# Infrared Limb Observations of Cloud

Submitted for the degree of Doctor of Philosophy
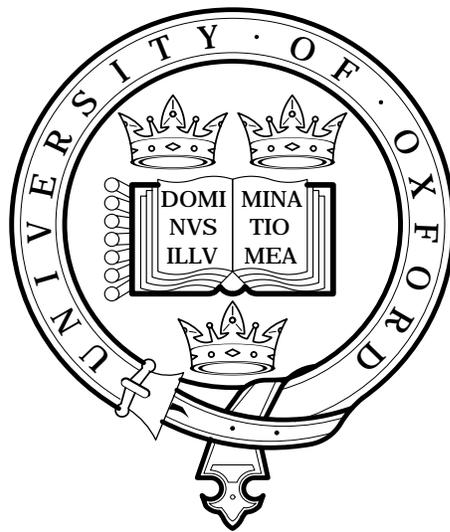
Graham Bruce Lachlan Ewen

Balliol College, University of Oxford

Trinity Term, 2005

graham

graham

To my parents

Unfortunately what is little recognized is that the most worthwhile scientific books are those in which the author clearly indicates what he does not know; for an author most hurts his readers by concealing difficulties.
~Evariste Galois (1811-1832)

Acknowledgements

Contents

Chapter 1

Clouds

1.1 Introduction

Clouds are complicated. Throughout history, clouds have captured the imagination of the great philosophers, from the ancient Greeks[1] to Descartes[2] they have proved a source of intrigue because of the optical phenomena they produce, and a source of concern due to their association with weather systems, rain and snow. Inherent complexities can be found in the current understanding of cloud formation, dissipation and composition, as well as their role in meteorological processes. Since the 1950s, the overall goal of cloud science has been focused on a quintessentially modern concern: the complex role of clouds within global climate change.

Clouds can significantly moderate the Earth's radiation budget as they can have a warming or cooling effect on the planet, depending on both the physical location and the microphysical properties of the cloud. Due to the variability of cloud, the overall effect of clouds on the radiation budget is not well understood. In addition, it has been recognised for decades that one of the largest uncertainties surrounding predictions of future climate is due to the representation of cloud processes in global atmospheric models (Randall et al., 2003).

To improve our current knowledge of cloud radiative properties and the impact on global climate it is imperative to obtain a global cloud climatology inclusive of the geophysical location, microphysical composition and optical properties. Considering the large spatial and temporal scales, the most feasible means of completing such a task is to derive the information from satellite observations.

The process of inferring the state of the atmosphere from a set of remote measurements is known as "retrieval", and it is the primary objective of this thesis to develop a *retrieval* method to determine cloud properties from measurements made by infrared limb sounding instruments. Limb sounders are extremely useful because of their high sensitivity to cloud

---

[1]See, for example,*Meteorologica*, Aristotle, 340 BC.
[2]See *Discours de la Methode*, René Descartes, 1637.

1

in the instrument line-of-sight, particularly to optically thin cloud, i.e. cirrus. The infrared region is of interest because of the importance of cloud radiative properties in this region. The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) is chosen because of its high spectral and vertical resolution. Furthermore, the method focuses on determining cloud properties for cirrus. The reasons for these choices will become clearer as this chapter and the next progress.

The purpose of this chapter is to introduce the background knowledge required for studying cirrus and to contextualise the importance of this research.

1.2 Overview: Formation and Classification

A cloud is a suspension (in the air) of minute water particles, either in the form of droplets or of ice crystals. Only about 0.0012 % of Earth's water is in the atmosphere, predominantly in the form of water vapour, which is a trace gas and over 90 % of the water vapour in the atmosphere is contained within the lowest 5 km of the troposphere. Clouds form when gaseous water in the atmosphere cools to its condensation or deposition temperature (called the *dew point* or *frost point*), or when enough water vapour is added to the air to cause condensation or deposition.

Generally, air which is condensed into clouds is cooled by mechanical uplift. This can happen by convection, at weather fronts, or by orographic (mountain) uplift. In each case, the rising air is forced to expand due to pressure dropping off with height in the atmosphere. Such expansion consumes heat energy and results in a fall in temperature.

In addition to cooling and sufficient moisture, water vapour needs *cloud condensation nuclei* (CCN) to form clouds. CCN are tiny particles, usually sub-micron size, about which cloud droplets coalesce. They are estimated to number on average in the air from 100 to 1000 per cubic centimeter. CCNs are usually considered a class of atmospheric aerosol and can be naturally or anthropogenically produced. Examples include: dust or clay particles, soot or black carbon particles from grassland or forest fires, sea salt from ocean wave spray, various by-products of fossil fuel burning, and sulphate aerosols from volcanic activity. To act as CCN, particles must be *hygroscopic*, i.e. they absorb water readily from their surroundings and accelerate the condensation process.

At any given time, clouds can cover up to 75 % of our planet. They are highly variable, and as such there exists a myriad of variations over time, space and composition, thus rendering the objective classification of clouds a daunting task. The first classification system (upon which our modern system is still based) was introduced in 1802 by an English pharmacist named Luke Howard in a presentation to the Askesian Society[3]. Howard's system grouped

---

[3] A social club for scientific thinkers, established in 1796 in London and founded by William Allen.

Height Group	Form	Genus	Species
LOW–LEVEL (<2 km all regions)	Cumuliform	Cumulus (Cu)	humilis mediocris congestus fractus
		Cumulonimbus (Cb)	capillatus calvus
	Stratiform	Stratocumulus (Sc)	floccus castellanus stratiformis lenticularis
		Stratus (St)	nebulosus fractus
MID–LEVEL (2–4 km polar) (2–7 km temperate) (2–8 km tropical)		Nimbostratus (Ns)	
		Altostratus (As)	
	Cumuliform	Altocumulus (Ac)	castellanus floccus stratiformis lenticularis
HIGH–LEVEL (3–8 km polar) (5–13 km temperate) (5–18 km tropical)	Cirriform	Cirrus (Ci)	uncinus spissatus floccus castellanus
	Stratiform	Cirrostratus (Cs)	nebulosus fibratus
	Cumuliform	Cirrocumulus (Cc)	castellanus floccus stratiformis lenticularis

Table 1.1: Cloud classification by cloud-base height above ground. Derived from World Meteorological Organisation (WMO) Cloud Classification (pp. 13-15 WMO, 1975).

clouds based on two criteria: their morphology and the height of the base of the clouds above the ground. He introduced a binomial nomenclature designated by a pair of Latin names; one defines cloud genus, and the second indicates cloud species. Howard observed that there were three basic cloud *genera*: thin, wispy clouds which he called *cirrus* (from the Latin for fibre or hair), thicker, puffy clouds he named *cumulus* (from the Latin for heap or pile) and sheet-like clouds which he named *stratus* (adapted from the Latin for layer). Howard also introduced the term *nimbus* (Latin for rain cloud) for precipitating clouds. The modern form of classification, with ten cloud types, was reached around 1880. The ten cloud types are: *cirrus, cirrostratus, cirrocumulus, altocumulus, altostratus, nimbostratus, stratocumulus, stratus, cumulus* and *cumulonimbus*.

By convention, the cloud types are classified into three cloud-base height regimes (with some overlap) in the troposphere. The cloud-base height classification is given in table 1.1 and figure 1.1 gives a pictorial representation of cloud type heights and morphology.

1.3 Introduction to Cirrus

It is now well known that cirrus permanently covers about 30 % of the Earth's surface (Wylie et al., 1994). Depending on geographic location, cirrus is present between 5 % and 65 % of the time. The impact of nonblack, high thin cirrus is recognised as one of the most important factors influencing Earth's Radiation Budget (Liou, 1986) (see Section 1.6.3 for
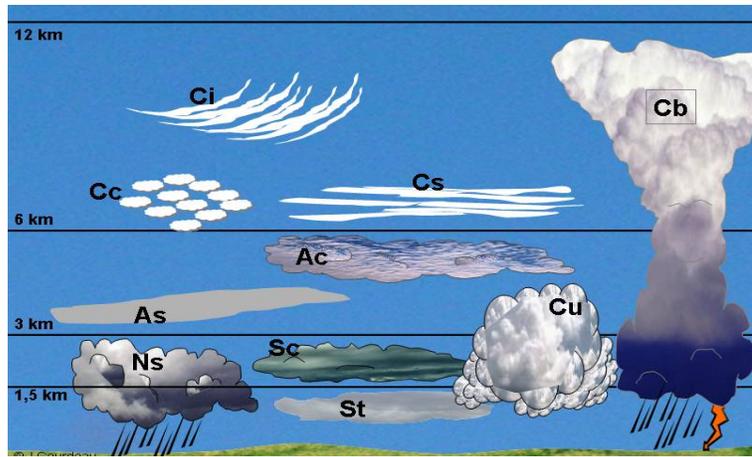
Figure 1.1: Pictorial representation of cloud genera by typical shape and average approximate base altitude (Gourdeau, 2004).

further discussion).

The morphological definition of cirrus is (WMO, 1975):

> *Detached clouds in the form of white, delicate filaments or white or mostly white patches or narrow bands. These clouds have a fibrous (hair-like) appearance, or a silky-sheen, or both.*

However, a more appropriate definition of cirrus for modern atmospheric research would be: *high, cold clouds composed entirely of ice crystals.* Defined as such, cirrus includes featureless ice clouds that lack visual structure and thus do not fit the conventional classification. Further, the definition encompasses a wide array of types and a vast range of macrophysical and microphysical properties. A thorough attempt to summarise the physical properties of cirrus clouds can be found in Dowling and Radke (1990).

Cirrus is unlike water cloud in two important ways:

- Cirrus is normally optically thin ($\tau \lesssim 1$) and nonblack, whereas water clouds are optically thick ($\tau \gg 1$) and can be considered black-bodies.

- Cirrus is composed of nonspherical ice crystals with a large range of possible sizes and shapes, as opposed to water clouds which are always composed of round drops.

The average properties of cirrus are summarised in table 1.2.

The existence of thin, tenuous cirrus lacking visual structure has been established for some time (e.g. Barnes, 1980; Heymsfield, 1986; Sassen and Cho, 1992). This cirrus is commonly referred to as "subvisual cirrus" and currently has no Latin designation nor is it included in conventional cloud classification. Subvisual cirrus is defined as cirrus with vertical optical depths of less than 0.03 at (a visible wavelength of) 0.694 μm. Such clouds will appear invisible against the blue sky to the naked eye. Subvisual cirrus are widespread (though mainly found at the tropics), forming at (and sometimes above) the tropopause (see Winker

Property	Mean	Observed range	Range ratio
Thickness	1.5 km	0.1–8 km	80:1
Altitude	9 km	4–20 km	5:1
Concentration	30 l^{-1}	10^{-4}–10^4 l^{-1}	10^8:1
Ice content	0.025 gm^{-3}	10^{-4}–1.2 gm^{-3}	10^4:1
Size (length)	250 μm	1–8000 μm	8000:1
Shape	Variable	Highly variable	Large

Table 1.2: Cirrus cloud summary (as given in Lynch, 1996).

and Trepte, 1998) and they are virtually undetectable with current passive nadir satellite sensors.

There is an additional cirrus type not recognised by the WMO cloud classification, known as "contrail cirrus". A contrail is a visible cloud that forms behind an aircraft, mainly due to water vapour emissions in the exhaust of a jet engine in cold ambient air (Schumann, 1996). Persistent contrails may last for hours and grow into spreading cirrus cloud (i.e. cirrostratus). These contrail cirrus can form where natural cirrus would not normally occur, and although the global coverage is small, there is potentially a high climatic importance.

1.4   Cirrus Formation and Types

It is possible for ice to form in most cloud types; however, only cirrus and cirrostratus are always ice crystal clouds. Ice particles form in the atmosphere when water vapour sublimes or when water drops freeze, both processes being called *glaciation*. There is still significant uncertainty in the understanding of ice-formation processes in cirrus.

According to DeMott (2002), at cirrus temperatures (−30 to −80 °C) nucleation of ice particles can occur by homogeneous nucleation and by the heterogeneous catalytic action of *ice nuclei* (IN). Homogeneous freezing is spontaneous freezing of pure water or spontaneous ice-embryo formation within soluble aerosol particles in various states of dilution. Aerosol solutions are believed to consist predominantly of mixtures of H_2SO_4, HNO_3, NH_3, and H_2O. Generally, condensation occurs first and then the cloud droplets supercool to subzero temperatures before freezing to ice. Homogeneous freezing of a pure liquid drop occurs when statistical fluctuations of the molecular arrangement of the water produce a stable, icelike structure that can serve as an ice nucleus (Rogers and Yau, 1989). Theoretically, this is expected to occur in pure liquid water at temperatures below about −36 °C (depending on particle size). The presence of a solute has the effect of lowering the freezing rate compared to an equal-sized pure droplet, but freezing rate increases rapidly with decreasing temperature. Homogeneous nucleation depends only on the water activity of the solution (Koop et al., 2000) which, under atmospheric conditions, can be expressed

as the relative humidity[4] with respect to water. An increase in relative humidity of the solution is expected to cause it to freeze at temperatures below about −40 °C. Only highly concentrated, small droplets will remain in liquid form high in the troposphere (associated with low air temperatures and low humidities). When the formation of cirrus is driven by convection, cloud droplets can remain in liquid form until they homogeneously freeze. Homogeneous nucleation of deposition (i.e. when vapour molecules form a stable ice embryo by chance collisions) is unlikely to occur in the atmosphere due to this requiring extreme conditions of supersaturation[5].

Through heterogeneous nucleation, there is potential for ice nucleation in some cirrus to be initiated at warmer initial temperatures and lower relative humidities than is possible by homogeneous nucleation. Heterogeneous nucleation takes place when supercooled liquid freezes around an ice nucleus, or vapour is deposited on a deposition nucleus. More specifically there are four hypothesised mechanisms (DeMott, 2002) of heterogeneous ice-nucleation:

- Deposition nucleation: ice embryos form from the vapour on an insoluble surface or a soluble but anhydrous particle.

- Condensation freezing: the soluble component of a mixed particle causes liquid condensation while the insoluble component causes freezing instantaneously.

- Immersion freezing: freezing by an insoluble particle previously immersed in a liquid drop.

- Contact freezing: freezing initiated by a solid aerosol particle collision with a liquid drop.

IN differ greatly from CCN and they are far less abundant, with concentrations of 10^{-3} IN cm^{-3} compared to 10^3 CCN cm^{-3}. Clay particles are one of the most important groups of IN, in particular, kaolinite, which has the capability of initiating ice at −9 °C. The ever-present nature of aerosol bodies in the upper troposphere is well established (Murphy et al., 1998) and thus heterogeneous nucleation is an important process in cirrus formation. However, *in situ* observations (e.g. DeMott et al., 1998) support the long-held belief that due to the very low concentrations of IN in the upper troposphere, homogeneous nucleation by the freezing of small liquid-solution droplets is the dominant mechanism of ice formation (Heymsfield and Miloshevich, 1993). Interestingly, more recent observations have shown that homogeneous nucleation dominates cirrus formation in southern latitudes

---

[4]Relative humidity is the ratio of the actual water vapour pressure to the saturation water vapour pressure at the prevailing temperature (usually expressed as a percentage). It does not define the water content of the air unless the temperature is given.

[5]Supersaturation is represented by values of relative humidity greater than 100 percent.

but heterogeneous nucleation becomes more important in the polluted northern latitudes (Haag et al., 2003), thus emphasising the association of cirrus formation with the presence of man-made aerosol in the upper troposphere, and therefore having strong implications for human influence on climate. Modelling studies (Gierens, 2003) have implied that there may be a critical concentration of ice nuclei for heterogeneous freezing to dominate cirrus formation, which also requires increased updrafts. The nucleation process can have important consequences for the macrophysical properties (e.g. optical thickness) of the cirrus cloud.

In general, cirrus ice crystals form as liquid solution (haze) droplets rise in the atmosphere and homogeneously freeze at temperatures below $-40\,°C$. The liquid droplets enter into an environment where the actual water vapour pressure is equal to or slightly greater than the equilibrium water vapour pressure over liquid water (usually relative humidities are between 100–105 %). Therefore, as the equilibrium water vapour pressure over ice is much lower than over water, the environment is supersaturated with respect to ice. Environments with high supersaturation with respect to ice are favourable for the rapid growth of ice crystals by diffusion and deposition. A vapour pressure gradient develops between the droplets and newly nucleated crystals. Due to this gradient, water vapour moves from the higher pressure surrounding the droplets to the lower pressure surrounding the crystals. Thus, the ice crystals grow at the droplets' expense. This process creates subsaturation with respect to water, and the droplets evaporate to maintain water saturation, making additional water vapour available for ice crystal growth. The ice particles grow rapidly and can reach dimensions of hundreds or thousands of micrometres (compared to $\leq 15\ \mu m$ for their parent drops or droplets typical to water clouds). If the liquid droplets eventually disappear, by evaporation or freezing, the water vapour pressure decreases to equilibrium relative to ice, and crystal growth by diffusion and deposition halts. However, it can continue by aggregation and collision. Crystal growth is an important factor in determining the microphysics of the cloud and is discussed further in Section 1.5.

Evidently, ice cloud formation requires conditions of supercooling and supersaturation. In the upper troposphere, the amount of water vapour is too low to produce much ice, and the vertical motions in most clouds are gentle. Thus, the properties of the ice crystals strongly depend on the cloud-generating mechanism. Cirrus cloud-generating mechanisms can be classed into four basic natural categories and one anthropogenic (Sassen, 2002). The breakdown is given in table 1.4.

Contrail-cirrus

The rapid formation of a condensation trail in the wake of a jet engine is due to the moisture and nuclei supplied by the exhaust, usually sulphur-based haze particles which

Category	Mechanism
Synoptic (jet stream, frontal, etc.)	Top-down generation
Injection cirrus	Thunderstorm anvil
Mountain-wave updraft	Orographic, terrain-induced
Cold trap	Tropopause-topped thin layer
Contrail-cirrus	Rapid cooling of aircraft exhausts

Table 1.3: Breakdown of cirrus clouds by generating mechanism. As in Sassen (2002).

homogeneously freeze during the mixing process. Under ambient conditions, the contrail ice particles can spread, acting as ice-embryos, and cause the subsequent development of cirrostratus.

Synoptic

The synoptic cirrus category covers a wide range of the common varieties of cirrus clouds that form in situ in the upper troposphere. Normally synoptic cirrus originates as a stable layer of liquid droplets, which then rises and subsequently cools in response to weather disturbances (e.g. frontal overriding). Updrafts range from the centimetre per second scale typical of gradual (i.e. frontal) uplift to the metre per second scale of the single convective uncinus cell. The scenario of cirrus uncinus formation by a weakly convecting generating cell is described by Heymsfield (1975). As the cloud droplets rise, they will homogeneously freeze, usually from the top of the cloud down. Large ice particles fall rapidly away from their generating cells within the water clouds, leading to the familiar fibrous patterns. In addition, the particle precipitation and evaporation processes can cause additional cirrus to develop below the original generating cell (Sassen, 2002).

An example of jet stream cirrus is shown in figure 1.2. The jet stream is a narrow zone of high-speed winds (reaching up to 100–250 km h^{-1}) typically found at altitudes of 8 to 12 km at mid-latitudes between 20° to 55° north and south. They result from temperature contrasts between polar and tropical regions. The high-speed winds can cause long-lived cirrus to form which can extend over thousands of kilometres with a typical thickness of only 1 or 2 kilometres.

Mountain-wave updraft

Orographic cirrus involves ice nucleation in very strong updrafts. Orographic cirrus is predominantly generated by very large amplitude vertically propagating gravity waves, produced by strong winds flowing over high mountains (Queney, 1948). The large vertical displacement of the air parcels causes the rapid updraft. The ice crystals which form can be advected by the wind for considerable distances, and thus lead to orographic cirrus with

Figure 1.2: Jet stream cirrus crossing Cape Breton Island in the Maritime Provinces of Eastern Canada. As seen from NASA Space Shuttle 29th April 1991 (Hamilton, 1997).

horizontal coverage on the order of a thousand kilometres.

Injection cirrus

Like the orographic case, ice formation within thunderstorm anvils occurs in strong updrafts of a few metres or more per second. Anvil cirrus is formed by deep convective storms that create vertically propogating buoyancy waves that extend above the convection to the tropopause and into the stratosphere (Potter and Holton, 1995). Convectively-raised aerosols from the boundary layer and mid-troposphere will cause ice nucleation in the maintenance stage, with ice crystals forming from the homogeneous freezing of haze (i.e solution) particles entrained into convective updraft cores (Fridland et al., 2004). As the storm develops, ice crystals will be ejected upward from the convective core and lofted by the buoyancy waves. Larger ice particles ($> 1000 \ \mu$m) in the outflow will precipitate out, leaving the smaller crystals to spread and develop into the anvil cirrus, which then thins and can eventually detach from the convective core. The main convective updrafts are deflected horizontally at the tropopause, causing the ice crystals at that level to spread out. Also, it is argued that concentrated interaction with thermal radiation at the cloud vertical boundaries creates thin frontal tongues of well-mixed cold and warm air that spread the anvil cirrus laterally, leaving the interior comparatively unperturbed (Scorer, 1978; Garrett et al., 2005). Figure 1.3 shows an image of a typical cirrus anvil top, still attached to the convective core beneath.

Figure 1.3: Cirrus anvils as seen from the International Space Station 17th June 2002 (Klinedinst, 2003).

Cold trap

Recent observations show that anvil cirrus in the tropics often forms in conjunction with a distinct upper layer of much more tenuous cirrus at the tropopause above the thunderstorm anvil (Garrett et al., 2004; Garrett et al., 2005). Garrett et al. (2004) hypothesise that such long-lived thin tropopause cirrus does not originate from convective detrainment, but instead originates as stratiform pileus clouds that form near the tropopause ahead of vigorous convective uplift. The pileus are then predicted to be penetrated by the storm convection and moistened through turbulent mixing. Then, once the storm subsides, they are sustained by radiative cooling due to the presence of the anvil layer beneath.

These tenuous tropical cirrus layers fall into the class of "cold trap" cirrus. Within a "tropical tropopause layer" (TTL) located near or within the tropopause, air is decoupled from the convection-dominated lower troposphere, somewhat acting as a "trap". It is in the TTL that the cold trap layers of cirrus are often found. They are classed as such due to their common characteristics: they are uniquely cold (-70 °C to -90 °C), and high (15–20 km), conditions which usually only occur in the tropics. Regardless of specific origin, cold trap layers require a high-altitude moisture flux and turbulent conditions or wave action to initiate crystal growth. They are often observed as "remnants" of other cloud formation mechanisms such as orographic, cirrus outflow and diffuse contrails. The most likely cause of cold trap cirrus is due to the interaction of gravity waves with moisture fields (of supercooled haze droplets) near the tropopause that are saturated with respect to ice. Cold trap cirrus layers have only been observed recently (Heymsfield, 1986; Winker and Trepte, 1998; Santacesaria et al., 2003; Sunil Kumar et al., 2003), they are narrow,

Category	Abbreviation	τ range	Description
Ultrathin	UTTC	$\sim 10^{-4}$	Invisible (a factor of 300 below visibility)
Subvisual/subvisible	SVC	< 0.03	Invisible against the blue sky
Thin	TTC	0.03–0.3	Translucent, retains a blueish colour
Opaque	-	0.3–3.0	Usually appears white

Table 1.4: Subclassification of cirrus by visible optical depth and transparency (c.f. table 2.1 Sassen, 2002).

layered structures ranging in thickness between a few hundred metres to 1 kilometre thick, and with large horizontal extents up to several thousand kilometres.

The identification of cold trap cirrus has led to the subclassification of cirrus by their transparency or optical depth (at 0.694 μm as in the definition of subvisual cirrus) which is given in table 1.4. The cold trap cirrus encompasses "thin tropopause cirrus" (TTC), with optical depths $\tau < 0.3$, subvisual (or subvisible) cirrus (SVC), with $\tau < 0.03$ and also "ultrathin tropopause cirrus" (UTTC), with $\tau \sim 10^{-4}$, observed over the Indian Ocean by Luo et al. (2003). The final subclass of cirrus which the other cloud forming mechanisms contribute to is "opaque cirrus", with $0.3 \leq \tau \leq 3.0$. Due to the ubiquitous nature of subvisible cirrus in the tropics and its extremely cold temperatures, it is widely believed to be important in the greenhouse effect and to possibly play a role in the dehydration of the lower stratosphere (to be discussed further in section 1.6.3).

1.5 Cirrus Microphysics

Cirrus clouds are composed of ice crystals with various non-spherical shapes. The sizes of the ice crystals can range from less than 10 μm to about 4000 μm (Dowling and Radke, 1990; Heymsfield and McFarquhar, 2002), and the typical geometry of the crystals can range from single pristine shapes such as hexagonal ice columns or plates, to complex aggregates of columns or bullets (Baran, 2004). The current understanding of sizes and shapes of cirrus cloud ice crystals is due mainly to in situ measurements. There has been a wealth of in situ measurement studies, predominantly by aircraft mounted instruments, and are too numerous to mention in detail here. Campaigns of note include:

- The early studies of Weickmann (1947) to characterise cirrus crystals over Germany.

- The First International Satellite Cloud and Climatology (ISCCP) Regional Experiments (FIRE I and II) over Wisconsin and Kansas (e.g. Kinne et al., 1992; Arnott et al., 1994; Heymsfield and Miloshevich, 1995).

- The SUbsonic aircraft: Contrail and Cloud Effects Special Study (SUCCESS), which

was a NASA field campaign over Oklahoma and Kansas from April 8th to May 15th, 1996 (e.g. Goodman et al., 1998; Heymsfield et al., 1998; Lawson et al., 1998).

- The Central Equatorial Pacific Experiment (CEPEX) which attempted to characterise anvil outflow cirrus (e.g. Chen et al., 1997; McFarquhar and Heymsfield, 1997).

- The European Cloud and Radiative Experiment (EUCREX) over the Atlantic ocean and mid-latitude Europe (e.g. Raschke et al., 1998; Buschmann and Albers, 2000).

- The Cirrus Regional Study of Tropical Anvils and Cirrus Layers - Florida Area Cirrus Experiment CRYSTAL-FACE, which was a six aircraft, intensive campaign during July 2002 (e.g. Garrett et al., 2003; Heymsfield et al., 2004)

The reader is directed to Heymsfield and McFarquhar (2002) for a review of in situ microphysical data.

Microphysical properties are taken here to mean the size and shape of the ice particles and their volume concentration. The volume concentration can be expressed as an ice water content (IWC), i.e. mass of ice per unit volume, or number density, i.e. number of particles per unit volume. The radiative properties of cirrus clouds (i.e. optical depth, albedo, emissivity and transmissivity) are dependent on the microphysical properties (Stephens et al., 1990; Mitchell and Arnott, 1994), and thus the characterisation of cirrus microphysics is of critical importance to a study of the role of cirrus in the climate system (see Section 1.6.3).

1.5.1 Ice Crystal Growth

Obviously, the microphysics of the cloud will be influenced by the nucleation, growth and evaporation of its constituent ice crystals. Once ice is initiated a crystal will begin to grow. Crystal growth depends on the location of a crystal with respect to the edge of the cloud and the intervening optical thickness; evaporation depends on larger scale (dynamic) processes (Hallett et al., 2002). Crystals grow by diffusion and deposition. The growth of a crystal by diffusion is analogous to the growth of a water droplet by condensation with added complication due to the nonspherical shape. A crystal growing by diffusion requires a supersaturated and supercooled environment and can be considered as depositional growth from either vapour or liquid. Diffusional growth of ice crystals has been described by Rogers and Yau (1989). Crystal growth can be approximately described by the generalised growth equation

$$\frac{dm}{dt} = 4\pi C D(\rho_s - \rho_e) \tag{1.1}$$

where m is the particle mass, D is the environment diffusivity, ρ_s is the vapour or solute concentration at equilibrium with the crystal surface and ρ_e is the vapour or solute con-

centration at equilibrium with the environment. C represents the size and shape of the particle[6]. For a sphere, $C = r$ and equation 1.1 reduces to the growth equation for a water droplet. Equation 1.1 assumes surface conditions are uniform and ignores radiative effects. Ambient conditions determine not only the growth rate, but also the habit that a growing crystal takes. The treatment of crystal growth as a process controlled by diffusion of water vapour cannot explain the different growth habits of ice. A molecular-kinetic approach is required to explain these effects (see Furakawa and Kobayashi, 1978). In its solid form water forms a hexagonal structure and it is this underlying structure that gives many forms of ice crystals their hexagonal symmetry. The surface of a growing crystal is made up of flat terraces of different heights terminating at ledges and separated by steps. This is shown schematically in figure 1.4. Kinetic growth is controlled by the rate at which molecules can attach to the lattice. When a water molecule impinges on the surface, energetically it is easier to add a molecule to a ledge (i.e. the basal plane) than to begin a whole new terrace (i.e. growth along the c-axis), thus causing crystal growth by the lateral motion of surface steps (along the basal plane), which in turn is due to the diffusional movement of molecules along terrace sites and into ledge sites. The molecularly smooth layer will also grow but at a slower rate, and particularly if there are surface defects. These kinetic growth effects are controlled by temperature and supersaturation. The occurence of dendritic habits and symmetries other than hexagonal is due to imperfections (i.e. basal plane stacking faults) in the hexagonal lattice structure. These imperfections introduce regions of cubic symmetry and if such a layer is thick enough, nucleation can occur on the surface, complicating the shape of the crystal growth. However, it is still difficult to explain the variety of shapes and their exact dependence on the ambient conditions (see Hallett et al., 2002). The dependence of crystal habit on both temperature and supersaturation over ice is shown in figure 1.5. In any case, the classical theory of crystal growth (equation 1.1) can be used for approximate growth calculations if the crystal habit is specified.

Further crystal growth can occur by *accretion* or *aggregation*. The term accretion is often reserved to describe the capture of supercooled droplets by an ice-phase precipitation particle. The supercooled droplet freezes on contact and sticks to the original ice crystal forming a coating of rime, leading to rimed crystals or graupels. As ice crystals grow by accretion, rapid freezing of droplets can result in shattering to form multiple tiny ice particles. This process is called splintering and is a secondary source of IN. This can then speed up the accretion process and glaciation of the cloud.

Aggregation is the process by which ice crystals stick together as they grow and collide. Liquid molecules on the outer surface of the crystals serve to increase bonding between two

---

[6]C is actually the electrical capacitance in *cgs* units, i.e. centimetres. For a sphere, this value differs from the electrical capacitance measured in Farads by a factor of $1/4\pi\epsilon_0$ (where ϵ_0 is the relative permittivity of free space).
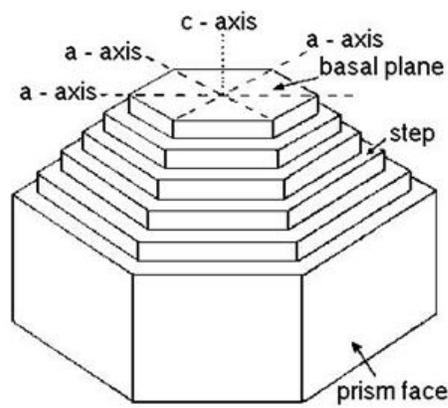
Figure 1.4: Ice crystal surface.
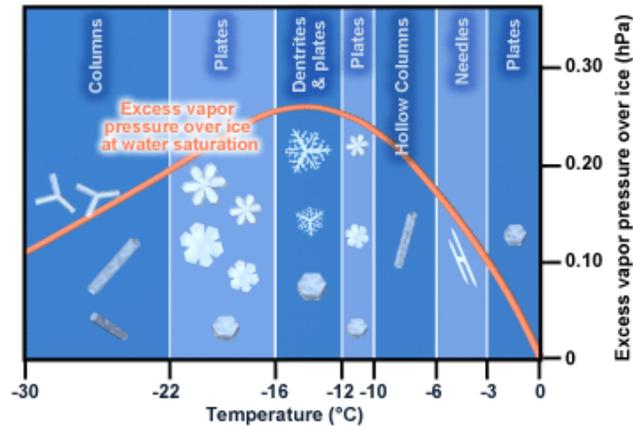
Figure 1.5: Common crystal habit and formation conditions (COMET Program, 2005).

colliding crystals. Crystals with dendritic characteristics can also mechanically interlock to form larger aggregates (Pruppacher and Klett, 1978), and are thus more likely to form aggregates than pristine crystals. The growth of crystals by accretion and aggregation depends on the size of the particles, the fall speed and their collision/collection efficiency (see Rogers and Yau, 1989).

1.5.2 Ice Crystal Shapes and Sizes

Typically, ice cloud can be composed of several different particle habits, and in some cases the particle shape is related to the relative vertical height within the cloud. Cirrus particle habit is known to depend on the type of formation mechanism and temperature and thus in effect geographic location (Heymsfield and McFarquhar, 2002).

Mid-latitude cirrus

In mid-latitude cirrus and cirrostratus, crystal habit has been shown to depend on whether the cirrus was formed from a convective cell or from slow updraft, and on the atmospheric

temperature and thus height. Heymsfield and Platt (1984) described crystal shapes in mid-latitude cirrus to be dominated by spatial crystals (e.g. bullet rosettes) above $-40\,°C$ and by hollow or solid columns below $-50\,°C$, while hollow columns and hexagonal plates predominated near the cloud top throughout. In the range $-40\,°C$ to $-50\,°C$ convective cirrus contained predominantly spatial crystal forms and stable cirrus contained predominantly hollow columns. Heymsfield et al. (1990) made further studies of mid-latitude cirrus during the FIRE campaign. The study found similar results to Heymsfield and Platt (1984), but found a larger proportion of the cloud particles were spatial aggregate crystals. In fact, aggregates were found at temperatures as low as $-56\,°C$; previously aggregation was not believed to occur below $-20\,°C$. However, the resolution of the PMS 2D probe used was only $25\,\mu m$, and thus was not able to resolve particle shapes below $\sim 125\,\mu m$. Particles $> 1000\,\mu m$ were frequently sampled and it was estimated that particles in the range 500–$800\,\mu m$ made up between $10\,\%$ and $30\,\%$ of the ice mass. However, the small shapes were shown to dominate the total ice particle concentrations and were classed by default as columns. There was limited confidence in this classification, but the evidence of aggregation at cold temperatures would imply that the constituent crystals of the aggregates were small columns, and therefore the crystals with sizes below the resolving limit are likely to be columns. Arnott et al. (1994) and Korolev et al. (2000) also reported on crystal shapes in mid-lattitude cirrus using the PMS 2D probe. Korolev et al. (2000) suggested that aggregates dominated in most cases, however this study neglected particles $< 125\,\mu m$ (due to the imager resolution) and didn't sample at temperatures less than $-45\,°C$. The recent study of mid-latitude frontal cirrus by Whiteway et al. (2004) is not as limited by imager resolution (the Cloud Particle Imager (CPI) instrument used has an improved resolution of $10\,\mu m$) and the results are more in agreement with Heymsfield and Platt (1984). Columnar crystals were found throughout the cloud while bullet rosettes were found mainly at the top. This suggests that the arms of the rosettes may break off as the crystals fall and thus being detected as columns. Small irregular crystals were found at the cloud base where sublimation will cause the crystal edges to round. Also, due to the size of the irregular crystals being close to the order of the CPI resolution, diffraction effects make the shapes difficult to distinguish. Figure 1.6 shows one of the figures from Whiteway et al. (2004) in which the crystal shape detected during an aircraft flight through the frontal cirrus is shown by the actual CPI images.

Tropical anvil cirrus

The shape of ice cloud crystals in the tropics is not as well known. There is a lack of in situ measurements of low-latitude anvil cirrus due to the logistic difficulties associated with their height and remoteness. In tropical anvils, particles can be aggregated and rimed
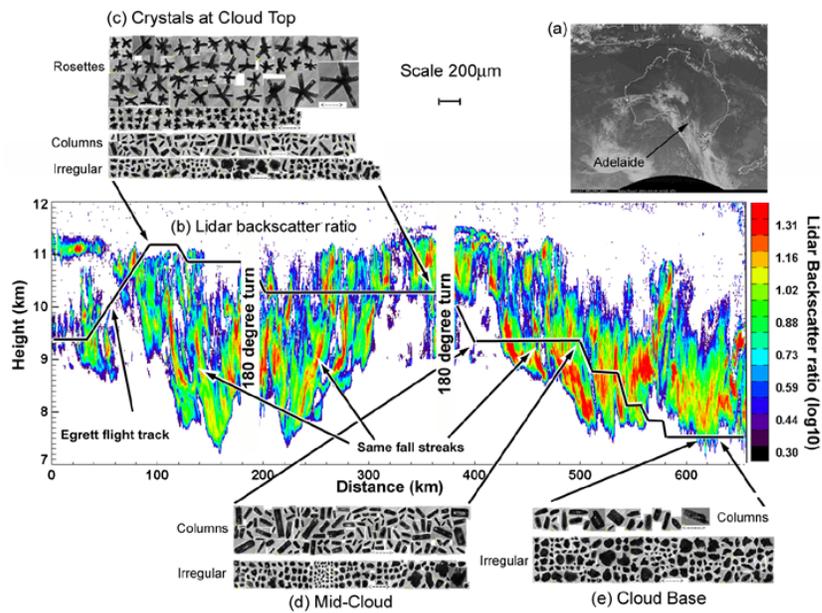
Figure 1.6: (a) Satellite IR image showing the cirrus cloud system above Southern Australia. (b) Lidar measurements on-board the King Air aircraft while flying directly below the Egrett aircraft. Ratio of total backscatter to that expected for molecular backscatter only. (c, d, e) Images of crystals recorded by the CPI on-board the Egrett. As figure 1 in Whiteway et al. (2004).

and generally can have a more complex structure than in mid-latitude cirrus. The high amounts of moisture and the turbulence associated with vertical motions within the anvil gives the greater variety and complexity of crystal shapes. During CEPEX, McFarquhar and Heymsfield (1996)[7] measured cirrus associated with blow-off anvils and found that aggregates and highly irregular shapes were frequently found, but that the size of the crystals decreased with height and were more likely to be single crystals (such as plates and columns) for sizes less than 100 μm. These small crystals tend to dominate the ice water content in the upper, colder parts of the cirrus, making up more than half the mass and can make up more than half the extinction. Similar results were found by Whiteway et al. (2004), measuring anvil outflow cirrus off the north of Australia. The results can be seen in figure 1.7. The anvil outflow measured was mainly composed of hexagonal plates and columns and large crystal aggregates that originated from the convective core. It was found that the large aggregates fell to the lower regions of the outflow cirrus cloud while the single crystals and small aggregates remained at the top. Both the studies of McFarquhar and Heymsfield (1996) and Whiteway et al. (2004) show the importance of sedimentation in determining the distribution of crystal size within anvils.

Evidence of sedimentation was also found during CRYSTAL-FACE. Garrett et al. (2005) found that tropical anvils over Florida were composed mainly of small crystals < 50 μm in the first few kilometres at the anvil top, and then increasingly large ice crystals > 1000 μm

---

[7]Also see Heymsfield and McFarquhar (1996) and McFarquhar and Heymsfield (1997).
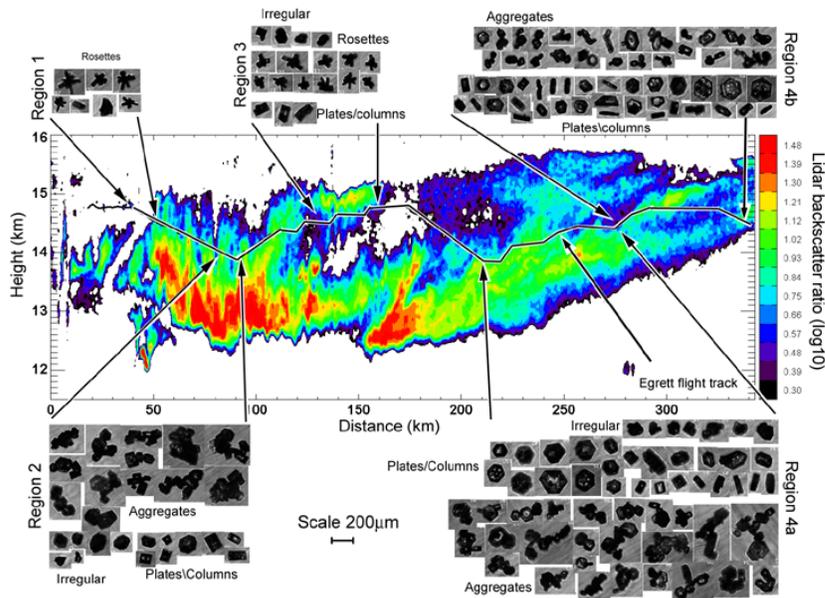
Figure 1.7: Lidar backscatter (ratio of total to molecular) measurements of the outflow cirrus cloud. Crystal images are shown within various sections of the cloud along the Egrett flight path (solid black line). As figure 3 in Whiteway et al. (2004).

at greater depths of the anvil. The small crystals were found to be quasi-spherical in shape, with a few pristine hexagonal plates. The larger ice crystals were primarily aggregates of plates that had undergone varying degrees of evaporation. The initial evolution of the compact, irregular or quasi-spherical crystals in anvils is difficult to speculate on. It may be the case that these particles are in the early stage of growth, being liquid drops that have just homogeneously frozen and have yet to develop into a defined habit. However, the Garrett et al. (2005) study also used the CPI and the problem of imaging resolution for the small particles somewhat prevents a confident definition of small particle shapes. Baumgardner et al. (2005) and Chepfer et al. (2005) also made in situ measurements during CRYSTAL-FACE, but using an optical particle spectrometer sensitive to crystals in the range 5–45 μm. The optical particle spectrometer measures scattered radiation from individual particles, and shape is deduced by the comparison of measured phase functions to calculations for idealised shapes. It was found in both studies that the majority (90 % in the case of Baumgardner et al. (2005)) of small crystals were in agreement with the theoretical model for a mixture of bullet rosettes, plates and hollow columns. The spherical models could not explain the measurements in Chepfer et al. (2005) and only approximately 10 % of the crystals measured by Baumgardner et al. (2005) were characterised as ice spheres. Garrett et al. (2003) were able to derive an effective radius r_e for anvil cirrus layers, while bypassing the issue of poor accuracy of in situ probe measurements for ice crystals < 50 μm, by making direct bulk measurements of the mass concentration of condensed water W and

the extinction coefficient β. The r_e was derived, using the expression

$$\beta = \frac{3W}{2\rho r_e} \tag{1.2}$$

where ρ is the bulk ice density. The effective radius was shown to clearly decrease with temperature, reaching approximately 5 μm at -75 °C and going as high as 30 μm at freezing. The asymmetry parameter g was also measured and shown to be weakly dependent on temperature. The values of g measured were between 0.70 and 0.78 for temperatures between -55 °C and -20 °C (no measurements were made below -55 °C). Such low values of g are expected for small pristine crystals (e.g. hexagonal columns or prisms) but g had no clear dependence on r_e (although it roughly seemed to decrease with r_e), so the small value could also be explained by surface roughness or crystal aggregation.

The best present description of ice crystal shapes in tropical anvils may be that there are typically two modes of size distribution, with small crystals < 100 μm (and with an effective radius ~ 5–30 μm) located near the anvil top and large crystals (with an effective radius ~ 100–200 μm) found below that level with size increasing with temperature. The small crystals are likely to be a mixture of bullet rosettes, columns, plates and small aggregates and the large crystals are likely to be aggregates (and possibly aggregate chains) of plates and/or columns.

Thin tropical cirrus

The thunderstorm anvils studied in Garrett et al. (2005) were often found to coexist with an upper (detached) layer of long-lived thin tropopause cirrus (TTC). The TTC was found to have a near-identical horizontal extent to the anvils. Within the TTC the size distribution was mono-modal, and the particles measured usually did not exceed 100 μm. CPI images showed columns, spheres and irregularly shaped ice crystals, with most columns and spheres found at the bottom of the TTC. The optical depth of this TTC was approximately 0.2. This was a particularly interesting study as previously all that was really known of thin and subvisible cirrus microphysics at the tropopause was from the in situ study of Heymsfield (1986). Heymsfield (1986) measured subvisible tropopause cirrus layers (approximately 0.5 km thick, located within a double tropopause at approximately -80°C) near Kwajalein, Marshall islands using a high-altitude WB-57F aircraft equipped with a particle measuring system axially scattering spectrometer probe (ASSP), which sized in the range 2–30 μm, and a one-dimensional cloud particle probe sizing in the range 20–300 μm. In addition, excellent information about the crystal shape was obtained with a formvar replicator since the small crystals did not break up on impact. The maximum dimensions of the particles measured were in the range 1–50 μm, much smaller than frontal cirrus or thicker, more

easily detectable tropopause cirrus (frequently $> 100~\mu$m). The crystals measured by the replicator were reported to be an approximate 50 % mixture of trigonal plates and hexagonal columns. Also, Sassen (1991) derived measurements of subvisible cirrus particle size as ~ 10–$30~\mu$m from lidar depolarization measurements. Particle sizes are expected to be even smaller for UTTC, and Luo et al. (2003) derived a mean radius of $\sim 5~\mu$m for aspherical ice crystals. It can be inferred that, for colder and optically thinner cirrus, the ice crystals are likely to be increasingly small pristine faceted crystals.

Contrail cirrus

The crystals measured in subvisible cirrus are very similar to those found in the core of contrail cirrus. Crystals in contrail cores were sampled during SUCCESS (Lawson et al., 1998) and were reported to be small particles in the range 1–20 μm, whereas the crystals in the contrail periphery were much larger, with most crystals in the range 20–350 μm and the number of crystals in the range 1–20 μm being 1 or 2 orders of magnitude lower than in the core (Heymsfield et al., 1998). Generally, crystals in the periphery were columns and bullet rosettes, and Goodman et al. (1998) used an ice particle replicator to sample contrail core particles, reporting mainly hexagonal columns and plates and a few triangles.

1.5.3 Ice Water Content

Accurate knowledge of the amount of water condensed in clouds, and its vertical distribution, is important for modelling the radiative properties of clouds. Within cirrus, most of the IR absorption and much of the reflectivity in the visible is due to small crystals (Baker, 1997). Increasing emphasis is being placed on the contribution of small crystals to the ice water content (IWC) of cirrus.

There are two main methods of measuring IWC in cirrus clouds: directly measuring the amount of water in solid form, or by integrating the measured size distribution (the third moment of the size distribution approximately represents the IWC). Usually a hygrometer type instrument is used to measure IWC directly, e.g. the counterflow virtual impactor (CVI) uses a Lyman-alpha hygrometer after cloud particles are separated from water vapour and heated within an inlet (Weinstock et al., 1994). The hygrometer-based instruments are usually sensitive to all particle sizes. However, an estimate of the contribution of small crystals to IWC can be obtained by comparison to an instrument which is sensitive only to small particles, such as the particle volume monitor (PVM) (Gerber et al., 1994). The PVM is sensitive to particles $< 50~\mu$m and measures IWC directly by using spatial filtering of forward-scattering light weighted by the second and third moment of the particle distribution.
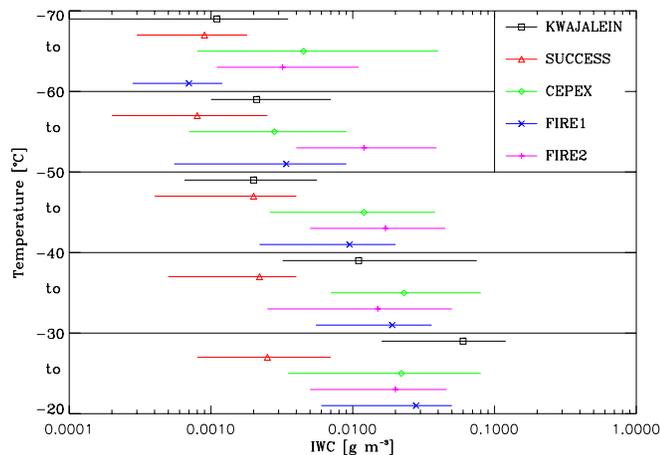
Figure 1.8: Ice water content as a function of temperature for different measurement campaigns. Symbols and horizontal bars represent median values and quartiles of distributions. Adapted from Heymsfield and McFarquhar (2002).

Measurements of IWC from size distributions should be viewed cautiously, particularly at temperatures below $-50\,°$C. This is mainly due to the lack of instrument sensitivity to small particles $\lesssim 75\ \mu$m. Most of the in situ projects mentioned at the beginning of section 1.5 made measurements of size distributions and have been discussed to some extent already in section 1.5.2. Results have been presented in summary by both Dowling and Radke (1990) and Heymsfield and McFarquhar (2002).

Generally, in mid-latitude cirrus, number concentrations decrease exponentially or in a power-law form with increasing size. Tropical cirrus exhibits similar behaviour but the decrease is not quite so rapid. In both mid-latitude and tropical cirrus, the largest particles in a size distribution increase with temperature. However, tropical cirrus at temperatures above $-40\,°$C have been observed to contain larger particles than mid-latitude cirrus at equivalent temperatures. In addition, for tropical cirrus at temperatures below $-60\,°$C (and particularly at about $-80\,°$C) small crystals typically dominate over the contributions of larger crystals (e.g. Heymsfield, 1986; Heymsfield et al., 1998). At such temperatures, the tops of tropical anvils rarely have particles greater than $100\ \mu$m, and thin and subvisible cirrus are composed mainly of particles less than $50\ \mu$m.

A summary of IWC data collected from various in situ campaigns, measuring different types of cirrus at a variety of locations and times of year, is given in figure 1.8. For reference, IWC of large cumulus clouds is in the range 0.3–$1.0\ \mathrm{gm}^{-3}$ and for stratus clouds IWC is in the range 0.1–$0.3\ \mathrm{gm}^{-3}$.

Measurements of mid-latitude synoptic cirrus (FIRE I and II) show a decrease in IWC with decreasing temperature, as expected. IWC ranges from $0.1\ \mathrm{gm}^{-3}$ above $-40\,°$C to 10^{-4} below $-60\,°$C. Heymsfield et al. (1990) showed that IWC derived from size distributions is dependent on particle habit and that IWC derived as such are accurate to within a factor of

two. In addition, for the mid-latitude cirrus sampled at lower altitudes, particles < 125 μm made up 10–12 % of the estimated mass, at the higher altitudes 20–35 % and at the highest altitude, virtually all the mass was in particles < 125 μm.

The IWCs measured for deep convective anvils in the tropics during CEPEX are much larger than the other datasets which measured cirrus not associated with convection. Additional measurements of convective anvils not included in the comparison in figure 1.8 were made over Florida by Garrett et al. (2005). The instrumentation used in the study was sensitive to small particles and the IWC reported was quite high, with maximum values reaching 0.3 g m^{-3}, and the contribution to IWC of small particles throughout the cloud was estimated to be 20 %. Surprisingly the IWC was relatively uniform vertically. Heymsfield et al. (2004) also used the same direct IWC measurements during CRYSTAL-FACE and were able to derive a power law relationship for IWC and the maximum dimension of particles, (for particles < 200 μm) appropriate for cold anvil cirrus which can be used to improve estimates of IWC from past data sets, where direct measurements were unavailable. Garrett et al. (2005) also measured the IWC of a thin tropopause cirrus layer above the anvil with an average of 0.00125 g m^{-3}; the number concentration was always below 500 l^{-1} and the effective radius of the particles was $\sim$ 10 μm. These figures are similar to the thin cirrus measured in the vicinity of Kwajalein, Marshall Islands (e.g. Heymsfield and McFarquhar, 1996; McFarquhar et al., 2000). The IWC generally decreased with temperature, and the measured values ranged from about 0.1 g m^{-3} above -30 °C right down to 10^{-4} g m^{-3} for subvisible cirrus measured at -85 °C (Heymsfield, 1986). Further, Santacesaria et al. (2003) measured subvisible cirrus over the West Indian Ocean, measuring ice water path for cirrus at -86°C with a particle mode around 5–6 μm, resulting in IWCs of 10^{-4} to 10^{-3} g m^{-3}.

1.6 Importance of Clouds in Climate Variability

The present understanding of the direct role of clouds in long-term climate change is incomplete. Clouds are integral to Earth's atmosphere system, regulating both the temperature and moisture structure. In addition, clouds can affect how Earth's climate responds to natural and anthropogenic influences. In order to quantify these effects, researchers have found it useful to separate the problem into two main parts:

- The impact of cloud on the incoming and outgoing radiation at the top of the atmosphere (TOA), i.e. how cloud modulates Earth's radiation budget.

- The mutual interaction of clouds and the climate system, i.e. the cloud-climate feedback problem.
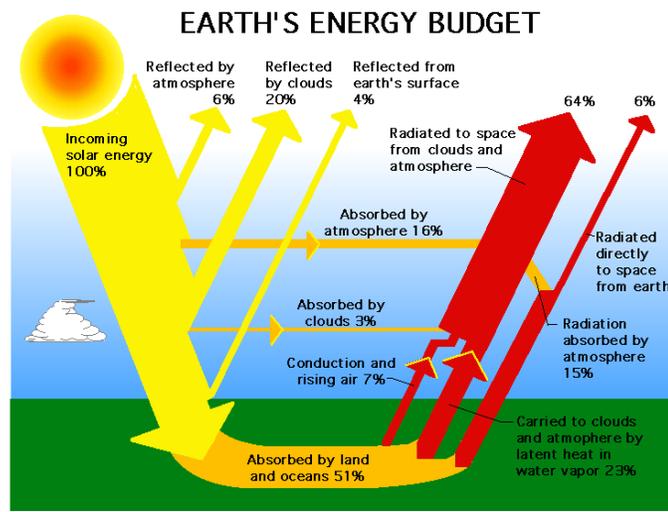
Figure 1.9: Earth's energy budget (Wong, 2002). Showing the average absorption and scattering of incoming solar radiation, as well as the net thermal radiation emitted.

1.6.1 Clouds in Earth's Radiation Budget

The radiation budget represents the balance between incoming energy from the Sun and outgoing thermal (longwave) and reflected (shortwave) energy from the Earth (Figure 1.9). The energy balance determines the planet's temperature and, ultimately, climate. Earth continually receives solar radiation. Of all the solar radiation that reaches the TOA, approximately 30 % is reflected back into space. This 30 % is called Earth's albedo. The solar radiation absorbed by Earth's surface, clouds, and the atmosphere is reradiated in all directions, usually at longer wavelengths (i.e. thermal radiation as opposed to sunlight). If the amount of incoming radiation is in equilibrium with the amount of radiation reflected and reradiated to space the Earth's temperature will remain stable.

To quantify the net effect of clouds on the radiation budget it is useful to define *cloud radiative forcing* (CRF). It is the difference between the radiative budgets at the TOA for cloudy and clear conditions. The radiative budget at the TOA is usually measured in terms of the net downward radiative flux N such that

$$N = Q - F \tag{1.3}$$

where Q is the incoming flux and F is the outgoing flux. Therefore, the CRF is given by the difference between the net downward radiative flux with clouds and that without, i.e.

$$\mathrm{CRF} = N - N_{\mathrm{clear}} \, . \tag{1.4}$$

The CRF varies with the wavelength of the electromagnetic radiation of interest. Normally, it is useful to distinguish between the interaction of clouds with radiation in the shortwave

(SW) region, with wavelengths less than 4 μm (and greater than 0.2 μm), and radiation in the longwave (LW) region, with wavelengths greater than 4 μm (and less than 100 μm). It is generally accepted (e.g. Albrecht, 1992) that the LW cloud forcing is positive, since without absorption by clouds there is increased LW emission to space, and that the SW forcing is negative, since the net absorption of solar radiation is increased by the absence of clouds decreasing the Earth's albedo.

Both models and satellite observations have been used to define cloud forcing. Arking (1991) gives a review of various data sources and methods used to measure CRF. To help understand such studies it is useful to introduce simple models for both the SW and LW CRF (see Schneider, 1996).

The SW CRF can be represented as:

$$\text{CRF}_{\text{SW}} = A_c \mu S \left[r_s - r_c - \frac{t_c r_s r_c}{1 - r_s r_c} \right] \tag{1.5}$$

for a given spatial domain, where A_c is the fractional cloud cover, μ is the cosine of the solar zenith angle, S is the solar constant, r_s is the surface reflectivity or albedo, r_c is the cloud reflectivity, and t_c is the transmissivity of the cloud. The first term inside the brackets in equation 1.5 represents the scattering of solar radiation back into space by the clear sky and surface. The second term inside the brackets denotes reflection to space by clouds, and the final term represents the transmission through the clouds, multiple reflections between the clouds and the surface and transmission back through the cloud to space.

The LW CRF can be represented as:

$$\text{CRF}_{\text{LW}} = A_c \left(\epsilon \sigma T_c^4 - \alpha \sigma T_s^4 \right) \tag{1.6}$$

for a given spatial domain, where ϵ is the cloud emissivity, α the cloud absorptivity, σ is the Stefan-Boltzman constant, and T_c and T_s are the cloud and surface temperatures, respectively. The first term in equation 1.6 represents the infrared radiation emitted by the clouds and the second term denotes the amount of of infrared radiation emitted by the surface which is absorbed by the clouds.

Obviously, the SW CRF will be small if the cloud transmissivity is small and the surface and cloud reflectivity are comparable. Thus, the effect of cloud cover depends primarily on the optical thickness of the cloud, but it also depends upon particle size and ice/water phase. In most places the SW CRF is strongly negative as cloud reflectivity is much larger than surface reflectivity. However, it is possible for the SW CRF to be positive if the cloud reflectivity is smaller than the surface reflectivity, which can occur when thin clouds form over bright surfaces such as snow or ice.

In the LW the CRF depends primarily upon cloud top temperature, which is a function of

cloud top height. The LW CRF is generally positive as the cloud top temperature is almost always colder than the surface. The size of the forcing with respect to a global mean will depend strongly on global cloud cover and cloud top height statistics. For thin clouds, the LW CRF will also depend on cloud emissivity, which is a function of optical thickness.

These simple models come with many caveats, but it is valid to make the general conclusion that the net CRF (i.e. $CRF_{NET} = CRF_{SW} + CRF_{LW}$) is negative as the solar radiation effect will dominate. Satellite observations of instantaneous radiative fluxes at the TOA make it possible to corroborate such a conclusion.

Ramanathan et al. (1989) introduced a method to measure CRF from the NASA Earth Radiation Budget Experiment (ERBE) scanner measurements using a selective algorithm to detect clear-sky scenes. Ramanathan et al. (1989) obtained the geographic distribution of the LW, SW and the net CRF parameters as well as global, annual averages. The global means returned were $31.3 \, W \, m^{-2}$ for LW CRF and $-48.4 \, W \, m^{-2}$ for the SW CRF, which gives a net CRF of $-17.3 \, W \, m^{-2}$. The study of Harrison et al. (1990) also employed a similar method to ERBE data and agreed with these figures, and Kiehl et al. (1994) found the net CRF from ERBE data to be $-19 \, W \, m^{-2}$. The accuracy of the method used by Ramanathan et al. (1989) depends strongly on the ability of the cloud detection algorithm, and additionally, even in correctly identified clear-sky cases the ambient conditions are not likely to be the same conditions where clouds are present (Arking, 1991).

The latest results from ERBE and the Clouds and the Earth's Radiant Energy System (CERES) sensor[8] aboard NASA's Terra satellite (Herring and Kaufman, 2005) indicate that in the global mean, the solar radiative effect dominates, and the cooling due to the presence of clouds is a function of season and ranges from approximately -13 to $-21 \, W \, m^{-2}$. While these values may seem small, they should be compared with the $4 \, W \, m^{-2}$ heating predicted by a doubling of atmospheric concentration of carbon dioxide. The variance of the net CRF, both seasonally and zonally, provides interesting insights into the role of clouds in the climate system, and recent studies have focused more on the geographic distribution of cloud radiative effects (e.g. Kiehl, 1994; Rajeevan and Srinivasan, 2000; Shupe and Intrieri, 2004). The influence of clouds on the radiation field and the LW fluxes will depend on their infrared radiative properties, which are modulated by their physical location and composition (Liou, 1986). In particular, the tropical radiative energy budget has been shown to be highly variable with results indicating that it is caused by changes in tropical mean cloudiness (Wielicki et al., 2002). These changes are not well predicted by current climate models.

In order to fully understand the impact of clouds on Earth's radiation budget, the knowledge of the *sensitivity* of the CRF to changes in cloud properties must be improved. One method

---

[8]The next generation ERBE sensor.

is to reconstruct global observations of Earth's radiation budget using complex radiative transfer models which can fully account for the radiative effects of clouds, and to then investigate the response to cloud variations (see Zhang et al., 2004). In such studies it is imperative to input accurate cloud climatology data (see Section 1.6.3) to obtain good agreement between model results and satellite observations. Investigation of the sensitivity of the CRF to changes in cloud properties links directly into the problem of evaluating cloud-climate feedbacks.

1.6.2 Cloud-Climate Feedback

The influence of clouds on climate via their impact on the planetary radiation balance has now been established. However, clouds can also influence climate by releasing latent heat of condensation, by transporting heat and moisture upwards from the surface, and by precipitating water to the surface. In turn, climate can influence clouds. The prevailing atmospheric conditions will directly affect the types of clouds that form, their vertical and horizontal distribution, their composition and their radiative and hydrological properties. For example, a warmer climate yields clouds that are more likely to be composed of liquid water than of ice crystals. Also, theoretically, warmer clouds have higher liquid water concentrations. There is also observational evidence to support this, for example, NASA's Tropical Rainfall Measuring Mission (TRMM) satellite observed that tropical storm clouds are denser and cover wider horizontal areas over warmer waters (Del Genio and Kovari, 2002). Hence, if warm clouds are denser and have increased cover then they will tend to reflect more solar radiation than colder clouds. This is an example of cloud-climate feedback.

In the context of the climate system, a feedback process is one in which a variable responding to a perturbation in net flux changes in such a way as to further modify the flux. It is a common misconception that because clouds cool the present climate they will act to moderate global warming, i.e. a negative cloud feedback. It is, in fact, the change in CRF associated with a change in climate that constitutes cloud feedback (Cess et al., 1996). Difficulty in observing and simulating cloud properties presents significant barriers to generating simple arguments for how cloud feedback might work. However, to understand cloud feedback it is useful to introduce a simple model (see Albrecht, 1992) to attempt to explain qualitatively how changes in the clouds can affect global climate change.

Consider a model which represents the balance between the incident solar radiation absorbed by the earth system and the infrared radiation emitted to space with a corresponding equilibrium surface temperature, T_s. Thus, climate can change in response to a change in the amount of absorbed SW or emitted LW radiation. If an external forcing $\triangle Q$ is introduced, e.g. due to changes in atmospheric CO_2, then the radiative balance will be

perturbed and the surface temperature will equilibriate to a new value so that the LW emissions to space at the TOA and the changes in net solar input to the Earth system will balance. Denote the response of the planetary radiation balance to the forcing by $\triangle N$, such that $\triangle N = -\triangle Q$. The change in surface temperature $\triangle T_s$ due to the forcing can be expressed as

$$\triangle T_s = \frac{K_0}{1-f}\triangle N = K_f \triangle Q \tag{1.7}$$

where K_0 is the gain (or sensitivity) of the climate system with no feedbacks, f is a feedback factor, and K_f is the gain (or climate sensitivity) with feedback. Positive feedbacks ($f > 0$) will amplify the response, warming climate; negative feedbacks($f < 0$) damp the response, cooling climate.

The response of clouds to changes in temperature is not fully understood, so it is difficult to determine whether clouds represent a negative or a positive feedback.

Now, consider if the feedback factor f in equation 1.7 represents the feedback for all processes in the climate system and the feedbacks are linear and independent (generally valid at the global scale (Colman et al., 1997), then $f = \sum f_j$ for variable j, e.g. water vapour, surface albedo, etc.. The feedback factor for cloud processes, f_c say, is then $f_c = \sum f_{c,i}$ where $f_{c,i}$ is the feedback factor for cloud property i. Thus, the cloud feedback factor is given by

$$f_c = K_0 \sum_i \frac{\partial N}{\partial C_i}\frac{dC_i}{dT_s} \tag{1.8}$$

where N is the net radiation at the TOA and C_i refers to cloud property i. From equation 1.8 it is clear that cloud feedbacks depend on both how the net radiation changes with cloud properties and how the cloud properties change with temperature. However, it still remains to express cloud-climate feedback in terms of a change in CRF due to a change in temperature.

Following from equation 1.4, express the response of the radiation balance as the sum of the change in the clear-sky net flux, $\triangle N_{\text{clear}}$, and the change in CRF, $\triangle\text{CRF}$, i.e. $\triangle N = \triangle N_{\text{clear}} + \triangle\text{CRF}$. Define the clear-sky climate sensitivity $K_{\text{clear}} \equiv \triangle T_s/\triangle N_{\text{clear}}$, then the *cloud-climate feedback* $\triangle\text{CRF}/\triangle T_s$ can be expressed as

$$\frac{\triangle\text{CRF}}{\triangle T_s} = \frac{1}{K_{\text{clear}}} - \frac{1}{K_f} \tag{1.9}$$

(see Cess et al., 1990) where $K_f = -\triangle T_s/\triangle N$ from equation 1.7 is the total sky sensitivity. It will be either positive or negative depending on whether the CRF increases or decreases with surface temperature. A positive cloud-climate feedback increases the climate sensitivity above the clear-sky value, and a negative feedback decreases K_f below K_{clear}. It is still unknown as to whether cloud feedback is positive or negative and this lack of a

quantitative understanding of cloud feedback on the climate is widely acknowledged to be the major obstacle confronting credible prediction of climate change. Essentially, there is a two-pronged attack to tackling the problem:

- The use of satellite data to sample a large number of cloud types, measuring the global distribution, their physical properties (such as thickness and droplet size) and the simultaneous atmospheric state. Then by stratifying the dataset of cloud properties with respect to the measures of atmospheric state (e.g. surface temperature and height profile of atmospheric temperature) it can be determined how the cloud properties change with changes in atmospheric state. This data can then be used to determine cloud feedbacks by using observations of TOA radiation and then determining how it varies with cloud properties or using high fidelity radiative transfer models to calculate the net radiation changes with changes in cloud properties and/or atmospheric state.

- The use of global circulation models (GCMs) to simulate changes in climate to investigate the effective change of the cloud feedback in the model.

The first method provides further motivation for the work undertaken in this thesis, in that current cloud climatologies are somewhat incomplete. The latter method highlights something of a self-perpetuating, cyclical problem, in that GCMs have a reduced accuracy in predicting climate due to a lack of understanding of cloud feedbacks, but they are needed to narrow the uncertainty in the impact of feedbacks on climate. From the various existing GCMs, the range of assessed model sensitivity (traditionally defined as the equilibrium response to a doubling of CO_2) was quoted in the third IPCC[9] assessment report (IPCC, 2001) as between 1.5 to 4.5°C. The differences in cloud feedbacks are largely responsible for the spread in sensitivity.

Because clouds change rapidly over short time and space intervals, they are notoriously difficult to simulate in computer models. Cloud processes are simulated in GCMs through a number of cloud *parameterisations*. Parameterisations are statistical relationships which describe the interaction of small scales with larger scales and include a number of idealisations and approximations. It is widely accepted that cloud parameterisations in current GCMs are an inadequate representation of the behaviour of clouds in the real atmosphere. However, the cloud parameterisation problem is overwhelmingly complicated (see, e.g. Randall et al., 2003), and among other processes they must include: precipitation (via collisions and coalescence), cumulus convection, stratiform cloudiness, interactions between cloud types, turbulence, radiative properties and microphysical processes, including formation and evaporation/sublimation.

---

[9]The International Panel on Climate Change.
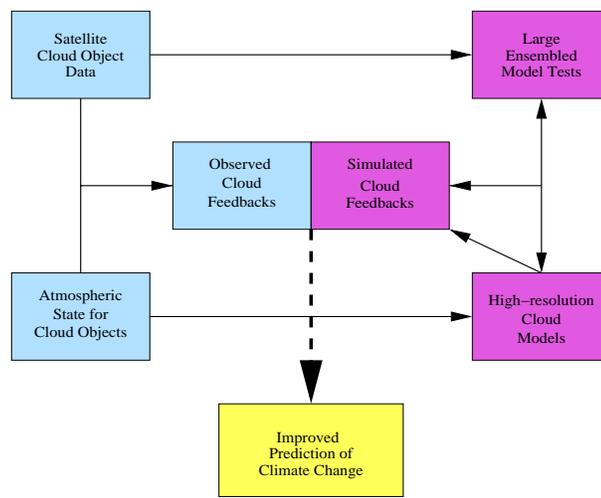
Figure 1.10: A schematic of the approach for cloud object observation and modelling to understand cloud feedbacks (Xu et al., 2005).

Given the complexity of the cloud parameterisation problem it is not surprising that there is a wide range of sensitivity to cloud feedbacks in different GCMs. Zhang (2004) gives a good general comparison of cloud feedbacks in GCMs. One of the most notable studies was by Cess et al. (1990), in which cloud feedbacks from 19 GCMs were compared using fixed season experiments. Taking the impact of clouds on water vapour, lapse rate and surface temperature feedbacks into account, Colman (2003) interprets cloud feedback for Cess et al. (1990) as ranging from approximately -0.6 to $1.5\,\mathrm{W\,m^{-2}\,K^{-1}}$. The variation in cloud feedbacks obtained indicates that the models have significant physical disagreements regarding cloud behaviour.

Recently, a third way to improve the understanding of cloud feedbacks has been suggested. Xu et al. (2005) propose a new methodology based on an integrated observational data analysis and high-resolution modelling approach (see figure 1.10). Satellite data is analysed to generate large samples of "cloud objects", grouped by cloud-system type, and then matched in space and time to the atmospheric state data. A large ensemble of cloud objects is then used to derive the partial derivatives of cloud properties versus atmospheric states. The atmospheric state is also used to drive simulations of high-resolution cloud models and the simulated cloud feedbacks can be analysed and compared to the observed cloud object analysis to further improve the high-resolution cloud model. The need for using large ensembles of cloud objects reinforces the need for high resolution data on cloud properties, which indicates the importance of cloud climatology studies.

1.6.3 Cloud Climatology and Cirrus Effects

According to Arking (1991), in addition to cloud amount, there are essentially three classes of parameters which determine a cloud's effect on the radiation field:

1. the environment above and below the cloud, including aerosol distributions, the temperature and humidity profiles and surface optical properties

2. the macrophysical structure of the cloud, i.e. horizontal extent, optical thickness, temperature, and the vertical and horizontal inhomogeneities

3. the microphysical structure, i.e. particle size, phase and shape, and the statistical distribution of these parameters within the macroscale volume.

Therefore, the exact effect that clouds have on Earth's radiation budget depends not only on how clouds of different types are distributed over the Earth, but also on how the above variables that characterise the clouds are distributed over the Earth and how they vary over time.

A cloud climatology is a database which contains a time series of the geographic distribution of cloud types and properties. Therefore, what is required for studies of the cloud radiation budget and for cloud feedbacks is a complete climatology of the parameters outlined above. Significant cloud feedbacks can result from changes in global mean cloudiness as small as 1% per decade, or regional change of 1% per year (Xu et al., 2005). Thus it is important that any cloud climatology used be as accurate as possible.

Several datasets on clouds have been compiled (e.g. Wylie et al., 1994), but by far the largest effort is the International Satellite Cloud Climatology Project (ISCCP; Rossow and Schiffer, 1991). ISCCP began collecting cloud frequency data from satellite instruments in 1983 as part of the World Climate Research Programme (WCRP). ISCCP uses global coverage data collected from most of the world's geostationary satellites (GOES-E, GOES-W, GMS, INSAT and METEOSAT) and also from the NOAA polar orbiting satellites. The data from the NOAA satellite comes from the nadir viewing Advanced Very High Resolution Radiometer (AVHRR). The ISCCP data products are listed in full in Rossow and Schiffer (1999). The products include: cloud amount and distribution information, total cloud properties (including cloud top temperature, cloud top pressure, cloud particle size and cloud optical thickness), cloud-type information, surface properties, radiances and atmospheric properties. ISCCP uses two channels from each satellite instrument to evaluate the radiance data: a thermal channel near 11 μm to detect clouds and estimate their height from the radiating temperature, and a visible channel near 0.6 μm to estimate cloud transmissivity (during the day) to better estimate the cloud top height. Further cloud properties are inferred by using the ISCCP cloud analysis algorithm (Rossow et al., 1985), which employs a single-layer, plane-parallel cloud model with a specified particle-size distribution. Once a satellite image pixel is flagged as cloudy, the cloud type is then classified by the derived cloud top pressure and optical thickness. The cloud classification scheme is shown in figure 1.11.
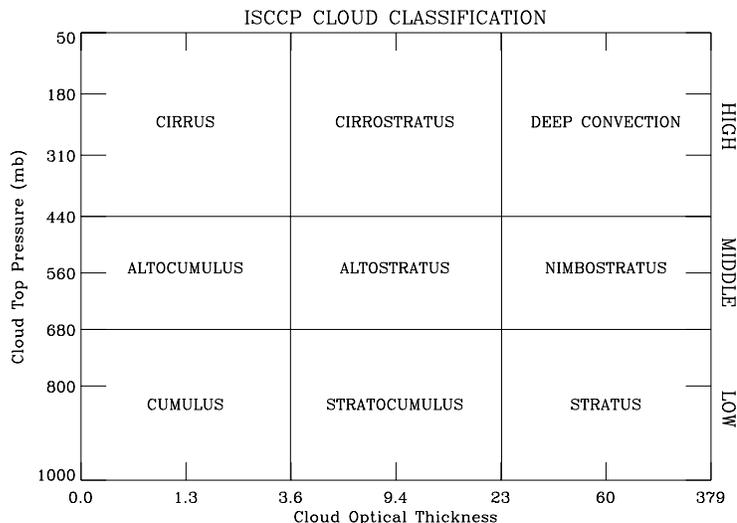
Figure 1.11: Cloud-type definitions used in the ISCCP-D series datasets (Rossow and Schiffer, 1999). All low and middle cloud types are separated into liquid and ice types; all high clouds are ice.

ISCCP is widely considered as the first reference (particularly for climate modelling purposes) concerning cloud climatology data. ISCCP compares favourably with a number of cloud climatology datasets and quantifies global variation in cloud properties to mesoscale resolution, 30 km on a 3-hourly basis (although little if any information on cloud vertical structure can be resolved). However, ISCCP's major shortcoming, which has been clearly established by several studies, is that ISCCP underdetects upper tropospheric cirrus (Liao et al., 1995 (1); Liao et al., 1995 (2); Jin et al., 1996; Wylie and Menzel, 1999). This is due both to a lack of detection sensitivity at low optical depths and to misidentifying cirrus layers which lie above lower-level cloud deck. This leads to an underestimate of cirrus cloud amount by at least $10\,\%$, with the greatest error over the tropics.

