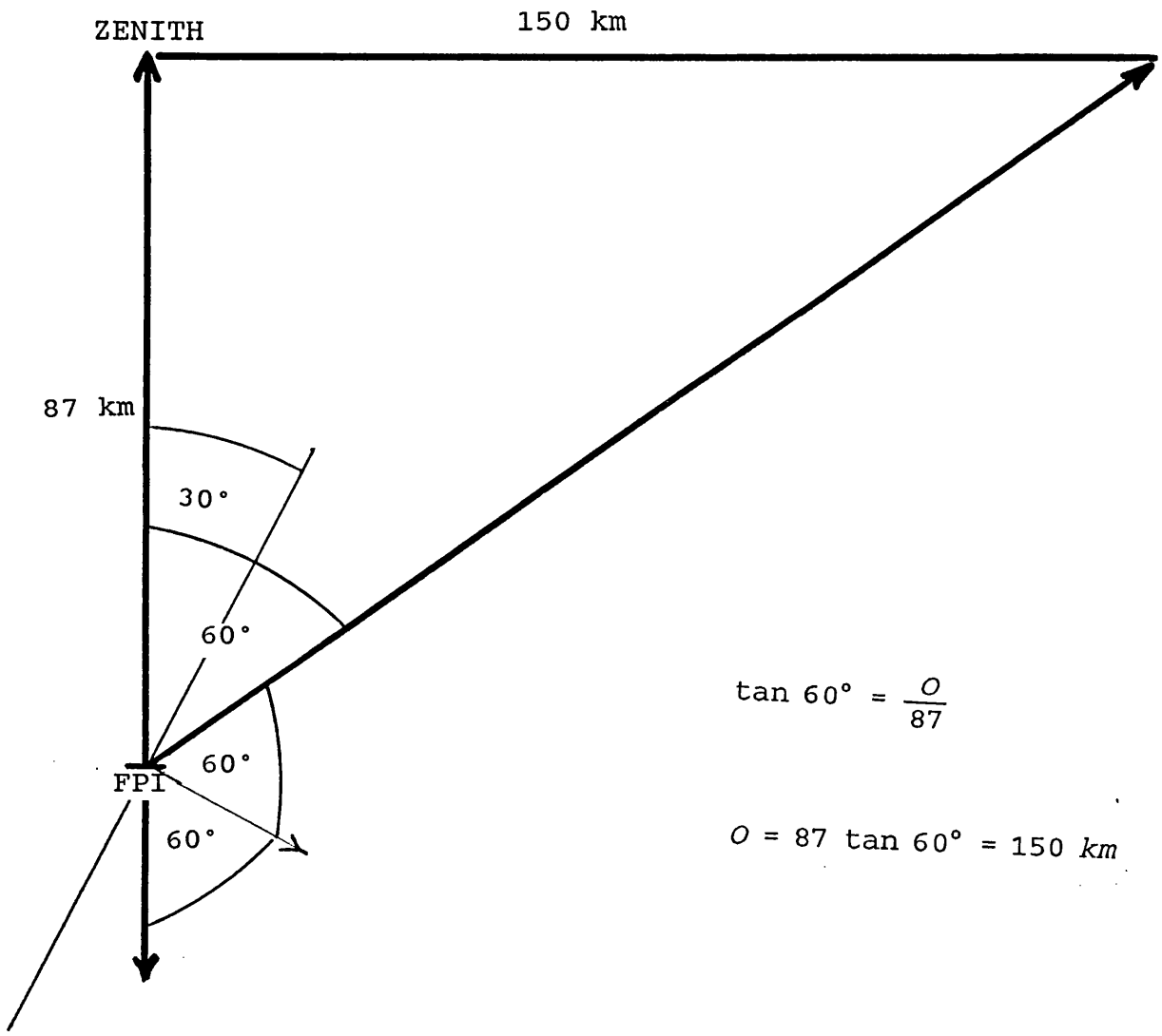

UNIVERSITY COLLEGE LONDON ATMOSPHERIC PHYSICS LABORATORY
UTAH 8430A FPI REDUCED TO R**2 DATA
4 57 43 ON 2/ 6/92 DIRECTION: SOUTH EAST
INTENSITY, NUMBER OF PHOTONS
4712
4123
3534
2945
2356
1767
1178
589
0
0
25
50
75
100
125
150
175
200
225
250
BIN NUMBER, RADIUS SQUARED

Figure 5.6

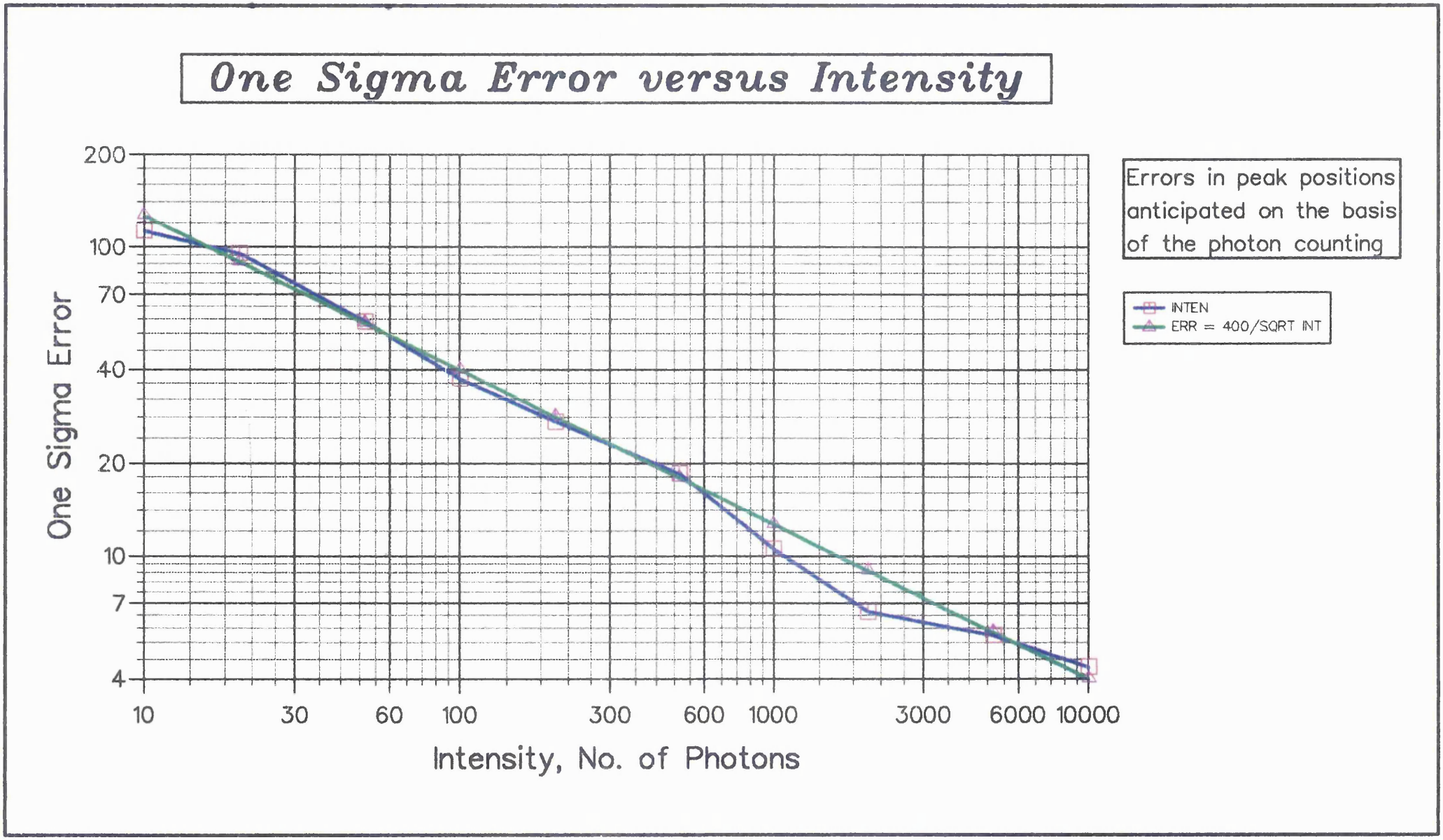


$$\tan 60° = \frac{O}{87}$$

$$O = 87 \tan 60° = 150\ km$$

Figure 5.7

Figure 5.8


One Sigma Error versus Intensity
200
100
70
40
20
10
7
4
One Sigma Error
10
30
60
100
300
600
1000
3000
6000
10000
Intensity, No. of Photons
Errors in peak positions anticipated on the basis of the photon counting
INTEN
ERR = 400/SQRT INT

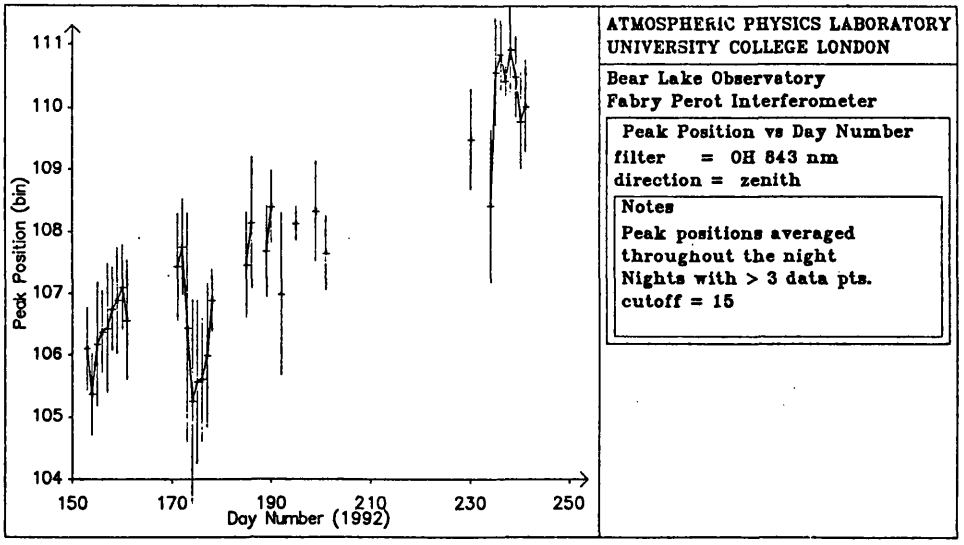


Figure 5.9a

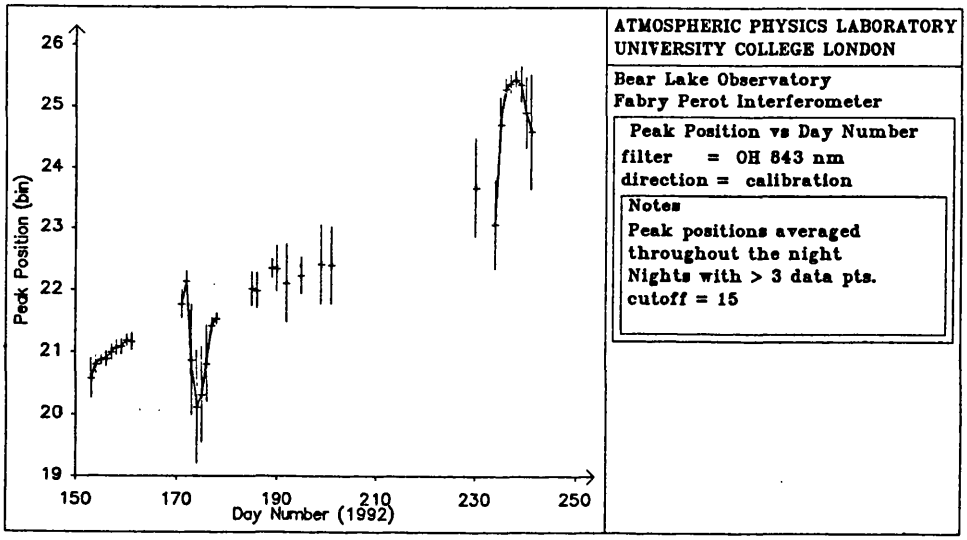


Figure 5.9b

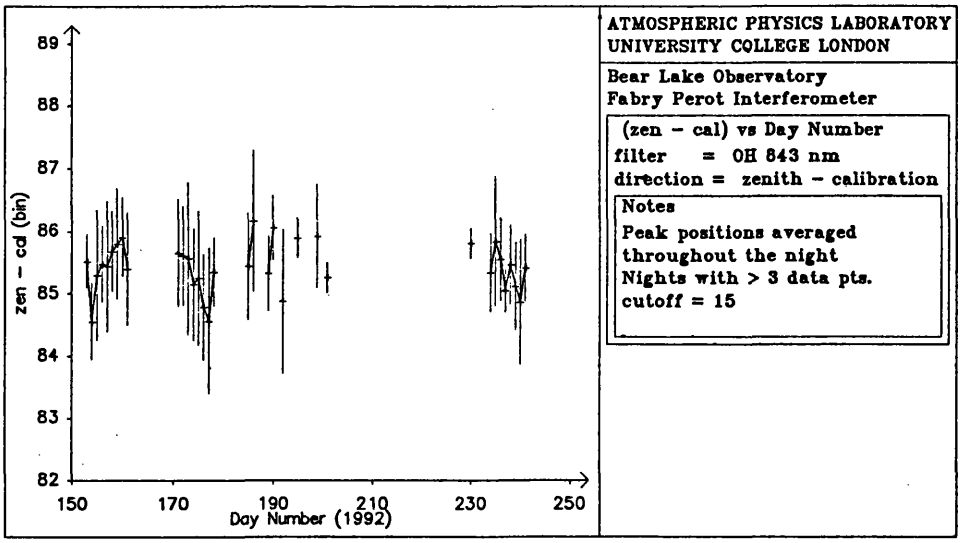


Figure 5.9c

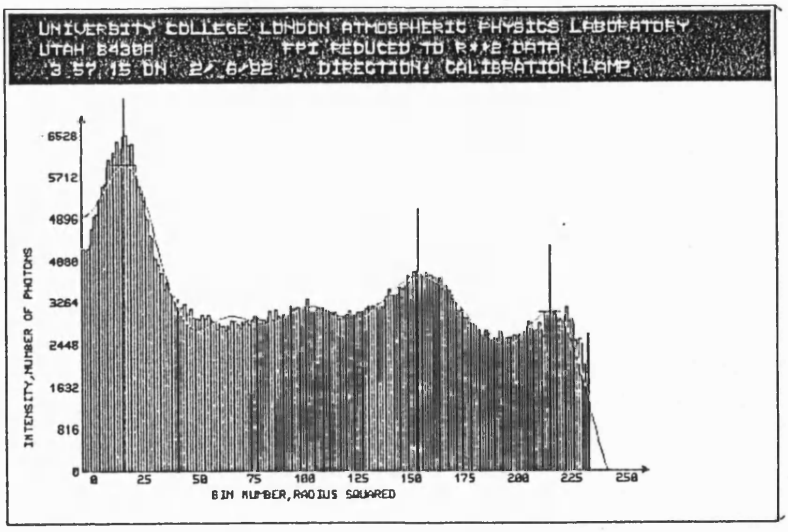


Figure 5.10a

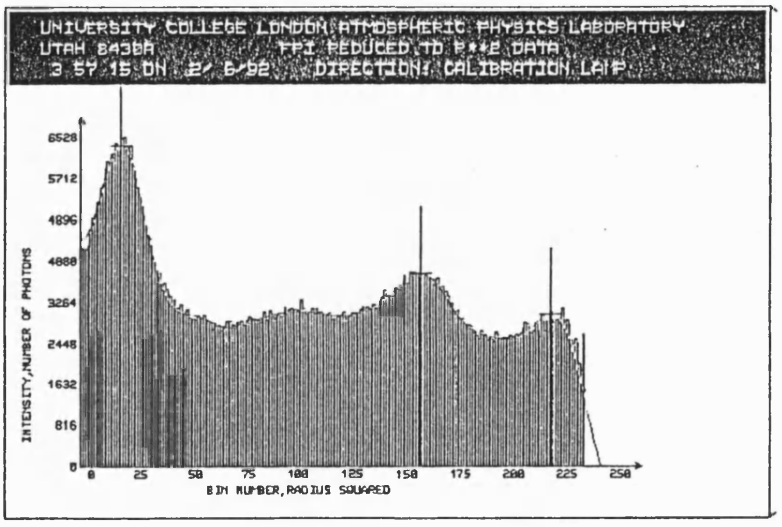


Figure 5.10b

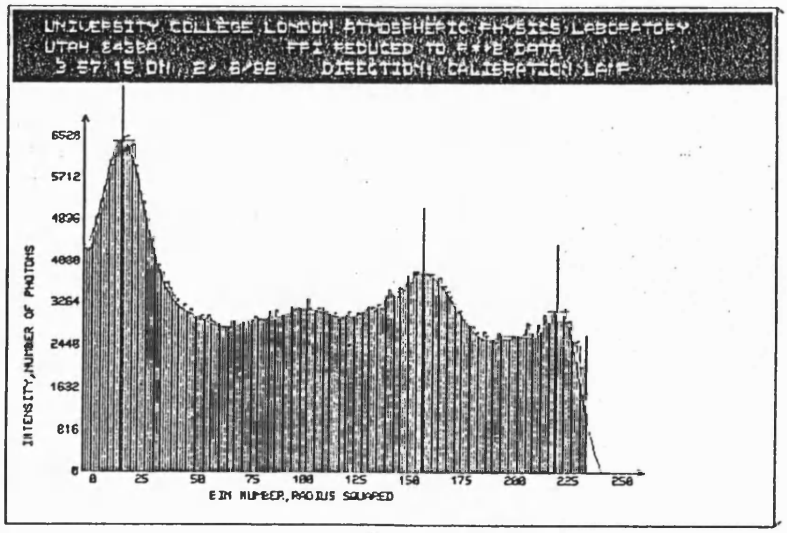


Figure 5.10c

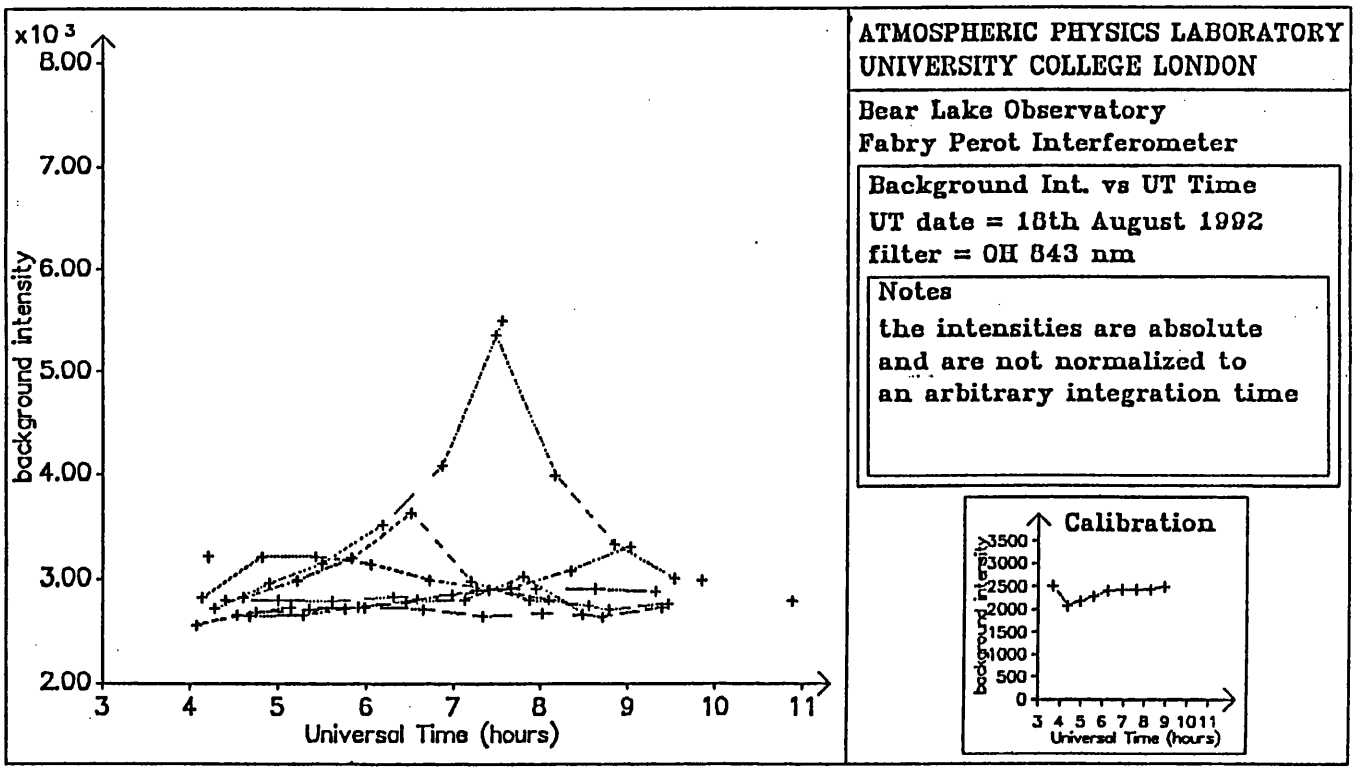


Figure 5.11

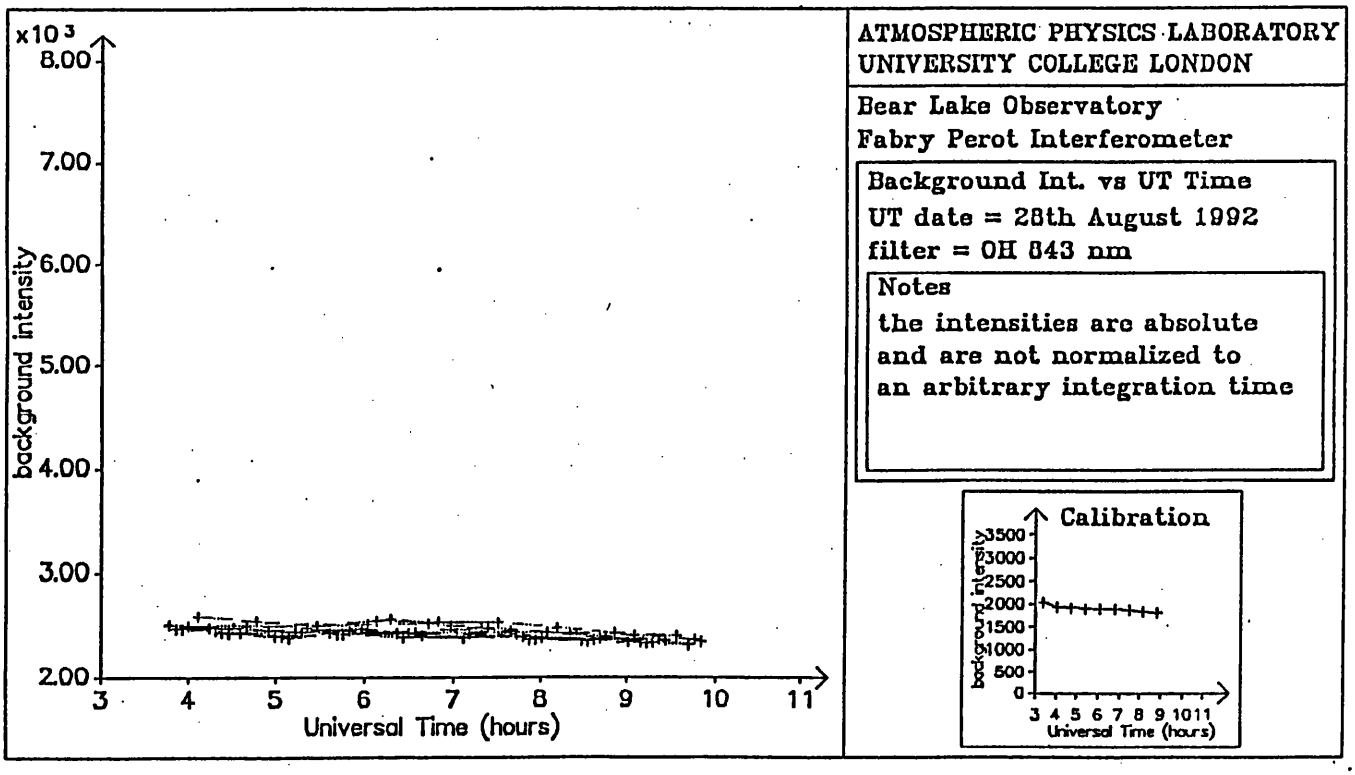


Figure 5.12

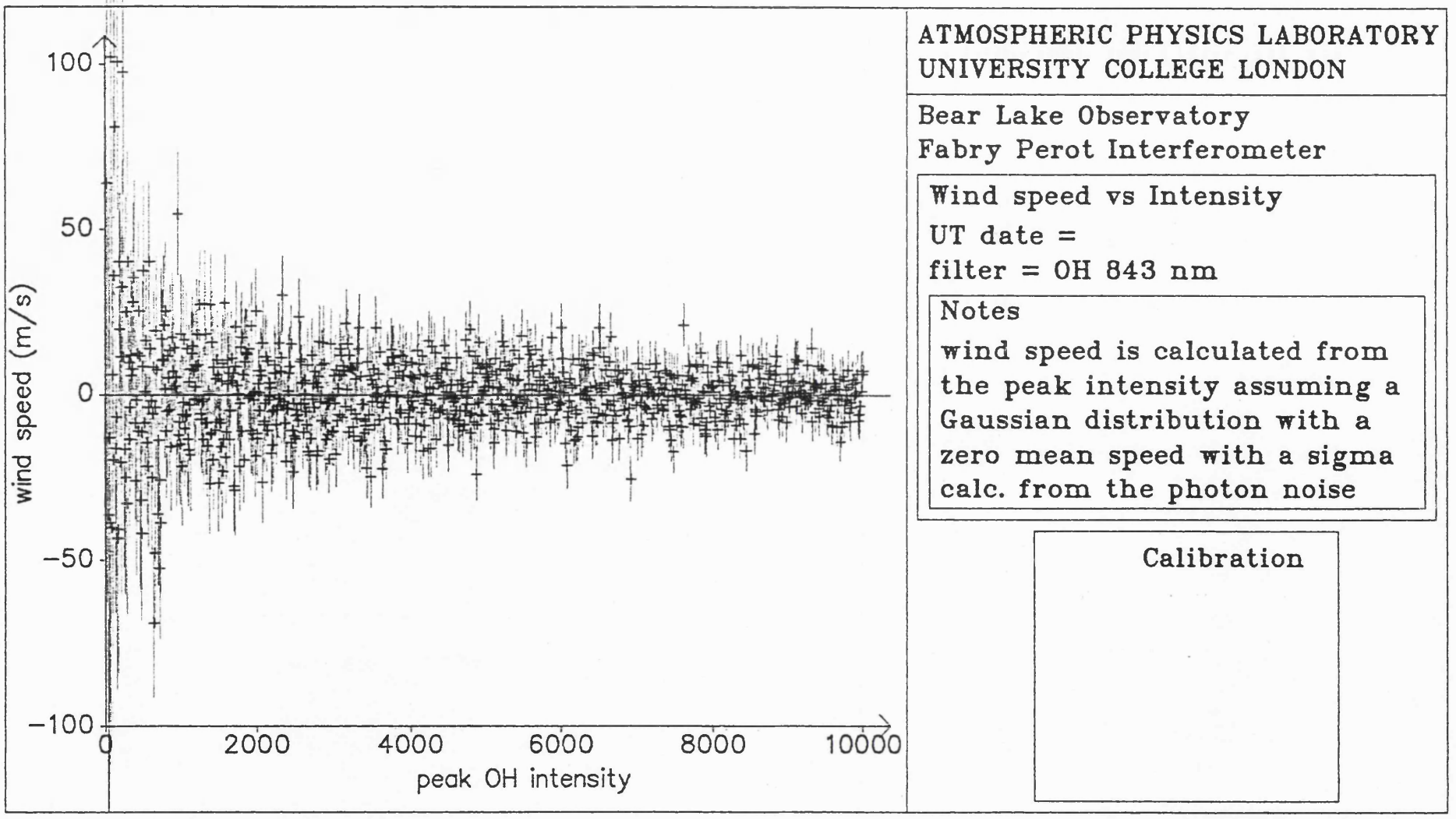

100
50
0
−50
−100
wind speed (m/s)
0
2000
4000
6000
8000
10000
peak OH intensity
ATMOSPHERIC PHYSICS LABORATORY
UNIVERSITY COLLEGE LONDON
Bear Lake Observatory
Fabry Perot Interferometer
Wind speed vs Intensity
UT date =
filter = OH 843 nm
Notes
wind speed is calculated from the peak intensity assuming a Gaussian distribution with a zero mean speed with a sigma calc. from the photon noise
Calibration

Figure 5.13

Chapter 6

6 Aeronomy of the Middle Atmosphere

6.1 Introduction

One of the most important dynamical properties of the atmosphere is its ability to support wave motions. These waves are of many different types. In this chapter, only waves that contribute significantly to the variability of the atmospheric parameters in the Middle Atmosphere will be discussed. Thus, the discussion of acoustic waves will be excluded, since these have frequencies comparable to or higher than the buoyancy frequency and violate the hydrostatic relation of Equation (1.6). Also, those waves which are modified by electrodynamic effects become important in the thermosphere, will be omitted.