Jin et al. (1996) compared ISCCP climatologies of high-level clouds to datasets obtained from the High Resolution Infrared Sounder (HIRS) which was part of the payload on the NOAA polar orbiters. HIRS is a 19-channel infrared radiometer and is used to obtain cloud properties via the CO_2-slicing method (see Wylie et al., 1994). Since the multiwavelength CO_2-slicing analysis of HIRS can infer cloud-top pressures independently of the transmissivity of the cloud, this technique can detect very thin cirrus clouds that might be missed or confused with some broken low-level clouds by a thresholding technique such as that used in ISCCP (Wylie and Menzel, 1989). Jin et al. (1996) found that the geographical and seasonal variations in the two climatologies were well-correlated, but that HIRS high-level clouds were larger than the ISCCP values by about $12\,\%$. The differences occur in the thin cirrus category, i.e. for visible vertical optical thickness $\tau_{vis} < 1.3$, particularly over oceans. The differences are due to the increased thin cirrus detection sensitivity of HIRS compared to ISCCP. The minimum value of τ_{vis} that can be detected by ISCCP for

high-level clouds is estimated to be 0.1–0.2 over the ocean, and $\tau_{vis} \approx 0.2 - 0.4$ over land (Rossow et al., 1993). HIRS allows for the detection of thinner cirrus, $\tau_{vis} \approx 0.1$. In the analysis of Jin et al. (1996), when the optically thinnest high-level clouds ($\tau_{vis} < 0.5$) were removed from the HIRS results, the zonal-mean difference compared to ISCCP was reduced to less than 2%. This difference was concentrated over the ocean, and can be put down to the large field-of-view of the HIRS instrument compared to AVHRR which results in a slight overestimate of cloud fraction.

Figure 1.12 shows the ISCCP 10-year global mean cirrus frequency from 1989–1999 and figure 1.13 shows the HIRS global mean cirrus frequency from 1989–2000. It can be seen from the two figures that the ISCCP and HIRS are in good agreement with zonal and relative trends of cirrus amount, but the frequency of cirrus occurrence is much higher in the HIRS dataset, thus showing the higher detection sensitivity of HIRS. Wylie and Wang (1997) compare the HIRS and SAGE II climatologies. The two data sets generally agree on the locations of clouds; however, SAGE II reported higher frequencies by 12 to 22 percent. This is primarily due to the differences in sensor sensitivity.

Liao et al. (1995 (1)) compared ISCCP to global high-level clouds identified in the Stratospheric Aerosol and Gas Experiment II (SAGE II) occultation measurements (see section 2.3). Because the SAGE II instrument directly views the Sun through the atmosphere over very long path lengths ($\sim$ 200 km), the measurements are very sensitive to the presence of particle scattering. Quantitatively, SAGE II only needs an integrated horizontal optical thickness of > 0.16 to exceed the background aerosol and gas extinction, which is equivalent to an extinction coefficient of 0.0008 km^{-1}. Assuming the cloud to be 75 km in horizontal extent, with a vertical extent of 2.5 km (see Liao et al., 1995 (2)), such a cloud would have τ_{vis} of 0.005. Thus, SAGE II is able to detect subvisible cirrus and the detection sensitivity is 20 times that of ISCCP. Liao et al. (1995 (1)) showed that compared to the more sensitive SAGE II results about one third of all high-level optically thin clouds ($\tau_{vis} < 0.1$) are missed by ISCCP. If cirrus with optical thickness < 0.05 (equivalent to an extinction of 0.008 km^{-1}) is neglected, then there is good agreement between ISCCP and SAGE II for zonal monthly mean high-level cirrus amounts. In addition, SAGE II has the ability to vertically resolve cloud structure, and Liao et al. (1995 (1)) found that the global mean thin cirrus amount is about 9% for clouds that do not have any other high-level cloud below them (i.e. isolated thin cirrus), and 12% total thin cirrus amount.

SAGE II had a much higher sensitivity to thinner clouds and could detect subvisible layers. Figure 1.14 illustrates the detection sensitivity of SAGE II, showing a single year of global seasonal mean frequency of occurrence of clouds at 16.5 km, where only near tropopause cirrus and particularly optically thin cirrus is expected. The SAGE II quarterly mean values for 1989 in figure 1.14 are representative of the 6-year climatology as in Wang et al.
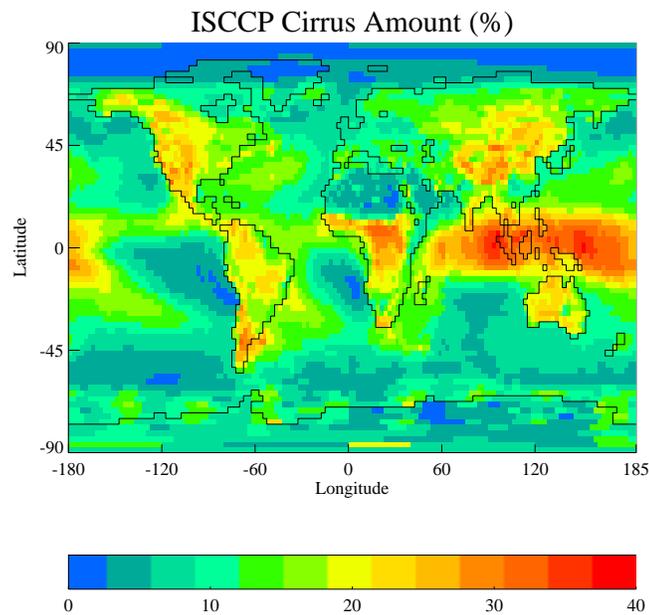
Figure 1.12: A 10-year global mean cirrus amount compiled from ISCCP data. The data is gridded on a 2.5 by 2.5 degree resolution (Dean et al., 2005)).

(1996).

Further evidence that ISCCP underestimates optically thin cirrus comes from in situ studies (e.g. McFarquhar et al., 2000; Luo et al., 2003). During CEPEX McFarquhar et al. (2000) noted that a spatially thin layer of cirrus, with both base and top above 15 km, was observed in the central Pacific tropics 29 % of the time, with an average τ_{vis} estimated at 0.01. Clear sky was below the thin cirrus for 36 % of the time. Such laminar cirrus was observed frequently over the tropics by Winker and Trepte (1998) during the 10 days of the Lidar in Space Technology Experiment (LITE). LITE data confirmed the SAGE deductions and showed thin cirrus layers near the tropopause with horizontal extents of up to 2700 km and with thicknesses from a few hundred meters to one kilometre. Also, Sunil Kumar et al. (2003) showed that ground observations in Gandaka, India measured subvisible cirrus about 32 % of the time during 6 months of measurements. The figures for the frequency of occurrence of optically thin cirrus in these studies are more in agreement with the SAGE II climatology (Wang et al., 1996) than with ISCCP.

As ISCCP underestimates optically thin high clouds, numerous studies on cloud effects on the radiation budget which use ISCCP cloud fields (e.g. Chen et al., 2000) will be missing the contribution of thin and subvisible cirrus. In addition, further errors will be introduced because comparisons to SAGE II data have shown ISCCP also underestimates the cloud top height of moderate to thin cirrus (Liao et al., 1995 (2)), and recent studies (Sherwood et al., 2004) of aircraft-mounted lidar measurements have shown that thermal satellite imagery, such as that used by ISCCP, will also underestimate cloud top heights of optically
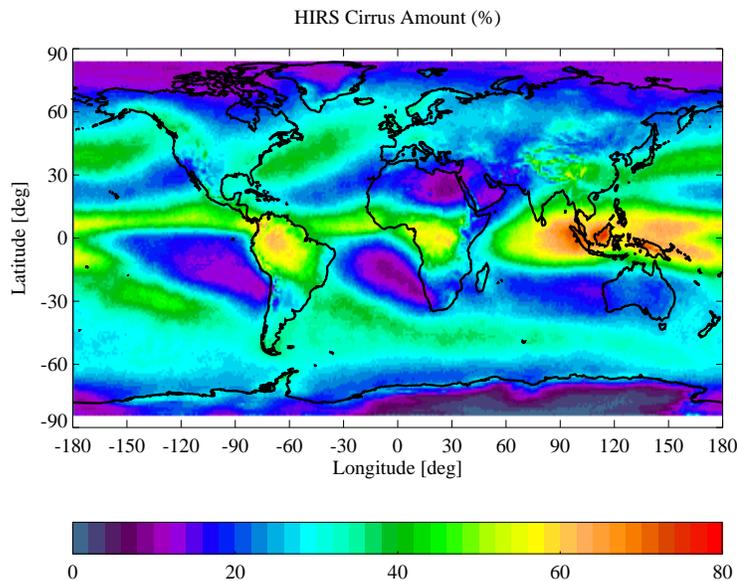
Figure 1.13: An 11-year summary of global mean cirrus amount compiled from HIRS data (1 degree resolution). Data sourced from the University of Wisconsin HIRS Cloud Climatology (Wylie, 2004).

thick deep convective cirrus.

Studies comparing model results to ISCCP observations have generally shown that GCMs underpredict thin cirrus (e.g. Fowler and Randall, 1999; Dean et al., 2005). Thus, the deficiency of thin cirrus in ISCCP serves to highlight the difficulty of including thin cirrus effects in GCM studies of future climate. One of the reasons for this difficulty is the lack of understanding about how thin cirrus is created, and thus the dynamic processes in the GCM which create the cirrus will be poorly resolved (see Del Genio, 2002).

The effect of thin and subvisible cirrus on the radiative budget of the Tropics is not negligible, and their other impacts, such as enhancing upper-troposopheric vertical motions and the lower-stratospheric water vapour, are also significant.

High-level ice clouds cause most of the annual mean changes in the global TOA LW radiative fluxes (Chen et al., 2000). Although the effect of cirrus is not as large as the other ice clouds, it is still very effective at trapping longwave radiation. Increasing the amount of optically thick cirrus lowers the radiating temperature and thus traps heat below the cloud. In the thermal infrared, ice is highly absorptive and it reradiates about half the energy it receives from the Earth back downward. Thus, on average cirrus clouds produce net heating of the planet. However, for some cirrus the LW effect does not always dominate over the SW, and it is possible for cirrus to also be a strong reflector of incoming solar radiation. The effect which dominates depends on both the size and the shape of the crystals. Large crystals efficiently scatter sunlight into the forward direction, whereas smaller crystals will have enhanced backscattering (Takano and Liou, 1989); thus higher albedos will be
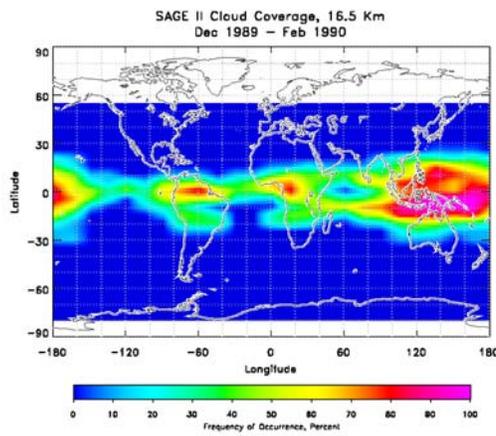
(a) December 1989 - February 1990
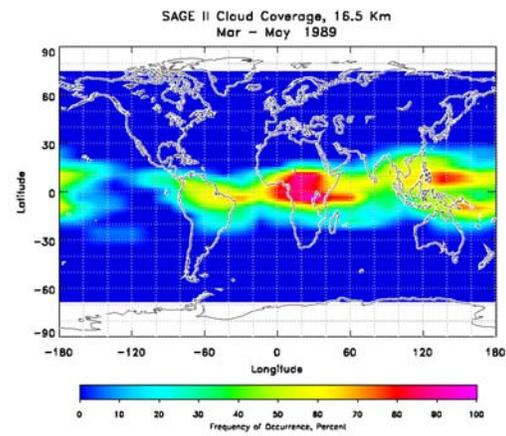
(b) March - May 1989
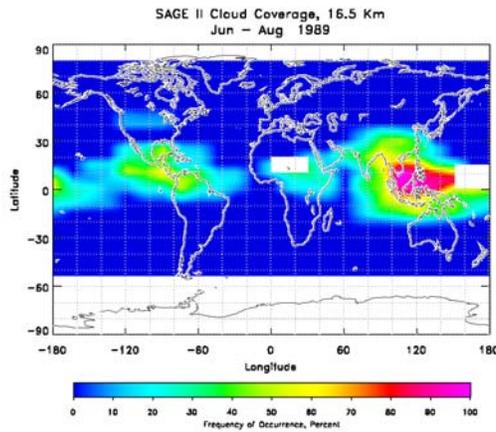
(c) June - August 1989
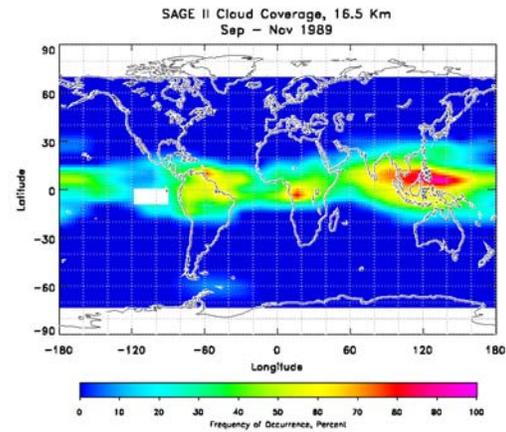
(d) September - November 1989

Figure 1.14: The SAGE-II cloud (high cirrus) frequency at 16.5 km for 1989 by seasonal global mean (Thomason, 2003).

promoted in cirrus with smaller particles. Cirrus clouds found high in the troposphere tend to be composed of smaller crystals than those at higher temperatures, and thus tend to backscatter more efficiently. However, cirrus at cold temperatures also tends to have lower ice water content, which causes decreased optical thickness and albedo. Equally, crystals with a larger projected area for the same volume, i.e. spatial crystals and those with rougher surfaces as opposed to pristine faceted crystals, will scatter more isotropically and lead to increased optical depth, albedo and emissivity for the same ice water path (see Garrett et al., 2003). The importance of crystal microphysics in determining the radiative fluxes at the TOA is the main reason that it is so important to obtain global distributions of size and shape from satellite remote sensing instruments.

Satellite information suggests that the albedos of tropical cirrus are greater than cirrus outside of the tropics and that albedos often exceed 30 % and sometimes are as high as

60 % (Harrison et al., 1990; Ramanathan and Collins, 1991). The reasons for this are not fully understood. Predominantly, cooling occurs due to cirrus associated with convection. This can take the form of thin cirrus layers overlying convective towers or anvils (e.g Garrett et al., 2005), or deep convective cells with an abundance of small crystals near the top (see Heymsfield and McFarquhar, 1996). However, such regions cover relatively small areas and Massie et al. (2002) suggest that ninety percent of the cirrus clouds near the tropopause are located outside regions of deep convection. Kiehl (1994) shows that high cloud coverage in the tropical Pacific is directly related to sea surface temperature (SST), increasing from nearly 0 % coverage at 299 K to 30 % coverage at 302 K. In addition, Ramanathan and Collins (1991) show that more LW radiation emitted by the surface is trapped when SST exceeds about 300 K, which signifies an increase in upper level cloudiness. It may be that deeper and more extensive cirrus (and stratiform) regions are produced when SST $\gtrsim$ 300 K, and that this is responsible for albedos which exceed 30 %. It is also likely that a high proportion of tropical cirrus is formed due to convective blow-off (see Massie et al., 2002) which (due to larger crystals precipitating out) is often associated with higer optical depths and albedo. Thus, when determining the radiative effects of cirrus it is important to distinguish between convective, optically thick, optically thin and subvisible cirrus.

Subvisible cirrus are ubiquitous over tropical regions and are not only radiatively important, but also play an important role in the dehydration of air entering the tropical lower stratosphere (Hartmann et al., 2001; Luo et al., 2003). Subvisible cirrus usually occurs within 2 km of the tropopause. If thin layer clouds occur in the upper troposphere in otherwise clear conditions they will be strongly heated by the absorption of upwelling longwave radiation. However, ice clouds near the tropopause can provide cooling if there is underlying convective cloud which is sufficiently cold. McFarquhar et al. (2000) calculated cloud radiative forcings using in situ microphysical properties of subvisible cirrus with an average τ_{vis} of 0.01 and obtained an average of 1.58 W m^{-2}, of which 2.19 W m^{-2} occurred in the infrared and -0.61 W m^{-2} in the solar. Jensen et al. (1996) hypothesised two viable formation processes for subvisible cirrus:

- Energetic convection (e.g. cumulonimbus clouds) transport large amounts of ice water to the upper troposphere and generate extensive cirrus outflow anvils. Crystals bigger than $\sim$ 20–30 μm will precipitate out of these anvils (in a matter of hours), leaving behind an optically thin layer of small crystals.

- Subvisible cirrus can also be generated by slow, synoptic scale uplift of humid layers. Given the very cold temperatures at the tropopause, the uplift can generate the ice supersaturation levels required for homogeneous nucleation. In this case, the crystals do not grow large enough to precipitate, so the low fall velocities allow them to remain

in the narrow supersaturation layer for at least a day.

The aridity of the stratosphere has long been established, but the exact role subvisible cirrus plays in stratospheric dehydration is not fully understood (see Sherwood and Dessler, 2000). The first method of subvisible cirrus formation allows a long enough time for radiative cooling to take place in air passing through the tropopause, thus forming ice crystals which sediment out and therefore dehydrate air before entering the lower stratosphere. In the latter method of formation the cirrus layer is generated away from convective regions. The radiative energy absorbed drives lifting of the cloud layer and the cloud may exist for days. The persistence of the non-convective cirrus layer allows enough time for radiative cooling to dehydrate slowly ascending air as it passes through the tropopause and into the stratosphere. Effectively, the presence of the subvisible cirrus ice crystals in the narrow supersaturated layer will freeze dry air passing through it vertically.

1.7 Summary of Research Impetus

Section 1.6 showed the importance of cloud in climate variability. In particular, the significant effect of cirrus on Earth's TOA radiative budget was established and the complexity of predicting cirrus-climate feedback was discussed. Cirrus and in particular subvisible cirrus also play a vital role in the vertical distribution of atmospheric water vapour.

Sections 1.4 and 1.5 have shown the complexity of cirrus formation, the complexity of the mechanisms of ice crystal growth, and the large variety of resultant ice crystal shapes and sizes, thus illustrating the incompleteness of the current understanding of cirrus and hence the difficulty of predicting its occurence and composition. As the radiative properties of cirrus depends on its microphysics as well as the physical location, in order to include cirrus contributions in future predictions of climate, it is important to fully understand the effects of cirrus on past and present climate, and how they vary with the macrophysical and microphysical properties. The main method for improving the knowledge of such effects is through the use of accurate cloud climatology data in both radiative budget and cloud feedback studies. Section 1.6.3 outlined the important properties to be included in cloud climatology, and showed the problems with current climatologies. In particular, the fact that ISCCP underpredicts thin cirrus and also has an associated uncertainty in cloud heights and microphysics. LITE and SAGE II both provide improved measurements, but LITE sampled the planet too sparsely to produce a true global climatology and SAGE II also had limited spatial and temporal sampling.

The hypothesis to be addressed in this thesis is that infrared limb sounding of cloud properties can improve current cirrus climatology data. Therefore, the primary objective is to develop a cirrus property retrieval method for MIPAS and to characterise its capabilities.

The following chapter will familiarise the reader with the cirrus limb viewing problem, then introduce the MIPAS instrument and the cirrus retrieval scheme.

Chapter 2

Cirrus in the Limb View

2.1 Introduction

The focus of this chapter is the motivation and techniques used to observe high level clouds with limb viewing instruments. Also, the chapter will introduce MIPAS and the method used to derive cloud information from MIPAS infrared limb measurements. It is beyond the scope of this chapter to give a detailed general review of satellite instruments and respective techniques used for remote sensing of clouds. For such material the reader is directed to Rossow (1989) and for more recent texts, to Ewen (2002) or Yang and Baum (2003). Moreover, an introductory review of space based measurements of cirrus properties can be found in Minnis (2002).

2.2 Satellite Remote Sensing of Cirrus Clouds

Remote sensing (or sounding) is defined as a means of obtaining information about an object without coming into contact with it. Satellite remote sensing involves gathering measurements about features on the Earth's surface or properties of the Earth's atmosphere from orbiting spacecraft. The two types of sensor systems commonly carried onboard satellites are known as *passive* and *active*. *Passive* remote sensing techniques are used to observe the extinction, emission and scattering of natural radiation. The source of which can can be the Sun, stars, and the Earth's surface and atmosphere. The most common passive sensor is a filter radiometer. An *active* remote sensing system uses an artificial source to propagate its own electromagnetic radiation and measure the intensity of the return signal. Examples of active sensors include lidar and microwave radar. Active remote sensors are principally deployed on land, ships, and aircraft platforms and recently there has been an increasing use (or planned use) aboard satellites. Previously, there were several drawbacks for using active sensors in space, such as onboard power requirements and high fabrication

costs in order to ensure the technology was rugged enough for space travel. However, as more is known about the radiation source *a priori*, active sounders have the potential to resolve the location of atmospheric properties better than passive methods. This potential may be realised in the future with the launch of CALIPSO and CloudSat planned for October 26th 2005, two experimental satellites which will use radar and backscatter lidar respectively to sound the vertical structure of clouds and cloud properties from space (see Hu et al., 2001; Miller and Stephens, 2001).

Satellite data has been employed for cloud observations since the first satellite images were returned (e.g. Krueger and Fritz, 1961). However, the progress from using passive satellite sensing for mere observations to the recovery of physical and radiative properties of cloud has been slow, owing to the complexity of the cloud interaction with the radiation field of the atmosphere. Many different techniques have been developed to determine cirrus properties from passive satellite observations. Predominantly, these techniques focus on a limited number of parameters to quantify cirrus clouds. These include the spatial coverage, top and base altitude, top and base temperatures, optical depth, vertical ice water path and the effective particle size and shape. Techniques are developed specifically for instrument data sets from different parts of the electromagnetic spectrum and different observation geometries. There are two main viewing geometries used by satellite observations: limb and nadir. Limb observing instruments scan the atmospheric layers above the horizon and provide good vertical resolution. Nadir viewers look downward at the surface, observing the atmospheric volume directly below the instrument, providing good horizontal resolution. With respect to measuring optically thin cloud, limb observations have a greater inherent detection sensitivity as compared to nadir observations (Dowling and Radke, 1990; Liao et al., 1995 (1); Liao et al., 1995 (2)) due to the small elevation angle clouds are viewed at, which maximizes their opacity. Both the SAGE II and HALogen Occultation Experiment (HALOE) instruments have demonstrated the detection sensitivity of limb sounders to even tenuously thin (i.e. subvisible) cirrus (Wang et al., 1996; Massie et al., 2003).

Radiances observed by passive nadir sensors of visible, near-infrared and mid-infrared wavelengths have been widely used to retrieve cloud properties (e.g. Nakajima and King, 1990; Platnick et al., 2003). Infrared radiation in the atmospheric window region (i.e. 8–13 μm), contains a wealth of information on the properties of ice clouds (Yang et al., 2001). Various algorithms have been suggested for inferring the microphysical properties of ice clouds from high spectral resolution infrared observations within the 8–13 μm atmospheric window (Ackerman et al., 1990; Chung et al., 2000; Huang et al., 2004). Currently, such algorithms do not exist for high resolution infrared limb spectra. However, Kahn et al. (2002) compared limb radiative transfer simulations to high resolution spectra of nongas-absorption transmission from Atmospheric Trace Molecule Spectroscopy Experiment (ATMOS) so-

lar occultation data, and found that calculations performed with nonspherical ice particle scattering models fit the observed spectra more accurately than simulations based on the approximation of spherical particles. Although large separation was noted in the quality of fit for varying ice models, Kahn et al. (2002) were unable to assess whether crystal size distribution could be retrieved. Various studies have noted the potential of high resolution infrared limb spectra to derive cirrus microphysical properties (e.g. Milz et al., 2004), but todate the technique has not been exploited.

2.3 Satellite Limb Sounding

Viewed from the side, the Earth looks like a flat circle, and the atmosphere appears like a halo around it. This glowing halo is known as the limb. Breathtaking views of the atmosphere looking through the limb can be obtained from satellites and space shuttles, an example of which can be seen in figure 1.3. Towers of convective cloud mark the extent of the troposphere. The stratospheric layer is easily identified by the strongly scattered blue sunlight which fades dramatically to the black of space as the atmosphere thins.

Although impressive, camera-like remote sensing at the Earth's limb does not return the depth of information required to derive atmospheric composition. Generally, satellite limb sounding instruments are non-imaging and take measurements of some property which is dependent on height in the atmosphere. The reference to "sounding" is an application of an old nautical term; the investigation of the state of a medium at different depths (originally applied to the ocean). The two main types of measurements that limb viewing instruments take are *emission* and *occultation*.

Occultation techniques view the Sun (although the moon and bright stars can also be used) directly through the atmosphere, measuring the amount of absorption of radiation at ultraviolet, visible or infrared wavelengths. Solar occultation uses a similar viewing geometry to the emission technique and can infer vertical profiles of a number of trace constituents (including ozone) with high vertical resolution. The limitation of the technique is the spatial coverage, as measurements can only be made at sunrise and sunset events which yields only two profiles per orbit. However, the coverage can be improved by using lunar and stellar occultation measurements in addition to solar.

The limb emission technique looks through the limb with cold space as a background and measures the longwave radiation (infrared or microwave) thermally emitted in the atmosphere along the line-of-sight. As the horizontal limb paths are very long there is a large opacity involved, hence limb emission sounders are useful for measuring temperature and trace constituents. Although the horizontal resolution is limited, the spatial coverage is much better than the solar occultation technique as the atmosphere can be sampled many
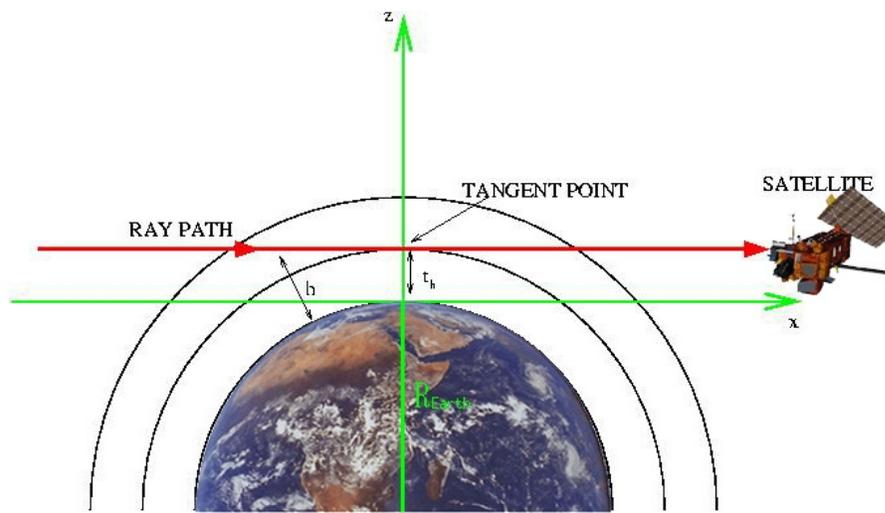
Figure 2.1: Satellite limb viewing geometry. The tangent height of the limb ray path is labelled t_h.

times per orbit both day and night.

The limb viewing geometry is illustrated in figure 2.1. A limb radiometer receives radiation emitted by the atmosphere along a ray path which may be identified by the tangent point. The tangent point is the point at which the ray path (i.e. the line-of-sight of the instrument) passes closest to the surface. The tangent height is usually thought of as the altitude of the atmosphere nominally viewed by a limb sounder. The majority of the emission originates in the few kilometers of the atmosphere immediately above the tangent point because of the exponential decrease of atmospheric density and pressure. This leads to the inherently high vertical resolution of limb sounders.

The limb emission technique is most useful in the upper troposphere and above. The reasons for this are twofold:

- The horizontal resolution is poor. The narrow layer of the atmosphere from which most of the emission originates is often known as the tangent layer and is usually of the order of 200 km in the horizontal. Interpretation of changes in the atmospheric state over this distance is problematic.

- The presence of clouds in the instrument field of view. Clouds in the line-of-sight can act as grey bodies of significant opacity (particularly water clouds) and thus alter the emerging radiation, introducing serious problems sensing atmospheric temperature and gaseous profiles below the cloud level.

In clear sky cases (when cloud is not present in the line-of-sight), the atmosphere can be considered to be non-scattering and as such all of the radiance detected by an infrared limb sounder can be considered to originate from above the tangent height. However, when cirrus clouds are present in the line-of-sight they have a semitransparent appearance

in the infrared and introduce a multiple scattering effect which effectively means that the instrument will detect radiance originating from below the tangent height. Although the presence of cirrus may preclude accurate retrieval of temperature and atmospheric constituents, valuable information may be inferred about the cloud itself. As scattering is dependent on cloud microphysics (Stephens, 1980), this effect will give indirect information about the cloud particles. This is a key point to appreciate throughout this thesis and is discussed further in the following section when considering previous limb observations of cirrus.

2.4   Previous Measurements

Predominantly, limb sounders that are used to obtain information on tenuous cirrus clouds have been occultation instruments, e.g. SAGE II, HALOE, ATMOS (Woodbury and Mc-Cormick, 1983; Wang et al., 1996; Hervig and McHugh, 1999; Kahn et al., 2002). However, Mergenthaler et al. (1999) and Spang et al. (2002) have demonstrated that cirrus features are frequently observed by infrared emission instruments.

The majority of cirrus studies using occultation data have been limited to detecting the presence of cloud at a given viewing tangent height, without attempting to infer microphysical properties. The SAGE II and HALOE instruments are well suited to the retrieval of aerosol extinction. HALOE has 4 infrared channels at 2.45, 3.40, 3.46 and 5.26 μm and SAGE II has 7 channels at 0.365, 0.448, 0.453, 0.525, 0.6, 0.94 and 1.2 μm covering the visible and near-infrared. These occultation instruments can detect subvisible cirrus as regions of enhanced extinction near the tropopause which do not obscure the solar beam. Extinction due to cirrus can be notably larger than extinction due to aerosols, and this characteristic has been used to detect cirrus with SAGE II measurements (e.g. Woodbury and McCormick, 1983; Liao et al., 1995 (1)). Wang et al. (1996) distinguished among background aerosol, subvisual cirrus, and opaque clouds by ranges of extinction at 1.2 μm: between 1×10^{-6} and 2×10^{-4} km^{-1} for background aerosol and between 2×10^{-4} and 2×10^{-2} km^{-1} for subvisual cirrus. A similar method can be used with HALOE extinction at 5.26 μm, for which Hervig and McHugh (1999) suggest a threshold for cirrus extinction between 3×10^{-4} and 8×10^{-4} km^{-1}. In both cases extinction values of $\sim 10^{-2}$ km^{-1} or greater are classified to be opaque clouds. The SAGE II and HALOE measurements measure transmission values between unity and ~ 0.03 and retrieve the aerosol extinction coefficient. However, extinction values greater than $\sim 2 \times 10^{-2}$ km^{-1} over an effective path length of ~ 300 km can yield a transmission too close to zero for the instrument to record a measurement.

In the presence of volcanic aerosols, cirrus and aerosol extinctions retrieved from occul-

tation instruments can overlap and the magnitude of extinction alone is insufficient to separate cloud events from aerosol. Additionally, the magnitude of cloud extinction can be in considerable error due to the measurement geometry and the inhomogeneous spatial distribution of cirrus in the line-of-sight (Hervig and McHugh, 1999). Kent et al. (1997) and Hervig and McHugh (1999) suggested that the shape of the extinction spectra should be used to distinguish between cloud and background aerosol for SAGE II and HALOE respectively. For clouds with an effective particle radius much larger than 10 μm the extinction is relatively constant over wavelength regardless of the particle size distribution. Spectral uniformity can therefore be used as an indication of cloud presence, as aerosols generally have small particle sizes and often exhibit distinct spectral character depending on chemical composition, e.g. sulphate aerosol. The reported extinction is also a useful measure which can give an indication of the optical thickness of cloud, but due to the complicated geometry, scattering, and inhomogeneous path sample volume there is a large associated error (usually an underestimate). This problem, together with the flat extinction spectra expected for typical cirrus cloud effective radii (i.e. $\gtrsim 10$ μm), means that the information on cloud microphysics in the occultation measurements is limited. The SAGE II and HALOE data have, however, proved very useful for global high level cloud frequency studies (see section 1.6.3), particularly because of the high resolution vertical sampling (0.5 km and 0.3 km respectively), e.g. figure 2.2 shows the zonal distribution of occurrence frequencies of subvisual cirrus for 1993–1998 from Massie et al. (2002) for HALOE 3.46 μm data. The occultation view, (where geographic coverage allows) has proven particularly useful for observing PSCs, due to their low opacity and high altitude above much of the background aerosol, and has been employed notably by Lee et al. (2003) to data observed by the Improved Limb Atmospheric Spectrometer (ILAS).

When opaque clouds are present the solar beam is blocked in the occultation view, whereas the thermal emission measurement will generally observe more radiation. In addition, when tenuously thin clouds are present the thermal emission measurement will be enhanced due to scattering, whereas the occultation measurement is likely to be reduced by scattering solar radiation out of the line-of-sight. Thus, intuitively the thermal emission technique has a greater potential range of sensitivity for differentiating between cloud optical depth than the occultation technique.

The study of Spang et al. (2002) used the CRyogenic Infrared Spectrometers and Telescopes for the Atmosphere (CRISTA) instrument to focus on the observation of subvisible cirrus around and above the tropopause, and showed that at tropical latitudes it occurred with frequencies > 40 % and up to 20 % above the tropopause (at altitudes of about 18 km). However, CRISTA had relatively short measurement periods of only one week in November 1994 and August 1997. As CRISTA scanned from the top of the atmosphere downward,
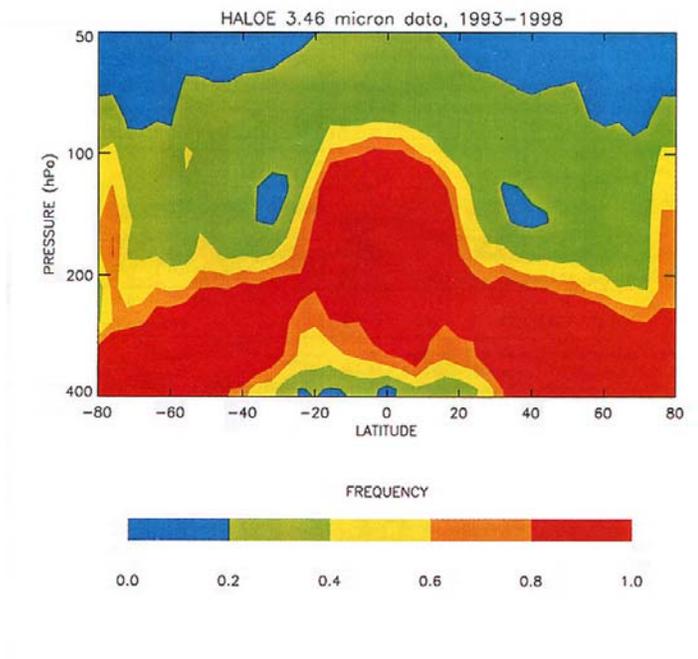
Figure 2.2: Zonal distribution of occurrence frequencies of subvisual cirrus for 1993–1998, using 3.46 μm HALOE extinction data for the extinction range 0.001 to 0.02 km^{-1}. As given in plate 1 of Massie et al. (2002).

the rough cloud top height was specified by the first tangent height in which cloud was detected. Figure 2.3, taken from Spang et al. (2002), shows the global cloud top height distributions from the entire set of radiance profile measurements ($\sim 43,000$ profiles) from August 1997. The minimum observable tangent height for CRISTA-2 during August 1997 was between 8 and 10 km. The total number of profiles for the August 1997 mission is comparable to a measuring period of 3.5 years for an occultation experiment. A belt of high clouds ($\gtrsim 16\,\mathrm{km}$) is clearly observed over the tropics, as is the expected area of reduced cloud activity in the subtropics (e.g. indicated by the cloud-free regions over Australia). The observations are also in good agreement with the August data from the 11-year HIRS climatology (1989–2000) in figure 3.9 (chapter 5). Spang et al. (2002) noted that the cloud indication scheme did not allow the definitive distinction between optically thick and thin clouds, but the wide spectral range of the CRISTA measurements had the potential to determine microphysical parameters of cloud type.

Mergenthaler et al. (1999) used data from the Cryogenic Limb Array Etalon Spectrometer (CLAES) emission measurements to study the seasonal evolution and spatial distribution of cirrus in the upper troposphere. The study used the 12.8 μm channel in an empirical threshold method to separate cloud events from the transient background aerosol. The cloud frequencies obtained by Mergenthaler et al. (1999) were comparable to the SAGE II climatology of Wang et al. (1996), which shows that CLAES data contain information
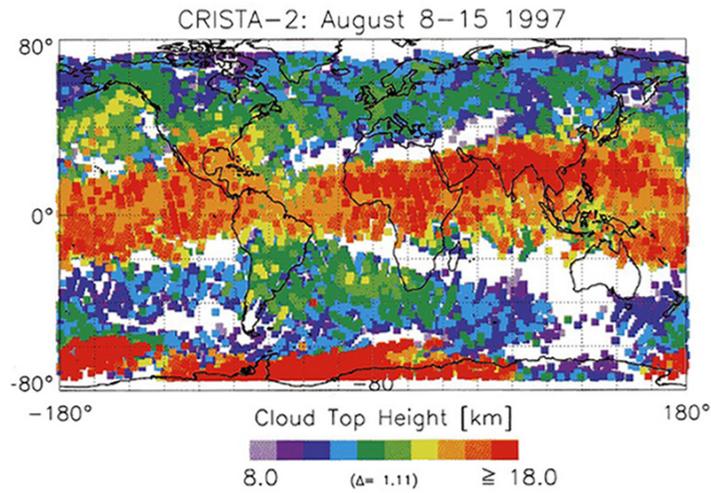
Figure 2.3: CRISTA-2 global map of the cloud top heights detected by the cloud identification scheme. Latitude range is −80 to +80 °N and time interval is 8–15 August 1997. As given in figure 3 of Spang et al. (2002).

describing subvisual cirrus in addition to thicker cirrus. It was noted that scattering of upwelling thermal radiation into the CLAES line-of-sight could significantly enhance the CLAES ability to detect thin cirrus. It was estimated that the radiation emanating toward CLAES from an ice cloud could be enhanced by 25 % or more by scattering (and the size of the enhancement can increase with particle size).

The ability of CLAES to detect subvisual cirrus was exploited in a comparison to coincident relative humidity measurements from the Upper Atmospheric Research Satellite (UARS) Microwave Limb Sounder (MLS) instrument in a study by Sandor et al. (2000). It was shown that there was a strong correlation between high frequency of thin and subvisual cirrus around the tropopause and high relative humidity with respect to ice.

The development of high fidelity space pointing limb sounders such as the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) and the High Resolution Dynamic Limb Sounder (HIRDLS) presents an opportunity to further investigate the potential of the limb thermal emission technique to yield more detailed information on cirrus than is possible from occultation alone, and with greater global coverage than previous emission instruments. Observations of optically thin cloud made by both MIPAS and HIRDLS will measure radiation emitted from the cloud and also a radiance component which is transmitted from below the cloud due to multiple scattering from cloud particles. Effectively, tropospheric radiation can be scattered into the line-of-sight and, in the case of the MIPAS measurements, cause additional absorption features in the spectra. Early studies of MIPAS data have demonstrated this for both PSCs (Höpfner et al., 2002) and for cirrus (Spang et al., 2004).

Höpfner et al. (2002) were able to demonstrate that the PSC signal in the MIPAS spectra

could be fitted using a forward code to calculate single scattering into the line-of-sight, assuming Mie scattering and that the PSCs were composed of spherical sulphuric ternary solution (STS) droplets in a lognormal size distribution. An approximate fit to the MIPAS spectra was obtained, through which the mode radius could be implied, without using any fit cost minimisation. For particles > 1 μm the scattered radiance accounts for a significant part of the continuum signal and leads to absorption line features. The subsequent modelling study by Höpfner (2004) showed that the scattered radiance has a strong dependence on particle size. This is strong evidence to suggest that if the dependence of scattering on particle size can be fully characterised it may be possible to retrieve the microphysical properties of PSCs from the MIPAS spectra.

Milz et al. (2004) have been able to model signatures of thin cirrus clouds inclusive of single scattering similar to, but not fitting, MIPAS measurements of thin cirrus. Likewise, Spang et al. (2004) were able to simulate the absorption line features observed in the MIPAS cirrus observations by a simple approach of adding an arbitrary percentage of a nadir calculation to the limb case. As scattering only becomes efficient when the particle sizes are equal to or greater than the wavelength of the incident radiation, showing that absorption lines in cirrus observations are due to scattering (as in Milz et al., 2004; Spang et al., 2004) implies that the detection of these lines gives indirect information on the size distribution of the observed cloud. However, neither Spang et al. (2004) nor Milz et al. (2004) were able to characterise any relationship between the scattering effect observed in the MIPAS measurements and the microphysical properties of cirrus.

2.5 The MIPAS Instrument

The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) was launched in March 2002, as part of the core payload of the European Space Agency Environmental Satellite (ENVISAT) polar-orbiter. Figure 2.4 shows a schematic of the ENVISAT satellite configuration including the positions of the payload instruments.

MIPAS is a Fourier transform spectrometer (FTS) designed for the monitoring of trace gas species by measuring high-resolution gaseous emission spectra at the Earth's limb. It is designed to operate in both the day and night parts of ENIVSAT's orbit, measuring in the near to mid infrared, ranging from 685 cm^{-1} to 2410 cm^{-1} (4.15 to 14.6 μm), where trace gases such as O_3, H_2O, HNO_3, CH_4 and N_2O have important emission features.

An FTS is an adaptation of the Michelson interferometer. The standard Michelson interferometer consists of a light source, a beamsplitter, two mirrors and a detector configured as shown in figure 2.5. In the inteferometer, a collimated beam from the light source is divided into two by the beamsplitter (usually a half-silvered mirror inclined at 45° to the
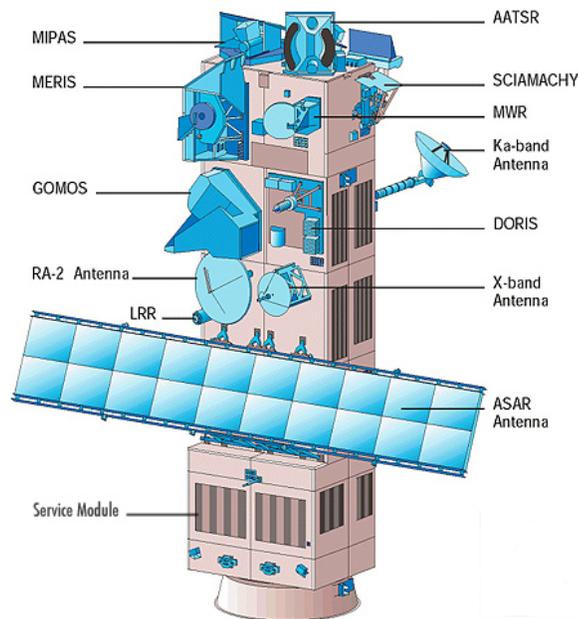
Figure 2.4: The ENVISAT satellite configuration. The position of each of the instruments is shown clearly.

incoming beam) and sent to the two mirrors. These mirrors reflect the beams back along the same paths to the beamsplitter, where they interfere. The output signal depends on the wavelength of the light and the optical path difference between the beamsplitter and each of the two mirrors. If the optical path difference between the two beams is zero or a multiple of the wavelength of the light then the output will be bright, but if the optical path difference is an odd multiple of half the wavelength of the light then the output will be dark. If the two mirrors are equidistant from the beamsplitter then a monochromatic light source will produce a 2D pattern on the output detector screen consisting of a bright spot surrounded by a series of alternating dark and bright rings (as shown in figure 2.5) similar to an Airy disk pattern.

In the FTS one of the mirrors is movable (or both are movable), thus allowing the path difference between the two arms of the interferometer to be altered. The movable mirror is scanned in the direction parallel to the light beam causing the output to alternate between bright and dark fringes. If the light source is monochromatic, then the signal recorded at the output will be modulated by a cosine wave. However, if the source is not monochromatic then the output signal will be the Fourier transform of the spectrum of the input signal, i.e. the interferogram.

Figure 2.6 is a schematic of MIPAS. As can be seen from the schematic, MIPAS is a complex instrument comprised of several sub-systems, of which the most technically innovative is the interferometer.

Radiation enters the instrument through the aperture in the front-end optics (at the bottom of figure 2.6). The front-end optics has two scanning mirrors which are used to select the
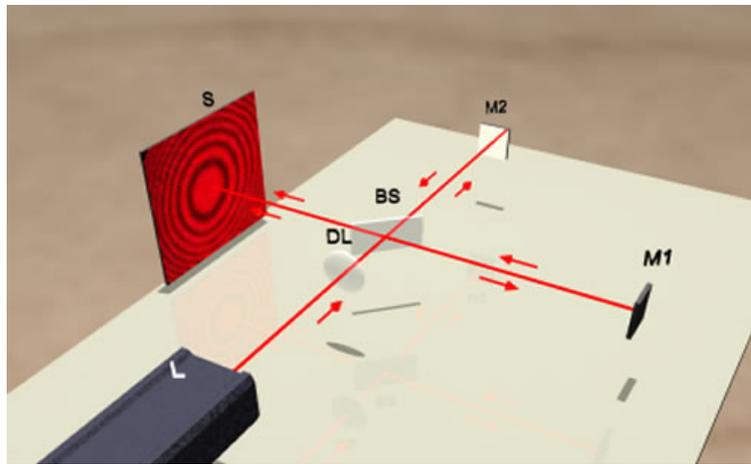
Figure 2.5: 3D view of the optical layout of a typical Michelson interferometer (Michael, 2005).
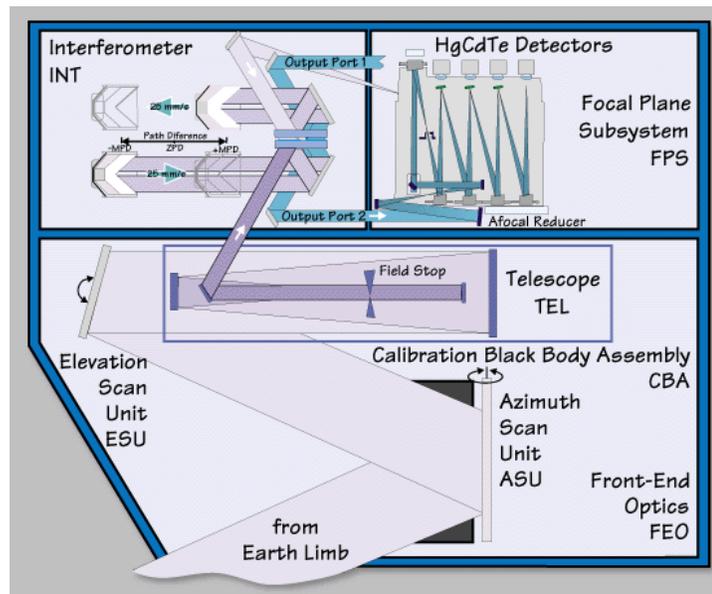
Figure 2.6: Schematic view of MIPAS optical layout (ESA, 2005).

target line-of-sight. The elevation scan mirror is used to point at various altitudes and is capable of correcting for variations in the orbital altitude and the Earth's geoid geometry. The azimuth scan mirror is used to point at different angles to the side or rear of ENIVSAT and can also be used to point at the internal calibration blackbody. The input radiation is directed into the telescope and collimator. The telescope has an internal field stop which is used to select the instrument field-of-view (FOV). Upon leaving the telescope the radiation is collimated and redirected into the interferometer. Inside the interferometer the two retro-reflector mirrors are moving over a 100 mm path at a constant velocity of 25 mm s^{-1}. The output signal is split into two beams, each directed to an output port. The output beams are then directed to a series of eight detectors, cooled by a pair of synchronized Stirling cycle coolers. To record a useful interferogram, the modulated output has to be sampled at very regular optical path difference intervals (the required sampling accuracy for MIPAS is
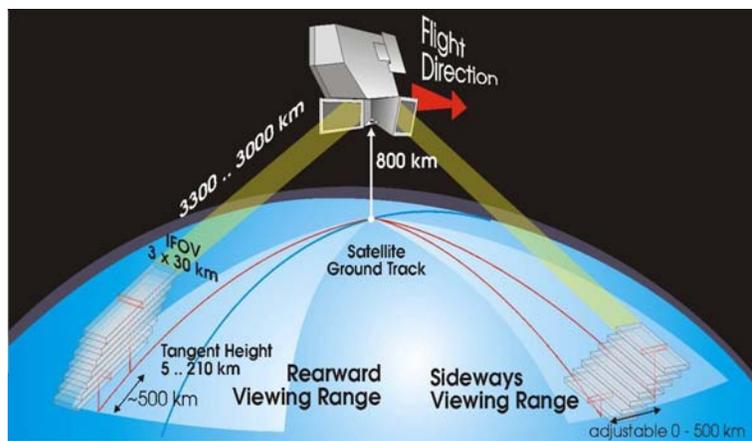
Figure 2.7: MIPAS viewing geometry.

about 30 nm). The detectors record the interferograms for later post-processing to obtain the spectral distribution of the input radiation from the limb.

As mentioned previously, the MIPAS front-end optics allows data to be collected from various altitudes and various positions in the atmosphere. The MIPAS viewing geometry is illustrated in figure 2.7. The nominal scan mode is in the anti-flight direction but MIPAS is also capable of pointing perpendicularly to the flight track. In nominal mode the vertical range is approximately 6–68 km, with 17 elevation scans making up an elevation profile. Vertical spacing between scans is 3 km for 6–42 km, and then increases to 6 km spacing between 42–60 km and 8 km between 60 km and the highest scan (figure 2.8). It takes about 4.5 seconds to scan the complete spectrum at a given elevation measurement, so a complete profile measurement takes approximately 80 seconds, giving over 1000 profiles per day. The horizontal measurement range is nominally 330 km between the first (highest) and last (lowest) scan tangent points, and the distance along-track between two measurements at the same altitude is approximately 500 km. At the tangent point, the Instantaneous FOV of MIPAS is about 3 km in elevation by 30 km in azimuth.

The ENVISAT spacecraft is in a Sun-synchronous polar orbit with a mean operating altitude of 800 km. The nominal orbit has a repeat cycle of 35 days and an orbital period of 100.6 minutes, with an inclination of 98.54 degrees. Due to the inclination of the orbit, to obtain full global coverage pole to pole when MIPAS is in the nominal viewing mode it is allowed to operate with a rearward oscillating scanning geometry in the azimuth within a 35° wide range. Therefore, the tangent points of the elevation scan measurements do not exactly overpass the ground footprint of the orbit path. MIPAS is also capable of scanning within a 30° wide azimuth area in the anti-sun direction. However, the sideways viewing range is only used for specialised observing modes. See ESA (2000) for further details on observing modes.

The high-resolution spectra output from MIPAS non-continuously covers the range of ther-
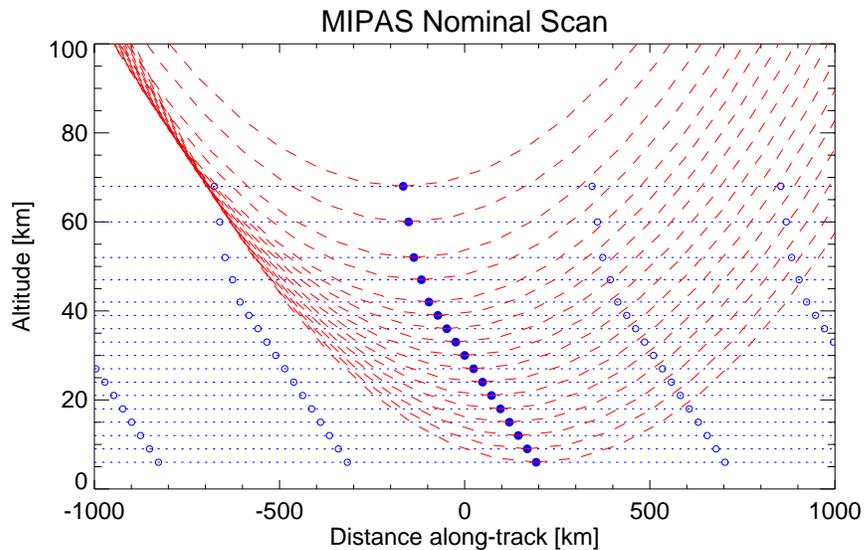
Figure 2.8: The tangent paths (dashed lines) and tangent points (solid dots) for nominal MIPAS limb scans with the satellite to the right and moving right. Open dots show tangent points for successive profiles. The figure is courtesy of Dr Anu Dudhia (University of Oxford) *via* private communication.

Band	Wavenumber range (cm^{-1})
A	$685 - 970$
AB	$1020 - 1170$
B	$1215 - 1500$
C	$1570 - 1750$
D	$1820 - 2410$

Table 2.1: The spectral ranges of the MIPAS Bands.

mal infrared from 685 to 2410 cm^{-1} by measuring over five spectral bands (see table 2.1). The unapodised MIPAS spectra have a spectral resolution of 0.025 cm^{-1}, achieved by a maximum optical path difference of the FTS of 200 mm. Apodised spectra are produced by applying a Norton-Beer 'strong' apodisation (Norton and Beer, 1976) in optical path difference space (i.e. to the inteferogram). The spectral bin width of the apodised spectra is also 0.025 cm^{-1}, but the narrowest resolvable line width is 0.035 cm^{-1}.

2.6   Outline of Methodology

Design of retrieval methods is an essential process in the design of a remote sounding instrument. From a thorough retrieval study, the instrument design can be optimised by identifying additional parameters to measure and by specifying the spectral region and spectral accuracy to make measurements which will yield the most information. In the case of MIPAS, the instrument was designed for the retrieval of vertical profiles of temperature, pressure and trace gases. The retrieval of cloud properties was not a consideration at

the instrument design stage. Therefore, it is likely that MIPAS may not be the optimal instrument for cirrus retrievals, but it is the objective of this thesis to design a retrieval method that can be used operationally to the best of the instrument's capability and be validated to show that cirrus properties can be retrieved.

2.6.1 The Inverse Problem of Atmospheric Sounding

In his 1964 book, "Use and Abuse of Statistics", W.J. Reichmann writes:

> *There are many quantities which we would like to measure but which, for a variety of reasons are actually immeasurable. Where so much is uncertain because we cannot undertake the actual measuring, it is necessary to devise some means whereby we can measure the probability that inferences about immeasurable quantities are correct.*

Although talking about making inferences about real quantities through statistical sampling, Reichmann's point transfers directly to the field of remote sounding of the atmosphere, capturing the essence of the inverse problem. Generally, when making a remote measurement of the atmosphere it is often a complicated function of the parameter that is actually required. The inverse problem is the question of finding the best representation of the required parameter given the measurements made, together with any appropriate prior information. In terms of Reichmann's statement, a retrieval scheme is the means *devised* to implement the inversion of the complicated function in order to obtain (or *retrieve*) the required parameter from the indirect measurement. Associated with the inverse method is the relationship between the true state of the system and that retrieved, which can conveniently be expressed in terms of probability (see section 6.2).

The inverse problem can be thought of conceptually as a question of constructing and subsequently solving a set of simultaneous linear or non-linear equations. Complications arise in the presence of experimental error in some of the measurements and/or in the presence of approximations in the formulation of the equations. The inverse problem can be highly complex in atmospheric sounding, hence the advent of the subject of *retrieval theory* to provide methods to solve inverse problems. Chapter 6 presents the details of retrieval theory as applied to the work undertaken in this thesis. However, it is necessary to introduce some of the core concepts (following Rodgers, 2000) at this stage as they will be referred to frequently.

2.6.2 Precursor to Retrieval Theory

In order to make a formal statement of the inverse problem, first consider the measurement of m quantities (e.g. radiances) assembled into a vector $\mathbf{y}$, the *measurement vector*, and
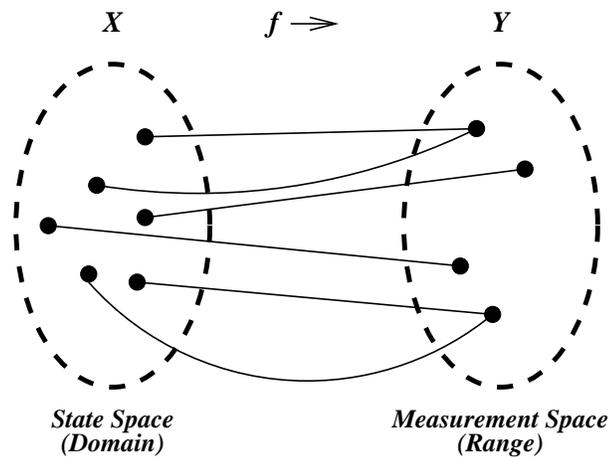
Figure 2.9: The act of measurement is equivalent to a mapping, where the domain is the state space and the range is the measurement space. The measurement is often a many-to-one mapping, as shown.

the n parameters describing the state (i.e. the atmosphere in the case of remote sounding) assembled into a vector $\mathbf{x}$, the *state vector*. The quantities to be retrieved are represented by the state vector.

When considering systems of simultaneous equations, it is useful to introduce the concept of *vector spaces*[1]. For inverse problems of remote sounding it is useful to define two vector spaces:

- The *state space* with dimension n, within which each conceivable state is represented by a point, or equivalently by a vector from the origin to the point.

- The *measurement space* with dimension m, within which each conceivable measurement is represented by a point or vector.

The act of measurement is equivalent to a mapping from the state space into the measurement space. The measurement mapping can be denoted by $\mathbf{f} : \mathbf{X} \longrightarrow \mathbf{Y}$, where $\mathbf{X}$ is the state space $(\mathbf{X} \subseteq \mathbb{R}^n)$, $\mathbf{Y}$ is the measurement space $(\mathbf{Y} \subseteq \mathbb{R}^m)$, and $\mathbf{f}$ can be thought of as a function such that for every $\mathbf{x} \in \mathbf{X}$, there is an object $\mathbf{f}(\mathbf{x}) \in \mathbf{Y}$. The measurement can be represented conceptually by a simple mapping diagram as in figure 2.9. In the absence of measurement error, the measurement is a many-to-one (or sometimes a one-to-one) relation (as shown in figure 2.9).

The inverse problem is therefore that of finding an appropriate inverse mapping from the measurement space back into the state space or finding a numerical method to approximate the inverse mapping. This is non-trivial as there is usually measurement error and the measurement mapping is mostly many-to-one and is often a non-linear relation.

---

[1] Roughly speaking, a vector space is a system of vectors defined over the elements of an arbitrary field of scalars which obey the set of ten vector space axioms. For further details on vector spaces and fields see, for example, Whitelaw (1996).

For each state vector there is a corresponding ideal measurement vector, $\mathbf{y_I}$, determined by the physics of the measurement. The physics are described formally by the *forward function* $\mathbf{f(x)}$:

$$\mathbf{y_I} = \mathbf{f(x)}. \tag{2.1}$$

In practice, there is always an experimental (measurement) error ϵ so, more generally,

$$\mathbf{y} = \mathbf{f(x)} + \epsilon. \tag{2.2}$$

The measurement errror ϵ is the difference between the real and perfect measurements. If the physics of the forward function are too complex to evaluate explicitly (e.g. in the case of a multiple scattering cloudy atmosphere), it is necessary to approximate the forward function to adequate accuracy. This is known as the *Forward Model* and is used to calculate synthesised radiances, $\mathbf{\hat{y}}$, and can be represented by

$$\mathbf{\hat{y}} = \mathbf{F(x)}. \tag{2.3}$$

This represents an estimate to the forward function, i.e. $\mathbf{F(x)} \simeq \mathbf{f(x)}$, and results in a forward model error. Assuming a perfect (or near perfect) forward model

$$\mathbf{y} = \mathbf{F(x)} + \epsilon. \tag{2.4}$$

To relate the retrieved state to the true state, we operate on the measurement with some *Inverse* or *Retrieval Method* $\mathbf{R}$

$$\mathbf{\hat{x}} = \mathbf{R(y, x_a, c)} \tag{2.5}$$

where $\mathbf{\hat{x}}$ is known as the retrieval. The vectors $\mathbf{x_a}$ and $\mathbf{c}$ together comprise parameters that do not appear in the forward function, but do affect the retrieval, and may be subject to an uncertainty; $\mathbf{x_a}$ is a vector of prior knowledge (*a priori* data) corresponding to the state vector and $\mathbf{c}$ contains any other parameters not included in the forward function.

The *retrieval method* involves some form of optimal estimation (described in chapter 6) in order to obtain a solution $\mathbf{\hat{x}}$ (as close to $\mathbf{x}$ as possible) by minimising the errors and the difference between the measurement and the synthesised measurement, $\mathbf{y} - \mathbf{F(\hat{x})}$ (which will also be weighted by the measurement error).

2.6.3   Design of Retrieval Methods: Thesis Outline

According to Rodgers (2000), the process of design and optimisation of an observing system should begin by specifying the measurement requirements, and selecting the style of instrument and the observing geometry, and then include:

1. Designing a forward model for the instrument, algebraically

2. Identifying the forward model parameters and state vector elements

3. Building a numerical forward model and the corresponding derivative model

4. Building a numerical retrieval method

5. Instrumenting the retrieval method with diagnostics

6. Selecting *a priori* information both atmospheric and instrumental

7. Determining whether the proposed measurements are adequate to determine the proposed target parameters

8. Optimising the retrieval characteristics with respect to the instrument design.

Essentially, (with the exception of step 8) this outlines the basic procedure required to carry out a retrieval study. Therefore, for a pre-existing instrument, such as MIPAS, to develop a retrieval of a particular set of target parameters which were not considered at the instrument design stage, steps 1 to 7 in the above process should be followed, with the caveat that step 7 will be:

- Establishing whether the measurements are adequate to determine the required target parameters.

However, before it can be determined whether or not cirrus properties can be retrieved from MIPAS, it must be established that the MIPAS measurements are sensitive to the presence of cirrus and that cloud can be detected accurately. This preliminary study is covered in chapter 3.

The main aim of this thesis is to assess the feasibility of retrieving cirrus properties from MIPAS measurements, thus the structure of the thesis will loosely follow the process outlined above:

- Steps 1 to 3 will be covered in chapter 4 (with the derivative model further discussed in chapter 7)

- The theory required to implement steps 4 and 5 is covered in chapter 6

- The discussion of the implementation of steps 4 and 5 is given in chapter 7

- Steps 6 and 7 are also covered in chapter 7

- It is implicit in covering step 7 that a study of the sensitivity of the measurements to the state vector elements (with respect to the forward model) should be carried out, which is essentially done in the channel selection section in chapter 7.

In addition, depending on the response of the forward model to changes in the state vector elements, the retrieval may not always return the correct solution as it is a numerical method. The focus of chapter 5 is to present the suboptimal and *ad-hoc* methods which were developed to "guide" the retrieval. Essentially, chapter 5 can be thought of as part of steps 4 and 7.

As MIPAS is already producing real data from the atmosphere, the results from applying the retrieval method designed are compared to results obtained from a different instrument[2] in chapter 8, by way of validation. As it is not possible to actually determine whether the retrieval is consistent with the state, because the true state is unknown, comparison with independent measurements is the best way to confirm that the retrieved quantities are as accurate as possible (within the limits of the associated errors) and that the relationship to the true state is well understood.

Finally, the conclusions on the feasibility of retrieving cirrus properties from MIPAS and whether this can improve current current cirrus climatology data are given in chapter 9.

---

[2]The Advanced Along Track Scanning Radiometer (AATSR) in this case.

Chapter 3

MIPAS Cloud Detection

3.1 Introduction

In the application of any retrieval scheme, it is important to ensure that the retrieval is being applied to an appropriate set of measurements. Therefore, this chapter is devoted to the detection of cloud in limb measurements. The process of cloud detection is equally as important for operational retrievals of atmospheric gases from MIPAS, in order to avoid attempting to retrieve trace gas information from cloud contaminated measurements.

This chapter introduces two existing methods for cloud detection, then presents a brief validation of these methods as applied to MIPAS data, compared against each other and against external datasets. In addition, this chapter also covers the developments of a new supplementary test to enhance the detection of optically thin cirrus. This new detection test is empirically based and has been derived from a Principal Components Analysis (PCA) applied to a large sample set of MIPAS limb emission spectra.

By establishing the measures used for cloud detection at the outset, the achievable detection sensitivity of MIPAS, with respect to microphysical and optical cloud properties, can be quantified upon the development of the forward model in chapter 4. Furthermore, modelling the cloud detection measures is also useful for the study in chapter 5 to estimate cloud properties without carrying out a full retrieval.

3.2 Cloud Detection Methods

Cirrus clouds affect limb sounding measurements through enhanced absorption and emission. In addition, there is a significant enhancement of the measurement due to scattering of radiation from altitudes below the tangent height into the line of sight by cirrus particles (see chapter 4). Operational MIPAS trace gas retrievals rely on a simple radiance ratio for the detection of clouds in the upper troposphere (Spang et al., 2004). Alternatively, a

CI MIPAS band	MW1 (cm^{-1})	MW2 (cm^{-1})	CI threshold value	CRISTA defined altitude range(km)
CI-A	788.20–796.25	832.3–834.4	1.8	8–60
CI-B	1246.3–1249.1	1232.3–1234.4	1.2	8–50
CI-D	1929.0–1935.0	1973.0–1983.0	1.8	8–32

Table 3.1: Cloud indices for MIPAS (Spang et al., 2004). The height range is extendable to lower altitudes (~ 6 km) for mid and high latitudes.

simple radiance threshold method, can be used to detect enhanced radiances in the limb measurements due to cirrus. This section briefly describes and evaluates each of these cloud detection methods.

3.2.1 Colour Index for Cirrus Detection

The Spang et al. (2004) method for MIPAS cloud detection was originally developed for the CRISTA instrument (Spang et al., 2002), and is based on a simple approach of using the ratio of mean radiances from two different wavelength regions labelled MW1 and MW2. The ratio can be used to detect clouds as the two regions have different sensitivites to cloud in the FOV. The ratio MW1/MW2 is known as the Colour (or Cloud) Index (CI), and there are 3 possible CIs for MIPAS, situated in detector bands A, B and D. The MIPAS CIs are listed in table 3.1, with CI-A being the standard CI developed for CRISTA. However, as the cirrus retrieval (see chapter 7) focuses on limb spectra measurements in the atmospheric window region ($\sim$750–1250 cm^{-1}), only CI-A is of interest here.

For CI-A, the first wavelength region 788–796 cm^{-1} is dominated by CO_2 and weaker ozone emissions, it is therefore relatively insensitive to the presence of optically thin cloud (see figure 3.1). However, the second wavelength region 832–834 cm^{-1} is highly sensitive to the presence of cloud as the only other radiance contribution in this region is from weak emissions by ozone and CFC-11. Therefore, the ratio MW1/MW2 can be expected to be large in cloud free cases and close to unity for optically thick cloud.

The variation in the ratio can be described physically by considering the expression for radiance arriving at a limb sounder along a ray path (ignoring scattering). The radiance from a limb path is dominated by the region of the atmosphere in the kilometre or two above the tangent height (see figures 2.1 and 2.8). This region corresponds to a section of the ray path about the tangent point (with length dependent on tangent height), and the radiance arriving at the limb sounder can be approximated by the radiance along this section of the path. The atmospheric gases along the section can be considered at constant temperature and at constant concentration. Therefore, if the transmittance at the start of the section is considered to be 1, and the transmittance at the end of the section considered
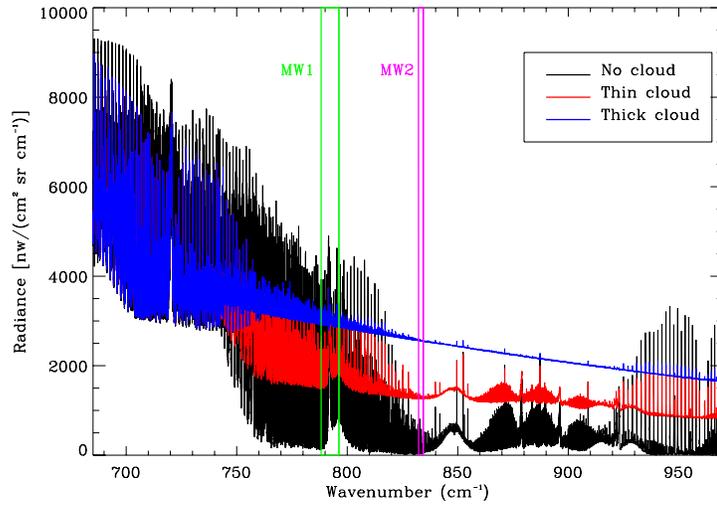
Figure 3.1: Modelled MIPAS band A limb radiance for cloud free, thin cloud (β_{ext} =0.002 km^{-1}, C_{depth} = 1 km, C_{top} = 13 km) and thick cloud (β_{ext} = 0.008 km^{-1}, C_{depth} = 1 km, C_{top} = 13 km) cases, at a 12 km viewing tangent height for a climatology equatorial atmosphere. The microwindow regions used for the Colour Index (CI-A) cloud detection are highlighted.

to be $\mathcal{T}$, then the monochromatic radiance along the path section is given by

$$R = B \int_{\mathcal{T}}^1 d\mathcal{T} = B(1 - \mathcal{T}) \tag{3.1}$$

where B is the planck function of the emitting atmosphere (for a constant temperature), and the transmittance $\mathcal{T} = e^{-\tau}$, where τ is the optical depth of the section of the ray path. Neglecting scattering, the optical path length is defined as

$$\tau(s_1, s_2) = \int_{s_1}^{s_2} \kappa(\bar{\nu}, s) \, ds \tag{3.2}$$

for the path between s_1 and s_2, where $\kappa(\bar{\nu}, s)$ is the absorption coefficent along the section of the path with respect to radiation of wavenumber $\bar{\nu}$. If cloud is present, the total absorption coefficient at point s along the path section can be given by the additive combination of the cloud absorption coefficient $\kappa_{cld}(\bar{\nu}, s)$ and the gas volume absorption coefficient $\kappa_{gas}(\bar{\nu}, s)$, i.e. $\kappa(\bar{\nu}, s) = \kappa_{gas}(\bar{\nu}, s) + \kappa_{cld}(\bar{\nu}, s)$. Therefore, equation 3.2 gives the total optical depth of the path as

$$\begin{aligned}
\tau &= \int \kappa_{gas}(\bar{\nu}, s) \, ds + \int \kappa_{cld}(\bar{\nu}, s) \, ds \tag{3.3} \\
&= \tau_{gas} + \tau_{cld} \tag{3.4}
\end{aligned}$$

where τ_{gas} is the optical path length of the atmospheric gases over the section and τ_{cld} is the optical path length of the cloud over the section.

Then equation 3.1, gives the mean radiance for MW1 as

$$R_1 = B_1(1 - e^{-(\tau_{gas,1}+\tau_{cld,1})}) \tag{3.5}$$

where B_1 is the mean Planck function value for MW1, and the optical paths for atmospheric gases $\tau_{gas,1}$ and for cloud $\tau_{cld,1}$ are assumed constant (with wavenumber) across MW1. Similarly for the mean radiance R_2 of MW2.

It can be assumed that $\tau_{gas,1}$ and $\tau_{gas,2}$ are reasonably well known, i.e. from the clear sky limb transmission calculation. It can also be assumed that $\tau_{cld,1} \approx \tau_{cld,2}$ as the refractive index of ice (which determines the cloud volume absorption coefficient) does not change significantly between the two microwindows. Therefore, in a non-scattering environment the ratio of the two microwindows can be expressed as:

$$\text{CI-A} = \frac{B_1}{B_2}\frac{(1 - e^{-\tau_{gas,1}}e^{-\tau_{cld}})}{(1 - e^{-\tau_{gas,2}}e^{-\tau_{cld}})}. \tag{3.6}$$

Therefore if optically thick cloud is present $e^{-\tau_{cld}} \longrightarrow 0$ and CI-A $\longrightarrow$ B1/B2, which is ≈ 1 for typical cirrus temperatures. If no cloud is present then $e^{-\tau_{cld}} = 1$ and the ratio is dominated by the difference between $\tau_{gas,1}$ and $\tau_{gas,2}$. Since $e^{-\tau_{gas,1}} < e^{-\tau_{gas,2}}$ then CI-A is large, i.e. CI-A > 4. When optically thin cloud is present CI-A<4. The relationship between CI-A and cloud optical depth is only weakly dependent on temperature, i.e. $< 1\%/\text{K}$ between 10 and 30 km (Spang et al., 2004). The detection sensitivity is governed by the threshold value used for CI-A. Spang et al. (2004) analysed measured spectra in conjunction with forward calculation using the RFM (Dudhia, 2004) to assign a threshold value. It was concluded that CI-values between 2 and 4 are related to optically very thin clouds or a situation where a cloud only partly fills the FOV. Therefore, the threshold value was set to flag cloudy measurements for CI-A values of less than or equal to 1.8 (table 3.1).

3.2.2 Window Radiance Threshold for Cirrus Detection

The use of radiance threshold tests for cloud detection is well known for nadir measurements in the infrared. A typical gross cloud check in the nadir is to use an infrared theshold test, and if the measured brightness temperature is below a certain threshold temperature the pixel is rejected as cloud-contaminated. For example, the gross cloud check for the Advanced Very High Resolution Radiometer (AVHRR) (Saunders and Kriebel, 1988) uses the 12 μm brightness temperature channel because clouds have a greater optical depth at these wavelengths. When clouds of significant optical depth are present they act as a grey-body in the FOV and their effective emitting temperature is lower than the expected surface brightness temperature. The MIPAS spectra in the window region are used to
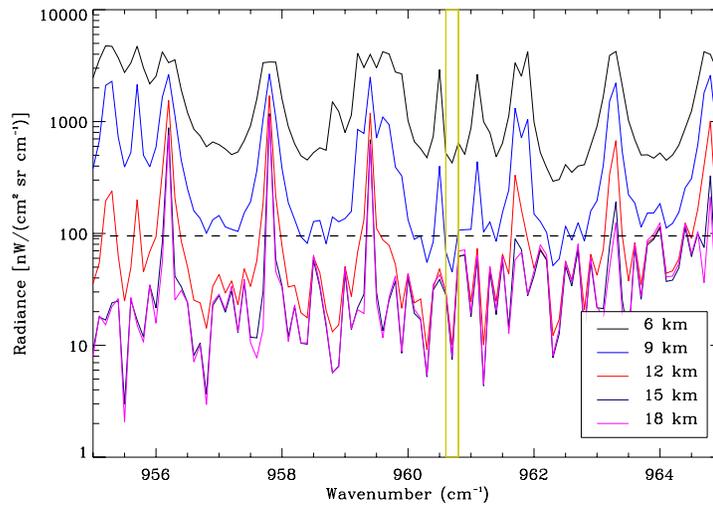
Figure 3.2: Modelled MIPAS spectra (inclusive of FOV effects) in the region 955–965 cm^{-1}, at various viewing tangent heights, for a typical equatorial atmosphere. The dashed line indicates the radiance value of the 9 km spectra at 960.7 cm^{-1} plus the Noise Equivalent Spectral Radiance. The yellow box highlights the 960.7 cm^{-1} region.

detect cloud effects by a similar principle. However, the expected clear-sky radiances will be low for the limb view out to cold space (as opposed to high for nadir views of the surface), and therefore clouds will be detected in limb view measurements when they are above a certain threshold radiance. The difficulty in both cases is defining an appropriate threshold value.