Atmospheric waves can be classified in various ways, according to their physical or geometrical properties. In the first place, they can be categorized according to their restoring mechanisms: hence buoyancy, or internal gravity, waves (more commonly known as "gravity-waves") (see Section 6.2) owe their existence to stratification, while inertio-gravity waves result from the combination of stratification and the Coriolis effects. Planetary, or Rossby (see Section 6.3), waves result from the Beta-effect or, more generally, from the Northward potential vorticity gradient.

A second type of classification is to distinguish forced waves, which must continually be maintained by an excitation mechanism of given phase speed and wave-number, from free waves, which are not maintained. Examples of forced waves include thermal tides (see Section 6.4), which are induced by the diurnal variation in the solar heating, while examples of free waves include Rossby waves.

A further classification results from the fact that some waves can propagate in all directions , while others may be trapped (or evanescent) in some directions. Therefore, under certain circumstances horizontally propagating planetary waves can be trapped in the vertical, while equatorial waves can propagate vertically and zonally, but are evanescent with increasing distance from the equator (briefly discussed in Section 6.5).

Section 6.6 will discuss the data obtained from the IFPI situated at Bear Lake, Utah investigating the aforementioned wave classifications during June, July, and August of 1992.

6.2 Gravity Waves

It is now believed that the level of gravity-wave activity in particular, determines the mean state of the mesosphere. The waves manifest themselves in the wind, temperature, density, pressure, ionization, and airglow fluctuations in the 80-120 km altitude range, and the amplitudes are so large (see Section 6.2.1) that they may dominate at these altitudes. Internal

gravity-waves are also likely to be significant common contributors to the total wind field in the lower mesosphere and the stratosphere, but observations at these levels are sparse at present. Instruments carried by rockets have enabled the density and temperature structure to be measured with excellent height resolution, while FPIs have allowed direct mesospheric wind determination by measuring Doppler shifts of the near IR emission of the hydroxyl molecule. Partial reflection, meteor radars [Muller, 1970][165] and, more recently, lidar measurements of temperature and density profiles, have enabled the temporal behaviour of waves in the mesosphere to be better understood.

It was mentioned in Section 6.1 that the pure internal gravity waves owe their existence to buoyancy restoring forces, while inertio-gravity waves are due to the combined effects of buoyancy and Coriolis forces. Figures 6.1a,b identify the aforementioned "gravity-wave modes". Hence, the propagation of a gravity-wave can be described by the combination of the First Law of Thermodynamics, namely the Conservation of Energy, together with the Continuity Equation (Equation 6.1), and a modified version of the Momentum Balance Equation (Equation 6.2).

$$\nabla \rho V = - \frac{\partial \rho}{\partial t} \qquad \textbf{(6.1)}$$

$$\frac{dV}{dt} = 2V \Lambda \Omega - \frac{1}{\rho} \nabla p + g + F \qquad \textbf{(6.2)}$$

where ρ = the density of a parcel of air

V = the velocity of the parcel

Ω = the angular velocity of frame of reference

∇p = the pressure gradient

g = the acceleration due to gravity and includes the centrifugal term = $g' - \Omega \wedge (\Omega \wedge r)$

r = the radius vector

F = the frictional force

In order to obtain a reasonably simple expression describing the propagation of a gravity-wave in the atmosphere, the following assumptions must be made:-

1) ignore motion or gradients along the y direction (North-South),
2) assume that the horizontal scale of the wave is sufficiently small so that the Coriolis term may be neglected by comparison with the other terms,
3) neglect the friction term,
4) assume that the unperturbed atmosphere is at rest.

The components of the Momentum Equation (6.2) are:-

$$\frac{du}{dt} + \frac{1}{\rho}\frac{\partial p}{\partial x} = 0 \qquad \textbf{(6.3)}$$

$$\frac{dw}{dt} + \frac{1}{\rho}\frac{\partial p}{\partial z} + g = 0 \qquad \textbf{(6.4)}$$

The Continuity Equation (6.1) becomes:-

$$\frac{1}{\rho}\frac{dp}{dt} + \frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0 \qquad \textbf{(6.5)}$$

and the First Law of Thermodynamics for adiabatic motion means that the potential temperature θ remains constant:-

$$\frac{d \ln \Theta}{dt} = 0 \qquad \textbf{(6.6)}$$

The potential temperature can be expressed in terms of ρ, and p such that:-

$$\frac{d \ln \Theta}{dt} = \frac{1}{\gamma}\frac{d \ln p}{dt} - \frac{d \ln \rho}{dt} \qquad \textbf{(6.7)}$$

where γ = the ratio of the specific heats at constant pressure and constant volume

For the unperturbed atmosphere neither θ, p nor ρ vary in the horizontal direction.

Let perturbations be introduced in Equations (6.3), (6.4), (6.5) and (6.7), such that:- u = u′, w = w′, $\rho = \bar{\rho} + \rho'$, $p = \bar{p} + p'$. Utilizing the Hydrostatic Equation:-

$$\frac{\partial p}{\partial z} = - \frac{p}{H} \qquad \textbf{(6.8)}$$

and the Equation of State:-

$$\frac{\bar{p}}{\bar{\rho}} = g H \qquad \textbf{(6.9)}$$

where H = the scale height

Neglecting the products of primed quantities, the four Equations (6.3), (6.4), (6.5), and (6.7) become for the unknowns u′, w′, $\rho'/\bar{\rho}$, and $p'/\bar{p}$:-

$$\frac{\partial u'}{\partial t} + g H \frac{\partial}{\partial x} \left(\frac{p'}{\bar{p}}\right) = 0 \qquad \textbf{(6.10)}$$

$$\frac{\partial w'}{\partial t} + g \left(\frac{\rho'}{\bar{\rho}}\right) + g H \frac{\partial}{\partial z} \left(\frac{p'}{\bar{p}}\right) - g \left(\frac{p'}{\bar{p}}\right) = 0 \qquad \textbf{(6.11)}$$

$$\frac{\partial u'}{\partial x} + \frac{\partial w'}{\partial z} - \frac{w'}{H} + \frac{\partial}{\partial t} \left(\frac{\rho'}{\bar{\rho}}\right) = 0 \qquad \textbf{(6.12)}$$

$$- B w' + \frac{\partial}{\partial t} \left(\frac{\rho'}{\bar{\rho}}\right) - \frac{1}{\gamma} \frac{\partial}{\partial t} \left(\frac{p'}{\bar{p}}\right) = 0 \qquad \textbf{(6.13)}$$

In solving these equations for the unknowns, first assume an isothermal atmosphere in which H is a constant, in which case

$$B = \frac{(\gamma - 1)}{\gamma H}$$

and is also constant.

Therefore, the wave solution for each of the four unknowns is given by:-

$$\exp(\alpha z)\ \exp i(\omega t + kx + mz) \qquad \textbf{(6.14)}$$

in which the first exponential has been introduced to allow for variation of the amplitude with altitude. When m = 0, α = 1/2H, this condition has no phase variation in the vertical and are known as external waves; the most important of these are surface waves, whose energy is localized at a boundary (or discontinuity) in the atmosphere similar to surface waves on the ocean. These waves may be the result of a violent disturbance ie nuclear explosion, volcanic eruption, jetstreams, tropospheric storms. However, when m $\neq$ 0, then internal waves are generated. The following dispersion relation is deduced:-

$$m^2 = k^2 \left(\frac{\omega_B^2}{\omega^2} - 1\right) + \frac{(\omega^2 - \omega_A^2)}{c^2} \qquad \textbf{(6.15)}$$

where k = the horizontal wavenumber

m = the vertical wavenumber

c = the velocity of sound

$$\omega_B^2 = \frac{(\gamma - 1)g}{\gamma H} \qquad \textbf{(6.16)}$$

Equation (6.16) is the square of the Brunt-Vaisala frequency for

the isothermal atmosphere.

$$\omega_A = \frac{1}{2}\sqrt{\left(\frac{\gamma g}{H}\right)} = \frac{c}{2H} \qquad (6.17)$$

is known as the acoustic cut-off frequency.

The dispersion relation (Equation (6.15) is illustrated in Figure 6.2, where the curves of m = 0 are plotted on a ω-k diagram. The two regions where m^2 is positive is apparent, the higher frequency region ($\omega > \omega_A$ when k = 0) describes acoustic waves, and the lower frequency region, gravity-waves.

If the atmosphere is now assumed to be non-isothermal, Equation (6.15) may still be employed as a reasonable approximation provided that $\omega_B < \omega_A$ as is normally the case throughout the atmosphere.

6.2.1 A Simple Model of Breaking Gravity Waves

Owing to the presence of the exp (z/2H) factor in Equation (6.14), the linear, non-dissipative theory predicts that the velocity and geopotential disturbances will grow with altitude. At some height the non-linear terms that have been neglected will become important, and thus the linear theory will break down.

A physical picture of this break-down can be understood by

considering a set of surface fluctuations at various levels, which are undulating as an internal gravity-wave propagates vertically through them. Figure 6.3 is a schematic diagram of this situation. In the lower mesosphere, the surface fluctuation (a) has a gentle sinusoidal variation, as predicted by linear theory. For gravity-waves of periods much less than a day, the effects of the radiative relaxation are small, and in the absence of other diabatic processes, (a) is an isentrope (a surface of constant θ). In the middle mesosphere, the surface fluctuation (b) is also sinusoidal, but the gravity-wave has a larger amplitude than (a); linear theory still holds, and (b) can also be taken as an isentrope. In the upper mesosphere, however, non-linear effects become important, leading to the rapid and irreversible deformation of surface contours (c), followed by turbulence, small-scale mixing, and dissipation.

The process described here is known as gravity-wave breaking [Garcia and Soloman, 1985][114], by analogy with the overturning and breaking of oceanic surface waves on a shelving beach. It also has points in common with the phenomenon of planetary wave breaking.

6.3 Planetary Waves

A well documented group of atmospheric waves is a class of zonally travelling structures of global scale, which have periods of a few days. Such waves are thought to be examples of free travelling planetary (Rossby) waves, or global normal modes and,

unlike thermally driven tides described in Section 6.4, they are apparently not maintained by travelling forcing effects.

Tropospherically forced planetary waves are confined primarily to the winter hemisphere. Satellite observations show that the waves reach their maximum amplitudes at or just above the stratopause, with maximum temperature fluctuations of the order 6 K reached at latitudes of 60° [Barnett, 1980][166]. Under favourable conditions in the Northern Hemisphere, planetary waves with periods between 5- and 30-days have been observed to penetrate up to altitudes of 110 km, producing oscillations with vertical amplitudes of between 2 and 10 km [Brown and John, 1979; Cavalieri **et al**., 1974; Fraser, 1977][167,168,169]. However, during the summer, the forced planetary waves are excluded from the Upper Atmosphere by the stratospheric westward winds (easterlies). However, it is possible for planetary scale transient motions, associated with normal modes or with baroclinic instabilities of the stratospheric jets to be observed in the mesosphere. Spectral analyses of extensive data sets suggest that the transient wave energy tends to maximize in the local summer [Vincent, 1984][170]. A range of wave periods has been identified, but the most commonly reported periods fall into three well defined intervals which are illustrated in Table 4.1.

Often, these intervals are referred to as the "16-day", "5-day", and "2-day" sinusoidal oscillations respectively, although precise determination of these periods involved is often not possible. Comparative studies made at different longitudes,

indicate the waves are westward travelling, with the best determinations of wavenumber being made for the 2- and 5-day waves. The inadequate period lengths of many data sets has restricted the "16-day" wave studies, although this situation will improve as continuous radar and lidar (ALOMAR, see Chapter 7) observations become more freely available.

It is usually assumed that these transient oscillations are caused by Rossby-gravity normal modes generated in the lower atmosphere. These modes are normally evanescent (trapped) in the vertical. However, in altitude regions with westward winds, and equatorward temperature gradients, the waves can be locally propagating in the vertical direction.

Each mode has a well defined structure with respect to latitude and longitude, however, in practice, a positive identification has not always been possible from observations made in the 80-120 km region, because of the inadequate geographic coverage. Hirota **et al.**, [1983][171] reported a correlation between a 5-day wind oscillation in the mesosphere with a westward travelling wavenumber-one wave in the stratosphere. Satellite radiance studies have shown evidence for both 2- and 5-day waves with temperature amplitudes tending to maximize in the lower mesosphere [Rodgers and Prata, 1983][172] with values of 0.5-0.8 K. Chapter 4 discusses further the recent temperature oscillations obtained by the Bomem Interferometer in 1991/2.

The most extensively studied oscillation is the 2-day wave, and

its seasonal and spatial behaviour is now well established. The wave is usually observed in the summer hemisphere, reaching maximum amplitudes at ~90 km in July/August in the Northern Hemisphere [Rodgers and Prata, 1983][172], and in January in the Southern Hemisphere [Craig and Elford, 1981][173]. Both radar wind measurements, and satellite observations show that the wave amplitudes are maximum in the Southern Hemisphere [Vincent, 1984; Rodgers and Prata, 1983][170,172]. A hemispheric difference in the wave period has been noted by Vincent [1984][170]. Northern Hemisphere observations give periods near 51 hours [Craig **et al.**, 1983][174], while periods nearer to 48 hours have been reported in the Southern Hemisphere by Craig and Elford [1981][173].

6.4 Atmospheric Thermal Tides

Atmospheric tides are global-scale daily oscillations, which are primarily forced by diurnal variations of the heating due to the absorption of the solar ultra-violet radiation by atmospheric water vapour and ozone. The solar and lunar gravitational forcing that produces ocean tides is much less important for the atmosphere, and emphasis will be placed on thermally driven tides. Further, attention will be restricted to migrating tides, which move with the Sun. Non-migrating tides are associated, for example, with topography and geographically fixed tropospheric heat sources. Migrating tides have been studied extensively; they can propagate through great depths of the atmosphere and can attain large amplitudes at some heights, especially the thermosphere (see Figure 1.2). Even in the mesosphere, tidal

temperature amplitudes can easily exceed 10 K. Hence, satellite data can be mis-interpreted when sampled data only occurs at a finite solar time.

6.4.1 Summary of the Main Results of Atmospheric Tidal Theory

The basic ideas of the tidal theory are quite simple, although some of the mathematical details are a little complicated.

The first point to notice, is that the solar heating is only active during the day, with a time variation similar to that illustrated in Figure 6.4. The Fourier analysis of this curve will include a steady component, a diurnal component (with a twenty-four hour period), a somewhat smaller semi-diurnal component (with a twelve hour period), and higher order harmonics. The response of the atmosphere to this heating can likewise be decomposed into a steady part, diurnal, semi-diurnal and higher-frequency oscillations. Kelvin in 1882 raised the question of why the semi-diurnal surface pressure oscillation (~1 mb in amplitude) [Lindzen, 1967][175] is larger and more regular than the diurnal. The answer is best explained by describing the response of the whole Middle and Lower Atmosphere to the diurnal and the semi-diurnal heating, as currently understood from theory and observation.

It transpires that the forced semi-diurnal tidal oscillation has a large vertical wavelength greater than 100 km. This allows it to be excited efficiently by the deep ozone heating region

present in the stratosphere [Lindzen, 1966][176]. Moreover, it can easily propagate to the ground, and hence cause a surface pressure fluctuations. It is regular, because the ozone heating is regular, and its latitudinal structure is fairly uniform.

The diurnal tide, on the other hand, has a more complex behaviour. Between 30°N and 30°S it can propagate vertically, with a wavelength of about 28 km [Butler and Small, 1963; Leovy, 1964][177,178]. Polewards of 30°, however, it is trapped in the vertical, close to its forcing region. As a result, the vertically propagating modes forced within the deep region of ozone heating tend to interfere destructively, and thus have small amplitudes at the ground, while the trapped modes forced by ozone never reach the ground at all. However, the comparatively shallow water vapour heating region in the troposphere [Siebert, 1961][179] can excite this tide quite effectively, although this heating is somewhat intermittent in space and time. The resulting surface pressure oscillation is also small and intermittent.

6.4.2 Classical Tidal Theory

The structure of the atmospheric tides has been modelled using "classical" tidal theory which involves the solution of linearized equations of continuity, momentum, and energy modified by the time dependent solar heating [Siebert, 1961; Chapman and Lindzen, 1970][179,180]. The basic Laplace equation [Lindzen, 1967][175] is used with suitable boundary conditions. The classical tidal

theory assumes that the mechanical forcing and dissipation are assumed to be zero, and the restriction to migrating tides means that all variables are taken to depend only on the local time, latitude, and height. Thus, for example, $J' = J'(\lambda + \Omega t, \Phi, z)$, since $t + \lambda/\Omega$ is the local time at longitude λ. J' denotes the derivation of J from the zonal mean net heating $\bar{J}$. As seen from Figure 6.4, the zonal mean solar heating is non zero; however, it will be assumed to be balanced by a zonal mean infra-red cooling, so that $\bar{J} = 0$.

Since J' is periodic, it can be expanded in Fourier harmonics:-

$$J' = Re \sum_{s=1}^{\infty} J^s (\Phi, z) \exp is(\lambda + \Omega t) \qquad \textbf{(6.18)}$$

where the coefficients J^s are assumed known from a radiative-photochemical calculation. This may not be a very accurate assumption: for example, the ozone distribution, which partly determines the heating, will depend to some extent on the tidal response to the heating. The s = 1 component is the diurnal heating, and corresponds to $\sigma = -1/2$; the s = 2 component is the semi-diurnal heating and corresponds to $\sigma = -1$; higher terms are seldom considered. The numbers s, and σ, according to the notation of Flattery [1967][181], represent:-

s = the zonal wavenumber

σ = the meridional wavenumber

The eigenfunction solutions are commonly known as the Hough functions. These functions θ^s_σ describes the latitudinal structure, and the corresponding eigenvalues h^s_σ, also known as the equivalent "depth", give the vertical structure for a given background temperature distribution. Positive or negative values of σ denote the gravitational or rotational modes respectively. Basically, s gives the number of cycles per day, and the number of nodes between the poles is given by $(s - \sigma)$, with the exception of the nodes at the poles. For example, a (2,4) mode has a period of twelve hours and two nodes $(s - \sigma = -2)$. According to the season, the solar heating is greatest at the equator, therefore modes with an antinode at the equator, rather than a node, are preferentially excited [Hargreaves, 1979][182]. The most important Hough functions are illustrated in Figure 6.5.

However, the classical tidal theory is a broad general approximation. In reality, the assumptions break down since solving the tidal equation of motion analytically is no longer viable; hence, several "non-classical" models have been developed in order to solve the equations numerically.

On ignoring the aforementioned model assumptions, the following consequences result:-

I) molecular viscosity, thermal conductivity, and ion drag become more important at altitudes above 100 km. The inclusion of the dissipative terms results in height dependent Hough functions, known as the Hough mode extensions [Forbes and Hagan,

1982][183]. The exponential growth of propagating modes is slowed down by dissipation, thus, the amplitudes and phases of the winds and temperatures reach constant values asymptotically at an altitude range of 150-200 km, or even decrease [Hong and Lindzen, 1976; Forbes, 1982a; 1982b; 1987b][184,185,186,187].

II) Lindzen and Hong [1974][188] indicated that the inclusion of the mechanical forcing and temperature gradients results in the generation of the higher order modes achieved through "mode coupling". It has been proposed that this may be the dominant mechanism in the generation of the higher order modes [Walterscheid **et al.**, 1980; Aso **et al.**, 1981; Fesen **et al.**, 1986][189,190,191].

Often studies neglect the diurnal modes, because they are complex to model for the following two reasons:-

i) the complexities of turbulent diffusion cannot be overlooked for this mode,

ii) the shorter vertical wavelengths require a higher vertical resolution than most contemporary models can facilitate [Forbes,1987a][192].

Forbes and Hagan [1979][193] overcame ii) by using an "Equivalent Wave f-Plane Formalism" to simulate the (1,1) mode. It was found that the (1,1) mode appeared dominant at low latitudes below a height of 100 km where the meridional temperature gradients are

weak, and the mean wind flow is small compared with the phase velocity. Under these conditions, the phase coupling is small, and the assumptions of the classical tidal theory hold well. Moreover, the (1,1) mode was found to have an exponential growth in amplitude and phase progression, achieving a peak amplitude at 105-110 km. However, above 110 km, the diurnal tide is heavily damped as a result of the increase in the molecular diffusion coefficient in the upper thermosphere, and the absence of the tidal heat and momentum sources which can coherently excite this mode due to its short vertical wavelength [Forbes, 1982a; 1982b; 1987a; Vial, 1986][185,186,192,194].

Forbes [1987a][192] suggested that the diurnal tidal oscillations observed between 90-150 km are generated equally by upward propagating waves, and by in-situ solar EUV excitation. In propagating modes, the phase varies with altitude, hence a vertical wavelength, λ_z, exists. λ_z is defined as the distance over which the phase changes by a factor 2π. However, if λ_z is less than or equal to the depth of the atmosphere where the majority of the heating occurs, then the wave will suffer from self destruction, and therefore become weak. This confirmed the argument proposed by Butler and Small [1963][177]. In the Forbes model, for the diurnal tide, $\lambda_z \leq 30$ km.

Equation (6.19) represents the vertical wavelength [Wilkes, 1949; Chapman and Lindzen, 1970][195,180]:-

$$\lambda_z = \frac{2\,\pi\,H}{\sqrt{\left(\frac{H}{h}\left[\frac{(\gamma - 1)}{\gamma} + \frac{dH}{dz}\right] - \frac{1}{4}\right)}} \qquad \textbf{(6.19)}$$

where H = the scale height at a height z

h = h^s_σ = the equivalent depth of the tidal mode

γ = the adiabatic index = ratio of the specific heats at constant pressure and constant volume

From theoretical calculations, it was proposed [Vial, 1986][194] that the seasonal variation in the latitudinal structure of the wind amplitudes and phases is due to the anti-symmetric modes (2,3) and (2,5). Forbes [1987a][192] challenged Vial, as he doubted that the (2,5) mode would be so dominant in the real atmosphere. Instead, the latitudinal variation in the vertical structure of the diurnal tidal winds, and temperatures may result from superposition of the diurnal (1,1) mode, and an evanescent mode of the form (1,-2). In the Forbes model [1987a][192], the evanescent mode is apparent at high latitudes at heights below 100 km, where there is no amplitude nor phase progression with altitude.

As the waves propagate upwards, no damping is experienced. The amplitudes become larger in order to conserve the energy density as the particle density decreases. In order to maintain a constant energy density of $(\rho u^2)/2$, the amplitude of the air velocity, u, must vary proportionally with $\sqrt{\rho}$, where ρ is the air density. Eventually, the amplitudes would become too large, and

unstable, or heavily damped. Therefore, the waves will break [Garcia and Soloman, 1985][114]. Instability results when the wave amplitudes are approximately equal to the speed of sound, or when the wave phase velocity approximates to the background wind speed [Lindzen, 1970][196]. This is the general behaviour of waves generated in the troposphere, that propagate upwards and break in the mesosphere and lower thermosphere (see Section 6.2.1) releasing energy as heat. However, the horizontal structure of the semi-diurnal modes with a vertical wavelength greater than 40 km are not significantly affected by this dissipation below ~100-110 km [Forbes and Hagan, 1982][183]. Moreover, the damping due to the molecular or eddy dissipation, may occur at heights between 90-100 km, but the horizontal shape remains unaltered [Lindzen, 1971][197].

Finally, the influence of the semi-diurnal tides in the Middle Atmosphere are characterized by at least four semi-diurnal modes:- (2,2), (2,3), (2,4), and (2,5), whose relative amplitudes and phases vary with the height and latitude. The following table summarizes the various characteristics of the semi-diurnal modes.

Table 6.2

Dominating Mode	Height /km	Description
(2,2)	< 60	Preferentially excited by H_2O and O_3.
(2,2)	50 < h < 70	Exponential growth checked because of background temperature causes this tide to become trapped and also, this mode couples into higher order modes ie (2,3), (2,4), (2,5), thus losing energy.
(2,2),(2,3), (2,4)	70 < h < 90	Modes join together in equal quantities.
(2,4)	90 → 120	May have contributions from (2,2) and (2,5) modes.
(2,4),(2,5)	> 120	Due to their short wavelength, these modes are damped out, because molecular dissipation is exponentially increasing with altitude.
(2,2)	> 140	This mode is dominant again with addition contributions from (2,4) and (2,5) modes.

6.5 Equatorial Waves

It has been known for many years that large-scale, equatorially confined wave motions propagate vertically and zonally through the Middle Atmosphere. These waves have periods of a few days, and are of planetary scale in the zonal directions, but are trapped within about 15°N and 15°S of the equator. The earliest observations of such waves were from radiosondes in the lower stratosphere. More recently, rocketsondes and satellites have detected them in the upper stratosphere and mesosphere as well. Waves of this kind are particularly significant for the dynamics of the Middle Atmosphere, since they are believed to play a

central role in forcing the equatorial quasi-biennial oscillation (QBO), and semi-annual oscillation (SAO).

6.6 Results from the Utah IFPI

The IFPI installed at the Centre for Atmospheric and Space Science (CASS), Utah State University (USU) was built and developed at UCL. The present collaboration between UCL and CASS is to make continuous, long term observations of the mesosphere using an automated instrument.

This objective is achieved by employing the IFPI to measure the Doppler shifts of the infra-red emission from the OH (6,2) Meinel band at a wavelength of 8430 Å. This emission is part of the Meinel band, which is composed of the rotational-vibrational bands of the excited hydroxyl radical, OH^* [Meinel ,1950][198]. These emissions are a dominant feature of the night airglow [Krasovskij and Shefov, 1965][199]. Rocket- and satellite-borne instrumentation have determined that the emission region is centred near 87 km which corresponds to the upper mesosphere, covering a region of 6 km Full Width Half Maximum (FWHM) [Baker and Stair, 1988][131].

6.6.1 Data Interpretation

The data collected between 1st of June 1992 and 31st of August 1992 were initially screened as discussed in Chapter 5. Out of a possible 92 nights, 14 nights were finally selected when the

instrument was fully functional and the skies were clear with no deleterious contamination from moonlight.

On these nights, the OH intensities varied from approximately 500 to 2500 photon counts per bin. Some nights the OH intensities remained relatively constant, whilst on others, large fluctuations were noted. The OH intensities are reported in Section 6.6.1.1.

The mesospheric winds recorded at the end of August were consistent with a semi-diurnal tidal mode. However, those measured in June and July showed no such tidal mode. The wind measurements are discussed in Section 6.6.1.2.

The mesospheric winds determined from the Utah IFPI were compared with those derived from the Sheffield meteor wind radar. The results from this comparison are presented in Section 6.6.1.3.

Section 6.6.1.4 addresses the important points raised in the previous sections.

6.6.1.1 Summer Mesospheric Hydroxyl Intensities as Determined from the BLO IFPI

The OH intensities, as determined from the height of the OH peaks in the reduced profiles, vary from 500 to 2500 photon counts per bin (pcpb), with the majority of the results lying predominantly in the range 1000 to 1500 pcpb.

Sample results from the 14 good observing nights are shown in Figures 6.6 to 6.12. Two intensity plots are presented for each night. On the first plot, the results from all nine viewing directions, together with the calibration, are displayed as a function of time on separate graphs. On the second plot, the results from the eight viewing directions at a zenith angle of 60° are all displayed as a function of time on the same graph.

No consistent trends are immediately apparent in the OH intensity profiles. For example the nights of the 24th and 25th of August show possible evidence of a semi-diurnal tidal variation, the nights of 28th and 29th of August show a similar decreasing trend throughout the night which could still be consistent with the semi-diurnal tidal variation, whereas the nights of the 26th and 27th of August have generally increasing intensities throughout the night and are not consistent with such a semi-diurnal tidal variation.

The night of the 5th of June 1992 is interesting and clearly anomalous. The OH intensity (see Figure 6.6) is relatively high for most of the night (1500 - 2000 pcpb). A dramatic drop in intensity is observed around 05:55 UT in all directions. The intensity maintains a low value of approximately 500 pcpb for ~30 minutes, and then with dramatic rapidity returns to the range 1500 - 2000 pcpb. A second smaller minimum of ~1000 pcpb is observed two hours later at 07:55 UT. Although this minimum is not so well defined as the first it is quite clearly a perturbation on the mean value.

6.6.1.2 Summer Mesospheric Hydroxyl Winds as Determined from the BLO IFPI

Sample results from the 14 good observing nights are shown in Figures 6.6 to 6.12. The horizontal wind speeds occasionally reach 100 ms^{-1}, but the majority of the wind speeds are 50 ms^{-1} or less.

Two wind plots are presented for each night. On the first the horizontal wind speeds from all nine viewing directions, together with the calibration peak positions and respective errors, are plotted as a function of time on separate graphs. On the second the results from the eight viewing directions at a zenith angle of 60° are all plotted as a function of time on the same graph.

The wind profiles collected from the six consecutive observing nights from 24th to 29th of August 1992 are self-consistent, in that the velocity components determined from opposite viewing directions are approximately equal and opposite. Interestingly, the results display a similar behaviour each night and are indicative of a semi-diurnal tidal wind variation. Although some degree of variability is evident from night-to-night, the average profiles clearly illustrate the general pattern (see Figure 6.13).

Semi-diurnal tidal modes may be represented by a simple harmonic dial consisting of a wind vector rotating in a clockwise direction, having a period of 12 hours (see Figure 6.14). The

velocity component in any given viewing direction can then be determined by resolving the wind vector along that direction.

The general pattern for the end of August is consistent with a semi-diurnal tidal mode with wind vector rotating from south-eastward at 04:00 UT to north-westward at 10:00 UT.

No such tidal variation is apparent in the wind profiles available for the 3 good nights in June and the 4 good nights in July.

The night of the 5th of June is anomalous in the wind profiles as well as the intensity profiles as discussed in Section 6.6.1.1. The majority of the winds are less than 20 ms^{-1}. However, large and divergent winds are observed between 05:50 UT and 06:20 UT, corresponding to the period of low intensity.

6.6.1.3 Comparison of the Winds Derived from the Utah IFPI with those Determined from the Sheffield Meteor Wind Radar

The observations were made using the Sheffield meteor wind radar, located at (53.4°N, 1.5°W). The aerial system used in this experiment consists of two identical 7-element Yagi arrays pointing in the directions NW and SW [Muller, 1970][165].

Some typical wind profiles from the meteor wind radar are shown in Figures 6.15a and b. These results clearly show the presence of a semi-diurnal tide, with a "2-day" and "5-day" tide

superposed (see Section 6.3).

For the semi-diurnal tidal mode observed by the Sheffield instrument, the wind vector points northward at approximately 05:00 LT (and again at 17:00 LT), and rotates at a uniform rate in a clockwise direction completing one rotation in 12 hours.

In order to compare the Sheffield results with those from the Utah IFPI, it should be noted that the Sheffield results are plotted as a function of Local Time (LT) whereas the Utah IFPI results are plotted as a function of Universal Time (UT). The semi-diurnal tide is expected to be a function of the real solar time. Since BLO is located at 111.4° W, and the Earth rotates 15° every hour, there is a 7 hr 25 min difference between the two observing sites. Thus, for consistent phase a semi-diurnal tidal feature observed at time T UT from BLO should be apparent at time T - ΔT UT from the UK, where ΔT = 7 hr 25 min. In the summer, UT is one hour behind LT (BST) so that the observations can be compared directly by subtracting 6 hr 25 min from the Utah UT time.

Therefore, during the Utah night (04:00 UT - 10:00 UT), the wind vector associated with the semi-diurnal tidal mode recorded by the Sheffield meteor wind radar, rotates from south-eastward to north-westward, in good agreement with the results obtained from the end of August from the Utah IFPI.

6.6.1.4 Discussion

The IFPI data collected from the six consecutive nights at the end of August appear to be consistent with a semi-diurnal tidal mode with wind vector rotating from south-eastward to north-westward as anticipated from the Sheffield study. These results are also consistent with those obtained from the BLO IFPI, immediately after its installation, in September 1989 [Aruliah, 1991][162]. Due to unforeseen instrumental difficulties, the IFPI was not operational during the summers of 1990 and 1991, and hence there are no further data sets for comparison.

The Sheffield results clearly show the presence of a semi-diurnal tide, with a "2-day" and "5-day" tide superposed, during the months of June, July and August. Although the IFPI results show evidence of the same semi-diurnal tidal mode at the end of August, no such mode is present in the June or July data.

The IFPI records winds from a region of the atmosphere centred at 87 km, whereas the Sheffield meteor wind radar measures winds centred at 95 km. It is therefore possible that the two regions may experience different tidal modes. For example, in the height range 70-90 km near the solstices, the asymmetric (2,3) mode is present in approximately equal contributions as the symmetric (2,2) and (2,4) modes [Kato, 1980][200] (see Table 6.1), so this mode will interfere with the symmetric tidal modes to produce a wind variation showing no distinct tidal wind mode. This could explain the lack of any obvious semi-diurnal tidal modes in the

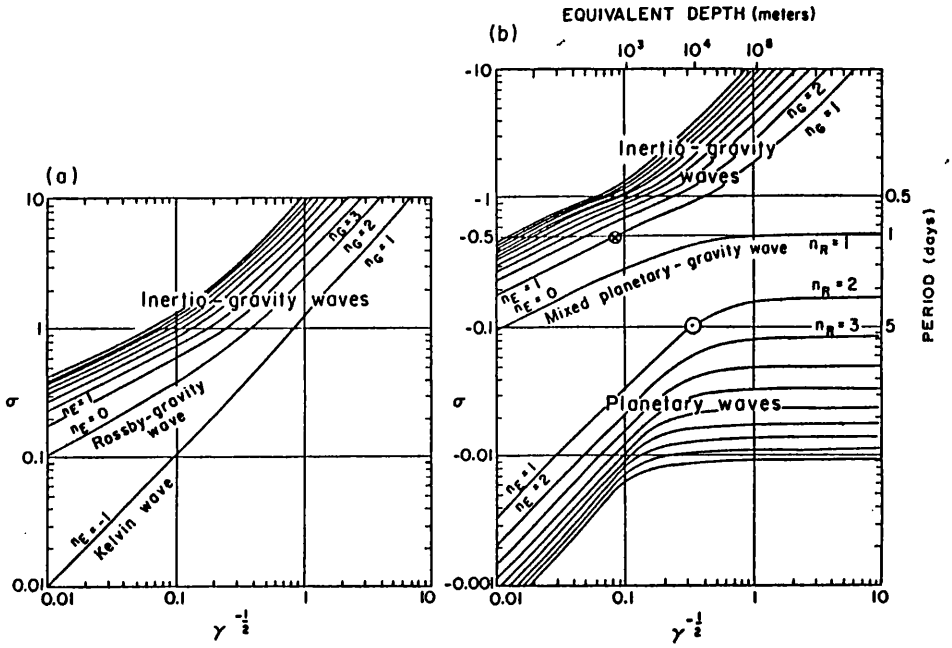
June and July IFPI data. Near the equinoxes, however, the asymmetric tidal modes are not excited as strongly, which could explain why a semi-diurnal tidal variation is then observed in August and September. The height region explored by the Sheffield meteor wind radar is dominated by the (2,2) mode at all seasons (Muller, personal communication), which could explain the consistency of the Sheffield results, and the disparity with the IFPI results in June and July. Further, more detailed analyses including subsequent years will be required before these conclusions can be truly validated.

It should be noted that the Sheffield meteor radar data appears highly regular as a result of temporal and spatial smoothing, and consequently, does not fully map the dynamic vertical and time variations, between 77 km and 120 km. The results collected from the IFPI are not smoothed and are more indicative of the "real" atmospheric conditions at any given time.

The night of 5th of June is anomalous in that the winds appear to be large and diverging between 05:50 UT and 06:20 UT, corresponding to the period of low intensity discussed in Section 6.6.1.1. Such a dramatic feature did not repeat itself again during the limited number of nights of high quality data available. This striking display has been tentatively ascribed to an upward propagating gravity wave. H.Muller (personal communication) gave further confirmation that this feature observed on the 5th of June could possibly be the result of a propagating gravity wave. However, a larger data set with other

complementary data sets are required before such a conclusion can be truly verified. If such a feature is "real", then similar features would be expected to be seen from the summer 1993 BLO data.

The intensity resolution is not adequate enough to determine whether Bands and/or Billows structures such as those described in Section 1.6.6 can be observed in the OH intensity profiles. However, it is hoped that an all-sky camera with a OH filter (see Chapter 7) will eventually be installed and operated in order to verify cloud-free nights and whether OH structures are present during any period of observation.



Figures 6.1a and b

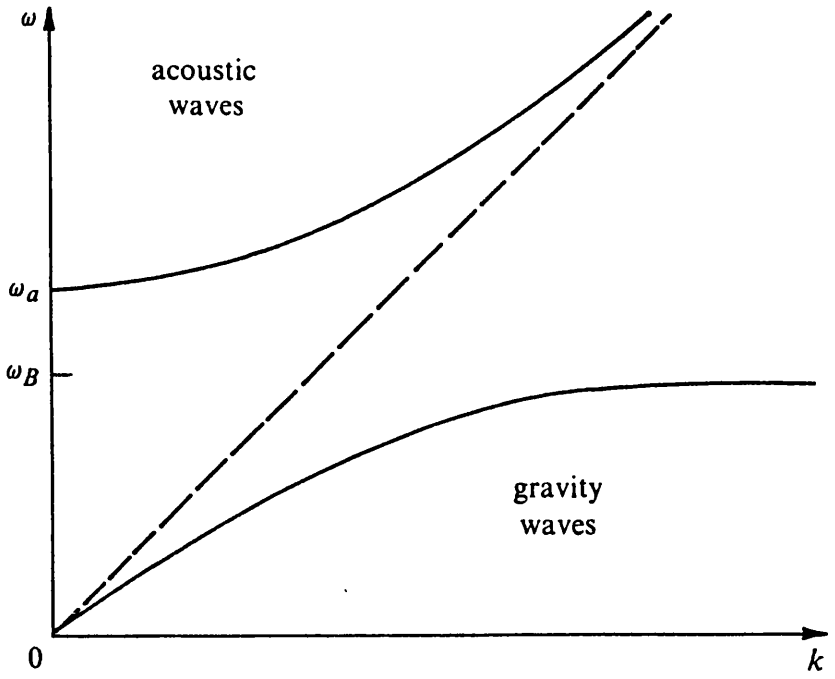


Figure 6.2

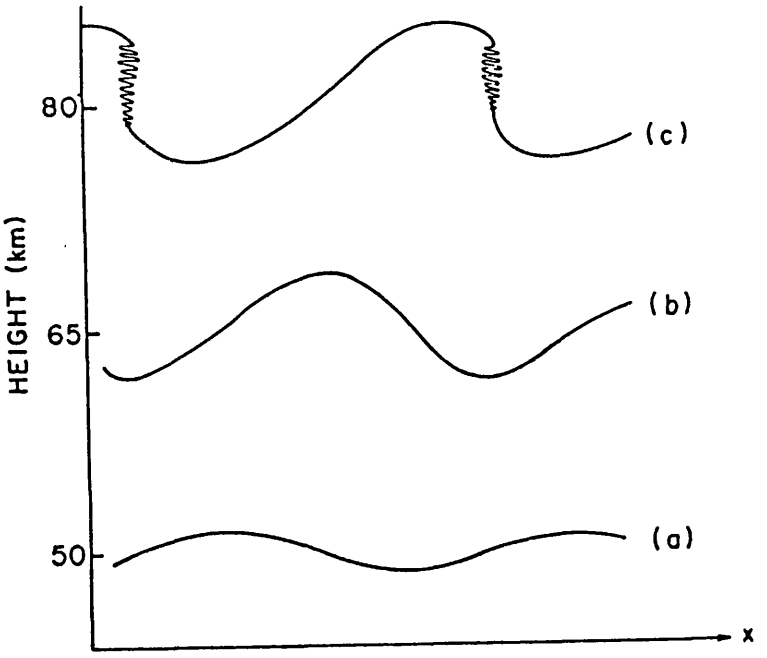


Figure 6.3

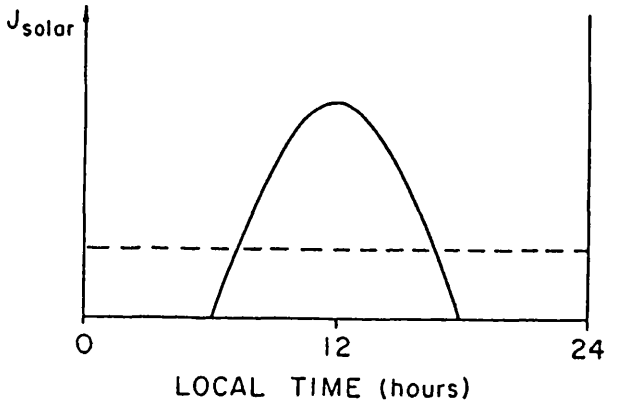


Figure 6.4

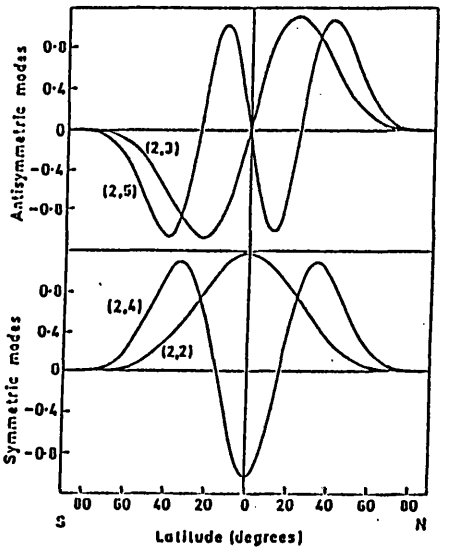
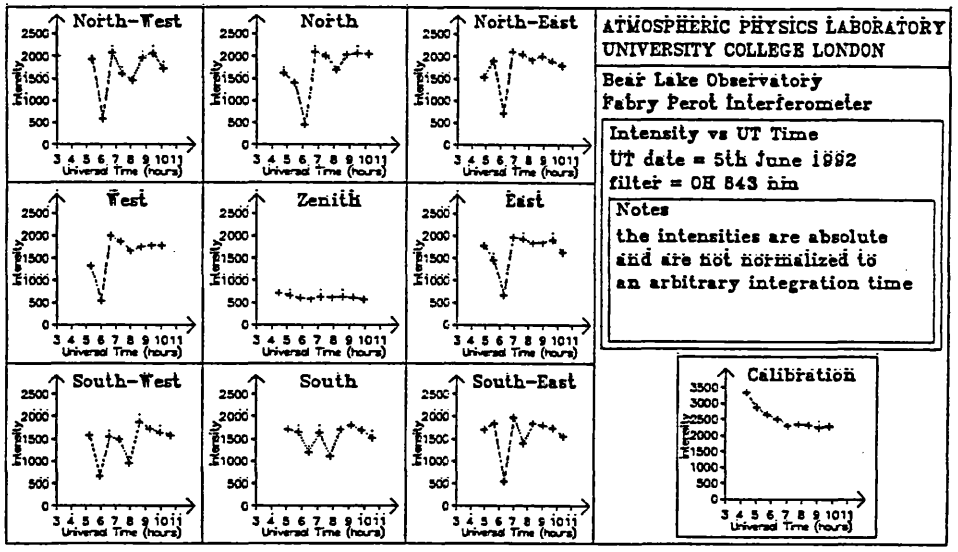
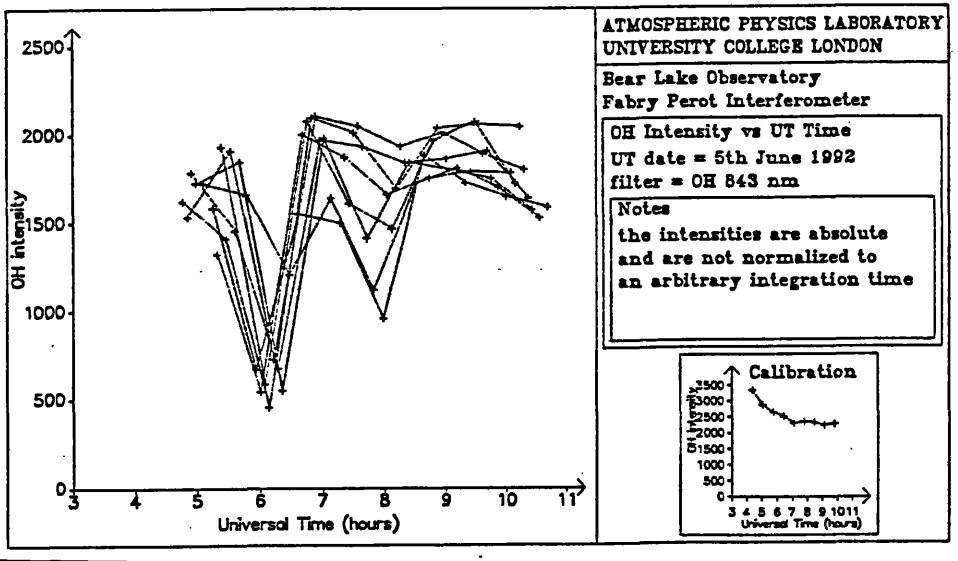