One of the most transparent spectral regions in the MIPAS A band is between 940 and 970 cm^{-1} and there is a narrow microwindow at approximately 960 cm^{-1} which has little influence from gas emission lines (see figure 3.1).The expected clear-sky MIPAS spectra (calculated by the RFM) in the 955–965 cm^{-1} region, for various tangent heights, assuming a typical equatorial atmosphere, is given in figure 3.2. It can be seen from figure 3.2 that the lower viewing tangent heights have a higher expected radiance due to observing warmer, denser parts of the atmosphere. For each of the modelled tangent heights, the spectra at 960.7 cm^{-1} is dominated by the background signature (and therefore aerosols and cloud when present) and the adjacent CO_2 emission lines are quite weak. Thus, 960.7 cm^{-1} is selected as the spectral location of the Window Radiance Threshold Test.

In order to determine a threshold value it is important to understand the sensitivity of the measurement to the presence of cloud. Without considering scattering effects at this stage (which, if significant, will enhance the radiance), the 960.7 cm^{-1} radiance that would be measured by MIPAS (inclusive of FOV effects) at various tangent heights (i.e. 6, 9, 12, 15 and 18 km) is modelled in the presence of cloud of various optical thickness. For each tangent height, it is assumed that there is a uniform cloud layer, with a given extinction

coefficient β_{ext}, which is 1 km thick with the cloud top 1 km above the viewing tangent height. The vertical optical thickness of the cloud is given by $\tau_{cld} = \triangle z . \beta_{\text{ext}}$, where $\triangle z$ is the vertical depth of the cloud. Figure 3.3 shows the results of the modelling. At all tangent heights, the cloud begins to have an impact on the limb radiance for optical thickness greater than 10^{-5}. However, the threshold radiance only begins to surpass values above the MIPAS noise level[1] over the clear-sky radiance at extinctions $\sim 10^{-4}$ km^{-1}. For tangent heights of 9 km or greater, an appropriate threshold value can be obtained by using the modelled radiance for an optical thickness of 10^{-4} at 9 km (shown on figure 3.3). This value will vary with latitude, and figure 3.3 only shows modelling for equatorial atmospheres. A global value of 125 nW/(cm^2 sr cm^{-1}) has proven to be fairly robust for tangent heights > 6 km as it is high enough to avoid detecting spectral noise fluctuations. Scattering at such values of optical thickness will enhance the radiance, and therefore, setting the threshold at this value allows detection of clouds with lower optical thickness in the real data. A more thorough investigation into the sensitivity of this threshold value, inclusive of scattering and cloud variability, with respect to cloud properties is undertaken in section 5.2. For clouds of tangent heights less than 9 km, the threshold can be set at a higher value, i.e. 300 nW/(cm^2 sr cm^{-1}). The radiance at the lower tangent heights will vary considerably and therefore, the high value should account for this and ensure that positive identifications will have cloud present in the field of view, which is desirable for the cloud retrieval. However, it will be less reliable from the point of view of ensuring that spectra which fail the test will be cloud free, as required for MIPAS gas retrievals.

The threshold values can be compared to real MIPAS data to obtain an idea of how appropriate they are. Figure 3.4 shows histograms of the radiance at 960.7 cm^{-1} for $\sim 30,000$ profile measurements for the month of August 2003. The statistics are for profiles at all latitudes (-90 to 90 °N) and sub-divided (by the Level 1B reported tangent height) into 3 km altitude bins from 4.5 km to 22.5 km centred on the nominal MIPAS limb scan tangent heights. Limb scans with tangent altitudes greater than this range are of no interest as cirrus will not form that high in the atmosphere. In fact tangent altitudes above ~ 20 km are only of interest in the polar regions where there is a possibility that PSCs may be present. On each of the histograms, the dashed line represents an approximate global average clear-sky frequency (i.e. free from both cloud and aerosol), such that the area under the histogram to the left of the dashed vertical line is equal to the clear-sky frequency value.

Figure 3.4(a) for the ~ 6 km tangent heights has a much higher spread of radiance values than for the higher limb scans, thus showing the high variability due to the higher temperature, pressure and atmospheric gas concentrations at lower height in the atmosphere.

---

[1]The MIPAS band A Noise Equivalent Spectral Radiance is ~ 50 nW/(cm2 sr cm^{-1}) - see section 7.2.5.
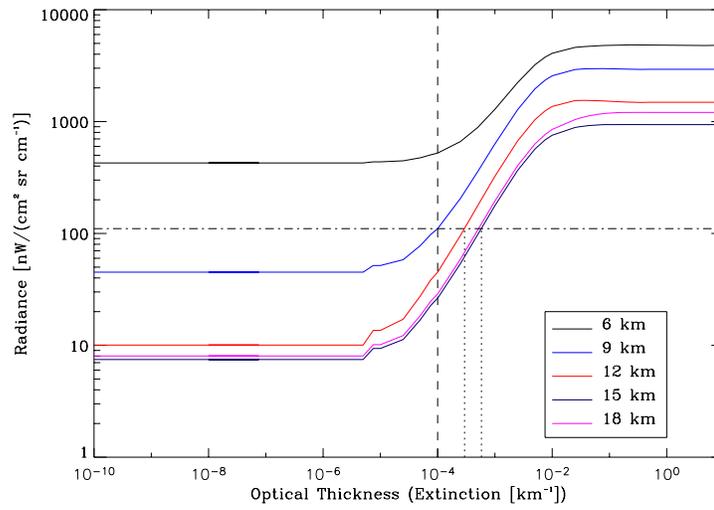
Figure 3.3: Modelled MIPAS radiances (inclusive of FOV effects) at 960.7 cm^{-1}, i.e. cloud detection threshold radiances, for clouds of various optical depths. The dashed line is $\tau = 10^{-4}$, the dot-dash line is the radiance at 9 km at $\tau = 10^{-4}$, and the two dotted lines are the optical thicknesses at the higher tangent heights where the radiance is equal to the dot-dash line.

Figures 3.4(b)–3.4(f) each have a peaked, Gaussian like distribution over the lower range of radiance values, which extends into a more random distribution over the higher values. In each of these figures the clear-sky frequency line is situated near the upper limit of the gaussian like distribution. The main peak of each of the histograms is highly correlated to the clear-sky profiles, but there is no clear cut-off between what appears to be the clear-sky values and what appear to be the cloudy values. From figure 3.3, it is obvious that this is due to the presence of aerosol and optically thin cloud observed in the MIPAS data (which will always be the case), which probably accounts for the limb scans contributing to the higher values of the bell curve.

It can be seen that the clear-sky frequency line is higher than the threshold value in both figures 3.4(a) and 3.4(b) for the $\sim$ 6 km and $\sim$ 9 km tangent heights respectively. Initially, this suggests that the threshold test is overly sensitive and will detect too much cloud. However, by considering the MIPAS viewing geometry (see figure 2.8) if there is high cirrus of large horizontal extent present then all tangent heights below the cloud top height will be affected by cloud in the FOV and will measure enhanced radiance. Effectively, with either the threshold or colour index cloud detection method, cloud will be flagged at these lower tangent heights without necessarily being present. This explains the higher than expected number of measurements at 6, 9, and 12 km with radiances above the detection threshold. In order to account for clear sky cases flagged as cloudy due to the presence of upper level cloud deck, a detection method dependent on the higher tangent height spectra would need to be introduced. However, this is a complicated relationship, as can be seen in section 5.2.

The clear-sky frequency line is lower than, but close to the $125\,\mathrm{nW/(cm^2\ sr\ cm^{-1})}$ threshold value in figures 3.4(c)–3.4(f). Although not conclusive, this implies that statistically the threshold test effectively flags cloud contaminated measurements at these tangent heights. This is further investigated by comparing the results from the threshold test to those from the colour index method in section 3.4.

3.3 A Meaningful Frequency of Cloud Occurrence Derived from Limb Measurements

In order to evaluate the merits of each of the cloud detection methods, the two methods are applied to a month of data and the results are investigated (in the following section), comparing against each other and against external datasets.

To evaluate cloud presence from similar methods and/or measurements over a period of time, the most obvious statistic to use is the actual occurrence of cloud at a given time and place. However, as two different sets of measurements are not necessarily co-located spatially and temporally, a direct comparison is not always possible. Instead, the number of times cloud is present in a given area/location as a percentage of the total number of measurements of that area/location over the measurement period is a useful statistic to compare between datasets. This is known as the *frequency of cloud occurrence.*

On first impression, it may appear that this is a rather straightforward quantity, but like any statistic, the meaning of the frequency of cloud occurrence is open to abuse if the variables measured are not clearly and precisely defined. One cannot measure anything unless one understands its limits, and it is important to ensure that anyone who is to utilise the derived measurement has a precise image of the quantity measured.

The first important point to make clear about the frequency of cloud occurrence derived from satellite limb measurements, is that it measures the percentage of measurements over a time period which have cloud present in them, not the percentage of the time period that cloud is present. The second point to note follows the argument from the previous section concerning the interpretation of the meaning of a positive cloud detection in a limb scan at a given altitude. It is often the case that cloud at higher altitudes can cause enhanced radiances at lower viewing tangent heights as the viewing ray path will pass through the cloud. This is a subtle point to understand and is best explained by considering an example. Consider the MIPAS elevation profile with a mean tangent point location at a latitude of $1.25\,°\mathrm{N}$ and a longitude of $36.5\,°\mathrm{N}$ from orbit 504 from the 16th of May 2002. The Window Threshold value and the Colour Index at each tangent height of the profile is given in figures 3.5(a) and 3.5(b) respectively. It can be seen that both methods detect cloud at all tangent
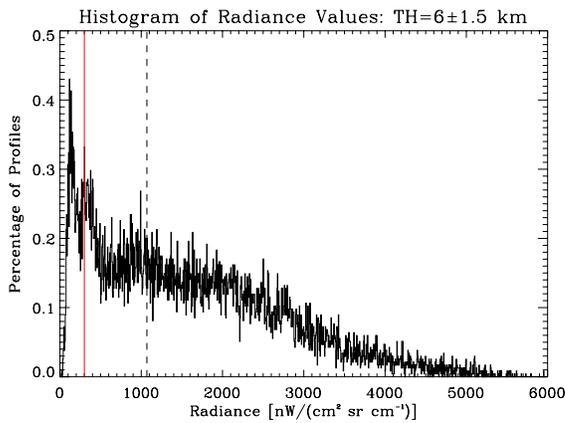
(a) 6±1.5 km. Clear-sky frequency ≈ 40%.
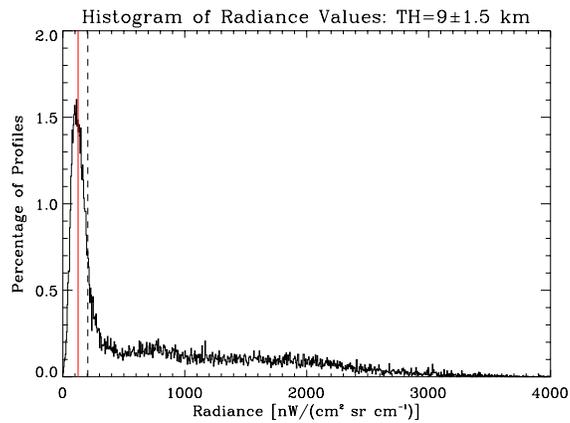
(b) 9±1.5 km. Clear-sky frequency ≈ 40%.
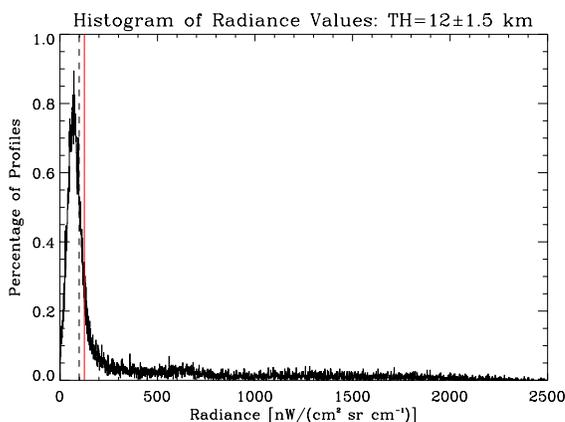
(c) 12±1.5 km. Clear-sky frequency ≈ 50%.
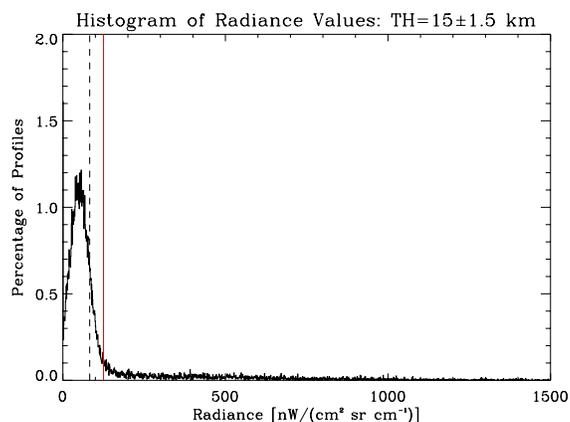
(d) 15±1.5 km. Clear-sky frequency ≈ 70%.
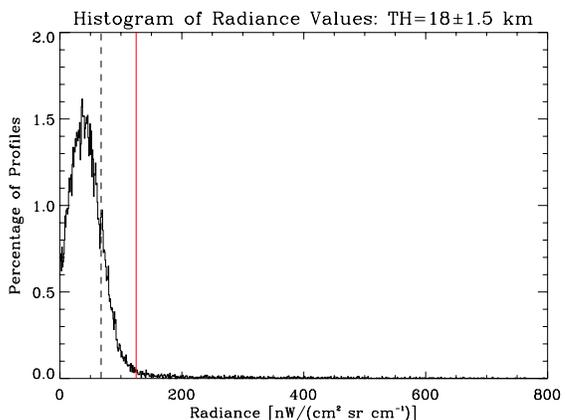
(e) 18±1.5 km. Clear-sky frequency ≈ 80%.
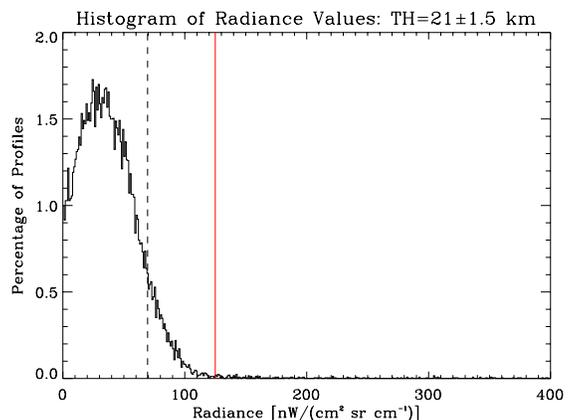
(f) 21±1.5 km. Clear-sky frequency ≈ 90%.

Figure 3.4: Frequency of MIPAS 960.7 cm^{-1} radiance values at the expected tangent height levels for all latitudes measured during August 2003. The overplotted dashed line indicates the clear-sky frequency line and the threshold value is overplotted in red.

heights below the highest tangent height that cloud is detected at, i.e. 19.3 km. In section 4.8 of chapter 4, the spectra from the 13.54 km limb scan of the same elevation profile is modelled to within good agreement (using McClouds_FM) with a cloud top at 18.75 km and a cloud depth of 0.7 km. Therefore, although cloud is detected at the lower tangent heights, the likelihood is that it is not actually present at these altitudes. With these simple cloud detection methods it is not possible to conclusively determine whether or not cloud is present at tangent heights below the first tangent height that cloud is detected in – MIPAS scans from the top of the atmosphere downward. Therefore, if a *cloud top event* is defined as the first limb scan that cloud is detected in, then only cloud top event limb scans can be considered to contain definite information on cloud presence at that altitude.
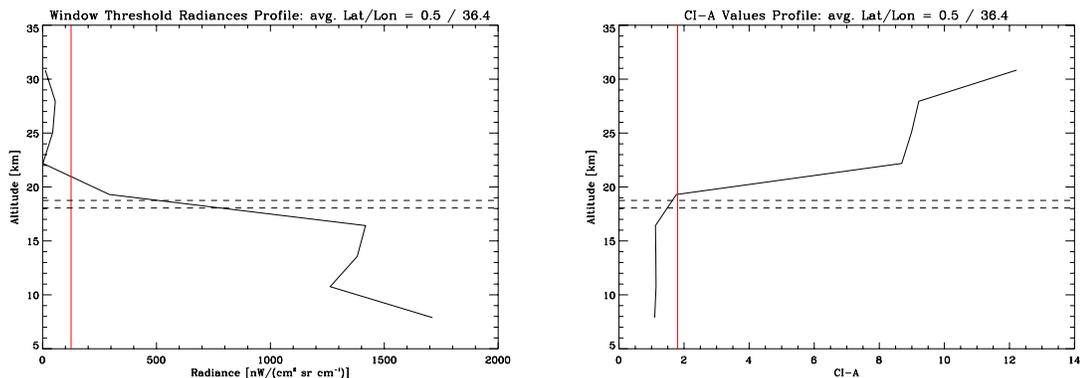
(a) Profile of radiance values at 960.7 cm^{-1}. The Window Threshold value is plotted in red.

(b) Profile of band A Colour Index values. The Colour Index threshold for cloud detection is plotted in red.

Figure 3.5: The Window Threshold radiance values (a) and the band A Colour Index values (b) for the same profile of MIPAS orbit 504 studied in section 4.8. In both cases the cloud top and bottom height used in section 4.8.1 to match the (McClouds_FM) modelled to the measured spectra for the 13.54 km tangent height is plotted as the dashed line.

The frequency of cloud occurrence f_c can be obtained by first calculating the number of measurements in which a valid cloud presence test can be performed. The number of MIPAS overpasses (i.e. limb measurements) in a defined latitude-longitude grid box at altitude $z(i)$ with $i = 1 \ldots n$ from the lowest level of the atmosphere in the analysis, is given by $N_{\text{raw}}(i)$ and the number of overpasses in the grid box at altitude $z(i)$ that pass the cloud detection test is given by $N_c(i)$. Assuming that all limb scans below a cloud top event are cloud contaminated (i.e. they pass the cloud detection test), then in a defined latitude-longitude grid box at altitude $z(i)$ the number of cloud top events $N_{\text{cte}}(i)$ can be obtained from

$$N_{\text{cte}}(i) = N_c(i) - N_c(i + 1), \tag{3.7}$$

and the number of valid measurements in which a cloud presence test can be performed (i.e. the number of overpasses which do not have a cloud top event at any altitude above it in the profile), $N(i)$ is given by

$$N(i) = N_{\text{raw}}(i) - N_c(i+1). \tag{3.8}$$

The frequency of cloud occurrence $f_c(i)$ in a defined latitude-longitude grid box at altitude $z(i)$ is then given by

$$f_c(i) = \frac{N_{\text{cte}}(i)}{N(i)} \times 100 \quad [\%]. \tag{3.9}$$

By defining the data as outlined above, the limits on the frequency of cloud occurence can be conveniently set, such that the maximum frequency of cloud occurence is the case in which all levels below cloud top events are assumed to have cloud present, i.e.

$$f_c^{\text{max}}(i) = \frac{N_c(i)}{N_{\text{raw}}(i)} \times 100 \quad [\%] \tag{3.10}$$

and the minimum frequency of cloud occurence is the case in which all levels below cloud top events are assumed to be cloud-free, i.e.

$$f_c^{\text{min}}(i) = \frac{N_{\text{cte}}(i)}{N_{\text{raw}}(i)} \times 100 \quad [\%]. \tag{3.11}$$

The frequency of cloud occurrence defined in equation 3.9 is similar to that used in Spang et al. (2002). However, it is more meaningful here with the addition of defined limits on the frequency due to the uncertainty of cloud presence in the levels below a cloud top event. The limits give an important indication of how the sample size that the statistic is calculated from decreases with tangent height.

Furthermore, it is also possible to define a probability density function of the altitude $z(i)$ of a cloud top event in a given latitude-longitude grid box, i.e.

$$P_{\text{cte}}(i) = \frac{N_{\text{cte}}(i)}{\sum_i^n N_{\text{cte}}(i)}. \tag{3.12}$$

Note that there is a subtle difference between P_{cte} and the probability density function of cloud top height, in that the former only gives the probability of the tangent altitude that a cloud top event occurs at, rather than the probability of the actual cloud top being at a particular height. Also, P_{cte} can not account for occurences of multiple levels of cloud decks.

The P_{cte} for the MIPAS August 2003 data (derived using the Colour Index test), binned for all longitudes in 5° latitude bins with altitude levels defined by the expected MIPAS tangent height levels, is given in figure 3.6. The behaviour of P_{cte} appears to be separated
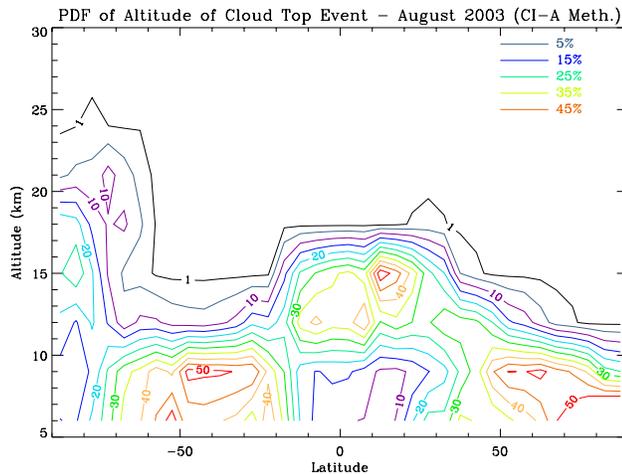
Figure 3.6: The probability density function of cloud top event occuring at nominal MIPAS tangent height altitudes for August 2003.

into 4 regimes: Southern polar latitudes, Southern mid-latitudes, tropical and equatorial latitudes, and for Northern latitudes above ~ 30 °N. There is a maximal probability of cloud above 15 km at Southern polar latitudes, clearly due to PSCs, and at Southern mid-latitudes between approximately 60 °S to 20 °S P_{cte} peaks for cloud top events at about 9 km. Of particular interest is the broad peak in probability in the tropical region for cloud top events to occur between 12 and 15 km, with a slightly sharper peak in cloud top event occurrence at 15 km between 10 °N and 20 °N. For latitudes higher than ~ 30 °N the probability of a cloud top event occurring generally increases lower in the atmosphere, with evidence of an enhanced probability at approximately 9 km from 45 °N to 60 °N. P_{cte} could potentially be used as an *a priori* cloud top height estimate under the assumption that cloud top event can be used as a proxy for cloud top height. A simple approach would be to find the expected value and standard deviation of the cloud top event in each latitude bin from P_{cte}, the expected value can be used as the *a priori* cloud top height estimate and then assuming a Gaussian distribution can be fitted for each bin with the same standard deviation will give an *a priori* uncertainty estimate.

The definition of the frequency of cloud occurrence can also be extended to limb occultation measurements. This is a necessary precursor for the comparison of MIPAS and limb occultation measurements in the following section. Section 2.4 discussed how occultation sounders (such as SAGE II and HALOE) can detect cirrus as enhanced extinction which does not obscure the solar beam. A cloud top event can be defined for occultation limb sounders as the first tangent height to detect such an enhanced extinction in a downward limb scan profile. However, as for MIPAS, occultation limb scans below the cloud top event do not contain definite information on cloud presence. This is because the lower limb scans extinction is enhanced by the higher level cloud, particularly if the cloud top event has a
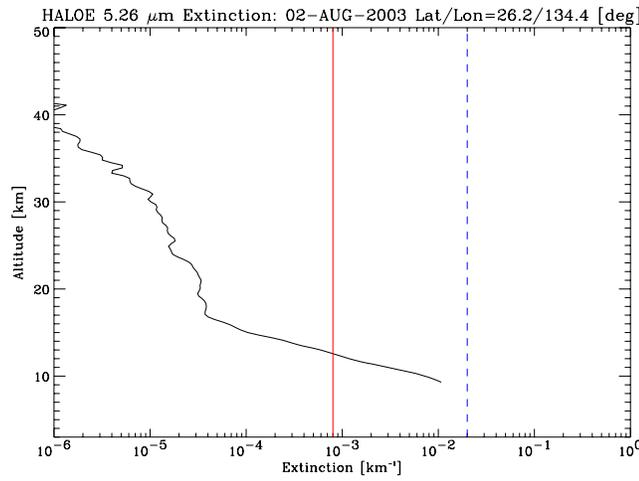
Figure 3.7: Profile of HALOE 5.26 μm extinction from 2nd August 2003 at approximately 26 °N and 134 °E. The red line indicates the cirrus extinction threshold (8×10^{-4} km^{-1}) and the blue line is the approximate upper limit of extinction (2×10^{-2} km^{-1}) before the limb transmission is too low for HALOE to record a measurement.

high extinction, i.e. an opaque cloud is present. Further, occultation measurements have additional uncertainty due to the fact that the extinction values are derived from measuring transmission. As mentioned in section 2.4, if the cloud extinction values are greater than $\sim 2 \times 10^{-2}$ km^{-1} the transmission will be too close to zero for the instrument to record a measurement. This is often symptomatic of limb scans below a cloud top event. This effect can be clearly seen in the HALOE 5.26 μm extinction profile given in figure 3.7. Theoretically, HALOE scans in the altitude range from 3 km to 90 km with a vertical stepping resolution of 0.3 km. In the profile in figure 3.7 it can be seen that the extinction below the 9.3 km limb scan is too high for HALOE to record a measurement. The cloud top event can be seen to occur at 12.3 km. For the purposes of determining cloud frequency, occultation limb scans which are below a cloud top event or do not record a measurement are considered to contain no definite information on cloud presence.

The formulae for the frequency of cloud occurence in equations 3.9 to 3.10 can be generalised for both limb and occultation limb measurements by considering a simple set theory explanation.

Let A be the finite set of all limb measurements in a particular height range for a particular instrument dataset, and let the subset of A which contains all cloud top event limb measurements be named T. Further, the subset of A which contains all limb measurements which are considered to contain no definite information on cloud presence (i.e. limb measurements which are below a cloud top event or which do not record a measurement) is named N and the subset of A which contains all clear-sky limb measurements is given by S. The set of all limb measurements is defined as the union of all of its disjoint subsets,

i.e. $A = T \cup N \cup S$. The number of elements in (i.e. the cardinal number of), for example, set A is denoted by $|A|$.

The frequency of cloud occurence f_c for any set of limb measurements in the specific height range is then given by

$$f_c = \frac{|T|}{|T \cup S|} \times 100 \quad [\%]\,, \tag{3.13}$$

the maximum frequency of cloud occurence $f_c^{\max}$ is obtained by assuming all limb measurements in N have cloud present in them, i.e.

$$f_c^{\max} = \frac{|T \cup N|}{|A|} \times 100 \quad [\%]\,, \tag{3.14}$$

and the minimum frequency of cloud occurence $f_c^{\min}$ is obtained by assuming all limb measurements in N do not have cloud present in them and is given by

$$f_c^{\min} = \frac{|T|}{|A|} \times 100 \quad [\%]\,. \tag{3.15}$$

3.4   Comparison and Validation of Detection Methods

To evaluate the two cloud detection methods they were applied to the MIPAS data from the entire month of August 2003 (staying consistent with the dataset investigated in section 3.2.2). Both the geographic distribution and the height distribution of the frequency of cloud occurrence are compared between datasets and evaluated against appropriate external datasets.

The latitudinal and longitudinal variation of cloud frequency given by each detection method is shown in figure 3.8. The MIPAS measurements were binned by expected altitude, on a 5° latitude by 10° longitude grid. Figure 3.8 shows the geographic variation of cloud frequency from MIPAS measurements with a viewing tangent height in the range 13.5 ± 3 km. For reference, the August average cloud frequency over 10 km, according to the HIRS 11-year climatology (Wylie, 2004), is given in figure 3.9.

The geographic distribution of high cloud frequency is similar for both detection methods, with a high frequency over the Southern polar region, due to PSCs, and a wide band of high cloud amount centred about the equator, with three particularly enhanced regions: over the West coast of the United States, over central Africa and over the East of the Indian Ocean and South-East Asia. Figures 3.8(a) and 3.8(b) are both very similar in geographic distribution and magnitude of cloud frequency to the HIRS August average in figure 3.9, with the exception of the Southern polar region, which shows a far higher frequency over the East in the HIRS data which is not so pronounced in the MIPAS results. However, figure 3.8(a) has a higher frequency in both the Northern and Southern mid-latitude regions

than 3.8(b), as well as an increased latitude spread of the main tropical regions of cloud enhancement. Therefore, by comparing to figure 3.9, which has regions of significant cloud frequency over Eastern Asia and Continental North America at mid-latitudes, the Window Threshold method results match the climatology better than the Colour Index method results. Furthermore, as the Colour Index method has much lower cloud frequencies in these regions, it can be concluded that the Window Threshold method is more sensitive to high cloud and quite probably, optically thin cloud, and that this sensitivity returns more geographically realistic high cloud amounts (as compared to the HIRS climatology).

In addition to cloud frequency, the detection methods can be used to return a nominal cloud top height. For a given profile, if a cloud top event occurs, then the cloud top height can be estimated to be the same as the tangent height of the cloud top event. The tangent height altitude returned in the MIPAS Level-1B (see ESA, 2000) pointing information is used for accuracy. This reported cloud top height can differ by as much as ~ 3 km from the actual cloud top height due to the uncertainty of cloud position in the MIPAS FOV and the uncertainty in the pointing information. The results for August 2003 are given in figure 3.10 for both detection methods.

From figures 3.10(a) and 3.10(b) it can be seen that the Window Threshold method, due to being more sensitive to cloud in the FOV, generally returns higher cloud top heights over all regions, particularly Northern mid-latitudes. The only exception to this is over Western areas of the Southern Hemisphere polar region, where the Colour Index method detects very high PSCs. Due to their sparse number density, and small particle sizes, PSCs have a very small enhancement on the absolute value of limb radiances, but they have a significant effect on the spectral shape, effectively flattening the emission spectra. Hence, the Colour Index method is equally as sensitive as the Window Threshold method for the detection of high level ($\gtrsim 20$ km) PSCs.

The MIPAS cloud top heights in figures 3.10(a) and 3.10(b) can be compared to cloud top heights derived from SAGE-II[2] , SAGE-III[3] and HALOE[4] measurements for the same month. The cloud top heights derived from the occultation instruments are determined in the same way as for MIPAS; by the maximum tangent height that clouds are detected at in a given profile. However, the SAGE instruments sample every 0.5 km vertically with a vertical FOV of 0.5 km and HALOE samples every 0.3 km vertically with a vertical FOV of 1.6 km, so the cloud tops should be more accurate. The SAGE-II and SAGE-III cloud tops are given in figure 3.11 and the HALOE cloud tops are shown in figure 3.12. The superior spatial and temporal sampling of MIPAS compared to both SAGE and HALOE is immediately evident from the figures. Where SAGE-II returns ~ 350 profiles

[2]SAGE II data was obtained online from Thomason (2004).
[3]SAGE III data was obtained online from Thomason (2005).
[4]HALOE data was obtained online from Remsberg and Deaver (2005).

Cloud Freq. of Occurence: TH=13.5+/-3.0km, Aug. 2003 (Thresh. Meth.)

Cloud Freq. of Occurence: TH=13.5+/-3.0km, Aug. 2003 (CI-A Meth.)

(a) Window Threshold Method

(b) Colour Index Method
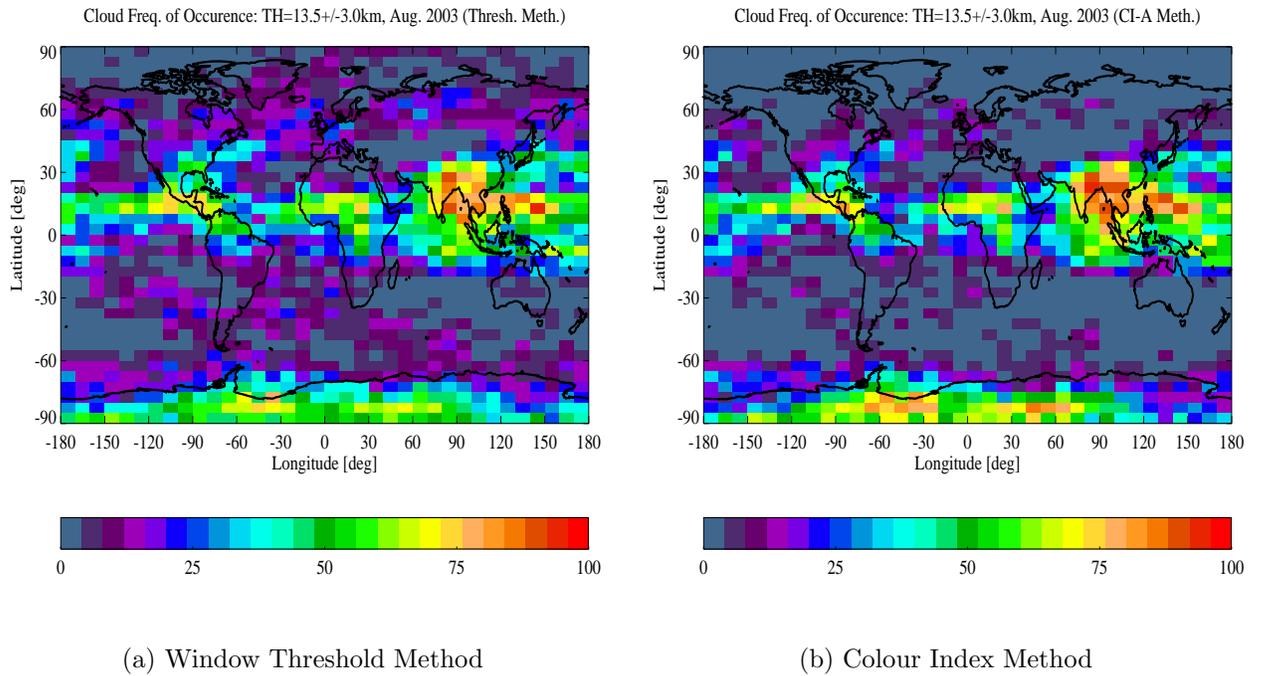
Figure 3.8: MIPAS derived cloud frequency of occurence in the range 13.5 ± 3 km. Variation with respect to latitude versus longitude, on a $5° \times 10°$ grid.

HIRS Cloud Amount over 10 km (%)
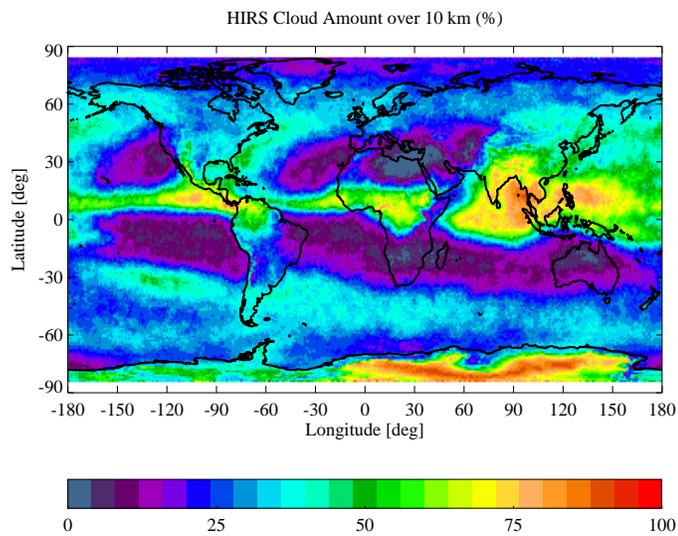
Figure 3.9: Average cloud frequency over 10 km during August from 11-years of HIRS data.

(a) Window Threshold Method   (b) Colour Index Method
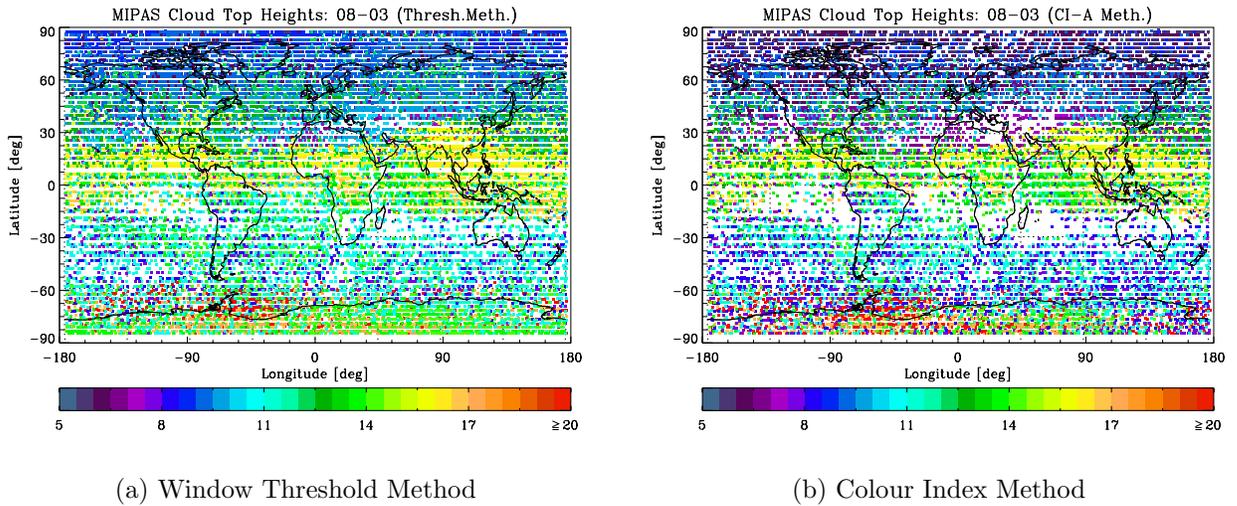
Figure 3.10: Approximate cloud top height from MIPAS for August 2003 ($\sim 30,000$ profiles). Each profile in which cloud is detected ($\sim 20,000$ profiles), is plotted by a square of which the colour indicates the nominal cloud top height in kilometeres.

per month and SAGE-III and HALOE both return ~ 750 profiles per month, MIPAS will make approximately 80 times as many profile measurements as SAGE-II and 40 times as many measurements as SAGE-III or HALOE. In addition, because the SAGE and HALOE instruments are solar occultation sounders, they only make two measurements per orbit and depending on the type of orbit the satellites are in, the sunrise and sunset events will have a very limited geographic spread. This is particularly evident in the SAGE-III data which is limited to one narrow latitude band per hemisphere. Therefore, any cloud data returned by MIPAS is particularly important because of this improvement with respect to temporal and spatial sampling over the state-of-the-art data widely used for climatology studies.

The SAGE cloud top heights seem in reasonable agreement with the MIPAS cloud tops (more so with the Window Threshold results in figure 3.10(a)), generally within 2 km with a high bias. The HALOE cloud top heights also show a similar pattern to the MIPAS data, with evidence of the same band of enhanced cloud top heights through the tropics, although the MIPAS cloud tops are generally slightly higher. However, due to the sparse sampling (both temporally and spatially) of the occultation instruments, it is not possible to make any conclusions on the global validity of the cloud top height comparison to MIPAS. In order to make a statistically valid comparison, the occultation measurements are binned by latitude and altitude and a vertical profile of the frequency of cloud occurrence is obtained (following the formalism in section 3.3). This can then be compared to the MIPAS cloud frequency (see section 3.3) obtained by binning the MIPAS measurements over the same
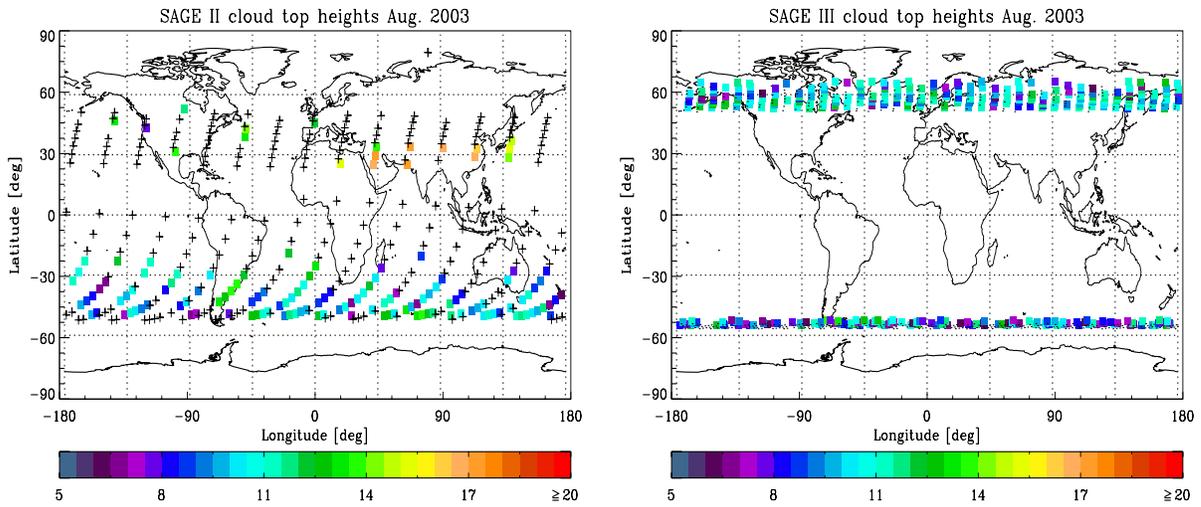
(a) August 2003 SAGE II cloud top heights. Crosses indicate profiles where no cloud is detected.

(b) August 2003 SAGE III cloud top heights. Periods indicate profiles where no cloud is detected.

Figure 3.11: Cloud top heights in kilometres derived from the SAGE instruments. The SAGE-II cloud information is derived from the Species file Cloud bit (Langley ASDC, 2003) and the SAGE-III cloud tops are derived from the Cloud Product file (Wofsy et al., 2003).
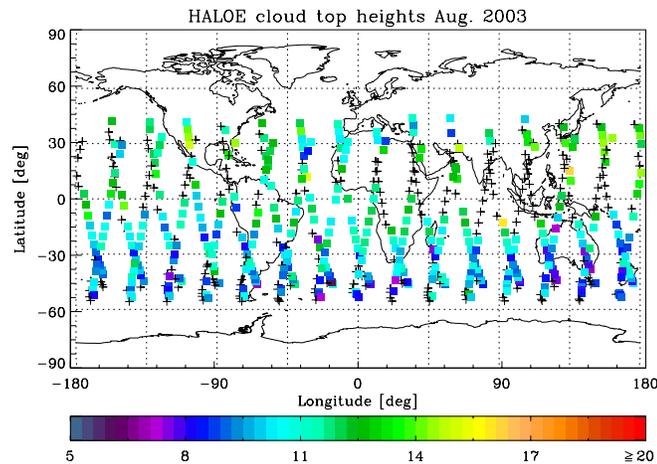
Figure 3.12: Cloud top heights in kilometres derived from HALOE. The cloud detection was determined using the algorithm of Hervig and McHugh (1999). Crosses indicate a cloud-free profile or a profile where no tropospheric measurements were recorded.

latitude range.

For the SAGE II measurements, only the Southern Hemisphere measurements are considered as there are too few cloudy profiles in the Northern Hemisphere. The SAGE II profiles are considered in the range 30 °S to 55 °S. The HALOE measurements have much greater coverage than both the SAGE datasets, but the highest concentration of data points is in the 30 °S to 55 °S latitude range, which is convenient for a secondary comparison to the MIPAS and SAGE cloud data in the same region. The SAGE III profiles in the latitude range 50 °N to 65 °N are considered to obtain a Northern Hemisphere comparison to MIPAS. The results from the comparisons between the MIPAS cloud frequency and the SAGE II, HALOE and SAGE III cloud frequencies are shown in figures 3.13, 3.14 and 3.15 respectively. For each case the cloud frequency for the appropriate latitude band derived from the HIRS August average is also plotted[5].

Intuitively, the actual cloud frequency with respect to altitude should increase monotonically as altitude decreases, as in the HIRS climatology cloud frequencies (overplot in blue) in figures 3.13, 3.14 and 3.15. However, due to the definition of cloud frequency derived from limb instruments, as given in section 3.3, the information contained in limb measurements below a cloud top event is considered indefinite with regard to cloud presence. Therefore, the calculated cloud frequency low in the atmosphere is obtained from a smaller sample set and can be expected to be lower than the true cloud frequency because it cannot account for multi-layer cloud. The problem of limb measurements being uncertain below cloud tops is evident from the increasing uncertainty in the cloud frequency range lower in the atmosphere. This trend can be seen in all of the cloud frequency profiles in figures 3.13, 3.14 and 3.15 derived from limb instruments. Further, the Southern Hemisphere cloud frequencies in figures 3.13 and 3.14 actually drop below the HIRS climatology frequency lower in the atmosphere, which is unexpected as the limb view is more sensitive to optically thin cloud. The trend is further evident by studying the maximum cloud frequencies, which exhibit the montonically increasing relationship with decreasing height because in contrast to the calculated cloud frequency, each measurement below a cloud top event is considered to contain information and it is assumed that cloud is present. The maximum cloud frequency values at 6 and 9 km are well in excess of the HIRS values, showing the limb measurements are more sensitive to cloud, though confident detection is limited lower in the atmosphere. However, the high HIRS frequencies may be artificially enhanced slightly due to the large instrument view footprint of roughly 20 km by 20 km (see Wylie et al., 1994), giving a positive detection for a measurement when the footprint is only partially filled.

The SAGE cloud frequencies peak at 9 km, then fall off at the 6 km values, dropping

---

[5]The HIRS cloud frequency is provided on a pressure height grid. This is converted to height in kilometres by using the pressure and density information from the appropriate climatology atmosphere, i.e. Remedios (1999).
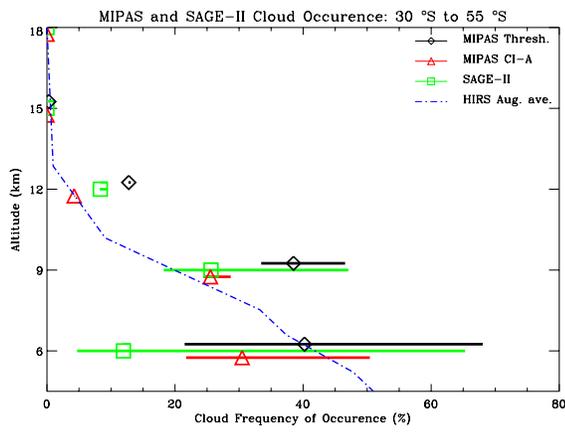
Figure 3.13: Comparison of MIPAS and SAGE II cloud frequency with respect to altitude for August 2003 in the latitude band 30 °S to 55 °S. The values at 6,9,12 and 15 km are plotted slightly offset for clarity. Each symbol represents the cloud frequency of a data set (given in the key) and the maximum and minimum frequency range is indicated by the length of the corresponding colour bar.
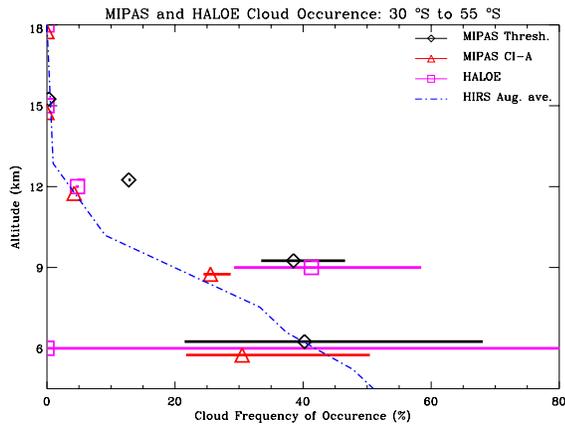
Figure 3.14: Comparison of MIPAS and HALOE cloud frequency with respect to altitude for August 2003 in the latitude band 30 °S to 55 °S.
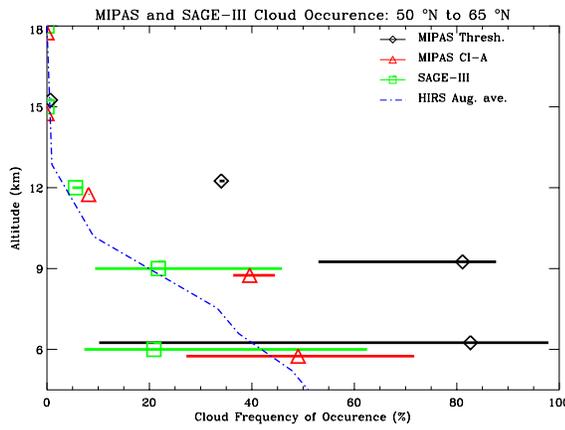
Figure 3.15: Comparison of MIPAS and SAGE III cloud frequency with respect to altitude for August 2003 in the latitude band 50 °N to 65 °N.

below the climatology frequency. This can be explained because SAGE has a high vertical resolution and a narrow vertical FOV, which means that there is a high chance of sampling cloud at a given level in the atmosphere, and thus, levels lower in the atmosphere have a higher chance of being below a cloud top event. Therefore, the size of the sample set of limb measurements in which a valid cloud test can be applied is reduced lower in the atmosphere, and the remaining measurements are more likely to be cloud-free. The sample set size is further reduced as SAGE limb scans which do not record a measurement (due to the solar beam being blocked) are discarded, and in many cases, these measurements are actual cloud top events but there is no data to confirm this. Therefore, when analysing the datasets the SAGE cloud frequencies below a certain cut-off altitude should be disregarded (see Kent et al., 1997). For tropical regions this cut-off altitude is in the region of ~ 9 km, but the level will decrease with latitude as the tropopause level decreases. For the datasets studied in this section, a cut-off level of ~ 7 km is appropriate.

The HALOE cloud frequency in figure 3.14 exhibits the same behaviour as the SAGE II cloud frequency in figure 3.13, and by similar reasoning the HALOE cloud frequencies below ~ 7 km should also be disregarded. However, the HALOE cloud frequency peaks at a much higher value than the SAGE profile, and is enhanced over both the MIPAS derived cloud frequencies at the same altitude. This could be an artefact of HALOE having a greater FOV than SAGE, or possibly HALOE is more sensitive to optically thin cirrus than the other two instruments. The range of HALOE frequencies span the Window Threshold frequency range but the Colour Index derived frequency range is too low to overlap the HALOE or Window Threshold range.

3.4.1 Discussion

In both the Northern Hemisphere (figure 3.15) and Southern Hemisphere (figures 3.13 and 3.14) samples, application of the Window Threshold cloud detection method to the MIPAS data returns significantly enhanced cloud frequency (to a possibly over-inflated value) over that obtained by applying the Colour Index method. This is a somewhat expected result as the Window Threshold method was developed as a specifically conservative cloud detection test (i.e. to "over-detect" cloud), and the threshold of 1.8 for the Colour Index method was selected knowing that it would miss some optically thin cloud (Spang et al., 2004) which can have Colour Indicies between 2 and 4, because clear sky Colour Index values can overlap. Considering the cloud frequencies derived from the occultation sounders to only be valid above 9 km, the SAGE II and HALOE cloud frequencies compare favourably to the MIPAS cloud frequencies derived from both methods (see figures 3.13 and 3.14). However, interestingly, between $\sim$9–15 km the SAGE II cloud frequency shows an enhancement over the Colour Index cloud frequency and a slight reduction compared to the Window Threshold

cloud frequency. The HALOE cloud frequency also exhibits a similar relationship between 9 and 12 km, figure 3.14 showing it is more in agreement with the higher MIPAS cloud frequency obtained from the Window Threshold detection method. This implies that the limb occultation sounders are more sensitive to optically thin cirrus high in the troposphere than MIPAS is capable of using the Colour Index detection method alone.

The SAGE III cloud frequency in figure 3.15 above 9 km is closer to the Colour Index derived cloud frequency. However, the SAGE III cloud detection algorithm (Wofsy et al., 2003) has more caveats than the SAGE II and HALOE algorithms, in that cloud presence is said to be indefinite if the spectral characteristics are too ambiguous to distinguish between cloud and aerosol. Data flagged as such was discarded from the cloud frequency calculations, but if included as cloud would have caused the SAGE III cloud frequency to be slightly higher than the Colour Index derived frequency. Furthermore, the SAGE III maximum cloud frequency is higher than the maximum cloud frequency derived from the Colour Index method, which further indicates that the actual cloud frequency is expected to be higher than the Colour Index derived values.

It can be concluded that by being less sensitive to optically thin cloud, the Colour Index method will miss a high percentage of cloud occurences compared to the Window Threshold method or to a limb occultation sounder, and this will adversely affect the quality of any MIPAS cloud data with respect to cloud climatology studies. The Window Threshold method has shown that MIPAS is theoretically more sensitive to optically thin cloud than SAGE II or SAGE III and at least equivalently sensitive to HALOE. As MIPAS has far superior spatial and temporal coverage than the occultation sounders, but a coarser vertical resolution, it is important that a cloud detection test is used which utilises the full capability of the instrument in order to maximise the possible improvement over current climatology by using cloud statistics derived from MIPAS data. The following section outlines a novel approach which was used to develop a supplementary test to the Colour Index method to improve the detection sensitivity to optically thin cloud, but without being as coarse as the Window Threshold method and thus having a higher confidence level.

3.5 A Cloud Signature from Principal Components Analysis

Principal Components Analysis (PCA) is a standard statistical technique (Murtagh and Heck, 1987) used for finding patterns in and summarising data of high dimension. Essentially, it is a mathematical procedure that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called principal components.

The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible.

PCA has rarely been used in connection with remote sensing measurements of cloud. Irwin and Dyudina (2002) applied PCA to spectra of Jupiter made by Galileo/NIMS, and the resultant set of basis functions could be used to represent the variance of real NIMS spectra. The second principle component was found to represent almost 10 % of the variance and it was representative of increased reflectivity from high jovian clouds. PCA has also been employed by Kahn et al. (2004) to assess the variability of a set of simulated TOA high-resolution infrared emission spectra of synthetic cirrus clouds with respect to microphysical and bulk cloud quantities.

This section comprises the first known attempt to apply PCA to infrared limb emission spectra with the objective of finding a component which captures the variability of limb spectra attributable to the presence of cloud in the limb path.

The advantages of applying PCA to investigate cloud presence, is that the full spectrum can be used without incurring prohibitive computational costs. In addition, the analysis is entirely empirical and thus does not depend on any assumptions about the physical nature of clouds in the measurement line-of-sight.

Although PCA is a standard technique, the application to a given problem is specific. The approach taken to apply PCA to MIPAS limb measurements and a formal description of the problem is given in appendix A, with a brief introduction in the following section.

3.5.1 PCA of the MIPAS Data

Consider repeated measurements (e.g. limb scan spectra) of some particular properties (e.g. radiance measured at different wavenumbers) of a physical object (e.g. the atmosphere viewed at a particular tangent height). Label the number of measurements m and the number of properties n, thus giving a data set as m points in n-dimensional space.

In $\mathbb{R}^n$, the space of objects, PCA searches for the best-fitting set of orthogonal axes to replace the set of n axes in this space given by the initial data. The principal components can therefore be understood as a set of basis vectors in n-dimensional space chosen so that the maximum object-to-object variation in the data belongs to the subspace formed by the least number of basis vectors.

The sample set under consideration is the same as in the previous sections of this chapter, i.e. all MIPAS band A limb scans from August 2003. The limb scans for the entire month were binned into 10 degree latitude bins for all longitudes and then subsequently binned by the Level 1B tangent altitude into the expected altitude ranges up to the lower stratosphere, i.e. 3 km altitude bins from 4.5 km to 28.5 km centred on the nominal MIPAS limb

scan tangent heights. Because limb emission spectra are highly dependent on atmospheric temperature and pressure, which in turn depends on latitude and height in the atmosphere, by binning the data as such, the temperature and pressure should be relatively stable in a given bin. Thus, the binning strategy should isolate the variance in the spectra attributable to cloud and atmospheric composition.

For the purposes of this work, MIPAS limb scans in the same latitude and tangent height bins are considered repeated measurements and the (Band A) radiance measurements from 750 cm^{-1} to 970 cm^{-1} at the 0.025 cm^{-1} MIPAS resolution are considered the measured properties. Therefore, there are m limb scans considered per height bin in a given latitude bin[6], and n (where, $n = 8801$) spectral points. For each sample bin, the PCA considers the $m \times n$ matrix of these values, and following the method in appendix A, the set of basis vectors for the data set is found by solving the eigenvalue-eigenvector problem for the corresponding correlation matrix.

3.5.2   Variance Captured by Resultant Principle Components

For each tangent height bin of each latitude bin the application of PCA yields a set of p (where $p \leq m$) principle components $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p$ of length n, i.e. the same length as the limb spectra, which are the eigenvectors of the correlation matrix for the sample bin (see appendix A). In addition, a corresponding set of p eigenvalues are also obtained. The eigenvalues are given by $\lambda_1, \lambda_2, \ldots, \lambda_p$, and the corresponding principle components are ordered by decreasing order of the magnitude of the eigenvalues. The ith principle component will capture $P_{\text{var}}(i)$ percent of the variance of the m spectra in the given bin, such that

$$P_{\text{var}}(i) = \frac{\lambda_i}{\sum_{j=1}^{p} \lambda_j} \times 100 \quad [\%]. \tag{3.16}$$

It was found that for each bin the first three principle components accounted for the majority of the variance, and the higher order components had relatively insignificant values of P_{var}, mainly accounting for the remaining spectral noise.

The total percentage variance captured by the first three principle components (i.e. $\sum_{i=1}^{3} P_{\text{var}}(i)$) for each altitude bin as a function of latitude bin is given in figure 3.16(a), and the individual percentage variance captured by the first, second and third principle components is given in figures 3.16(b), 3.16(c) and 3.16(d) respectively.

The significance of the PCA result becomes immediately apparent by comparing figure 3.16 to the frequency of cloud occurence with respect to altitude and latitude, given in figure 3.17(a). The total percentage variance captured by the first three principle com-

---

[6]Typical values of m per height bin per latitude bin approximately ranged between 1200 and 2000 limb scans.
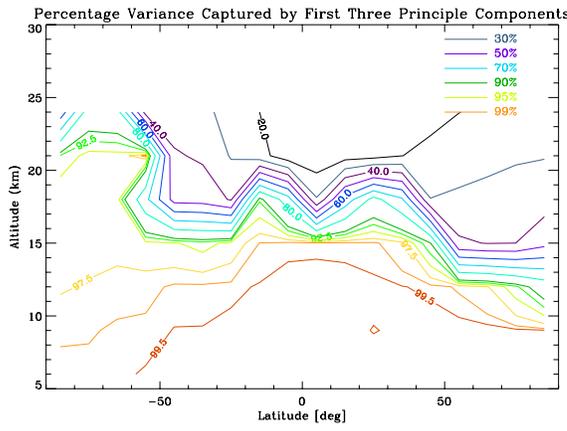
(a) Total percentage variance captured by the first three principle components, $\sum_{i=1}^{3} P_{\mathrm{var}}(i)$.
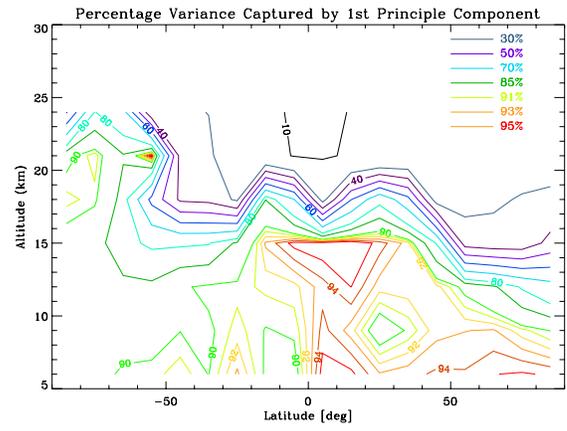
(b) Percentage of variance captured by the first principle component, $P_{\mathrm{var}}(1)$.
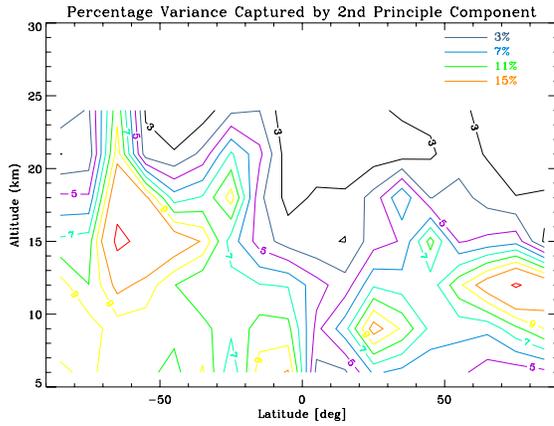
(c) Percentage of variance captured by the second principle component, $P_{\mathrm{var}}(2)$.
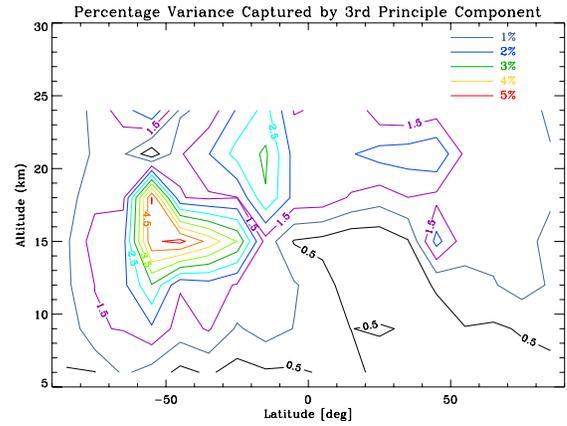
(d) Percentage of variance captured by the third principle component, $P_{\mathrm{var}}(3)$.

Figure 3.16: Percentage of the variance of the limb spectra per bin captured by the calculated principle components, with respect to the altitude and latitude of the sample bin.

ponents is strongly correlated to the cloud frequency. The patterns of the two quantities share a strong similarity[7], with peaks in captured variance corresponding directly to cloud occurrence. Both the pattern values become significant at around the tropopause level, generally increasing lower in the atmosphere. Also, both patterns are enhanced in the lower stratosphere at Southern polar latitudes, due to PSCs. However, the percentage variance in figure 3.16(a) doesn't replicate the subtleties of the cloud frequency in 3.17(a), such as the dramatic drop in cloud frequency in the sub-tropic regions around 10 °S to 20 °S and 25 °N to 35 °N, and the peak in cloud frequency at high altitudes between 10 °N and 20 °N which drops off lower in the troposphere. This is an expected result, as the

---

[7]Note that the percentage variance is sampled on a 10°latitude scale whereas the cloud frequency is given on a 5°scale.
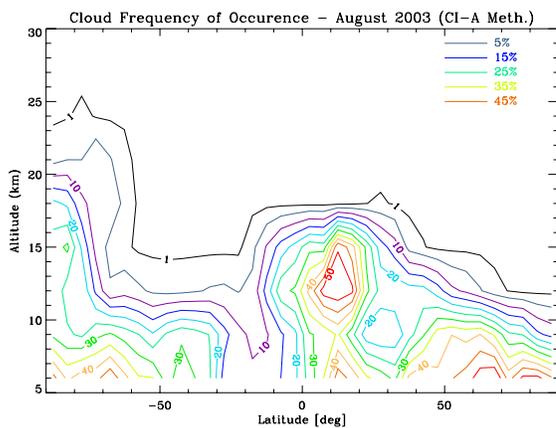
(a) Frequency of cloud occurrence; August 2003.
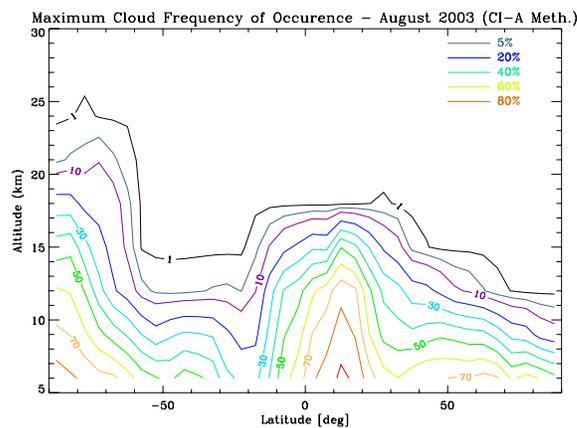
(b) Maximum frequency of cloud occurrence; August 2003.

Figure 3.17: Frequency of cloud occurrence in MIPAS spectra at a given altitude as a function of latitude. Derived using the Colour Index cloud detection method from limb spectra sampled on a 5° latitude and 3 km altitude grid.

spectra included in the PCA were not screened for cloud top events, whereas the below cloud top spectra were rejected from the analysis (see equation 3.13) used to produce the cloud frequency in figure 3.17(a). If the percentage variance of figure 3.16(a) is compared to the maximum cloud frequency in figure 3.17(b), in which the below cloud top spectra is included in the frequency analysis (see equation 3.14), the difference in the patterns is less obvious although still clearly apparent. Note that the patterns are not expected to match exactly because the variance in the limb spectra is not merely due to whether or not cloud is present but it is also due to the variability of the microphysical properties of any cloud present and its position in the FOV. Therefore, the variance accounted for is expected to be higher than the cloud frequency. The correspondence between patterns implies that principal components are highly likely to capture a spectral cloud signature.

Out of the first three principle components alone, the percentage variance captured by the first component appears to correspond most strongly with the presence of cloud. The pattern of variance captured by the first principle component in figure 3.16(b) has a strong match to the cloud frequency in figure 3.17(a), replicating the subtle features (unlike the variance of the first three components combined in figure 3.16(a)). The regions of low cloud frequency in the sub-tropic regions clearly corresponds to drops in the variance captured although the reduction is relatively less significant. This strong match implies there may be some information about cloud top measurements in the first principle component.

The variance accounted for by the second principle component shown in figure 3.16(c), varies considerably. The pattern of variance has little structure but it does appear anti-correlated with the cloud frequency in figure 3.17(a). The reason for this could be that

the second component is also significantly correlated with the atmospheric temperature and density. This correlation is more evident where the measurements are not dominated by cloud presence. The percentage variance captured by the third component has a much different pattern to the cloud frequency, with $P_{var}(3)$ peaking between 12–15 km in the 10 °S to 60 °S region, which is the region in which the cloud frequency is minimal.

If cloud occurs in these regions it is most likely to be optically very thin, and it is known that the Colour Index method is somewhat insensitive to optically thin cirrus, so even though the cloud frequency doesn't show much detected cirrus in this region of (figure 3.17(a)) it may be the case that the third principle component is tuned to the variance arising from spectral features specific to tenuously thin cloud. This is also implied by the general decrease in the variance captured lower in the atmosphere, where optically thin cloud is less likely to occur. However, from the simple comparison of the percentage variance accounted for by each of the principle components to the frequency of cloud occurrence, the first component is the strongest candidate to capture a spectral cloud signature, due the high percentage of variance accounted for and the strong correlation between variance captured and cloud frequency.

In order to understand the significance of the strong correlation between the principle components and the presence of cloud, the spectral variability of the principle components themselves and the similarity to the spectral effect of clouds on emission spectra is investigated.

3.5.3 Significance of the Spectral Variability of Principle Components

It would be uneconomical and unwarranted to present the spectral analysis of the principle components from each of the 126 (7 altitude × 18 latitude) sample bins, as the volume of data is greater than the information contained in the results. For this reason, only examples typical of the whole sample set are focussed on here.

For this section, the latitude bin between 0 °N and 10 °N is considered. The average of the spectra in the 12 ± 1.5 km tangent height bin is shown in figure 3.18(a), and the first three calculated principle components are shown in figures 3.18(b), 3.18(c) and 3.18(d). The higher order principle components account for less than 0.5 % of the variance each, and essentially look like noise spectra.

The averaged spectrum in figure 3.18(a) is heavily influenced by cloud presence (e.g. the raised continuum level), and looks similar to the modelled thin cloud 12 km limb spectra in figure 3.1. Both tropical anvil cirrus and optically thin cirrus are common at 12 km in such regions. From figure 3.17(a), the cloud frequency at 12 km in that region is $\sim 45\,\%$ and
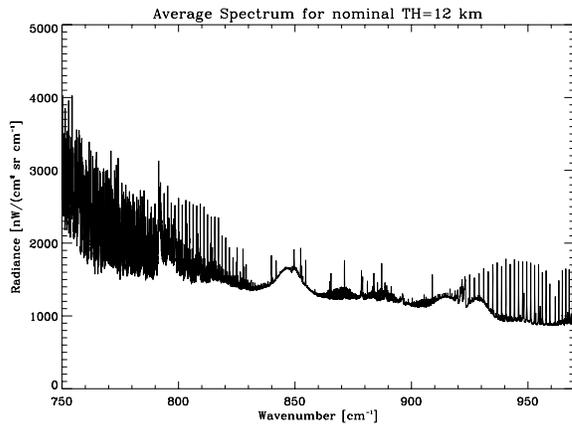
(a) Average of all spectra in the sample bin.
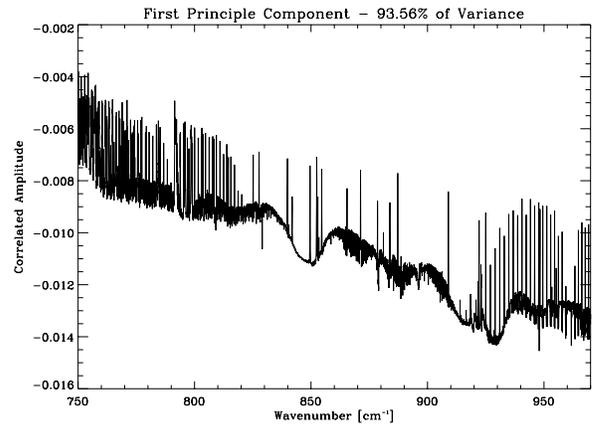
(b) First principle component, which captures 93.56 % of the variance.
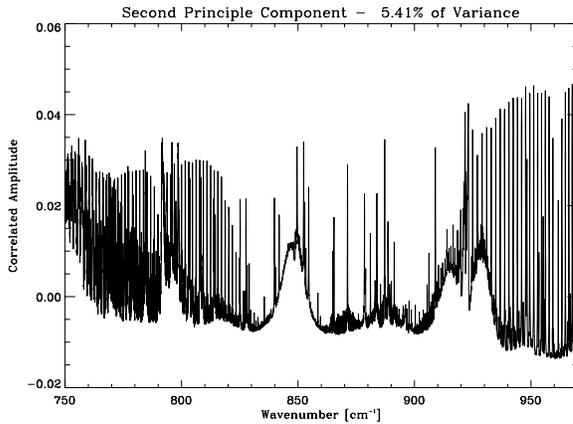
(c) Second principle component, which captures 5.41 % of the variance.
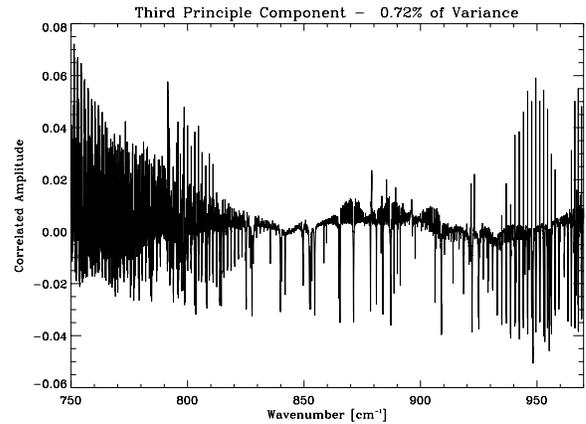
(d) Third principle component, which captures 0.72 % of the variance.

Figure 3.18: August 2003 PCA results for the 12 ± 1.5 km tangent height bin of latitude range 0 °N to 10 °N.

from figure 3.17(b) the percentage of cloud contaminated spectra is $\sim 60\,\%$. The first principle component captures the majority of the variance ($93.56\,\%$) associated with the presence of cloud and the variability of the microphysical properties. Furthermore, it can be seen from figure 3.18 that the second and third principle components have spectrally average base values near zero, but the first principle component has a significantly non-zero base continuum level. On appearance alone, the first principle component seems to have a spectral structure similar to that expected from a cloudy spectra.

The limb measurements themselves can be approximated using the principle components as in equation A.8. Essentially, a contribution coefficient scales each principle component and it is added to the average spectrum to obtain the approximation. The contribution coefficient is obtained from equation A.9. The first principle component is obviously the most important contribution to this approximation. Intuitively, if the average spectrum in figure

3.18(a) is considered to be similar to moderately thin cirrus (the average state between clear-sky and cloudy spectra, weighted by the cloud occurrence frequency), then to obtain thick cirrus, which will have a higher continuum baseline, a large negative coefficient will be required to give a positive contribution due to the first principle component and increase the continuum level. To obtain a spectra for thin cirrus for which the spectra is enhanced by scattering from below the cloud, then a negative coefficient will also be required. A small positive coefficient may be required for tenuously thin cirrus, to bring the continuum level down and a large positive coeffcent will be required to approximate clear-sky spectra to bring the continuum down toward zero. In order to confirm and visualise what is intuitively obvious about the contribution coefficent of the first principle component, c_1, it is plotted against both the band A Colour Index and the Window Threshold radiance in figures 3.19(a) and 3.19(b) respectively. However, the relationship between the contribution coefficent of the second and third principle components (c_2 and c_3) and the cloud type is far less obvious. The relationships between c_2 and c_3 and the cloud detection measures are shown in figures 3.19(c)–(f).