Figure 6.5

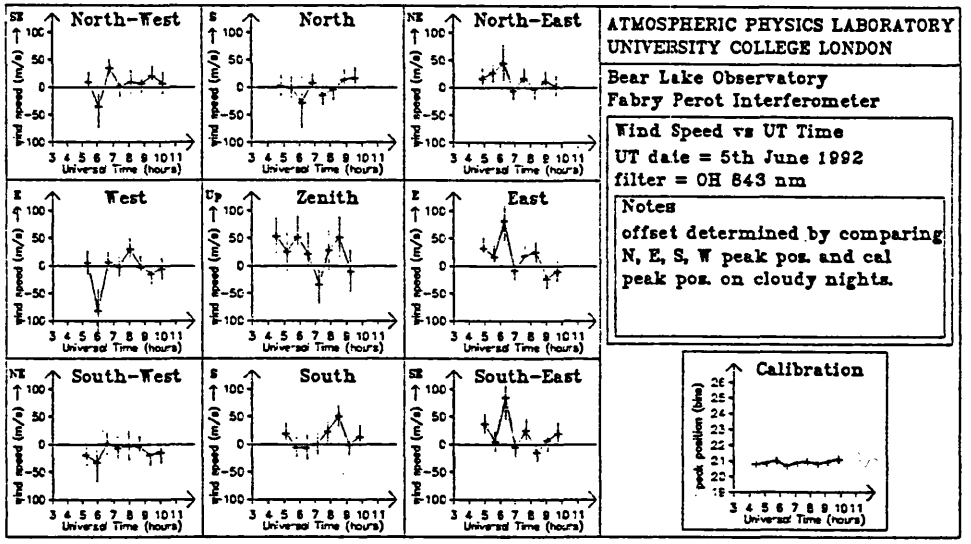
Figures 6.6a, b, c, and d



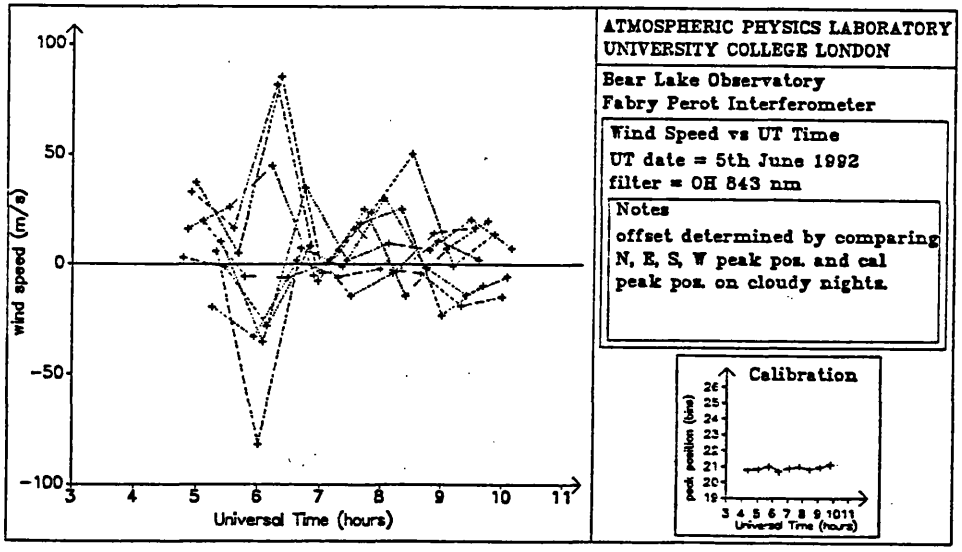
a



b

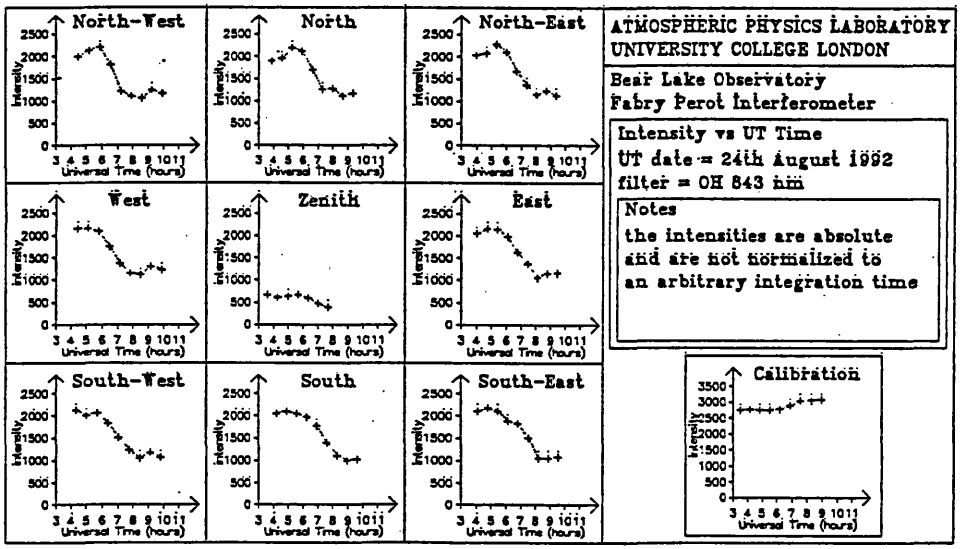


c

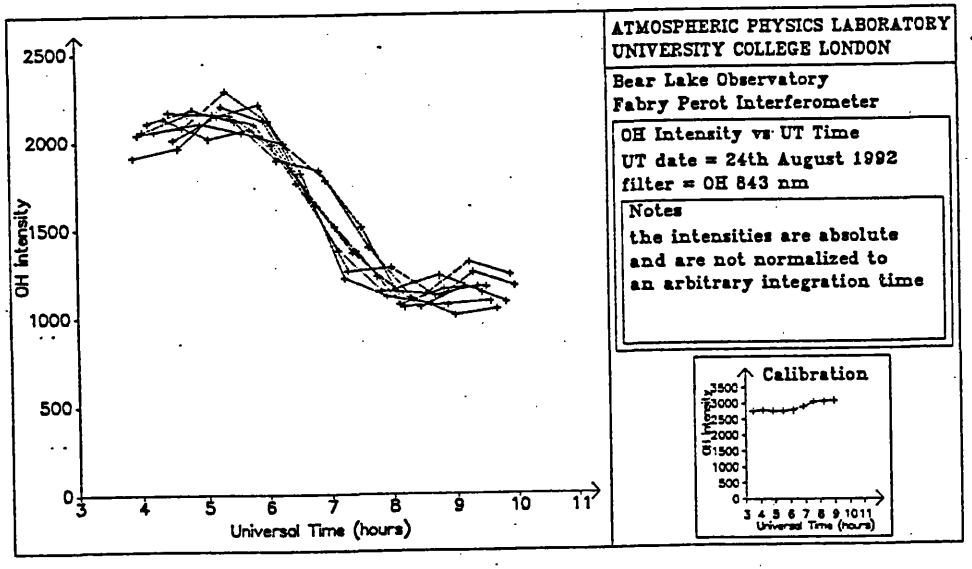


d

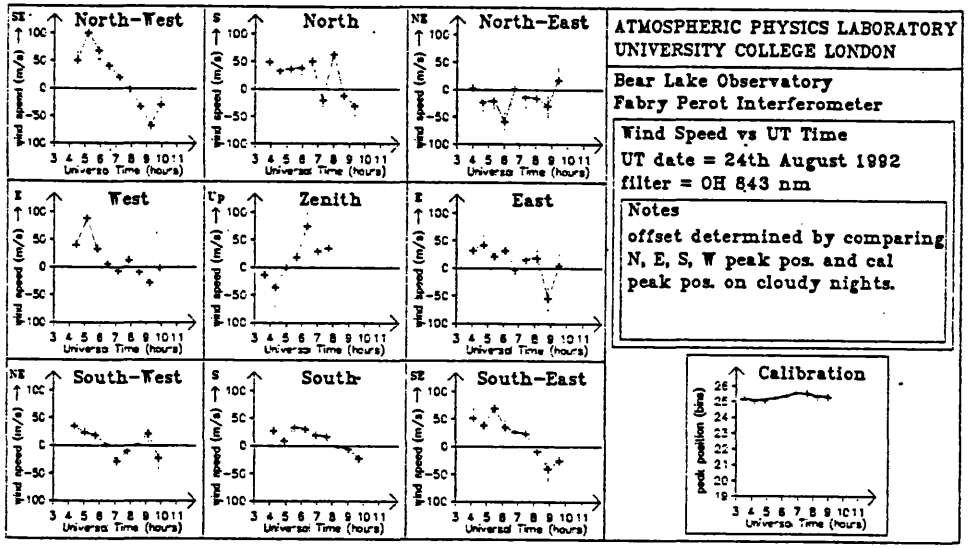
Figures 6.7a, b, c, and d



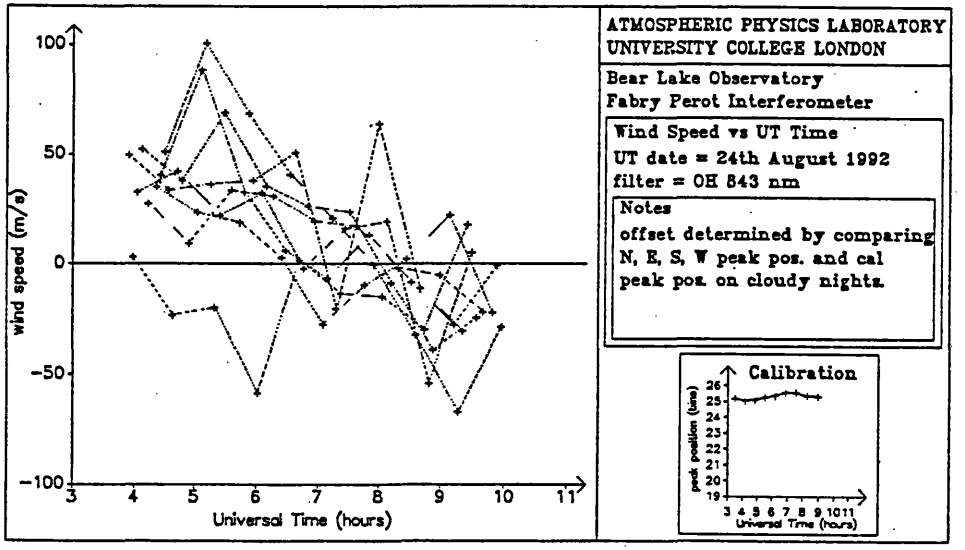
a



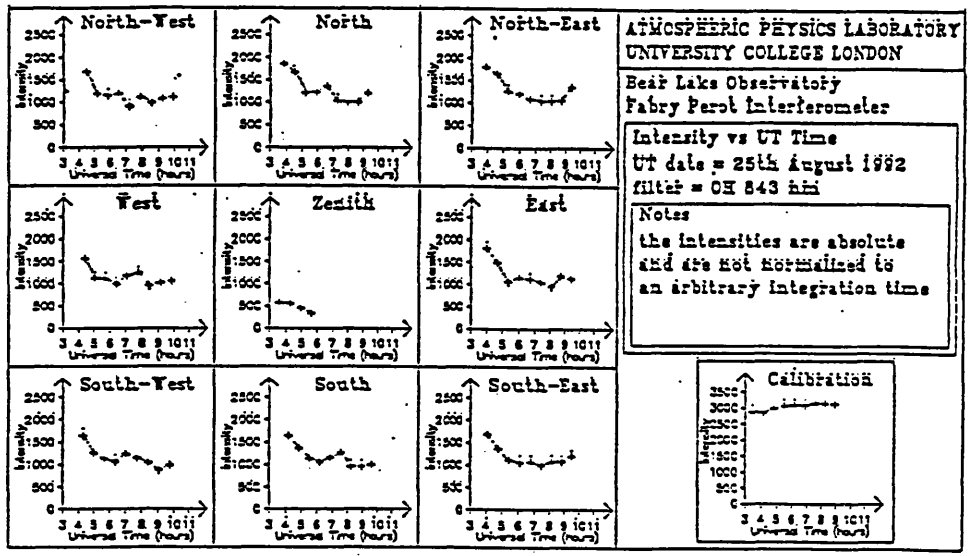
b



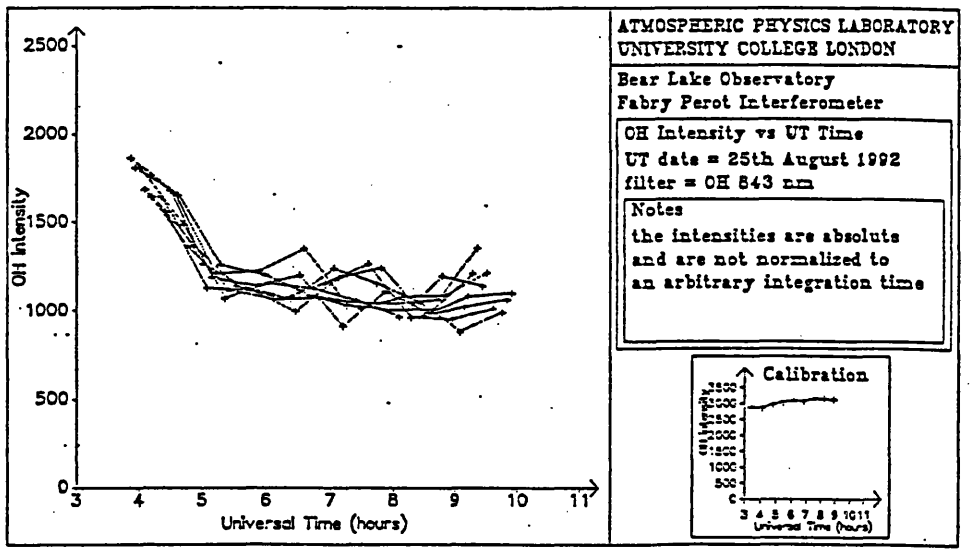
c



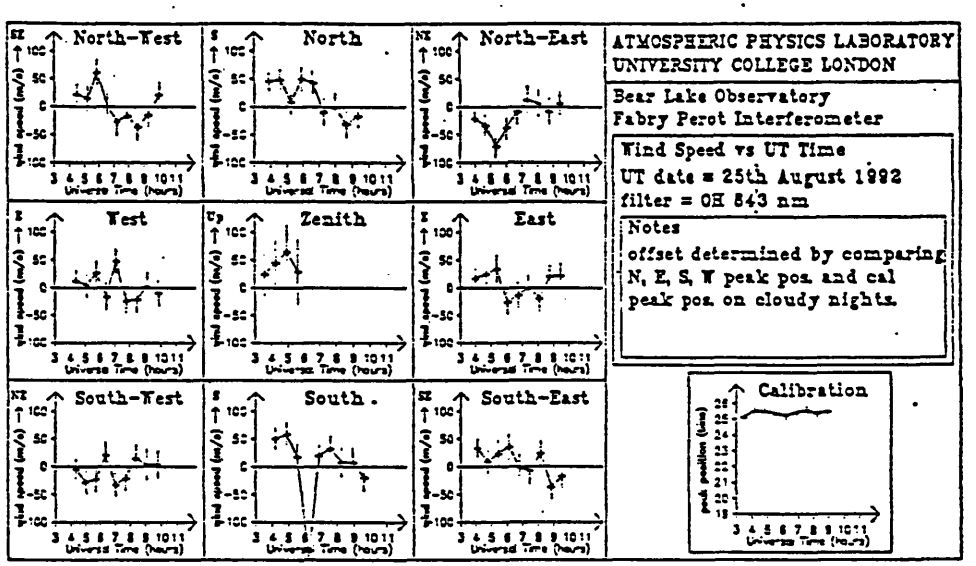
d



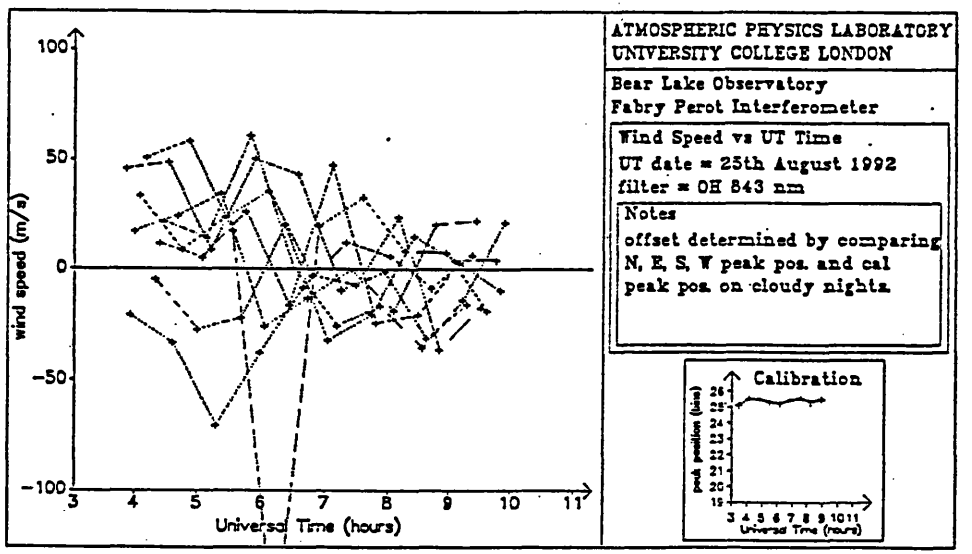
a



b

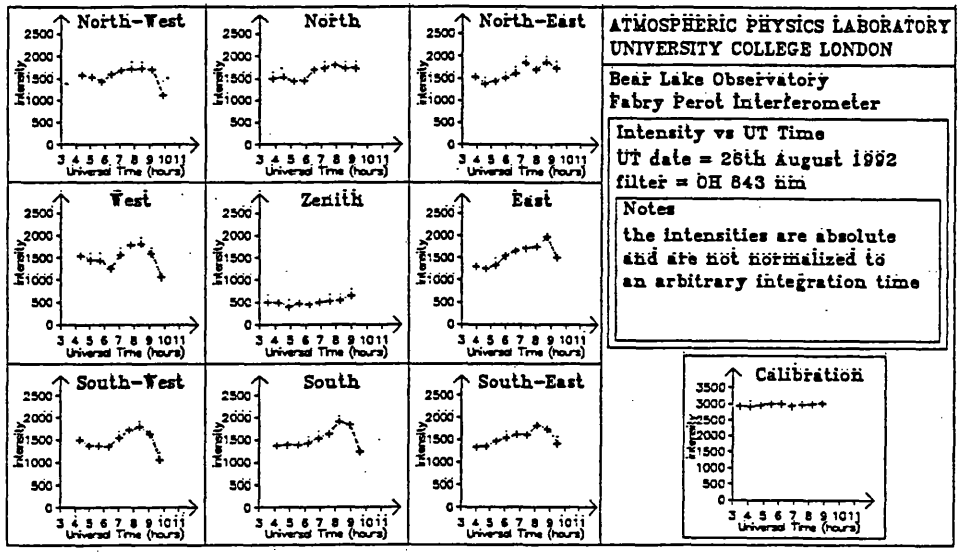


c

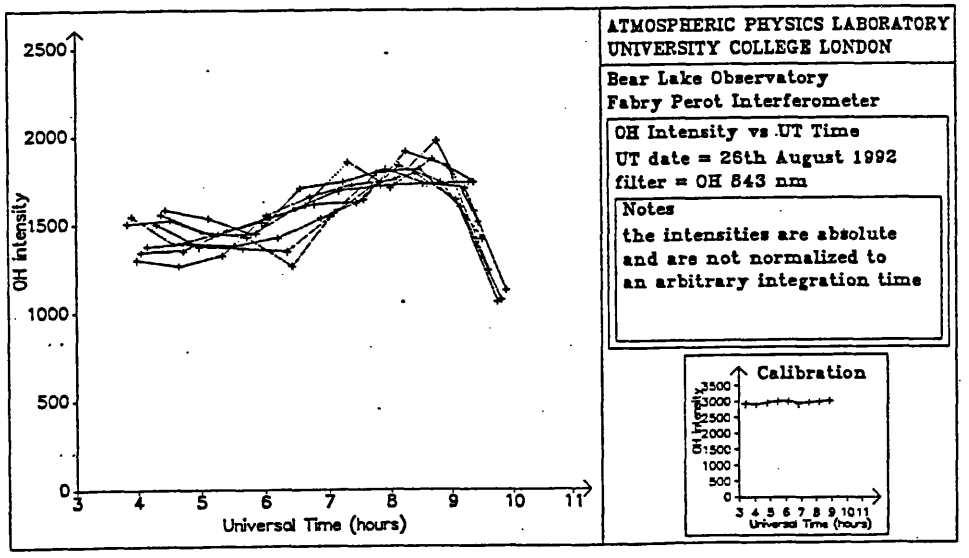


d

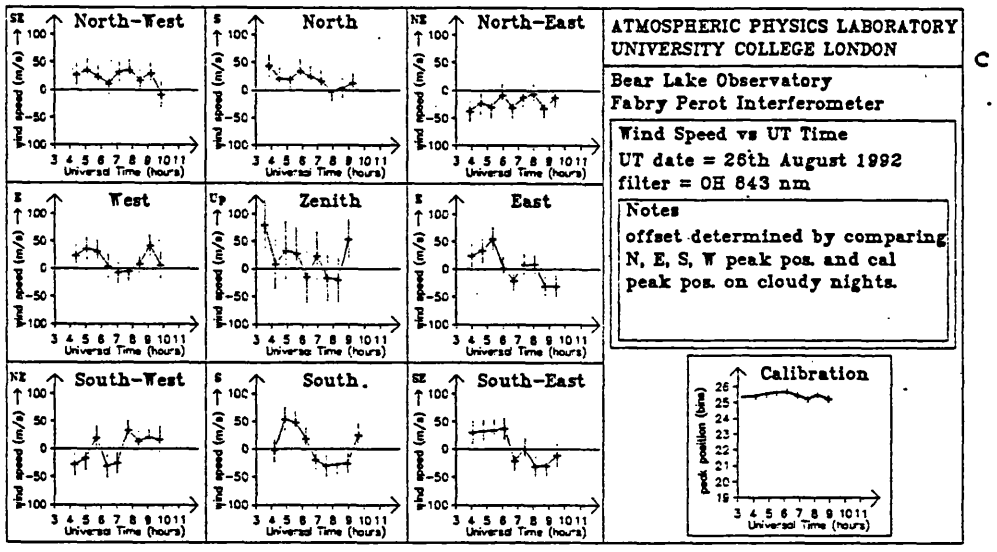
Figures 6.8a, b, c, and d



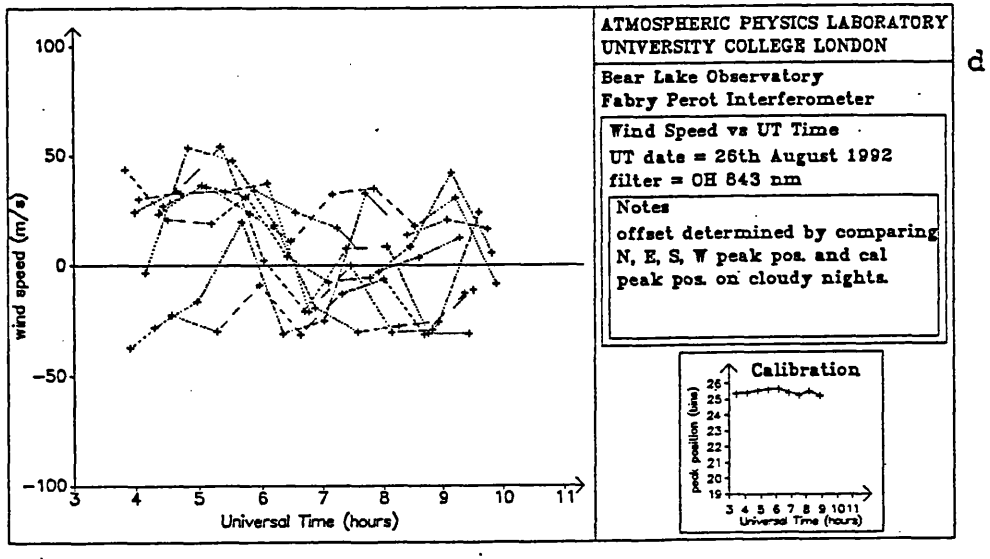
a



b



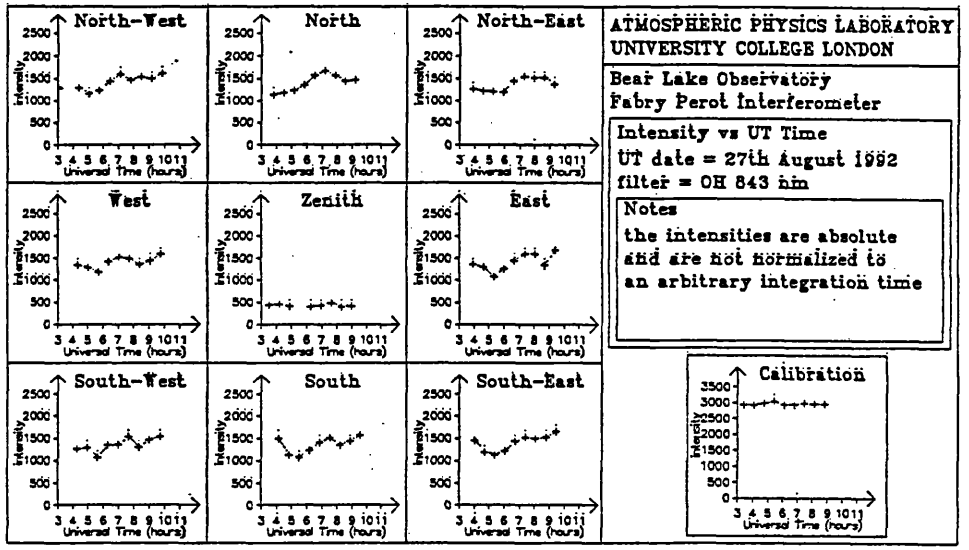
c



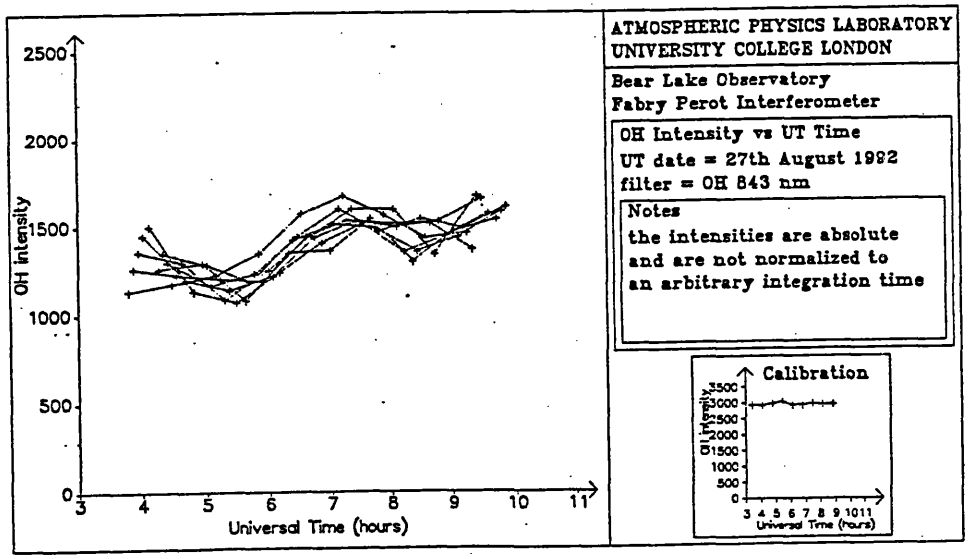
d

Figures 6.9a, b, c, and d.

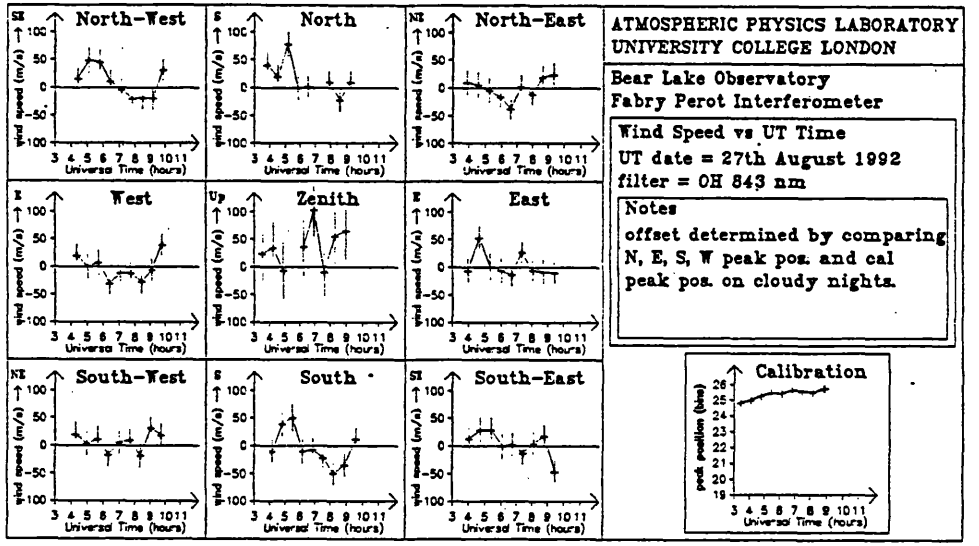
Figures 6.10a, b, c, and d



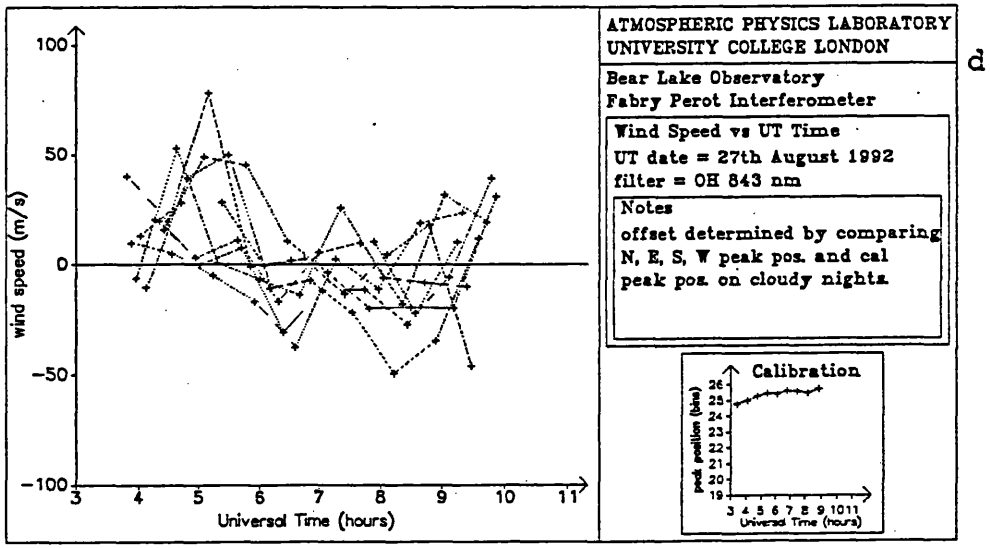
a



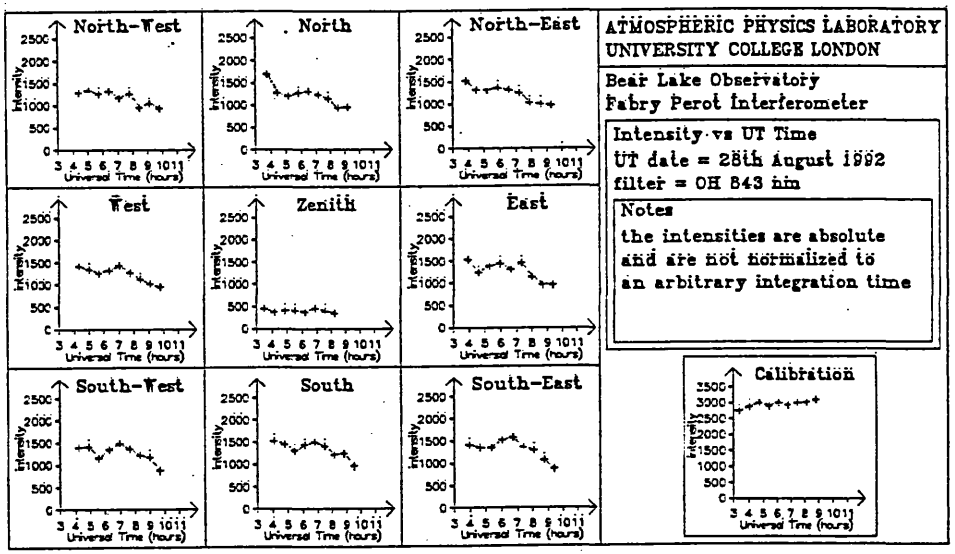
b



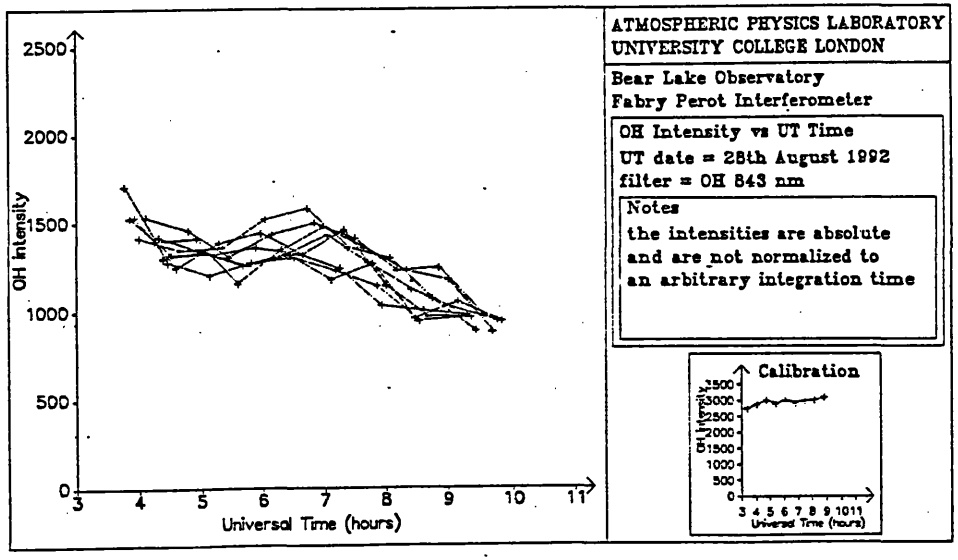
c



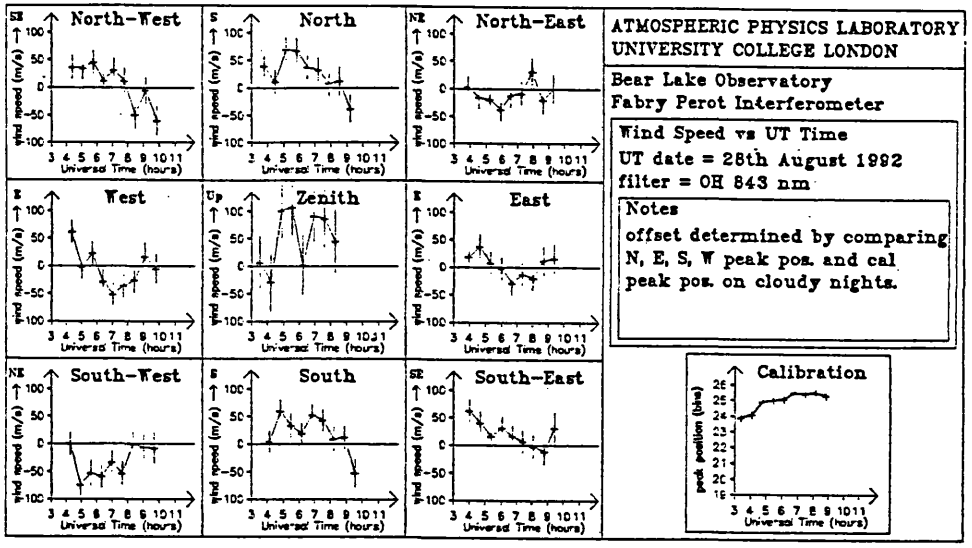
d



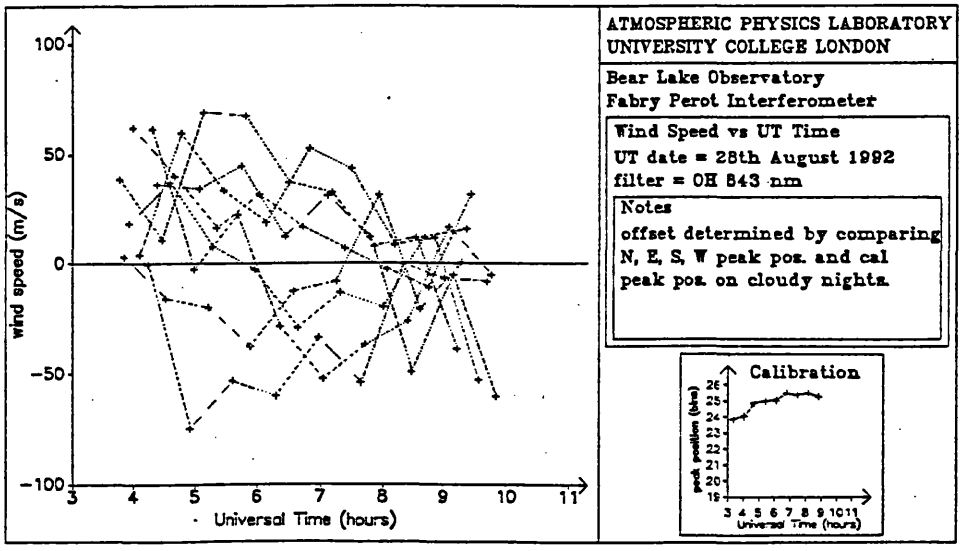
a



b



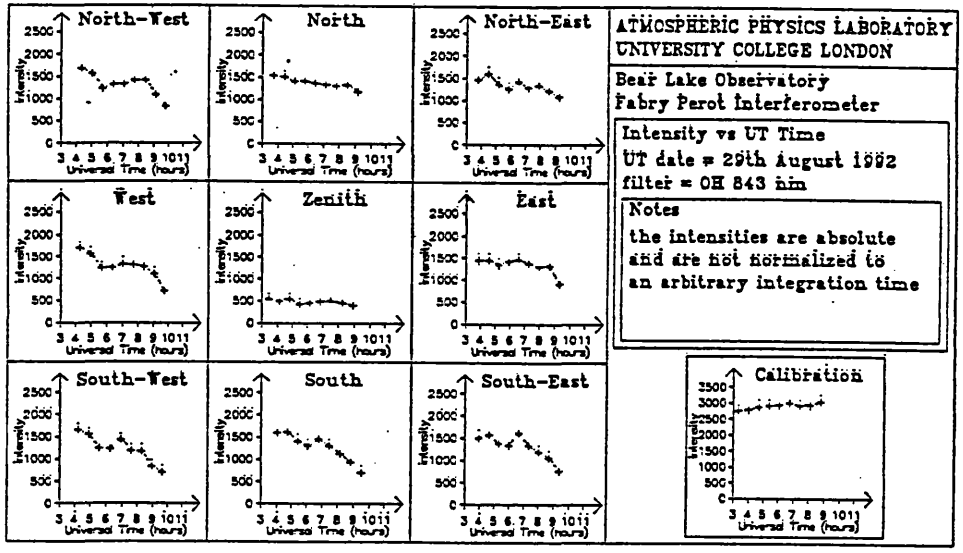
c



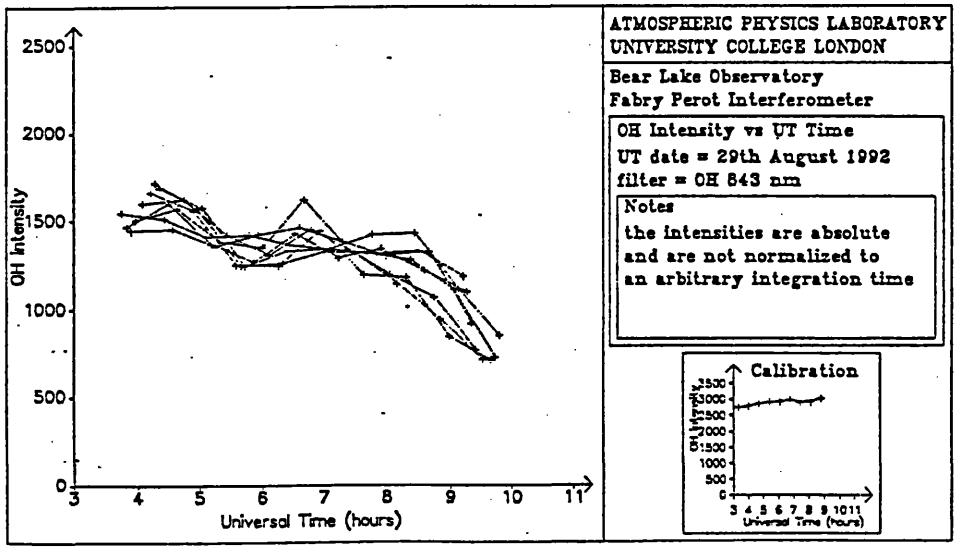
d

Figures 6.11a, b, c, and d

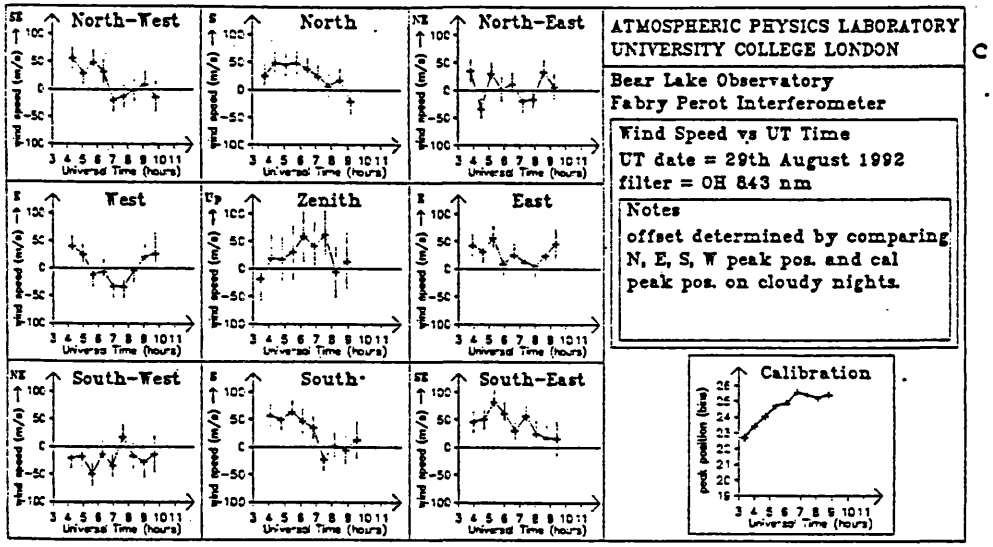
Figures 6.12a, b, c, and d



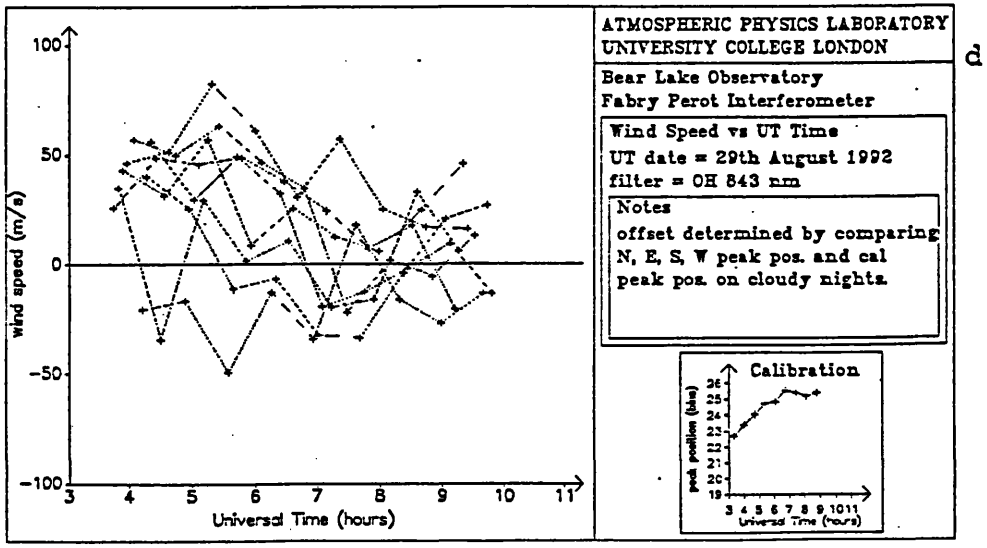
a



b



c



d

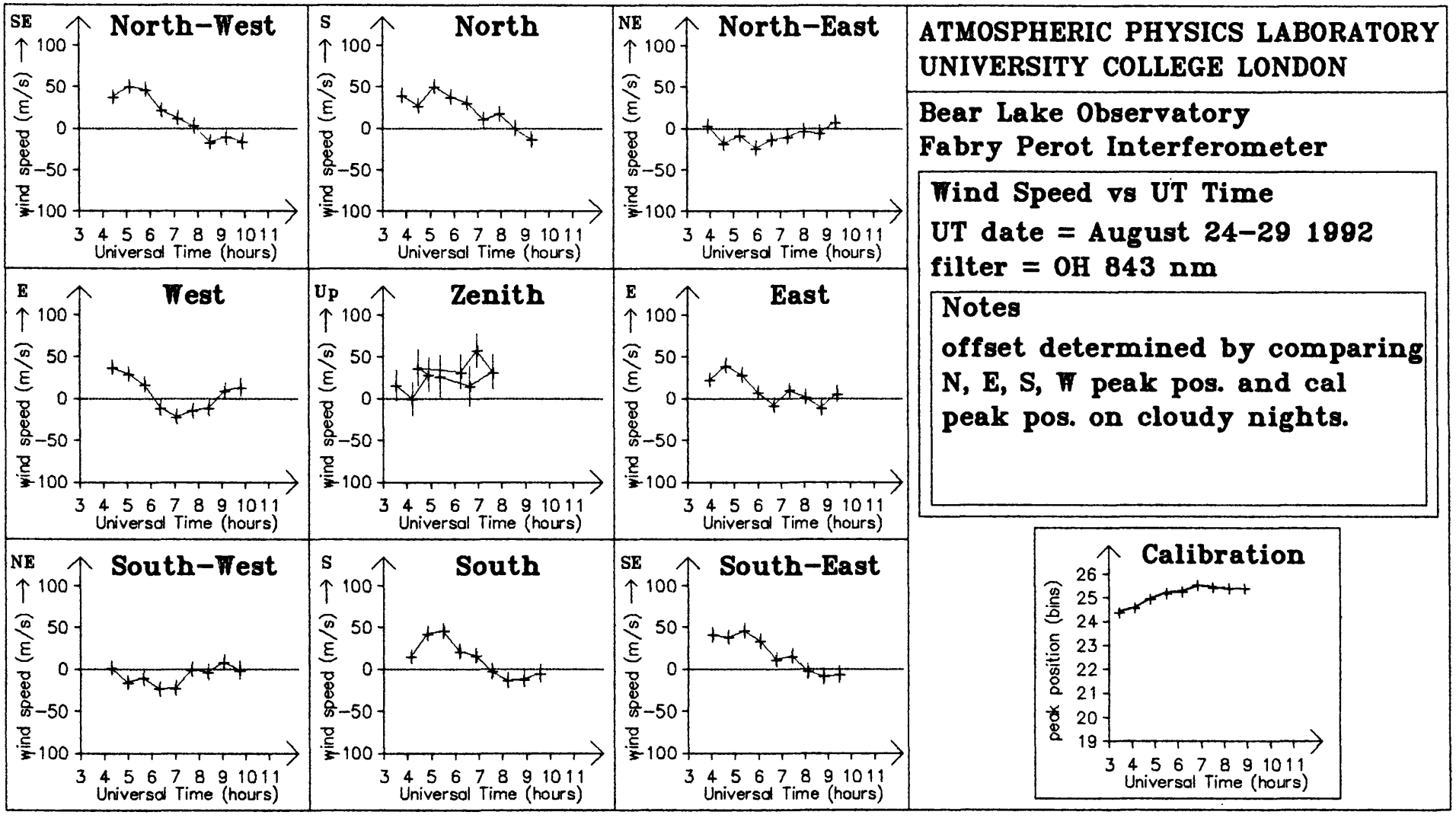
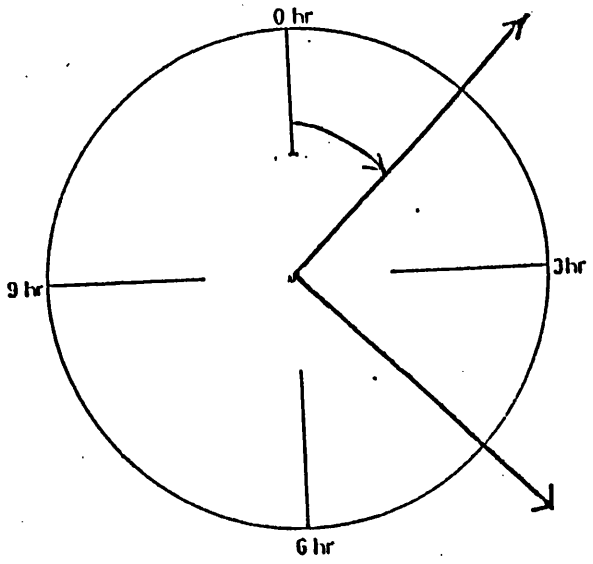

SE
North-West
wind speed (m/s)
100
50
0
−50
−100
3 4 5 6 7 8 9 10 11
Universal Time (hours)
S
North
NE
North-East
E
West
Up
Zenith
E
East
NE
South-West
S
South
SE
South-East
ATMOSPHERIC PHYSICS LABORATORY
UNIVERSITY COLLEGE LONDON
Bear Lake Observatory
Fabry Perot Interferometer
Wind Speed vs UT Time
UT date = August 24-29 1992
filter = OH 843 nm
Notes
offset determined by comparing
N, E, S, W peak pos. and cal
peak pos. on cloudy nights.
Calibration
peak position (bins)
26
25
24
23
22
21
20
19

Figure 6.13

12 hour tide



SW +ve component

NW +ve component

Figure 6.14

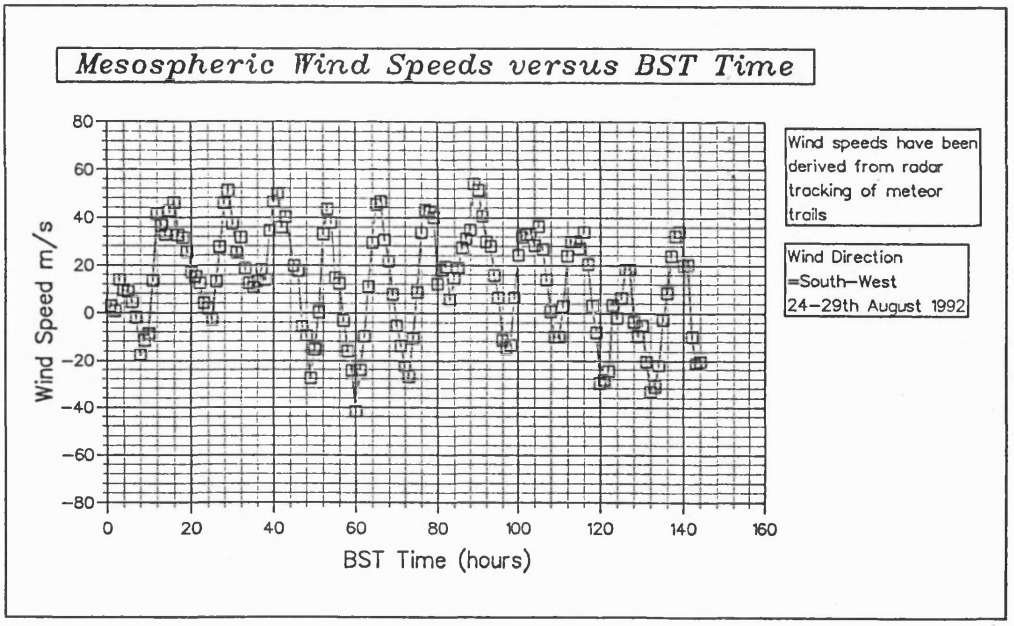


Figure 6.15a

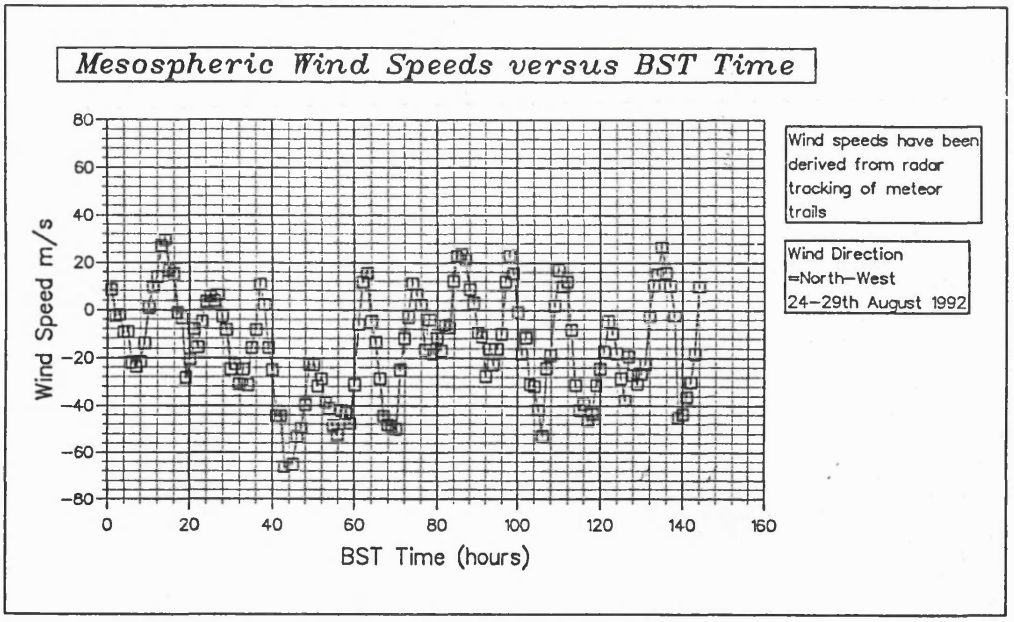


Figure 6.15b

Chapter 7

7.1 Introduction

This chapter summarizes the principle results discussed in Chapters 3, 4, and 6, respectively, and will consider the future work proposed to investigate the Middle Atmosphere.

Section 7.2 will be sub-divided into three sections each discussing the respective results obtained in Chapters 3,4,and 6. Section 7.2.1 will summarize the results obtained from the APL rocket-borne instrumentation flown during the absence, and presence of Noctilucent Clouds (NLC). APL is considering participation in future proposed campaigns. The results yielded by the Bomem MI, and possible new directions of investigation will be discussed in Section 7.2.2. Finally, Section 7.2.3 will review the recent wind and intensity results from the Imaging Fabry-Perot Interferometer (IFPI) at Bear Lake Observatory (BLO), Utah, during the summer period in 1992.

Current investigations, such as the Upper Atmosphere Research Satellite, are providing the first direct, global-scale measurements of the horizontal wind field in the Upper Atmosphere. Some preliminary results are discussed in Section 7.3.

Lastly, in Section 7.4, Future Work, the proposed future investigations by the APL in the mesosphere, lower thermosphere,

and ionosphere will be highlighted.

7.2 Principle Results

7.2.1 Principle Results Obtained by the APL Rocket-borne Instrumentation

The APL has been involved with a series of NLC related campaigns centred around rocket flights, on which various experiments were flown into the vicinity of a NLC. The same photodetectors were flown on NASA rocket payloads during the Middle Atmosphere ElectroDynamics (MAED) Campaign in 1986, and the recent Noctilucent Cloud 1991 (NLC-91) Campaign in 1991.

The APL photometers are instruments with a small field-of-view, which scan the entire sky laterally from the payload, utilizing the payload spin and cone. As a result of the spin and coning motions of the payload, these instruments provide two-dimensional imaging of the NLC layer/s from below 60-80 km, and from above when the payload is over 90 km. They also detect the transition as the payload passes through an aerosol layer.

The APL photometric devices have proved to be an excellent method of detecting the presence of NLC (see Figures 3.16a and b). Both Figures 3.16a and b, illustrate the characteristic change in gradient, negative and positive (upleg/downleg) respectively. These figures are indicative of a traversal through an aerosol layer, illustrating the expected change in gradient

(negative/positive) of the photometer reading. During the flight, the transition is marked on the upleg by a sharp, 1-3 km, decrease in the lowest signal observed in the spin/cone periods, and the reversal is true for the downleg flight. From Figure 3.16a, the single NLC display was a strong, stable layer located at an altitude of 83 +/- 2 km. However, there were suggestions that there were more than one aerosol layer as observed during the CAMP (Cold Arctic Mesopause Project) Campaign in 1982. This is an understandable presumption on looking at Figure 3.16b. However, these postulates are erroneous. The glitches observed are probably due to the NLC being in the process of cloud dispersion and having a "patchy" appearance. To confirm that the cloud was non-stratified, the temperature profile illustrated (see Figure 3.19) indicated only one minimum. Also, plotting the density profile logarithmically gave no further indication of the cloud structure. Although not conclusive, the NLC display observed (26th of July 1986) during the MAED Campaign was most probably only one layer. The MISU results also confirm that only one NLC layer occurred at 82.2 km.

Because of the success obtained during the MAED Campaign, the APL photometers were utilized on both Mother Sections of each EFIELD payload flown during the NLC-91 Campaign. Prior to the launches, the APL photometers passed all pre-flight test runs adequately. However, the data return from each flight were poor due to the attitude problems encountered with the Mother payloads (see Section 2.5). The EFIELD payloads suffered the most difficulties. Salvo B (launched 1st of August 1991) developed attitude problems

when the Daughter Section collided with the Mother Section shortly after its separation. The collision resulted in an exaggerated coning, and severe attitude problems. Salvo A (launched 9th of August 1991) developed even more severe attitude difficulties despite the fact that modifications had been made to the timing sequence to prevent another collision during the upleg flight (see Table 2.3).

The preliminary results from salvo B have not been illustrated graphically, as very little can be discerned from the plots. This is the result of the severe coning encountered during the flight. This flight did not meet the minimum NASA launch success criteria, and has been classified a failure by NASA.

Although, during salvo A, the EFIELD A payload suffered similar attitude problems as experienced by EFIELD B payload, the scanning APL photometers did detect an aerosol layer on both the upleg and downleg flights (see Figure 3.20a and b). With the problems encountered the data presented in Chapter 3 are in their preliminary state. Figure 3.20a indicates a feature at a height of 74 km. Naturally, there was some concern with such a result especially when the height of a NLC is usually ~82 km. On re-plotting, and seeking advice, the feature still occurred at 74 km. However, on the downleg flight (see Figure 3.20b), a similar feature to that seen in Figure 3.20a occurred at 82 km. The author is certain that the latter feature is "real", and may indeed indicate the presence of the confirmed NLC at an altitude of ~85 km. This result is consistent with the preliminary results

of Hale, Pfaff and Witt presented at the NLC-91 Workshop, NASA/GSFC, 28-29 August 1992. However, the upleg feature may be spurious. The results will need many months of detailed analysis, and interpretation before an irrefutable conclusion can be drawn.

7.2.1.1 Future NLC Rocket Campaigns

Future planning is under way to repeat the NLC-91 Campaign, with NLC-93, which is mainly a European program, and NLC-94, another joint USA-European program. As yet, these proposed campaigns have not been funded by the agencies.