From figure 3.19(a) it can be seen that there is a strong relationship between c_1 and the Colour Index. The points form a reasonably well-defined, asymptotic curve, with large negative values of c_1 for Colour Index values near 1, which is indicative of opaque cirrus. The curve rapidly increases in c_1 over a small range of Colour Index values and then levels off at a high positive coefficient for Colour Index values greater than ~ 4, which is indicative of clear-sky spectra. The relationship between c_1 and the Window Threshold values in figure 3.19(b) is practically linear, with a slight spread most likely due to measurement noise. The linear relationship between c_1 and the Threshold radiance indicates that cloud information based on the values of c_1 will be scalable and highly sensitive.

Figures 3.19(c)-(f) show that although there is clearly evidence of a weakly logarithmic increase in c_2 and c_3 with respect to Colour Index and a vaguely linear relationship to the Window radiances, there is a large spread in the points. Thus, cloud information contained within c_2 and c_3 is likely to be dependent on the inter-relationship between coefficients. Take c_3 for example, as it has less spread than c_2, the relationship to c_1 is shown in figure 3.20. The blue line overplot in figure 3.20 is representative of measurements of scattering cirrus and the red line is representative of clear-sky measurements. This can be understood by thinking of how a thin cirrus limb spectra enhanced by scattering can be approximated:

- As the first principle component has a negative continuum level (see figure 3.18(b)), a negative c_1 is required to increase the continuum line over the average.

- In order to add the absorption lines which would be caused by the scattering of radiation into the line-of-sight, a positive c_3 is required. This is because the third
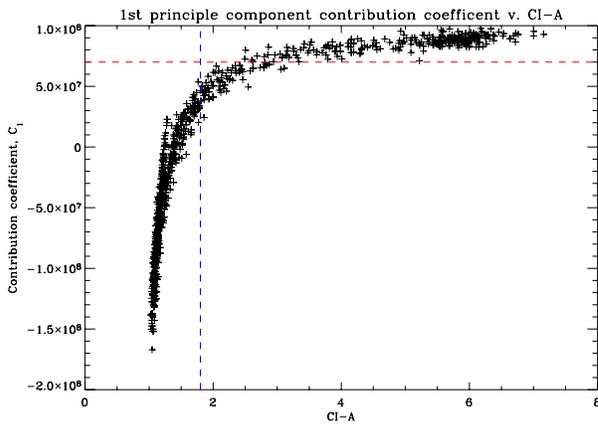
(a) Contribution coefficent c_1 plotted against the Colour Index. CI-A=1.8 is in blue and $c_1 = 7 \times 10^7$ is in red.
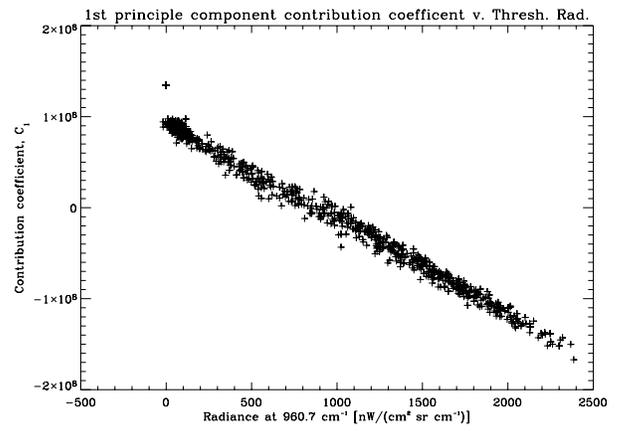
(b) Contribution coefficent c_1 plotted against the Window Threshold radiance.
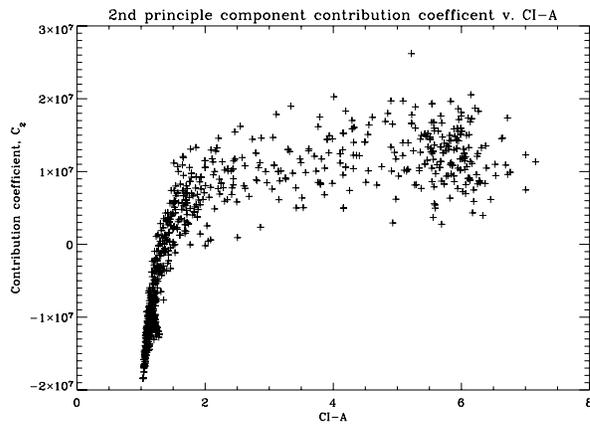
(c) Contribution coefficent c_2 plotted against the Colour Index.
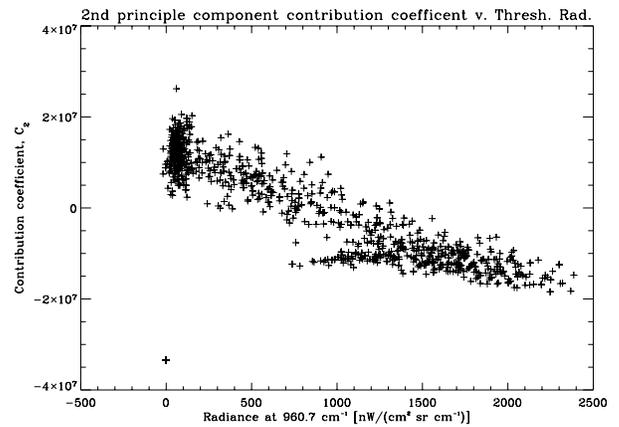
(d) Contribution coefficent c_2 plotted against the Window Threshold radiance.
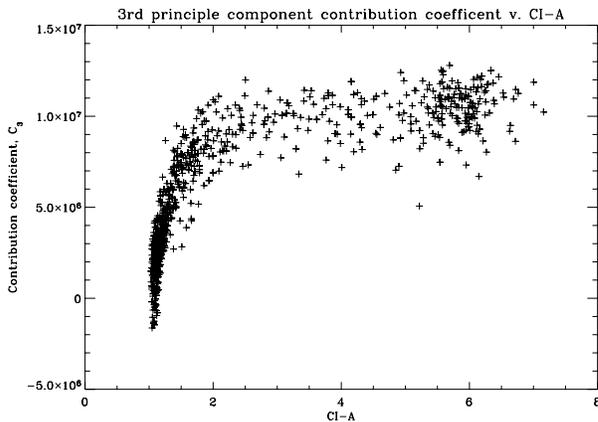
(e) Contribution coefficent c_3 plotted against the Colour Index.
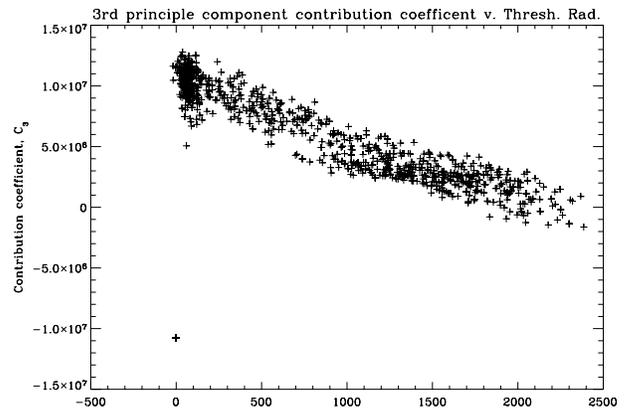
(f) Contribution coefficent c_3 plotted against the Window Threshold radiance.

Figure 3.19: The contribution coefficents for the first three principle components to approximate each individual limb scan in the 12 ± 1.5 km sample bin, plotted against both the band A Colour Index and the radiance at the Window Threshold of the limb scan.
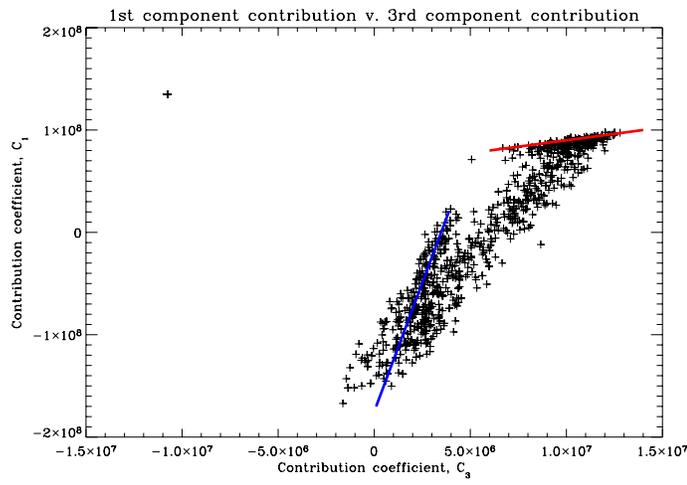
Figure 3.20: The contribution coefficient of the first principle component c_1 versus the contribution coefficient of the third principle component c_3.

principle component has inverted water vapour lines (see figure 3.18(d)).

By the same logic, a clear-sky measurement requires both a positive c_1 and a negative c_3. Figure 3.21 plots the spectral region between 940 and 950 cm^{-1} from an unapodised limb measurement corresponding to a point on the blue line such that $c_1 = -8.71 \times 10^7$ and $c_3 = 1.3 \times 10^6$. The clearly visible water vapour and CO_2 absorption features in this region can only be explained by photons scattered into the limb view from below the cloud. It is not clear how to interpret the measurements represented by the broadly linear continuum of points between the blue and red lines in figure 3.20. As the aim of this chapter is to improve on the information on thin cloud obtained by the existing detection methods, the focus is on the use of only the first principle component to determine thin cloud sensitivity as it has the strongest cloud signature. Although the investigation into the relationship between coefficients is not continued further, it has potential for future work.

The relationship of c_1 to the Colour Index follows the same asymptotic trend in all sample bins in which the first component accounts for above about 50 % of the variance. However, in some sample bins the relationship is inverted, with the clear-sky values subtending a negative value of c_1, and high positive values of c_1 approaching a Colour Index of 1. This is easily explained because in these cases the continuum baseline of the first principle component is positive non-zero, but the sign of the vector is not significant. The magnitude of c_1 subtended by the high Colour Index values in the inverted relationships is similar to the value subtended in spatially adjacent sample bins.

A cloud detection scheme can be introduced by setting a limit on c_1. For example, the red line in figure 3.19(a) separates the cloudy spectra below it from the clear-sky spectra above it. However, the prime requisite in developing an additional detection method is that it improve the differentiation between clear-sky and cloud, particularly where the
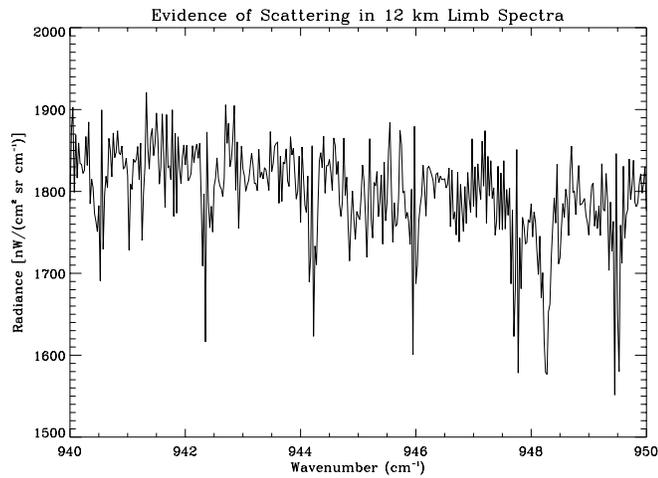
Figure 3.21: Evidence of absorption lines due to scattering in the 12 km limb spectra approximated using $c_1 = -8.71 \times 10^7$ and $c_3 = 1.3 \times 10^6$.

Colour Index has problems, i.e between 1.8 and 4, where tenuously thin and clouds which partially fill the MIPAS FOV go undetected. Therefore, the value of the limit on c_1 (e.g. the red line in figure 3.19(a)) is set by maximising the difference in the average spectrum calculated from the limb measurements with Colour Index values between 1.8 and 4 on each side of the line. Further, the percentage population of points in each region should be at worst a 60 to 40 percentage split so that maximising the difference in the spectra is statistically significant. Comparing the c_1 threshold to the Colour Index threshold which is plotted in the blue line in figure 3.19(a), it can be seen that for Colour Index values less than 1.8, the new test yields no further information. For Colour Index values greater than 1.8 and less than 4, the Colour Index test would not detect a large number of cloudy measurements represented by the points to the right of the blue line and below the red line in the figure. This effect is shown in figure 3.22, by comparing the average spectrum from the measurements in each of the three possible cases, i.e.

1. Flagged as cloud by the Colour Index and c_1 threshold: to the left of the blue line and below the red line in figure 3.19(a).

2. Flagged as clear-sky by the Colour Index and cloud by the c_1 threshold: to the right of the blue line and below the red line) in figure 3.19(a) (and with CI-A < 4).

3. Flagged as clear-sky by both methods: to the right of the blue line and above the red line in figure 3.19(a) (and with CI-A < 4).

By comparing figure 3.22 to figure 3.1, it can be seen that the three cloud detection cases above have a strong similarity to the modelled spectra of the three regimes of no cloud, thin cloud and thick cloud with line colours corresponding. This implies that the Colour Index method is only efficient at detecting optically thick cloud confidently, and that using
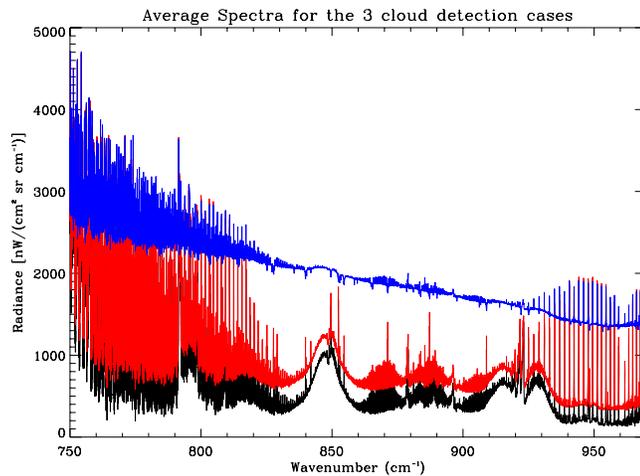
Figure 3.22: The average spectra for the three cloud detection cases at 12 km. The blue line is the average of the case 1 spectra ($\sim$ optically thick cloud), the red line is the average of the case 2 spectra ($\sim$ optically thin cloud), and the black line is the average of the case 3 spectra ($\sim$ clear-sky).

the c_1 threshold test can effectively separate optically thin cloud spectra from clear-sky spectra (as evidenced by the enhanced case 2 average spectra over the case 3 average spectra). The Window Threshold radiance for the case 2 and case 3 average spectra is 339.23 nW/(cm^2 sr cm^{-1}) and 135.61 nW/(cm^2 sr cm^{-1}) which further implies that the case 3 average spectrum is very close to being typical of a clear-sky measurement. The percentage population split was 53 % case 2 and 47 % case 3. There may still be some cloud contaminated measurements flagged as clear by the contribution coefficent limit, but the method offers a considerable improvement over the Colour Index method which would not have detected a significant number of measurements with optically thin cirrus or partially filled FOV cirrus measurements.

3.5.4   Detecting Optically Thin Cloud via the Principle Component Contribution

In the previous section it was shown that for a given limb scan in a sample bin in which the first principle component captures a significant amount of variance, the contribution coefficient c_1 can be used effectively as an indicator of whether or not cloud is present in the limb view. However, as the principle components and therefore the resultant contribution coefficents are empirical functions derived for each sample bin, the relationship between c_1 and the cloud presence (or a derived measure of cloud presence such as the band A Colour Index) will be dependent on the latitude and altitude of both the limb measurement and the empirical principle components used. Essentially, any cloud test based on the contribution coefficient must to some extent be tailored to each sample bin location; this is non-desirable,
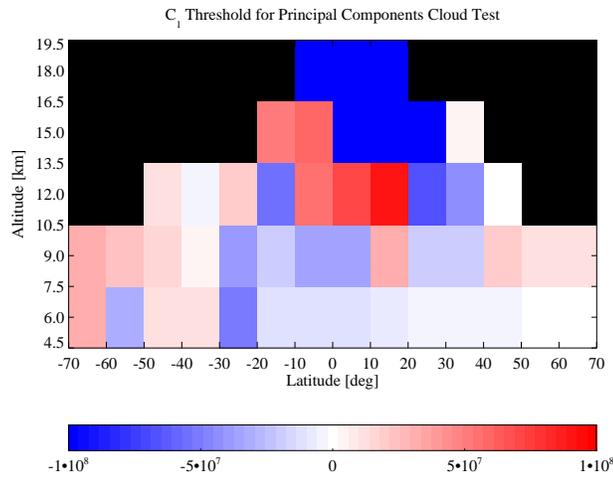
Figure 3.23: The threshold limit c_1 by sample bin for application of the principal component cloud test. The region in black is where the test is not applicable due to the variance associated with the first principle component being too low. The threshold value of the bin centred at (15 °N, 15.0 km) and the three bins centred at 18 km exceeds the lower limit of the colour scale. The actual value of the (15 °N, 15.0 km) bin is -1.4×10^8 and the 18 km bin values are from South to North: -1.6×10^9, -8×10^9, and -5.5×10^9.

but unavoidable.

Firstly, only those sample bins for which the value of $P_{\mathrm{var}}(1) > 50\,\%$ are considered for application of the test. Secondly, the polar regions are neglected as they are complicated due to the presence of PSCs, i.e. only 70 °S to 70 °N is considered.

For each sample bin a limit c_1^{lim} is set on the contribution coefficent of the first principle component for the limb spectra (given in figure 3.23). Following the same method as the previous section, the limit is set such that for all limb spectra with $1.8 \leq$ CI-A ≤ 4, the percentage split between the number of spectra with $c_1 < c_1^{\mathrm{lim}}$ and the number of spectra with $c_1 > c_1^{\mathrm{lim}}$ should be 60 to 40 or better. In addition, the difference between the average spectrum of the spectra with $c_1 < c_1^{\mathrm{lim}}$ and the average spectrum of the spectra with $c_1 > c_1^{\mathrm{lim}}$, should be maximised.

The *principle component cloud test* is then applied by applying the following steps in order:

1. Determine the contribution coefficient c_1 for the limb measurement, using the principle component and average spectrum for the appropriate latitude and altitude sample bin.

2. Determine the spectral average value of the first principle component for the sample bin, given by

$$\bar{\mathbf{u}}_1 = 1/n \sum_{j=1}^{n} u_{1,j}. \tag{3.17}$$

3. From the tabulated values (shown graphically in figure 3.23), determine the sample

bin contribution limit for the test, i.e. c_1^{lim}.

4. If $\bar{\mathbf{u}}_1 < 0$ then the measurement is flagged cloudy if $c_1 > c_1^{\text{lim}}$, or else if $\bar{\mathbf{u}}_1 > 0$ then the measurement is flagged cloudy if $c_1 < c_1^{\text{lim}}$.

Applying the test to the MIPAS data for the month of August 2003 used in the comparison to the SAGE II and HALOE data in section 3.4 yields the cloud frequencies (at 6, 9, 12 and 15 km) given in figure 3.24. It can be seen that application of the principle component cloud test yields a cloud frequency at 9 and 12 km more in line with the SAGE II derived cloud occurrence than from the Colour Index method applied alone. The cloud occurrence derived from the principle component cloud test is now significantly enhanced over the Colour Index derived cloud frequency, due to the enhanced sensitivity of the detection test to optically thin cloud at high levels, while remaining more selective than the Window Threshold method. Of the three cloud detection methods applied to MIPAS data, the range in cloud frequency at 9 km returned from the principle component cloud test is in best agreement with the two external datasets.

However, the principle component cloud test does have several minor drawbacks, in particular that it is more complicated to apply (mainly due to the altitude and latitude dependence) than either of the existing methods, and thus it may be more prone to user error. To limit the error, the test need only be applied for high altitude cases in which optically thin cirrus is likely and when the Colour Index method fails to detect cirrus but returns an Index in the range 1.8 to 4. In addition, the principle component cloud test has been developed having applied the PCA only to the MIPAS data from August 2003, and it is likely that the principle components in each sample bin will also be dependent on season, as the atmospheric temperature and density can be expected to vary. The PCA was only applied to a month of data in order to avoid a more computationally onerous singular value decompostion (used to solve the eigenvalue-eigenvector problem - see section A.3) when the number of spectra per sample bin is far greater than the number of spectral points, i.e. $m \gg n$. However, for this thesis, high optically thin cirrus is of most interest, which generally occurs in the tropics. As the upper tropospheric temperature and pressure is more stable seasonally than for other regions, it is assumed that the pressure and temperature conditions throughout the month of August are typical of the annual variation and therefore that the derived principle components can be expected to vary little seasonally. Thus, in this thesis, for MIPAS limb measurements not taken from August 2003, the principle component cloud test will only be applied for high limb measurements > 10 km in the tropics, between 30 °S and 30 °N.
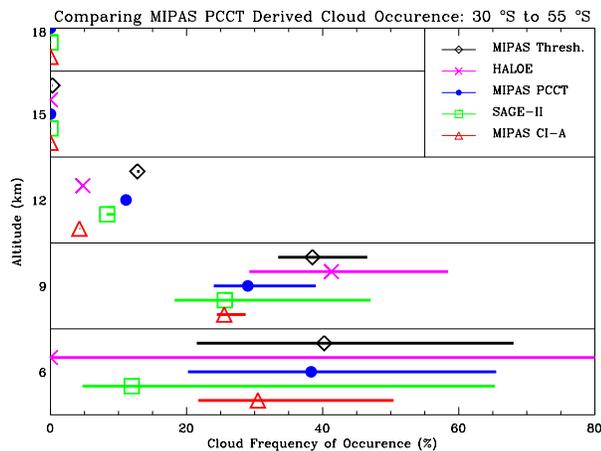
Figure 3.24: Cloud occurrence derived from various methods and instruments for limb measurements in the region 30 °S to 55 °N from August 2003. Three different methods are shown which were applied to MIPAS: Window Radiance Threshold test, Colour Index method, and the Principle Component Cloud Test. Also shown are the cloud occurrences obtained from SAGE II and HALOE.

3.6   Conclusions

The two existing methods of MIPAS cloud detection were introduced, and in section 3.4 it was shown that both methods when applied to a month of data yield cloud frequency of occurrence results in line with climatology (taking the expected increased sensitivity in the limb into account). It was also shown in section 3.4, from comparison to occultation measurements from the same time period, that the Window Threshold test may have a propensity to over-detect high thin cloud and the Colour Index method alone does not optimise the sensitivity of MIPAS and misses a significant percentage of high thin cloud. The simulation of MIPAS measurements in section 5.2 with full consideration of scattering effects, also show that the Colour Index method will not detect tenuously thin cirrus. As a consequence of this result, the PCA derived cloud detection test was developed in section 3.5.

The principal component cloud test was shown to detect much of the optically thin cirrus that the Colour Index method will miss. However, the principal component cloud test is dependent on both altitude and latitude, and is thus a less general test than either of the existing methods. This is not necessarily a limitation of the method, as it merely reflects the complexity of the physics involved when tenuously thin cirrus or cirrus which partially fills the FOV is measured. The aim of this chapter was to design an improved cloud detection test, to ensure optically thin cloud measurements are succesfully detected and input to the retrieval. The supplementary test developed achieves this aim, giving an improved sensitivity to optically thin cirrus over the Colour Index method alone and a more reliable test than the coarse Window Radiance Threshold method. In addition, this

chapter has shown that there is great potential in the PCA approach to cloud detection and that if more time were dedicated to investigate the detailed results then a superior cloud detection test may be possible. This would require further investigation of the inter-relationship between the principle components, and applying the PCA to a wider dataset, with a larger (or more selectively chosen) spectral range, or with a larger coverage with respect to latitude and altitude. However, the size of the dataset will be limited by the computational effort required to calculate the singular value decomposition.

For the purposes of this thesis, the principal component cloud test as outlined in sections 3.5.3 and 3.5.4 will be used as a detection test for optically thin cloud, applied to limb measurements in the tropical region between 30 °S and 30 °N for which $1.8 \leq$ CI-A ≤ 4.

Chapter 4

The Monte Carlo Cloud Scattering Forward Model

4.1 Introduction

In order to determine cloud properties from limb emission measurements the first requirement is a radiative transfer model to simulate the scattering of atmospheric thermal radiation by cirrus clouds. Such a model must be able to model spherical geometry and multiple-scattering radiance. Accurate radiative transfer models which satisfy these requirements for simulation of infrared limb emission spectra at high spectral resolution are not yet publicly available. Recent studies on limb emission measurements have relied on simpler calculations to model the scattering features caused by cloud. Spang et al. (Spang et al., 2004) investigated the effect of adding an arbitrary 10% radiance contribution of a nadir spectrum to a line-by-line calculated limb spectrum. They found similar spectral features in their model spectra compared to observations by MIPAS but the continuum signal was matched poorly with discrepancies of up to 40%. The more rigorous study by Höpfner (Höpfner, 2004) on modelling the effects of PSCs on limb spectra included Mie single scattering into a forward line-by-line code. The method was successful in modelling the scattering features in MIPAS balloon spectra of PSCs (Höpfner et al., 2002), and was able to match the baseline continuum signal to within about 15% of the measured spectra. Other than not accounting for multiply scattered photons the drawbacks of this method is the time consuming calculation of the scattering source function and the necessity to calculate the incident radiation for a variety of angles for each atmospheric path segment. The Monte Carlo Cloud Scattering Forward Model (McClouds_FM) was developed for this thesis to accurately model infrared limb emission measurements of cirrus clouds, inclusive of multiple-scattering, spherical geometry and accurate cloud ice particle scattering properties. McClouds_FM is the first radiative transfer model designed specifically for this purpose, and

is a valuable tool for improving the understanding of the role played by multiple scattering in the analysis of limb measurements of cirrus. In addition, it is particularly powerful as it is able to model high resolution infrared spectral limb measurements. The goal of McClouds_FM was that it should be able to replicate such limb measurements, e.g. from MIPAS, to within instrument noise and model error. Nominally, McClouds_FM should be able to reproduce limb measurements to within 5% of the continuum baseline level and spectral average.

This chapter describes the development of McClouds_FM following Ewen et al. (2005).

4.1.1 Definition of the Forward Model

A forward model used to simulate indirect measurements (e.g. atmospheric radiances at the satellite) of clouds can be represented by

$$\hat{\mathbf{y}} = \mathbf{F}(\hat{\mathbf{x}}_{\mathbf{c}}, \hat{\mathbf{b}}) \tag{4.1}$$

where $\hat{\mathbf{x}}_{\mathbf{c}}$ and $\hat{\mathbf{b}}$ are the best estimates of the cloud and atmosphere parameters. The vector $\hat{\mathbf{x}}_{\mathbf{c}}$ is known as the state vector and contains all parameters necessary to characterise the cloud field: cloud particle effective radius, total number density of cloud particles, cloud top height and vertical depth of the cloud layer. The vector $\hat{\mathbf{b}}$ is the set of parameters required to characterise the atmosphere: these are critical to the resultant synthesised radiance but are chosen not to be included in the state vector (e.g. molecular cross sections).

The forward model must combine an estimate of the atmospheric radiance with the physics of the measurement. McClouds_FM breaks the problem down into two steps: computing the radiative transfer through the atmosphere, then modulating the radiance to take into account the effects of the observing instrument e.g. by modelling the field-of-view. In cloudy atmospheres, scattering and molecular absorption are involved in the radiative transfer which makes the physics of the forward function very complex. McClouds_FM also splits the radiative transfer calculations into two regimes. The model is separated into a scattering domain, i.e. the cloud volume, and a non-scattering domain, i.e. the clear sky. In the clear sky domain line-by-line radiative transfer is used to calculate the emission and transmission along the ray paths. In the cloud domain a Monte Carlo *reverse* scattering model is used to simulate photon trajectories from a pencil beam from the detector arriving at the cloud in the backward direction. Effectively, McClouds_FM uses Monte Carlo radiative transfer to trace all photon ray paths backwards from the satellite then calculates the reciprocal radiative transfer along each path to determine the total radiance arriving at the detector.

4.2 Assumptions for McClouds_FM

This section outlines the basic assumptions made in order to simplify the radiative behaviour of the cloud-atmosphere system and to limit the number of parameters required to quantify the cloud in the model. The main consideration when making these assumptions is to preserve accuracy while making the forward model fast enough so that it can be used to retrieve cloud properties from limb emission observations.

4.2.1 Single Scattering and Microphysical Properties

Cirrus clouds are composed principally of ice crystals with various non-spherical shapes. The sizes of the ice crystals can range from less than 10 μm to about 4000 μm (Heymsfield and Miloshevich, 2003), and the typical geometry of the crystals can range from single pristine shapes such as hexagonal ice columns or plates, to complex aggregates of columns or bullets (Baran, 2004). There are many different definitions of cirrus, but for the purposes of this model cirrus clouds are considered to be composed entirely of (randomly oriented) non-spherical ice crystals and to occur at high altitudes. In addition, the cirrus is assumed to be homogeneous, i.e. the cloud is composed of a single crystal habit, the cloud is spatially contiguous, and that there exists a size distribution which is representative of the entire cloud volume.

A single crystal habit is assumed in order to limit the degrees of freedom for the retrieved solution in the measurement inversion and to reduce the complexity and computation time of the scattering calculations in the forward model. There is a wealth of evidence from *in situ* aircraft measurements showing that cirrus with relatively small particle sizes (i.e. < 150 μm) are frequently observed to have a dominating crystal habit (Heymsfield and McFarquhar, 2002). Additionally, it is common for single-view instruments (such as MODIS: The MODerate Resolution Imaging Spectrometer) to make some assumption about crystal shape in their retrieval of cirrus properties (e.g. Huang et al., 2004).

As McClouds_FM incorporates multiple scattering effects it is important to have accurate single scattering properties for the cloud particles. The use of the correct ice crystal scattering properties is key to reducing errors during the retrieval of optical thickness or microphysical properties of cirrus from remote measurements (Huang et al., 2004; Baran et al., 1999). Due to the non-sphericity of ice crystals, the theoretical description of their single-scattering properties is far more complicated than for their spherical counterparts. The single scattering properties used in McClouds_FM are the extinction coefficient, β_{ext}, single-scatter albedo, ω_0, asymmetry parameter, g, and the scattering phase function. Therefore, McClouds_FM is flexibly configured to use any non-spherical crystal shape assuming that the single scattering properties input are accurately calculated.

Throughout this thesis, unless otherwise stated, hexagonal columns are assumed since for subvisual and non-convective laminar layer cirrus it is likely that such cloud consists of small pristine faceted ice crystals of less than 100 μm. For such small ice crystals there is no current evidence to suggest that these are not hexagonal. In addition hexagonal columns are a basic constituent of cirrus ice crystals and have been shown to reproduce certain well-known cirrus optical phenomena (Platt and Dilley, 1984).

The single scattering properties used for hexagonal columns are calculated as in Baran and Francis (2004) for wavelengths between 5 μm and 16 μm and single particles of maximum dimensions ranging from 3.0 μm to 3500 μm. T-matrix methods (Mishchenko, 1991; Havemann and Baran, 2001) are applied to hexagonal columns to calculate total optical properties (i.e. β_{ext}, ω_0 and g) and phase functions for maximum dimensions up to about 40 μm. For sizes from 40 μm to about 175 μm T-matrix methods were applied to the equal area circular cylinder to compute the optical properties and phase function. Then finally the T-matrix electromagnetic solutions are supplemented for maximum dimensions greater than about 175 μm by the Complex Angular Momentum (CAM) approximation. The CAM method is used to calculate the total optical properties but cannot be used to calculate the phase function, which has been calculated using a piecewise linear extension of the Henyey-Greenstein phase function due to Baran et al. (2001). The single scattering properties for the varying sizes of individual crystals are integrated over cirrus size distributions derived from *in situ* measurements (Fu, 1996) to give the bulk cloud scattering properties. The size distributions are described by the effective radius, r_e, and the total number density, N (i.e. the total number of cloud particles per unit volume). In order to avoid ambiguity in the meaning of an effective dimension for a cloud composed of ice crystals, the definition of effective radius r_e used in this thesis will be

$$r_e = \frac{3}{4} \frac{\text{IWC}}{\rho \sum n_j A_j} \tag{4.2}$$

as given by Francis et al. (1999), where IWC is the ice water content (mass per unit volume), n_j is the crystal concentration in the jth size bin, and A_j is the mean cross-sectional area of this bin, and ρ is a reference density of 1 g cm^{-3}. The total number density is defined as

$$N = \int_0^\infty n(L) \, dL \simeq \sum n_j \tag{4.3}$$

where L is the maximum particle dimension. The IWC can be expressed as

$$\text{IWC} = \rho_{\text{ice}} \sum n_j V_j \tag{4.4}$$

where V_j is the mean volume of a crystal in the jth size bin and ρ_{ice} is the density of ice.
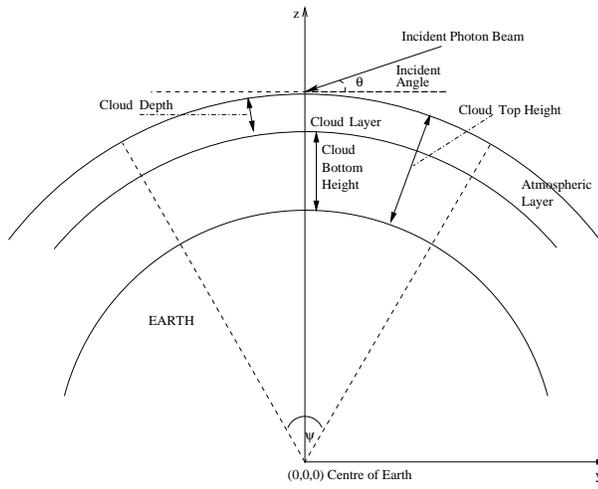
Figure 4.1: 2-D Slice of cloud/Earth geometry.

If the size distribution is normalised to a different total number density, the effective radius, single scatter albedo, asymmetry parameter and phase function are unchanged. However, the extinction coefficient and the IWC change by a factor equal to the ratio of the new number density compared to the original. At a given wavelength the bulk single scattering properties for a cirrus cloud are characterised by the crystal habit, the effective radius of the polydispersion of cloud ice particles and the total number density.

4.2.2 Cloud Geometry

The cloud geometry is approximated as a uniform depth spherical shell layer in the atmosphere, with a finite extent given by the angle ψ that the cloud edge subtends at the centre of the Earth. The geometry is then defined by the cloud top height, C_{top} (measured radially from the Earth's surface) and the vertical (radial) depth of the cloud layer, C_{depth}. The cloud volume is perhaps best described in Earth centred spherical coordinates as the locus of points $\{(r, \phi, \theta)\}$ such that

$$R_{\text{Earth}} + C_{\text{top}} - C_{\text{depth}} \leq \; r \; \leq R_{\text{Earth}} + C_{\text{top}} \tag{4.5}$$
$$0 \leq \; \phi \; \leq \frac{\psi}{2}$$
$$0 \leq \; \theta \; \leq 2\pi$$

where R_{Earth} is the radius of the Earth. The geometry of cloud in the model is illustrated in figure 4.1, where the cloud is rotationally symmetric about the z-axis. In addition, for simplicity in pencil beam simulations, the detector limb-view path is assumed to intersect the centre of the cloud top, i.e. the point where the cloud top intersects the z-axis in figure 4.1.

The assumption of a spatially contiguous cloud is a necessary assumption as a disconnected

cloud domain would add a further degree of complexity to the coupling of the clear sky and scattering calculations, as well as having the implication that there could be an unquantifiable number of permutations of broken cloud field in the measurement inversion. This assumption is often valid, particularly for tropical and subvisible cirrus that has been observed in very thin unbroken layers which have a horizontal extent of the order of hundreds of kilometres (Dowling and Radke, 1990; Heymsfield, 1986; Lynch and Sassen, 2002).

The assumption of cloud as a uniform layer is often made in cloud retrieval schemes from remote measurements. There is also much observational evidence to support such an approximation for cirrus, e.g. Heymsfield (1986), Winker and Trepte (1998).

The finite extent of the cloud is chosen not to be included in the cloud state vector because limb-viewing instruments have a large uncertainty in cloud horizontal coverage along atmospheric paths (Liao et al., 1995 (1)). Therefore, as the measurement inversion is unlikely to be able to retrieve the horizontal extent, it is instead included in the atmosphere state vector, $\hat{\mathbf{b}}$, and defined as the arc length of the cloud top derived from ψ and R_{Earth}.

In terms of equation 4.1 the cloud state vector is given by $\hat{\mathbf{x}}_\mathbf{c} = [r_e, N, C_{\text{top}}, C_{\text{depth}}]$ for a given crystal habit.

4.2.3 Atmospheric Model

A one-dimensional spherically symmetric atmosphere is assumed in the model, with an altitude grid resolution of 1 km. It is necessary to assume that the atmospheric properties are well known so that the line-by-line radiative transfer calculations of gaseous absorption and emission give accurate results. The temperature and pressure profiles can be taken from climatology data, ECMWF data or from retrieved profiles. Gas concentration profiles are taken from climatology, however a chemical transport model may give more representative information on trace gases. The source of atmospheric data used depends on the use of the model and the accuracy of the model output required.

In equation 4.1 the atmosphere state vector, $\hat{\mathbf{b}}$, contains the profiles of temperature, pressure and all relevant gases. In addition, the tangent height, t_h, of the limb viewing ray path (assumed to be well known) is also included in $\hat{\mathbf{b}}$ as is the temperature and emissivity of the Earth's surface, the radius of the Earth, the spectral range to be modelled, $[\nu_0, \nu_n]$ and the spectral resolution, $\nu_{\text{res}} = (\nu_n - \nu_0)/(n - 1)$ where n is the number of spectral points to model (i.e. the length of the synthesised measurement vector).

4.3 Reverse Monte Carlo Scattering Model

The Monte Carlo method has long been employed to solve the difficult problem of physically realistic light transport in a cloudy atmosphere (Plass and Kattawar, 1968; Weinman and

Davies, 1978; Davis et al., 1979; Harshvardhan et al., 1981).

Limb sounding only requires a very limited subset of outward propagation paths. Thus, the reverse Monte Carlo method is chosen because it allows all computational effort to be concentrated on calculating radiances for the desired line-of-sight. Furthermore, the reverse Monte Carlo model is chosen over alternative plane parallel methods, for example, discrete ordinates methods (e.g. Stamnes et al., 1988), as these are far more onerous to solve for the limb viewing geometry and have a computationally expensive requirement to calculate the whole radiation field. The reversed Monte Carlo method has been shown to be adept at dealing with the spherical atmosphere geometry necessary for limb sounding (Collins et al., 1972) and has been used recently for modelling limb measurements of backscattered solar radiance, without specific account for clouds (Oikarinen et al., 1999), and for modelling microwave limb sounding measurements of cloud (Liu et al., 1996; Davis et al., 2004).

In McClouds_FM, the radiative transfer process in the cloud is considered as the stationary Markov chain whose states are photon interactions with cloud particles. Each interaction can have one of two outcomes: absorption or scattering of the photon, i.e. the photon path terminates or the photon leaves the interaction travelling in a new direction.

When simulating a pencil beam (in the backward direction from the detector), every photon in the beam enters the cloud at the same angle as given by the viewing geometry (which is defined by the tangent height, t_h, of the viewing ray path). Upon entry, each photon is traced through the cloud by determining the distance between interactions and the direction after scattering by generating random numbers[1] to sample the appropriate probability distributions until the photon is either absorbed or exits the cloud.

The Monte Carlo algorithm used is given in appendix B, together with a description of the appropriate probability distributions which are determined by the single scattering properties.

As the single scattering properties are dependent on the wavelength of infrared radiance the reverse scattering model holds for monochromatic radiation. If a high spectral resolution is required, then it is not practical for the scattering calculations to be done monochromatically at each spectral point. However, if the wavelength range to be simulated is relatively narrow (i.e. $< 0.25~\mu$m) then the scattering properties can be assumed constant across this range. Alternatively, if the range is wider, then the model is run at sampled points for which the single scattering properties have been calculated and then interpolating in the scattering output domain yields finer spectral resolution for McCloudS_FM calculations (see section 4.5). At present, the single scattering properties are sampled at 0.25 μm, as the complex refractive index of ice varies relatively smoothly at this resolution.

---

[1]Currently, McClouds_FM uses the *ran2* random number generator given in Press et al. (1992). The generator has a period length of approximately 2.3×10^{18} and is a combination of two Lehmer generators and a Bays-Durham shuffle to break up any serial correlations.

Horizontal Extent of Cloud (km)	Percentage of Lost Photons
10-20	1-3
20-50	0.2-1
50-80	0.05-0.2
80-110	0.01-0.05
> 110	< 0.01

Table 4.1: Typical percentage of photons which exit the cloud side after scattering in a backward simulation. Figures are for a range of typical cirrus cloud optical properties ($\tau \simeq 0.1$) and viewing geometry, with a cloud depth of 1 km. The number of lost photons is weakly dependent on viewing geometry and decreases with physical cloud depth.

Each pencil beam started at the detector uses a large number of photons ($\gtrsim 10^6$), so that there is good statistical convergence and the model is still adequately fast. However, it is too demanding on computer memory and processing to store the exact path for each photon and calculate the radiance along it. Therefore, McClouds_FM makes a statistical estimate of the radiance from the photons scattered into the line-of-sight by focussing on a limited number of properties of the backward photon trajectories. First of interest is if the photon is non-divergent, i.e. neither scattered nor absorbed in the cloud domain. Secondly, if the photon is divergent, then it is determined if it exits the cloud base, the cloud top, or if it is absorbed in the cloud. Divergent photons escaping out of the cloud side are assumed lost as they are too complicated to account for in the model: the total number of photons is re-normalised to account for this. If the horizontal extent is large (see section 4.7.1) and the vertical extent is relatively thin, it is found that very few photons exit the side of the cloud. An idea of the size of the error incurred from discarding the lost photons is given in Table 4.1. It can be seen that the induced error will be negligible for clouds with horizontal extent greater than 50 km. For photons which exit the cloud base or top, the exit angle with respect to the local horizontal (of the exit surface) is calculated. Only the polar angle is of interest as a spherically symmetric atmosphere is assumed and the scattering is azimuthally isotropic. Finally, the terminal point of photons absorbed by the cloud are stored.

To reverse the simulated trajectories, the calculated photon exit angles out of the cloud top and bottom can be considered as the *entry* angles of photons entering the cloud and then being scattered into the line-of-sight, and the terminal location of photons absorbed inside the cloud can be considered the initial location of photons emitted from the cloud. Thus it can be seen that in the reverse of the simulation, there are four possible sources of radiation entering the (pencil-beam) line-of-sight from the cloud arriving at the detector: non-scattered from the direct limb path, scattered into the line-of-sight from below the cloud, scattered into the line-of-sight from above the cloud and emitted from within the
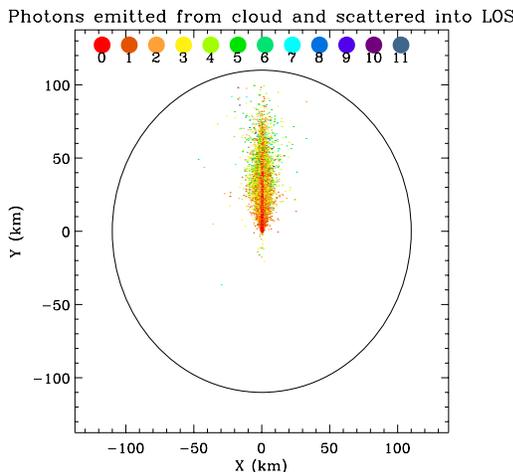
Figure 4.2: The horizontal location of photons emitted from the cloud and scattered into the line-of-sight. The colour of each point indicates the number of scattering interactions each photon undergoes before entering the line of sight. Approximately 30 % of emitted photons are scattered more than once. The satellite is in the negative y direction, with $x = 0$ km.

cloud and scattered into the line-of-sight.

Detailed information on the photon trajectories can be used to investigate the spread of photon source locations. As an example, scattering calculations were carried out for a cloud composed of hexagonal crystals with $\hat{\mathbf{x}}_{\mathbf{c}} = [9.2 \ \mu\text{m}, 0.5 \ \text{cm}^{-3}, 17.5 \ \text{km}, 0.75 \ \text{km}]$ and for a viewing tangent height of 16.4 km and incident radiation of wavelength 10.75 μm ($\sim 930 \ \text{cm}^{-1}$). Figure 4.2 shows the locations of photons emitted from the cloud and the degree of scattering each undergoes before leaving the cloud domain. For this particular cloud no photons escape from the cloud sides in the backward simulation, which can be seen from the emission points being located near the centre of the cloud. There is also a substantial degree of multiply scattered photons entering the limb path. Figure 4.3 shows the equivalent vertical distribution of these photons and thus that the emission of photons from the cloud (that reach the limb) decays rapidly with height inside the cloud.

4.4 Radiative Transfer in the Clear Sky Domain

The radiative transfer calculations used in the clear sky domain in McClouds_FM are computed by the Reference Forward Model (RFM). The RFM is a general line-by-line radiative transfer model based on GENLN2 (Edwards, 1992) developed[2] by Dr Anu Dudhia at the University of Oxford (Dudhia, 2004).

The RFM performs the radiative transfer calculations (and ray tracing, inclusive of refraction if required) with user-specified measurement geometry, atmospheric model and

[2]Further details on the physical approximations used by the RFM can be found in Glatthor et al. (1999).
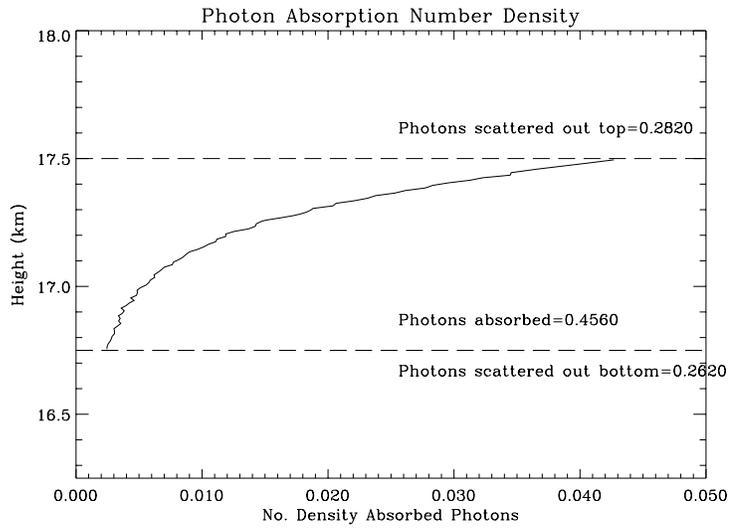
Figure 4.3: The vertical distribution of photons absorbed inside the cloud in the backward simulation. This is equal to the vertical distribution of photons emitted from the cloud and scattered into the line-of-sight in the reverse simulation.

spectroscopic data.

The radiance of photons arriving at the cloud bottom and cloud top is calculated by the RFM. This is done by the option in the RFM which allows the radiance arriving at an observer position in the atmosphere to be calculated for a given 'elevation' viewing angle i.e. the angle with the local horizontal at that altitude. This radiance calculated by the RFM is denoted, R_{obs}, and it can be described as a function of observer height (which will be C_{top} or C_{bot}), elevation angle, and wavenumber.

The RFM is also used to calculate the transmission and emission from the exit point from the cloud to the satellite, denoted $T_{\text{cld-sat}}(C_{\text{top}}, \phi_{\text{exit}}, \bar{\nu})$ and $R_{\text{cld-sat}}(C_{\text{top}}, \phi_{\text{exit}}, \bar{\nu})$ respectively. Where the exit angle, ϕ_{exit}, is the angle of incidence (measured with the respect to the local horizontal) of the viewing ray path with the cloud top.

As line-by-line calculations are being used to simulate measurement spectra, the convention is to define infrared radiation in terms of wavenumber, $\bar{\nu} = \frac{1}{\lambda}$, as opposed to wavelength, λ.

4.5   McCloudS_FM Calculations

For a given cloud in the model, i.e. with state parameter $\hat{\mathbf{x}}_{\mathbf{c}} = [r_e, N, C_{\text{top}}, C_{\text{depth}}]$, the formulae used to combine the RFM calculations with statistical measures of the scattering output are given by:

1. Upwelling radiance from below the cloud redirected into the field-of-view;

$$R_{\text{bot}}(\bar{\nu}) = \sum_{i=1}^{N_\phi} \Phi_{\text{bot}}(i, \bar{\nu}) R_{\text{obs}}(C_{\text{bot}}, \phi_i - \frac{1}{2}\phi_{\text{bin}}, \bar{\nu}) \tag{4.6}$$

where $\Phi_{\text{bot}}(i, \bar{\nu})$ is the angular distribution of photons entering the cloud from its bottom and being scattered into the line-of-sight. $\Phi_{\text{bot}}(i, \bar{\nu})$ gives the fraction of photons entering from the bottom of the cloud which have entry angle ϕ in angle bin i i.e. $\phi \in [\phi_{i-1}, \phi_i)$. $N_\phi = 90°/\phi_{\text{bin}}$ is the number of entry angle bins (ϕ_{bin} being the entry angle bin-size). $\Phi_{\text{bot}}(i, \bar{\nu})$ is determined from the Monte Carlo scattering calculations and can be thought of as the discrete pdf of bottom entry angle.

2. The downwelling radiance from above the cloud redirected into the field-of-view, R_{top} can be similarly defined

$$R_{\text{top}}(\bar{\nu}) = \sum_{i=1}^{N_\phi} \Phi_{\text{top}}(i, \bar{\nu}) R_{\text{obs}}(C_{\text{top}}, \phi_i - \frac{1}{2}\phi_{\text{bin}}, \bar{\nu}) \tag{4.7}$$

where $\Phi_{\text{top}}(i, \bar{\nu})$ is the discrete pdf of top entry angle.

3. Emission from within the cloud directed into the field-of-view;

$$R_{\text{cloud}}(\bar{\nu}) = \sum_{i=1}^{N_h} E_h(i, \bar{\nu}) B[T(h_i - \frac{1}{2}h_{\text{bin}}), \bar{\nu}] \tag{4.8}$$

where $E_h(i, \bar{\nu})$ is the emission density as a function of height. $E_h(i)$ gives the fraction of emitted photons which are emitted from a height h in height bin i i.e. $h \in [h_{i-1}, h_i)$. $N_h = C_{\text{depth}}/h_{\text{bin}}$ is the number of emission height bins (h_{bin} being the entry angle bin-size). $E_h(i, \bar{\nu})$ and can be thought of as the discrete pdf of the source height of an emitted photon. $B[T, \bar{\nu}]$ is Planck's blackbody function which is calculated for the temperature $T = T(h_i - \frac{1}{2}h_{\text{bin}})$ at the mid-point of each height bin, derived from the temperature profile for the atmosphere. Planck's function is given by

$$B[T, \bar{\nu}] = \frac{c_1 \bar{\nu}^3}{\exp[c_2 \bar{\nu}/T] - 1} \tag{4.9}$$

in units of nW cm^{-1} sr^{-1}, where $c_1 = 1.1911 \times 10^{-3}$ nW cm^2 sr^{-1} and $c_2 = 1.439$ K cm for $\bar{\nu}$ in units of cm^{-1} and T in Kelvin, K.

4. Radiance which will exit from the cloud in a pencil beam in the direction of satellite;

$$R_{\text{exit}}(\bar{\nu}) = f_{\text{bot}}(\bar{\nu}) R_{\text{bot}}(\bar{\nu}) + f_{\text{top}}(\bar{\nu}) R_{\text{top}}(\bar{\nu}) + f_{\text{cloud}}(\bar{\nu}) R_{\text{cloud}}(\bar{\nu}) \tag{4.10}$$

where $f_{\text{top}}(\bar{\nu})$ ($f_{\text{bot}}(\bar{\nu})$) is the fraction of photons entering the cloud from its top (bottom) and scattered into the line-of-sight, and f_{cloud} is the fraction of photons emitted from the cloud entering the line-of-sight. Note that $f_{\text{top}}(\bar{\nu}) + f_{\text{bot}}(\bar{\nu}) + f_{\text{cloud}}(\bar{\nu}) = 1$.

5. Finally the radiance at the satellite is calculated;

$$R_{\text{sat}}(\bar{\nu}) = f_0(\bar{\nu})R_{\text{limb}}(t_h) + (1 - f_0(\bar{\nu}))\left[R_{\text{exit}}(\bar{\nu})T_{\text{cld}-\text{sat}}(C_{\text{top}}, \phi_{\text{exit}}, \bar{\nu}) + R_{\text{cld}-\text{sat}}(C_{\text{top}}, \phi_{\text{exit}}, \bar{\nu})\right]$$
(4.11)

where $R_{\text{limb}}(t_h)$ is the radiance from emission along the limb viewing ray path with tangent height, t_h and calculated by the RFM, and $f_0(\bar{\nu})$ is the fraction of total photons in the pencil beam arriving at the satellite that are non-divergent or zeroth order, i.e. they are not scattered or absorbed in the cloud domain and pass directly through.

The main advantage of this method is that we retain the information on where the photons are coming from. We can see the proportion of the radiance arriving at the satellite that is from the cloud (emitting as a grey-body), and the proportion that has been scattered into the line-of-sight from above and below the cloud.

In terms of formulating McCloudS_FM as a function as in equation 4.1, the synthesised radiance measurement vector can be written as $\hat{\mathbf{y}} = \{\hat{y}(\bar{\nu}_0), \hat{y}(\bar{\nu}_1), \ldots, \hat{y}(\bar{\nu}_n)\}$ where $[\bar{\nu}_0, \bar{\nu}_n]$ is the spectral range and $\bar{\nu}_i = \bar{\nu}_0 + i\bar{\nu}_{\text{res}}$ for $i = 0, 1, \ldots, n$, where $\bar{\nu}_{\text{res}}$ is the spectral resolution. The synthesised measurement radiance, is given by

$$\hat{y}(\bar{\nu}_i) = R_{\text{sat}}(\bar{\nu}_i).$$
(4.12)

As stated in section 4.3, if the spectral range required is relatively large then interpolation in the scattering output domain is required. If the single scattering properties for a given r_e and N have been calculated for a set of sampled wavenumbers i.e. $\{\bar{\nu}_{s_j} \in [\bar{\nu}_{s_0}, \bar{\nu}_{s_m}] : \quad \bar{\nu}_{s_j} = \bar{\nu}_{s_0} + j\bar{\nu}_{s_{\text{res}}}; \quad \frac{1}{\bar{\nu}_{s_{\text{res}}}} = 0.25\,\mu\text{m}; \quad j = 0, 1, \ldots, m\}$ such that $[\bar{\nu}_0, \bar{\nu}_n] \subseteq [\bar{\nu}_{s_0}, \bar{\nu}_{s_m}]$, then the scattering calculations yield the discrete functions of two variables $\Phi_{\text{bot}}(i, \bar{\nu}_{s_j})$, $\Phi_{\text{top}}(i, \bar{\nu}_{s_j})$ and $E_h(i, \bar{\nu}_{s_j})$. These functions are then interpolated from the grid of $\bar{\nu}_{s_j}$'s at resolution $\bar{\nu}_{s_{\text{res}}}$ onto the grid of $\bar{\nu}_i$'s at resolution $\bar{\nu}_{\text{res}}$ to yield the discrete distributions $\Phi_{\text{bot}}(i, \bar{\nu}_j)$, $\Phi_{\text{top}}(i, \bar{\nu}_j)$, and $E_h(i, \bar{\nu}_j)$. These functions are then used in the empirical formulae (equations 4.6, 4.7, 4.8, 4.10 and 4.11) outlined above to obtain the synthesised measurement vector.

4.5.1 Gaseous Absorption and Emission Within the Cloud Domain

Due to computing limitations it is not practical to calculate the gaseous absorption and emission for each photon path inside the cloud domain. Thus, in order to include this in the model calculations several approximations must be made and this will result in additional forward model error.

McClouds_FM is currently configured to account for gaseous absorption and emission inside the cloud domain in one of three ways. The first being not at all and the McClouds_FM calculations are as set out previously. The second is to include only gaseous absorption and emission for photons entering the line-of-sight from below the cloud by making a crude approximation (for the purposes of the line-by-line calculations) of an infinitesimally thin cloud (often used in nadir sounding forward models), so C_{top} replaces C_{bot} in equation 4.6 and the other equations remain unaltered.

The third method to account for gaseous absorption and emission inside the cloud domain uses the information from the scattering calculations on the total path length of each photon trajectory inside the cloud domain. From the atmospheric profiles of temperature, pressure and gas concentration, and using the RFM and the approximation that the temperature, pressure and gas concentrations are constant within the cloud domain it is possible to calculate the amount of absorption and emission for a path of a given length in the domain. However, it is still too computationally demanding to do this for every photon trajectory, so a statistical measure of path length is used. Photons scattered into the line-of-sight from below or above the cloud are binned by entry angle, and for each set of photons from a given entry angle bin the mean path length inside the cloud domain is calculated, denoted $L_{\phi_{\text{bot}}}(i, \bar{\nu})$ or $L_{\phi_{\text{top}}}(i, \bar{\nu})$. Likewise, photons emitted from the cloud and scattered into the line-of-sight are binned by emission height and for each height bin the mean path length inside the cloud domain of photons with that emission height is calculated as is the standard deviation, denoted $L_h(i, \bar{\nu})$. From the path length information the gaseous transmission and emission inside the cloud domain are calculated. This is then included in the McCloudS_FM calculations by changing equation 4.6 to

$$R_{\text{bot}}(\bar{\nu}) = \Phi_{\text{bot}}(i, \bar{\nu}) R_\phi^{L_{\text{bot}}}(i, \bar{\nu}) + \sum_{i=1}^{N_\phi} \Phi_{\text{bot}}(i, \bar{\nu}) T_\phi^{L_{\text{bot}}}(i, \bar{\nu}) R_{\text{obs}}\left(C_{\text{bot}}, \phi_i - \frac{1}{2}\phi_{\text{bin}}, \bar{\nu}\right) \quad (4.13)$$

where $R_\phi^{L_{\text{bot}}}(i, \bar{\nu})$ is the radiance of gaseous emission for the mean path length, $L_{\phi_{\text{bot}}}(i, \bar{\nu})$ of the photons which have entry angle ϕ in angle bin i i.e. $\phi \in [\phi_{i-1}, \phi_i)$ and $T_\phi^{L_{\text{bot}}}(i, \bar{\nu})$ is the gaseous transmission along the mean path length given by $L_{\phi_{\text{bot}}}(i, \bar{\nu})$. $R_{\text{top}}(\bar{\nu})$ can be similarly defined by changing bot to top functions in equation 4.13. Also, equation 4.8 must

be changed to

$$R_{\text{cloud}}(\bar{\nu}) = E_h(i,\bar{\nu})R_h^L(i,\bar{\nu}) + \sum_{i=1}^{N_h} T_h^L(i,\bar{\nu})E_h(i,\bar{\nu})B[T(h_i - \frac{1}{2}h_{\text{bin}}),\bar{\nu}] \qquad (4.14)$$

where $R_h^L(i,\bar{\nu})$ is the radiance of gaseous emission for the mean path length, $L_h(i,\bar{\nu})$ of the photons which have emission height h in height bin i i.e. $h \in [h_{i-1}, h_i)$.

It is actually found that there is only a very small difference (usually $\sim 0.5\%$ in magnitude) between the radiance at the satellite calculated from the second and third methods of including gaseous emission and absorption in the cloud domain. As the second method requires far less computation time, the error margin is seen as acceptable to use method two to save on processing. However, it is more difficult to quantify the error incurred on line shape from the second method.

4.6 Forward Model Errors

It is important to quantify the accuracy of the forward model. However, when analysing the error budget it is important to distinguish between the error in quantifying the quality of the simulation given the assumptions about the cloud state, and the error incurred by the assumptions about the cloud state. Both of these error sources are then in turn independent of the error due to the sensitivity of the model to the atmosphere state parameters, i.e. the forward model parameter error (see section 6.4.2). The model parameter error gives an uncertainty on the modelled radiance due to the uncertainty on the atmospheric state parameters.

This section will only consider the errors incurred by the forward model approximations to the physics of the radiative transfer for a cloud type as given in the model. The errors due to the assumptions (section 4.2) about cloud type are essentially a problem for the measurement inversion and the accuracy of retrieved properties, which will not be considered here.

Table 4.2 lists the sources of errors in the model and the corresponding error in the simulated radiance. The difficulty in expressing the errors in this way is that the absolute value of the error is dependent on the both the input cloud state vector and the atmosphere state vector, as well as varying spectrally. The error values given in table 4.2 are root-mean-square percentage errors in the at satellite radiance for 850–970 cm^{-1} and averaged over a range of cloud types (with a typical cirrus vertical optical depth of $\tau \simeq 0.1$), over a vertical profile of viewing geometries and using a tropical climatology for the atmospheric parameters.

Table 4.2 includes a simple photon counting error which is included to represent the Poisson

Error Source	Average Percentage Error
Poisson statistics	0.3 %
Discrete angle bins - $\phi_{bin} = 1°$	0.1 %
Discrete height bins - $h_{bin} = 10$ m	0.05 %
Gaseous absorption in cloud (method 1)	0.7 %
Gaseous absorption in cloud (method 2)	0.35 %
Gaseous absorption in cloud (method 3)	0.2 %

Table 4.2: Forward model errors expressed as an average of the root mean square percentage error. The error values are obtained by averaging over an ensemble of simulations for a range of cloud types and by averaging spectrally.

statistics uncertainty in the number of photons per angle or height bin. The value given is for $\phi_{bin} = 1°$ and $h_{bin} = 10$ m. Table 4.2 also includes the error in the simulated radiance due to the use of discrete angle and height bins, the error value is for $\phi_{bin} = 1°$ and $h_{bin} = 10$ m, and is obtained by comparing against simulations with $\phi_{bin} = 0.001°$ and $h_{bin} = 0.01$ m. The error incurred is negligible. Finally, table 4.2 gives the error incurred from using the various methods to calculate the gaseous absorption for the photon paths inside the cloud domain (see section 4.5.1). The error values given are for the various methods applied with $\phi_{bin} = 1°$ and $h_{bin} = 10$ m from comparison to calculations applied with $\phi_{bin} = 0.001°$ and $h_{bin} = 0.01$ m using method 3, so that there are very few photons per bin and thus the average path length per bin is more representative than for the larger bin size. The root-mean-square error for gaseous absorption is seen to be very small for each of the three methods, with the error mainly centred on the main gas lines, i.e. H_2O and CO_2.

4.7 Considerations for Simulating MIPAS Measurements

Cirrus measurements made by limb observers are complicated by the non-continuous areal nature of clouds, and additional problems are manifested in the instrument field-of-view (FOV), the limb viewing and orbital geometry and the assumption of a one-dimensional (spherically symmetric) atmosphere in the measurement inversion. In this section, a discussion is given of the pertinent problems arising from the MIPAS measurements in order to develop the method to take the observing instrument into account in the forward model.

4.7.1 Effective Horizontal Cloud Extent

Limb sounders may be able to detect cloud somewhere along the very long atmospheric path length, but they cannot determine how much of the path length is occupied by the

cloud.

If a spherical atmosphere is assumed, composed for example of spherical atmospheric shells, say 1 km thick, then the layer containing the tangent point of the limb path contributes the most information to the measurement and (around the mid-troposphere) has a volume sample length of approximately 200 km. In limb occultation detection the measurement is of cloud extinction along the ray path, so a common method of dealing with uncertainty in horizontal coverage is to assume that the cloud field is restricted to the length of the tangent layer volume sample length, or to assume an effective horizontal extent, for example 75 km (Liao et al., 1995 (2)). For infrared limb emission to incorporate limb scattering a similar approximation of horizontal coverage can be made to account for the effective scattering volume.

It can be seen from table 4.1 that all clouds with horizontal extent greater than 110 km 'look' the same to a pencil beam incident on the centre of the cloud top, in that a negligible number of photons is scattered out of the sides of the cloud, hence the scattering output is not influenced by the horizontal extent. Therefore, by setting the horizontal extent as 200 km, in line with the volume sample length in the tangent layer, it has the effect of modelling the scattering from a pencil beam for all cloud with horizontal extent greater than 110 km. Effectively this models optically thin cloud without distinct (sharp) side boundaries (horizontal edges), which is physically more realistic. This model limit on cloud extent has several implications if applied to the physics of the limb measurement.

4.7.2   The Instrument Field-of-View

MIPAS has a nominal vertical FOV of 3 km at the tangent point. At a tangent height of 10 km, for example, the FOV 'Footprint' on a cloud layer at a height of 12 km would be approximately 90 km (see figure 4.4).
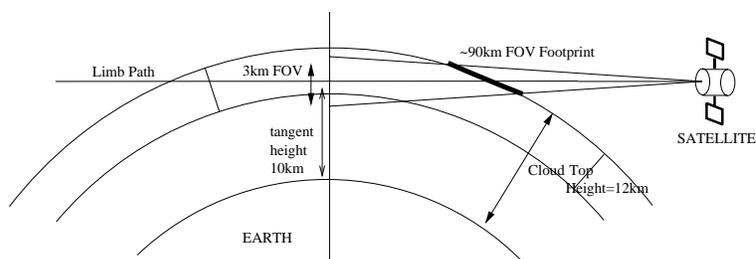
Figure 4.4: FOV 'Footprint'.

When cloud is present in the limb view it is very difficult to determine how much of the FOV 'Footprint' is filled by the cloud. Thus, it further complicates the model if FOV effects are to be included. If the simulated radiance of a pencil beam at the satellite is a function of wavenumber and tangent height and is given by $\hat{y}(\bar{\nu}, t_h)$, then the radiance at

the satellite taking the FOV into account can be given by

$$\hat{y}_{\text{FOV}}(\bar{\nu}, t_h^c) = \frac{1}{N+1} \sum_{i=-N/2}^{N/2} W_{\text{FOV}}(i)\, \hat{y}(\bar{\nu}, t_h^c + i\delta_h) \qquad (4.15)$$

for $N+1$ beams in the FOV (N being even), where t_h^c is the tangent height at the centre of the FOV, $W_{\text{FOV}}(i)$ is the normalised FOV response weighting function and $\delta_h = \frac{h_{\text{FOV}}}{N}$, where h_{FOV} is the nominal vertical extent of the FOV at the tangent point. The simplest case is when the entire FOV 'Footprint' is cloud filled (as in Figure 4.4). This will occur if the horizontal extent of the cloud is of the order of or greater than the 'Footprint' and the viewing geometry is correct. In this case the instrument measurement can be approximated by a pencil beam in the centre of the FOV simulated by McClouds_FM, or modelled more accurately by using equation 4.15 where each pencil beam in the FOV is simulated by McClouds_FM.

Problems begin to be encountered when the FOV 'Footprint' is not filled. This can occur in three ways: if the cloud horizontal extent is less than the FOV Footprint, or if the cloud is of large horizontal extent but the viewing tangent height is too high or too low for the FOV 'Footprint' to be entirely cloud filled. When the cloud is of large horizontal extent but the viewing geometry is incorrect for the FOV 'Footprint' to be entirely cloud filled, the FOV effect can be accounted for by using clear sky radiative transfer for the pencil beams in the FOV which do not emerge from the cloud. There is a large uncertainty when modelling cases in which the cloud extent is less than the FOV 'Footprint' or the viewing tangent height is too low due to lost photons. As a result of these errors, a measurement inversion is likely to be insensitive to cloud position within the FOV (or more specifically to have multiple solutions with high error). Therefore, when simulating MIPAS data, the case where the FOV 'Footprint' is not entirely filled is only modelled when the viewing tangent height is too high. Furthermore, the effective horizontal extent of the cloud is assumed to be greater than the Footprint and of the order of 200 km, in line with the limit of accuracy mentioned in section 4.7.1. Therefore, this work is focussed on modelling limb observations of cirrus of large horizontal extent. And any comparisons to MIPAS data are on the assumption that the observed cirrus is of large horizontal extent.

This assumption should not be considered a drawback, as cirrus is usually much greater in horizontal extent than vertical extent. For example, the Lidar In Space Technology Experiment (LITE) (Winker and Trepte, 1998) detected laminar cirrus with a horizontal extent of $<$ 100 km up to 2700 km and thickness of less than 200 m in an altitude region of 14 to 18.7 km. Orographic and convective cirrus have frequently been observed with horizontal extent on the order of hundreds of kilometres, (e.g. Kahn et al., 2003), and jet stream cirrus can cross continents nearly unbroken with a vertical depth of only a few

kilometres.

Cirrus clouds can be considered contiguous on average, and the horizontal extent of an average cirrus cloud mass is greater than about 20 or 30 kilometres in at least one dimension and can range smaller or larger than this by two orders of magnitude (Dowling and Radke, 1990).

The assumption that any *observable* cirrus is of a large horizontal extent also resolves the uncertainty of the cloud limb position. If the cloud top height is higher than tangent height then it can be assumed that the limb observation is 'seeing' the cloud top, i.e. the ray path intersects with the cloud top.

4.8   Model Comparisons to Data

In order to validate the McClouds_FM simulated satellite measurements the model output was compared to real MIPAS measurements of cirrus. The atmospheric window region was of main interest as the transparency of the atmosphere allows the photons from the scattered ray paths to originate from lower in the atmosphere, i.e. from warmer temperatures, and thus the evidence of scattering should be more apparent. In this study attention was particularly focussed on modelling the absorbing H_2O lines due to scattering, in the region 940-950 cm^{-1} (following Höpfner et al., 2002; Spang et al., 2004).

4.8.1   MIPAS data

The MIPAS data used in this case study is from orbit 504 from the 16th of May 2002.

For cirrus cloud observations, attention was focused on elevation profiles from tropical regions where it is well known that high cirrus frequently occurs, as does subvisible cirrus near the tropopause (Woodbury and McCormick, 1986; Wang et al., 1996; Spang et al., 2002). The Principle Component Cloud Test introduced in section 3.5 was used to determine the presence of cloud.

The elevation profile selected for use in this case study has a mean tangent point location at a Latitude of 1.25 °N and a Longitude of 36.5 °E . The measurement spectra of two elevation scans from (Level 1B) tangent heights of 13.54 km and 16.42 km can be seen in figures 4.5 and 4.6 respectively. The inverted spectral lines due to scattering of tropospheric radiation into the FOV can be clearly seen at ~ 948.25 cm^{-1}.

The scattering features under investigation, i.e. the H_2O absorption lines and the inverted CO_2 side-lobes, are spectrally very narrow and as MIPAS is a Fourier transform spectrometer it is important to be sure that any such features in the measured spectra cannot be due to instrument line shape (ILS) artefacts. Therefore, only the *apodised* MIPAS spectra are
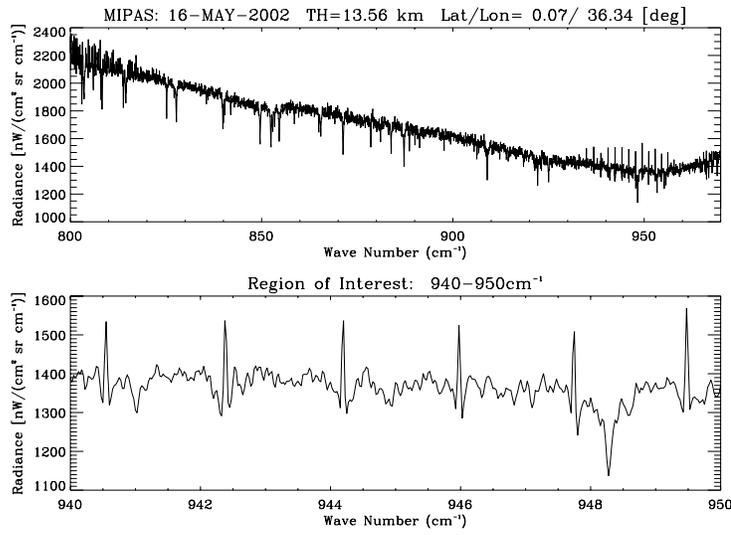
Figure 4.5: MIPAS Spectra from Orbit 504 Elevation Scan at Tangent Height 13.54 km.
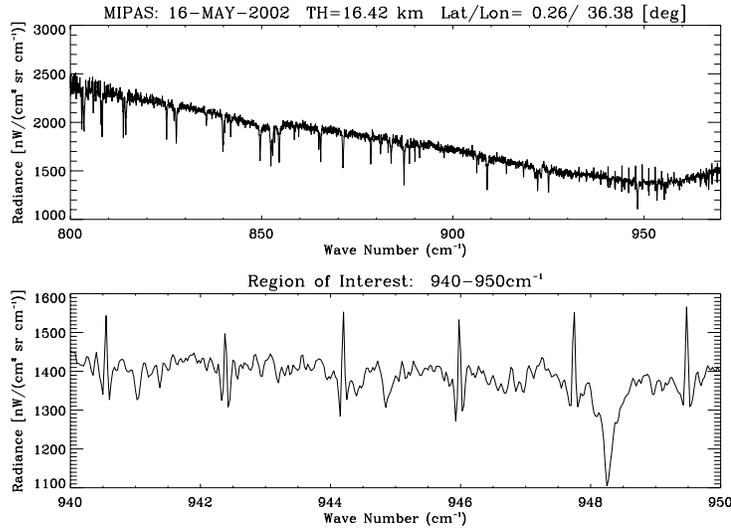
Figure 4.6: MIPAS Spectra from Orbit 504 Elevation Scan at Tangent Height 16.42 km.

used for the comparisons to the model output. The MIPAS apodised spectra are produced using an ILS with the Norton-Beer 'strong' apodisation applied (Norton and Beer, 1976).

4.8.2 Initial Comparisons to Data

The measurement spectra from the elevation scans at 13.56 km and 16.42 km are denoted by the vectors $\mathbf{y}_{13}$ and $\mathbf{y}_{16}$ and are composed of $N_{\bar{\nu}} = ((950 - 940)/0.025) + 1$ spectral points. For each measurement vector in turn, a cloud state ($\hat{\mathbf{x}}_{\mathbf{c}} = [r_e, N, C_{\text{top}}, C_{\text{depth}}]$) was estimated (assuming hexagonal column crystal habit) and McClouds_FM was used to simulate the spectra (with a 3-beam FOV) obtaining $\hat{\mathbf{y}}_{13}$ and $\hat{\mathbf{y}}_{16}$ respectively. After each simulated spectra was generated, the goodness-of-fit to the measurement was tested and

if the fit did not meet a certain criteria the cloud state was altered systematically and the spectra simulated again. This process was iterated (over a finite set of cloud property permutations) until a "best-fit" spectra was obtained.

The goodness-of-fit was measured by the quantity

$$\chi^2 = \sum_{i=1}^{N_{\bar{\nu}}} \frac{|y(\bar{\nu}_i) - \hat{y}(\bar{\nu}_i)|^2}{\epsilon_y(\bar{\nu}_i)^2} \tag{4.16}$$

where $y(\bar{\nu}_i)$ is the MIPAS measured radiance at the spectral point $\bar{\nu}_i$, $\hat{y}(\bar{\nu}_i)$ is the modelled radiance value from McClouds_FM at the spectral point $\bar{\nu}_i$, and $\epsilon_y(\bar{\nu}_i)^2$ is the uncertainty in $y(\bar{\nu}_i)$ from the measurement error, i.e. the noise equivalent signal radiance which in Band A is 50 nW/(cm^2 sr cm^{-1}). As a general rule, for a given number of degrees of freedom, df, one wants the "reduced χ^2" (χ^2/df) to be approximately equal to one ($\chi^2 \sim df$). A reduced χ^2 that is much greater than one indicates a poor fit. In this case, if we assume that for the vectors $\mathbf{y}$ and $\hat{\mathbf{y}}$, each spectral point in the vector is independent from all other points in the vector, then the number of degrees of freedom is $df = N_{\bar{\nu}} - 1$. The "best-fit" spectra is such that it minimises the χ^2 value.

In order to compare the McClouds_FM output against the apodised MIPAS data, the simulated spectra must also be apodised by post-processing. This post-processing involves convolving the spectra simulated by McCloudS_FM at high resolution ($\bar{\nu}_{\text{res}} = 0.001$ cm^{-1}) for each of the elevation scans, with an ILS function[3] for MIPAS band A measurements, then down-sampled to the resolution of the measurement spectra.

The "best-fit" search process was carried out with the gas concentration profiles taken from reference atmosphere climatology (Remedios, 1999), and the temperature and pressure profiles obtained from ECMWF data. The viewing tangent heights of 13.56 km and 16.42 km are taken from the nominal MIPAS level 1B tangent heights derived from the MIPAS pointing information. They are assumed to be accurate here (see section 4.8.3 for further comment).

It was found that the best-fit spectra, were obtained from cloud state vectors of

$$\hat{\mathbf{x}}_{13} = [11.75 \,\mu\text{m}, \, 2.75 \,\text{cm}^{-3}, \, 18.75 \,\text{km}, \, 0.5 \,\text{km}]$$

and

$$\hat{\mathbf{x}}_{16} = [7.85 \,\mu\text{m}, \, 4.76 \,\text{cm}^{-3}, \, 18.75 \,\text{km}, \, 0.7 \,\text{km}]$$

respectively.

Figure 4.7(a) and 4.7(b) show the breakdown (for the central beam of the FOV in the $\hat{\mathbf{x}}_{13}$ and $\hat{\mathbf{x}}_{16}$ simulations respectively) of the radiance from the various sources which is scattered

---

[3]Obtained via private communication from Dr Chiara Piccolo, University of Oxford.
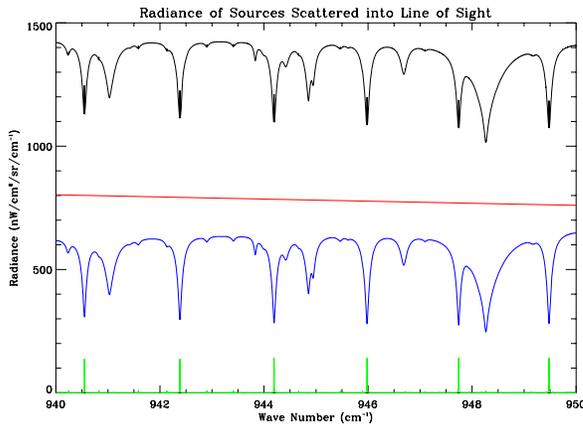
(a) Breakdown of McCloudS_FM simulated spectra for $\hat{\mathbf{x}}_{13}$.
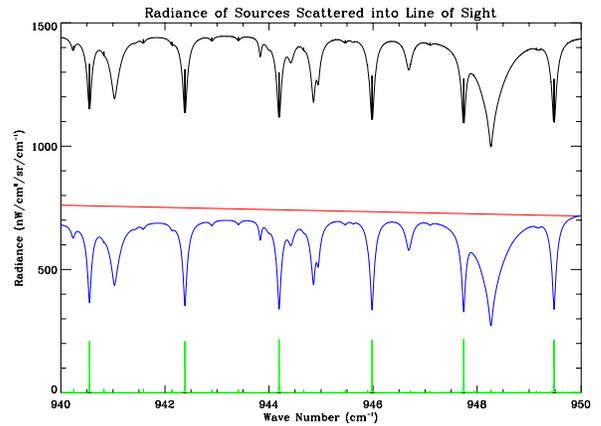
(b) Breakdown of McCloudS_FM simulated spectra for $\hat{\mathbf{x}}_{16}$.

Figure 4.7: McCloudS_FM simulated spectra using ECMWF temperatures. In blue is $f_{\mathrm{bot}}R_{\mathrm{bot}}$, the radiance scattered into the line-of-sight from photons from below the cloud. In green is $f_{\mathrm{top}}R_{\mathrm{top}}$, the radiance scattered into the line-of-sight from photons from above the cloud. In red is $f_{\mathrm{cloud}}R_{\mathrm{cloud}}$, the radiance of photons emitted from the cloud and scattered into the line-of-sight. In black is $R_{\mathrm{exit}} = f_{\mathrm{bot}}R_{\mathrm{bot}} + f_{\mathrm{top}}R_{\mathrm{top}} + f_{\mathrm{cloud}}R_{\mathrm{cloud}}$.

into the line-of-sight i.e. the quantities, $f_{\mathrm{bot}}R_{\mathrm{bot}}$, $f_{\mathrm{top}}R_{\mathrm{top}}$, $f_{\mathrm{cloud}}R_{\mathrm{cloud}}$ and R_{exit} as given in equations 4.6, 4.7, 4.8, and 4.10.

The simulated radiance arriving at the satellite, $\hat{\mathbf{y}}_{13}$ and $\hat{\mathbf{y}}_{16}$ is graphed in figures 4.8(a) and 4.8(b) respectively, with the MIPAS measurement overplot as well as the radiance obtained from the same simulation but with only single scattering. Both figures 4.8(a) and 4.8(b) illustrate clearly that without multiple scattering included in the simulation the radiance can be grossly underestimated by around 15–20 %. In addition, it is possible to see that the H_2O absorption features are recreated in the simulation and from figures 4.7(a) and 4.7(b) it can be seen that the only source of these absorption features is from radiance scattered into the line-of-sight from below the cloud, which accounts for approximately 40 % of the radiance exiting the cloud.

Finally, figures 4.9(a) and 4.9(b) give the percentage difference between the apodised radiance at the satellite and the apodised MIPAS measurement spectra. The root-mean-square percentage difference and root-mean-square noise percentage are both overplot on the figures. It can be clearly seen that the rms percentage difference is less than the rms measurement noise percentage. Therefore, it is possible to obtain convergence for the cloud state values from which McClouds_FM can successfully model MIPAS spectra within measurement noise. Furthermore, the percentage residual in figures 4.9(a) and 4.9(b) is less than the noise level in ~ 85 % of the simulated spectral window. The outlying points being centred around the H_2O absorption features, but do not exceed a difference of 13 % from
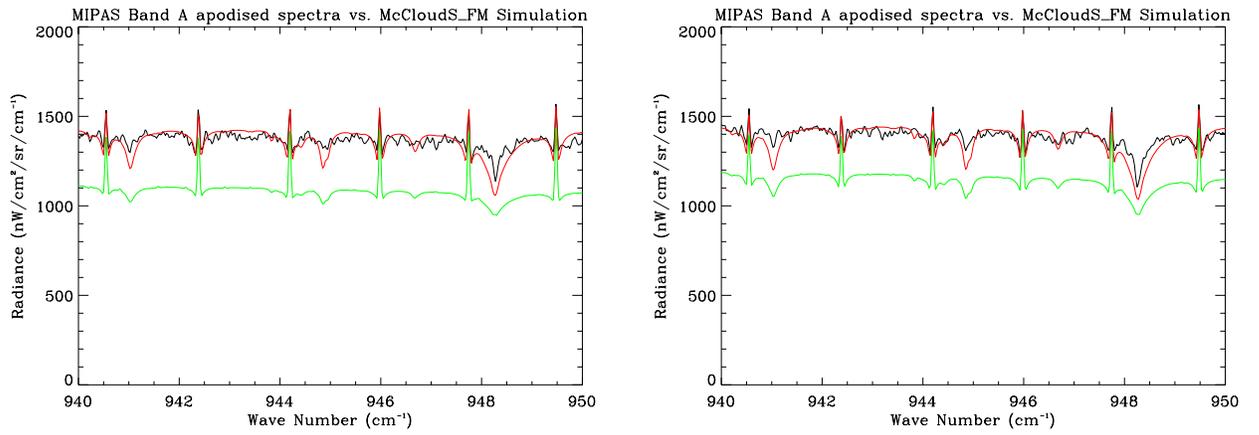
(a) Simulation for $\hat{\mathbf{x}}_{13}$, with $\mathbf{y}_{13}$ in black.

(b) Simulation for $\hat{\mathbf{x}}_{16}$, with $\mathbf{y}_{16}$ in black.

Figure 4.8: McCloudS_FM simulated apodised spectra using ECMWF temperatures: the red line is the multiple scattering case and single scattering case is the green line. The measurement spectra is overplot in black.

the measurement in either case.

4.8.3 Discussion

Interestingly, the cloud top was found to be slightly higher than the tropopause by approximately 0.5 km. This may be due to the uncertainty in the MIPAS pointing (i.e. the tangent height) or due to errors in the ECMWF temperature profile. Also, the cloud vector for the $\hat{\mathbf{y}}_{13}$ simulation has a smaller effective radius and larger number density than the cloud vector for the $\hat{\mathbf{y}}_{16}$ best-fit simulation. It is not unreasonable to expect that this is a physical difference. Due to the integration time of each tangent height measurement the horizontal distance between tangent points can be up to 30 km, therefore it is possible that the cloud is horizontally inhomogeneous on such a scale.

The "best-fit" spectra discussed above are only best-fit in the sense that they have the lowest χ^2 value for a set of spectra generated from a finite set of cloud property permutations. Moreover, it is not to say that there are no other permutations which could generate a simulated spectra which would have a smaller χ^2 and model the spectra more accurately than in the cases above. This is important to realise as it points the way to adopting an optimal estimation or 'retrieval' technique in order to use the inherent information in the forward model in a more efficient manner to iterate to the point in the infinite 4D cloud state space which would give the best-fit simulated spectra. This argument could also be extended to explain the difference between the $\hat{\mathbf{x}}_{13}$ and $\hat{\mathbf{x}}_{16}$ cloud vectors.

The radiance in the McClouds_FM simulations is too low around the H_2O absorption features compared to the MIPAS measurements. Thus, the absorption features are slightly
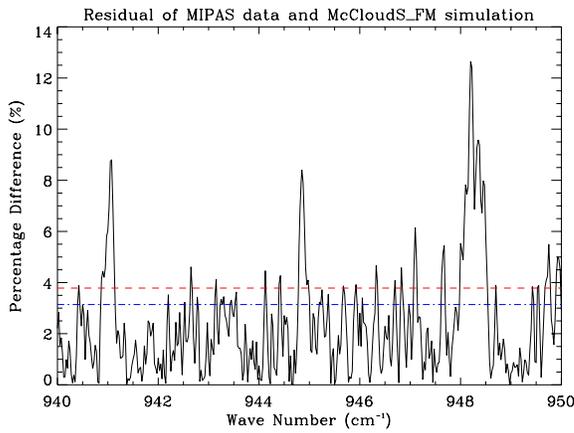
(a) The percentage difference between the simulated apodised radiance for $\hat{\mathbf{x}}_{13}$ and the 13.54 km MIPAS measurement.
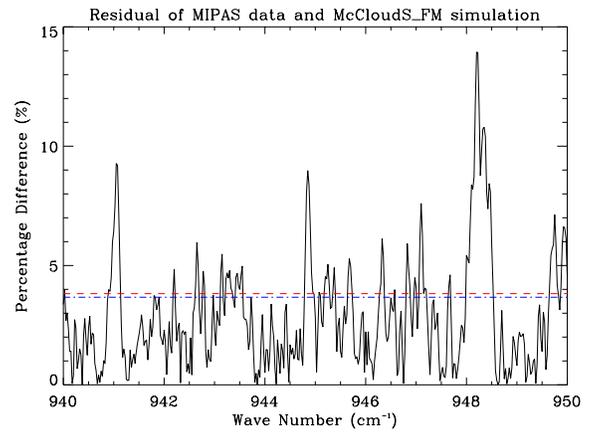
(b) The percentage difference between the simulated apodised radiance for $\hat{\mathbf{x}}_{16}$ and the 16.42 km MIPAS measurement.

Figure 4.9: McCloudS_FM simulated spectra using ECMWF temperatures. The percentage difference between the simulated apodised radiance and the MIPAS measurement. The red line is the root mean square uncertainty due to measurement noise and in blue is the root mean square percentage difference.

too broad, which suggests that McClouds_FM may be overestimating the radiance scattered in from below the cloud, which is the source of the absorption lines. However, a more likely source of this error could be due to the use of climatology data for the gas concentration profiles. If the H_2O amounts are slightly inaccurate in the atmospheric model it will lead to discrepancies between model output and the data at the major lines.

The simulations were carried out without gaseous absorption (and emission) within the cloud domain. This was again done for simplicity to reduce the number of calculations required. However, the lack of gaseous absorption could explain both the missing fine structure in the simulations and the slightly larger discrepancy around the absorption features. The small numbers in table 4.2 for error due to gaseous absorption inside cloud, corroborate this explanation, as they are rms errors and (as mentioned in section 4.6) the main error is centred about the main gas spectral lines, i.e. H_2O and CO_2.

Another reason that the simulation may not match the measurement as well at the absorption features is that the viewing tangent height returned by the MIPAS pointing information[4] is only certain to within $\sim \pm 0.5$ km. Changing the tangent height in the simulations will change the scattering geometry (though only to a small degree as the radius of curvature in the spherical geometry is large). This in turn can alter the magnitude of the

---

[4]At the design stage of MIPAS the goal was to determine the geometric limb height by pointing information from the spacecraft with a standard deviation below 600 m. However, in the first few years after MIPAS launched the error on absolute tangent altitude was assumed to be up to 1.5 km. However, the tangent heights have since been retrospectively validated (and corrected where necessary) to confirm an error on tangent altitude of $\sim \pm 0.5$ km or better.

effect of warm photons scattering into the line-of-sight from below the cloud and hence the size of the absorption features. An example of this effect is given in figure 4.10. It can be seen that by differing the tangent height by 1 km the simulated radiance at the satellite can differ by up to ∼5 %, but interestingly the differences are smallest at the absorption features. This could prove a large error source in the quality of fit of the McClouds_FM simulation to the MIPAS data. However, this is a parameter error and is essentially a problem for the measurement inversion, as the cloud state vector can be changed to give a better simulation for the correct tangent height.
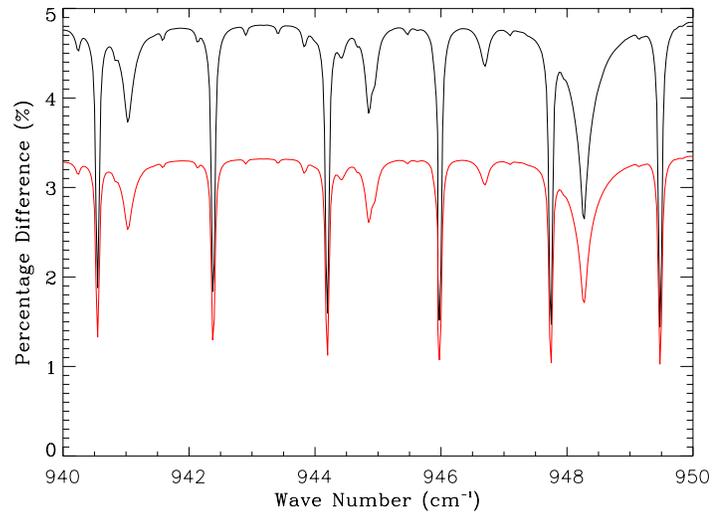
Figure 4.10: Percentage residual between a pencil beam simulation for $\hat{\mathbf{x}}_{16}$ at a tangent height of 16.42 km and simulations with a tangent height of 15.42 km (in red) and 17.42 km (in black).

4.9 Detection Sensitivity

Now that it has been shown that McClouds_FM can model cirrus contaminated MIPAS measurements well, using realistic cloud fields, McClouds_FM can be used to investigate the potential cloud detection sensitivity of MIPAS. By simulating the radiance across the detection threshold at $\bar{\nu} = 960.7$ cm^{-1} for a wide of range of optical depths of cloud, the minimum optical depth that can be detected can be investigated. Figure 4.11 shows the radiance at the threshold for cloud effective radius versus vertical optical depth, for a simulation with the cloud top height at 18.0 km and the viewing tangent height at 16.42 km. The figure clearly shows that cloud is detectable (above measurement noise) for optical depths of 0.002. This high sensitivity to optical depth is due to the multiple scattering enhancement. In addition, an interesting result is that the multiple scattering effect actually peaks at an optical depth of about 0.1 (increasing slightly at higher effective radii), then decreases and levels off at higher optical depths. This is because at low optical

depths (< 1.0) photons from lower in the atmosphere are scattered into the line-of-sight, and as the optical depth increases up to 0.1, the number of ice crystals increases and hence the strength of multiple scattering increases causing a radiance increase until a peak effect of scattering is reached. As the cloud becomes optically thicker the radiance decreases as the satellite detector sees less of the photons from below the cloud which are now being absorbed within the cloud. Eventually, as the optical depth increases further the satellite is only seeing the photons from the cloud top and the radiance levels off and in effect the satellite is only sensitive to the cloud top temperature.

The enhancement of multiple scattering over single scattering for the same cloud scenarios that produced figure 4.11, can be seen in figure 4.12. It can be seen that the peak multiple scattering effects occur at optical depths of around 0.1 and can cause radiances of up to 25 % greater than single scattering alone. This is a powerful result and clearly illustrates the importance of including multiple scattering in the radiative transfer.

4.10 Conclusions

The results from McClouds_FM have shown that a realistic cloud field can be found such that the MIPAS spectra with cirrus in the FOV can be well modelled to within instrument noise. Additionally, the root mean square difference of the modelled spectra and the limb measurements is below 5 %, thus meeting the initial aim for spectral accuracy requirements. Therefore, under the assumptions made in the model it can be expected that the cloud state parameters are a good first order approximation to the real cirrus. Further, the H_2O absorption lines are replicated in the McClouds_FM simulated spectra and (from figures 4.7(a) and 4.7(b)) it can be seen that these absorption features are clearly due to tropospheric radiation from below the cloud, and below the limb path being scattered into the instrument line-of-sight.

The results have also shown that simulated spectra under single scattering into the line-of-sight significantly underestimates the radiance arriving at the satellite. This clearly states the case for the inclusion of multiple scattering within cloudy regions in the limb radiative transfer calculations.
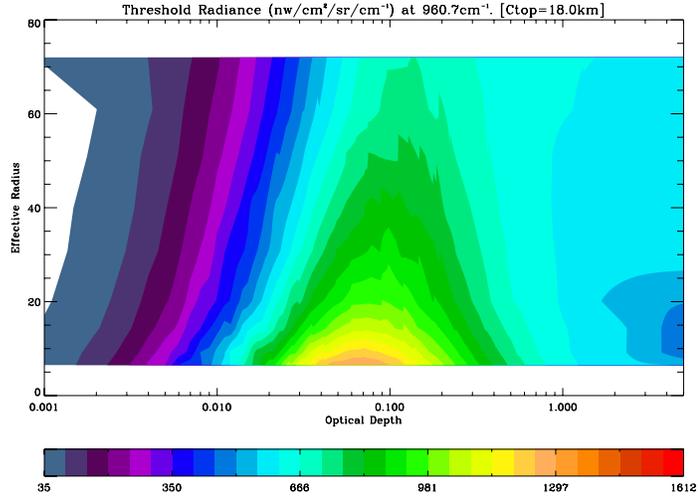
Figure 4.11: McClouds_FM radiance at 960.7 cm^{-1} for various cloud types with cloud top at 18.0 km and a viewing tangent height of 16.42 km. Radiance is a function of effective radius versus vertical optical depth.
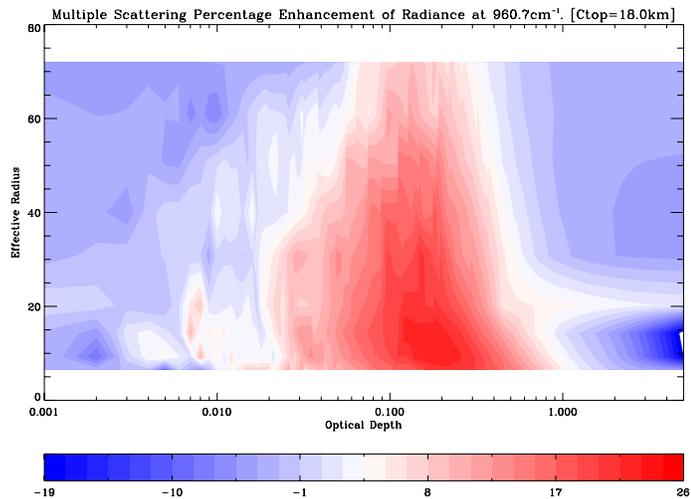
Figure 4.12: Percentage enhancement of multiple scattering over single scattering McClouds_FM calculated radiance at 960.7 cm^{-1} for various cloud types with cloud top at 18.0 km and a viewing tangent height of 16.42 km.

Chapter 5

Initial Estimates of Cloud Properties

In most instances, the optimal method to derive information from a set of measurements is to use a full nonlinear retrieval which minimises a cost function based on all of the available data and the appropriate *a priori* information (see chapter 6). However, it has been shown in chapter 4 that the forward model developed for retrieving cloud properties from MIPAS measurements is time consuming, due to the complex scattering calculations, and that it is a many-to-one, nonlinear mapping and as such, any cost function will have a complicated topology. In this case, a full retrieval is both likely to take a long time and to find a false minimum. In order to combat these problems, the aim of this chapter is to develop a suboptimal method to generate a good first guess for the retrieval. Furthermore, an algorithm to generate a first guess is essential if the retrieval is to be automated.

Practical retrieval methods often use the *a priori* as the first guess for the retrieval. However, this can result in the retrieval being biased by the *a priori* if the *a priori* covariance error is not large enough and, in the non-linear case, it also does nothing to address the problem of the retrieval finding a false minimum (or not converging at all). By developing an *ad-hoc* method to "point" the first guess to the neighbourhood of state space that the true solution is expected to lie (given the measurement), the retrieval will be less inclined to return a solution corresponding to a local minimum.

In order to develop an effective *ad-hoc* method, which is not computationally intensive, the inverse problem must be simplified by reducing the dimensionality of the problem, i.e. by reducing the dimension of the measurement space of interest, and reducing the dimension of the state space to be investigated. The dimension of measurent space of interest is restricted by selecting spectral properties or measures of the MIPAS spectra instead of considering the full spectrum. However, by doing so, the information content of the particular spectral properties is likely to be less than that of the full spectrum. The spectral properties selected are those used for cloud detection as described in chapter 3, and the properties introduced in the following section which are derived from the scattering features which are typically

present when cirrus is in the MIPAS line-of-sight.

The suboptimal method for generating a first guess is developed by investigating the sensitivity of the spectral properties to the cloud properties as modelled by McClouds_FM. By using the same forward model as will be used in the retrieval it allows the state space to be investigated (which is the purpose of section 5.2). McClouds_FM uses four state parameters to describe the cloud in the MIPAS FOV, and it is particularly difficult to visualise the dependence of the spectral properties on four variables. Also, it will be computationally time consuming to calculate the model results for a state space sampled at high resolution in four dimensions. Therefore, it is necessary to reduce the number of parameters of interest at one time (this problem is dealt with in section 5.2), and often it would be superfluous to do otherwise because of the reduced information content of the simplified measures of the spectral properties. The *ad-hoc* method for solving the inverse problem based on the sensitivity of McClouds_FM is described in section 5.3, which is then applied to real MIPAS data in section 5.4.

5.1 Detection of Cirrus Scattering Effects

Scattering is dependent on cloud microphysics, and therefore, the detection of scattering effects in the measured limb spectra will give indirect information on the size distribution of the observed cloud. A fast and simple detection of scattering effects in the spectra on a global basis would be helpful for the analysis and classification of cloud distributions (Spang et al., 2004). The analysis and classification of the cloud distributions, requires an understanding of the dependence of the scattered radiation measured at the limb detector to the cloud microphysics and the geometric properties of the cloud. This is investigated in section 5.2, while this section introduces the first requirement of a simple method to detect and quantify the scattering effects in the limb spectra.

The absorption lines in measured spectra due to radiation scattered into the line-of-sight has been discussed extensively throughout this thesis. In particular, figures 4.5 and 4.6 showed clear examples of absorbing H_2O lines and absorption related inverted CO_2 side-lobes in MIPAS spectra. In section 4.8.2, the modelling results (in figures 4.7(b) and 4.7(a)) showed that the absorption effects can only be ascribed to warm tropospheric radiation scattered into the line-of-sight from below the cloud.

The method for the detection of scattering effects focuses on the two gas species separately; first on the detection of particular H_2O lines and then on select CO_2 side-lobes.
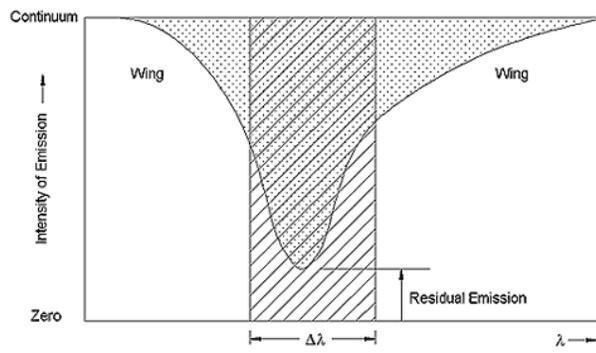
Figure 5.1: Equivalent line width schematic. The area enclosed by the absorption line is given by the dotted region, and the equvalent area enclosed by a pure black line is given by the diagonally lined region. The equivalent line width is labelled $\triangle\lambda$.

5.1.1 H$_2$O Absorption Features

There are several H$_2$O lines which are potential candidates for use in measuring the scattering effect. In particular, at $\sim$803 cm^{-1}, $\sim$825 cm^{-1}, and $\sim$948 cm^{-1} strong H$_2$O absorption features are typical of cloud contaminated spectra from high in the troposphere. There are two quantities which can be used to measure the absorption effect: the *band* or *line depth* and the *equivalent line width*.

The band depth of an absorption feature is basically the relative difference between the absorption band and the continuum. If the average radiance of the absorption band is given by R_{abs} and the continuum emission is given by R_{cntm} then the band depth D_{abs} is given by:

$$D_{\mathrm{abs}} = \frac{R_{\mathrm{abs}} - R_{\mathrm{cntm}}}{\frac{1}{2}(R_{\mathrm{abs}} + R_{\mathrm{cntm}})}. \tag{5.1}$$

The equivalent width is the width that a pure black line would have to have in order to enclose the same area underneath the 100 % emission level (i.e. the continuum), that the actual line encloses (see figure 5.1). The difficulty in calculating both quantities comes in fitting an appropriate continuum line.

For each of the absorption lines, two microwindows in the region close to either side of the feature are selected as a *buffer* region in which the spectra is relatively insensitive to gaseous emission and absorption. The continuum is taken to be the average of the radiance in the buffer. The absorption band radiance is then taken as the average radiance of the measurements from a narrow microwindow situated at the expected spectral location of the scattering feature. Figure 5.2(a) shows the microwindows for the absorption feature at $\sim$803 cm^{-1}, with example MIPAS spectra. The buffer microwindows and absorption band microwindow for each of the absorption lines are given in table 5.1. The band depth given by equation 5.1 for each of the three absorption lines is known as the Scattering Effect Index (SEI) and the three lines are labelled 1,2 and 3 with the increment corresponding to

SEI	Absorption Band MW (cm^{-1})	Continuum Buffer MWs (cm^{-1})		Region of Interest for EqW (cm^{-1})
SEI-1	803.5–803.6	802.2–802.3 and	803.8–803.95	803.25–803.9
SEI-2	825.1–825.2	824.6–824.8 and	825.3–825.5	824.8–825.4
SEI-3	948.2–948.3	947.2–947.4 and	948.5–948.7	947.925–948.575

Table 5.1: The Scattering Effect Indices for MIPAS.

the wavenumber. Negative values of SEI immediately indicate an absorption feature. This method is similar to that used by Spang et al. (2004) to indicate absorption features in cloud contaminated MIPAS spectra.

For each of the SEIs, if negative, the corresponding equivalent line width (EqW) is calculated. The area of the absorption line is the area enclosed by the fitted continuum line and the first points either side of the absorption band where the continuum line and the spectra intersect. If there is no intersection either side, then the enclosed area is bounded by the continuum line and the spectral points (either side of the abosprtion band) for which the radiance is closest to the continuum level, within a region of interest, which is also given in table 5.1. An example of the area enclosed by the absorption line for SEI-1 is given in figure 5.2(b).

The absorption features of interest are spectrally very narrow, and as such both the Scattering Effect Indices and equivalent line widths are sensitive to noise. Generally, the spectra should be apodised before the calculations are carried out. It can be seen from figure 5.2 that the H_2O line is clearly absorbing in the cloudy spectra, and emitting in the clear-sky spectra. The H_2O lines used to calculate SEI-2 and SEI-3 are not as strongly emitting in clear-sky spectra and as such SEI-2 and SEI-3 are more affected by noise for clear-sky spectra (CI-A > 4).

Figures 5.3(a) and 5.3(b) show the values obtained from the MIPAS data from August 2003, for measurements in the range 30 °S to 30 °N at tangent heights between 10 and 17 km (where high tropical cirrus is expected). Comparing the figures, the influence of noise on the SEI-2 values can be seen clearly for the clear-sky spectra with high CI-A values, such that the noise causes negative SEI-2 values where there should not be absorption features. The relationship of SEI-3 to CI-A is similar to that of SEI-2, but the noise problem is not quite so pronounced. Figure 5.3(a) shows that SEI-1 has a reasonably well constrained relationship to CI-A. In addition, a large number of spectra can be seen to have negative values of SEI-1, in fact 15 % of all spectra are negative and 43 % of spectra in which cloud is detected (by the Principal Components Cloud Test) have negative values of SEI-1 (the figure is 50 % for cloudy spectra detected by the Colour Index method alone). These figures are higher for SEI-2 (as can be seen in figure 5.3(b)) due to the influence of noise at high CI-A values, with 21 % of all spectra having negative SEI-2 and 61 % of all cloud contaminated
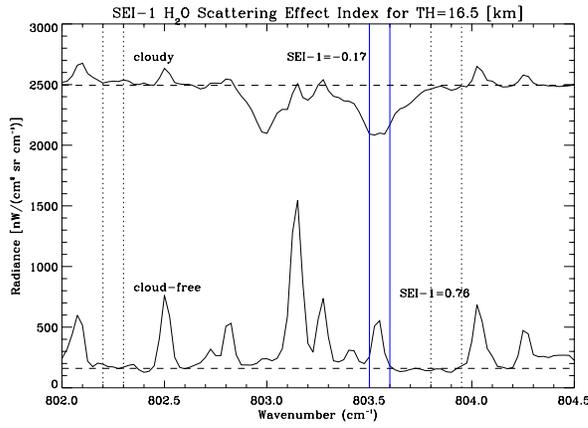
(a) The absorption band is clearly indicated by the region between the blue lines and the continuum buffer regions by the regions between the two sets of dotted lines. The fitted continuum base lines are plotted by the dashed lines.
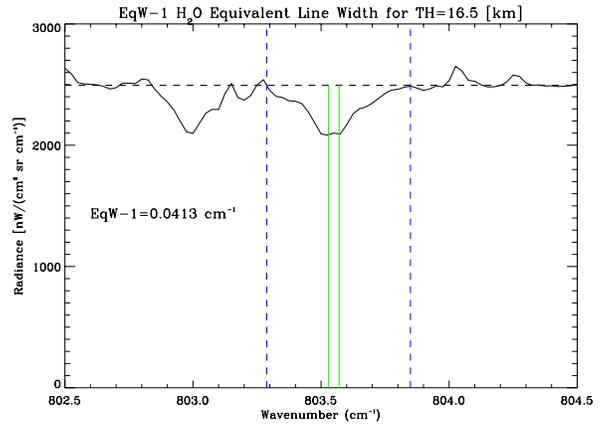
(b) The area of the absorption line is enclosed by the two dashed blue lines, the dashed continuum line, and the spectra. The equivalent line width is shown by the distance between the two green lines.

Figure 5.2: Figure (a) shows the MWs used to calculate SEI-1 with example radiance for a cloudy (lat=1.25 °N; lon=36.5 °E) and clear (lat=-14.5 °N; lon=35.2 °E) limb measurement at ∼16.5 km from MIPAS orbit 504. Figure (b) shows the area enclosed by the absoprtion line of the same cloudy spectra, and shows the calculated equivalent line width (EqW-1).
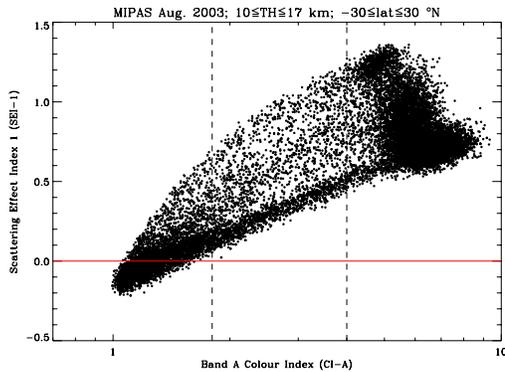
(a) Scattering Effect Index 1 (SEI-1) versus CI-A.
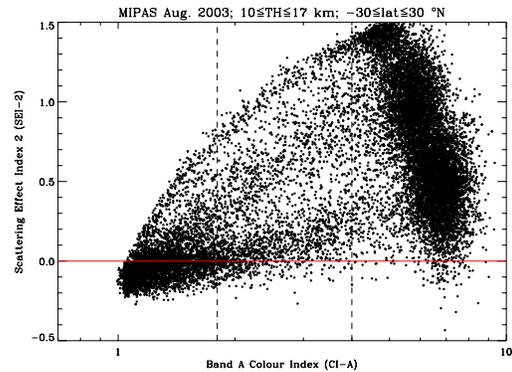
(b) Scattering Effect Index 2 (SEI-2) versus CI-A.

Figure 5.3: The Scattering Effect Indices calculated from MIPAS spectra from August 2003 in the region of interest for high tropical cirrus, i.e. $10 \leq t_h \leq 17$ km and $-30 \leq$ lat ≤ 30 °N.
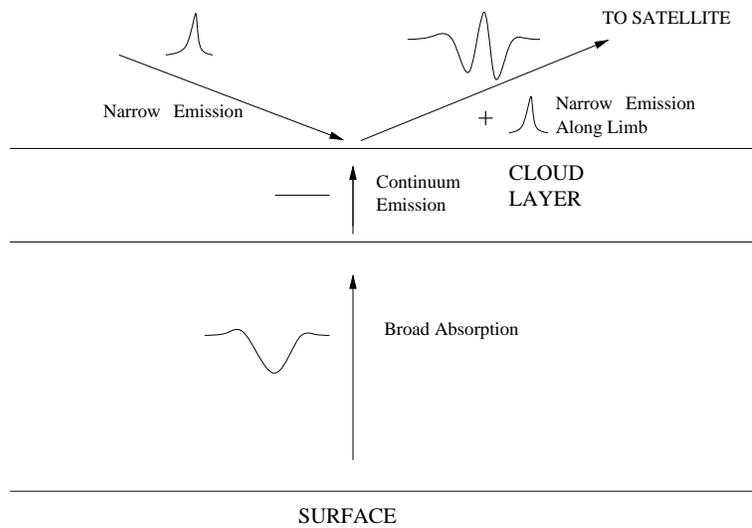
Figure 5.4: Expected absorption and scattering features in limb measurements.

spectra. As the Scattering Effect Indices contain similar information, the empirical method developed focuses on the use of SEI-1 to avoid noise problems as much as possible.

5.1.2 CO_2 Absorption Features

Like the H_2O lines discussed previously, CO_2 absorption features are caused by tropospheric radiation scattered into the line-of-sight. However, unlike the H_2O absorption features, which are generally absorption lines, the CO_2 features have a side-lobe structure. An illustrative explanation of this is given in figure 5.4. Essentially, the scattering features are due to broad absorption features being scattered into the line of sight from below the cloud which are duly added to narrow emission features from along the limb or scattered into the line-of-sight from above the cloud. Hence, the spectra measured at the satellite at particular CO_2 lines is shaped like a sinc function with a narrow emission line peak and absorption side bands either side.

Clear examples of CO_2 absorption features have been shown previously in cloud contaminated MIPAS limb spectra (see figures 4.5 and 4.6) at $\sim$944 cm^{-1} and $\sim$946 cm^{-1} and there is an additional strong absorption feature at $\sim$968 cm^{-1}. These features are spectrally much narrower than the H_2O absorption features, and the MIPAS spectra must be apodised before they are analysed. Moreover, as the features are narrow and sinc-like, the conventional measures of band depth and equivalent line width are not as appropriate. Instead, two very simple measures are used, though they both still depend on fitting a continuum line. The same method as the previous section is used to fit the continuum line, and two continuum buffer regions either side of the absorption feature are selected from which the average radiance is calculated and assumed to be a close fit to the continuum radiance R_{cntm}. Although the conventional band depth isn't used, a measure of

SLI	Absorption Side Band Spectral Points (cm^{-1})	Continuum Buffer MWs (cm^{-1})		
SLI-A	944.125, 944.25–944.3	943.8–943.9	and	944.4–944.5
SLI-B	945.9–945.95, 946.025	945.65–945.75	and	946.15–946.25
SLI-C	967.625–967.675, 967.75–967.775	967.4–967.5	and	967.9–968.0

Table 5.2: The CO_2 absorption Side Lobe Indices (SLIs) for MIPAS.

PK	Region of Interest (cm^{-1})
PK-A	944.1–944.25
PK-B	945.9–946.0
PK-C	967.65–967.75

Table 5.3: The CO_2 emission peak ratios (PKs) for MIPAS.

the amount or depth of absorption is required. Due to the narrow spectral width, spectral points where absorption side bands are expected are selected and an average absorption side band radiance is calculated, known as R_{sb}. Then the Side Lobe Index (SLI) is given by

$$\text{SLI} = \frac{R_{\text{cntm}} - R_{\text{sb}}}{\frac{1}{2}(R_{\text{cntm}} + R_{\text{sb}})} \tag{5.2}$$

which is very similar to the equation for the band depth (equation 5.1), but in this case a positive value indicates an absorption feature. In addition, a measure of the emission peak is required, as the emission line is very narrow, the ratio of the maximum radiance within the expected spectral region of the peak R_{pk} to the continuum radiance is taken as the relative magnitude of the emission peak. This value is known as PK, such that

$$\text{PK} = \frac{R_{\text{pk}}}{R_{\text{cntm}}}. \tag{5.3}$$

For the CO_2 absorption features at $\sim$944 cm^{-1}, $\sim$946 cm^{-1} and $\sim$968 cm^{-1}, the spectral regions for calculating the side band absorption as well as the continuum buffer regions are given in table 5.2. The spectral regions for calculating the maximum emission peak radiance are given in table 5.3. Figure 5.5 shows the microwindows at $\sim$944 cm^{-1} used to calculate SLI-A and PK-A for example MIPAS spectra.

As the SLI and PK values are calculated without averaging over many measurements, they are prone to error due to measurement noise. However, this is unavoidable because of the narrowness of the features. Figure 5.6 shows the values of SLI-A and PK-A obtained from the MIPAS data from August 2003, for measurements in the range 30 °S to 30 °N at tangent heights between 10 and 17 km. The CO_2 absorption features are perhaps not quite as sensitive to scattering as the H_2O features, as only 12 % of all measurements have positive SLI-A values and 40 % of cloud contaminated measurements have positive SLI-A.
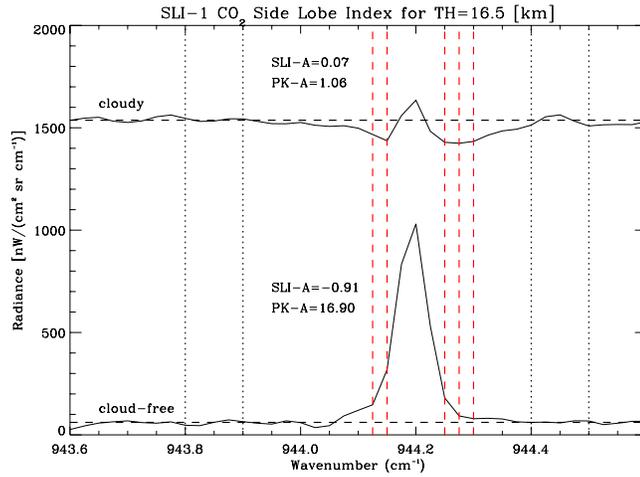
Figure 5.5: Spectral regions used to calculate SLI-A for a cloudy (lat=1.25 °N; lon=36.5 °E) and clear (lat=-14.5 °N; lon=35.2 °E) limb measurement at ∼16.5 km from MIPAS orbit 504. The continuum buffer regions are indicated by the dotted lines and fitted continuum lines for the cloudy and clear spectra are plotted by the upper and lower black dashed lines respectively. The spectral points used to calculate the side band absorption average are indicated by the red dashed lines, and the region of interest for selecting the maximum peak radiance is $941.1 \leq \bar{\nu} \leq 941.25$ cm^{-1} which is roughly the region bounded by the two side lobe bands.
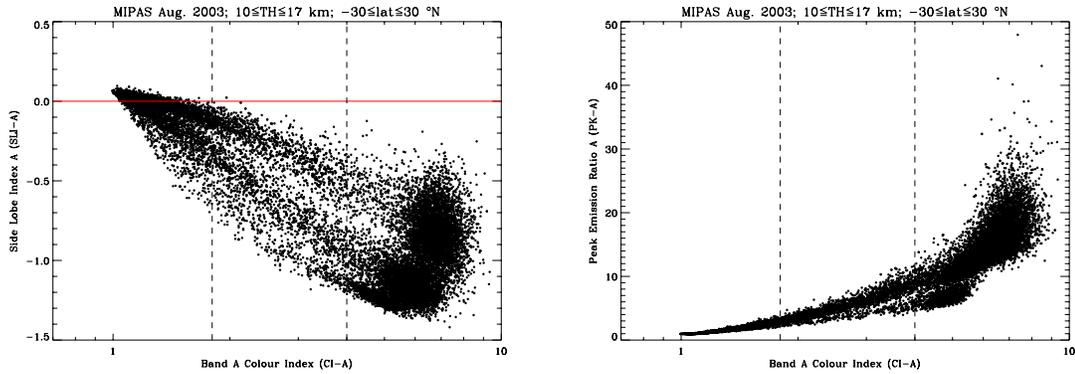
(a) Side Lobe Index A (SLI-A) versus CI-A.

(b) Peak Emission Ratio A (PK-A) versus CI-A.

Figure 5.6: Side Lobe Index A and Peak Emission Ratio A calculated from MIPAS spectra from August 2003 in the region of interest for high tropical cirrus, i.e. $10 \leq t_h \leq 17$ km and $-30 \leq$ lat ≤ 30 °N.

However, although the SLI-A values are in gross terms less sensitive to scattering effects than the SEI-1 values, if the scattering effect peaks for a certain size range of particles the SLI-A values may be a stronger indicator of the particle size. In addition, the values of PK-A can be seen to increase exponentially with CI-A, and thus it may be of use in increasing the confidence of clear-sky detection. The values of SLI-B and SLI-C, and PK-B and PK-C have a near identical relationship to SLI-A and PK-A, respectively, with CI-A.

5.2 Modelled Cloud Indices and Scattering Effects

The dependence of the modelled detection and scattering measures on the cloud state parameters was investigated. There are 14 properties derived from the measurement space (3 SEIs, 3 EqWs, 3 SLIs, 3 PKs, the window threshold radiance and CI-A), and 4 state parameters. However, the objective of the investigation was to derive some empirical relationship between the measures and state parameters, and without reducing the dimensionality of the problem this would involve approximating a solution to a system of 14 (non-linear) simultaneous equations in 4 unknowns. This would be an ineffcient use of time, considering the complexity of the problem and that the relationship is only required to obtain a gross estimate of the cloud state, as it is the job of the retrieval to find the optimal solution. Therefore, for the purposes of the sensitivity study the measures of interest were restricted to[1]: SEI-1, SLI-A, CI-A and the threshold radiance (so the number of measurements are equal to the number of parameters in the cloud state vector and the problem is, in theory, well-posed). The principal component contributions (see chapter 3) were not simulated as this would involve simulating spectra from 750 to 970 cm^{-1}, which is considerably more computationally costly. In addition, the principal components are dependent on latitude and altitude and this would add a further degree of complexity to the problem.

Deriving an empirical function of 4 variables is an onerous task, thus the number of dependent variables investigated was restricted by sampling in terms of cloud top height C_{top}, effective radius r_e and infrared/thermal optical thickness τ at 12 μm. The thermal optical thickness is given by

$$\tau = \beta_{\text{ext}}(r_e, N).C_{\text{depth}} \tag{5.4}$$

where β_{ext} is the extinction calculated at 12 μm which is a function of the effective radius and the number density N (for a given crystal habit - see section 4.2.1). By investigating the sensitivity of the modelled measures in terms of τ, the 4 dimensions of the state space are explored, but the dependence need only be thought of in terms of three variables. When the initial estimate of the cloud state is calculated, τ must be transformed back into terms

---

[1]The Equivalent Line Widths are not useful for investigating the state space as they are only meaningful when the H$_2$O absorption lines are present (i.e. when SEI-1 is negative).

of N and C_{depth} for a given r_e (section 5.3).

In section 4.1.1 (see equation 4.1) it was stated that the synthesised measures are not only dependent on the cloud state vectors, but also on the set of atmosphere parameters, $\hat{\mathbf{b}}$ (see section 4.2.3). This includes profiles of temperature, pressure, trace gas concentration and also the tangent height of the viewing limb path t_h. Therefore, the relationship between the detection and scattering measures and the cloud state is likely to also depend (weakly) on $\hat{\mathbf{b}}$. However, for the purposes of the investigation an equatorial/tropical climatology atmosphere (Remedios, 1999) was assumed for all calculations as high tropical cirrus (30 °S $\leq$ lat $\leq$ 30 °N) is the focus of this thesis. Although the MIPAS viewing tangent height (reported by the Level-1B data) can differ from from the nominal MIPAS tangent heights by up to 1.8 km, for the same cloud top heights relative to the tangent height, the synthesised measures are not expected to vary significantly with small changes in the t_h value input in $\hat{\mathbf{b}}$. Therefore, the synthesised measures are only calculated at the nominal tangent heights of interest, i.e. 12, 15 and 18 km. As the dependence of the measures on tangent height is weak, calculating the measures over a large range of tangent height values would prove to have a high time cost for a poor return on information, as well as further complicating the inverse estimation of the cloud properties. It is assumed that the relationships between the measures and cloud parameters (with respect to the relative cloud top height) at each nominal tangent height holds over the uncertainty range.

For each tangent height, the modelling was carried out over a range of cloud properties. Hexagonal column crystals were assumed and the effective radius was sampled over nine values: 6.5, 9.2, 14.3, 20.2, 30.8, 40.0, 51.0, 60.9, and 72.0 μm, which describe the cirrus size distributions derived from *in situ* measurements due to Fu (1996) (see section 4.2.1). The thermal optical depth was sampled over the range $0.001 \leq \tau \leq 1.0$, and the cloud top height was sampled over the range $t_h - 1.5 < C_{\text{top}} \leq 19$ km. As the range of cloud top heights of interest include values for which the MIPAS FOV will be partially filled, the FOV effects must be adequately accounted for (see section 4.7.2). However, the computational cost of using a high number of pencil beams to sample across the FOV is not cheap. Therefore, the choice of the FOV function and the number of pencil beams used is determined by a trade off between the computational cost and the physical accuracy. For the modelling discussed in this chapter, a 7 beam FOV with a 3 km trapezoidal FOV response function is used.

The results for the 12 km tangent height measures of SEI-1, SLI-A, CI-A and the threshold radiance over the range of τ and C_{top} are given in figures 5.7 and 5.8 for $r_e = 9.2$ μm and $r_e = 51.0$ μm respectively.

Immediately apparent from both figures 5.7 and 5.8 is that the modelled values are considerably different depending on whether the cloud top height is above or below the tangent height. This is particularly apparent for the cloud top heights within the FOV, i.e.

$t_h - 1.5$ km $< C_{\text{top}} < t_h + 1.5$ km. Of particular interest is that for the smaller effective radius simulations (figures 5.7(a) and 5.7(d)), indication of scattering features (i.e. negative values of SEI-1 and positive values of SLI-A) is only seen for the cloud top heights above the tangent height. There is also a similar trend for the 51.0 μm effective radius simulations, but with the SEI-1 and SLI-A values both tending to zero for higher optical thickness and cloud tops above the tangent height (see figures 5.8(a) and 5.8(d)). Furthermore, it can be clearly seen that for both the 9.2 μm and 51.0 μm simulations, at the same optical thickness, the threshold radiance tends to be lower and the Colour Index tends to be higher when the cloud top is below the tangent height (which is true over the range of effective radii). This result is explained by the measurement becoming increasingly dominated by the presence of cloud as a greater percentage of the FOV is filled (see figures 5.7(b) and 5.8(b)). Therefore, generally there will be a higher detection sensitivity to cloud top heights above the nominal viewing tangent height, i.e. when a greater portion of the FOV is filled. With respect to detection sensitivity, it can be seen from the threshold radiance and colour band simulations that the cloud detection has a stronger dependence on the cloud optical thickness than on the cloud top height. From figures 5.7(c) and 5.8(c), the detection sensitivity limit with respect to optical thickness is variable, but the rough limit for the 9.2 μm effective radius is $\tau \approx 0.003$ under the principal component cloud test and $\tau \approx 0.015$ for the Colour Index method, and the limit for the 51.0 μm effective radius is $\tau \approx 0.0005$ under the principal component cloud test and $\tau \approx 0.0015$ for the Colour Index method. This is further evidence that the Colour Index method will tend not to detect tenuously thin cirrus effectively (further to chapter 3). Generally, for low optical thickness clouds, those with a larger effective radius of the constituent ice particles tend to give a higher radiance and lower Colour Index than for clouds with a smaller effective radius.

The simulated scattering and detection measures at 9, 15 and 18 km show very similar trends with respect to effective radius, thermal optical thickness and relative tangent height. From the results in figures 5.7 and 5.8, there is considerable (discontinuous) variability in SEI-1 and CI-A with respect to both optical thickness and cloud top height when the cloud top is within the FOV. This is due to the approximation of the FOV response function by a (small) finite number of pencil beams modelled by McClouds_FM. For cloud top heights above the field of view (i.e. $C_{\text{top}} > t_h + 1.5$ km), the values of SEI-1, CI-A and SLI-A are less sensitive to the cloud top height and begin to depend only on the thermal optical thickness. However, this is not true of the threshold radiance which peaks in magnitude at cloud top heights at approximately $t_h + 2$ km such that the FOV is entirely filled and the geometry will still be such that the ray path lengths through the cloud are still long. The threshold radiance then begins to decrease as the cloud top height increases, then levelling off before increasing again slightly >17 km when the cloud top enters the stratosphere

and the atmospheric temperature increases again. The threshold radiance also peaks with respect to optical thickness at $\tau \sim 0.1$, as in section 4.9, which coincides with the peak in scattering effect (figures 5.7(a) and 5.8(a)) which increases the radiance by scattering a greater number of photons into the line-of-sight from below the cloud.

As the Colour Index and scattering indices only vary weakly with cloud top height when the cloud top is above the FOV (i.e. the FOV is filled), the values can be averaged with respect to cloud top height in order to investigate the variance of the values with respect to optical thickness and effective radius. Figure 5.9 shows the average values of SEI-1 calculated at tangent heights of 9, 12 and 15 km for the respective cloud top ranges: $t_h + 1.5 < C_{\text{top}} \leq t_h + 4.5$ km. The results for 18 km are not investigated as it is unlikely that clouds will form above the FOV range at 18 km, i.e. above 19.5 km, in the tropical region. The relationship between r_e, τ and SEI-1 shows little variation between the different tangent heights, although the magnitude of the scattering effect decreaes at the lower tangent heights.

It is clear that (at all tangent heights) the dependence of SEI-1 on effective radius and optical thickness is complex and non-linear[2], though the general trend at each of the tangent heights is that absorption features (evidenced by negative SEI-1) are stronger for smaller effective radii. This implies that smaller crystals enhance scattering to a larger degree, which is probably because smaller crystals tend to have a higher single scatter albedo at such wavelengths and their phase function has a weaker forward peak which means the scattering can be more diffuse. In addition, the absorption effects peak in the approximate region of thermal optical thicknesses around 0.1. For photons travelling upward, they are more likely to diverge into the line-of-sight if they undergo a greater number of scatters. This requires a greater density of cloud particles or a greater vertical thickness of cloud, which in effect requires a significant thermal optical thickness. However, as can be seen in figure 5.9, if the optical thickness increases above this optimum level for scattering, the number of photons absorbed inside the cloud increases and the scattering effect is diminished.

Further information can be obtained by comparing the measurements between each of the tangent heights at which the cloud will be detected. However, only measurements between adjacent tangent heights should be compared because of the assumptions in McClouds_FM about the horizontal extent of the cloud. From figure 2.8, the horizontal distance between adjacent ray paths at the same height in the atmosphere is $\sim$100 km, which is consistent

---

[2]The complexity of the SEI-1 relationship to optical thickness and effective radius can in part be explained by the difference in the size distributions assumed. The size distributions are derived from *in situ* studies (Fu, 1996), and the total number density is not constant with respect to effective radius. Neither does the total number density vary smoothly with effective radius, and it peaks at effective radii of 30.8, 51.0 and 72.0 μm. It can therefore be seen from figure 5.9 that high total number density seems to be correlated with positive/high values of SEI-1.
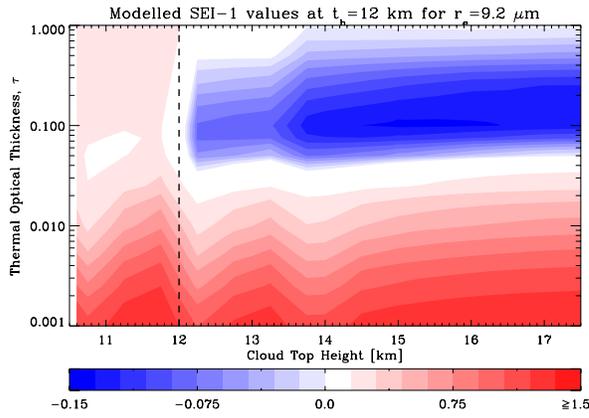
(a) Scattering Effect Index 1 (SEI-1)
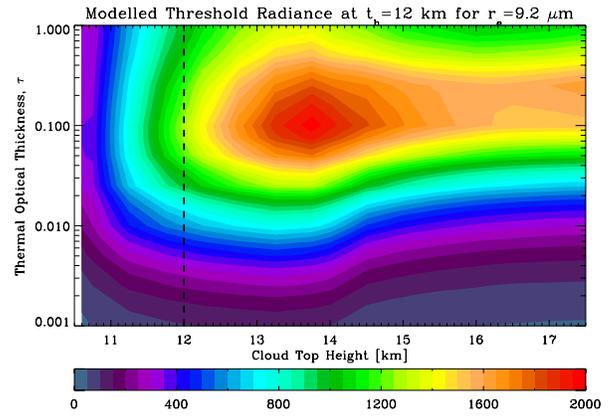
(b) Window Threshold Radiance
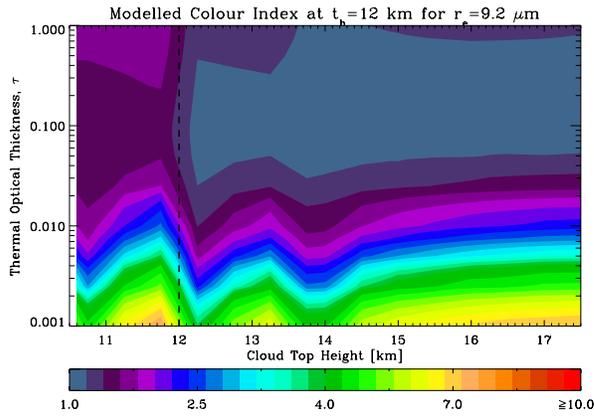
(c) Band A Colour Index (CI-A)
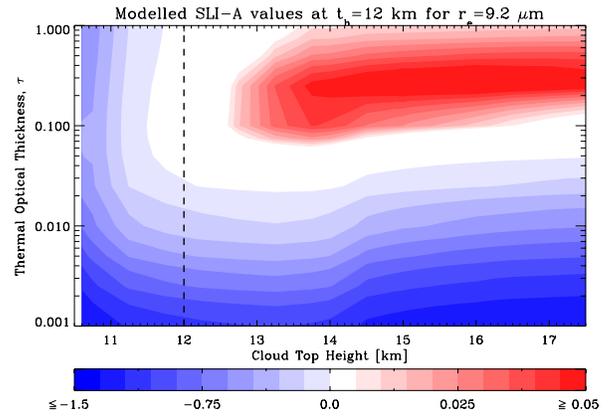
(d) Side Lobe Index A (SLI-A)

Figure 5.7: Detection and scattering measures modelled by McClouds_FM for a viewing tangent height of 12 km, an effective radius of 9.2 μm, $0.001 \leq \tau \leq 1.0$, and $10.5 < C_{\text{top}} \leq 17.5$ km.
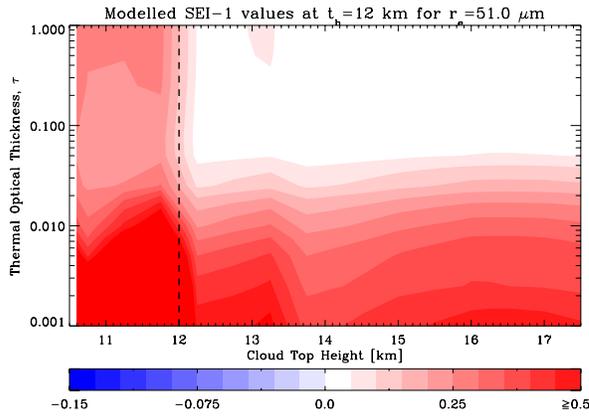
(a) Scattering Effect Index 1 (SEI-1)
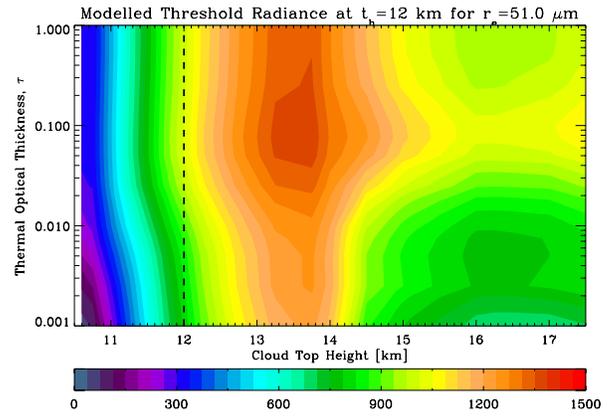
(b) Window Threshold Radiance
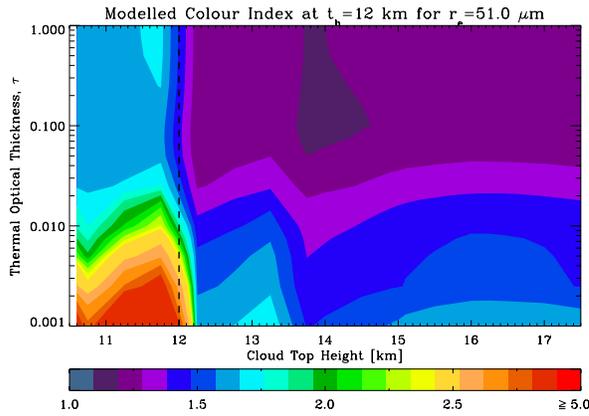
(c) Band A Colour Index (CI-A)
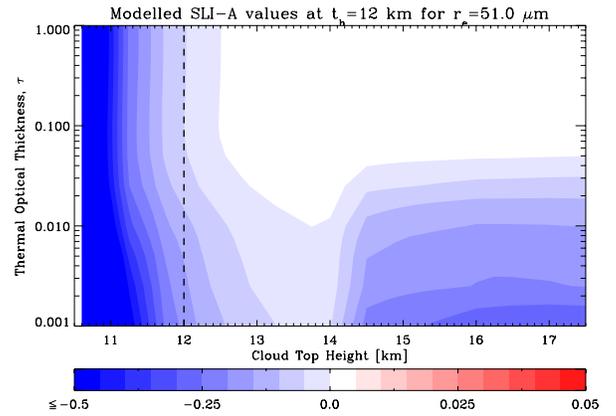
(d) Side Lobe Index A (SLI-A)

Figure 5.8: Detection and scattering measures modelled by McClouds_FM for a viewing tangent height of 12 km, an effective radius of 51.0 μm, $0.001 \leq \tau \leq 1.0$, and $10.5 < C_{\text{top}} \leq 17.5$ km.
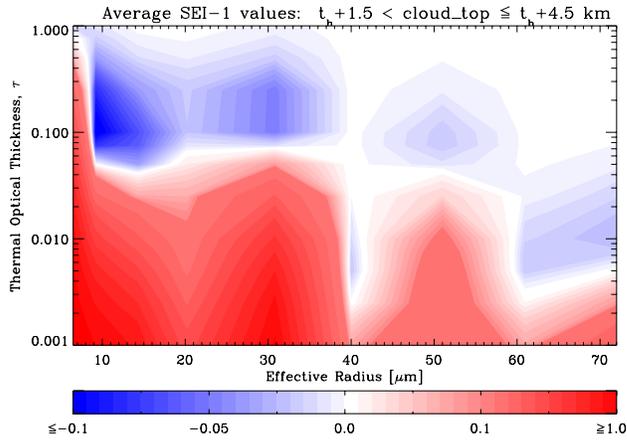
(a) Averaged SEI-1 values at a 9 km tangent height for $t_h + 1.5 < C_{\text{top}} \leq t_h + 4.5$ km.
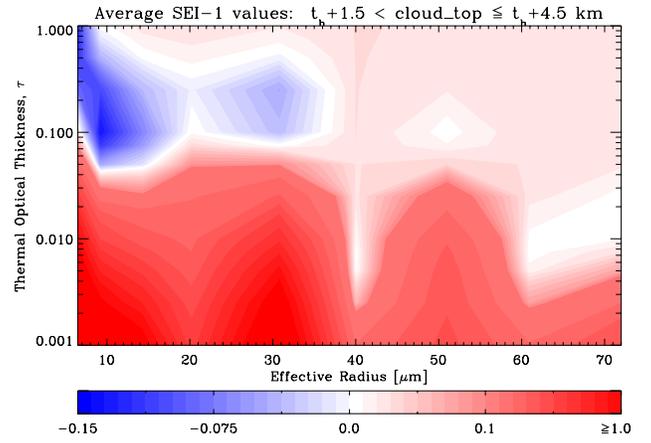
(b) Averaged SEI-1 values at a 12 km tangent height for $t_h + 1.5 < C_{\text{top}} \leq t_h + 4.5$ km.
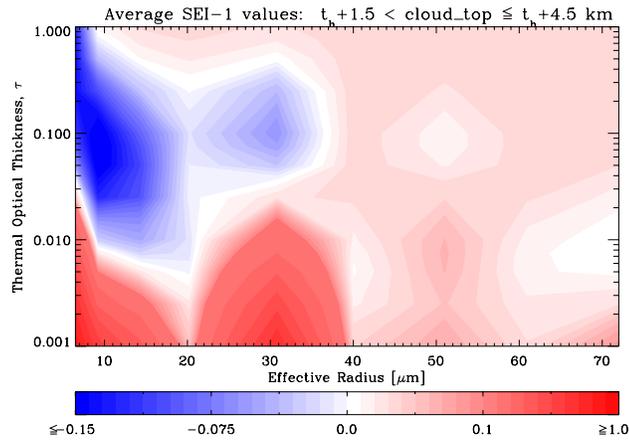
(c) Averaged SEI-1 values at a 15 km tangent height for $t_h + 1.5 < C_{\text{top}} \leq t_h + 4.5$ km.

Figure 5.9: The modelled values of SEI-1 averaged over the cloud top heights in the first 3 km above the FOV for each of the tangent heights at 9,12 and 15 km.

with the assumptions of cloud horizontal extent. It is difficult to interpret the physical meaning of ratios of the scattering effects and Colour Index at different tangent heights. However, the ratio of the threshold radiance of the limb scan cloud is first detected in and the radiance of the limb scan below can be more easily interpreted. Figures 5.10(a) and 5.10(b) give the ratio of threshold radiance at t_h=15 km and t_h=12 km for effective radii of 9.2 μm and 51.0 μm respectively. It can be seen that the radiance ratio varies mainly with respect to cloud top height. This is because the FOV of the 12 km limb measurement is filled, thus the radiance will be high and fairly constant for cloud top heights over the range of the 15 km limb FOV and the 15 km limb radiance will be low when the cloud top is below 15 km, gradually increasing with cloud top height (i.e. the FOV becomes increasingly filled). As the cloud top height increases above the 15 km FOV, the ratio will tend to approximately unity as the cloud top height increases. This is because the path length of both of the limb paths decreases, limited by the vertical cloud depth and the viewing geometry. As scattering has a greater effect on longer path lengths, the scattering effect decreases with cloud top height and the radiance of each limb scan tends to its respective limit (see figures 5.7(b) and 5.8(b)). Figures 5.10(a) and 5.10(b) show that the ratio of threshold radiances from two limb scans can give a good indication of cloud top height within the highest limb scan FOV range. Although the radiance ratio is still dependent on effective radius, it is particularly useful for the estimation of cloud top height because the ratio does not depend on atmospheric temperature, unlike the absolute value of the threshold radiance.
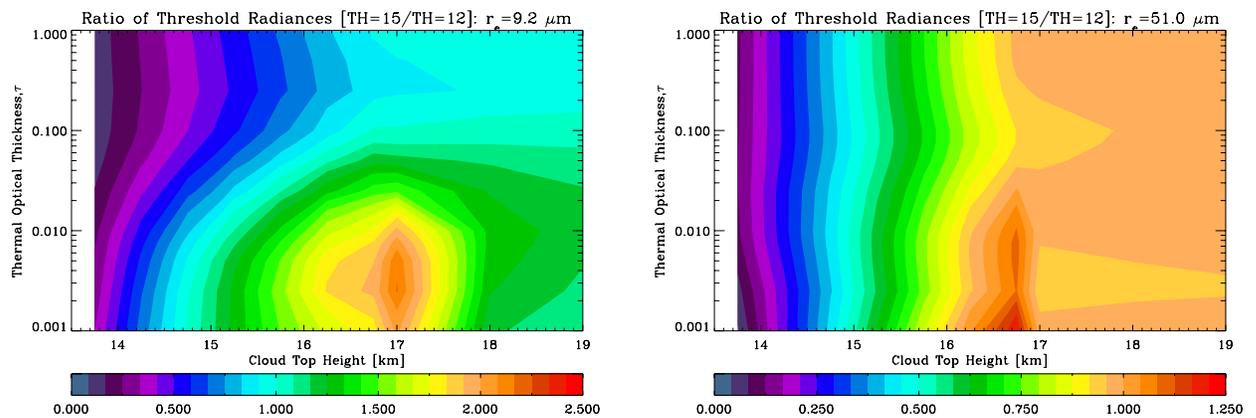
(a) For effective radius r_e=9.2 μm.

(b) For effective radius r_e=51.0 μm.

Figure 5.10: Ratio of the threshold radiance at 15 km over the radiance at 12 km for cloud top heights in the range detectable in both limb scans.

5.3 Estimating Cloud Properties

In order to estimate the properties of cloud detected in a MIPAS profile, the results from the modelling in the previous section are used to obtain an approximate solution to the inverse problem. The detection and scattering measures are calculated from the real MIPAS data and the values are then compared to the modelled values in order to obtain an idea of the cloud properties. In doing so, the measurement error is neglected[3], as are the differences between the climatology atmosphere and the true atmosphere, and it must be assumed that a realistic range of cloud properties has been simulated and that the resultant model output is representative of the range of values for real cirrus (and therefore that McClouds_FM is an accurate representation of the physics). However, the difficulty lies in determining the cloud properties from the detection and scattering measures. It is clear from the previous section that the detection and scattering measures have a non-linear dependence on the cloud properties of interest (i.e. τ, r_e and C_{top}), and as such the inverse function mapping the domain of modelled scattering and detection measures to the range of cloud properties is not an injection (i.e. a one-to-one map). Thus, the problem of estimating cloud properties is ill-conditioned.

The problem can be tackled by separating the treatment of one of the cloud variables and then by making an assumption about this variable given the measurements, the other two variables can be inferred if given some external constraints. The best candidate to initially separate and obtain an approximate solution for is the cloud top height. In figures 5.7 and 5.8 it can be seen that the values of SEI-1, CI-A and SLI-A at a given tangent height vary weakly with cloud top height when it is above the FOV. Thus, the dependence on effective radius and optical thickness can be determined with some confidence (using the results in figure 5.9) if the cloud top height can be determined to be above the FOV. This is a simple problem, as the measurements from the whole MIPAS profile can be used to determine the approximate cloud top height. The cloud will be above the FOV in all measurements at tangent heights below the first limb scan the cloud is detected in. For the reasons stated in the previous section on the assumptions of horizontal cloud extent, the tangent height immediately below the cloud top event limb scan is used to obtain the scattering and detection measures and the corresponding modelling results for that tangent height are used to estimate effective radius and thermal optical depth. Furthermore, this bypasses the problems in the modelling results when cloud top heights are in the FOV due to the approximation of the FOV response function.

From the accurate detection of cloud presence (i.e. the principle component cloud test) the cloud top height can be estimated within 3 km given the tangent height and FOV width

---

[3]This is only a problem for the threshold radiance of the measured spectra, as SEI-1, SLI-A and CI-A are all calculated from averaging over microwindows.

of the first limb scan that cloud is detected at in a given profile. From figures 5.7 and 5.8, it is clear that there is no general trend using any combination of SEI-1, SLI-A, CI-A and the threshold radiance which can determine cloud top height to better than 3 km for all values of r_e and τ.

Once the cloud top height has been roughly determined, the appropriate set of modelling results can be used to map the value calculated from the MIPAS limb scan to the effective radius and optical thickness. The cloud top height can then be re-estimated using the modelling results once values for r_e and τ have been obtained. In the method used, only the Scattering Effect Index modelling results were considered as they were more sensitive to r_e and τ then CI-A or SLI-A. The value of SEI-1 from the limb scan immediately below the cloud top event is calculated and compared to the SEI-1 model results for the corresponding tangent height which have been averaged over the first 3 km above the FOV (see figure 5.9). From figure 5.9, it can be seen that for a given value of SEI-1, there may not be a unique r_e and τ pair to which it corresponds[4]. At this point, external constraints are required to determine the most plausible values of r_e and τ for the first guess. From the set of all pairs of r_e and τ which give the same value of SEI-1, the pair chosen as the first guess is that which is considered to be most likely by the external constraints. If the external constraint is in the form of a joint probability density function (*pdf*), given by $\mathbf{P}(r_e, \tau)$, then the probability of particular values r_e and τ occuring is $\mathbf{P}(r_e, \tau)\mathrm{d}r_e\mathrm{d}\tau$. The first guess is taken to be the pair of r_e and τ which has the highest probability of occurrence.

As the problem is ill-conditioned, any constraints are desirable, but, from chapter 1 it is clear that there is large uncertainty in typical values of both effective radius and optical thickness for cirrus. Moreover, the joint dependence of r_e and τ is unclear. Therefore, $\mathbf{P}(r_e, \tau)$ is better determined by transforming from r_e and τ to IWC, for which several *in situ* campaigns have made direct measurements of cirrus, which are more accurate than other *in situ* measurements of particle size (see section 1.5.3). IWC can be expressed as a function of r_e and τ for an assumed crystal habit and known size distribution, i.e. IWC[r_e, τ], such that

$$\mathbf{P}(r_e, \tau) = \mathbf{P}(\text{IWC}[r_e, \tau]), \tag{5.5}$$

where $\mathbf{P}(\text{IWC}[r_e, \tau])$ is the *pdf* of cirrus IWC. The IWC for the size distributions and optical depths used in the modelling study is given in figure 5.11. For a given pair of r_e and τ there is a unique value of IWC.

For the purposes of this thesis the external constraint on IWC is assumed to be a temperature dependent Gaussian *pdf* (with respect to the logarithm of IWC), derived from the

---

[4]This is a clear example of the problem of multiple solutions in the cirrus retrieval problem. The reason for this is the complex behaviour of SEI-1 with respect to optical thickness and effective radius, which is influenced by the size distributions assumed (Fu, 1996) in which the total number density varies significantly with the effective radius.
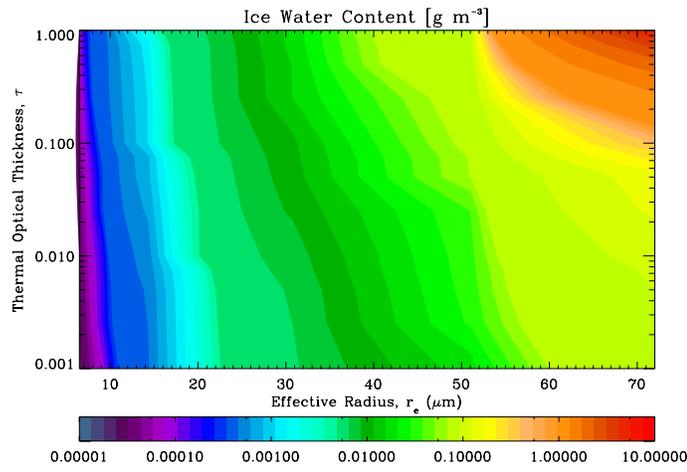
Figure 5.11: The cirrus ice water content (IWC) as a function of infrared optical thickness and the effective radius for the 9 size distributions derived from Fu (1996).
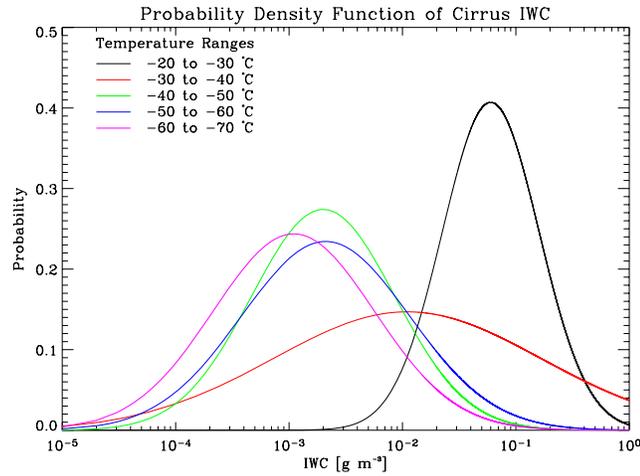
Figure 5.12: The *pdf* of cirrus IWC, **P**(IWC), to be used as the external contraints on IWC in the initial estimate.

in situ measurement campaign over Kwajalein, Marshall Islands (Heymsfield and McFarquhar, 1996; McFarquhar et al., 2000) as discussed in section 1.5.3 (see also figure 1.8). The Kwajalein *in situ* results were chosen as the cirrus sampled was more typical of high thin tropical cirrus than the other campaigns discussed in section 1.5.3. Figure 5.12 gives the *pdf* of cirrus IWC, **P**(IWC), derived from the Kwajalein data, to be used as the external contraints. The *pdf*s for five temperature ranges are given in figure 5.12. The temperature range to be used in the initial estimate is determined from the rough cloud top height and the corresponding temperature in the climatology temperature profile input to the modelling described in the previous section.