The APL have been invited to re-fly the same photometers flown during NLC-91 Campaign in the NLC-94 Campaign. The APL has accepted this invitation.

7.2.2 Principle Results Obtained by the Bomem Michelson Interferometer

During 1991, and 1992, a closer look was made of the climatology of the OH during the NLC season. The climatological data indicated the presence of large scale planetary disturbances, and presently the correlation between these, and the occurrence of NLC are being examined.

With respect to the NLC-91 Campaign, the Bomem MI suggested that there were no NLC displays on the 1st of August 1991, the launch night of salvo B. Therefore, the results obtained from the

APL photometers were spurious as a consequence of the exaggerated payload spin and coning during this launch.

During the NLC display observed on the 9th of August 1991, the launch night of salvo A, the Bomem MI indicated no significant changes in either the temperature, or radiances.

The nightly averages of the integrated band radiances discussed in Chapter 4, seemed to increase after a NLC display, although there was no significant trend during the display. But it should also be noted that the integrated band radiances have also increased at times when NLC were not present. The corresponding nightly averages of the rotational temperatures also indicated a temperature decrease after the NLC display, but this was found not to be always the case. The rotational temperature can be seen to decrease before the NLC display. Figures 4.16, and 4.17 clearly illustrate the aforementioned trends.

One strong trend was reported upon; the integrated radiance during a NLC was less intense than that following a NLC display. Considering the corresponding rotational temperatures around local midnight (see Figures 4.18, and 4.19), this time series reveals several long period fluctuations in the hydroxyl radiance and temperature which appear to be correlated with the appearance of NLC. These fluctuations are interpreted to be the signature of large, planetary-scale disturbances that are responsible for changing the local mesospheric air mass or wind field over the MI.

An accurate assessment of the period of the disturbance was not possible, since the observations were carried out at different geographic locations. At intervals of several days, the MI was relocated northward, as the solar depression angle permitted, in order to make coordinated observations with EISCAT and the NLC-91 rocket flights and CUPRI.

From the investigation of the climatology of the high latitude, mesosphere OH during the 1991 summer NLC season, a new direction of investigation has become clear. A closer examination should be made of the combination of low infra-red radiance, and low temperature. Does this combination allow NLC still to be seen? See Section 7.2.2.1, Future Proposed Bomem Investigations.

During the summer season, the temperature of the mesosphere changes very little on days when NLC are present compared to when clouds are absent. Many of the infra-red OH measurements illustrated that the temperature near the mesopause to be above the "frost" point during periods when NLC were present [Espy and Witt, 1991; Taylor **et al**., 1989; Shefov, 1967; Harrison, 1973][132,91,133,90].

There has been much speculation surrounding the possibility that the presence of waves in the mesosphere supports the NLC formation process. Since gravity waves are assumed to cause the Band structures observed in NLC displays, the structured temperature profiles with wave-like oscillations [Philbrick **et al**., 1984][128] have led to the postulation that gravity waves might

be necessary for the formation of NLC by providing growth regions in the cold "troughs" of the waves. This provides a mechanism whereby particle nucleation and growth could occur even though the measured OH temperature, which is generally an average over time and altitude, is actually above the "frost" point. However, there are as yet no observations that unambiguously relate the appearance of the NLC to the level of gravity wave activity, or to the phase and amplitude of the tides, and planetary waves present in the polar summer mesosphere.

Another feature of the gravity wave mechanism is that the upward component of acceleration associated with these waves could counteract particle sedimentation rates, and allow the particles to grow larger than the models presently allow [Turco **et al.**, 1982; Jensen and Thomas, 1988][61,103].

However, the average particle thus levitated would be exposed to both hot and cold temperatures. Since the time scales for cloud growth are of the order of hours to days, while the sublimation rates rapidly approach seconds for temperatures near 150 K, the wave may actually destroy the cloud. Nevertheless, since gravity wave periods were short compared to the NLC particle growth times, the wave activity must stop once particles had nucleated or the gravity waves would lead to rapid ice particle ablation. These wave motions in the mesosphere could explain the variability in the NLC observations (see Chapter 1). The effect of these planetary waves on the nucleation, and growth of mesospheric clouds has not yet been explored in models.

7.2.2.1 Future Proposed Bomem Investigations

A future proposal known as the Arctic Noctilucent Clouds Campaign-93, will remotely sense the Arctic summer mesosphere using infra-red optical techniques. This will be a unique opportunity of using both aircraft-borne, and the ground-based MI to measure the OH radiance, and determine both the spatial, and temporal characteristics of the wave fields. This information will then be correlated with stereo-photographic measurements of the cloud presence, height, and structure in order to deduce the relationship between the formation of mesospheric clouds, and the presence of small-, and large-scale waves.

Using the USU MI, observing OH in the near infra-red (NIR) spectral region, has the advantage that the reduced light scattering in this spectral range provides the possibility of measuring the OH radiance and rotational temperature from the ground with the Sun only 3°-5° below the horizon. Therefore, on the aircraft, due to the decrease of solar scattering with altitude, the existing instrumentation should be capable of meeting these future developments of measuring OH radiance and rotational temperature as far North as 65°-70°. Obviously to increase the range of these instruments, polarizing filters would be able to improve the rejection of the solar scattered light.

7.2.2.2 Future Bomem Studies

Although the knowledge of the behaviour of OH (3,1) has been

greatly improved by ground-based Bomem observations, particularly from the NLC-91 Campaign, many important questions still remain to be answered:-

1) will the combination of low OH radiance, and low temperature allow NLC to be still observed?

2) how does the OH respond to the occurrence of upward propagating waves?

3) is there an influence from geomagnetic activity on OH temperature?

4) how does the OH radiance, and temperature vary as a function of latitudes, and is there a seasonal dependence?

With respect to 4), this investigation would not be possible at high latitude observation sites alone. Also, the observations would need to be on a long-term routine basis on all clear nights.

At the time of writing, the Arctic Noctilucent Clouds Campaign-93, and the TIMED (Thermosphere-Ionosphere-Mesosphere Energetics and Dynamics) mission are the only proposed future major investigations measuring OH and examining the spatial, and temporal characteristics of the wind fields. This information will then be correlated to the measurements of mesospheric cloud presence, height, and structure in order to procure a

relationship between the formation of mesospheric clouds, and between the formation of mesospheric clouds, and the presence of small-, and large-scale waves.

7.2.3 Principle Results Obtained by the Imaging Fabry-Perot Interferometer (IFPI)

The IFPI located at BLO, Utah provides a unique opportunity, providing continuous wind, and intensity profiles of the upper mesosphere and the lower thermosphere (MALT) throughout the year.

Despite the limited data available from the summer period of 1992, the data collected from six consecutive observing nights at the end of August is consistent with a semi-diurnal tidal mode, with wind vector rotating from south-eastward to north-westward during the observing night. Further, these observations are in surprisingly good agreement with those determined from the Sheffield radar meteor data taken in the UK during the same periods of observation.

A dramatic feature occurred on the 5th of June 1992, which was not consistent with a semi-diurnal tidal mode. During the 5th of June 1992, the winds observed in the opposite viewing directions were moving in opposite directions. This anomalous result has been tentatively ascribed to the passage of a gravity wave, possibly associated with a meteorological disturbance in the troposphere.

Although, this feature did not occur again during the limited summer data, it is believed to be a possible gravity wave, but until the data collected during the summer 1993 has been analyzed this dramatic feature can not be truly verified.

The analysis of the upper mesosphere data from BLO, Utah is still in its initial phase of development, and it is planned that with future data sets, the reduced data will be fitted with sinusoidal functions to examine the periodicities more closely, and to verify the presence of gravity waves.

7.2.3.1 Future IFPI Studies

It is anticipated that, in the future, the analysis of a larger data base will provide more conclusive information regarding the dynamic forcing - due to tides, planetary waves, and gravity waves - driving the chemistry yielding mechanisms for both fine, and large scale structures in the mesosphere.

However, an in-situ instrument must be employed to identify the minor species concentrations, and the geophysical conditions must be explored to comprehend the coupling between the chemistry, and the dynamics in the MALT.

Observations of the OH (6,2) emission band show considerable promise for investigating the properties of the wave motions propagating through the MALT.

For the BLO IFPI, the neon calibration lamp was replaced by a frequency-stabilized He-Ne laser in November 1992. Temperatures, in addition to winds, should be available from the BLO IFPI from subsequent date. With OH intensities, wind speeds, and temperatures, the horizontal wave structure induced by gravity waves will be investigated with additional confidence.

An all-sky camera with a OH filter would provide an excellent method for remote sensing the two-dimensional nature of the induced wave structure over a large geographic area, and with a high temporal resolution. These observations would give detailed information on the frequency of occurrence, horizontal extent, and direction of the propagating of the waves. However, during the summer months, to date there have been no simultaneous IFPI, and all-sky imaging investigations of the vertical and horizontal wave structure induced by tides, and gravity waves over BLO.

The present account of meteorological conditions logged at Logan, once per night is inadequate. It should be emphasized that BLO is forty miles away from Logan, and is situated on the other side of the Wasatch Mountains. Therefore, an on-site optical technique is required at BLO to monitor the weather conditions.

The detailed error analysis discussed in Section 5.5.4, stresses strongly the need of at least 1000 counts peak intensity ie peak intensity above the background intensity, in order to achieve a satisfactory wind determination. From the summer of 1992, the measured peak intensities varied from 500 to 2500, and on this

basis of the analysis, it has been recommended that the integration time be doubled, and the intermediate viewing directions, NE, SE, SW, and NW be eliminated, therefore maintaining the same overall cycle time. Table 5.9 illustrates clearly the improvement in the propagation of errors compared to those summarized in Table 5.8.

It is believed that the aforementioned suggestions will enhance the quality of the data collected, reduce errors and hence, improve the confidence in the derived wind speeds from the BLO, Utah and elsewhere.

7.3 Current Investigations

Upper Atmosphere winds can be measured directly by ground-based radars and lidars, and by balloon- and rocket-borne sounding instruments. Such techniques permit study of small-scale dynamical processes and long-term monitoring of atmospheric dynamics over fixed deployment sites; however, they are intrinsically limited to a few geographic locations. For investigations of the global-scale dynamical processes that determine the distribution of ozone and other Upper Atmosphere constituents, satellite data are required.

7.3.1 Wind Measurements by UARS

The UARS (Upper Atmosphere Research Satellite) is providing the first direct, global-scale measurements of the horizontal wind

field in the Upper Atmosphere. With the aid of theoretical investigations, and numerical modelling, the UARS data will shed new light on fundamental questions in stratospheric and mesospheric dynamics:-

1) the relative importance of various types of wave motion as a function of altitude,
2) the relative roles of tropospheric energy fluxes , and in-situ energy generation by solar heating,
3) the factors affecting the breakdown of polar winter circulation patterns in the stratosphere, and the mechanisms responsible for the warm winter and cold summer mesopause.

Two UARS instruments - a High Resolution Doppler Imager (HRDI), and the Wind Imaging Interferometer (WINDII) provide direct observations of wind velocity through measurements of the Doppler shifts of selected emission, and absorption lines. These shifts are measured in two different directions, yielding two components of the wind velocity relative to the spacecraft; the true wind velocity can then be calculated from the observing geometry, and a knowledge of the spacecraft velocity. The spacecraft motion, combined with vertical scanning by the instruments produces a three-dimensional global map of the Upper Atmosphere wind field.

Since atmospheric emission and absorption characteristics vary strongly with altitude, a number of atomic and molecular spectral features are utilized to obtain a wide range of altitudes. In

both experiments, the Doppler shift that arises from spacecraft motion must be separated from the shifts caused by atmospheric motions.

7.3.2 High Resolution Doppler Imager (HRDI)

At altitudes between 10-40 km, the HRDI observes the Doppler shifts of spectral lines within the atmospheric band system of molecular oxygen to determine the wind field. There are no sharp emission lines in the radiance of the Earth's limb at such altitudes, but the molecular oxygen contain absorption bands in the near IR that appear as deep absorption features in the brilliant spectrum of sunlight back-scattered by the atmosphere (Rayleigh scattering). A triple etalon FPI, serves as a high resolution spectral filter. HRDI uses the absorption features present in the spectrum of the sunlight back-scattered by the atmosphere to measure Doppler shifts from which wind data for the stratosphere and upper troposphere are determined to an accuracy of 5 ms^{-1}.

At altitudes greater than 60 km, HRDI observes emission lines of neutral and ionized atomic oxygen in the visible and the near infra-red spectral regions by the same interferometric technique. Unlike the molecular absorption lines, the emission lines are observable both day and night. These measurements will furnish the wind field in the mesosphere and thermosphere to an accuracy of 15 ms^{-1}.

7.3.3 Wind Imaging Interferometer (WINDII)

WINDII utilizes the emission lines for the basic Doppler shift measurements. In addition to lines of neutral and ionized atomic oxygen, these include two lines of the OH molecule, and a molecular oxygen line. WINDII obtains wind and temperature measurements both day and night at altitudes above 80 km [Shepherd **et al.**, 1993][201] .

The WINDII spectral filter is a high resolution Michelson interferometer. The instrument consists of a telescope, the interferometer, and a detector array. The telescope views 45° and 135° from the spacecraft velocity vector simultaneously. In normal operation, the detector provides a vertical resolution of about 4 km and a horizontal resolution of ~20 km. Wind velocity accuracy within 10 ms^{-1}, in the altitude range between 80 km and 300 km.

7.4 Future Work

The future holds enormous promise for ground-based, and satellite-borne remote-sensing instrumentation.

With the pioneering measurements to be made by the proposed ALOMAR and TIMED projects, it is anticipated that a better understanding of the Middle Atmosphere, will be gained. Sections 7.4.1, and 7.4.2 will briefly outline the scientific value of each project, together with the proposed role of the Atmospheric

Physics Laboratory (APL).

7.4.1 Arctic Lidar Observatory for Middle Atmosphere Research (ALOMAR)

7.4.1.1 Overall Description of the Project

ALOMAR will be a major new international lidar facility for observing the Arctic Middle Atmosphere. The Observatory will be located North of the Arctic Circle on the Ramnan mountain (379 m) on the island of Andoya in Northern Norway (69°N, 16°E).

ALOMAR will measure the concentrations of aerosols, and minor constituents including ozone, wind velocities, momentum fluxes, temperature and density profiles throughout the atmosphere within the height range of 10 km to 110 km, with high accuracy, and good vertical, and time resolution. It will be built by a consortium of research groups, contributing a range of essential scientific, and technological experience to the design, building and commissioning of the necessary complement of Rayleigh/Raman, resonance fluorescence, and ozone lidars.

The proposed lidar system is anticipated to enable high resolution studies of the thermal structure, and aerosol constituents throughout the Middle Atmosphere. These studies would encompass:-

1) the mean and variable thermal structure due to, or in

association with, stratospheric warming, and the polar stratospheric clouds (PSC) credited with ozone depletion via heterogeneous surface chemistry,

2) dust layer structure, and morphology in the stratosphere, and mesosphere,

3) the structure distribution, and associated dynamics of NLC, and PMC near the polar summer mesopause.

ALOMAR will provide a much needed link between the Lower, and Middle Atmosphere, contributing very high spatial, and temporal resolution of winds, and temperatures, yielding a direct measure of the wave forcing of the mean flow over a wide range of heights, and permit joint lidar and radar studies permitting even greater capabilities. In addition, the new proficiency of the lidar installation for simultaneous temperatures, and dynamics measurements will extensively profit investigations of the Middle Atmosphere turbulence.

7.4.1.2 The Role of the Atmospheric Physics Laboratory (APL), UCL

The technical role of the APL is the development, fabrication and commissioning of the Doppler wind and temperature channels (DWTC) for the Rayleigh/Raman lidars. Both channels will measure winds, and temperatures from the Doppler shift, and Doppler broadening of the Rayleigh back-scattered signal. This is scheduled to be completed in two phases.

In phase 1, to be completed by mid-1994, DWTC1 will be built. This channel will examine the relatively intense returned signal from the upper troposphere, and stratosphere (10 - 60 km). This will consist of a double-etalon Fabry-Perot Interferometer (FPI) utilizing proficient photomultiplier detectors. Initially, DWTC1 will utilize all of the available returned signal. However, after phase 2, 20% of the total received signal will be diverted into this channel, providing an adequate signal to obtain wind velocities, and temperatures up to an altitude of 55 km.

In Phase 2, to be completed by mid-1995, DWTC2 will be built. This will consist of a double etalon FPI, optimized to use a ring-anode Imaging Photon Detector (IPD). 80% of the returned signal will be guided into this channel. DWTC2 will utilize the back-scattered signal more effectively, and will make measurements of wind velocities, and temperatures from 30 km to 85 km.

This philosophy will allow adequate overlap in altitude between the two systems, ensuring continuity without losing high-altitude capability. By providing the equipment in two phases ALOMAR will always have a wind measuring capability, while cross-calibration can be ensured for the entire facility. The etalons and detectors for phase 2 are distinct from phase 1, and each is a specialized and optimized system.

The scientific role of the APL will be concerned with the exploitation of the wind, and temperature observations of the

Middle Atmosphere. Observations which will be of the greatest interest will be those describing the upward propagation of gravity, planetary waves, and tides into the upper stratosphere and mesosphere, and the interactions between disturbances of the Middle Atmosphere origin, and those of magnetospheric origin within the mesopause region, and the lowest parts of the thermosphere. Studies of the roles of various local, and large-scale processes in the generation of particulates in the mesosphere, and of lower ionospheric layers will also be of great interest.

7.4.2 Thermosphere-Ionosphere-Mesosphere Energetics and Dynamics Mission (TIMED)

7.4.2.1 Overall Description of the Project

TIMED has been proposed to carry out the first comprehensive investigation of the physical, and chemical processes acting in the terrestrial mesosphere, lower thermosphere and ionosphere (MLTI) within the altitude range of 60 km and 180 km. This region represents the "cross-roads" in the flow of mass, momentum, and energy through the coupled MLTI and ionosphere-magnetosphere system. It is a region which is poorly understood, primarily due to the difficulty in performing in-situ measurements at these altitudes. This is a transition region, where atmospheric temperatures, and thermal gradients reach extreme values, where turbulent mixing gives way to molecular diffusive flow, where the composition changes from molecular to atomic, and where the

photochemical and electrodynamical processes become predominant.

TIMED will be an exploratory mission to characterize, and understand the interplay of the dynamics, energetics, and radiation, and neutral, and ionized chemistry within this region. This mission will complement the NASA UARS (Upper Atmosphere Research Satellite) ($tratospheric), and other MALT projects. TIMED will contribute to the study of anthropogenically induced atmospheric changes. Specifically, the mesosphere is sensitive to increasing levels of methane, and carbon dioxide. It is critical to establish a quantitative base-line of the region at the earliest possible date, and to utilize new observations to understand, and monitor the changing processes.

The TIMED satellite mission is proposed to incorporate two identical spacecraft, one in a near polar orbit, and the other in an orbit of moderate inclination, to provide extensive coverage of MLTI parameters in altitude, latitude, local time and season. The spacecraft are proposed to be launched into orbits with 95°, and 49° inclination, respectively. Both spacecraft are proposed to have on-board propulsion capabilities to permit both circular, and eccentric orbit phases, and to maintain periods where the perigee of the spacecraft may be safely maintained as low as 130 km, for significant periods. The anticipated operating lifetime of each spacecraft is forty-eight months.

7.4.2.2 The Role of the Atmospheric Physics Laboratory (APL), UCL

The University of Michigan requested the participation of the APL in developing a science investigation for TIMED, based on a FPI instrument. One FPI is proposed to be flown on each of the two payloads:-

1) TIMED-H (TIMED-HIGH) will be in a near polar orbit of 95° inclination.
2) TIMED-L (TIMED-LOW) will be in a medium inclination orbit of 49°.

The TIMED FPI Investigation will make measurements of the atmospheric dynamics, temperature, and certain minor concentrations by day, and by night throughout the altitude range between 60 km, and 300 km. The FPI is a core instrument for the TIMED mission. The region between 60 km and 180 km is of particular relevance to TIMED, and the FPI will be optimized for measurements of wind, and temperature within that region. The simulated performance of the FPI is now based on actual intensity, and Doppler wind, and temperature measurements within the region performed by the HRDI (High Resolution Doppler Imager), and WINDII (Wind Doppler Interferometer) instruments of the recently launched UARS spacecraft. The instrument will be able to observe the daylit, and the night-time regions, and will be capable of achieving typical vector wind errors of 5 ms^{-1} with a horizontal resolution of 100 km to 200 km, and a vertical resolution of 0.5 scale height. The FPI will also be adept at

collecting remote-sensor measurements of ion drifts in the auroral oval, polar cap, and in the sunlit hemisphere.

The TIMED FPI is a double etalon FPI, using a state-of-the-art CCD (Charge Coupled Device) detector. The high resolution etalon will have a fixed gap, while the low resolution etalon will be capacitance stabilized. A filter wheel will allow the selection of an appropriate spectral emission feature on the dayside or nightside.

The FPI data sets, comprising of global maps of wind, and temperatures as a function of altitude, will be acquired for each orbit, or during the appropriate phases of the orbit, when the TIMED spacecraft is in its eccentric orbit phase of operations. These data sets, combined with those from the highly complementary data sets of other instruments on TIMED, will be utilized to deduce the parameter sets which will represent the data products used to evaluate, describe, and interpret the atmospheric, and ionospheric phenomena which are the primary scientific objectives of the FPI investigation.

Appendix 1

Determination of the Look Direction of the Photodiode on a Rocket Payload

Assume that the rocket payload with the body-fixed Cartesian coordinates defined by the z_B-axis is aligned with the longitudinal (roll-) axis, and x_B- and y_B-axes in the plane are perpendicular to it. Rotation about the x_B-axis is defined as the yaw, and rotation about the y_B-axis is defined as the pitch of the body. It should be noted that in some contexts, especially in aerodynamic literature, the longitudinal axis is defined as the body-fixed x-axis. This does not change the generality of the geometric considerations. Assume further that the photodiode mounted on the payload is characterized by its internally fixed coordinate system (x_I, y_I, z_I). To find the orientation of the instrument-fixed coordinate axes, it is first necessary to project these axes onto the body-fixed system and then, if the orientation of the payload in space is known, onto the space-fixed system. Similarly, the angles between a celestial body such as the Sun are obtained from the direction cosines of the instrument, and reference direction.

The projection of one coordinate system onto another is most readily performed by using three-dimensional matrices representing the rotation of θ degrees about a fixed axis.

As an example, Equation (A1.1) represents the rotation in the

positive sense about the y-axis, given by the matrix:-

$$\begin{vmatrix} \cos\Theta & 0 & -\sin\Theta \\ 0 & 1 & 0 \\ \sin\Theta & 0 & \cos\Theta \end{vmatrix} \qquad \textbf{(A1.1)}$$

From Equation (A1.1), the elements of this matrix r_{ij} are such that the rows and columns containing the rotation or pivot axis x_k are all zero except for r_{kk}. The other elements are defined from the argument of the rotation, assumed positive in the clockwise direction when the unit vector along the pivot axis is viewed against the origin.

The projection of the photodiode-fixed system is performed by the successive rotations about the body-fixed z_B-axis first, then about the body-fixed x_B-axis, and finally, in cases where it is required, the photodiode-fixed z_I-axis. The transformation matrix representing the entire procedure is:-

$$R_{B\rightarrow I} = \begin{vmatrix} r_{XX} & r_{XY} & r_{XZ} \\ r_{YX} & r_{YY} & r_{YZ} \\ r_{ZX} & r_{ZY} & r_{ZZ} \end{vmatrix} = \begin{vmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{vmatrix} \begin{vmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

(A1.2)

Now, it only remains to define the orientation of the body-fixed axes in a space-fixed reference system. Here it is assumed that

the reference system is the local topocentric (T) system centred at the rocket launcher, with z_T-axis pointing in the zenith direction, the y_T-axis pointing towards the geographic North, and the x_T-axis pointing towards the geographic East. The orientation of the body-fixed system is normally obtained from a gyro system, or from the orientation with reference to the geomagnetic field measured by a two-axis magnetometer. In the latter case, an additional independent reference, for example the Sun's azimuth is required. In the topocentric system, the rotations of the body-fixed coordinates are:-

1) the roll phase about the z_B-axis,
2) the elevation,
3) the zenith angle about the x_B-axis, and the azimuth about the z_T-axis.

However, it should be noticed that the roll movement is normally positive in the right hand sense, that is, the roll phase, κ, is positive in an anti-clockwise direction when the roll axis is viewed towards the backward direction. Similarly, the elevation θ is counted as positive, the zenith angle as negative, in the anti-clockwise direction. The azimuth angle, Φ, on the other hand, is considered as positive in the clockwise direction. The corresponding rotation matrices for the topocentric (T) to body (B) fixed transformation are thus:-

$$R_{T\rightarrow B} = \begin{vmatrix} \cos\kappa & \sin\kappa & 0 \\ -\sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} \cos\Theta & 0 & \sin\Theta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\kappa \end{vmatrix} \begin{vmatrix} \cos\Phi & -\sin\Phi & 0 \\ \sin\Phi & \cos\Phi & 0 \\ 0 & 0 & 1 \end{vmatrix} \qquad \textbf{(A1.3)}$$

$$R_{T\rightarrow B} = R(-\kappa)\ R(-\Theta)\ R(\Phi) \qquad \textbf{(A1.4)}$$

Since the rotation matrices are unitary, the transformation from the body (B) fixed to the topocentric (T) system ie represented by the following matrix:-

$$R_{B\rightarrow T} = (R_{T\rightarrow B})^{T} \qquad \textbf{(A1.5)}$$

where the superscript denotes the transpose of the matrix.

Appendix 2

Mathematical Details of the CAMP Data Analysis

The light scattering properties of air molecules are well established. It is known that, the received energy from the scattering of sunlight is directly proportional to the number of molecules along the line-of-sight of the instrument ie proportional to the local ambient pressure.

In twilight conditions, the radiation reflected by the Earth, and by the underlying atmosphere is negligible. If the atmospheric pressure profile is assumed, the measured radiance at the cloud base can be related to that produced by a certain number of molecules, thus, the light-scattering properties of the cloud particles can be expressed in the terms of the molecular scattering at a certain reference level.

When the Sun is above the horizon, this method can be used if an appropriate correction for the reflected radiation is applied. This correction was found from tabulations.

Solar radiation, which is scattered by the air molecules or by particulate matter, such as that encountered in NLC, is polarized. The nature of the scattering medium is characterized by the degree of polarization, and the colour index of the scattered radiation [Witt, 1968][110]. These parameters are related to the scattered radiance components vibrating parallel $\|$, and

perpendicular to ⊥, the scattering plane. This plane is defined by the photometer axis, and the position of the Sun. The degree of linear polarization is:-

$$P = \frac{(I_{\perp} - I_{\parallel})}{(I_{\perp} + I_{\parallel})} \qquad \textbf{(A2.1)}$$

and the colour ratio is:-

$$C = \frac{(I_{\perp} + I_{\parallel})_1}{(I_{\perp} + I_{\parallel})_2} \qquad \textbf{(A2.2)}$$

where the subscripts 1, and 2 refer to two different wavelengths

From Equation (A2.1):-

$$I = I_{\perp} + I_{\parallel} \qquad \textbf{(A2.3)}$$

$$Q = I_{\perp} - I_{\parallel} \qquad \textbf{(A2.4)}$$

are the first two Stoke parameters of the scattered radiation [Chandrasekhar, 1960][202], which are functions of the scattering

angle. This is the angle between the incident, and scattered beams of radiation as illustrated in Figure 3.4. Figure 3.5 illustrates the relationship between the scattering-plane frame of reference, and the direction of the optic axis of the photometer.

The instantaneous direction of the optical axis, in this case of an axially looking photometer, is in turn determined by the dynamical coordinates, of the spin-stabilized rocket, as shown in Figure 3.6. Here:-

Z = the longitudinal axis of the payload
S = the spin vector
Σ = the axis of the precession cone
V = the velocity vector
ξ = the semi-angle of the precession cone
α = the angle of attack
l = the precession rate
ω = the spin rate

In general, $l \gg \omega$. In certain cases, usually as a result of an asymmetric distribution of mass in the payload, the Z axis does not coincide with the spin axis but performs a precession about it. This effect, nutation, can lead to an undesired additional modulation of the output signal from a polarization photometer. This modulation has the same rate as the rocket spin whilst the polarization causes a modulation at twice this frequency.

This ultimate background light seen by any rocket-borne photometric device consists of starlight, diffuse galactic light, and the light scattered by solid particles in the interplanetary space (zodiacal light). For the proper assessment, and correction for these components, the direction of the optical axis must be known in terms of the sundry astronomical systems. The necessary coordinate transformations, and the method of obtaining the extra-terrestrial background correction are discussed in the literature [Wood and Witt, 1978][203].

In a molecular Rayleigh-scattering atmosphere, the degree of linear polarization at a scattering angle, θ,is given by the expression [Chandrasekhar, 1960][202]:-

$$P(\Theta) = P(90°)\ \frac{\sin^2 \Theta}{(1 + P(90°)\ \cos^2 \Theta)} \qquad \textbf{(A2.5)}$$

where P(90°) = 0.945 [Penndorf, 1952][204]

The colour ratio is defined by:-

$$C_{12} = \frac{(n_1 - 1)^2\ (\lambda_2)^4}{(n_2 - 1)^2\ (\lambda_1)^4} \qquad \textbf{(A2.6)}$$

where n_1 and n_2 = the refractive indices of air at the wavelengths λ_1 and λ_2 respectively.

A rotating photometer-polarizer combination, as discussed with respect to the axially mounted photometer on a spinning rocket yielded a sinusoidally modulated output when exposed to the polarized sky radiance. The output signal was modulated at twice the roll-rate of the rocket (see Equation (A2.7)). The DC component of the output signal is proportional to the time average of the radiance, that is I, while the amplitude of the modulation is proportional to Q, as previously defined:-

$$B_{obs} = \frac{1}{2} \left[(k_1 + k_2)\, I(\Theta) + (k_1 - k_2)\, Q(\Theta) \cos 2\Psi \right] \qquad \textbf{(A2.7)}$$

where k_1 and k_2 = the principal transmittances of the polarization filter

Ψ = the angle between the larger (k_1) transmission axis of the filter, and a direction perpendicular to the scattering plane

$$\Psi = \frac{2\,\pi\,(t - t_0)}{T_{rot}} \qquad \textbf{(A2.8)}$$

where t = the time

t_0 = the time of an arbitrary maximum of the signal

T_{rot} = the period of one rotation of the polarizer

As aforementioned, I and Q are both functions of the scattering angle, and depend on the number of scatterers in the line-of-

sight ie the slant geometric thickness of the scattering medium in the direction of observation. They vary with the changing attitude of the rocket. Both this variation and the change of the radiance with altitude are normally slow compared with the rate of modulation resulting from the rocket spin. There may be occasions when this is not the case, such as when the rocket traverses a very shallow light scattering layer.

The output signal is determined from Equation (A2.7), with $T_{\perp}/T_{\parallel}$ = 34, and k_1/k_2 = 50.

From consecutive maxima, and minima of the observed signal, the time or altitude variation of both I and Q, have been established.

The actual observed signal exhibits statistical fluctuations arising from various noise sources, and it also contains background noise due to the photomultiplier dark current, starlight, stray illumination from parts of the photometer or the payload, and possible airglow contamination. The background signal can be measured at altitudes where the scattered light contribution is negligible, ie close to the apogee. It was necessary to perform the background measurement at every scattering angle covered by the precession of the payload, since the stray light component is partially polarized, and also, varies with this angle. Primarily, the signal must be smoothed to reduce errors caused by the aforementioned noise sources, before the background can be subtracted. The telemetry signal was

digitized, thus smoothing was performed using the numerical techniques developed at MISU [Witt **et al**., 1973][205]. The numerical procedures yields I and Q as functions of time or altitude in fixed time increments. The same procedure can be used to eliminate periodically varying background signals caused by sunlight reflected into the photometer from parts of the payload.

The magnitude of the polarized component Q does not contain contributions from dark current or starlight, because these are modulated by the rocket spin.

The polarization and colour ratio of the light scattered by solid particles was derived from the smoothed, and background-corrected profiles of I and Q according to the following procedure. Molecular scattering is denoted with the suffix M, airglow emission by A, and scatter due to particles by D.

$$I = I_M + I_A + I_D \qquad \textbf{(A2.9)}$$

$$Q = Q_M + Q_A + Q_D \qquad \textbf{(A2.10)}$$

where I, and Q are the Stokes parameters.

From Equation (A2.1):-

$$P = \frac{(Q_M + Q_A + Q_D)}{(I_M + I_A + I_D)} = \frac{(\frac{Q_M}{I_M} + \frac{I_A}{I_M} * \frac{Q_A}{I_A} + \frac{I_D}{I_M} * \frac{Q_D}{I_D})}{(1 + \frac{I_A}{I_M} + \frac{I_D}{I_M})} \qquad \textbf{(A2.11)}$$

Substituting $Y = I_A/I_M$, and $X = I_D/I_M$ Equation (A2.11) becomes:-

$$P = \frac{P_M + Y\,P_A + XP_D}{1 + X + Y} \qquad \textbf{(A2.12)}$$

Similarly, from Equation (A2.2), the colour ratio becomes:-

$$C = \frac{C_M + Y_2\,C_A + Y_2\,C_D}{1 + Y_2 + X_2} \qquad \textbf{(A2.13)}$$

Similarly, the colour ratio can be defined for the polarized component Q. This yields an additional, independent source of scientific information.

Appendix 3

Tabulated Summary of the 1992 Summer BLO Data Set

From June to August, 1992, inclusive, the following table, Table A3.1, summarizes the complete number of days available for data analysis, and interpretation.

Table A3.1

Date /92	Day Number	Met. Conditions		Lunar Appearance			
		APL *	MET **	Time / UT	EL / °	AZ / °	PHA / %
2/6/	153	CLR	CLR	5	-	-	-
				7	-	-	-
				9	-	-	-
3/6/	154	CLR	CLR	5	-	-	-
				7	-	-	-
				9	-	-	-
4/6/	155	CLR	CLR	5	4.4	290.8	12
				7	-	-	-
				9	-	-	-

5/6/	156	CLR	CLR	5	11.3	278.9	21
				7	-	-	-
				9	-	-	-
6/6/	157	OVC	OVC	5	17.5	266.4	31
				7	-	-	-
				9	-	-	-
8/6/	159	LUN	BRK	5	26.6	239.2	54
				7	5.9	260.8	55
				9	-	-	-
24/6/	175	LUN	SCT	5	-	-	-
				7	-	-	-
				9	17.0	91.8	40
25/6/	176	CLR	OVC	5	-	-	-
				7	-	-	-
				9	11.6	80.9	30
26/6/	177	CLR	BRK	5	-	-	-
				7	-	-	-
				9	5.4	70.3	21
27/6/	178	CLR	CLR	5	-	-	-
				7	-	-	-

				9	-	-	-
4/7/	185	CLR	CLR	5	1.3	275.6	19
				7	-	-	-
				9	-	-	-
5/7/	186	CLR	CLR	5	6.6	262.9	29
				7	-	-	-
				9	-	-	-
6/7/	187	CLR	OVC	5	11.3	250.3	40
				7	-	-	-
				9	-	-	-
8/7/	189	LUN	SCT	5	18.7	224.9	62
				7	0.6	246.6	63
				9	-	-	-
9/7/	190	LUN	CLR	5	21.2	212.0	72
				7	5.9	235.5	73
				9	-	-	-
18/8/	230	LUN	CLR	5	10.9	85.0	80
				7	32.7	105.2	80
				9	52.0	134.8	79
19/8/	231	LUN	SCT	5	5.6	74.8	72

				7	27.3	93.5	71
				9	48.5	117.4	70
21/8/	233	LUN	SCT	5	-	-	-
				7	13.2	72.9	51
				9	34.7	90.9	51
22/8/	234	OVC	OVC	5	-	-	-
				7	5.0	63.5	41
				9	25.5	81.0	40
23/8/	235	CLR	SCT	5	-	-	-
				7	-	-	-
				9	15.3	72.5	29
24/8/	236	CLR	CLR	5	-	-	-
				7	-	-	-
				9	4.4	64.8	20
25/8/	237	CLR	CLR	5	-	-	-
				7	-	-	-
				9	-	-	-
26/8/	238	CLR	CLR	5	-	-	-
				7	-	-	-
				9	-	-	-

27/8/	239	CLR	CLR	5	-	-	-
				7	-	-	-
				9	-	-	-
28/8/	240	CLR	CLR	5	-	-	-
				7	-	-	-
				9	-	-	-
29/8/	241	CLR	CLR	5	-	-	-
				7	-	-	-
				9	-	-	-

* APL as determined by the background intensity.

** MET courtesy of Dr D.Jensen, Director of Climatology, USU.

EL = elevation

AZ = azimuth

PHA = phase

BRK = breaking cirrus clouds

CLR = clear, cloud-free

LUN = lunar appearance

OVC = overcast, cloudy

SCT = randomly scattered cirrus

\- = no moon

Dates not tabulated above, 31st of July to the 15th of August,

1992, were unavailable as a result of a high photon count, in some cases greater than 12k. The thermo-electric cooler of the IPD had failed so that the IPD was operating at room temperature with a very high thermionic emission.

Bibliography

Born, M., and E.Wolf, "Principles of Optics",
Electromagnetic Theory of Propagation Interference, and
Diffraction of Light,
6th Edition, Pergammon Press, 1987.

Brasseur, G., and S.Soloman, "Aeronomy of the Middle Atmosphere",
Atmospheric Sciences Library,
2nd Edition, D.Reidel Publishing Co..

Dawson, A., "Global climate change",
1st Edition, Oxford University Press, 1992.

Gadsden, M., and W.Schroder, "Noctilucent Clouds",
Springer-Verlag, Berlin Heidelberg, 1989.

Hall, H.E., "Solid State Physics",
1st Edition, John Wiley and Sons, 1987.

Hanwell, J., "Atmospheric Processes",
1st Edition, Unwin Hyman, 1980.

Hargreaves, J.K., "The Upper Atmosphere and Solar Terrestrial
Relations",
An introduction to the aerospace environment,
Van Nostrand Reinhold International Student Edition, 1979.

Hecht, E., and A.Zajac, "Optics",
Addison-Wesley Publishing Company,
6th Edition, 1980.

Herandez, G., "Fabry-Perot Interferometers",
Cambridge Studies in Modern Optics 3, 1966.

Houghton, J.T., "The Physics of Atmospheres",
2nd Edition, Cambridge University Press, 1986.

Jones, M.H., "A Practical Introduction to Electronic Circuits",
2nd Edition, Cambridge University Press, 1985.

Kato, S., "Dynamics of the Upper Atmosphere",
Development in Earth and Planetary Sciences 01,
Center for Academic Publications Japan,
D.Reidel Publishing Co., 1980.

Lindzen., R.S., "Dynamics in Atmospheric Physics",
1st Edition, Cambridge University Press, 1990.

Meinel, A., and M.Meinel, "Sunsets, twilights, and evening
skies",
1st Edition, Cambridge University Press, 1991.

Pitt, V.H., "The Penguin Dictionary of Physics",
Penguin Publication, 1976.

Rawer, K., "Winds and Turbulence in Stratosphere, Mesosphere and Ionosphere",
North-Holland Publishing Co. Amsterdam, 1968.

Rees, M.H., "Physics and chemistry of the upper atmosphere",
Cambridge Atmospheric and Space Science Series,
1st Edition, Cambridge University Press, 1989.

Sze, S.M., "Physics of Semiconductor Devices",
2nd Edition, A Wiley-Interscience Publication, 1981.

Webb, W.L., "Structure of the Stratosphere and Mesosphere",
Vol. 9, International Geophysics Series,
Academic Press, 1966.

References

1) Chapman, S., 1931,
"Some phenomena of the upper atmosphere",
Bakerian Lecture, Proc.R.Soc., London Ser. A 132, p.353-374.

2) Goldberg, R., and G.Witt, 1977,
"Ion Composition in a Noctilucent Cloud Region",
J.Geophys.Res., Vol.82, No.19, p.2619-2627.

3) Gadsden, M., 1982,
"Noctilucent Clouds",
Space Science Reviews, 33, p.279-334.

4) Fogle, B., and B.Haurwitz, 1966,
"Noctilucent Clouds",
Space Science Reviews, 6, p.278-340.

5) Witt, G., 1962,
"Height, structure and displacements of noctilucent clouds",
Tellus 14, 1, p.1-18.

6) Fogle, B., 1964,
Geophys.Inst.Rep. No.UAG R-158, University Of Alaska.

7) Fogle, B., 1966,
Geophys.Inst.Rep. No.UAG R-177, University of Alaska.

8) Ludlam, F.H., 1957,
"Noctilucent Clouds",
Tellus 9, 3, p.341-364.

9) Grishin, N.I., 1955,
"On the structure of noctilucent clouds",
Meteorologiia i Gidrologiia Leningrad No.1, p.23-28,
(in Russian).

10) Backhouse, T.W., 1885,
"The luminous cirrus clouds of June and July",
Met.Mag., Vol.20, p.133.

11) Arago, F., 1854,
Complete Works, 4, p.73.

12) Scultetus, H.R., 1949,
"Altere Beobachtungen von Leuchtstreifen",
Z.fur Met. 3, p.272.

13) Jesse, O., 1896,
"Die Hohe der leuchtenden Nachtwolken",
Astron.Nachr., Kiel, Vol.140, p.161.

14) Astapovich, I.S., 1925,
"Noctilucent Clouds",
Izvest.Akad.Nauk.SSSR, Su.Geof.i.Geogr. 2, p.183-204,
(in Russian).

15) Stormer, C., 1933,
"Height and velocity of luminous night clouds in Norway 1932",
Oslo University Publ. 6, p.1-45.

16) Vestine, E.H., 1934,
"Noctilucent Clouds",
J.Roy.Astron.Soc.Can., 28, p.249-272, and p.303-317.

17) Paton, J., 1949,
"Luminous night clouds",
Met.Mag., Vol.78, p.354-357.

18) Grishin, N.I., 1956,
"Research of continuous spectrum of noctilucent clouds",
All Union Astr.Geod.Soc.Bull.19, p.3-16,
(in Russian).

19) Currie, B.W., 1962,
J.Roy.Astron.Soc.Can., 56, p.141.

20) Fiocco, G., and G.Grams, 1969,
"Optical Radar Observations of Mesospheric Aerosols in Norway during the Summer 1966",
J.Geophys.Res., Vol.74, No.10, p.2453-2458.

21) Grahn, S., and G.Witt, 1971,
Rep. AP-4, MISU.

22) Hansen, G., M.Serwazi, and U.von Zahn, 1989,
"First detection of a noctilucent cloud by lidar",
Geophys.Res.Lett., Vol.16, No.12, p.1445-1448.

23) Theon, J.S., W.Nordberg, and W.S.Smith, 1967,
"Temperature Measurements in Noctilucent Clouds",
Science, Vol.157, p.419-421.

24) Nordberg, W., and W.S.Smith, 1964,
"The rocket grenade experiment",
NASA TN D-2107.

25) Nordberg, W., L.Katchen, J.S.Theon, and W.S.Smith, 1965,
"Rocket observations of the structure of the mesosphere",
J.Atmos.Sci., Vol.22, p.611-622.

26) Taylor, M.J., M.A.Hapgood, and D.A.R.Simmons, 1984,
"The effect of atmospheric screening on the visible border of noctilucent clouds",
J.A.T.P., Vol.46, No.4, p.363-372.

27) Khvostikov, I.A., 1964,
"Upper Atmospheric Layers",
Leningrad, Gidrometeoizdat 606,
(in Russian).

28) Sharonov, V.V., 1960,
"Problems of noctilucent clouds climatology",
Meteorol.Issl. 6, p.5-22,
(in Russian).

29) Gadsden, M., 1975,
"Observations of colour and polarization of noctilucent clouds",
Ann.Geophys. 31, p.507-516.

30) Avaste, O.A., M.Gadsden, and G.M.Grechko, 1988,
"The coloured edge of noctilucent clouds",
J.A.T.P., Vol.50, No.7, p.591-599.

31) Witt, G., 1960a,
"Polarization of light from noctilucent clouds",
J.Geophys.Res., Vol.65, No.3, p.925-933.

32) Witt, G., 1960b,
A note to the paper by G.Witt,
"Polarization of light from noctilucent clouds",
J.Geophys.Res., Vol.65, No.7, p.2199-2202.

33) Willmann, Ch.I., 1962,
"On the polarization of light from noctilucent clouds",
Trudy Sovech.Sereb.Oblakam., III, p.29-54 (Tallinn),
(in Russian).

34) Tarasova, T.M., 1962,
"Polarization of the light flux of noctilucent clouds",
Trudy Sovech.Sereb.Oblakam., III, p.55-67 (Tallinn),
(in Russian).

35) Vasilyev, O.B., 1962,
"Results of the absolute photometry and polarimetry of noctilucent clouds",
Trudy Sovech.Sereb.Oblakam., III, p.14-28 (Tallinn),
(in Russian).

36) Stormer, C., 1935,
"Measurement of luminous night clouds in Norway 1933 and 1934",
Astrophysica Norwegica 1, p.87-114,

37) Fogle, B., and Y.Gotaas, 1963,
Geophys.Inst.Rep. No. UAG R-157, University of Alaska.

38) Fogle, B., 1965c,
"Noctilucent clouds over Punta Arena, Chile",
Nature (London), Vol.207, p.66.

39) Schilling, G.F., 1964,
"Forbidden Regions for the Formation of Clouds in a Planetary Atmosphere",
J.Geophys.Res., Vol.69, No.17, p.3663-3667.

40) Thayaparan, T., 1990,
"Studies of Polar Mesosphere Summer Echoes by high resolution EISCAT VHF radar and ground-based optical measurements",
Thesis of Master of Science, University of Oslo.

41) Khvostikov, I.A., 1952,
"Noctilucent Clouds",
Priroda, Moscow, No.5, p.49,
(in Russian).

42) Burov, M.I., 1959,
"Methode photogrammetrique pour la determination de l'altitude des nuages argentes",
Trudy ovech.Sereb.Oblakam., I, p.92-111 (in Tallinn),
(in Russian).

43) Burov, M.I., 1966,
"Determination of parameters of noctilucent clouds",
Meteorologicheskiye Issledovaniya, 12, p.33-46,
(in Russian).

44) Dirikis, M.A., S.V.Evdokimenko, and J.L.Francmann, 1966,
"Determination of space coordinates of noctilucent clouds",
Meteorologicheskiye Issledovaniya, 12, p.47-51,
(in Russian).

45) Gadsden, M., and W.Schroder, 1989,
"Noctilucent Clouds",
Springer-Verlag New York Berlin Heidelberg.

46) Bessonova, T.D., 1963,
"Research on the climatology of Noctilucent Clouds",
Article No.6, p.23-63,
Academy of Sciences, USSR,
(in Russian).

47) Sharonov, V.V., 1965,
"Aurorae and Airglow",
No.11, p.48, USSR Publishing House "Nauka", Moscow.

48) Meinel, A.B., B.Middlehurst, and E.Whitaker, 1963,
"Low latitude noctilucent cloud of 15 June 1963",
Science, Vol.141, p.1176-1178.

49) World Meteorological Organization, 1966,
International Noctilucent Cloud Observation Manual, p.1-43.

50) Taylor, M.J., 1986,
"TV Observations of mesospheric wave structures",
in Collection of Works of the International Workshop of Noctilucent Clouds, edited by O.A.Avaste, p.204, and p.153-172,
Academy of Sciences of Estonian SSR, USSR.

51) Theon, J.S., W.S.Smith, and W.E.McGovern, 1969,
"Wind Measurements in Noctilucent Clouds",
Science, Vol.164, p.715-716.

52) Schmidlin, F.J., H.S.Lee, and W.Michel, 1991,
"The Inflatable Sphere A:A Technique for the Accurate Measurement of Middle Atmosphere Temperatures",
J.Geophys.Res., Vol.96, No.D12, p.22,673-22,682.

53) Gavine, D., 1987,
"Noctilucent clouds over Western Europe during 1986",
Met.Mag., Vol.116, p.386-388.

54) Fogle, B., 1964a,
"Noctilucent clouds in the Southern Hemisphere",
Nature, Vol.204, p.14-18.

55) Fogle, B., S.Chapman, and C.Echols, 1965,
Geophys.Inst.Rep. No.UAG R-162, University of Alaska.

56) Avakyan, S., O.A.Avaste, Ch.I.Willmann, V.Kovalyonok, A.Lazarev, and V.Savinykh, 1986,
"NLC Observations in Equatorial and Low Latitudes Carried Out by Main Team of the 5th Expedition of the Orbital Station "Salyut-6",
in Collection of Works of the International Workshop of Noctilucent Clouds, edited by O.A.Avaste, p.121-130,
Academy of Sciences of the Estonian SSR, USSR.

57) Hesstvedt, E., 1961,
"Note on the nature of noctilucent clouds",
J.Geophys.Res., Vol.66, No.6, p.1985-1987.

58) Hesstvedt, E., 1962,
"On the possibility of ice clouds formation at the mesopause",
Tellus 14, 3, p.290-296.

59) Hesstvedt, E., 1964,
"On the water content of the high atmosphere",
Geofys.Publ. 25, p.1-18.

60) Hesstvedt, E., 1969b,
"Noctilucent clouds observations and deduction:water vapour in the stratosphere and mesosphere",
Ann.IQSY., Vol.1969, p.23.

61) Turco,R.P., O.B.Toon, R.C.Whitten, R.G.Keese, and D.Hollenbach, 1982,
"A study of mesospheric rocket contrails and clouds produced by liquid-fuelled rockets",
Solar Power Review.

62) Vestine, E.H., and D.Deirmendijan, 1961,
"Some remarks on the nature and origin of noctilucent cloud particles",
Ann.IGY., Vol.11, p.6-13.

63) Witt, G., 1969,
"The nature of noctilucent clouds",
Space Res. IX, p.157-169,
North Holland Publishing Comp., Amsterdam.

64) Hemenway, C.L., R.K.Soberman, and G.Witt, 1964,
"Investigations of noctilucent cloud particles",
Tellus 16, p.84-88.

65) Thomas, G.E., 1991,
"Mesospheric clouds and the physics of the mesopause region",
Reviews of Geophys., Vol.29, No.4, p.553-575.

66) Roddy, A.F., 1986,
"The Physics of Noctilucent Cloud Formation",
in Collection of Works of the International Workshop of Noctilucent Clouds, edited by O.A.Avaste, p.33-57,
Academy of Sciences of the Estonian SSR, USSR.

67) Bjorn, L.G., E.Kopp, U.Herrmann, P.Eberhardt, P.H.G.Dickinson, D.J.Mackinnon, F.Arnold, G.Witt, A.Lundin, and D.B.Jenkins, 1985,
"Heavy Ionospheric Ions in the Formation Process of Noctilucent Clouds",
J.Geophys.Res., Vol.90, No.D5, p.7985-7998.

68) Arnold, F., 1980,
"Ion-induced nucleation of atmospheric water vapour at the mesopause",
Planet.Space.Sci., Vol.28, p.1003-1009.

69) Bjorn, L.G., and F.Arnold, 1981,
"Mass spectrometric detection of pre-condensation nuclei at the arctic summer mesopause",
Geophys.Res.Lett., Vol.8, No.11, p.1167-1170.

70) Kohlrausch, 1887,
W.Zur.Hohe.der Wolken,
Ann.d.Physik u.Chemie,
Neue Folge, Vol.31, p.1047-1048.

71) Foerster, W., 1906,
Von der Erdatmosphare zum Himmelsraum,
Leipzig, Germany.

72) Vasilyev, O.B., Zhuravlev et al., 1965,
"Nocturnal Luminous Clouds and Optical Anomalies Connected with the Tungusska Meteorite",
USSR Publishing House "Nauka", Moscow.

73) Wegner, A., 1926,
Zusatz.Met.Z., Braunschweig, Vol.43, p.103.

74) Spangenberg, W.W., 1949,
"Uber die leuchtenden Nachtwolken 1932-1941",
Wetter und Klima, 2, p.15.

75) Fessenkov, V.G., 1949,
"The mass of the atmospheric residue of the Sikhote-Alin meteorite",
Dok.Akad.Nauk. SSSR, Vol.66,
Translation T 133 R,
Defence Sci.Inf.Service, Canada, 1954.

76) Fogle, B., 1965b,
"Noctilucent clouds over North America",
Nature (London), Vol.207, No.4998, p.696-698.

77) Simmons, D.A.R., and D.A.McIntosh, 1983,
"An analysis of NLC over Western Europe during the period 1966 to 1982",
Met.Mag., Vol.112, p.289-298.

78) Paton, J., 1967,
"Noctilucent Clouds over Western Europe during 1967",
Met.Mag., Vol.97, p.174-176.

79) Thomas, G.E., R.D.McPeters, and E.J.Jensen, 1991,
"Satellite observations of polar mesospheric clouds by the SBUV radiometer:Evidence of solar-cycle dependence",
J.Geophys.Res., Vol.96, No.D1, p.927-939.

80) Gadsden, M., 1990,
"A secular change in noctilucent cloud occurrence",
J.A.T.P., Vol.52, No.4, p.247-251.

81) Garcia, R.R., S.Soloman, R.G.Roble, and D.W.Rusch, 1984,
"A numerical response of middle atmosphere to the 11 year solar cycle",
Planet.Space.Sci., Vol.32, No.4, p.411-423.

82) Thomas, G.E., J.J.Olivero, E.J.Jensen, W.Schroder, and O.B.Toon, 1989,
"Relation between increasing methane and the presence of ice clouds at the mesopause",
Nature, Vol.338, p.490-492.