The pair of r_e and τ selected as the initial estimate will have the maximum probability for the corresponding IWC. As **P**(IWC$[r_e, \tau]$) is Gaussian, it is possible that two values of IWC could have the same probability (if they are equidistant from the mean IWC). If such a case

occurs, or the probability of all possible IWC's is zero, then the r_e and τ corresponding to the smallest IWC is selected, as high thin cirrus with low IWC is expected in the tropical regions.

Once r_e and τ have been selected, the cloud top height is then re-estimated by using the modelling results (given r_e and τ) of the ratio of the threshold radiance of the cloud top event limb scan and the limb scan immediately below. The radiance ratio from the MIPAS measurements is calculated and the cloud top height is found by using a linear interpolation on the modelled ratio's and corresponding cloud top heights. This will give the cloud top height assuming the nominal tangent heights, and the actual cloud top height is estimated by using the Level-1B MIPAS tangent height and assuming the same relative difference between cloud top height and tangent height.

Finally, to generate the input to the retrieval, τ must be transformed to the cloud state parameters N and C_{depth}. Changes in N and C_{depth} have a very similar effect on the model output, so the parameters are effectively correlated. In order to determine N and C_{depth}, external constraints must be used. As total number density depends on the size distribution and r_e, a very simple constraint is used to determine C_{depth} and N is determined from

$$N = \frac{\tau N_{\text{orig}}}{C_{\text{depth}} \beta_{\text{ext}}(r_e, N_{\text{orig}})} \qquad (5.6)$$

where N_{orig} is the total number density of the size distributions used to generate the modelling results (derived from Fu, 1996) and $\beta_{\text{ext}}(r_e, N_{\text{orig}})$ is the extinction coefficient at 12 μm, calculated for the given value of r_e and the original size distribution.

The cloud depth prior constraint is derived by fitting cloud depth as a Gaussian function dependent on the cloud top height. The maximum cloud depth is assumed to be the mean cirrus cloud depth given in table 1.2 (i.e. 1.5 km, from climatology), and this is expected to occur at the mean cirrus cloud top height also given in table 1.2 (i.e. 9 km). A full width at half maximum of 14 km is assumed. The estimated cloud depth function is given in figure 5.13.

The method of determining N and C_{depth} from τ is simplistic, but due to the non-linearity of McClouds_FM it would be too time consuming (and complex) to attempt a more accurate estimation of N and C_{depth} by considering the parameters separately in the modelling (and it still would not ensure that a false minimum wouldn't be found). However, it would be useful in future work to investigate the dependence of the modelled parameters on N and C_{depth} separately. Although the derived values may not be as close to the true solution as would be hoped, it is ultimately the job of the retrieval to determine the best estimate of the cloud state parameters. The values obtained are a good first guess, as they are typical of high thin cirrus (from external constraints) and the MIPAS scattering and detection

indices and the modelled indices are consistent.
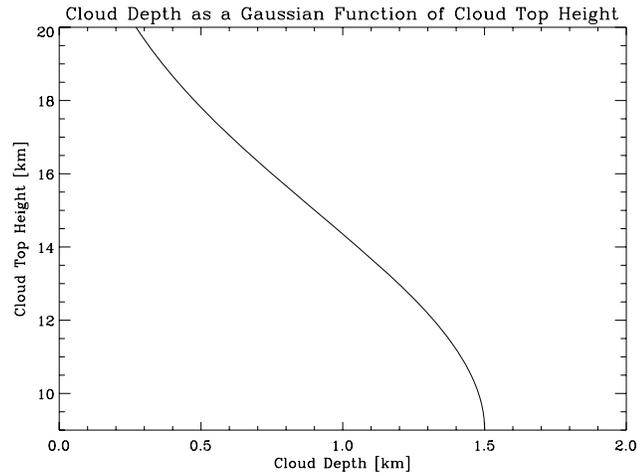
Figure 5.13: Cloud depth as a Gaussian function of cloud top height (with FWHM=14 km), used for the initial estimation of cloud depth.

As the initial estimate method is a complex process, the procedure described above is summarised as a flow diagram in figure 5.14.

5.4 Cloud Statistics from MIPAS using Estimation Methods

The initial estimate algorithm was applied to the MIPAS data from August 2003 for latitudes in the range 30 °S to 30 °N ($\sim$ 9000 profiles) and only considering nominal tangent heights $\geq$ 9 km and clouds detected above $\gtrsim$ 10.5 km.

The principle component cloud test (see chapter 3) was first applied to the dataset. The results from the cloud detection are presented in figure 5.15, which gives the percentage of MIPAS profiles per bin, of a 5° latitude by 10° longitude grid, in which at least one limb scan above $\gtrsim$ 10.5 km has cloud present. This statistic can simply be thought of as a high cirrus ($\gtrsim$ 10.5 km) frequency of occurence. High cirrus was detected in $\sim$ 45 % of all the MIPAS profiles of interest. The enhanced ($>$ 50 %) areas of high cirrus occur in the expected regions:

- Over central Africa spanning to the coast of Western sub-Saharan Africa (with a slight coastal high possibly associated with biomass burning).

- Over the East Indian Ocean and the Indonesia and Malaysia region.

- Over the Central Americas spanning out to the mid-Pacific between aproximately 10 to 20 °N.

INPUT STEPS OUTPUT

START

Perform cloud detection on each limb scan for all tangent heights less than or equal to 30 km

Input MIPAS profile of limb scans: 0,..,i,..,16 with nominal tangent heights: TH_0,...,TH_i,...,TH_{16} and Level 1B tangent heights: T_0,...,T_i,..,T_{16}

The cloud top event limb scna is the highest limb scan that cloud is detected at. Say this is limb scan i with level 1B tangent height T_i

ROUGH ESTIMATE
The cloud top height is in the range T_i +/− 1.5 km

ASSUME
The FOV of limb scan i−1 is cloud filled

Determine the value of SEI−1 for the limb scan i−1

STORE
SEI−1 for limb scan i−1

Compare SEI−1 for scan i−1to the SEI−1 modelled results for the range TH_i +/− 1.5 km

Find all possible values of effective radius, Re and optical depth, Tau, where the modelled SEI−1 matches the measured value. Then convert the Re and Tau values to IWC (see figure 5.11)

STORE
All possible Re and Tau values and corresponding unique IWC

The a priori pdf of IWC i.e. P(IWC) taken from in situ studies

From the set of possible solutions, find the IWC with maximum probability. The corresponding Re and Tau values give the initial estimate.

INITIAL ESTIMATE
Effective radius, Re Optical depth, Tau

Refine cloud top height estimate to within 3km for given Re and Tau

Calculate the ratio of the window threshold radiance in limb scan i to that in limb scan i−1. Compare this ratio to the modelled ratio (for equivalent nominal tangent heights) as a function of cloud top height

Using a linear interpolation

CALCULATE
Relative cloud top height w.r.t nominal tangent heights, CT

ASSUME
The relative cloud top height holds for Level 1B TH

Calculate cloud top height, Ctop using:
Ctop= TH_i + CT

INITIAL ESTIMATE
Cloud top height, Ctop

The assumed a priori function of Cdepth dependent on Ctop

Estimate cloud depth, Cdepth assuming the a priori relationship to Ctop as given in figure 5.13.

INITIAL ESTIMATE
Cloud top height, Ctop

From the estimates of Re, Tau, Ctop and Cdepth, the total number density can be calculated using equation 5.6.

INITIAL ESTIMATE
Total number density, N

END

OUTPUT

INITAL ESTIMATE
CLOUD STATE
[Re, N, Ctop, Cdepth]

Figure 5.14: Flow diagram of the initial estimate algorithm. The algorithm is applied under the assumption that the modelled relationships between the cloud parameters and the scattering and detection measures are not strongly dependent on viewing tangent height. Hence, the modelled relationships for the nominal tangent heights are assumed to hold at the actual level 1B tangent heights, i.e. ± 1.5 km of the nominal tangent height.

- A high between approximately 10 to 20 °N over the Pacific to the East of the Phillippines.

There is also a prominent, near 100 % band of cirrus occurrence between approximately 5 °S and 15 °S over the mid-Eastern Indian Ocean.

The results from applying the initial estimate algorithm to the cloudy profiles and averaging over the 5° latitude by 10° longitude grid, is given in figure 5.16 for the four cloud state parameters. The initial estimate fields of ice water content and optical thickness (both infrared and visible[5]) is also given in figures 5.17 and 5.18 respectively.

What is immediately obvious from each of the initital estimate fields for all the cloud properties is that there is a great deal of variability in the average cloud properties in the regions of low frequency of occurrence, e.g. 20 to 30 °S, particularly over the Pacific and Atlantic. Whether this is an indication of a physical characteristic of the cirrus types which are more likely to form in those regions or if it is simply a statistical artefact from averaging over less cloud measurements cannot be concluded, but it is more likely that the variability is due to statistical effects.

In the regions of enhanced high cirrus frequency (listed above) the universal trend in the cloud properties is fairly uniform, with $r_e \approx 15$ μm, $C_{\text{top}} \approx 15$ km, $0.4 \lesssim N \lesssim 0.7$, $C_{\text{depth}} \approx 0.75$ km, IWC≈ 0.1 g m^{-3} and $\tau_{\text{ir}} \approx 0.075$ ($\tau_{\text{vis}} \approx 0.2$). Little can be inferred by the cloud depth values as they depend only on the estimated cloud top height. The IWC values appear to peak at the mean of the external constraint *pdf*, but it is not clear whether the external constraint is biasing the IWC or if the external *pdf* and measurements are in agreement. The variability in the regions of low cloud frequency suggest that the IWC is not being biased, and hence the external constraint is representative of the average cirrus values. There are several subtle differences between the high frequency regions. In particular, over the Central Americas there is a slightly lower average cloud top height of ~ 13 km compared to over Central Africa (~ 16 km) and South-East Asia (~ 15 km). Interestingly this is correlated with a slightly lower effective radius in the high cloud top regions of $\lesssim 10$ μm, compared to an average effective radius ~ 15 μm over the Central Americas. There is also a slight corresponding trend in the total number density which tends to be lower over the Central Americas. There is very little variation between regions in optical thickness, C_{depth} or IWC. Intuitively, smaller effective radius would be expected at higher cloud top heights as smaller, pristine, faceted crystals are expected to be more common at the colder temperatures higher in the troposphere. The band of very high cirrus frequency between 5 and 15 °S over the mid-Eastern Indian Ocean and Northern

---

[5]The visible optical thickness was calculated by multiplying the infrared optical thickness by the ratio of the visible extinction coefficient over the infrared extinction coefficient (for a given r_e and N). The optical properties in the visible for the same size distributions, i.e. those derived from Fu (1996), are calculated by integrating the individual crystal optical properties given by Yang et al. (2000) over the size distributions.

Australia exhibits similar behaviour. More interestingly, the optical depth in this band is noticeably lower than for the other high frequency areas (see figures 5.18(a) and 5.18(b)) with a corresponding increase in cloud top height ~ 16 km which is near the tropopause in this region. From figure 5.18(b) the visible optical thickness τ_{vis} ranges from ~ 0.03 to 0.075, which is near sub-visual (i.e. $\tau_{\mathrm{vis}} < 0.03$). Considering the possible error on the initial estimate and the effect of averaging over a month, with some occurence of thicker, lower cirrus likely, this result is strongly suggestive that subvisible cirrus is ubiquitous in the East Indian region near Indonesia.
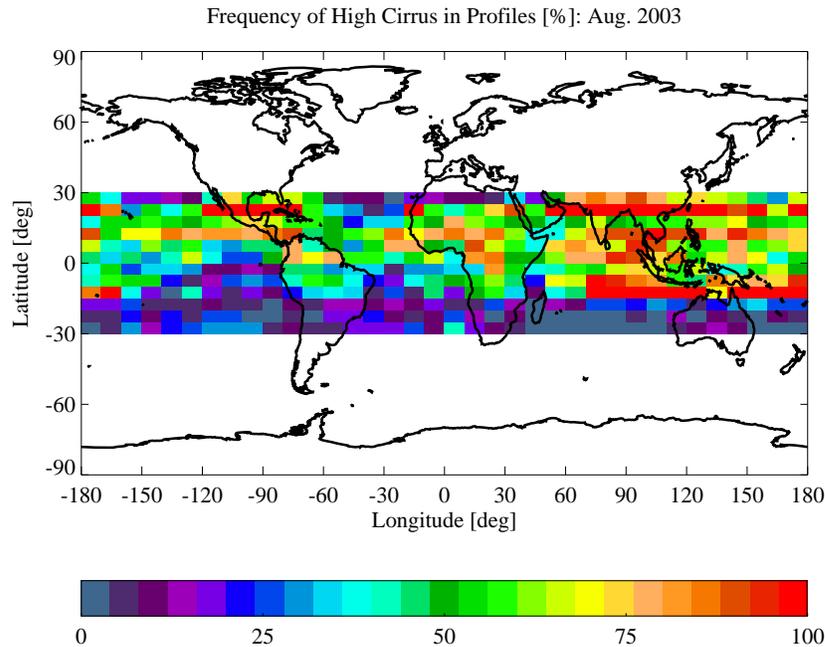
Figure 5.15: The percentage of MIPAS profiles in a 5° latitude by 10° longitude grid box for August 2003 in which cloud is present in at least one limb scan $\gtrsim 10.5$ km. The principle component cloud test (see chapter 5) was applied to detect cloud.

5.4.1 Discussion

From comparing the initial estimate results to the expected climatology cirrus values discussed in chapter 1, it appears that the initial estimates return realistic cirrus properties in the tropical region. No external validation of the initial estimate results for August 2003 was carried out (against existing climatology) because of the lack of error analysis and the difficulty in quantify the error of such an ad-hoc method. The initial estimates can however, be considered to be indicative of what the full retrieval will return (if the solution is close to the first guess). Further, when analysing the initial estimate results given in figures 5.16 to 5.18, it must be taken into consideration that the estimation method fits the MIPAS measurements under the phyiscal assumptions about the cloud given in section 4.2. This may
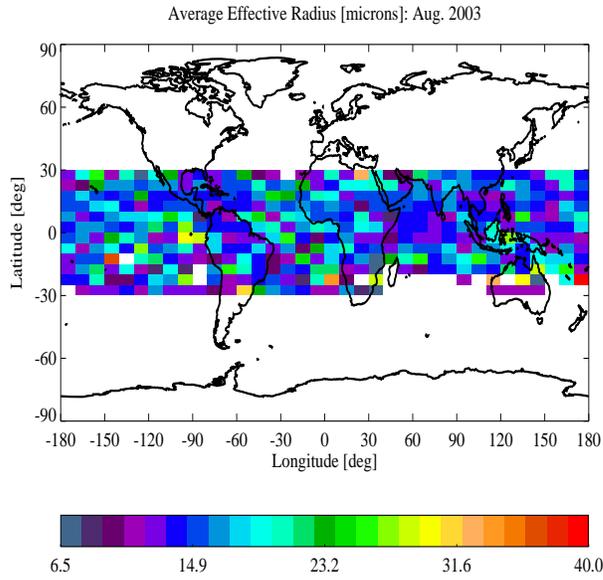
(a) Average initial estimate of cirrus effective radius, r_e (μm).
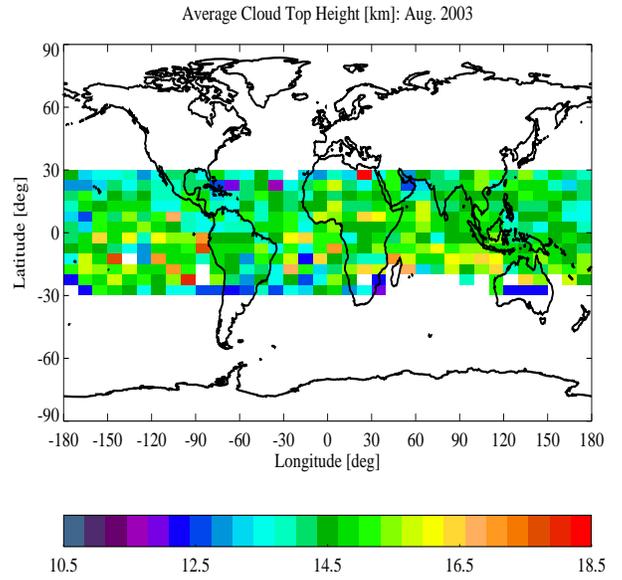
(b) Average initial estimate of cirrus cloud top height, C_{top} (km).
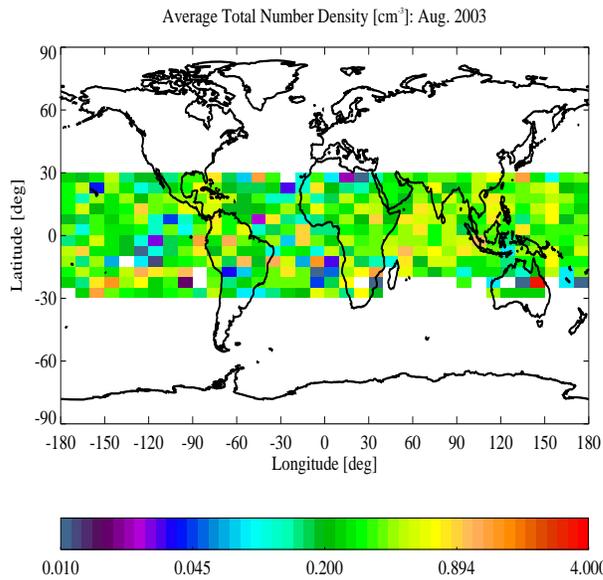
(c) Average initial estimate of cirrus total number density, N (cm^{-3}).
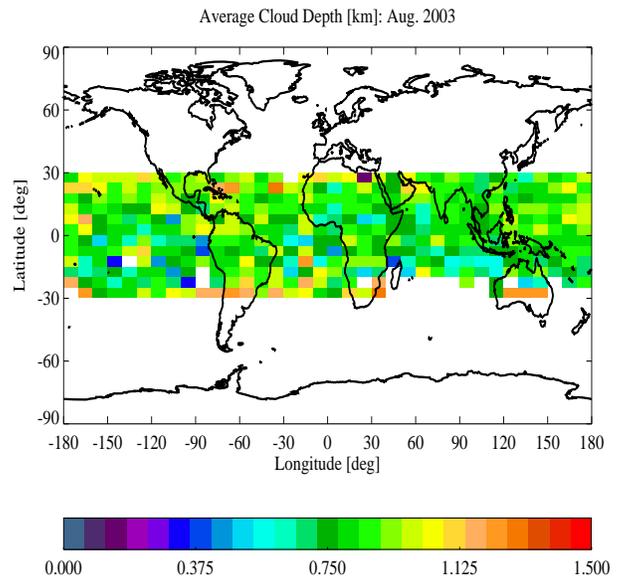
(d) Average initial estimate of cirrus vertical cloud depth, C_{depth} (km).

Figure 5.16: The initial estimates of cloud state parameters derived from MIPAS profiles in which cloud is detected above $\gtrsim$ 10.5 km and averaged over a 5° latitude by 10° longitude grid for August 2003.

Average Ice Water Content [g m⁻³]: Aug. 2003
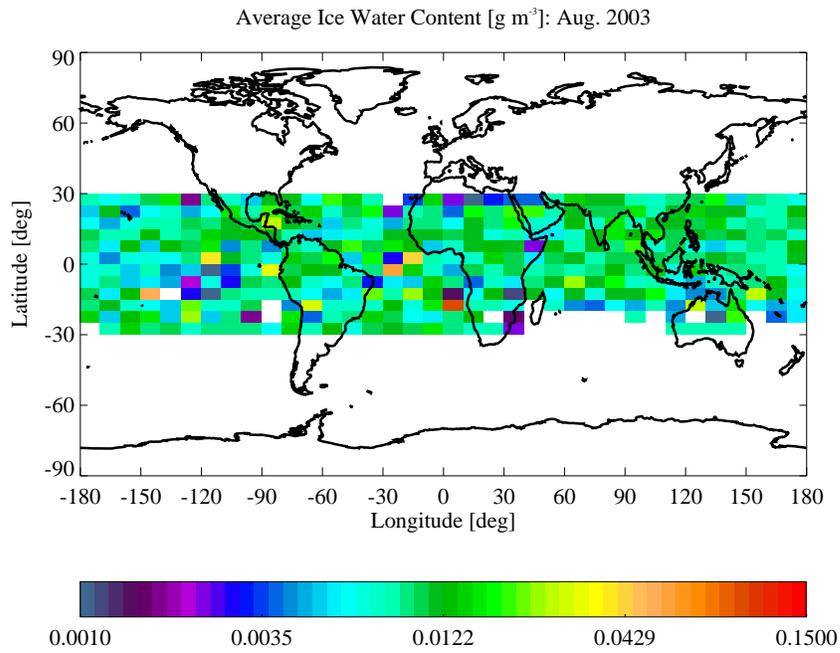
Figure 5.17: The initial estimates of cirrus ice water content (IWC in g m^{-3}) derived from MIPAS profiles in which cloud is detected above $\gtrsim$ 10.5 km and averaged over a 5° latitude by 10° longitude grid for August 2003.
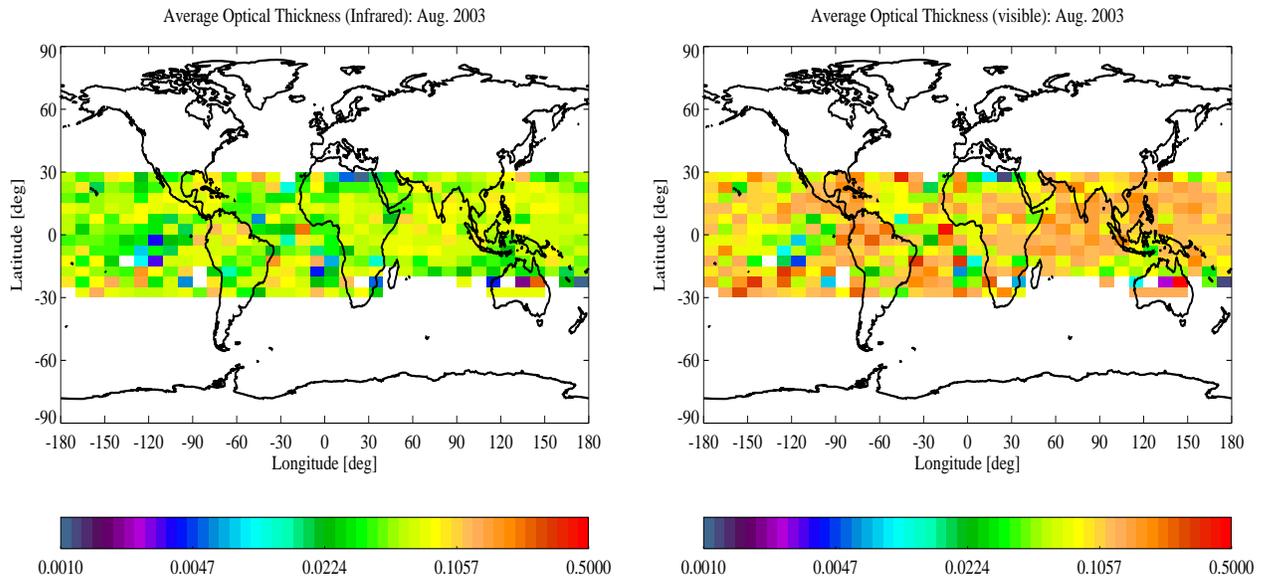
(a) Average initial estimate of infrared (thermal) optical thickness of cirrus (at λ=12 μm).

(b) Average initial estimate of visible optical thickness of cirrus (at λ=0.67 μm).

Figure 5.18: The initial estimates of cirrus optical thickness derived from MIPAS profiles in which cloud is detected above $\gtrsim$ 10.5 km and averaged over a 5° latitude by 10° longitude grid for August 2003.

partly explain the absence of any high average optical thickness regions indicative of deep convection. The results will also be biased by assuming the IWC *pdf* to be representative of optically thin cirrus. However, only cloud top heights above 10.5 km are considered and above this height convective anvils often form into laminar outflow. In addition, convective anvils often form in conjunction with a thin near-tropopause cirrus layer above it (Garrett et al., 2005). In these cases, the thin upper layer will be detected first in a MIPAS profile and will tend to dominate the measurements, due to scattering effects. Thus, an area of interest for future work would be to model the cloud under the same assumptions about the upper cirrus layer and assume a lower deck of black-body cloud.

In short, the fast methods developed in this chapter for detection of scattering effects and the characterisation of the cloud using the detection measures are not only useful to guide the first guess of the retrieval but can also be used to make a qualitative study of large MIPAS datasets without the high computational overhead of the full retrieval. The full retrieval results with error budget will be required for input to climate models and quantitative studies, such as cirrus radiation budget calculations. However, the qualitative initial estimate results would be ideally suited to studies dependent on analysing a high volume of data, such as Massie et al. (2000). Massie et al. (2000) studied the effect of the 1997 El Niño on the distribution of upper tropospheric cirrus by studying the HALOE extinction data from 1993 to 1998. The MIPAS data would be ideally suited to carry out a similar qualitative study over the measurement period to date, to investigate the moderate ENSO (El Niño/Southern Oscillation) of 2002/2003. The initial estimation method would allow the entire MIPAS dataset to be processed rapidly and the distribution of upper tropospheric cirrus[6] to be investigated as well as the relative differences in cloud top height, optical thickness and effective radius. Such a study is a high priority for future work.

---

[6]The spatial resolution investigated could be considerably higher than the 5°by 10° used in section 5.4, as seasonal or annual statistical means would be considered.

Chapter 6

Retrieval Theory for Atmospheric Sounding

6.1 Introduction

The theory of retrieval of atmospheric state parameters from remote measurements is well
understood, see e.g., Rodgers (1976), Carlotti (1988), Marks and Rodgers (1993), Rodgers
(2000).

The aim of this chapter is to give a succinct overview of retrieval theory and define some
of the key concepts, following Rodgers (2000), in order to familiarise the reader with the
background theory required to build a numerical retrieval method and thus required to de-
scribe its construction. The method developed for retrieving cloud properties from MIPAS
measurements is then detailed in the following chapter.

6.2 Basic Retrieval Theory

Following the general introduction to the concepts of retrieval theory in section 2.6.2, con-
sider again the measurement of m quantities (e.g. radiances) assembled into a measurement
vector $\mathbf{y}$, and the n parameters describing the state space (i.e. the atmosphere in the case of
remote sounding) assembled into the state vector $\mathbf{x}$. If the n parameters in the state vector
are unknown (which is usually the case in remote sounding), then the aim of a retrieval is
to obtain the best estimate of $\mathbf{x}$ given $\mathbf{y}$.

The need to implement a retrieval often arises when the physics of the forward function
are too complex to evaluate explicitly, thus precluding a simple inversion of the forward
function. In such cases the forward model which approximates the physics can be used to
infer the unknown state vector from the known measurement.

Following section 2.6.2, the retrieved state can be related to the true state by substituting

equation 2.4 into equation 2.5 to obtain

$$\hat{\mathbf{x}} = \mathbf{R}(\mathbf{F}(\mathbf{x}) + \epsilon, \mathbf{x_a}, \mathbf{c}) . \tag{6.1}$$

However, most practical inverse problems are nonlinear due to the nature of the forward model, and can be ill-conditioned or ill-posed. It is often necessary to linearise nonlinear problems to describe their behaviour or the information contained in the measurement.

6.2.1   Dealing with Measurement Error: A Bayesian Approach

A key issue in solving any retrieval problem is the proper treatment of experimental error or 'noise' which is present in all real measurements. For these purposes it is useful to describe experimental error in terms of probability density functions (*pdf's*) following a Bayesian approach as in Rodgers, 2000. The Bayesian approach is introduced to the general inverse problem to relate the measurement *pdf* to the *pdf* of the state.

For vector measurements, the measurement *pdf*, $\mathbf{P}(\mathbf{y})$, is defined over measurement space such that the probability that the true value of the measurement lies in a multidimensional interval $(\mathbf{y}, \mathbf{y} + \mathrm{d}\mathbf{y})$ in measurement space is given by $\mathbf{P}(\mathbf{y})\mathrm{d}\mathbf{y}$. For a good approximation of experimental error, the form of $\mathbf{P}(\mathbf{y})$ is taken to be Gaussian and can thus be written as

$$\mathbf{P}(\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^m |\mathbf{S_y}|^{1/2}}} \exp\left\{ -\frac{1}{2}(\mathbf{y} - \bar{\mathbf{y}})^T \mathbf{S_y}^{-1}(\mathbf{y} - \bar{\mathbf{y}}) \right\} \tag{6.2}$$

where $\mathbf{S_y}$ is the covariance matrix of $\mathbf{y}$ and must be nonsingular. The role of the covariance matrix of $\mathbf{y}$ in the vector Gaussian distribution (equation 6.2) is analogous to the role of the variance in a scalar Gaussian distribution. The covariance matrix $\mathbf{S_y}$ is defined such that the ij^{th} element, $S_{y,ij}$, is the covariance of the ith and j^{th} elements of $\mathbf{y}$ i.e.

$$S_{y,ij} = \text{cov}(y_i, y_j) = \varepsilon\{(y_i - \bar{y}_i)(y_j - \bar{y}_j)\}. \tag{6.3}$$

The diagonal elements of $\mathbf{S_y}$ are in fact the variances of the individual elements of $\mathbf{y}$, and the off diagonal elements give a measure of how strongly two different elements of $\mathbf{y}$ are correlated.

The act of measurement maps the state into the measurement space according to the forward function, which can be considered the mapping according to the forward model for the purposes of the inverse problem. However, in the presence of measurement error a *point* in state space maps into a *region* in measurement space determined by the probability density function of ϵ. Thus, the conceptual mapping diagram of the forward model will differ from that of the forward function in figure 2.9 and can be represented as in figure 6.1.
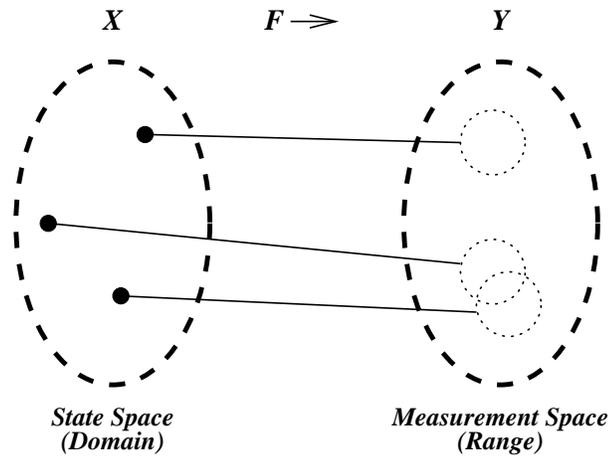
Figure 6.1: The forward model is equivalent to a mapping, where the domain is the state space and the range is the measurement space. In the presence of measurement error a point in state space maps into a region in measurement space.

The conditional *pdf* of $\mathbf{y}$ given $\mathbf{x}$, $\mathbf{P}(\mathbf{y}|\mathbf{x})$, is determined by the *pdf* of ϵ, and $\mathbf{P}(\mathbf{y}|\mathbf{x})\mathrm{d}\mathbf{y}$ is the probability that $\mathbf{y}$ is in the interval $(\mathbf{y}, \mathbf{y}+\mathrm{d}\mathbf{y})$ when $\mathbf{x}$ has a given value in state space. The measurement error is known only statistically and ϵ is assumed to be Gaussian about any point in measurement space, thus $\varepsilon\{\epsilon\} = 0$ and following from equation 6.2 the *pdf* of ϵ can be written as

$$-2\ln\mathbf{P}(\epsilon) = \epsilon^T\mathbf{S}_\epsilon^{-1}\epsilon + c_1, \tag{6.4}$$

where c_1 is a constant[1] and $\mathbf{S}_\epsilon$ is the measurement error covariance. By extension, if forward model error is negligible, then $\mathbf{P}(\mathbf{y}|\mathbf{x})$ can be expressed as,

$$-2\ln\mathbf{P}(\mathbf{y}|\mathbf{x}) = (\mathbf{y} - \mathbf{F}(\mathbf{x}))^T\mathbf{S}_\epsilon^{-1}(\mathbf{y} - \mathbf{F}(\mathbf{x})) + c_2 \tag{6.5}$$

were c_2 is a constant. The measurement error covariance is defined (for Gaussian statistics) as

$$\mathbf{S}_\epsilon = \varepsilon\{\epsilon\epsilon^T\} = \varepsilon\{(\mathbf{y} - \mathbf{F}(\mathbf{x}))(\mathbf{y} - \mathbf{F}(\mathbf{x}))^T\}, \tag{6.6}$$

where the diagonal elements of $\mathbf{S}_\epsilon$ are the variances of the individual elements of $\mathbf{y}$. Considering the inverse problem, if $\mathbf{y}$ is a given measurement, it could be the result of a mapping from anywhere in a region of state space described by some *pdf*, rather than from a single point. It is this *pdf* that is important for solving the inverse problem, and is known as the *posterior pdf*. The *posterior pdf* is denoted $\mathbf{P}(\mathbf{x}|\mathbf{y})$ and is known as the conditional *pdf* of $\mathbf{x}$ given $\mathbf{y}$, where $\mathbf{P}(\mathbf{x}|\mathbf{y})\mathrm{d}\mathbf{x}$ gives the probability that $\mathbf{x}$ lies in $(\mathbf{x}, \mathbf{x} + \mathrm{d}\mathbf{x})$ when $\mathbf{y}$ has a given value in measurement space.

Bayes' theorem describes the relationship between the two different conditional *pdf*'s, such

---

[1] $c_1 = 2\ln\left[\sqrt{(2\pi)^m}|\mathbf{S}_\epsilon|^{1/2}\right]$ and is constant because the determinant of $\mathbf{S}_\epsilon$ is a constant.

that

$$\mathbf{P}(\mathbf{x}|\mathbf{y}) = \frac{\mathbf{P}(\mathbf{y}|\mathbf{x})\mathbf{P}(\mathbf{x})}{\mathbf{P}(\mathbf{y})}, \tag{6.7}$$

where $\mathbf{P}(\mathbf{x})$ is the prior *pdf* of the state $\mathbf{x}$ (before the measurement is made), which is normalised, and $\mathbf{P}(\mathbf{y})$ is the prior *pdf* of the measurment (before it is made). $\mathbf{P}(\mathbf{y})$ is generally a normalising factor and is often not needed. The Bayesian approach to the inverse problem is basically a way of taking imperfect prior knowledge about the state and using the measurement to update this knowledge, taking into account experimental error. Any prior knowledge of the state is very useful in the inverse problem, and can be used to constrain the solution. If Gaussian statistics are used to describe prior knowledge of the state (although this is not very realistic), then $\mathbf{P}(\mathbf{x})$ is given by

$$-2\ln\mathbf{P}(\mathbf{x}) = (\mathbf{x} - \mathbf{x_a})^T\mathbf{S_a}^{-1}(\mathbf{x} - \mathbf{x_a}) + c_3, \tag{6.8}$$

where c_3 is a constant, $\mathbf{x_a}$ is the *a priori* value of $\mathbf{x}$, and $\mathbf{S_a}$ is the associated covariance matrix

$$\mathbf{S_a} = \varepsilon\{(\mathbf{x} - \mathbf{x_a})(\mathbf{x} - \mathbf{x_a})^T\}. \tag{6.9}$$

Using Bayes' theorem (equation 6.7) and equations 6.5 and 6.8 the *posterior pdf* is obtained:

$$-2\ln\mathbf{P}(\mathbf{x}|\mathbf{y}) = (\mathbf{y} - \mathbf{F}(\mathbf{x}))^T\mathbf{S_\epsilon}^{-1}(\mathbf{y} - \mathbf{F}(\mathbf{x})) + (\mathbf{x} - \mathbf{x_a})^T\mathbf{S_a}^{-1}(\mathbf{x} - \mathbf{x_a}) + c_4, \tag{6.10}$$

where c_4 is a constant.

The Bayesian approach itself is a general formulation of inverse methods, and characterises the class of possible solutions through assigning a *pdf* in state space.

6.2.2 Solving the Inverse Problem

The inverse problem generally does not have an exact solution, so the best estimate or optimal solution is selected from the class of possible solutions. In addition, the inverse problem is often ill-posed, and can produce unstable solutions with large errors. In order to constrain the problem to one that is well posed a solution is found by using *a priori* knowledge of the state. Even if the knowledge is imperfect, and the estimate of the state is a poor one with large errors, it can still improve the stability of the solution. Therefore, in this thesis, *a priori* information is always used when solving the retrieval problem.

There are a number of methods that can be used to select a state from an ensemble of possible solutions described by a probability function. The best solution, $\hat{\mathbf{x}}$ can be selected as the 'expected' state i.e.

$$\hat{\mathbf{x}} = \varepsilon\{\mathbf{x}\} = \int \mathbf{x}\mathbf{P}(\mathbf{x}|\mathbf{y})d\mathbf{x}, \tag{6.11}$$

or the 'most probable' state, where $\hat{\mathbf{x}}$ is such that

$$\left.\frac{d\mathbf{P}(\mathbf{x}|\mathbf{y})}{d\mathbf{x}}\right|_{\hat{\mathbf{x}}} = 0. \tag{6.12}$$

If the statistics are assumed to be Gaussian then the most probable state is the same as the expected state. This state is the one for which $\mathbf{P}(\mathbf{x}|\mathbf{y})$ is maximum, and is known as the *Maximum a Posteriori* (MAP) solution. Under the assumption of Gaussian statistics, $\mathbf{P}(\mathbf{x}|\mathbf{y})$ is given as in equation 6.10. Solving equation 6.12 is equivalent to setting the derivative of equation 6.10 to zero:

$$\nabla_{\mathbf{x}}\{-2\ln\mathbf{P}(\mathbf{x}|\mathbf{y})\} = -[\nabla_{\mathbf{x}}\mathbf{F}(\mathbf{x})]^T\mathbf{S}_\epsilon^{-1}[\mathbf{y} - \mathbf{F}(\mathbf{x})] + \mathbf{S}_a^{-1}[\mathbf{x} - \mathbf{x}_a] = 0. \tag{6.13}$$

Note that $\nabla_{\mathbf{x}}\mathbf{F}(\mathbf{x})$ is an $n \times m$ matrix as the gradient of a vector-valued function is a matrix-valued function. Giving the equation for the solution $\hat{\mathbf{x}}$ as:

$$-[\nabla_{\mathbf{x}}\mathbf{F}(\hat{\mathbf{x}})]^T\mathbf{S}_\epsilon^{-1}[\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}})] + \mathbf{S}_a^{-1}[\hat{\mathbf{x}} - \mathbf{x}_a] = 0. \tag{6.14}$$

It is convenient to define the *joint cost function*

$$\chi^2 = (\mathbf{x} - \mathbf{x_a})\mathbf{S_a}^{-1}(\mathbf{x} - \mathbf{x_a})^T + \epsilon^T\mathbf{S}_\epsilon^{-1}\epsilon, \tag{6.15}$$

where $\epsilon = \mathbf{y} - \mathbf{F}(\mathbf{x})$. The join cost function is equivalent to the posterior *pdf*, i.e. the right hand side of equation 6.10. Therefore, the MAP solution is found by minimising the cost function which is known as the 'optimal estimation' retrieval.

6.2.3 Weighting Function Matrix

The retrieval method is more conveniently dealt with if the measurement is considered a linear problem. Equation 2.4 can be represented as a linear system of equations if the forward model is linearised about some reference state $\mathbf{x_0}$, i.e.

$$\mathbf{y} - \mathbf{F}(\mathbf{x_0}) = \left.\frac{\partial\mathbf{F}}{\partial\mathbf{x}}\right|_{\mathbf{x_0}}(\mathbf{x} - \mathbf{x_0}) + \epsilon, \tag{6.16}$$

where $\left.\frac{\partial\mathbf{F}}{\partial\mathbf{x}}\right|_{\mathbf{x_0}}$ is the $m \times n$ Jacobian matrix

$$\begin{bmatrix} \left.\frac{\partial F_1}{\partial x_1}\right|_{\mathbf{x_0}} & \cdots & \left.\frac{\partial F_1}{\partial x_n}\right|_{\mathbf{x_0}} \\ \vdots & \ddots & \vdots \\ \left.\frac{\partial F_m}{\partial x_1}\right|_{\mathbf{x_0}} & \cdots & \left.\frac{\partial F_m}{\partial x_n}\right|_{\mathbf{x_0}} \end{bmatrix}$$

where $\left.\frac{\partial F_i}{\partial x_j}\right|_{\mathbf{x_0}}$ is the first order partial derivative of the ith element of the forward model $F_i(\mathbf{x})$ with respect to the jth component of the state vector x_j evaluated at $\mathbf{x_0}$. The matrix $\partial \mathbf{F}/\partial \mathbf{x}$ is sometimes simply called "the Jacobian" (Simon and Blume, 1994; Rodgers, 2000), but this can be somewhat confusing as the determinant of the Jacobian matrix is often called "the Jacobian" as well. In the context of atmospheric remote sounding $\partial \mathbf{F}/\partial \mathbf{x}$ is more appropriately known as the *weighting function matrix* $\mathbf{K}$, and is defined such that each element $K_{ij} = \partial F_i/\partial x_j$. The weighting function matrix represents the sensitivity of the measurement (if equation 2.4 holds) to changes in the state vector. Thus, equation 6.16 can be re-written as

$$\mathbf{y} - \mathbf{F}(\mathbf{x_0}) = \mathbf{K}|_{\mathbf{x_0}}(\mathbf{x} - \mathbf{x_0}) + \epsilon, \tag{6.17}$$

where $\mathbf{K}|_{\mathbf{x_0}} = \nabla_{\mathbf{x}}\mathbf{F}(\mathbf{x_0})$ is the weighting function matrix evaluated at $\mathbf{x_0}$.

Under a transformation of coordinates in both state space and measurement space, so that the origin of state space becomes $\mathbf{x_0}$ and the origin of measurement space becomes $\mathbf{F}(\mathbf{x_0})$, Equation 6.17 can be re-written as a simple linear problem

$$\mathbf{y} = \mathbf{K}\mathbf{x} + \epsilon \tag{6.18}$$

where $\mathbf{K} = \mathbf{K}|_{\mathbf{x_0}}$ and is considered constant (as the problem is linearised) in the neighbourhood of $\mathbf{x_0}$.

6.2.4   Solving Linear Problems

If the retrieval problem is linear (or has been linearised), then the measurement can be expressed by equation 6.18. If Gaussian error statistics are assumed (and forward model error is negligible), then by substituting equation 6.18 into equation 6.10 the posterior *pdf* is obtained:

$$-2\ln \mathbf{P}(\mathbf{x}|\mathbf{y}) = (\mathbf{y} - \mathbf{K}\mathbf{x})^T\mathbf{S}_\epsilon^{-1}(\mathbf{y} - \mathbf{K}\mathbf{x}) + (\mathbf{x} - \mathbf{x_a})^T\mathbf{S_a}^{-1}(\mathbf{x} - \mathbf{x_a}) + a_1, \tag{6.19}$$

where a_1 is a constant and the measurement error covariance is defined as

$$\mathbf{S}_\epsilon = \varepsilon\{\epsilon\epsilon^T\} = \varepsilon\{(\mathbf{y} - \mathbf{K}\mathbf{x})(\mathbf{y} - \mathbf{K}\mathbf{x})^T\}. \tag{6.20}$$

Equation 6.19 can then be written as

$$-2\ln \mathbf{P}(\mathbf{x}|\mathbf{y}) = (\mathbf{x} - \hat{\mathbf{x}})^T\hat{\mathbf{S}}^{-1}(\mathbf{x} - \hat{\mathbf{x}}) + a_2, \tag{6.21}$$

where a_2 is a constant, $\hat{\mathbf{x}}$ is the expectation value and $\hat{\mathbf{S}}$ the corresponding covariance.

Thus, the *posterior pdf* is also a Gaussian distribution with covariance

$$\hat{\mathbf{S}} = (\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} + \mathbf{S_a}^{-1})^{-1}. \tag{6.22}$$

From equation 6.14, minimising the cost function (where $\epsilon = \mathbf{y} - \mathbf{Kx}$) in the linear case yields

$$\hat{\mathbf{x}} = (\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} + \mathbf{S_a}^{-1})^{-1}(\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{y} + \mathbf{S_a}^{-1} \mathbf{x}_a), \tag{6.23}$$

which is equivalent to setting the right hand sides of equations 6.19 and 6.21 equal to one another. Thus, equation 6.23 gives the optimal estimation solution for the linear problem, which can be re-written as

$$\hat{\mathbf{x}} - \mathbf{x}_a = \mathbf{G}(\mathbf{y} - \mathbf{Kx_a}) \tag{6.24}$$

with total covariance $\hat{\mathbf{S}}$ given in equation 6.22, and where $\mathbf{G}$ is the retrieval *gain matrix* (or *contribution function matrix*) which is conveniently defined as

$$\mathbf{G} = (\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} + \mathbf{S_a}^{-1})^{-1} \mathbf{K}^T \mathbf{S}_\epsilon^{-1} = \mathbf{S_a} \mathbf{K}^T (\mathbf{K} \mathbf{S_a} \mathbf{K}^T + \mathbf{S}_\epsilon)^{-1}. \tag{6.25}$$

The optimal estimation retrieval is therefore

$$\hat{\mathbf{x}} = \mathbf{Gy} + (\mathbf{I}_n - \mathbf{GK})\mathbf{x}_a. \tag{6.26}$$

Following from equation 6.22, the total covariance can be re-written (see, e.g. Rodgers, 1976) in a form corresponding to equation 6.26, such that

$$\hat{\mathbf{S}} = \mathbf{S_a} - \mathbf{S_a} \mathbf{K}^T (\mathbf{S}_\epsilon^{-1} + \mathbf{K} \mathbf{S_a}^{-1} \mathbf{K}^T)^{-1} \mathbf{K} \mathbf{S_a} \tag{6.27}$$

$$= (\mathbf{I} - \mathbf{GK})\mathbf{S_a}. \tag{6.28}$$

6.2.5 Retrieval Quality

As the optimal estimation retrieval is simply a best estimate of the solution from the class of all possible solutions, it is useful to have some measures of the 'quality' of the estimate and the confidence in the retrieval method.

The most simple and obvious measure is from the diagonal elements of the error covariance matrix $\hat{\mathbf{S}}$. As a first order approximation the square root of the diagonal elements gives the retrieval error on the corresponding retrieved parameter, i.e. the standard deviation assuming Gaussian statistics. This value can be compared to the corresponding diagonal element in the *a priori* error covariance matrix to obtain an estimate of how much the uncertainty (or information) on the parameter has improved.

There are several other useful quantities used in the analysis of the retrieval parameters and their errors, and the reader is directed toward Rodgers (2000) for a thorough discourse. However, two useful quantities to be mentioned here (which describe the linearised problem) are the *degrees of freedom for signal* and the *information content* of the measurement.

The number of degrees of freedom of a measurement gives the number of independent pieces of information that can be determined from the measurement. In principle, the number of degrees of freedom can be determined from the rank of the weighting function matrix, p. However, the presence of experimental noise may render a component worthless if the error is very large. Thus, the number of useful pieces of information may be less than p. The number of independent pieces of information in a measurement made to better than measurement error is the number of singular values[2] of $\mathbf{S}_\epsilon^{-1/2}\mathbf{K}\mathbf{S}_\mathbf{a}^{-1/2}$ which are greater than unity.

In the Gaussian linear case the most probable state is the one which minimises the joint cost function (see equation 6.15). At the minimum the expected value of χ^2 is equal to the number of degrees of freedom, or the number of measurements, m. The concept of degrees of freedom can be made more precise by considering how many degrees of freedom of the measurement are related to the signal, and how many are related to noise. For the Gaussian linear case these correspond to the two parts of equation 6.15, therefore one can define the degrees of freedom for signal as

$$d_s = \varepsilon\{(\mathbf{x} - \mathbf{x_a})\mathbf{S_a}^{-1}(\mathbf{x} - \mathbf{x_a})^T\}, \tag{6.29}$$

and the degrees of freedom for noise as

$$d_n = \varepsilon\{\epsilon^T\mathbf{S}_\epsilon^{-1}\epsilon\}. \tag{6.30}$$

The degrees of freedom for signal can be expressed explicitly as

$$d_s = \mathrm{tr}([\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K} + \mathbf{S_a}^{-1}]^{-1}\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K}) = \mathrm{tr}(\mathbf{GK}) \tag{6.31}$$

where "tr" denotes the trace of a matrix. This leads to the definition of the *averaging kernel matrix* $\mathbf{A} = \mathbf{GK}$. The degrees of freedom for signal will approach the maximum as the averaging kernel approaches the limit of the identity matrix $\mathbf{I_n}$.

The information content of a measurement can be defined qualitatively as the factor by which knowledge of a quantity is improved by making a measurement. Generally, in retrieval methods an information theoretic approach is taken and information content should

---

[2]For a matrix $\mathbf{A}$, the square roots of the eigenvalues of $\mathbf{A}^*\mathbf{A}$, where $\mathbf{A}^* = \bar{\mathbf{A}}^T$ is the adjoint matrix, are called singular values.

be taken to be read as Shannon information content.

Suppose there is a (random) event E, then the amount[3] of *self-information* associated with the event is given by $-\log_2 \Pr[E] = \log_2\left(\frac{1}{\Pr[E]}\right)$ bits. If an event has probability 1, then its self-information is zero (since $\log_2 1 = 0$). Thus, a low-probability event has more self-information than a high-probability event.

Now, suppose there is a "random experiment", whose possible (disjoint) outcomes are $E_1, ..., E_k$, then the average information, or *entropy* of the experiment is:

$$-\sum_i^k \Pr[E_i]\log_2 \Pr[E_i]. \tag{6.32}$$

For a continuous *pdf* of an arbitrary variable x, i.e. $P(x)$, the entropy is defined as

$$S(P) = -\int P(x)\log_2[P(x)/M(x)]\,\mathrm{d}x \tag{6.33}$$

where M is a measure function[4] chosen so that P/M is dimensionless, and so that $P = M$ corresponds to a state of no knowledge of x. For a Gaussian distribution P with covariance $\mathbf{S}'$ the entropy $S(P) = \frac{1}{2}\ln|\mathbf{S}'|$.

If $P_1(x)$ describes knowledge before a measurement, and $P_2(x)$ describes it afterward, then the information content of the measurement is the reduction in entropy:

$$H = S(P_1) - S(P_2). \tag{6.34}$$

In the linear Gaussian case the information content of a measurement can be evaluated either in state space

$$H_s = S(P(\mathbf{x})) - S(P(\mathbf{x}|\mathbf{y})), \tag{6.35}$$

or in measurement space

$$H_m = S(P(\mathbf{y})) - S(P(\mathbf{y}|\mathbf{x})), \tag{6.36}$$

although both should yield the same value

$$H = -\frac{1}{2}\ln|\mathbf{I}_n - \mathbf{A}|. \tag{6.37}$$

Therefore, it can be seen that the measurement will approach maximum information as the averaging kernel approaches the limit of the identity matrix, giving the relationship between information content and degrees of freedom for signal.

---

[3]Note that information is quantified by "bits", which are essentially dimensionless units and are determined by the base of the logarithm used.

[4]Examples of measures include length, surface area, volume and solid angle.

6.3 Converging to a Solution: The Nonlinear Problem

The optimal estimation retrieval is obtained by minimising a joint cost function (equation 6.15). If the forward model is linear and the *a priori* information has Gaussian statistics, then the cost function will be quadratic in the state vector and the equations to be solved will be linear. However, the forward model is very rarely linear in the real case.

To account for non-linear retrieval problems it is necessary to use iterative numerical methods. Essentially, the retrieval iterates to a solution by linearising the problem about the value of the state at each iteration stage. Two of the most common iterative methods are Newtonian Iteration and the Levenberg-Marquardt Method, which are described subsequently.

6.3.1 Newton's Method

Newton's Method is essentially a root-finding algorithm. In the scalar case it uses the first few terms of the Taylor series of a function $f(x)$ in the vicinity of the root to iteratively improve the estimate of the root. With a good initial approximation to the root x_0 the algorithm is applied to obtain

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{6.38}$$

for $n = 1, 2, 3, \dots$. Unfortunately, the method can be somewhat unstable near a horizontal asymptote.

In the case of a vector function $\mathbf{g}$ of a vector $\mathbf{x}$, the algorithm is iteratively applied to obtain:

$$\mathbf{x}_{i+1} = \mathbf{x}_i - [\nabla_\mathbf{x}\mathbf{g}(\mathbf{x}_i)]^{-1}\mathbf{g}(\mathbf{x_i}). \tag{6.39}$$

In the retrieval problem, Newtonian iteration is used to find the root of the derivative of the joint cost function χ^2 (in equation 6.15) i.e. to find $\hat{\mathbf{x}}$ such that $\nabla_\mathbf{x}\chi^2|_{\hat{\mathbf{x}}} = \mathbf{0}$, which is equivalent to minimising χ^2. So, let $\mathbf{g} = \nabla_\mathbf{x}\chi^2$ and substitute[5] into equation 6.43, to write the algorithm for Newtonian iteration to minimise χ^2 as

$$\mathbf{x}_{i+1} = \mathbf{x}_i - [\nabla_\mathbf{x}^2\chi_i^2]^{-1}\nabla_\mathbf{x}\chi_i^2. \tag{6.40}$$

From equation 6.15 the derivative of the joint cost function can be written explicitly as

$$\nabla_\mathbf{x}\chi^2 = -[\nabla_\mathbf{x}\mathbf{F}(\mathbf{x})]^T\mathbf{S}_\epsilon^{-1}[\mathbf{y} - \mathbf{F}(\mathbf{x})] + \mathbf{S}_a^{-1}[\mathbf{x} - \mathbf{x}_a] \tag{6.41}$$

---

[5]Note that as $\mathbf{x}$ is an $n \times 1$ column vector the gradient of a scalar function with respect to $\mathbf{x}$ will be a vector-valued function. Therefore, $\nabla_\mathbf{x}\chi^2$ is also an $n \times 1$ column vector.

where $\epsilon = \mathbf{y} - \mathbf{F}(\mathbf{x})$, and thus the second derivative of the cost function[6]

$$\nabla_{\mathbf{x}}^2 \chi^2 = \mathbf{S}_a^{-1} + \mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} - [\nabla_{\mathbf{x}} \mathbf{K}]^T \mathbf{S}_\epsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x})] \tag{6.42}$$

where $\mathbf{K} = \nabla_{\mathbf{x}} \mathbf{F}(\mathbf{x})$. The term involving $[\nabla_{\mathbf{x}} \mathbf{K}]^T$ is a complicated object (i.e. a vector of matrices) and is often expensive to compute. Fortunately, the term is usually small if the problem is moderately linear, therefore it is ignored and the following algorithm is obtained for Newtonian iteration:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + (\mathbf{S}_a^{-1} + \mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_i)^{-1} [\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x_i})) - \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a)] \tag{6.43}$$

where $\mathbf{K}_i = \mathbf{K}|_{\mathbf{x}_i} = \nabla_{\mathbf{x}} \mathbf{F}(\mathbf{x_i})$.

Essentially, the iteration step to $\mathbf{x}_{i+1}$ in the Newtonian method in equation 6.43 is the same as solving for $\hat{\mathbf{x}}$ in equation 6.23 if the forward model is linearised about $\mathbf{x}_i$. Thus, omitting the $[\nabla_{\mathbf{x}} \mathbf{K}]^T$ term in equation 6.42 is equivalent to the assumption of linearity within the error bars of the retrieval.

6.3.2   Convergence Criteria

It is important to determine the correct convergence criterion for stopping the iteration. There are a suite of convergence tests that can be applied, however they are all concerned with checking the smallness of one of the reduction of the cost function (between iterations), the gradient of the cost function or the size of the step, in state space or measurement space. The absolute value of the cost function does not say anything about whether the iteration has converged, but it is a useful measure of the quality of the fit of the converged solution to the measurement.

A simple and computationally straightforward test is on the cost function being minimised. A simple threshold (e.g. < 1) on the change in χ^2 between iterations can be set and if the change is less than this the iteration is stopped.

It is often useful to set more than one criteria for convergence. So in addition to testing for minimising the cost function a suitable test for convergence is if the size of the step in state space is less than the retrieval error i.e.

$$d_i^2 = (\mathbf{x}_i - \mathbf{x}_{i+1})^T \hat{\mathbf{S}}^{-1} (\mathbf{x}_i - \mathbf{x}_{i+1}) \ll n. \tag{6.44}$$

Another simple test is to check if the change in $\hat{\mathbf{x}}$ is less than some threshold.

---

[6]Note that $\nabla_{\mathbf{x}}^2 \chi^2$ does not denote the scalar Laplacian but as $\nabla_{\mathbf{x}} \chi^2$ is an $n \times 1$ column vector, $\nabla_{\mathbf{x}}^2 \chi^2$ is the $n \times n$ matrix such that the ij^{th} element is $\partial^2 \chi^2 / \partial x_j \partial x_i$.

6.3.3   Levenberg-Marquardt Method

Newton's method is dependent on the fact that the algorithm is employed for values of the state near enough to the solution so that (even in the non-linear case) the cost function is quadratic or nearly quadratic and thus the solution can be found. However, if the iteration point is too far from the true solution, a quadratic may not be a good representation of the cost function at that point and therefore the step taken will be meaningless. Levenberg-Marquardt is a popular method for finding the minimum of a non-linear function and is introduced to improve convergence in the cases where the initial approximation is far from the true solution.

Levenberg-Marquardt combines Newton's method with the method of steepest descent, using Newtonian iteration near the solution and steepest descent far from the solution.

The method of steepest descent is an algorithm for finding the nearest local minimum of a function which presupposes that the gradient of the function can be computed. When applied to a 1-dimensional function $f(\mathbf{x})$ of the vector $\mathbf{x}$, the method of steepest descent takes the form of iterating

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \gamma^{-1} \nabla_{\mathbf{x}} f(\mathbf{x_i}). \tag{6.45}$$

starting from x_0 for some scalar value γ, minimising $f(\mathbf{x})$ in the direction of the local downhill gradient until some convergence criteria is reached.

If the method of steepest descent was employed solely to minimise the cost function (which is a scalar function of a vector) the solution would be obtained from iterating

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \gamma^{-1} \nabla_{\mathbf{x}} \chi_i^2 \tag{6.46}$$

with the step size between iterations scaled by γ^{-1}.

The Levenberg-Marquardt method can be used to find the minimum of χ^2 in the retrieval problem from the algorithm

$$\mathbf{x}_{i+1} = \mathbf{x}_i - [\nabla_{\mathbf{x}}^2 \chi_i^2 + \gamma \mathbf{I}_n]^{-1} \nabla_{\mathbf{x}} \chi_i^2, \tag{6.47}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix. The inclusion of the identity matrix in the $\gamma \mathbf{I}_n$ term may seem counter intuitive but is necessary because $\mathbf{x}$ is a column vector and so $\nabla_{\mathbf{x}}^2 \chi_i^2$ is an $n \times n$ matrix.

It is clear from equation 6.47 that as $\gamma \longrightarrow 0$, the Levenberg-Marquardt step tends to the Newtonian iteration method and as $\gamma \longrightarrow \infty$, it tends to the method of steepest descent. The values of γ are chosen to give steepest descent far from the solution and Newtonian iteration near the solution by monitoring χ^2:

- If χ^2 increases, γ should be increased (making the step size smaller) and $\mathbf{x}_i$ left

unchanged.

- If χ^2 decreases, γ should be decreased (making the step size larger) and $\mathbf{x}_i$ updated.

The factor by which γ is increased or decreased can be flexible so it is somewhat dependent on the problem, but is most often set to a factor of 10.

Applying the Levenberg-Marquardt method to the optimal estimation problem where the cost function is given in equation 6.15 gives the iteration:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + (\mathbf{S}_a^{-1} + \mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_i + \gamma \mathbf{I}_n)^{-1} [\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x_i})) - \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a)]. \qquad (6.48)$$

As elements of the state vector may have different magnitudes and dimensions, they should be scaled. A convenient means of doing this is to replace $\mathbf{I}_n$ with $\mathbf{S}_a^{-1}$ as a scaling matrix as it has the right dimensions. The Levenberg-Marquardt step then becomes:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + ((1+\gamma)\mathbf{S}_a^{-1} + \mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_i)^{-1} [\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x_i})) - \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a)]. \qquad (6.49)$$

The computational demands of the Levenberg-Marquardt method are little different to Newton's method, but more steps may be required as the iteration is usually started further from the solution (i.e. in cases where the a priori is poor) and this method is usually applied to more challenging problems.

6.3.4 Interpreting the Solution

It might seem that the concepts of retrieval quality discussed in section 6.2.5, i.e. information content of the measurement, degrees of freedom and the averaging kernel, will not apply to the iterative retrievals as the quantities are derived under the assumption that the forward model and retrieval method are linear over the range between the *a priori*, the true state and the retrieved state. However, as the prior state appears only linearly in the iterative retrieval method, the same solution would be obtained from a linear problem corresponding to linearisation at the retrieval (see pp. 86-87 Rodgers, 2000). Thus, the averaging kernel is a valid concept for interpreting the iterative optimal estimation retrieval. Also, the total error covariance matrix of the solution $\hat{\mathbf{S}}$ (as in equation 6.22) is valid if it is evaluated at the solution and regarded as if the problem is linear at the solution. The inverse of the total covariance matrix is the Fisher information matrix (see Frieden, 2004) which is a measure of the variance of the derivative of the likelihood of the solution (i.e. the posterior *pdf*). The simplest part of the information is contained in the diagonal elements of $\hat{\mathbf{S}}$. These are the variances of the elements of the state vector. The square roots of the diagonals give the standard deviation of the retrieval $\hat{\mathbf{x}}$, mainly due to measurement

noise. The off-diagonal elements give the correlation of errors between elements of the state vector, ideally these will be zero.

In short, for moderately non-linear problems, linearisation is adequate for the error analysis and information content study, but not for finding a solution.

6.4 Retrieval Errors

In order to determine the optimality of a retrieval method, a quantitative analysis of the sensitivity of the retrieval to the true state and the sensitivity of the retrieval to the various sources of error which propagate into the final product should be carried out.

6.4.1 Including Forward Model Error

The expressions derived for retrieval methods have been done so assuming that the forward function can be approximated by the forward model incurring only a negligible uncertainty. However, for complicated measurements there is usually an associated error with this approximation to the complete physics:

$$\triangle \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') = \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') - \mathbf{F}(\mathbf{x}, \mathbf{b}) \tag{6.50}$$

such that the measurement is a function of the state vector $\mathbf{x}$ and of some other set of parameters and $\mathbf{b}'$ represents the forward function parameters which are ignored in the construction of the forward model. Therefore, the relation of the true state to the retrieved state (in equation 6.1) becomes

$$\hat{\mathbf{x}} = \mathbf{R}(\mathbf{F}(\mathbf{x}, \mathbf{b}) + \triangle \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') + \epsilon, \hat{\mathbf{b}}, \mathbf{x}_a, \mathbf{c}). \tag{6.51}$$

The error analysis of the inverse method can be better understood if the problem is considered linear at the solution. The linear problem, including forward model error as well as measurement noise error can be written

$$\mathbf{y} = \mathbf{K}_x \mathbf{x} + \mathbf{K}_b(\mathbf{b} - \hat{\mathbf{b}}) + \triangle \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') + \epsilon \tag{6.52}$$

where $\mathbf{K}_x = \frac{\partial \mathbf{F}}{\partial \mathbf{x}}$ and $\mathbf{K}_b = \frac{\partial \mathbf{F}}{\partial \mathbf{b}}$. Subsituting this expression into equation 6.26, the solution is obtained

$$\hat{\mathbf{x}} = \mathbf{G}_y \mathbf{K}_x \mathbf{x} + (\mathbf{I}_n - \mathbf{G}_y \mathbf{K}_x)\mathbf{x}_a + \mathbf{G}_y \epsilon_y \tag{6.53}$$

where $\mathbf{G}_y = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{y}}$ is the sensitivity of the retrieval to the measurement, which is the same as

its sensitivity to measurement error, $\mathbf{x}$ is the true state and

$$\epsilon_y = \mathbf{K}_b(\mathbf{b} - \hat{\mathbf{b}}) + \triangle \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') + \epsilon. \tag{6.54}$$

Equation 6.53 represents a well behaved retrieval inclusive of forward model errors, forward model parameter errors and measurement noise error.

6.4.2 Error Analysis

An expression for the error in the retrieval $\hat{\mathbf{x}}$ as a difference from the true state can be obtained from equation 6.53, assuming the retrieval is well behaved, such that

$$\hat{\mathbf{x}} - \mathbf{x} = (\mathbf{A} - \mathbf{I}_n)(\mathbf{x} - \mathbf{x}_a) + \mathbf{G}_y \mathbf{K}_b(\mathbf{b} - \hat{\mathbf{b}}) + \mathbf{G}_y \triangle \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') + \mathbf{G}_y \epsilon \tag{6.55}$$

where the averaging kernel

$$\mathbf{A} = \mathbf{G}_y \mathbf{K}_x = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} \tag{6.56}$$

is the sensitivity of the retrieval to the true state. As $\mathbf{A} \longrightarrow \mathbf{I}_n$, the bias of the *a priori* on the retrieval is reduced.

The terms in equation 6.55 each correspond to a source of error propagated through the inverse method into the solution:

- The first term, $(\mathbf{A} - \mathbf{I}_n)(\mathbf{x} - \mathbf{x}_a)$, is known as the *smoothing error* or the *null-space error*, but will be referred to here as the *sensitivity error*.

- The second term, $\mathbf{G}_y \mathbf{K}_b(\mathbf{b} - \hat{\mathbf{b}})$, is known as the *model parameter error*.

- The third term, $\mathbf{G}_y \triangle \mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}')$, is known as the *forward model error*.

- The final term, $\mathbf{G}_y \epsilon$, is known as the *retrieval noise*.

These error sources are discussed in further detail by (Rodgers, 2000). The model parameter error is essentially the error in the retrieval due to the uncertainty in the model parameters. The term 'smoothing error' is not used in this thesis as it can be confusing when applied to state vectors composed of orthogonal quantities, but is useful when the state vector is an atmospheric profile (e.g. of temperature) with respect to height. The term 'null-space' error is often used as it is descriptive of the error that arises from a basic lack of information in the measurement system, but the term should be reserved for the error contribution from those components which lie in the null space[7] of $\mathbf{K}$. The retrieval sensitivity error is the error due to the lack of sensitivity of the observing system to the individual parameters of

---

[7]The null space is the undetermined part of state space and is the complementary subspace to the row space of $\mathbf{K}$, where the row space of $\mathbf{K}$ is the subspace spanned by the vectors forming the rows of $\mathbf{K}$.

the state vector. This is manifested in the correlations between state parameters, and in terms of the formalism, the degree to which the averaging kernel differs from an identity matrix. As the true state is not known, the actual sensitivity error cannot be determined. However, for the purposes of error analysis the retrieval can be regarded as an estimate of a state transformed by the averaging kernel, i.e. $\mathbf{x}' = \mathbf{A}\mathbf{x}$, rather than an estimate of the true state with an error contribution due to the retrieval sensitivity (or effectively, the lack there of), and the estimate of sensitivity error can be abandoned.

Considering forward model error and forward model parameter error (so the measurement is given by equation 6.52) gives the actual measurement covariance matrix of $\mathbf{y}$ as

$$\mathbf{S}_y = \varepsilon\{(\mathbf{y} - \mathbf{y_I})((\mathbf{y} - \mathbf{y_I})^T\} = \mathbf{S}_\epsilon + \mathbf{S}_{fx} + \mathbf{S}_{fb} \tag{6.57}$$

(as opposed to equation 6.6) where

$$\mathbf{S}_{fx} = \varepsilon\{[\mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') - \mathbf{F}(\mathbf{x}, \mathbf{b})][\mathbf{f}(\mathbf{x}, \mathbf{b}, \mathbf{b}') - \mathbf{F}(\mathbf{x}, \mathbf{b})]^T\} \tag{6.58}$$

and

$$\mathbf{S}_{fb} = \mathbf{K}_b \mathbf{S}_b \mathbf{K}_b^T \tag{6.59}$$

such that

$$\mathbf{S}_b = \varepsilon\{(\mathbf{b} - \hat{\mathbf{b}})(\mathbf{b} - \hat{\mathbf{b}})^T\} \tag{6.60}$$

is the covariance matrix of $\mathbf{b}$.

In the linearised problem with only measurement noise error (equation 6.22) the total error covariance matrix is the sum of the measurement error contribution $\mathbf{G}_y \mathbf{S}_\epsilon \mathbf{G}_y$ and the contribution from sensitivity error for an ensemble of states corresponding to the *a priori* $(\mathbf{A} - \mathbf{I}_n)\mathbf{S}_a(\mathbf{A} - \mathbf{I}_n)^T$ (which is seen from equation 6.25). If forward model error and forward model parameter errors are now included in the measurement error contribution then the total error covariance matrix is now

$$\begin{aligned} \hat{\mathbf{S}} &= \mathbf{G}_y \mathbf{S}_y \mathbf{G}_y + (\mathbf{A} - \mathbf{I}_n)\mathbf{S}_a(\mathbf{A} - \mathbf{I}_n)^T \tag{6.61} \\ &= (\mathbf{A} - \mathbf{I}_n)\mathbf{S}_a(\mathbf{A} - \mathbf{I}_n)^T + \mathbf{G}_y \mathbf{S}_\epsilon \mathbf{G}_y^T + \mathbf{G}_y \mathbf{K}_b \mathbf{S}_b \mathbf{K}_b^T \mathbf{G}_y^T + \mathbf{G}_y \mathbf{S}_{fx} \mathbf{G}_y^T. \tag{6.62} \end{aligned}$$

To simplify practical retrieval methods the assumption that $\mathbf{S}_y \approx \mathbf{S}_\epsilon$ can be made, i.e. that forward model error can be neglected. This was the assumption for deriving the mathematical expressions for the non-linear methods. To propagate all the error sources through the non-linear retrieval, $\mathbf{S}_\epsilon$ should be replaced by $\mathbf{S}_y$ (equation 6.57) at the linearised problem about each iteration. The error analysis above is valid for the non-linear case if the problem can be interpreted as linear at the solution within the error bounds of the retrieval.

Chapter 7

Cloud Property Retrieval from MIPAS Spectra

This chapter presents the practical application of retrieval theory (as presented in the previous chapter) to infer cloud properties from MIPAS spectra. This is the first known attempt to use an optimal estimation technique to retrieve cirrus properties from infrared limb sounder measurements.

7.1 Formulation of the Problem

In any inverse problem, the first step in finding the solution is to determine whether or not the problem is well-posed. The most common application of inverse theory to remote (limb) sounding of the atmosphere is to retrieve information, in particular temperature, pressure and trace gas concentration, as a function of height. These are continuous functions (of altitude) and as such they constitute an infinite number of pieces of information. Therefore, as an instrument can only take a finite number of measurements, the inverse problem of retrieving atmospheric profile information is fundamentally under-constrained, i.e. there are a greater number of unknowns than measurements ($m < n$). There are a number of methods that can be introduced to discretise the problem (see Rodgers, 2000), but fortunately, for the retrieval of cirrus properties it is possible to formulate the problem so that the unknown variables to solve for are single pieces of information, i.e. discrete objects. Thus, if the number of measurements is more than the number of discrete parameters to be solved for ($m > n$) then the problem becomes over-determined and if they are equal ($m = n$) then the problem is well-determined. The choice of measurement vector is therefore an important decision.

The retrieval scheme implemented to solve the problem (i.e. equation 6.51) of inferring cirrus properties from MIPAS measurements, uses McClouds_FM to perform the forward

model calculations, thus the scheme is referred to as McClouds_RT.

The following section describes the McClouds_RT implementation.

7.2 Methodology

An optimal estimation retrieval was used in McClouds_RT to determine the cloud state vector $\hat{\mathbf{x}}$ which is the most-likely estimate which best fits the measurements $\mathbf{y}$ for a given *a priori* $\mathbf{x}_a$ with covariance $\mathbf{S}_a$.

As McClouds_FM is nonlinear, the inverse problem is also nonlinear, thus an iterative method is required to find a solution by minimising the cost function, i.e. finding the root of the derivative of the posterior *pdf*.

7.2.1 The State Vector

The state should be represented in terms of as small a number of parameters as possible. In chapter 4 it was shown that the scattering properties of the cloud can be derived from the effective radius and the total number density (for an assumed crystal habit and size distribution) and the physical geometry of the cloud, under the assumption of a spherical layer in the atmosphere with a given horizontal coverage, can be fully described by the cloud top height and vertical depth of the cloud. Thus, the state vector contains the fundamental information required to calculate the physics of the measurement. However, the actual quantities of interest when describing cloud are cloud top height, effective radius, optical thickness and/or IWC. Optical thickness and IWC are functions of the state vector elements and can therefore be inferred from the retrieved parameters. These desired quantities are not directly as useful in the forward model. Therefore, inferring these properties as opposed to attempting to retrieve them in the state vector avoids having to make more complicated assumptions about the cloud representation, further constraining the problem.