83) Lubken, F-J., and U.von Zahn, 1991,
"Thermal Structure of the Mesopause Region at Polar Latitudes",
J.Geophys.Res., Vol.96, No.D11, p.20841-20857.

84) Clemesha, B.R., D.M.Simonich, and P.P.Batista, 1992,
"A long-term trend in the height of the atmospheric sodium layer:Possible evidence of global change",
Geophys.Res.Lett., Vol.19, No.5, p.457-460.

85) Fels, S.B., J.D.Mahlman, M.D.Schwarzkopf, and R.W.Sinclair, 1980,
"Stratospheric sensitivity to perturbations in ozone and carbon dioxide:Radiance and dynamical response",
J.Atmos.Sci., Vol.37, p.2265-2297.

86) Roble, R.G., and R.E.Dickinson, 1989,
"How will changes in carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere?",
Geophys.Res.Lett., Vol.16, No.12, p.1441-1444.

87) Rind, D., R.Suozzo, N.K.Balachandran, and M.J.Prather, 1990,
"Climate change and the middle atmosphere, Part 1, The coupled CO_2 climate",
J.Atmos.Sci., Vol.47, p.475-494.

88) Gadsden, M., and M.J.Taylor, 1992,
"Anweisungen fur die photographischen aufnahmen der leuchtenden nachtwolken - 102 years ago",
J.A.T.P., in press.

89) Hines, C.O., 1991,
"Noctilucent and mesospheric clouds",
paper presented at the Noctilucent Cloud Symposium,
Int.Union of Geod. and Geophys.Gen.Assem., Vienna, Austria.

90) Shefov, N.N., 1968,
"The Behaviour of Upper Atmosphere Emissions During High Meteoric Activity",
Planet.Space.Sci., Vol.16, p.134-136.

91) Harrison, A.W., 1973,
"Spectrophotometric Measurements of Noctilucent Clouds",
Can.J.Phys., Vol.51, No.4, p.373-377.

92) Taylor, M.J., A.P.von Eyken, H.Risbeth, G.Witt, N.Witt, and M.A.Cliverd, 1989,
"Simultaneous observations of noctilucent clouds and polar mesospheric radar echoes:Evidence of non-correlation",
Planet.Space.Sci., Vol.37, No.8, p.1013-1020.

93) Smyth, C.P., 1886,
"The silver-blue cloudlets again",
Nature (London), Vol.34, p.311-312.

94) Backhouse, T.W., 1886,
"The bright clouds and the aurora",
Nature (London), Vol.34, p.386-387.

95) Paton, J., 1951,
Letter to the editor,
"Simultaneous occurrence of aurora and luminous night clouds",
Met.Mag., Vol.80, p.145.

96) Bryne, F.D., 1964,
"Noctilucent Cloud and Aurora",
Met.Mag., Vol.93, p.121-122.

97) Schroder, W., 1965,
"Simultaneous sightings of aurora and noctilucent clouds",
Gerlands.Beitr.Geophys., Vol.74, p.471-473.

98) Murcray, W.B., 1957,
"A Possible Auroral Enhancement of Infra-red Radiation Emitted by Atmospheric Ozone",
Nature (London), Vol.180, No.4577, p.139-140.

99) Campbell, W.H., and M.H.Rees, 1961,
"A Study of Auroral Coruscations",
J.Geophys.Res., Vol.66, No.1, p.41-55.

100) Chamberlain, J.W., 1961,
"Physics of the Aurora and Airglow",
Acad.Press., NY, p.394-413.

101) Meriwether, J.W., 1991,
"Noctilucent cloud observations at Sondestrom, Greenland",
paper presented at the Noctilucent Cloud Symposium,
Int.Union of Geod. and Geophys.Gen.Assem., Vienna, Austria.

102) Donahue, T.M., B.Guenther, and J.E.Blamont, 1972,
"Noctilucent clouds in daytime:Circumpolar particulate layers near the summer mesopause",
J.Atmos.Sci., Vol.29, p.1205-1209.

103) Jensen, E.J., and G.E.Thomas, 1988,
"A growth-sedimentation model of polar mesospheric clouds:Comparisons with SME measurements",
J.Geophys.Res., Vol.93, No.D3, p.2461-2473.

104) Olivero, J.J., and G.E.Thomas, 1986,
"Climatology of polar mesospheric clouds",
J.Atmos.Sci., Vol.43, p.1263-1274.

105) Thomas, G.E., and J.J.Olivero, 1986,
"The heights of polar mesospheric clouds",
Geophys.Res.Lett., Vol.13, No.13, p.1403-1406.

106) Thomas, G.E., R.D.McPeters, and E.J.Jensen, 1991,
"Satellite observations of polar mesospheric clouds by the SBUV radiometer:Evidence of solar-cycle dependence",
J.Geophys.Res., Vol.96, No.D1, p.927-939.

107) Lee, H.S., and S.M.Sze, 1970,
"Silicon p-i-n Photodetector Using Internal Reflection Methods"
IEEE Trans.Electron Devices, ED-17, p.342.

108) Muller, J., 1978,
"Thin Silicon Film p-i-n Photodiodes with Internal Reflection",
IEEE Trans.Electron Devices, ED-25, p.247.

109) Thomas, G.E., 1984,
"Solar Mesosphere Explorer measurements of polar mesospheric clouds (noctilucent clouds)",
J.A.T.P., Vol.46, No.9, p.819-824.

110) Witt, G., 1968,
"Optical characteristic of mesospheric aerosol distributions in relation to noctilucent clouds",
Tellus 20, 1, p.98-114.

111) Bjorn, L.G., 1984,
"The cold summer mesopause",
Adv.Space.Sci., Vol.4, No.4, p.145-151.

112) Kofman, W., F.Bertin, J.Rottger, A.Cremieux, and P.J.S.Williams, 1984,
"The EISCAT mesospheric measurements during the CAMP campaign",
J.A.T.P., Vol.46, No.6/7, p.565-575.

113) Schmidlin, F.J., 1992,
"First observation of mesopause temperatures lower than 100 K",
Geophys.Res.Lett., in press.

114) Garcia, R.R., and S.Soloman, 1985,
"The Effect of Breaking Gravity Waves on the Dynamics and Chemical Composition of the Mesosphere and Lower Thermosphere",
J.Geophys.Res., Vol.90, No.D2, p.3850-3868.

115) Nastrom, G.D., B.B.Balsley, and D.A.Carter, 1982,
"Mean meridional winds in the mid- and high-latitude summer mesosphere",
Geophys.Res.Lett., Vol.9, No.2, p.139-142.

116) East, S.A., 1992,
"The Preliminary Results Obtained From The APL Photometers - NLC Characteristics and Composition",
NLC-91 Workshop, NASA/GSFC, Greenbelt, MD., 28-29 August 1992.

117) Hale, L., 1992,
Modelled Results of the Electric Fields in the NLC During The NLC-91 Rocket Campaign,
NLC-91 Workshop, NASA/GSFC, Greenbelt, MD., 28-29 August 1992.

118) Espy, P.J., C.R.Harris, A.J.Steed, J.C.Ulwick, R.H.Haycock, and R.Straka, 1988,
"Rocket-borne interferometer measurements of infrared auroral spectra",
Planet.Space.Sci., Vol.36, No.6, p.543-551.

119) Lowe, R.P., K.L.Gilbert, and D.N.Turnbull, 1991,
"High latitude summer observations of the hydroxyl airglow",
Planet.Space.Sci., Vol.39, p.1263.

120) Mies, F.H., "Calculated vibrational transition probabilities of $OH(X^2\pi)$",
J.Molec.Spectrosc., Vol.53, p.150-188.

121) Turnbull, D.N., and R.P.Lowe, 1983,
"Vibrational population distribution in the hydroxyl night airglow",
Can.J.Phys., Vol.61, p.244-250.

122) Pendleton, W.R., P.J.Espy, J.Hartman, M.Fetrow, and G.Sivjee, 1989,
"[NI] 3F ($^2P^\circ$ - $^2D^\circ$) Emissions in the aurora",
J.Geophys.Res., Vol.94, No.A3, p.2595-2604.

123) Shefov, N.N., 1961,
"On determination of the rotational temperatures OH bands",
in "Spectral, Electrophotometrical, and Radar Researches of Aurora and Airglow", Edited by V.I.Krassvosky,
Academy of Sciences, Moscow, p.1-7.

124) Pendleton, W.R., P.J.Espy, D.Baker, A.Steed, M.Fetrow, and K.Henrikson, 1989,
"Observation of OH Meinel (7,4) $P_{1,2}$(N"=13) transitions in the Night Airglow",
J.Geophys.Res., Vol.94, No.A1, p.505-510.

125) Gadsden, M., 1981,
"The silver-blue cloudlets again: nucleation and growth of ice in the mesosphere",
Planet.Space.Sci., Vol.29, No.10, p.1079-1087.

126) Jensen, E.J., G.E.Thomas, and B.B.Balsley, 1988,
"On the statistical correlation between polar mesospheric cloud occurrence and enhanced mesospheric radar echoes",
Geophys.Res.Lett., Vol.15, No.4, p.315-318.

127) von Zahn, U., and W.Meyer, 1989,
"Mesopause temperatures in polar summer",
J.Geophys.Res., Vol.94, No.D12, p.14647-14651.

128) Philbrick, C.R., J.Barnett, R.Grendt, D.Offermann, W.R.Pendleton, P.Schlyter, J.F.Schmidlin, and G.Witt, 1984, "Temperature measurements during the CAMP program", Adv.Space.Sci., Vol.4, No.4, p.153-156.

129) Fritts, D.C., S.A.Smith, B.B.Balsley, and C.R.Philbrick, 1988, "Evidence of gravity wave saturation and local turbulence production in the summer mesosphere and lower thermosphere during the STATE experiment", J.Geophys.Res., Vol.93, p.3850.

130) Arnold, F., and D.Krankowsky, 1977, "Water vapour concentrations at the mesopause", Nature (London), Vol.268, p.218-219.

131) Baker, D.J., and A.T.Stair, 1988, "Rocket Measurements of the Altitude Distributions of the Hydroxyl Airglow", Physica Scripta, Vol.37, p.611-622.

132) Espy, P.J., and G.Witt, 1991, "Climatology of high-latitude, mesospheric hydroxyl during the summer noctilucent cloud season", EOS 72, 374, paper SA51A-1, Fall AGU.

133) Shefov, N.N., 1967, Proceedings of the International Symposium on Noctilucent Clouds, Tallinn, p.186.

134) von Zahn, U., 1990, "Temperature and altitude of polar mesosphere in summer", Adv.Space.Sci., Vol.10, No.12, p.12233-12231.

135) Rosenfeld, Shk., 1986, "Relationship between observations of noctilucent clouds and internal gravity waves", Collection of Works of the International Workshop of Noctilucent Clouds, Edited by O.A.Avaste, Academy of Sciences of the Estonian SSR, Tallinn, Valgus, p.173.

136) Jensen, E.J., G.E.Thomas, O.B.Toon, 1989, "On the diurnal variation on noctilucent clouds", J.Geophys.Res., Vol.94, No.D12, p.14693-14702.

137) Andrews, D.G, J.R.Holton, and C.B.Leovy, 1987, "Middle Atmospheric Dynamics", Academic Press, San Diego, CA.

138) Vial, F., 1989, "Tides in the middle atmosphere", J.A.T.P., Vol.51, No.1, p.3-17.

139) Myabo, H.K., and O.E.Harang, 1988,
"Temperatures and tides in the high latitude mesopause region as observed in the OH night airglow emissions",
J.A.T.P., Vol.50, No.8, p.739-748.

140) Vincent,R.A., 1990,
"Planetary and gravity waves in the mesosphere and lower thermosphere",
Adv.Space.Sci., Vol.10, No.12, p.1293.

141) Houghton, J.T., 1986,
"The Physics of Atmospheres",
Second Edition, Cambridge University Press.

142) Thomas, G.E., 1991,
"Recent Developments in the Study of Mesospheric Clouds",
The World Space Congress, IAF/COSPAR, Washington DC, 28th August - 5th September 1992, paper C2-S308,
J.A.T.P., in press.

143) Carter, D.A., and B.B.Balsley, 1982,
"The summer wind field between 80 and 93 km observed by the MST radar at Poker Flat, Alaska (65°N)",
J.Atmos.Sci., Vol. 39, p.2905-2915.

144) Bevilacqua, R.M., W.J.Wilson, and P.R.Schwartz, 1987,
"Measurements of mesospheric water vapour in 1984, and 1985: Results and implications for middle atmospheric transport",
J.Geophys.Res., Vol.92, No.D6, p.6679-6690.

145) Rees, D., I.McWhirter, P.A.Rounce, and S.J.Kellock, 1980,
"Miniature imaging photon detectors",
J.Phys.E: Sci.Instrum., Vol.13, p.763-770.

146) Rees, D., I.McWhirter, P.A.Rounce, and F.E.Barlow, 1981,
"Miniature imaging photon detectors II. Devices with transparent photocathodes",
J.Phys.E: Sci.Instrum., Vol.14, p.229-233.

147) McWhirter, I., D.Rees, and A.H.Greenaway, 1982,
"Miniature imaging photon detectors III. An assessment of the performance of the resistive anode IPD",
J.Phys.E: Sci.Instrum., Vol.15, p.145-150.

148) Killeen, T.L., B.C.Kennedy, P.B.Hays, D.A.Symanow, and D.H.Ceckowski, 1983,
Image plane detector for the Dynamics Explorer Fabry-Perot interferometer",
Applied Optics, Vol.22, p.3503-3513.

149) Rees, D., P.A.Rounce, I.McWhirter, A.F.D.Scott, A.H.Greenaway, and W.Towlson, 1981,
"Observations of atmospheric absorption lines from stabilised balloon platform and measurements of stratospheric winds",
J.Phys.E: Sci.Instrum., Vol.15, p.191-206.

150) Rees, D., P.A.Rounce, P.Charleton, T.J.Fuller-Rowell, I.McWhirter, and K.Smith, 1982, "Ground-based Fabry-Perot observations supported by dynamical simulations with a three-dimensional, time-dependent thermospheric model", J.Geophys., Vol.50, p.202-211.

151) Smith, R.W., D.Rees, F.G.McGormac, and P.Charleton, 1986, "Two-station observations of thermospheric winds in the auroral zone and the polar cap", J.A.T.P., Vol.48, No.1, p.97-105.

152) Winser, K.J., A.D.Farmer, D.Rees, and A.L.Aruliah, 1988, "Ion-neutral dynamics in the high latitude ionosphere: First results from the INDI experiment", J.A.T.P., Vol.50, No.4/5, p.369-377.

153) Rees, D., A.H.Greenaway, R.Gordon, I.McWhirter, P.J.Charleton, and A.Steen, 1984, "The Doppler Imaging System: Initial observations of the auroral thermosphere", Planet.Space.Sci., Vol.32, No.3, p.273-285.

154) Batten, S., D.Rees, D.Wade, and A.Steen, 1988, "Observations of thermospheric neutral winds by the UCL Doppler Imaging System at Kiruna in Northern Scandinavia", J.A.T.P., Vol.50, No.10/11, p.861-888.

155) Csorba, I.P., 1979, "Recent advancements in the field of image intensification: the generation of a 3 wafer tube", Applied Optics, Vol.18, p.2440-2444.

156) Lampton, M., and C.W.Carlson, 1979, "Low distortion resistive anodes for two-dimensional, position-sensitive MCP systems", Rev.Sci.Instrum., Vol.50., p.1093-1097.

157) Meredith, N.P., 1990, "Narrow-Band Imaging and Doppler Imaging of Neutral and Artificial Gas and Plasma Clouds in the Interplanetary Medium and in the Earth's Magnetosphere", Ph.D. thesis, University College London.

158) Born, M., and E.Wolf, 1987, "Principles of Optics", Electromagnetic Theory of Propagation Interference, and Diffraction of Light, Sixth Edition, Pergammon Press.

159) Cooper, V.G., 1971, "Analysis of Fabry-Perot interferograms by means of their Fourier transforms", Applied Optics, Vol.10, No.3, p.525-530.

160) Batten, S.M., 1989,
"The Measurement of Neutral Winds in the Thermosphere using a Doppler Imaging System",
Ph.D. thesis, University College London.

161) Koopmans, L.H., 1974,
"The Spectral Analysis of Time Series",
Academic Press.

162) Aruliah, A.L., 1991,
"The Synoptic Variability of Thermospheric and Mesospheric Winds Observed Using a Fabry-Perot Interferometer",
Ph.D. thesis, University College London.

163) Aruliah, A.L., and D.Rees, 1993,
"Problems with Thermospheric Vertical Winds: Geomagnetic, Seasonal and Solar Cycle Dependence",
J.A.T.P., in press.

164) Meredith, N.P., 1991,
"The Study of the Effect of Photon Noise, and Number of Peaks on the Fits to the Fabry-Perot Fringes",
APL - Internal Report.

165) Muller, H., 1970,
"The Sheffield meteor wind experiment",
Quart.J.Roy.Met.Soc., Vol.96, No.408, p.195-197.

166) Barnett, J.J., 1980,
Satellite measurements of Middle Atmosphere temperature structure",
Phil.Trans.Roy.Soc.Lond., Vol.A296, p.41-57.

167) Brown, G.M., and J.I.John, 1979,
"Vertical penetration of planetary waves into the lower ionosphere",
J.A.T.P., Vol.41, p.379-385.

168) Cavalieri, D.J., R.J.Deland, T.A.Plemra, and R.F.Gavin, 1974,
"The correlation of VLF propagation variations with atmospheric planetary - scale waves",
J.A.T.P., Vol.36, p.561-574.

169) Fraser, G.J., 1977,
"The 5-day wave and ionospheric absorption",
J.A.T.P., Vol.39, p.121-124.

170) Vincent, R.A., 1984,
"MF/HF radar measurements of the dynamics of the mesopause - A review",
J.A.T.P., Vol.46, No.11, p.961-970.

171) Hirota, I., Y.Mackawa, S.Fukao, K.Fuykuyama, M.P.Saltzer, J.L.Fellous, T.Tsuda, and S.Kato, 1983,
"Fifteen-Day Observation of Mesospheric and Lower Thermospheric Motions with the Aid of the Arecibo UHF Radar",
J.Geophys.Res., Vol.88, No.C11, p.6835-6842.

172) Rodgers, C.D., and A.J.Prata, 1981,
"Evidence for a Travelling Two-Day Wave in the Middle Atmosphere",
J.Geophys.Res., Vol.86, No.C10, p.9661-9664.

173) Craig, R.L., and W.G.Elford, 1981,
"Observations of the quasi 2-day wave near 90 km altitude at Adelaide (35°S)",
J.A.T.P., Vol.43, No.10, p.1051-1056.

174) Craig, R.L., R.A.Vincent, S.P.Kingsley, and H.G.Muller, 1983,
"Simultaneous observations of the quasi 2-day wave in the Northern and Southern hemisphere",
J.A.T.P., Vol.45, No.8/9, p.539-541.

175) Lindzen, R.S., 1967,
"Thermally Driven Diurnal Tide in the Atmosphere",
Proc.Roy.Soc., Vol.A93, p.18-42.

176) Lindzen, R.S., 1966,
"On the Relation of Wave Behaviour to Source Strength and Distribution in a Propagating Medium",
J.Atmos.Sci., Vol.23, p.630-632.

177) Butler, S.T., and K.A.Small, 1963,
"The Excitation of Atmospheric Oscillations",
Proc.Roy.Soc., Vol.A274, p.91.

178) Leovy, C., 1964,
"Radiative Equilibrium in the Mesosphere",
J.Atmos.Sci., Vol.21, p.238.

179) Siebert, M., 1961,
"Atmospheric Tides",
Adv.Geophys., Vol.7, p.105.

180) Chapman, S., and R.S.Lindzen, 1970,
"Atmospheric Tides",
D.Riedel, Dordrecht.

181) Flattery, T.W., 1967,
"Hough Functions",
Tech.Rep.21, I11
Dep. of Geophys.Sci.,
Univ. of Chicago.

182) Hargreaves, J.K., 1979,
"The Upper Atmosphere and Solar Terrestrial Relations",
An introduction to the aerospace environment,
Van Nostrand Reinhold Co. Ltd..

183) Forbes, J.M., and M.E.Hagan, 1982,
"Thermospheric Extensions of the Classical Expansion Functions for Semidiurnal Tides",
J.Geophys.Res., Vol.87, No.A7, p.5253-5259.

184) Lindzen, R.S., and S.S.Hong, 1976,
"Solar Semi-Diurnal Tide in the Thermosphere",
J.Atmos.Sci., Vol.28, p.275-280.

185) Forbes, J.M., 1982a,
"Atmospheric Tides 1. Model Description and Results for Solar Diurnal Component",
J.Geophys.Res., Vol.87, No.A7, p.5222-5240.

186) Forbes, J.M., 1982b,
"Atmospheric Tides 2. The Solar and Lunar Semidiurnal Components",
J.Geophys.Res., Vol.87, No.A7, p.5241-5252.

187) Forbes, J.M., 1987b,
Handbook for MAP, Vol.23.

188) Lindzen, R.S., and S.S.Hong, 1974,
"Effects of Mean Winds and Horizontal Temperature Gradients on Solar and Lunar Semi-Diurnal Tides in the Atmosphere",
J.Atmos.Sci., Vol.31, p.1421-1466.

189) Walterscheid, R.L., J.G.Devore, and S.V.Venkateswaran, 1980,
"Influence of Mean Zonal Motion and Meridional Temperature Gradients on the Solar Semi-Diurnal Atmospheric Tide: A Revised Spectral Study With Improved Heating Rates",
J.Atmos.Sci., Vol.37, p.455-470.

190) Aso, T., T.Nonoyama, and S.Kato, 1981,
"Numerical Simulation of Semi-Diurnal Atmospheric Tides",
J.Geophys.Res., Vol.86, No.A1, p.388-400.

191) Fesen, C.G., R.E.Dickinson, and R.G.Roble, 1986,
"Simulation of the Thermospheric Tides at Equinox with the National Center for Atmospheric Research Thermospheric General Circulation Model",
J.Geophys.Res., Vol.91, No.A4, p.4471-4489.

192) Forbes, J.M., 1987a,
"Modelling the Propagation of Atmospheric Tides from the Lower to the Middle and Upper Atmosphere",
Physica Scripta, Vol.T18, p.240-248.

193) Forbes, J.M., and M.E.Hagan, 1979,
"Tides in the Joint Presence of Friction and Rotation: A f-Plane Approximation",
J.Geophys.Res., Vol.84, No.A3, p.803-810.

194) Vial, F., 1986,
"Numerical Simulations of Atmospheric Tides for Solstice Conditions",
J.Geophys.Res., Vol.91, No.A8, p.8955-8969.

195) Wilkes, M.V., 1949,
"Oscillations of the Earth's Atmosphere",
Cambridge University Press.

196) Lindzen, R.S., 1970,
"Internal Gravity Waves in Atmospheres With Realistic Dissipation and Temperature, I. Mathematical Development and Propagation of Waves into the Thermosphere",
Geophys.Astrophys.Fluid.Dyn., Vol.1, p.303-355.

197) Lindzen, R.S., 1971,
"Tides and Gravity Waves in the Upper Atmosphere",
Mesospheric Models and Related Experiments, p.122-130,
Edited G.Fiocco, D.Reidel Publishing Co..

198) Meinel, A.B., 1950,
"OH Emission Bands in the Spectrum of the night-sky",
Astrophys.J., Vol.111, No.1, p.434-444.

199) Krasovskij, V.I., and N.N.Shefov, 1965,
"Airglow",
Space Science Reviews, 4, p.176-198.

200) Kato, S., 1980,
"Dynamics of the Upper Atmosphere",
Development in Earth and Planetary Sciences 01,
Center for Academic Publications Japan,
D.Reidel Publishing Co..

201) Shepherd, G.G., G.Thuillier, W.A.Gault, B.H.Solheim, C.Hersom, J.M.Alunni, J.-F.Brun, S.Brune, P.Charlot, L.L.Cogger, D.-L.Desaulniers, W.F.J.Evans, F.Girod, D.Harvie, R.H.Hum, D.J.W.Kendall, E.J.Llewellyn, R.P.Lowe, J.Ohrt, F.Pasternak, O.Peillet, I.Powell, Y.Rochon, W.E.Ward, R.H.Wiens, and J.Wimperis, 1993,
"WINDII-The Wind Imaging Interferometer on the Upper Atmosphere Research Satellite",
in press.

202) Chandrasekhar, S., 1960,
"Radiative transfer",
Dover Publishing Inc., New York.

203) Wood, and G.Witt, 1978,
Rep AP-9, MISU.

204) Penndorf, R., 1952,
"Tables of the Refractive Index for Standard Air, and the Rayleigh Scattering Coefficient for the Spectat Region between 0.2 and 20.0 μ and their Application of Atmospheric Optics",
J.Opt.Soc.Am., Vol.47, p.47.

205) Witt, G., J.Dye, and N.Wilhelm, 1973,
"Measurement of scattered sunlight in the mesosphere during twilight",
Rep AP-8, MISU.

Acknowledgements

Looking for inspiration, information and help during the compilation of this thesis, I've met some really brillo people. Bear with me if I don't manage to mention all you wonderful people - early case of senile dementia, thats my excuse, and I'm sticking to it!! Whatever, thanks everyone. In the words of the great poet, "I owe you one."

I would like to a say a mega thank you to my mother, who encouraged me all the way. I would like to express my deep gratitude to Dr Dai Rees for providing me with the opportunity to travel abroad and take an active role in the NLC-91 Rocket Campaign. Cheers! I am also deeply grateful to Professor Georg Witt at Arrhenius Laboratory, Stockholm University, for numerous helpful discussions. Tack. I would also like to express a mega thank you to Dr Patrick Espy at Space Dynamics Laboratory, Utah State University, for all his extensive help and time with the Bomem OH data. Thanks Pat!

I would like to thank all the people who have assisted me in the field. I would like to express my appreciation to Dr Dick Goldberg and his assistant Patti Twigg, the resident data cruncher at Goddard Space Flight Center, and Frank Schmidlin. I would also like to thank Per-Erik Olsen from the Swedish Space Corporation for his help in locating long forgotten campaign documents. And a big thank you to Heinz Muller at Sheffield University. Next round at the B-Hive is on me!

Last but not least, I would like to express my appreciation to all at the Atmospheric Physics Laboratory who have helped me. In particular, I would like to mention Andrew Scott for his guidance and helpful comments; Jim Percival and the workshop gang for their help building and assembling the various photodetecting systems mentioned in this work; Conrad Winchester, Daniel Wade and John Harmer for software support; Nigel Meredith for proof reading all re-writes and missing "Brookside" to do so - merky buckets Nig!; Simon Mugleston for putting up with my black cloud days, and my "mad half hours" in our office - Cheers Plug!!; Dave Rooks for reprinting tons of photographs.

Also, I would like to thank S.Hall for collecting some obscure references.

My thesis was supported by grants from the UK Science and Engineering Research Council. Thanks in particular to Allan Webb.