In short, the quantities retrieved by McClouds_RT are:

- Effective radius r_e

- Number density N

- Cloud top height C_{top}

- Cloud vertical thickness C_{depth}

for an assumed crystal habit and under the assumptions stated in section 4.2 on cloud microphysics and geometry. Therefore, the state vector is chosen to be consistent with McClouds_FM, i.e. $\hat{\mathbf{x}} = [r_e, N, C_{\text{top}}, C_{\text{depth}}]$.

7.2.2 Atmospheric State Parameters

The atmospheric state parameters $\hat{\mathbf{b}}$ are chosen to be consistent with those required by the forward model (section 4.2.3). They are assumed to be well known and include:

- Atmospheric temperature and pressure profiles. This can be climatology or operational ECMWF data.

- List of significant absorbing gases in the spectral region of interest (which depends on the measurement). These are determined by using the '*' wildcard option (Dudhia, 2004) in the RFM clear-sky calculations used by McClouds_FM to select all absorbers above a chosen cut-off strength[1].

- Concentration of significant absorbing gases as a function of height in the atmosphere. These profiles come from climatology (Remedios, 1999) except for water vapour for which the ECMWF operational data can be used.

- The local radius of curvature of the Earth.

- The temperature and emissivity of the Earth's surface.

- Instrument line shape (ILS) data.

- The viewing tangent height of each limb scan, provided in the MIPAS Level-1B data (ESA, 2000).

As well as the parameters listed above, molecular spectroscopic data (i.e. absorption cross-sections) are supplied in the HITRAN-2000 database (see Rothman et al., 2003).

When cloud is in the instrument line-of-sight, the MIPAS spectra will be dominated by cloud effects and the sensitivity of the measurements to the fine structure of atmospheric temperature, pressure and absorbing gases is reduced. Therefore, although the assumption of these atmospheric parameters to be well known will rarely hold, the effect of the variations is comparatively small and so they are included in $\hat{\mathbf{b}}$ as opposed to the state vector. The retrieval will still be somewhat sensitive to these parameters and the effect on the solution can be investigated as part of the forward model parameter error term (see Section 6.4.2). As with all limb sounders, precise knowledge of the pointing of the line-of-sight is crucial for succesful data analysis. Figure 4.10 showed that the synthesised measurements for a given cloud type are sensitive to the viewing tangent height supplied, causing up to 5 % possible uncertainty in the measurement for 1 km errors in tangent height. The

---

[1]With optical strength greater than or equal to 12. The "optical strength" parameter is notionally $15 + log_{10}$(max optical depth) for a zenith path through a standard atmosphere so that a value 15 corresponds to optical depth 1 (hence "*(12)" excludes all absorbers with optical depths less than 10^{-3}). Pretabulated and approximate optical strength values of gas species are used by the RFM (Dudhia, 2004).

engineering tangent heights in the MIPAS level-1B data are supplied with an uncertainty giving a standard deviation of $\sigma_{t_h} = 0.6$ km. This is a large uncertainty and could prove problematic for retrievals which are strongly dependent on temperature profile structure. Retrieval schemes have been implemented to retrieve tangent altitude profiles (e.g. von Clarmann et al., 2002) constrained by the engineering pointing errors in joint retrievals with temperature and pressure. However, the problem with such retrieval schemes is the high correlation between temperature, pressure and the tangent altitudes, and that it cannot be applied in the presence of clouds. In McClouds_RT only one limb scan measurement is used and the tangent height is chosen not to be included in the state vector as it would be highly correlated to the cloud top height, which is of main interest and it would also considerably increase the computational effort required. The view taken is that the inversion problem is more tractable and more likely to succeed with the state vector comprised of only the essential cloud components. It is preferred to retrieve the cloud state for an assumed tangent height, and the error on the solution from the effect of pointing uncertainty in the retrieval can be determined from equation 6.62.

Alternatively, the error in pointing information could be improved before the tangent height is input to McClouds_RT. Conventional MIPAS pressure and temperature retrievals are not applicable to cloudy limb scans, however, they can be applied to the limb scans above the cloud level. Pressure retrievals (e.g. Carli et al., 2004) are quoted to give the limb pressure to within $\sim 5\%$. By using the pressure in conjunction with either the assumption of hydrostatic equilibrium or the ECMWF operational geo-potential height data (co-located temporally and geographically with the limb tangent point), the absolute altitude of the limb scan above the cloud contaminated measurement can be calculated. Then, using the fact that the relative pointing uncertainty between limb scans is only 300 m ($\sigma_{rt_h} = 0.15$ km), theoretically the altitude of the cloudy limb scan could be estimated to within $\lesssim$ 0.5 km (depending on pressure). However, this method of improving the tangent height estimate is not used in this thesis as it is often the case that the pressure retrievals have higher errors in the first 2 or 3 limb scans above the cloudy scan in the same profile. Once the error in the pressure retrievals, the error incurred converting pressure to altitude, and the error in the relative pointing is accounted for, the additional effort is not worth the very small improvement on the estimate of the limb tangent height.

7.2.3 Iterative Method

The Levenberg-Marquardt method was implemented for solving the retrieval problem. In addition, the improvement on the Levenberg-Marquardt method suggested by Fletcher (1971) was also implemented as a second approach.

Fletcher (1971) suggested that the use of the γ parameter itself could be altered to reduce the number of iterations required in the retrieval. The strategy used to update the value of γ in the retrieval can be modified to be based on the ratio of the change in the cost function computed properly to that computed with the linear approximation to the forward model. This ratio will be unity if the linear approximation is satisfactory, and negative if χ^2 has increased rather than decreased. The aim is to find a value of γ which restricts the new value of $\mathbf{x}$ to lie within linear range of the previous estimate. This strategy may avoid the retrieval making unnecessary steps which will only increase χ^2 further. The strategy is basically:

- If the ratio is greater than 0.75, reduce γ.

- If the ratio is less than 0.25, increase γ.

- Otherwise make no change.

- If γ is less than some critical limit, use zero.

Fletcher's strategy does not require a sizeable amount of further computation for a given step compared to the original Levenberg-Marquardt, and any additional computation will be traded off by the reduction in the number of unnecessary steps.

7.2.4 *A Priori* Information

A Priori represents the knowledge of the state before the measurement is made. The term *a priori* includes both a mean state $\mathbf{x}_a$ and its covariance $\mathbf{S}_a$. If the *a priori* distribution is non-Gaussian then the inverse problem will be nonlinear.

A Priori information is introduced to restrict the components of the solution that are in the regions of state space that the measurement is not sensitive to i.e. the null-space or near null-space of the weighting function matrix. It is important that the *a priori* constrain the solution in a reasonable way; if the constraint is too tight then the component of the retrieval will be biased towards the *a priori* state, and if it is too loose then the retrieval error for the component and its estimate may be too large if the measurement information is limited. The sensitivity of the retrieval to the *a priori* is given by:

$$\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}_a} = \mathbf{I}_n - \mathbf{A}. \tag{7.1}$$

Thus, the influence of the *a priori* on the retrieval is reduced as the averaging kernel approaches an identity matrix, which is consistent with high degrees of freedom for signal and high information content of the measurement.

A Priori information can be difficult to construct. Climatology is often used, as are *ad-hoc* approximations or smoothness constraints. For clouds, climatology gives a wide variation of occurences, types and properties, mostly with non-Gaussian distributions. There are two options of how to progess:

- Use very loose constraints, so that there is effectively no *a priori*, i.e. an arbitrary mean *a priori* state with near infinite variance.

- Construct the *a priori* under the assumption that the retrieval only operates on measurements of cirrus.

The latter option is chosen for the *a priori* information in McClouds_RT. As the retrieval is nonlinear, it is likely that there is more than one solution, so the retrieval problem is better constrained to solutions that are known *a priori* to be typical of cirrus clouds. From chapter 1 it is known that cirrus macrophysical and microphysical properties vary widely, depending on formation mechanism and geographic location among other factors. Therefore, the *a priori* information will still be a loose restriction due to the variability of cirrus. The loose constraint is beneficial as it will limit the bias of the *a priori* on the retrieval. Cirrus climatology can then be used to construct the *a priori*.

Justifying the use of cirrus climatology to construct the retrieval *a priori* depends on the interpretation of the true meaning of *a priori* information as it relates to the measurement. Before the measurement is made, there is no *a priori* knowledge if there is cloud in the MIPAS FOV or not. To include a frequency of cloud occurence in the *a priori* would involve a non-Gaussian distribution which would depend on numerous atmospheric and geographic conditions, which would make the retrieval ill-posed. However, if the measurement is considered a measurement of cirrus, then it must be necessarily true that cirrus is present before the measurement, hence prior knowledge of cirrus properties can be used and no information from the measurement itself has biased the *a priori*.

For the purposes of the retrieval, it is assumed that the measurement vector is an indirect measurement of cirrus. In fact, by simply using McClouds_FM as the forward model cirrus has already been assumed *a priori* due to the necessary assumptions in the model. Therefore, assuming cirrus is effectively a model parameter error. Of course it may happen that MIPAS measurements are supplied to McClouds_RT which do not have cirrus in the FOV and these are likely to either not converge or return spurious results. In order to avoid wasting computing time, and to ensure the measurements are consistent with the *a priori*, it is important to filter the measurements supplied to McClouds_RT to be cirrus measurements. The methods discussed in chapters 3 and 5 are used to detect cloud in the limb scans and classify it as cirrus before the limb scan measurement is supplied to McClouds_RT. Thus, the cloud detection and classification methods do use information

Cloud Property	Mean Value $\mathbf{x}_a$	Standard Deviation σ_a
r_e [μm]	40	35
N [cm^{-3}]	3.5	3.5
C_{top} [% of tropopause height]	70	30
C_{depth} [km]	1.5	1.25

Table 7.1: *A priori* mean values and standard deviations for cloud properties determined from limb scans with tangent heights within 60–80 % of the tropopause height.

Cloud Property	Mean Value $\mathbf{x}_a$	Standard Deviation σ_a
r_e [μm]	30	25
N [cm^{-3}]	2.0	2.0
C_{top} [% of tropopause height]	90	25
C_{depth} [km]	1.0	0.9

Table 7.2: *A priori* mean values and standard deviations for cloud properties determined from limb scans with tangent heights within 80–100 % of the tropopause height.

from the measurement, but this is only used to determine the fact that the *a priori* is appropriate, and the *a priori* data itself comes from an independent source and is not influenced by the measurement.

The cloud detection and first guess methods assume some knowledge of the limb scan tangent height (from the pointing information), by doing so and not attempting to retrieve tangent height it allows the tangent height of the limb scan to select the *a priori* information. In chapter 1 the results of in-situ cirrus measurements and approximate climatologies were discussed. The trends from chapter 1 are used to select *a priori* information. Firstly, only limb scans with tangent heights above 60 % of the local tropopause height are considered for cirrus to be present as high thin cirrus is of greatest interest (see section 1.4). Secondly, as cirrus properties are dependent on height in the atmosphere i.e. higher clouds are expected to be thinner, more tenuous and contain smaller particles, the *a priori* is determined by whether the tangent height (which is the tangent height of the limb scan selected by the method in chapter 5) is within 60–80 % of the tropopause level or within 80–100 %. The *a priori* mean values and standard deviation are listed in tables 7.1 and 7.2. These *a priori* values are comparable to the average properties of cirrus given in table 1.2. However, the values differ reasonably significantly as high thin cirrus is of most interest and thus the *a priori* values are influenced by the trends discussed in section 1.5 due to *in situ* studies. Cloud top height is dependent on local temperature profile and thus on geographic location. This is accounted for in the *a priori* by using a percentage of the tropopause height from the temperature profile used in the atmospheric state parameters. The higher cirrus clouds have a smaller expected radius, number density and cloud thickness. In both cases cirrus is

approximated *a priori* to be a spherical layer. A broad Gaussian distribution is a reasonable estimate for cloud top height, effective radius and cloud depth due to the high variability. Total number density is not well known (see section 1.5.3), but it is likely to peak at lower number densities (< 1), with a rapid decrease and have a long tail into the higher number densitites. Thus the distribution may not be well approximated by a Gaussian distribution, so the variance for number density is relatively higher than for the other state parameters. This high variance gives the retrieval a very loose constraint for number density, which is a fair reflection of the lack of *a priori* knowledge. The fact that the distribution goes into negative numbers with such a high variance is physically impossible for number density but irrelevant as far as the retrieval constraints are concerned[2].

The state parameters are probably correlated, e.g. smaller crystals with higher cloud top height and smaller cloud depth as discussed in section 1.5, but these relationships are a weak trend and are not well defined empirically or theoretically. In order to simplify the problem, the *a priori* covariance matrix $\mathbf{S}_a$ is assumed to be diagonal so that the correlations are neglected and the off-diagonal elements are set to zero. $\mathbf{S}_a$ will be of the form:

$$
\begin{bmatrix}
\sigma_a(r_e)^2 & 0 & 0 & 0 \\
0 & \sigma_a(N)^2 & 0 & 0 \\
0 & 0 & \sigma_a(C_{\text{top}})^2 & 0 \\
0 & 0 & 0 & \sigma_a(C_{\text{depth}})^2
\end{bmatrix}
\tag{7.2}
$$

where the σ_a values are taken from tables 7.1 and 7.2.

7.2.5 Choice of Measurement Vector and Measurement Noise

In cases where there are a large number of measurements it is not always possible or practical to use all the measurements, therefore a measurement vector must be selected. The choice of the measurement vector supplied to the retrieval can prove critical in ensuring the inverse problem is well-posed. The length of the measurement vector is particularly important, as is the sensitivity of the measurements to changes in the state parameters. The measurement noise will also affect the choice of the measurement vector.

MIPAS records atmospheric limb spectra at very high resolution (section 2.5) covering a wide range of the thermal infrared region (685–2410 cm^{-1}) which results in a very high density of information. There are roughly 60,000 spectral measurements over the five spectral bands (table 2.1) per limb scan. With 17 limb scans per profile, and only 4 state parameters to retrieve it is clear that McClouds_RT will be inclined to be overconstrained. Only spectral measurements from one limb scan are supplied to the retrieval, following the

---

[2]A physical constraint is applied in McClouds_RT to keep number density positive.

method outlined in chapter 5 to select the limb scan to be used. The spectral region of interest is restricted to the atmospheric window region (see chapters 2 and 5), i.e. 8–13 μm ($\sim$ 770–1250 cm^{-1}), as the cloud effects are expected to be dominant (i.e. a large magnitude cloud signature) in this region. Therefore, for the purposes of McClouds_RT only MIPAS band A and AB spectra are of interest. However, there are still over 17,000 spectral measurements in these two bands, and the retrieval will be grossly overconstrained. The number of spectral points of interest can be further reduced by discarding the part of the spectra below 750 cm^{-1} as the H_2O and CO_2 lines are very strong in this region (as can be seen from figure 3.1) and will dominate the measurement. Thus, the number of MIPAS spectral measurements appropriate for McClouds_RT to operate on is approximately 15,000 and are in the range 750–970 cm^{-1} and 1020–1170 cm^{-1}.

There are two methods to better constrain the retrieval problem by choice of the spectral channels to include in the measurement vector. The first, is to select a narrow contiguous region of the spectra where there is strong evidence of scattering effects due to cirrus (i.e. absorption features) and thus appears sensitive to the cloud state parameters. The second, is to select the channels based on the information content and the number of degrees of freedom for signal, following Rodgers (1996). This will optimise the retrieval method to use the measurements most effectively.

The latter method is clearly more complex than the first. The physics of the measurement must be understood in order to quantify the change in entropy before and after the measurement is made. This is practically evaluated through the weighting function of the forward model. If the forward model is nonlinear then the information content and degrees of freedom will be dependent on the state (and atmospheric state). Therefore, it is nontrivial to select the optimal channels for the retrieval method as it will theoretically depend on the true solution, which is unknown. However, channels can be optimised if the retrieval method can be considered linear within the error bars of the retrieval. The discussion of the method of optimising channel selection for retrieving cloud properties is left for section 7.4.

The first method was used as a simple initial approach to implement the retrieval scheme and investigate the resultant diagnostics and capability. Even though it is not optimal this intitial approach retrieval method can still return useful results. The spectral region used in chapter 4 (i.e. 940–950 cm^{-1}) to compare McClouds_FM simulated radiances with real MIPAS measurements was selected as a region sensitive to cloud state. It is a reasonable selection as it is sensitive to scattering effects, with H_2O absorption lines and CO_2 sidelobes due to photons scattered into the FOV from below the cloud. The initial approach is described in section 7.3.

The measurement noise of MIPAS is random noise due to the instrument and is charac-

MIPAS Band	NESR [nW/(cm^2 sr cm^{-1})]
A	50
AB	40
B	20
C	6
D	4

Table 7.3: NESR requirements for unapodised MIPAS spectra by band.
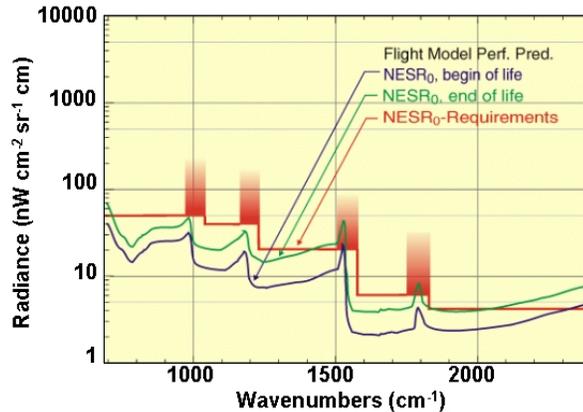
Figure 7.1: The NESR requirements (red) and the pre-flight predictions of NESR (red and green).

terised by the Noise Equivalent Spectral Radiance (NESR). The NESR requirements of MIPAS are listed in table 7.3 for unapodised spectra. The noise for unapodised spectra with a spectral bin width of 0.025 cm^{-1} will be uncorrelated. Figure 7.1 shows how the NESR requirements compare to the predicted noise. In-flight assesment of NESR is in good agreement with the predicted values (Birk and Wagner, 2003). It can be seen from figure 7.1 that the NESR requirements below 750 cm^{-1} are not well met. The high noise in this region of Band A is a further reason to neglect the spectra below 750 cm^{-1} for consideration in McClouds_RT.

As the NESR is uncorrelated for unapodised spectra the measurement error covariance matrix will be diagonal and can be expressed as:

$$\mathbf{S}_\epsilon = \begin{bmatrix} N_1^2 & 0 & \cdots & 0 \\ 0 & N_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & N_m^2 \end{bmatrix} \tag{7.3}$$

where m is the number of measurements in the measurement vector and N_i is the NESR of the ith measurement (given by table 7.3).

In section 4.8.1, it was stated that apodised spectra should be used to compare to model

MIPAS Band	NESR [nW/(cm^2 sr cm^{-1})]
A	32
AB	26
B	13
C	4
D	3

Table 7.4: NESR requirements for apodised MIPAS spectra by band.

output so that the spectrally narrow ILS artefacts are not confused with absorption features. Where available, apodised spectra should be supplied to McClouds_RT. The NESR values accounting for apodisation are given in table 7.4. Apodisation causes the error covariance matrix to have off diagonal elements. However, as the ILS is spectrally narrow the apodised noise correlation length is short and as the off-diagonal elements are relatively small they can effectively be ignored for simplicity. In the channel selection method the measurement vector elements are not necessarily spectrally adjacent. This creates a complicated error covariance matrix with respect to off-diagonal elements and thus makes a numerical inversion very difficult. It is thus necessary to ignore the noise correlations in the channel selection method. The error covariance matrix for apodised spectra is assumed to have the same form as in equation 7.3 with NESR values given in table 7.4.

7.2.6 Calculating Weighting Functions

The weighting function represents the sensitivity of the measurement to changes in the state vector and is dependent on state for non-linear forward models. Ideally, the weighting function can be derived analytically from the first derivative of the forward function with respect to the state vector. However, weighting functions are often calculated by numerical methods to simplify practical retrievals. Weighting functions are evaluated in McClouds_RT by perturbing, in turn, each of the parameters of the input state vector to the forward model, then differencing the resultant modelled spectrum from the original unperturbed modelled spectrum. This is a time-consuming process and the accuracy, like most numerical differentiation methods, is limited. If the perturbation is too small, the result can be affected by rounding errors and if the model is nonlinear then the weighting functions may be error prone if the perturbations are too large.

The sensitivity of the measurement to changes in the atmospheric state parameters, $\mathbf{K_b} = \frac{\partial \mathbf{F}}{\partial \mathbf{b}}$, can be evaluated by a similar method, where the perturbations are applied to the parameters of the input atmosphere and the tangent height. For a solution to be reached, $\mathbf{K}_x \gg \mathbf{K}_b$ must hold. In McClouds_FM (see chapter 4) the Monte Carlo scattering calculations will be unaffected by perturbations in the atmospheric temperature, pressure or

trace gas concentration, but the clear sky RFM calculations will. Therefore, due to the numerous RFM calculations required for the various elevation angles, the calculation of the weighting function with respect to the input atmosphere is very time-consuming, but limited mainly by the speed of the RFM, the number of trace gas species and the number of height/pressure levels assumed in the atmosphere model. However, both the Monte Carlo scattering calculations and RFM calculations (of $T_{\mathrm{cld-sat}}(C_{\mathrm{top}}, \phi_{\mathrm{exit}})$ and $R_{\mathrm{cld-sat}}(C_{\mathrm{top}}, \phi_{\mathrm{exit}})$ for equation 4.11) will be affected by a perturbation in tangent height, due to the change in viewing geometry of the photon propogation paths. The scattering calculations are more computationally intensive than the RFM clear sky calculations, so the weighting function with respect to the viewing tangent height model parameter is also time-consuming.

7.2.7 Forward Model Errors

There are two aspects of the forward model error to consider for McClouds_FM. The first error component is the error incurred from the difference between the forward model and the forward function given the assumptions about the cloud and atmosphere state, and the second error component is the error due to the difference between the real atmosphere and cloud and their representation in the model, e.g. the assumption of homogeneous cloud or the assumption of well mixed trace gases. The second error component is not easily quantifiable and it is neglected in this thesis. The first error component was discussed in section 4.6, the main points being:

- The forward model error depends on state.

- The model error can only be estimated as the forward function is too complex to know it accurately.

- The estimate is difficult and time consuming to calculate.

- The expected forward model error is small (see table 4.2).

For these reasons, the state dependent forward model error (as given in equation 6.50) and the corresponding covariance error contribution term ($\mathbf{G}_y \mathbf{S}_{fx} \mathbf{G}_y^T$ in equation 6.62) are not propogated through the iterations of McClouds_RT. However, a percentage 'mean' error (as in table 4.2) can be included at each iteration, which is a mean over an ensemble of states as a percentage of the mean spectral radiance. It is assumed to be spectrally uncorrelated, and can effectively be treated as a random error component[3] (i.e. by assuming $\mathbf{S}_\epsilon = \mathbf{S}'_\epsilon + \mathbf{S}_{fx}$ where $\mathbf{S}'_\epsilon$ is the error covariance due to measurement noise alone), which is generally smaller than the measurement error. This forward model error will be referred to

---

[3]The component of the error due to photon counting is in fact a random error with a Poission distribution.

as *random forward error* and propogating it through the retrieval will have a near negligible effect on the retrieved solution, but will scale the cost function and error analysis.

Similarly to the forward model error, the model parameter error $\mathbf{K}_b(\mathbf{b} - \hat{\mathbf{b}})$ and the corresponding covariance error contribution term $(\mathbf{G}_y\mathbf{K}_b\mathbf{S}_b\mathbf{K}_b^T\mathbf{G}_y^T$ as in equation 6.62) are not propogated through the retrieval iterations. This is because:

- The additional computational time required per iteration to compute $\mathbf{K}_b$ (which depends on state due to nonlinear effects) for all atmospheric parameters would make the retrieval untenable from a practical perspective.

- When cloud is present in the FOV, its effect dominates measurements in the window region and thus the sensitivity to $\hat{\mathbf{b}}$ will be small for most trace gas species, i.e. the magnitude of $\frac{\partial \mathbf{F}}{\partial \mathbf{b}}$ will be small. This low model parameter error is a further benefit of using atmospheric window region measurements.

For the purposes of error analysis it is useful to think of the additional model parameters vector as $\mathbf{b} = [\mathbf{b}_0, t_h]$ where $\mathbf{b}_0$ is the surface temperature and emissivity and the atmospheric temperature, pressure and trace gas information and t_h is the viewing tangent height of the MIPAS limb scan. This is useful because the calculations of the sensitivity of the measurements to $\mathbf{b}_0$ have a greater computational expense than the calculations of the sensitivity with respect to t_h. In this thesis, the model parameter error due to $\mathbf{b}_0$ is not investigated for these reasons and because the magnitude of $\mathbf{K}_{b_0}$ is expected to be small (at least with respect to the atmospheric gases), including error due to the uncertainty in $\mathbf{b}_0$ will have little effect on the retrieved solution, i.e. $\mathbf{K}_x \gg \mathbf{K}_{b_0}$. Alternatively, the retrieval can be thought of as the solution given that $\mathbf{b}_0 = \hat{\mathbf{b}}_0$. This, is an acceptable approach if the best possible estimates of $\mathbf{b}_0$ are taken from external datasets colocated both temporally and spatially (e.g. ECMWF operational data), so that $\mathbf{b}_0$ is well known and $\mathbf{S}_{b_0}$ is small. The retrieval error due to the uncertainty in the tangent height, i.e. the MIPAS pointing uncertainty, was investigated for the initial retrieval method and is discussed in the section 7.3.3.

7.3   Initial Approach

For the initial implementation of the retrieval, the measurement vector was chosen to be the spectral region of the MIPAS limb scans from 940 to 950 cm^{-1} as throughout this thesis it has been shown to be a region particularly sensitive to cloud scattering effects and thus cloud microphysical properties. Therefore, the measurement vector will be of length $m = 401$.

In order to evaluate the method a test case is presented. For the test case, both a retrieval on real MIPAS data and a retrieval on synthetic data are performed. The synthetic retrieval is useful because the true state is known. The retrievals were run using Fletcher's improved Levenberg-Marquardt algorithm, and hexagonal ice crystals were assumed in McClouds_FM.

7.3.1   Test Case

The test case selected is the same as that in section 4.8. Using orbit 504 from the 16th of May 2002, the measurement vector for the real apodised MIPAS data is taken from the 16.54 km tangent height of the profile with a mean tangent point location at a latitude of 1.25 °N and a longitude of 36.5 °E. The synthetic measurement vector is taken to be the "best-fit" modelled spectra (which, compared to the MIPAS spectra inclusive of the measurement noise and random forward error, has a cost function of $\chi^2 = 338.36$ and a reduced χ^2 value of 0.85), obtained from a cloud state vector of

$$\hat{\mathbf{x}}_c = [7.85\,\mu\text{m},\ 4.76\,\text{cm}^{-3},\ 18.75\,\text{km},\ 0.7\,\text{km}]$$

as given in section 4.8.2, with white noise added[4] to modelled spectra in line with the expected NESR (see tables 7.3 and 7.4). The retrievals for both cases were given an initial guess calculated using the method developed in chapter 5. The random forward error was propogated through the real case retrieval but not the synthetic retrieval as the synthetic measurements were assumed to have no forward error.

The results for the real and synthetic retrievals are given in tables 7.5 and 7.6 respectively, along with the initial estimate, the standard deviation $\hat{\sigma}$ of the retrieved parameters[5] and the theoretical uncertainty (i.e. standard deviation) as a percentage of the solution. The residuals (i.e. $\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}})$) for both the real and synthetic retrievals are given in figure 7.2. The values of the cost function χ^2 at the solutions were 88.82 for the real data retrieval and 135.47 for the synthetic retrieval, and the reduced χ^2 values[6] were 0.22 and 0.34. The averaging kernels are given in equations 7.4 and 7.5 respectively.

---

[4]Realistic noise is added to the synthetic measurement by adding the product of the NESR and a pseudo-random number between -1 and 1 to the synthetic radiance at each spectral point.

[5]The standard deviation is the square root of the diagonal of the total covariance matrix $\hat{\mathbf{S}}$ (equation 6.22).

[6]The reduced χ^2 values are calculated by χ^2/df where the degrees of freedom $df = N_{\bar{\nu}} - 1$ and $N_{\bar{\nu}}$ is the number of spectral points in the measurement vector. The reduced χ^2 values quoted are often less than one which is because the spectral points are not actually independent but correlated due to the forward model dependence on the cloud properties. Therefore, the assumed degrees of freedom is quite different to the degrees of freedom for signal d_s, which is less than or equal to the number of state parameters and in most cases in this thesis is at least one order of magnitude smaller than df. An alternative measure of the quality of retrieval fit considering model sensitivity is χ^2/d_s.

Cloud Property	$\mathbf{x}_0$	$\hat{\mathbf{x}}$	$\hat{\sigma}$	[%] Uncertainty
r_e	10.365	14.299	1.262	8.8
N	1.175	1.975	1.716	86.9
C_{top}	18.542	19.594	0.092	0.47
C_{depth}	0.414	0.544	0.476	87.6

Table 7.5: Retrieval results from test case using actual MIPAS measurements and assuming ECMWF atmospheric temperature and pressure data. Retrieval cost: $\chi^2 = 88.82$, $\chi^2/df = 0.22$.

Cloud Property	$\mathbf{x}_{true}$	$\mathbf{x}_0$	$\hat{\mathbf{x}}$	$\hat{\sigma}$	[%] Uncertainty
r_e	7.85	11.347	7.579	0.952	12.6
N	4.76	1.445	3.830	2.575	67.2
C_{top}	18.75	18.796	18.802	0.103	0.55
C_{depth}	0.7	0.386	0.862	0.554	64.2

Table 7.6: Retrieval results from test case using synthetic measurements and ECMWF atmospheric temperature and pressure data. Retrieval cost: $\chi^2 = 135.47$, $\chi^2/df = 0.34$.
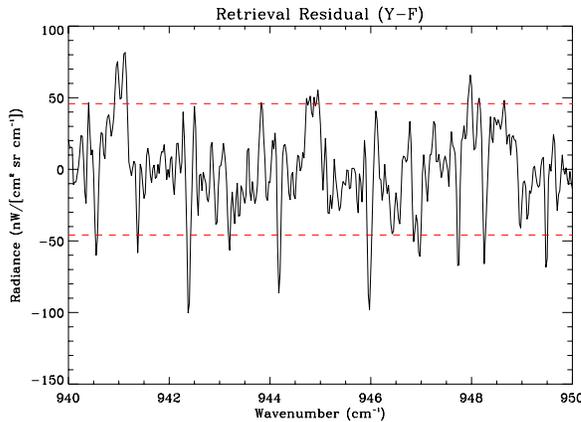
(a) Real case retrieval residual.
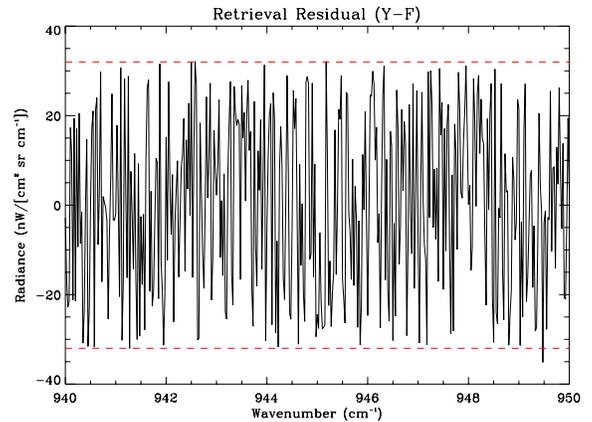
(b) Synthetic case retrieval residual.

Figure 7.2: Residuals (i.e. $\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}})$) from the real and synthetic initial approach retrievals. Overplotted in red is the root mean square measurement error (due to measurement noise and in the real case, random forward error).

$$\mathbf{A}_{\text{real}} = \begin{bmatrix} 0.9975 & -0.0727 & 5.314 \times 10^{-5} & -0.1990 \\ -4.652 \times 10^{-4} & 0.2638 & 0.0090 & 0.9246 \\ 1.164 \times 10^{-6} & 0.0307 & 0.9994 & -0.0428 \\ -2.579 \times 10^{-4} & 0.1872 & -2.533 \times 10^{-3} & 0.7200 \end{bmatrix} \quad (7.4)$$

$$\mathbf{A}_{\text{synth}} = \begin{bmatrix} 0.9985 & 0.0324 & 2.540 \times 10^{-3} & -0.0837 \\ 6.344 \times 10^{-4} & 0.4588 & 0.0299 & 1.758 \\ 2.799 \times 10^{-5} & 0.0168 & 0.9985 & -0.0569 \\ -1.085 \times 10^{-4} & 0.1163 & -6.684 \times 10^{-3} & 0.6212 \end{bmatrix} \quad (7.5)$$

Table 7.6 shows that the solution to which the real case retrieval converges is quite different to the "best-fit" cloud state in section 4.8. This is expected as the "best-fit" cloud state is only the best fit from a limited subspace of state space which was iterated to using a trial and error method (and the retrieved solution is outside the subspace which was tested). Moreover, the real case retrieval residual in figure 7.2(a) is near zero on average, with fluctuations of the order of the NESR plus the random forward error, and is thus much smaller than the "best-fit" residual in figure 4.9(b), which is also reflected in the lower cost function and thus the retrieval gives a better fit to the measurements. The residual in figure 7.2(a) shows that the forward modelled measurement at the solution appears to slightly underestimate the strength of the CO_2 emission lines and overestimate the strength of the absorption features.

The most significant difference between the retrieval and the "best-fit" cloud parameters is that the real case retrieved cloud top height is ~1 km higher. The retrieved cloud top is also above the tropopause, by over 1 km, according to the ECMWF temperature data. Such a high cloud top height is unexpected, but the retrieval uncertainty is very low. A possible explanation of this is given at the end of section 7.3.3. The remaining cloud parameters are close to what might be expected as typical values. The small effective radius of 14.299 μm is typical of thin tropopause level cirrus, and the error on r_e is relatively small, and is a considerable reduction on the *a priori* uncertainty. The values of N and C_{depth} have very high uncertainties, which can be explained by the averaging kernel in equation 7.4. The degrees of freedom for signal is $d_s = tr(\mathbf{A}_{\text{real}}) \approx 2.981$, with high sensitivity in the retrieval to r_e and C_{top}, but very low sensitivity to N and C_{depth}. This low sensitivity to N and C_{depth} causes the large uncertainty, but it is still an improvement over the initial *a priori* information. The lack of sensitivity to N is also shown by an apparent *a priori* bias on the retrieved value.

The synthetic retrieval result can be seen (from table 7.6) to be very close to the true cloud

state, i.e. well within the uncertainty on the retrieval. In addition, the residual (in figure 7.2) is essentially white noise of magnitude of the order of the NESR. This is a positive indication that, although the initial estimate is outside of the error range from the true state, it is close enough to the global minima for the retrieval to converge to a solution near the true state. The values of N and C_{depth} are within the error bounds of the true state but differ from it by 19.5 % and 22.9 % respectively. The larger discrepancy in the retrieved N and C_{depth} can be explained by the averaging kernel in equation 7.5. The diagonal elements of $\mathbf{A}_{\text{synth}}$ corresponding to N and C_{depth} are much less than 1 and the off diagonal elements (2$^{\text{nd}}$ row, 4$^{\text{th}}$ column and 4$^{\text{th}}$ row, 2$^{\text{nd}}$ column) are highly correlated. The degrees of freedom for signal is only $d_s = tr(\mathbf{A}_{\text{synth}}) \approx 3.077$, thus indicating that the retrieval is returning information on r_e and C_{top} effectively (as the corresponding diagonal elements are near unity), but the third degree of freedom is some parameter involving the correlated effect of N and C_{depth}. Hence, the high errors on N and C_{depth} can be explained because of the lack of retrieval sensitivity to them.

7.3.2 Nonlinearity

McClouds_RT is a non-linear retrieval, however, once a solution has been found, it can be checked whether the problem is near linear at the solution, so that a linear error and information content analysis can be used.

For synthetic retrievals, the true state is known and thus the error in the retrieval due to nonlinearity can be estimated by

$$\delta\hat{\mathbf{x}} = \mathbf{G}[\mathbf{F}(\hat{\mathbf{x}}) - \mathbf{F}(\mathbf{x}) - \mathbf{K}(\hat{\mathbf{x}} - \mathbf{x})] \tag{7.6}$$

where $\mathbf{x}$ is the true state (used to generate the synthetic measurement vector). If $\delta\hat{\mathbf{x}}$ is acceptable, e.g. less than the standard deviation, then near linearity can be assumed about the solution.

For real case retrievals, the true state is unknown and therefore $\delta\hat{\mathbf{x}}$ (equation 7.6) cannot be evaluated, but the degree of nonlinearity can be explored by evaluating $\delta\mathbf{y} = [\mathbf{F}(\hat{\mathbf{x}}) - \mathbf{F}(\mathbf{x}) - \mathbf{K}(\hat{\mathbf{x}} - \mathbf{x})]$ for values of $\mathbf{x}$ at the equivalent of one standard deviation away from $\hat{\mathbf{x}}$ using error patterns of $\hat{\mathbf{S}}$, i.e. for values $\mathbf{x} = \hat{\mathbf{x}} + \mathbf{e}_i$ where $\mathbf{e}_i = \lambda_i^{1/2}\mathbf{u}_i$ are the error patterns of $\hat{\mathbf{S}}$ and λ_i and $\mathbf{u}_i$ are the eigenvalues and eigenvectors of $\hat{\mathbf{S}}$. The significance of the nonlinear effects can be evaluated by comparing $\delta\mathbf{y}$ to the measurement error. Thus, the parameter $\zeta^2 = \delta\mathbf{y}^T\mathbf{S}_\epsilon^{-1}\delta\mathbf{y}$, which is essentially a χ^2, can be used to test the size of the nonlinear effects.

Corresponding Component	$\mathbf{u}_1$	$\mathbf{u}_2$	$\mathbf{u}_3$	$\mathbf{u}_4$
r_e	-0.1518	0.9779	0.0708	-0.1253
N	-0.9591	-0.1108	-0.1527	0.2108
C_{top}	0.0398	0.0121	-0.8887	-0.4566
C_{depth}	0.2354	0.1770	-0.4265	0.8552
Eigenvalue	$\lambda_1=3.1799$	$\lambda_2=1.5885$	$\lambda_3=3.978\times10^{-3}$	$\lambda_4=1.702\times10^{-5}$

Table 7.7: The eigenvectors, $\mathbf{u}_i$, and eigenvalues, λ_i, of the total covariance matrix of the real case retrieval.

Error Pattern	Nonlinearity Parameter
1	0.304033
2	0.153464
3	0.145705
4	0.150657

Table 7.8: The nonlinearity parameters for the error patterns in the real case retrieval (table 7.5).

For the synthetic retrieval in table 7.6, $\delta\mathbf{x}$ was evaluated and found to be

$$\delta\mathbf{x} = [0.04689\,\mu\text{m}, -0.04319\,\text{cm}^{-3}, -0.006718\,\text{km}, 0.05443\,\text{km}].$$

Thus, the error in in the synthetic retrieval due to nonlinearity is negligible compared to the error due to measurement noise and *a priori* covariance, so the problem can be considered nearly linear at the solution.

For the real case retrieval in table 7.5, the eigenvalues and eigenvectors of the total covariance matrix were calculated and are given in table 7.7. The ζ^2 parameter was calculated for each of the error patterns, and ζ^2/m gives a convenient nonlinearity parameter which compares the effects of the nonlinearity to the measurement noise (and random forward error). The nonlinearity parameter for each of the error patterns is given in table 7.8. The values of the nonlinearity parameter are all less than unity, indicating that within the range of the solution the effect of nonlinearity is less than that of measurement noise (and random forward error). Therefore, the problem can be considered near linear at the solution within the error bars of the retrieval, which, from experience, is generally the case.

7.3.3 Tangent Height Error

The results in section 7.3.1 were obtained without propogating model parameter error through the retrieval. However, the error due to the uncertainty in the tangent height (i.e. $\sigma_{t_h} = 0.6$ km) can be estimated at the solution such that the contribution to the total

Cloud Property	real			synthetic		
	$\hat{\sigma}$	$\delta\hat{\sigma}$	$\frac{\delta\hat{\sigma}}{\hat{x}}$ [%]	$\hat{\sigma}$	$\delta\hat{\sigma}$	$\frac{\delta\hat{\sigma}}{\hat{x}}$ [%]
r_e	1.295	0.033	0.23	0.967	0.015	0.20
N	1.718	0.002	0.10	2.578	0.005	0.13
C_{top}	0.098	0.006	0.03	0.103	4.1×10^{-4}	2.2×10^{-3}
C_{depth}	0.482	0.006	1.10	0.560	0.006	0.70

Table 7.9: The uncertainty in the retrieval inclusive of tangent height error (calculated at the solution), and the difference from the uncertainty in the retrieval exclusive of tangent height error.

covariance matrix is given by $\mathbf{G_y}\mathbf{k}_{t_h}\sigma_{t_h}^2\mathbf{k}_{t_h}^T\mathbf{G_y}^T$, i.e.

$$\hat{\mathbf{S}} = (\mathbf{A} - \mathbf{I}_n)\mathbf{S}_a(\mathbf{A} - \mathbf{I}_n)^T + \mathbf{G}_y\mathbf{S}_\epsilon\mathbf{G}_y^T + \mathbf{G_y}\mathbf{k}_{t_h}\sigma_{t_h}^2\mathbf{k}_{t_h}^T\mathbf{G_y}^T \qquad (7.7)$$

(c.f. equation 6.61) where $\sigma_{t_h}^2$ is the variance of the tangent height, the (m×1) column vector $\mathbf{k}_{t_h}$ is the weighting function with respect to tangent height for a given cloud state ($\hat{\mathbf{x}}$ in this case). Therefore, the new error estimates on $\hat{\mathbf{x}}$ (i.e. the standard deviations) are obtained from the square root of the diagonals of the new total covariance matrix inclusive of tangent height error, given in equation 7.7. For both the real and synthetic retrievals, the values of $\hat{\sigma}$ inclusive of tangent height error, the difference from the values without tangent height error (tables 7.5 and 7.6), and the equivalent difference as a percentage of the cloud parameters are given in table 7.9. In addition, replacing $\mathbf{S}_\epsilon$ in equation 6.15 with $\mathbf{S}_y = \mathbf{S}_\epsilon + \mathbf{k}_{t_h}\sigma_{t_h}^2\mathbf{k}_{t_h}^T$ to compute χ^2, gives a cost inclusive of tangent height error of 79.04 for the real retrieval and 70.24 for the synthetic case, and a reduced χ^2 of 0.20 and 0.18 for the real and synthetic cases respectively. The lower cost values show that by increasing the measurement error, the margin of error for fitting the measurements is increased and thus the cost function is decreased (due to the inverse measurement error covariance matrix in equation 6.15). Table 7.9 shows that the effect of including tangent height error in the total error covariance at the solution has only a very small increase in the percentage uncertainty in the retrieval. This could be considered negligible in comparison to the percentage uncertainty in the retrieval due to measurement noise alone (see tables 7.5 and 7.6).

Due to the small effect of tangent height uncertainty at the solution on the total covariance and due to the high computational cost required to calculate the weighting function with respect to the tangent height, the model tangent height parameter error is not propogated through the retrieval. In addition, the tangent height error is not propogated through the retrieval because of its nonlinear dependence on cloud state and the additional introduction of off-diagonal correlations to the total covariance matrix. As $\mathbf{k}_{t_h}$ depends on $\mathbf{x}$, including the error can bias the retrieval by effectively decreasing the cost at values of $\mathbf{x}$ at which

McClouds_FM is most sensitive to changes in tangent height. As this cost decrease is not necessarily a manifestation of a physical improvement in the estimate of the state vector, it has a complex and undesirable effect (which can cause non-convergence). An example of this effect is given in figure 7.3(a) which shows the spectral mean of the weighting function with respect to tangent height $|\bar{\mathbf{k}}_{t_h}|$ as it varies with changes in cloud top when the the the other cloud parameters are kept constant, in this case, the $\mathbf{x}_0$ values in table 7.5. The cost function at each cloud top with respect to $\mathbf{F}(\mathbf{x}_0) + \epsilon$, with and without including tangent height uncertainty, is given in figure 7.3(b).
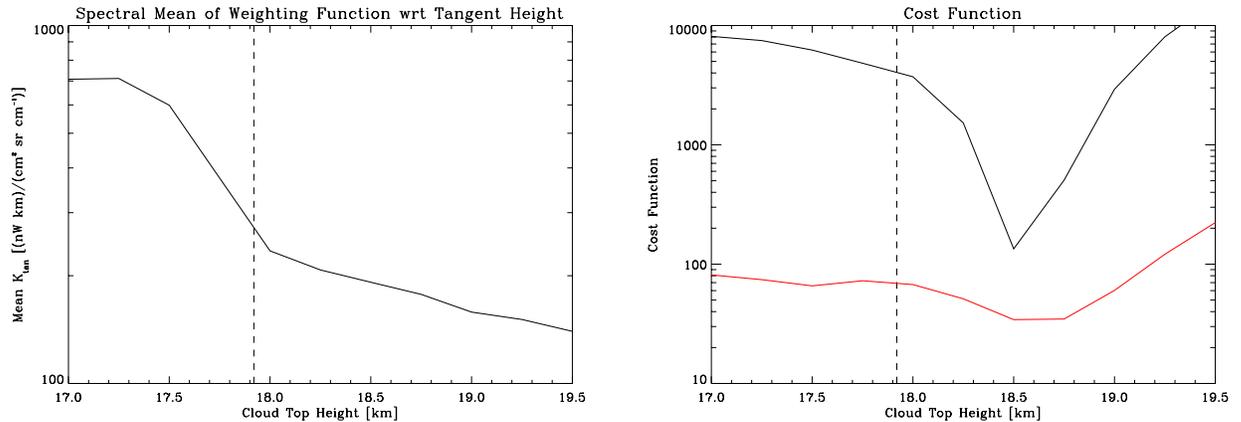
(a) Spectral mean of the weighting function with respect to tangent height, evaluated at various cloud top heights.

(b) Cost function of model measurements fit to $\mathbf{F}(\mathbf{x_0}) + \epsilon$, with (red) and without (black) including tangent height uncertainty.

Figure 7.3: The effect of tangent height uncertainty for varying cloud top heights, with the remaining cloud parameters r_e, N and C_{depth} constant and given by the $\mathbf{x_0}$ values in table 7.5. The viewing tangent height is 16.42 km and the dashed line is the height of the top edge of the FOV at the tangent point.

Figure 7.3(a) shows that for constant r_e, N and C_{depth} (in this case) the sensitivity of the measurements with respect to tangent height decreases as the cloud top height increases. There is significant sensitivity to tangent height at cloud top heights within the FOV, and figure 7.3(a) illustrates the reason that only limb scan measurements in which the FOV is fully cloud filled (i.e. cloud top heights greater than $\sim t_h + 1.5$ km) are used in the retrieval. Figure 7.3(b) shows the effect on the cost function of the nonlinear dependence of $\mathbf{k}_{t_h}$ on cloud top. When the tangent height uncertainty is included in the calculation of the cost function the cost values are greatly reduced (as expected), but it also introduces a local minima at approximately 17.5 km which is undesirable and could cause the retrieval to return a false solution. In addition, the global minima without tangent height uncertainty is clear and narrow about the true solution, whereas the global minima with tangent height uncertainty is broad and centred slightly higher than the true cloud top height which will bias the retrieval. In short, the nonlinear dependence of $\mathbf{k}_{t_h}$ on cloud state can introduce

complexities to the cost function topology if the tangent height error is propogated through the retrieval. These complexities can cause both a bias and additional iterations in the retrieval.

In general, if the cloud top height is above the FOV (as for all limb scans below the cloud top event scan) then the measurement becomes less sensitive to changes in tangent height (i.e. $\mathbf{k}_{t_h}$ decreases). As the sensitivity is small in this height range (and fairly constant with respect to cloud top height, e.g. figure 7.3(a)), the effect of the tangent height uncertainty on the retrieval total covariance (i.e. $\mathbf{G_y}\mathbf{k}_{t_h}\sigma^2_{t_h}\mathbf{k}^T_{t_h}\mathbf{G_y}^T$) is also small. Thus, instead of propogating tangent height error through the retrieval, an *ad-hoc* physical constraint is applied to the cloud top height in McClouds_RT so that it is above the FOV (± 0.5 km), and the small effect of the uncertainty on the retrieval can be estimated at the solution assuming linearity within the error bounds.

When dealing with model parameter error in the retrieval problem, the effect of the uncertainty on the retrieval covariance depends on the state/solution. For cloud tops above the FOV (which they must be by definition in McClouds_RT) perturbing the tangent height by a small amount has little effect on the simulated measurement (e.g. see figure 7.3(a)), and thus the tangent height uncertainty does not greatly affect the solution. However, this is not to say that the retrieval is actually insensitive to the pointing information. For example, consider the real retrieval for the test case above. If the tangent height is assumed to be one standard deviation lower, i.e. 0.6 km, and assuming the relative pointing information is still accurate, then McClouds_RT gives the solution in table 7.10, with the averaging kernel in equation 7.8. As the sensitivity to N and C_{depth} is low, only r_e and C_{top} are of concern here. The value of r_e is near identical to that in table 7.5, and the cloud top height is within the error bars of being σ_{t_h} lower than the cloud top height in 7.5. Placing the cloud top only slightly above the tropopause which is closer to the expected value. Considering the low sensitivity but the close agreement between N and C_{depth} for the retrievals at the two different tangent heights, it appears as if the retrieval is insensitive to the position of the tangent height, but the retrieval is actually sensitive to the cloud top height relative to the tangent height. However, this is merely conjecture and cannot be confirmed without validating the retrieval with further statisitics, but it does suggest that to estimate the true error due to the pointing uncertainty, the retrieval should be repeated assuming different tangent heights within the uncertainty and the range in the solutions accounted for in the quoted retrieval error.

Cloud Property	$\mathbf{x}_0$	$\hat{\mathbf{x}}$	$\hat{\sigma}$	[%] Uncertainty
r_e	10.365	14.235	1.350	9.5
N	1.175	1.562	1.421	91.0
C_{top}	17.642	18.964	0.243	1.3
C_{depth}	0.414	0.361	0.380	105.3

Table 7.10: Retrieval results from test case assuming $t_h = t_h - \sigma_{t_h}$, using actual MIPAS measurements and assuming ECMWF atmospheric temperature and pressure data. Retrieval cost: $\chi^2 = 298.23$, $\chi^2/df = 0.75$.

$$
\mathbf{A}_{\delta t_h} =
\begin{bmatrix}
0.9980 & -0.2448 & 5.242 \times 10^{-6} & 0.3550 \\
-2.781 \times 10^{-3} & 0.4393 & 0.0121 & 0.8212 \\
1.215 \times 10^{-7} & 0.0246 & 0.9979 & -0.0404 \\
7.407 \times 10^{-4} & 0.1508 & -3.642 \times 10^{-3} & 0.7772
\end{bmatrix}
\tag{7.8}
$$

7.3.4   Discussion

For the test case presented, the initial approach retrieval has been shown to return improved knowledge on the cloud state with a reduced relative uncertainty compared to the *a priori* information. It has also been shown that the nonlinearity in the retrieval is not a significant problem in either the real or synthetic test case, such that a linear error analysis and information content study are applicable at the solution. In addition, the effect of the tangent height uncertainty in the test case has been shown to have a negligible effect on the retrieval uncertainty and this is generally the case if there is a physical constraint on cloud top height to be above the FOV (of the limb scan measurement supplied to the retrieval).

Based on the test case and experience from additional retrievals (not presented), nonlinearity is not a problem once a solution is found, and although there is most often more than one minima, the number is unlikely to be of high order. Using the initial estimate method of chapter 5, has been shown for the test case to return a solution to the synthetic retrieval which is within the error bars of the true state. This suggests that the initial estimate is a close enough approximation to the global minima that the retrieval can iterate to an acceptable solution. Experience of other cases implies that by using the initial estimate method of chapter 5 and with a judicious starting value of the γ parameter in the Levenberg-Marquardt iteration step (in order to limit the size of the step in state space between iterations), the global minima can be closely approximated, considering the lack of retrieval sensitivity to N and C_{depth}. Therefore, the use of extravagant domain searches to avoid minor minima are not required.

The averaging kernels, in equations 7.4 and 7.5 show that the retrieval measures information on r_e and C_{top}, but the degrees of freedom for signal is $2.75 \lesssim d_s \lesssim 3.25$ (depending on the solution/state) which does not suggest the retrieval is sensitive to each of the four state parameters independently, i.e. that a change in state is not necessarily proportional nor does it directly correspond to a change in the retrieved state. It was found for the additional cases investigated (but not shown) that the initial approach retrieval, in general, gives $d_s \approx 3$, with one degree of freedom corresponding to each of r_e and C_{top}, and the third degree of freedom split (often almost equally) between N and C_{depth}. The large uncertainties in N and C_{depth} in the retrieval covariance matrix are due to this lack of sensitivity as is the difference from the true state in the synthetic retrieval. Essentially, the fact that $d_s < n$ for the initial approach, shows that the rank of the weighting function (at the solution) is $< n$ and thus a null space exists. The components of the solution which lie in null space cannot have been obtained from the measurements alone, hence the influence of the *a priori* on N and C_{depth}.

There are two explanations for the limited degrees of freedom in the initial approach retrieval:

- The forward model state parameters may not be the best parameters to be retrieved. As there are only $\sim$3 degrees of freedom for signal, it may be more appropriate to have a 3 element state vector, with r_e, C_{top} and either the product of N and C_{depth} or some function of the two parameters, such as optical thickness.

- The measurement vector used in the initial approach is not optimal and a higher information content and degrees of freedom for signal could be achieved by selecting an alternative subset of channels.

The first explanation is not investigated in this thesis (although it would be possible to reformulate McClouds_RT to do so), as it is more desirable to develop a retrieval method in which the degrees of freedom for signal is optimised. The information content of the MIPAS measurements and the optimal channels for cirrus property retrievals are investigated in the following section.

7.4 Channel Selection

In section 6.2.5 the information content and degrees of freedom for signal were introduced as figures of merit both to describe the quality of retrievals and to quantify the usefulness of remote measurements. The information content and degrees of freedom also provide single parameters which can be used in the optimisation of the selection of microwindows or subsets of channels for retrieval. Microwindow selection is a commonly applied technique

for operational retrievals of pressure, temperature and atmospheric gases from MIPAS (e.g. Dudhia et al., 2002; Ridolfi et al., 2000). However, as microwindows use spectrally adjacent measures, it gives a reasonable result but is not necessarily the best use of the data. The use of microwindows is, however, more efficient in terms of forward model calculations as opposed to the use of isolated spectral points since apodising calculated spectra requires that radiances are calculated at a number of spectrally adjacent points. For this thesis, the optimal selection of a subset of channels was used, and the forward model calculations were carried out over numerous microwindows which covered the regions in which the best channels were situated.

7.4.1 Theory

In principle, all possible combinations of channels should be considered, and the information content calculated for each combination in order to determine the optimal channel selection. However, such a method would not be computationally feasible. The simpler method of selecting channels sequentially is employed (following Rodgers, 1996). Before discussing sequential selection, it is important to reiterate that information content and degrees of freedom for signal are linear concepts. If the retrieval problem is non-linear, as is the case for MIPAS cloud retrievals, the linear theory can still be applied if the forward model is linear within the error bounds of the retrieved state. This is generally the case for a moderately non-linear problem. In the previous section the degree of non-linearity for the initial approach retrieval was assessed for the test case and shown to be acceptable within the error bounds. As the weighting function depends on state, the degree of non-linearity may not be acceptable over the whole of state space. For the purposes of channel selection, McClouds_FM is assumed to be near-linear locally. The channel selection will therefore only hold within the local neighbourhood of the arbitrary points in state space, and will thus need to be calculated over locations of state space representative of the range of typical cirrus properties. This will be discussed later.

For the purposes of discussing sequential estimation, consider the linear problem (equation 6.18) where the weighting function does not depend on state or a problem linearised about a point in state space such that the weighting function does not depend on state within the neighbourhood of that point.

If the original measurement error covariance matrix $\mathbf{S}_\epsilon$ is diagonal, then the equations in section 6.2.4 describing the linear problem can be applied to the individual measurements. Thus, by applying equation 6.26 to the individual measurement y_j, the retrieval can be estimated by

$$\mathbf{x}_j = \mathbf{g}_j y_j + (\mathbf{I}_n - \mathbf{g}_j \mathbf{k}_j^T)\mathbf{x_a} \tag{7.9}$$

where the $(n \times 1)$ column vector $\mathbf{k}_j$ is the jth weighting function, such that $\mathbf{k}_j^T$ is the jth row of the original weighting function matrix $\mathbf{K}$ associated with y_j, and (from equation 6.25) the gain function $\mathbf{g}_j$ for the individual measurement y_j is the $(n \times 1)$ column vector

$$\mathbf{g}_j = \mathbf{S_a}\mathbf{k}_j(\sigma_j^2 + \mathbf{k}_j^T\mathbf{S_a}\mathbf{k}_j)^{-1} \tag{7.10}$$

where σ_j is the jth diagonal element of the original measurement error covariance matrix, i.e. $\mathbf{S}_{\epsilon,jj}$. The error covariance for the retrieval from the individual measurement $\hat{\mathbf{S}}_j$ can be calculated from

$$\hat{\mathbf{S}}_j = (\mathbf{I}_n - \mathbf{g}_j\mathbf{k}_j^T)\mathbf{S_a} \tag{7.11}$$

(c.f. equation 6.28). The information content of the individual measurement is given by the difference in entropy before and after the measurement

$$H_j = \frac{1}{2}\ln|\mathbf{S_a}| - \frac{1}{2}\ln|\hat{\mathbf{S}}_j| = -\frac{1}{2}\ln|\mathbf{I}_n - \mathbf{g}_j\mathbf{k}_j^T|. \tag{7.12}$$

The method of selecting channels sequentially begins by calculating the information content spectrum, i.e. $\mathbf{H} = [H_1, \ldots, H_j, \ldots, H_m]$ where H_j is given in equation 7.12. The individual channel which maximises the information content is selected first, then at each stage the next channel selected from those not yet chosen is the channel that subsequently maximises the information content. Effectively, at each stage i the estimate of the state is sequentially updated by including the measurement y_i as the ith component of the measurement vector. The sequential estimate of the retrieval is obtained from

$$\mathbf{x}_i = \mathbf{g}_i y_i + (\mathbf{I}_n - \mathbf{g}_i\mathbf{k}_i^T)\mathbf{x}_{i-1} \tag{7.13}$$

where $\mathbf{k}_i^T$ is the row of the weighting function matrix corresponding to the measurement y_i and the sequentially updated gain function is

$$\mathbf{g}_i = \hat{\mathbf{S}}_{i-1}\mathbf{k}_i(\sigma_i + \mathbf{k}_i^T\hat{\mathbf{S}}_{i-1}\mathbf{k}_i)^{-1}. \tag{7.14}$$

The retrieval covariance for i channels is given by

$$\hat{\mathbf{S}}_i = (\mathbf{I}_n - \mathbf{g}_i\mathbf{k}_i^T)\hat{\mathbf{S}}_{i-1} \tag{7.15}$$

and thus the change in information content on introducing channel i is

$$\delta H_i = \frac{1}{2}\ln|\hat{\mathbf{S}}_{i-1}| - \frac{1}{2}\ln|\hat{\mathbf{S}}_i|. \tag{7.16}$$

Before any channels have been selected $\hat{\mathbf{S}}_0 = \mathbf{S}_a$. So for $i = 1$, equations 7.13 to 7.16 are

equivalent to the individual measurement case in equations 7.9 to 7.12, as the first channel selected is done so on the basis of the information spectrum values.

If there are a high number of channels to chose from, the selection of channels sequentially is still computationally intensive. The process can be made to be numerically more efficient, following Rodgers (1996), by re-writing equation 7.15 more conveniently:

$$\hat{\mathbf{S}}_i = \hat{\mathbf{S}}_{i-1} \left(\mathbf{I}_n - \frac{\mathbf{k}_i (\hat{\mathbf{S}}_{i-1} \mathbf{k}_i)^T}{[1 + (\hat{\mathbf{S}}_{i-1} \mathbf{k}_i)^T \mathbf{k}_i]} \right) . \tag{7.17}$$

The change in information content, equation 7.16, then becomes

$$\delta H_i = \frac{1}{2} \ln(1 + \mathbf{k}_i^T \hat{\mathbf{S}}_{i-1} \mathbf{k}_i) . \tag{7.18}$$

Thus, the determinant of the whole matrix is no longer required, and to test the information content as a figure of merit for each of the remaining channels requires only an evaluation of $\mathbf{k}_i^T \hat{\mathbf{S}}_{i-1} \mathbf{k}_i$, which is m^2 operations.

The expressions derived above do not take systematic errors, such as model parameter error and forward model error, into account in the measurement covariance. The inclusion of systematic errors in sequential estimation is discussed in Dudhia et al. (2002). However, these are not taken taken into account in this thesis, primarily due to computational cost.

7.4.2 Optimal Channels for Cirrus Retrievals

As the forward model is nonlinear, the weighting function depends on state and therefore the optimal channels depend on state. In theory, the true state would have to be known *a priori* to obtain the best subset of channels for the retrieval. However, as this is clearly not possible and would defeat the point of the retrieval, a range of cloud states are investigated and the optimal channels selected for each permutation. The initial estimate for a given MIPAS measurement can then be used to select the set of channels for the retrieval on the basis that the solution should be close enough to the first guess such that the selected channels are still optimal (or near optimal).

The same reasoning as employed in section 5.2 is used to choose the sampling of state space. State space is sampled in terms of cloud top height C_{top}, effective radius r_e and infrared optical thickness τ_{ir} at 12 μm (given by equation 5.4):

- The cloud top heights of interest will be 2, 3, and 4 km above the viewing tangent height.

- Hexagonal column crystals are assumed and the same size distributions as in section 5.2 are used such that the effective radius is sampled over nine values: 6.5, 9.2, 14.3,

20.2, 30.8, 40.0, 51.0, 60.9 and 72.0 μm.

- The optical thickness is sampled over the range $0.001 \leq \tau \leq 1.0$.

As the measurements will also depend on the additional atmospheric parameters, so too will the channel selection, although the influence will be of low order. The channel selection is only investigated at the nominal tangent heights which will be supplied to the retrieval, i.e. 9, 12 and 15 km. The channel selection is assumed to be constant over the tangent height uncertainty range, with respect to a relative cloud top height. In addition, as in section 5.2 a standard tropical/equatorial climatology atmosphere is assumed (Remedios, 1999) for temperature, pressure and atmospheric gas concentrations. Therefore, the channel selection was only investigated in this thesis for high tropical cirrus assumed to be in the region 30 °S to 30 °N.

For the channel selection, the information was used as the figure of merit to select the remaining channels and the measurement vector consisted of the radiances at each of the $m = 8601$ spectral points in the range 750–965 cm^{-1}. The spectral range had to be restricted due to computational limitations because the evaluation of the change in information content for each channel requires m^2 operations and the computer memory requirements limits the length of vectors and the size of the matrices to which the linear algebraic operations can be applied efficiently. The range 750–965 cm^{-1} was chosen as it has been the main region of interest throughout this thesis.

In general, the channel selection is successful in obtaining a figure for degrees of freedom for signal close to 4 while selecting less channels than used in the measurement vector in the initial approach.

Figure 7.4(a) plots the information spectrum for a cloud state with r_e=9.2 μm, $C_{\text{top}} = t_h + 2\,\text{km}$ and τ_{ir}=0.1 for a viewing tangent height of 15 km and figure 7.4(b) plots the total information of selected channels against the wavenumber of the most recently selected channel, on an inverted information scale. The tallest line is the first channel selected, and the shortest is the last. Only the first 200 channels selected are plotted and for these channels $d_s = 3.996$. The measurement at 765.875 cm^{-1} has the most information, but there is a range of channels in the surrounding region that have very similar amounts of information. This channel is selected first, and the state vector updated. A new information spectrum can then be computed for the remaining channels, and it is found that the information is decreased in the channels similar to 765.875 cm^{-1} and the second channel to be selected (as can be seen from figure 7.4(b)) is at 964.125 cm^{-1}. The sequence of remaining channels to be selected can be followed from figure 7.4(b) from the monotonically decreasing heights of the selected channel lines. Interestingly, there are no channels in the 940 to 950 cm^{-1} region (used in the initial approach retrieval) selected in the first 200 measurements.

Of the first 200 channels selected in figure 7.4(b), they appear to be grouped into (6) distinct regions (or meso-windows). This characteristic is extremely useful for limiting the computing resources required for the radiative transfer calculations in McClouds_FM. The radiative transfer calculations need only cover the regions in which the channels lie by using a small number of contiguous meso-windows, rather than covering the entire spectral range of 750 to 965 cm^{-1} which is time consuming to compute and requires large amounts of memory to store and manipulate the output (due to the high spectral resolution required).
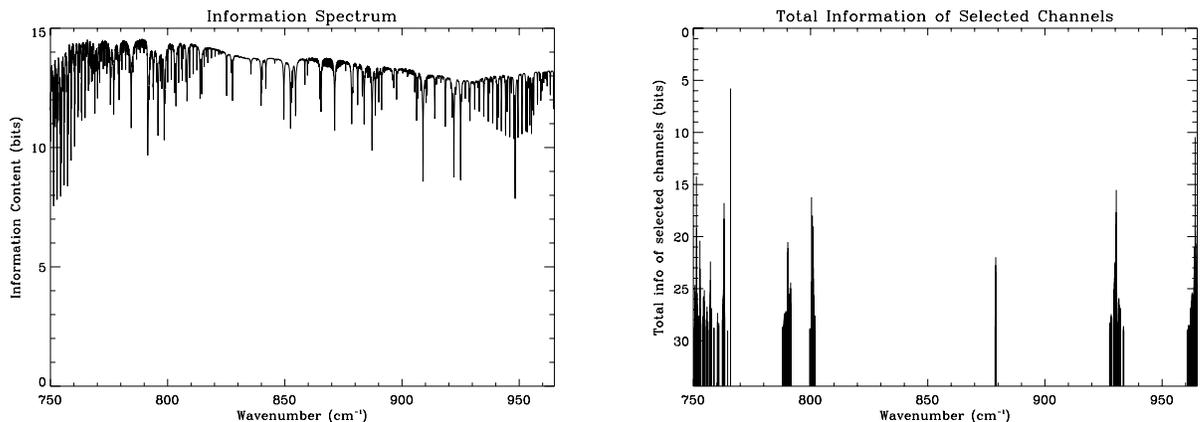
(a) Information content of individual channels plotted against wavenumber.

(b) Total information of selected channels plotted against the wavenumber of the most recently selected channels.

Figure 7.4: (a) The information spectrum and (b) the total information of sequentially selected channels for a cloud with r_e=9.2 μm, $C_{\text{top}} = t_h + 2$ km and τ_{ir}=0.1 for t_h=15 km.

The usefulness of the full spectrum can be evaluated with respect to retrieving cloud properties by plotting the information content against the number of channels selected. An example is given in figure 7.5(a) for several typical cloud types at a viewing tangent height of $t_h = 12$ km, with the degrees of freedom for signal plotted against the number of channels in figure 7.5(b). The cloud top height is $C_{\text{top}} = t_h + 3$ km and clouds of two different effective radii, 6.5 and 30.8 μm, are each plotted for two values of optical thickness, 0.05 and 0.5. Both the information content and the degrees of freedom for signal increase rapidly as the first three channels selected are added to the measurement vector. For all of the cloud types in the example, after selecting 3 channels the degrees of freedom for signal is ≈ 3, indicating that each of the measurements corresponds to an independent (orthogonal) piece of information. To obtain the fourth degree of freedom requires significantly more channels, although, for the larger effective radius and the lower optical thickness at the smaller radius, the 4$^{\text{th}}$ channel adds the majority of a degree of freedom, i.e. for 4 channels $d_s \gtrsim 3.75$. For the cloud type with $\tau = 0.5$ and $r_e = 6.5$ μm, after the 3$^{\text{rd}}$ channel is

selected, the degrees of freedom does not increase again signifcantly until over 20 channels have been added to the measurement vector. Likewise, the optically thicker cloud type for $r_e = 30.8$ μm also requires a higher number of measurements than the thinner cloud to achieve an equivalent value of degrees of freedom for signal. The information content (in figure 7.5(a)) also shows a similar trend, in that the optically thicker clouds, at both effective radii, show a much lower information content for the same number of channels selected. Figure 7.5(a) also shows that when the number of channels selected is greater than about 10 the information content is approximately linear in the logarithm of the number of channels selected and even after 5000 measurements does not appear to be levelling off. There is no simple explanation for this linearity (see, e.g. Rodgers, 1996).
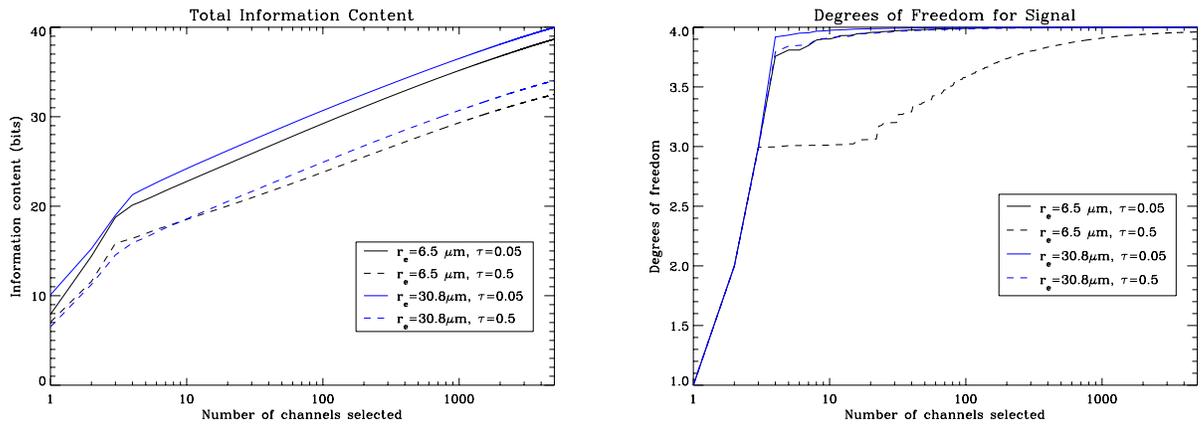
(a) Information content.　　　　　　　(b) Degrees of freedom for signal.

Figure 7.5: The information content (a) and the degrees of freedom for signal (b) plotted against the number of channels selected, for a cloud with $C_{\text{top}} = t_h + 3$ km at $t_h = 12$ km for r_e and τ_{ir} values as given in the plot keys.

The relationship between the figures of merit and the number of channels varies very little with cloud top and is similar at different tangent heights for the same relative cloud top heights and r_e and τ values. However, as shown in figure 7.5, the quality indicators are dependent on both the r_e and τ values for which they are calculated (assuming cloud top and tangent height are constant). This relationship is plotted in figures 7.6 and 7.7 for a viewing tangent height of 15 km and $C_{\text{top}} = t_h + 3$ km. The information content and degrees of freedom were calculated for a measurement vector after selecting the first 10 channels, plotted in figure 7.6, and after selecting the first 40 optimal channels, plotted in figure 7.7. Although these two plots do not show the exact dependence of the figures of merit on the number of channels (as such a plot would be difficult to visualise), by choosing two 'markers' of the number of channels, comparing the quality indicators in the two figures will give an indication of the trend. The first channel marker is chosen at 10 channels because (from

figure 7.5) the trends for typical cirrus values are evident at this stage and the information content begins to increase linearly with respect to the logarithm of the number of channels. The second marker is chosen at 40 channels somewhat arbitrarily, because it is ten times the number of state parameters and a tenth of the number of channels used in the initial approach retrieval and also, figure 7.5 shows that for the higher optical thickness cases the degrees of freedom only begins to increase signifcantly again after 10 channels once 40 channels are selected.
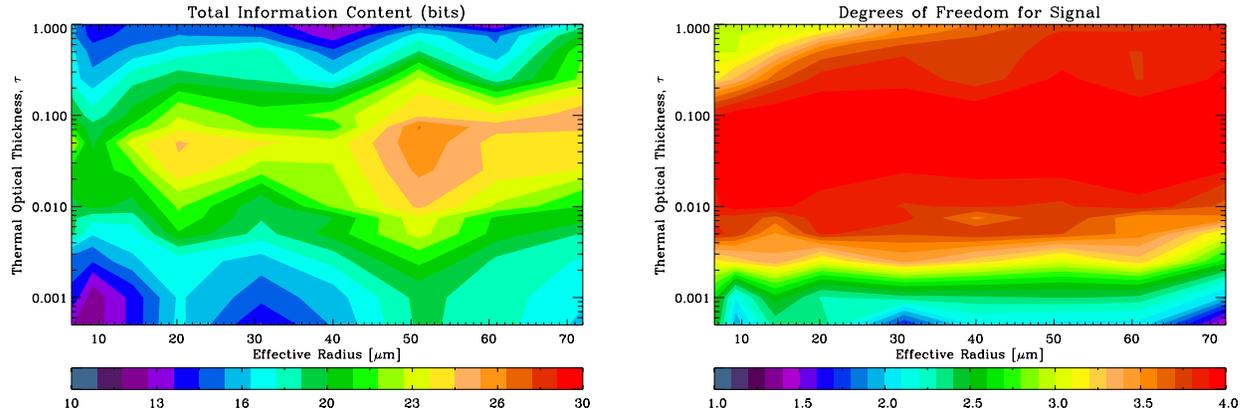
(a) Information content for first 10 optimised channels.

(b) Degrees of freedom for signal for first 10 optimised channels.

Figure 7.6: The total information content (a) and degrees of freedom for signal (b) after 10 channels have been included in the measurement vector, plotted as a function of effective radius and optical thickness at which the linearised channel selection was applied (for $t_h = 15$ km and $C_{\text{top}} = t_h + 3$ km).

The information content after 10 channels are selected (figure 7.6(a)) is greater than 10 bits everywhere, but does not vary very much with effective radius, with the exception of two peaks at approximately 20 and 50 μm. However, there is a strong dependence on optical thickness and the information content peaks in the region between about 0.02 and 0.1.

After 10 channels have been selected the degrees of freedom for signal (see figure 7.6(b)) is only weakly dependent on effective radius and, as intuitively expected, is much lower for the more tenuous cirrus. Below the limit of MIPAS cloud detection, i.e. $\tau_{\text{ir}} \lesssim 0.005$, it is found that $d_s \lesssim 3.5$. The degrees of freedom peak in the region $0.01 \lesssim \tau_{\text{ir}} \lesssim 0.1$, which is the same region in which the peak information content values occur and more importantly, the same region in which the multiple scattering effect peaks (see chapter 5). It is logical that the scattering effects, which are governed by the cloud microphysics, will have a large effect on the degrees of freedom. Thus, for high optical thickness, which is dominated by cloud emission, and for very low optical thickness, which has low extinction and thus low order scattering, the degrees of freedom for signal is much lower as the measurements are

less influenced by the cloud microphysics.

The absolute value of the information content after selecting 40 channels (see figure 7.7(a)) has increased with the greater number of channels in the measurement vector. It is essentially a linear scale increase as the dependence on cloud state is near identical to the dependence of the information content after 10 channels. However, the degrees of freedom for signal after 40 channels are selected shows a noticeable difference compared to the equivalent figure with only 10 channels. The degrees of freedom is still $\lesssim 3.5$ for optical thickness $\lesssim 0.005$, but $d_s \approx 4.0$ for all r_e and $\tau_{ir} \gtrsim 0.005$, excluding the region of state space where $\tau_{ir} \gtrsim 0.3$ and $r_e \lesssim 20.0~\mu m$. Where $d_s < 4.0$ the degrees of freedom for signal will increase slowly (toward 4.0) as additional channels are selected.

Figures 7.6 and 7.7 show that it is possible to optimise channel selection for detectable cirrus (i.e. $\tau_{ir} \gtrsim 0.005$) on the basis of maximising information content through sequential update, such that the degrees of freedom for signal will be near 4. In general, the number of channels required to achieve high degrees of freedom for signal is an order of magnitude less than the number of channels used in the initial approach retrieval, and is only a small fraction of all the available MIPAS spectral channels.
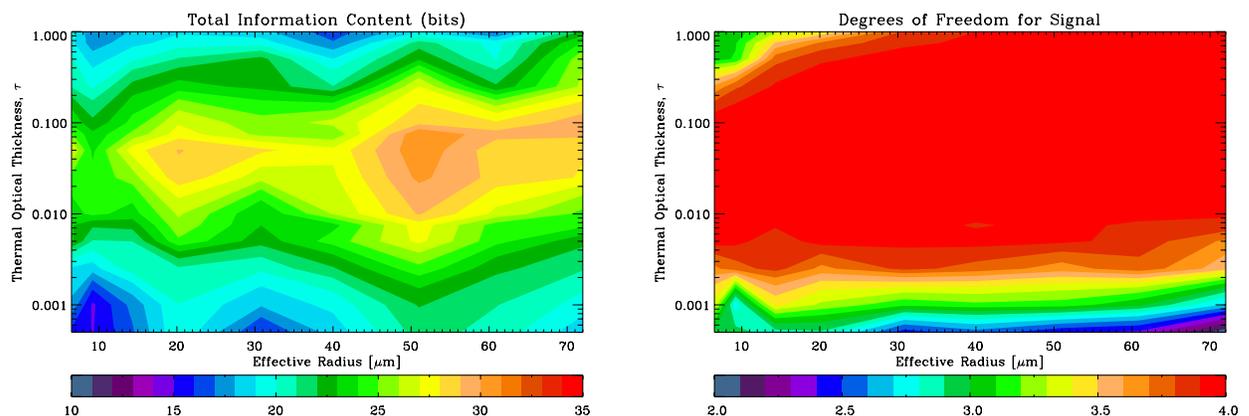
(a) Information content for first 40 optimised channels.

(b) Degrees of freedom for signal for first 40 optimised channels.

Figure 7.7: The total information content (a) and degrees of freedom for signal (b) after 40 channels have been included in the measurement vector, plotted as a function of effective radius and optical thickness at which the linearised channel selection was applied (for t_h=15 km and $C_{top} = t_h$+3 km).

7.4.3 Cirrus Retrieval Accuracy using Optimal Channels

The information content and degrees of freedom discussed above, are useful as an indicator of the quality of the retrieval and for the purpose of the channel selection, but they give no indication about the accuracy of the resulting retrieval. To quantify the expected retrieval

accuracy the error variance must be examined. The error variance was evaluated for two selections of channels, namely the first 40 channels and the first 400 channels. Figures 7.8 and 7.9 plot the error variances for the first 40 channels and first 400 channels, respectively.

(a) Effective radius standard deviation.

(b) Number density standard deviation.

(c) Cloud top height standard deviation.

(d) Vertical cloud depth standard deviation.
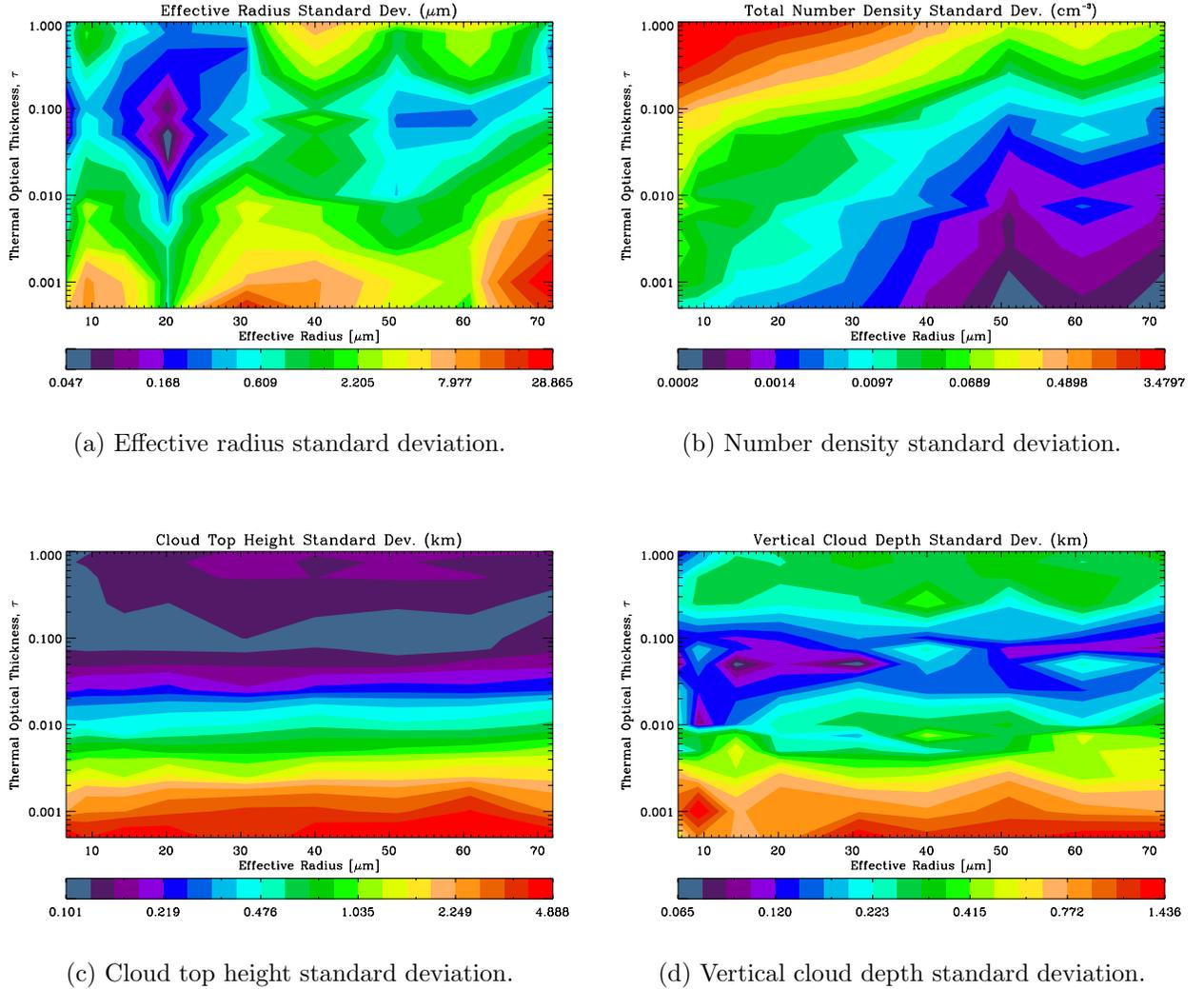
Figure 7.8: Square root of the variance (i.e. the square root of the diagonal elements of the total covariance matrix $\hat{\mathbf{S}}$) for each of the cloud state parameters using the first 40 channels selected for total information content. The square root of the variance, i.e. the standard deviation, is plotted as a function of the true optical thickness and effective radius assuming $t_h = 15$ km and the true $C_{\mathrm{top}} = t_h + 3$ km.

Figure 7.8 shows that even 40 channels can be expected to give a reasonably useful cloud retrieval. An expected accuracy of effective radius to within ~ 2 μm and cloud top height to less than ~ 0.5 km for all effective radii and for optical thickness $\gtrsim 0.01$. The vertical cloud depth and number density can also be expected to be retrieved relatively accurately. With the accuracy of cloud depth highest when multiple scattering has the greatest effect, and expectedly, the accuracy of the number density is worse for smaller effective radius where number density is higher and more variable. From the corresponding degrees of freedom
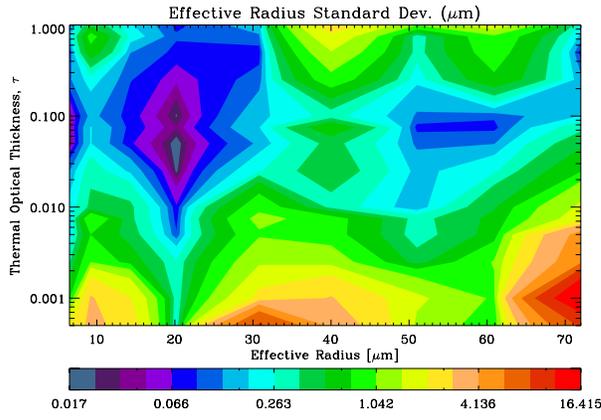
(a) Effective radius standard deviation.
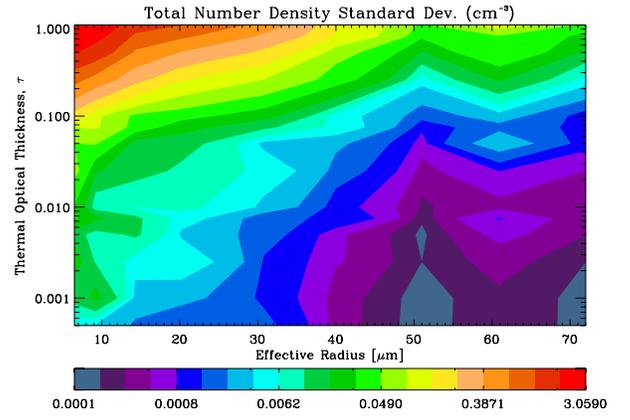
(b) Number density standard deviation.
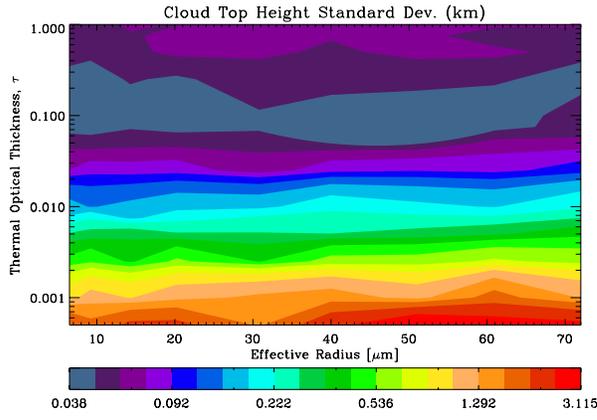
(c) Cloud top height standard deviation.
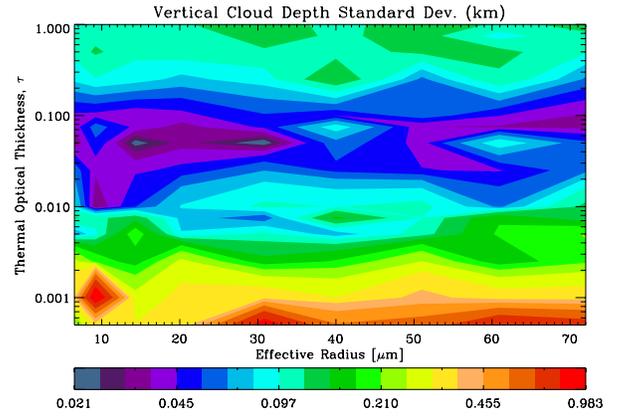
(d) Vertical cloud depth standard deviation.

Figure 7.9: As figure 7.8, but for the first 400 channels.

in figure 7.7(b), for $\tau_{\mathrm{ir}} \gtrsim 0.01$, the retrievals can be expected to have high 'resolution' (i.e. $\mathbf{A} \longrightarrow \mathbf{I}$ for $d_s \sim 4$), so that the information on the cloud parameters should be near independent.

At 400 channels the improvement is considerable. Figure 7.9 shows that the retrieval noise is much lower than at 40 channels, with $\sigma_{r_e} \lesssim 1\ \mu$m, $\sigma_{C_{\mathrm{top}}} \lesssim 0.2$ km, and $\sigma_{C_{\mathrm{depth}}} \lesssim 0.2$ km for all $\tau_{\mathrm{ir}} \gtrsim 0.01$. However, the number density error does not improve as much as for the other cloud parameters and it is still influenced by the dependence of number density on both effective radius and optical thickness.

The information content increases linearly with the logarithm of the number of channels, and the higher information content at 400 channels is used to improve the retrieval 'resolution' as well as the noise, such that the degrees of freedom for signal at 400 is an improvement over the value at 40 channels given in figure 7.7(b).

The results for the retrieval quality and the retrieval errors above are given for the opti-

mised channels at each of the sampled points in state space. The channels are based on information content which assumes linearisation about each of the points. In McClouds_RT the optimised channels to be used are selected from the first guess and it is found from experience that linearisation about the first guess is not valid for finding a solution, hence the use of the Levenberg-Marquardt method. If the retrieval iterates to a cloud state for which the channels in the measurement vector are no longer optimal in terms of information content, then it is possible that the retrieval solution may have less degrees of freedom for signal and higher retrieval errors than for the optimised case. Once a solution has been found the averaging kernel and the covariance matrix can be examined to determine if the measurement vector channels are no longer optimal, if so, the retrieval can be repeated, using the channels which would be optimal at the solution and using the solution as the first guess. This should improve the estimate of the cloud state and the sensitivity of the retrieval, while reducing the error bars.

The extent to which this is expected to be a problem can be determined by examining the dependence of the spectral location of the optimal channels on the cloud state. In general the location of the selected channels does not vary much with cloud top height and the trends are similar at different tangent heights for the same relative cloud states. There is variation in the selected channels with respect to effective radius and optical thickness of the linearisation point. Figure 7.10 plots the spectral location of optimal channels for a typical case, with $t_h = 15$ km and $C_{\text{top}} = t_h + 3$ km. Figure 7.10(a) plots the number of selected channels per wavenumber as a function of r_e for the first 500 channels selected, assuming a constant optical thickness of 0.5. It can be seen from the figure that the location of the channels is relatively constant with respective to r_e and thus the optimality of the selected channels is unlikely to change significantly with steps in r_e of the order of the expected retrieval error bars (see figures 7.8(a) and 7.9(a)). This is generally the case for a given constant optical thickness. However, if r_e is held constant then the location of the selected channels does vary with respect to the optical thickness of the linearisation point. This dependence on optical thickness can be seen in figure 7.10(b), which assumes a constant effective radius of 9.2 μm. The spectral location of the optimal channels does vary noticeably with optical thickness, however, bearing in mind figure 7.10(b) is presented using the logarithm of τ_{ir} as the y-axis, and that optical thickness will not vary much (i.e. much less than an order of magnitude) during the retrieval if the first guess values of N and C_{depth} are close to the solution, then the channels should remain near optimal during the retrieval. In any case, if upon finding the solution, the optical thickness has changed considerably from the first guess such that degrees of freedom is much lower than 4, then repeating the retrieval as described above will improve the retrieval quality.
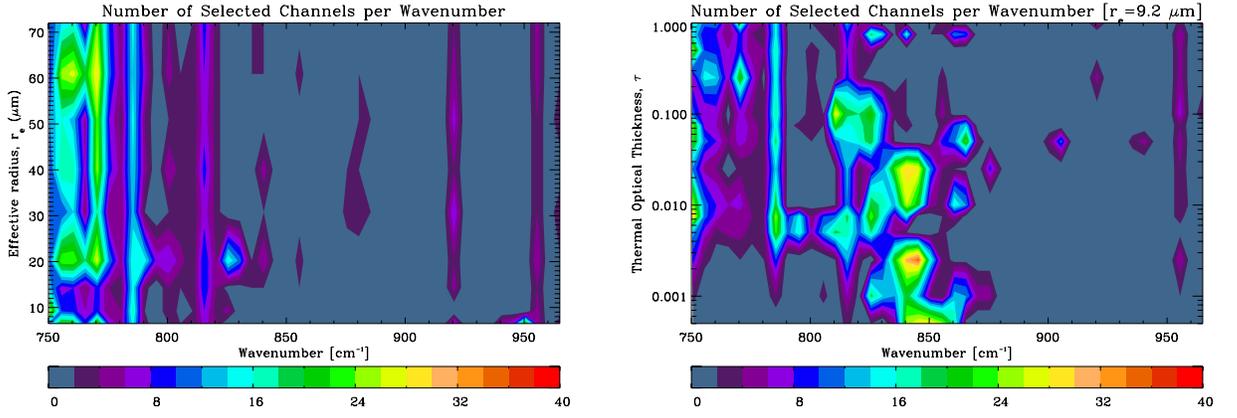
(a) Number of channels per wavenumber, plotted as effective radius against wavenumber.

(b) Number of channels per wavenumber, plotted as optical thickness against wavenumber.

Figure 7.10: Locations of the first 500 optimised channels. Example cases are plotted (for $t_h = 15$ km and $C_{\text{top}} = t_h + 3$ km) as the number of channels per wavenumber as a function of effective radius in (a), assuming $\tau_{\text{ir}} = 0.5$ is constant, and also as a function of optical thickness in (b), assuming $r_e = 9.2$ μm is constant.

7.4.4   Test Case Results

In order to demonstrate the effectiveness of the optimised channel retrieval it was applied to the same test case as used for the initial approach in section 7.3.1. The results for the real and synthetic retrievals are given in tables 7.11 and 7.12 respectively, with the corresponding averaging kernels in equations 7.19 and 7.20. The set of channels to be used in the retrieval was selected based on the first guess values in tables 7.11 and 7.12, and is the same set in both cases for which the selected channels are given in figure 7.4. The retrievals were run using only the first 100 selected channels. The retrieval residuals with respect to channel number are given in figure 7.11, where the channels have been sorted to increment with wavenumber.

$$\mathbf{A}_{\text{real}} = \begin{bmatrix} 0.9988 & -0.1876 & -1.388 \times 10^{-3} & 0.3468 \\ -1.201 \times 10^{-4} & 0.7800 & -9.828 \times 10^{-4} & 0.3999 \\ -4.499 \times 10^{-5} & -4.971 \times 10^{-3} & 0.9992 & 6.579 \times 10^{-3} \\ 4.496 \times 10^{-4} & 0.0810 & 2.632 \times 10^{-4} & 0.8516 \end{bmatrix} \tag{7.19}$$

$$\mathbf{A}_{\text{synth}} = \begin{bmatrix} 0.9997 & -6.349 \times 10^{-3} & 3.961 \times 10^{-4} & 0.0286 \\ -4.064 \times 10^{-5} & 0.9117 & 1.984 \times 10^{-3} & 0.1614 \\ 1.284 \times 10^{-5} & 0.0101 & 0.9997 & -0.0195 \\ 3.716 \times 10^{-5} & 0.0327 & -7.828 \times 10^{-4} & 0.9386 \end{bmatrix} \tag{7.20}$$
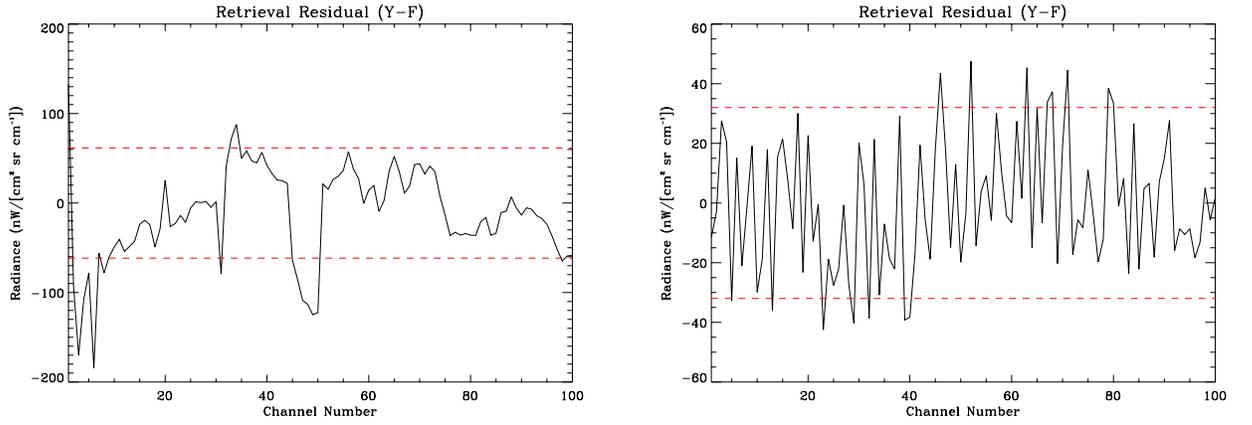
(a) Real case retrieval residual.

(b) Synthetic case retrieval residual.

Figure 7.11: Residuals (i.e. $\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}})$) from the real and synthetic optimised channel retrievals for the test case. Overplotted in red is the root mean square measurement error (due to measurement noise and, in the real case, random forward error).

Cloud Property	$\mathbf{x}_0$	$\hat{\mathbf{x}}$	$\hat{\sigma}$	[%] Uncertainty
r_e	10.365	10.284	0.846	8.2
N	1.175	1.138	0.880	77.3
C_{top}	18.542	17.694	0.124	0.71
C_{depth}	0.414	0.479	0.120	31.7

Table 7.11: Retrieval results from test case using actual MIPAS measurements, applying the channel optimisation method (where 100 channels were selected and at the first guess it's expected that $d_s = 3.992$) and assuming ECMWF atmospheric temperature and pressure data. Retrieval cost; $\chi^2 = 70.73$, $\chi^2/df = 0.71$, $\chi^2/d_s = 17.71$.

Cloud Property	$\mathbf{x}_{true}$	$\mathbf{x}_0$	$\hat{\mathbf{x}}$	$\hat{\sigma}$	[%] Uncertainty
r_e	7.85	11.347	8.236	0.428	5.2
N	4.76	1.445	3.842	0.594	15.5
C_{top}	18.75	18.796	18.755	0.081	0.43
C_{depth}	0.7	0.386	0.736	0.223	30.3

Table 7.12: Retrieval results from test case using synthetic measurements, applying the channel optimisation method (where 100 channels were selected and at the first guess it's expected that $d_s = 3.992$) and ECMWF atmospheric temperature and pressure data. Retrieval cost: $\chi^2 = 47.51$, $\chi^2/df = 0.0.48$, $\chi^2/d_s = 11.90$.

Corresponding Component	$\mathbf{u}_1$	$\mathbf{u}_2$	$\mathbf{u}_3$	$\mathbf{u}_4$
r_e	-0.6426	0.7364	0.1940	0.0844
N	-0.7181	-0.6053	-0.2019	0.2777
C_{top}	-0.0204	0.2538	-0.9484	-0.1890
C_{depth}	0.2663	0.1640	-0.1488	0.9381
Eigenvalue	$\lambda_1=1.6722$	$\lambda_2=0.0462$	$\lambda_3=0.0131$	$\lambda_4 = 1.173 \times 10^{-4}$

Table 7.13: The eigenvectors, $\mathbf{u}_i$, and eigenvalues, λ_i, of the total covariance matrix of the real case retrieval.

Error Pattern	Nonlinearity Parameter
1	0.6342
2	0.3578
3	0.1972
4	0.1153

Table 7.14: The nonlinearity parameters for the error patterns in the real case retrieval (table 7.11).

In order to confirm that a linear error analysis is applicable, the degree of non-linearity at the solution was calculated for both the real and synthetic retrievals. The eigenvalues and eigenvectors of the total covariance matrix from the real retrieval are given in table 7.13, form which the error patterns can be obtained. Following the same method as in section 7.3.2, $\delta\mathbf{y}$ was investigated for values of $\mathbf{x}$ at the equivalent of one standard deviation away from $\hat{\mathbf{x}}$ using the error patterns of $\hat{\mathbf{S}}$, and table 7.14 gives the nonlinearity parameter ζ^2/m which was calculated for each of the error patterns. The nonlinearity parameters in table 7.14 are less than unity, therefore the retrieval is near linear at the solution within the error bounds of the retrieval. The synthetic retrieval was also shown to be near linear at the solution by calculating $\delta\mathbf{x}$, as given by equation 7.6, which was found to be

$$\delta\mathbf{x} = [0.2697\,\mu\text{m}, -0.3615\,\text{cm}^{-3}, 0.0088\,\text{km}, 0.1078\,\text{km}],$$

i.e. considerably smaller (by approximately a factor of 2) than the retrieval uncertainty. The results from the real retrieval in table 7.11 differ very little (i.e. within the error bars) from the initial guess in all cloud parameters apart from the cloud top height which is 1 km lower. The errors on the solution are a great improvement from the *a priori*, but the errors are much larger than expected from the calculated retrieval accuracy (detailed in the previous section). The reason for this is that the degrees of freedom for signal is $d_s =3.630$ which is ~ 0.3 less than the value expected for the optimised channels. The loss of degrees of freedom is due to the high correlations between effective radius and number density,

effective radius and cloud depth, and number density and cloud depth, which can be seen from the averaging kernel in equation 7.19. The channels in the measurement vector are theoretically near optimal for the solution, but in the previous retrieval accuracy analysis the correlations in the cloud parameters were not accounted for. Therefore, repeating the retrieval is unlikely to yield an improved retrieval accuracy. Furthermore, from the retrieval residual, it can be seen that the modelled measurement at the solution has trouble fitting the measurement vector in channels 1–9 and 45–52, which are in the wavenumber regions ~ 760–770 cm^{-1} and ~ 875–880 cm^{-1} (see figure 7.4). The reason for this may be due to the spectral variance of the scattering properties, as the overall cost is low. Due to the problems with fitting the measurement over a wider spectral range, increasing the number of channels in the measurement vector does not significantly reduce the retrieval uncertainty. The spectral variance depends on the size distribution and crystal habit, therefore if the wrong size distribution or crytal shape is assumed in the cloud microphysics, then even if the effective radius and number density are correct the model may not be able to match the measured spectra across the entire range. It may be possible to improve the optimal channel selection by including correlations between cloud parameters in the *a priori* information, however, it will entail further study, either based on measurement campaigns or retrieval output statistics to obtain accurate covariance values. An investigation into the effect of the crystal habit and the shape of the size distribution on the modelled measurements and retrieved properties is left for future work.

For the real case the optimal channel retrieval has a greater number of degrees of freedom for signal and an improved retrieval accuracy over the initial approach retrieval results in section 7.3.1. Interestingly, the cloud top height is about 2 km lower in the optimised retrieval. This places the cloud top height at the tropopause, which is its most likely location if it is indeed cold trap cirrus. It may be that the optimal channels are sensitive to absolute cloud top height, whereas the initial approach retrieval is only sensitive to a cloud top height relative to the tangent height. A greater number of comparisons between the two approaches are needed before any conclusions on this can be drawn. If there is statistical significance then it suggests that the tangent height should be included in the state vector in the initial approach.

The synthetic retrieval results in table 7.12 differ significantly from the first guess and are, with the exception of number density, within the error bars of the true state. However, the number density is close to the true state and the difference is due to the lower sensitivity (the second diagonal element of equation 7.20) for number density compared to the other cloud parameters. The retrieval uncertainty is extremely low and the residual in figure 7.11(b) is simply white noise within the measurement NESR. The degrees of freedom for signal is 3.850 at the solution, which is near optimal for 100 channels and this high value

leads to a solution very close to the true state which shows that the retrieval method is self-consistent and the first guess is close enough to the solution for the retrieval to converge to the global minima.

7.5 Conclusion

An effective methodology has been developed to retrieve cloud properties from MIPAS *via* two optimal estimation approaches:

- Using a contiguous channel measurement vector from the atmospheric window region

- Using a set of isolated channels in the measurement vector which are optimised to maximise information content.

The optimised channel method in particular has been shown theoretically to be a powerful retrieval method, capable of returning high accuracy and high resolution information. The contiguous channel approach is also very useful with respect to the information returned, but it has the advantage that is much less computationally demanding than the optimised channel retrieval. At the time of writing, the optimal channel version of McClouds_RT takes ~ 100 minutes per iteration (using 5×10^5 photons per monochromatic McClouds_FM calculation performed over a 0.25 μm grid) running on a 1.1-GHz CPU with 1 GB of RAM. This CPU time is prohibitively slow and must be improved if processing large MIPAS datasets is to prove feasible.

For each of the methods, the test case presented (which is typical of the results obtained in general), showed that both methods are self-consistent (from the synthetic retrievals) and both methods are highly sensitive to the effective radius and cloud top height. The retrieval is not as sensitive to number density and cloud depth, which is reflected in the high uncertainties on these parameters in the retrieval. However, the optimised channel retrieval returned an acceptable accuracy and high degrees of freedom for signal. Sections 7.4.2 and 7.4.3 showed that theoretically, the optimised retrieval should be able to achieve a degrees of freedom for signal > 3.8 and retrieve (for typical thin cirrus values):

- Effective radius to within 2 μm

- Number density to better than 20 %

- Cloud top height to better than 0.2 km

- Cloud depth to better than 25 %.

The synthetic retrieval is able to achieve these error margins (with the exception of cloud depth), but the forward model has some trouble fitting the real measurements and some of

the cloud parameters are also correlated, which results in the real retrieval falling a little short of the expected retrieval accuracy. The degrees of freedom is also lower than expected but still higher than that achievable by the initial approach.

Generally, the first guess for the cloud state, using the method in chapter 5, can be expected to be close to the global minima as shown by the synthetic retrieval results. Further synthetic retrievals should be carried out in future to show statistical significance of this result. However, the retrieved number density is smaller than the true value in both synthetic test case retrievals. The initial guess number density is likely biasing the retrieval, due to the low sensitivity to number density, as seen from the averaging kernels. The number density is determined from the r_e and τ pair generated by the first guess method (see section 5.3). This is selected based on maximising the probability given an IWC *pdf*. The IWC *pdf* used in this thesis was chosen to be representative of very-thin high cirrus (see McFarquhar et al., 2000) and the mean is at relatively low IWC values. If the cloud is not well represented by this IWC *pdf*, i.e. it is thicker cirrus, then the first guess of number density will be artificially low. Further investigation should be carried out into the effect on the first guess of varying the IWC *pdf* and the subsequent effect on the retrieved number density.

Succinctly, this chapter has shown that it is possible to retrieve up to four independent pieces of information on cloud state to good accuracy from MIPAS measurements. Further processing of MIPAS data is required in order to fully validate McClouds_RT by comparing monthly or annual statistics to external datasets, or by comparing MIPAS results to independent measurements from either coincident satellite measurements or *in situ* measurements. Preliminary results following the latter approach are discussed in the following chapter.

Chapter 8

Validation of the Retrieval

8.1 Introduction

When McClouds_RT is applied to actual MIPAS data it is not possible to actually determine whether the retrieval is consistent with state, because the true state is unknown. The best way to validate the retrieved quantities is to compare them to independent measurements. With respect to remote measurements from satellites, independent instruments rarely make spatially and temporally coincident measurements. However, MIPAS is part of the ENVISAT payload which also includes the Advanced Along Track Scanning Radiometer (AATSR, see figure 2.4) which is a nadir viewing instrument and in tropical regions it will view the same part of the atmosphere in the nadir that the MIPAS limb view (in the anti-flight direction) will measure approximately 7 minutes after the satellite overpass.

A methodology for the retrieval of cloud properties from ATSR-2 and AATSR data has been developed by Watts (1995), for which the results are available[1]. At present the McClouds_RT processing time for a MIPAS profile is too slow for a practical data analysis over a long time scale (e.g. one month) to be included in this thesis. However, detailed comparisons between the MIPAS and AATSR cloud retrievals can be made on a case by case basis. This will not only facilitate a preliminary validation of the MIPAS retrieval, but for cases where the cirrus is too optically thin to be retrieved by AATSR it will demostrate the difference between the nadir and limb view capabilities for cirrus detection and characterisation.

8.2 The AATSR Instrument

The Advanced Along Track Scannning Radiometer (AATSR), being the successor to ATSR-2, is the latest in the series of the ATSR instruments which build on the heritage of the NOAA

---

[1]From Dr Sam Dean, University of Oxford, *via* private communication

AVHRR sensors. The ATSR instruments are second generation space radiometers which, like AVHRR, exploit multiple spectral channels (originally introduced for the retrieval of sea surface temperature). AATSR records in seven channels, i.e. 0.55 μm, 0.66 μm, 0.87 μm, 1.6 μm, 3.7 μm, 11.0 μm and 12.0 μm.

Each of the ATSR instruments have been designed for exceptional sensitivity and robust calibration stability. These features are achieved through the use of low-noise infrared detectors cooled by a pair of Stirling cycle coolers to near-optimum temperatures (i.e. approximately 80 K), and through continuous on-board radiometric calibration of the infrared channels against two stable, high-accuracy blackbody calibration targets, and for AATSR, calibration of the visible and near infrared channels with an on-board visible calibration system.

The novel feature of the ATSR instruments is the use of "along track scanning" to provide two observations of the same surface scene through differing amounts of atmosphere; the "along track" view passes through a longer atmospheric path and is more affected by the atmosphere than the nadir view. The AATSR's optical path is arranged to provide two curved swaths that give a dual view of the Earth's surface at different angles (see figure 8.1). The AATSR sensor first views the surface along the direction of the orbit track at an incidence angle of 55° as it flies toward the scene. Then, approximately 120 seconds later, the satellite passes directly over the scene and AATSR records a second observation at an angle close to the nadir. The two curved swaths comprising the AATSR field of view are both 500 km-wide with 555 pixels across the nadir swath and 371 pixels across the forward swath. The pixel size is $1.1\,\text{km} \times 1.1\,\text{km}$ at the center of the nadir scan and $1.5\,\text{km} \times 2\,\text{km}$ at the center of the forward scan. The AATSR cloud retrieval (Watts, 1995) uses the geolocated rectified data product which maps the acquired pixels onto a 512×512 grid with 1 km pixel size whose axes are the satellite ground-track and great circles orthogonal to the ground-track.

8.3   Cloud Properties from AATSR

The AATSR cloud retrieval developed by Watts (1995) also employs a methodology based on optimal estimation. Like McClouds_RT, the basic principle of the AATSR cloud retrieval is to maximise the probability of the retrieved state conditional on the value of the measurements and any *a priori* (i.e. to find the MAP solution, see section 6.2.2).

The basic state to be retrieved from the AATSR measurements is defined as

$$\mathbf{x} = [\tau_{\text{vis}}, r_e, p_c, f, T_s] \qquad (8.1)$$
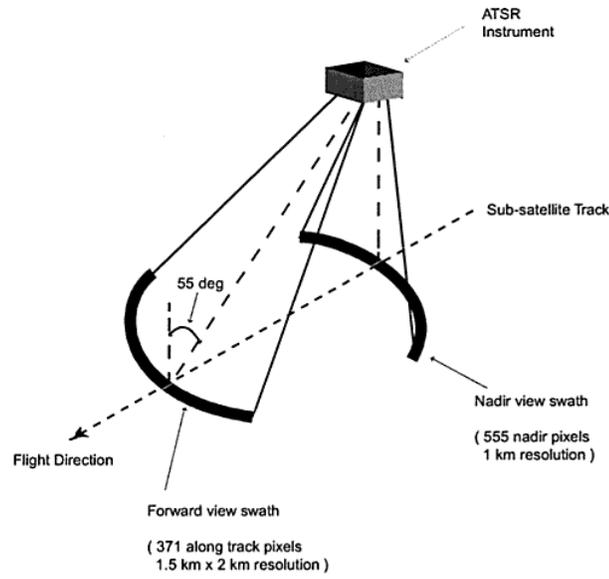
Figure 8.1: AATSR viewing geometry.

where:

- τ_{vis} is the *visible optical depth* which is defined using the optical depth at 0.55 μm (i.e. AATSR channel 1)

- r_e is the *particle size* which covers the cases for water particles where a well defined *effective radius* is available and also covers ice particles where the effective radius is defined by equation 4.2

- p_c is the *cloud top pressure*, which can be converted to a cloud top height in kilometres once a solution is found

- f is the *cloud fraction*

- T_s is the *surface temperature*.

The AATSR cloud retrieval can be applied to both water and ice clouds. The *cloud phase* is effectively determined by the retrieval, by using the phase for which the measurements can be modelled most closely.

The measurement vector is comprised of the full set of available measurements and can be defined for each pixel as:

$$\mathbf{y} = [y_{0.5}, y_{0.6}, y_{0.8}, y_{1.6}, y_{3.7}, y_{11}, y_{12}] \tag{8.2}$$

where a certain shorthand has been used to describe the channel information. The 0.55 μm, 0.66 μm, 0.87 μm and 1.6 μm channel measurements can be assumed to be expressed in terms of percentage reflectance (i.e. albedo) $R = 100 \times \frac{I\pi}{F_0}$, where I is the measured radiance

and F_0 is the solar irradiance. The 3.7 μm, 11 μm and 12 μm channels are assumed to be expressed in brightness temperatures, $T_B = B^{-1}(R, \bar{\nu})$, where B^{-1} is the inverse Planck function, R is the measured radiance and $\bar{\nu}$ is the effective central wavenumber of the channel.

The Levenberg-Marquardt iterative method is used to find a solution to the inverse problem assuming Gaussian statistics. The forward model used to simulate measurements within the iterative scheme is detailed in Watts et al. (1998). The forward model uses a plane-parallel approximation to represent the cloud and assumes it is infinitesimally thin. In addition to the state the forward model also requires the additional model parameters

$$\mathbf{b} = [\mathbf{T}(\mathbf{z}), \mathbf{H}(\mathbf{z}), \mathbf{R_s}, \varepsilon] \tag{8.3}$$

where $\mathbf{T}(\mathbf{z})$ is the atmospheric temperature profile and $\mathbf{H}(\mathbf{z})$ is the humidity profile similarly obtained from ECMWF operational data. $\mathbf{R_s}$ is a vector of surface reflectances defined at the visible and near-infrared channels and ε is a vector of surface emissivities defined at the infrared channels.

There is no real *a priori* data for the cloud parameters in the state vector and as such the *a priori* covariance errors are assumed to be infinite. There is, however, significant *a priori* information on the surface skin temperature. The value is supplied through ECMWF operational data and the error is set to 1 K for ocean scenes and 3 K for land scenes.

8.4 Comparison between MIPAS and AATSR

To intercompare measurements made by remote sounders, Rodgers and Connor (2003) outline a rigorous treatment to account for the differing characteristics of the observing systems, particularly their averaging kernels and error covariances. However, the averaging kernels and error covariance matrices were not available for the AATSR retrieval, and as the retrieved cloud states are comprised of essentially orthogonal quantities which are only slightly influenced by prior knowledge (i.e. a low sensitivity error, unlike atmospheric profiles obtained from remote sounding), the comparison between the MIPAS and AATSR cloud products can be made without the rigorous formal treatment. As such, the retrieved quantities (or functions of them) are directly compared and the errors on the individual state elements treated separately. It is also assumed that the cloud field will not change significantly in the 7 minute time lag between the satellite overpass and the MIPAS measurement.

This section presents the comparison of the MIPAS and AATSR retrievals for two case studies, giving an initial validation of McClouds_RT.

8.4.1 Coincident Measurements

In section 2.5 it was explained that the ENVISAT orbit has a 98.54° inclination and MIPAS operates with a rearward oscillating scanning geometry in the azimuth. The oscillating rearward scan is in place so that the data coverage of MIPAS is a near polar orbit. Therefore, the tangent points of the elevation scan measurements do not exactly overpass the ground footprint of the orbit path. The effect of the scanning geometry can be seen in figure 8.2. The MIPAS tangent points will only overpass the sub-satellite footprint, and thus by definition, the AATSR observation footprints, when the rearward azimuth angle is close to zero. The azimuth angle is zero as the spaceraft overpasses the equator and thus there is a narrow latitude band within which the MIPAS and AATSR observations will be coincident. It can be seen from the example orbits plotted in figure 8.2 that this latitude band is approximately 20 °S to 20 °N. This region of coincident measurements is well placed for comparisons between McClouds_RT and the AATSR retrieval as high tropical cirrus has been the focus of this thesis and it is well suited to be retrieved from MIPAS.
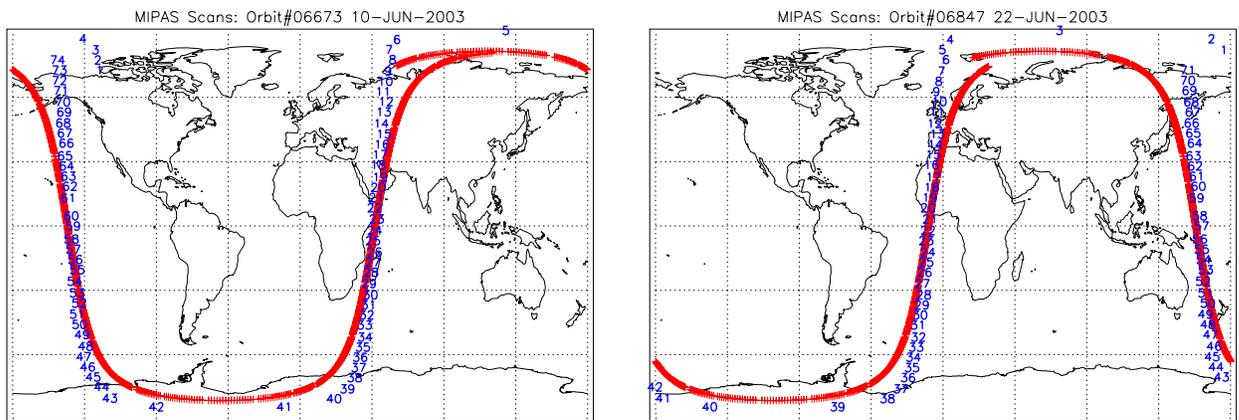
Figure 8.2: Maps of orbit ground paths (in red) and MIPAS elevation profile locations (in blue and numbered by profile), for orbits 6673 (left) and 6847 (right).

The two orbits shown in figure 8.2 are selected for the two case studies, which are:

1. Profile 23 (~ 0.5 °N, 48 °E) of MIPAS orbit 6673 from the 10th of June 2003 – *anvil outflow cirrus off the East African coast*

2. Profile 21 (~ 0.5 °N, -8 °E) of MIPAS orbit 6847 from the 22nd of June 2003 – *tenuous cirrus off the West African coast.*

Figure 8.2 shows that the satellite position (i.e. the nadir observation) and the MIPAS tangent points are coincident at the chosen elevation profiles. For each orbit the threshold radiances are plotted by profile against the average (for the orbit) tangent height in figure

8.3. Orbit 6673 has a region of particularly high cirrus in profiles 18 to 24 over the tropical region of East Africa, and an additional region of enhanced radiation in profiles 55 to 57, due to cloud presence as the satellite over-passes the equatorial region again. Orbit 6847 also displays the two regions of high cirrus presence in the tropical regions, in profiles 18 to 22 and 51 to 55, corresponding to the descending and ascending parts of the orbit. In addition, the presence of PSCs can also be seen in orbit 6847, with enhanced radiances in profiles 34 to 39 over the southern polar region.
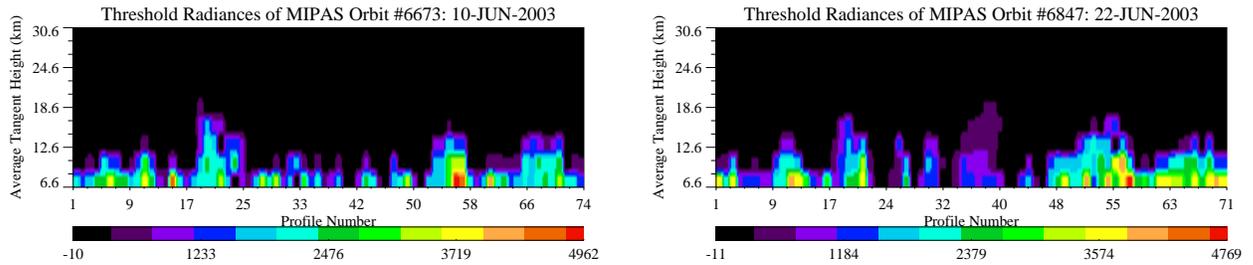
Figure 8.3: Threshold radiance values (in units of $nW/(cm^2 \; sr \; cm^{-1})$) of the limb scans below $\sim$30 km, plotted by elevation profile for orbits 6673 (left) and 6847 (right).

8.4.2 Anvil Outflow Cirrus off the East African Coast

Cloud is first detected[2] in the 16.28 km limb scan of profile 23, for which the limb spectra is given in figure 8.4(a). The continuum level of the spectra is very low, but still too high to be clear-sky at that altitude. Initial impressions suggest that the FOV is only partially filled, with the cloud top below the tangent height. Figure 8.4(b) then plots the limb spectra recorded by the 13.42 km limb scan from which the measurement vector for the retrieval is derived. The spectra in figure 8.4(b) is clearly dominated by cirrus presence, but looks very flat and it is only by closer inspection of the zoomed region of interest that the absorption side-lobes and H_2O lines can be seen. However, they are not particularly pronounced.

The majority of emission in a limb scan originates from the first few kilometres above the tangent height. Considering the spherical geometry, for tangent heights between 9 and 18 km the sample volume can be considered to be approximately 300 km in the flight path direction by 30 km in the orthogonal horizontal, centred on the tangent point. This estimated sample region can be used in the comparison to the nadir AATSR images to determine the approximate region of the image (and thus the region of the cloud field) that the limb observation is sensitive to. For the 13.42 km tangent height of profile 23, the MIPAS sample region is plotted (in black) over the coincident AATSR measurements in figure 8.5. Figure

---

[2]The principal components cloud test was used for cloud detection throughout this chapter.

8.5(a) shows the 0.55 μm visible reflectance channel and figure 8.5(b) shows the difference between the 11 μm and 12 μm thermal brightness temperature channels.

The visible reflectance channel can be thought of as an indicator for optically thick clouds or low level water clouds, which generally have a high albedo. Optically thick clouds are more evident over the ocean as it has a low albedo, whereas land (particularly desert and snow) has a much higher albedo which makes it more difficult to detect overlying cloud. In figure 8.5(a) most of the ocean has a low albedo, but to the west of the MIPAS sample region near the bottom of the image, there is a region of particularly high albedo (i.e. >35 %) associated with the core of a tropical convective tower. In the region surrounding the convective core there is an area of reflectance enhanced over the ocean, ranging between 15-20 %, for which the pattern is consistent with convective anvil cirrus. The bottom lower half of the MIPAS sample region is inhomogeneously filled by the cirrus anvil outflow. There is also a streak of optically thick cloud running (North to South) through the centre of the image which is probably cloud associated with a weather front.

The difference between the 11 and 12 μm brightness temperatures is known to be an indicator of the presence of medium- and high-level clouds, and especially thin cirrus. The difference may be large over optically thin cirrus cloud owing to emissivity differences at the two wavelengths (Saunders and Kriebel, 1988). The exact threshold for the thin cirrus test is a function of both the brightness temperature at 11 μm and the air mass in the line of sight, however, differences greater than ~ 2.5 K can be expected to indicate the presence of thin cirrus. Low and optically thick clouds will give very little difference between the two channels.

The 11 and 12 μm difference image in figure 8.5(b), shows that thin cirrus is highly probably over the majority of the ocean covered by the image. The frontal cloud and convective core show almost no difference between the two channels. The entire MIPAS sample region has a difference of $\gtrsim 2.5$ K and as such can be expected to be filled by cirrus, with thicker cirrus in the southern part, a transitional patchy area, then thin cirrus in the northern region.

McClouds_RT was applied to the 13.42 km limb measurements using the first 100 optimised channels to construct the measurement vector. The retrieval results are given in table 8.1. The degree of non-linearity was calculated and was found to be near linear about the solution within the error bounds. The averaging kernel is given in equation 8.5, and is almost ideal, with near zero off diagonal elements indicating only small correlations between cloud parameters, and the diagonal elements are all near unity, indicating a high retrieval sensitivity to changes in the true cloud state. The degrees of freedom for signal was 3.921, thus the cloud parameters are retrieved near independently and the retrieval uncertainty in table 8.1 has low values to correspond. The cost function was 36.62, and the residual in figure 8.6 shows that the model fits the measurements to considerably better
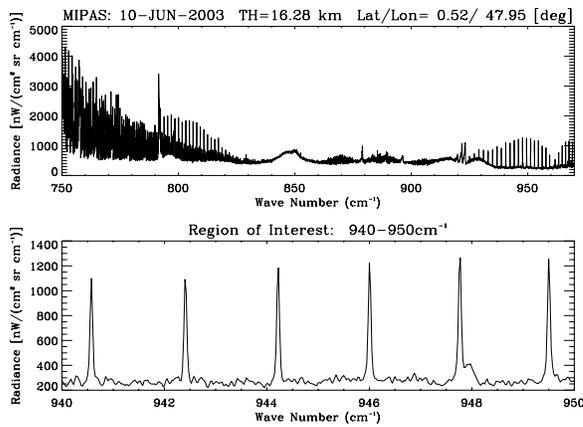
(a) Band A spectra for the 16.28 km limb scan.
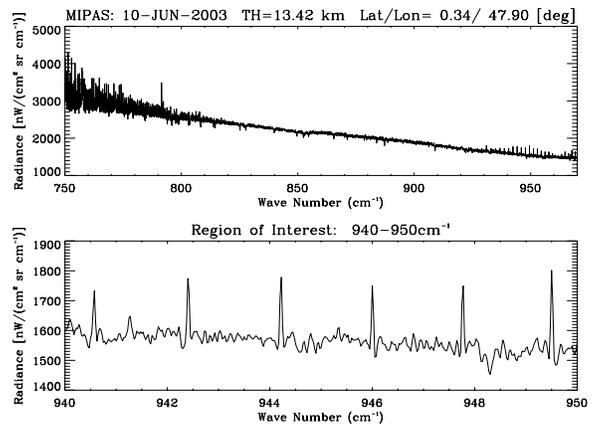
(b) Band A spectra for the 13.42 km limb scan.

Figure 8.4: Apodised MIPAS spectra from profile 23 of orbit 6673. (a) plots the spectra from the first limb scan cloud is detected in and (b) plots the spectra from the subsequent lower limb scan from which the measurement vector for McClouds_RT is derived.
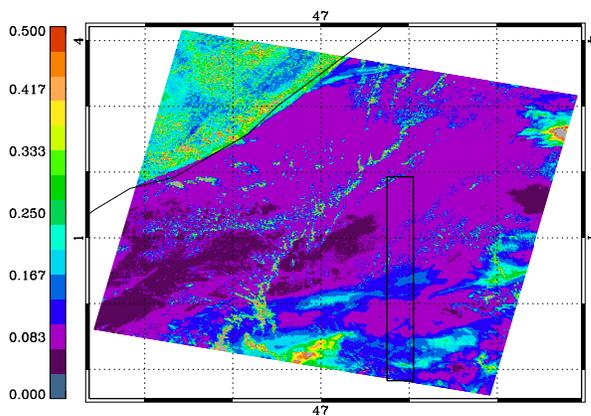
(a) The 0.55 μm visible reflectance channel.
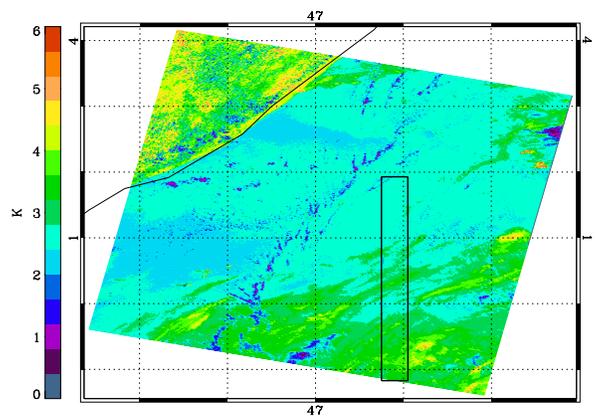
(b) The difference between the 11 μm and 12 μm thermal brightness temperature channels.

Figure 8.5: AATSR channel information for case study 1. The MIPAS sample region is overplotted in black.

Cloud Property	$\mathbf{x}_0$	$\hat{\mathbf{x}}$	$\hat{\sigma}$	[%] Uncertainty
r_e	23.792	20.261	0.366	1.8
N	0.101	0.504	0.184	36.5
C_{top}	15.340	14.710	0.291	2.0
C_{depth}	0.849	0.823	0.232	28.2

Table 8.1: Retrieval results from appliying the channel optimisation method (where 100 channels were selected and at the first guess it's expected that $d_s = 3.997$) and assuming ECMWF atmospheric temperature and pressure data. Retrieval cost: $\chi^2 = 36.62$, $\chi^2/df = 0.37$, $\chi^2/d_s = 9.34$.

Derived Cloud Property	$f(\hat{\mathbf{x}})$	$\hat{\sigma}_f$	[%] Uncertainty
τ_{ir}	0.228	0.050	21.9
τ_{vis}	0.340	0.074	21.9

Table 8.2: Thermal and visible optical thickness derived from the retrieved parameters in table 8.1 for the East Africa case study.

than the measurement error. The derived thermal optical depth (from equation 5.6) and the corresponding visible optical depth are given in table 8.2. The variance of optical thickness is calculated from:

$$\sigma_\tau^2 = \sigma_N^2 \left(\frac{\partial \tau}{\partial N}\right)^2 + \sigma_{c_d}^2 \left(\frac{\partial \tau}{\partial c_d}\right)^2 + 2\text{cov}(N, c_d) \left(\frac{\partial \tau}{\partial N}\right) \left(\frac{\partial \tau}{\partial c_d}\right) \tag{8.4}$$

where $c_d \equiv C_{\text{depth}}$ and $\text{cov}(N, c_d)$ is the covariance of N and C_{depth}. It can be seen that the percentage uncertainty in optical depth is smaller than the uncertainty in both N and C_{depth}, which is due to the negative covariance between N and C_{depth}.

$$\mathbf{A}_1 = \begin{bmatrix} 0.9998 & -0.0151 & -1.118 \times 10^{-3} & 0.0638 \\ -9.719 \times 10^{-5} & 0.9915 & 5.975 \times 10^{-4} & 0.0422 \\ -3.625 \times 10^{-5} & 3.025 \times 10^{-3} & 0.9958 & -0.0356 \\ 8.275 \times 10^{-5} & 8.549 \times 10^{-3} & -1.424 \times 10^{-3} & 0.9337 \end{bmatrix} \tag{8.5}$$

The McClouds_RT results were then compared to the AATSR cloud retrievals presented in figure 8.7. Only the cloudy pixels in which the AATSR retrieval has determined the phase as ice should be considered for comparison to the McClouds_RT results. These pixels are also filtered so that the AATSR retrieval cost is less than 10. From figure 8.7, it can be seen that the MIPAS sample region is determined to be only partially filled with ice cloud, in two distinct bands in the southern half of the sample region. The thin cirrus expected to be present from the difference between the 11 and 12 μm channels (figure 8.5(b)) is thought to be water cloud by the AATSR retrieval, which is shown in figure 8.7(d) (and
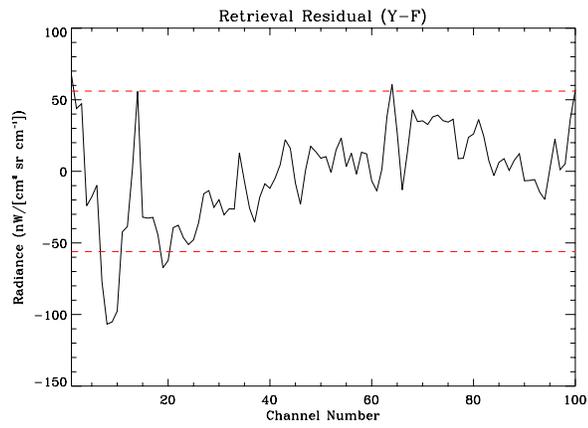
Figure 8.6: Residual (i.e. $\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}})$) from the optimised channel retrieval for the first validation case study.
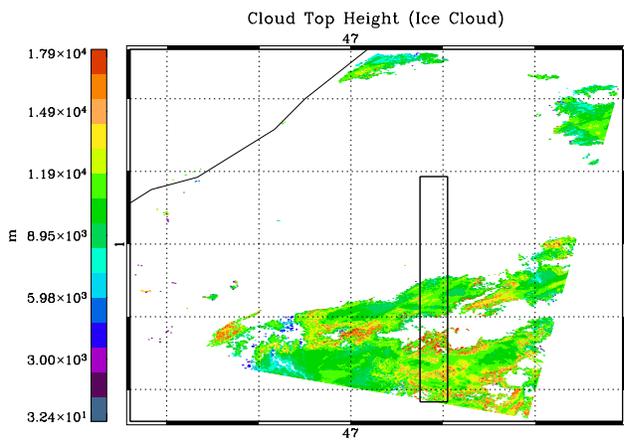

(a) Cloud top height for ice clouds.
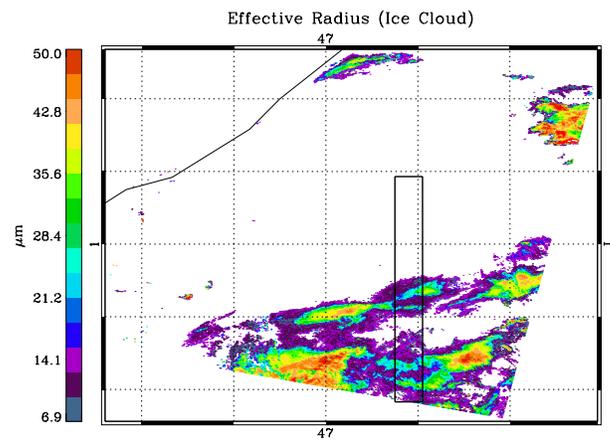
(b) Effective radius for ice clouds.
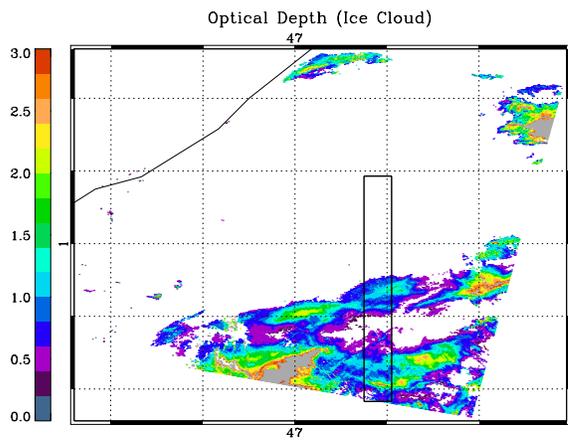
(c) Optical depth for ice clouds.
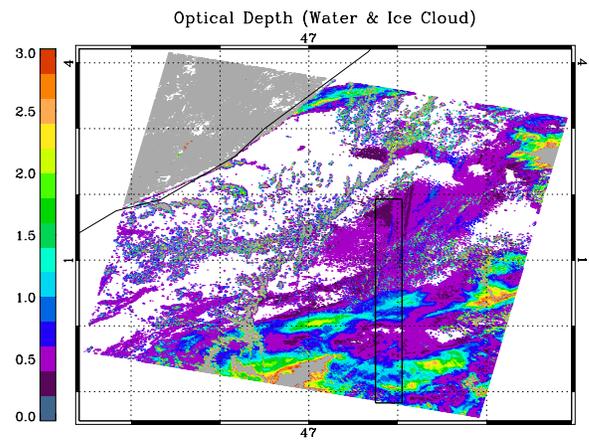
(d) Optical depth for water and ice clouds.

Figure 8.7: Results from the AATSR cloud retrievals for case study 1.

Cloud Property	mean value	standard deviation	[%] Uncertainty
cloud top height (km)	11.805	1.839	15.6
effective radius μm	16.635	6.781	40.8
τ_{vis}	0.908	0.337	41.5

Table 8.3: The AATSR cloud retrieval results for the East Africa case study; statistical means and corresponding standard deviations are given for ice cloud values within the MIPAS sample region.

accordingly has lower cloud top heights). The mean statistical values of cloud top height, effective radius and optical depth for ice cloud within the sample region are given in table 8.3.

The effective radius from the two retrievals is within reasonable agreement, differing by $\lesssim 4\mu$m where the statistical standard deviation on the AATSR retrieved effective radius is ~ 6.78 μm. The cloud top heights differ by ~ 3 km, but nadir imagers are known to underestimate cloud top height for thick convective cirrus (e.g. Sherwood et al., 2004) and thus the difference is not considered a significant disagreement. Also, there is a high variability of cloud top height within the MIPAS sample region, with the cloud edge areas having cloud top heights up to 15 km. Actual anvil cirrus is unlikely to have such high variability in cloud top height as it is generally laminar, even if it is inhomogeneous with respect to ice particle number density and size distribution. The regions of high cloud top coincide with low effective radius and low optical depth. The relationship between the ice cloud optical depth and cloud top height from the AATSR retrieval is plotted in figure 8.8(a). From the retrieval cost[3] plotted against optical depth in figure 8.8(b), it is clear that the forward model in the AATSR retrieval has difficulty fitting the measurements for optical depths less than ~ 0.5, which causes the high cloud top heights. Also, due to the approximation of a spherical layer cloud geometry in McClouds_FM it is unclear how the retrieved cloud top height will behave if the true cloud has a spatially variable cloud top height. It may be sensitive to a mean cloud top level, or possibly to the highest point. It is also quite possible that the true cloud field in this case is actually composed of two cloud layers in the sample region: a high unbroken thin cirrus layer (at ~ 15 km) with an underlying broken layer of thicker anvil blow-off (at ~ 11 km). A tenuous cirrus layer is known to frequently form above tropical cirrus anvils (Garrett et al., 2004; Garrett et al., 2005). The MIPAS measurement would be sensitive to the higher tenuous cirrus layer and the AATSR view would be more sensitive to the thicker inhomogeneous lower layer. The fact that the AATSR cloud retrieval has difficulty modelling the expected thin

---

[3]The AATSR retrieval cost is not normalised for the degrees of freedom (i.e. the number of channels) in this thesis. It can be seen from figure 8.8(b) that if the cost is divided by the number of channels it can return a value less than one. This is because the degrees of freedom for signal, i.e. the number of pieces of independent information that can be retrieved, is less than the number of channels.

cirrus region in the northern half of the sample region and determines it to be a water cloud suggests that lack of sensitivity to the tenuous upper layer may be the problem. The possibility of the multi-layer scenario would also explain the large differences in optical depth, with the McClouds_RT value consistent with thin cirrus and the AATSR value more typical of thicker anvil cirrus. However, if the cloud field is actually a single layer with cloud properties as given by the AATSR results, then the difference in optical depth could be due to the inhomogeneity of cloud in the sample region and the fact that it is only partially filled. It is unclear how the McClouds_RT retrieved optical thickness will behave corresponding to a partially filled sample volume when it is assumed homogeneous in McClouds_FM. If the relationship is linear then the fact that roughly one third of the sample region is filled may account for the (approximate) factor of three difference between the two optical depth values.

(a) Retrieved cloud top height versus optical depth.

(b) Retrieval cost function versus optical depth.
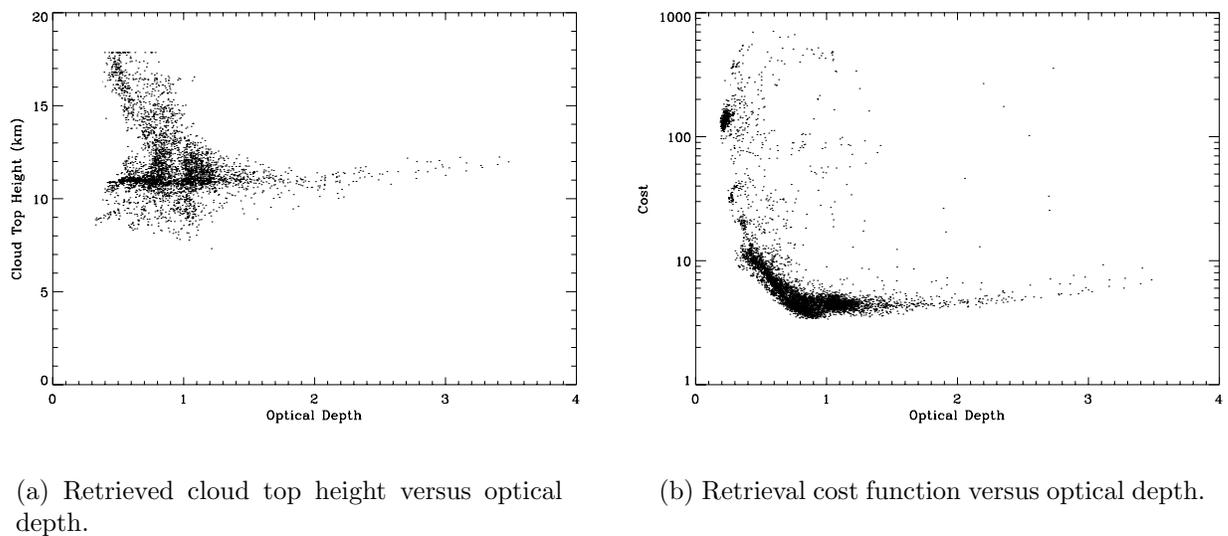
Figure 8.8: AATSR retrieval dependence on ice cloud optical depth. (a) plots cloud top height against optical depth and (b) plots the cost as a function of optical depth.

8.4.3 Tenuous Cirrus off the West African Coast

For the second case study, the measurements were selected on the basis of choosing a comparison where the optical depth of the cloud is at the limit of the AATSR detection sensitivity. The AATSR scene for the case study is shown in figure 8.9. From the visible channel (figure 8.9(a)) the cirrus partially filling the MIPAS sample region appears inhomogeneous and has a low reflectance, i.e. $\lesssim 25\%$, which suggests that the cirrus is optically thin. The difference between the 11 μm and 12 μm brightness temperatures (figure 8.9(b)) also shows that the cirrus in the Northern half of the sample region has a difference of ~ 2.5

K and is probably optically thin. There is also evidence of small patches of thin cirrus in the Southern half of the sample region.
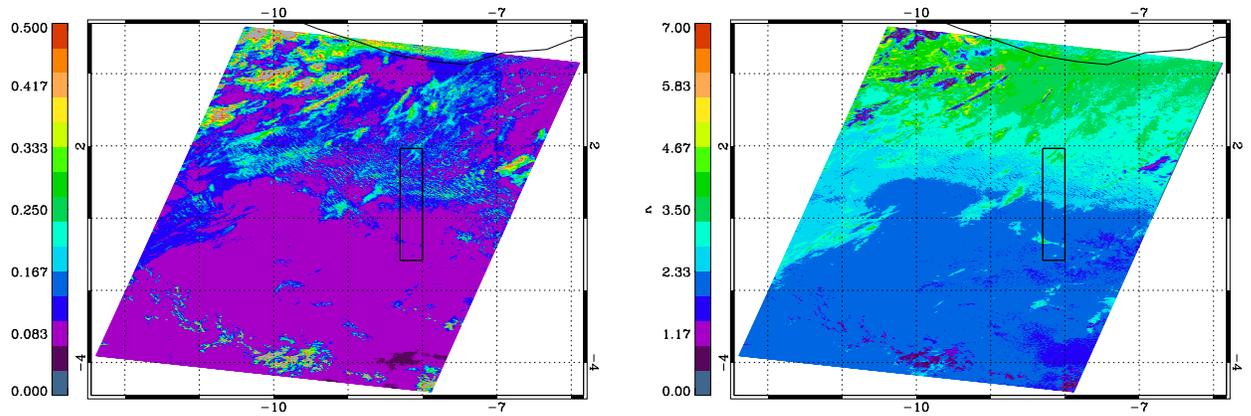
(a) The 0.55 μm reflectivity channel.

(b) The difference between the 11 μm and 12 μm brightness temperature channels.

Figure 8.9: AATSR channel information for case study 2. The MIPAS sample region is overplotted in black.

In the coincident MIPAS profile measurement, the first limb scan to detect cloud is at 13.30 km, for which the recorded limb spectra is plotted in figure 8.10(a). The subsequent limb scan in the profile, at 10.48 km, is used to construct the measurement vector for McClouds_RT. The spectra from the 10.48 km limb scan is given in figure 8.10(b).

The spectra from the 13.30 km limb scan has a very high continuum level, and the region of interest shows clear CO_2 emission lines (although they are not strongly peaked relative to the continuum) and there is strong evidence of H_2O absorption features. The spectra suggests strong scattering into the line-of-sight from below the cloud and from experience the cloud top is probably at or around the tangent height. The spectra from the lower limb scan at 10.48 km (figure 8.10(b)) also shows a greatly enhanced continuum level. However, the spectra looks like a Planck function with some atmospheric absorption. However, for the absorption to dominate over the limb emission to such an extent, and to obtain such a high continuum level, there must be a high percentage of photons arriving at the detector which have been scattered into the line-of-sight from below the cloud.

The results from McClouds_RT using the first 100 optimised channels to derive the measurement vector are given in table 8.4, with the corresponding averaging kernel in equation 8.6. The derived optical thickness in the infrared and visible is given in table 8.5. From the averaging kernel in equation 8.6, $d_s = 3.545$ which is much lower than the optimised channel value and is clearly due to correlations between effective radius and cloud depth, and cloud top height and cloud depth. The lack of retrieval sensitivity to cloud depth is the main problem, and this is reflected in the high error. The retrieved value of cloud depth

Cloud Property	$\mathbf{x}_0$	$\hat{\mathbf{x}}$	$\hat{\sigma}$	[%] Uncertainty
r_e	13.574	9.213	2.314	25.1
N	0.344	0.301	0.076	25.2
C_{top}	13.489	12.081	0.512	4.2
C_{depth}	1.128	1.580	0.592	37.5

Table 8.4: Retrieval results from appliying the channel optimisation method (where 100 channels were selected and at the first guess it's expected that $d_s =3.994$) and assuming ECMWF atmospheric temperature and pressure data. Retrieval cost $\chi^2 =316.4634$, $\chi^2/df =0.32$, $\chi^2/d_s =80.75$.

Derived Cloud Property	$f(\hat{\mathbf{x}})$	$\hat{\sigma}_f$	[%] Uncertainty
τ_{ir}	0.098	0.022	22.4
τ_{vis}	0.208	0.047	22.4

Table 8.5: Thermal and visible optical thickness derived from the retrieved parameters in table 8.4 for the West Africa case study.

itself, is much higher than would be expected for such a tenuous cirrus case, and although the visible optical depth is low, the high physical cloud depth is biasing τ_{vis} high. An optical thickness of ~ 0.2 will be at the limit of the AATSR detection sensitivity (expected to be between 0.1 and 0.2), and the tenuous nature of the cirrus can be clearly seen in the AATSR images in figure 8.9. The number density and effective radius appear to be sensible, as low values will be expected, but the cloud top height is quite low and such a low cloud top is not consistent with the 13.30 km limb measurement. It is possible that McClouds_RT has had to place the cloud top height slightly lower than expected as McClouds_FM is having trouble fitting the measurements. The residual in figure 8.11 shows the measurement has a consistent high bias and there is a very poor fit over the first 10 channels, reflected in the high cost function (i.e. $\chi^2=316.46$). In order to increase radiance, the retrieval may place the cloud top lower as the temperature is higher lower in the atmosphere and this will increase the radiance emitted from the cloud.

$$\mathbf{A}_2 = \begin{bmatrix} 0.9914 & 0.0404 & -0.0301 & -1.0426 \\ 2.585 \times 10^{-4} & 0.9986 & 0.0014 & 0.0447 \\ -9.737 \times 10^{-4} & 0.0071 & 0.9871 & -0.3710 \\ -1.351 \times 10^{-3} & 0.0091 & -0.0148 & 0.5676 \end{bmatrix} \tag{8.6}$$

The results from the AATSR retrieval are presented in figure 8.12. Immediately obvious from the results is that cloud in the sample region is highly inhomogeneous. Figure 8.12(d) shows the high variability of optical depth for all cloud retrieved, varying from ~ 0.2 in the South of the sample region to over ~ 1.5 in the north of the sample region. The optical

depth of the ice cloud with cost less than 10 is given in figure 8.12(c), clearly the retrieval is only fitting the measurements well in a small fraction of the cloudy pixels in the sample region. In these low cost retrievals there are two regions at the extremities of the sample region, with a low optical depth in the south region, consistent with the McClouds_RT results, and a high optical depth in the northern region. The low optical depth cloudy pixels are correlated with unrealisitically low cloud top heights for ice cloud (i.e. less than 3 km) in figure 8.12(a) and the high optical depth pixels have a cloud top height between 9 and 10 km. The remainder of the cloudy pixels in the sample region, with cost greater than 10, also have a cloud top height between 9 and 10 km. This cloud top height is in reasonable agreement with the McClouds_RT value, considering the propensity of the nadir retrievals to slightly underestimate cloud top height. The difference in cloud top height between the two retrievals is consistent with the same difference in the previous case study. For the effective radius in figure 8.12(b), the majority of the sample region, throughout the centre, has very small values of less than 5 μm. From figure 8.13, this region can be seen to be correlated with very high costs and the retrieved values are probably unreliable. The low cost region at the northern edge of the sample region has a range of effective radius between approximately 5 and 20 μm, which is more consistent with the McClouds_RT value. However, the low cost region at the southern edge of the sample region has effective radii reaching up to 50 μm, but due to the low cloud top heights which are not consistent with typical ice clouds, these values are also considered unreliable.

Discussion

The second validation case study does not have as good an agreement between MIPAS and AATSR as the first study. There are many possible reasons for this. One reason may be that the cloud in the sample region is inhomogeneous and possibly broken, and this means that the assumption of cloud horizontal homogoneity and uniformity in McClouds_FM will not hold and thus it will be difficult for the model to fit the measurements. The effects of inhomogoneity are not well understood, if inhomogeneous clouds enhance the multiple scattering into the line-of-sight it might explain why McClouds_FM was inclined to return lower radiance values.

The McClouds_RT residual in figure 8.11 shows that the model can fit the measurements well in the majority of channels but has major difficulty in the first 10 channels. This may be a problem due to the spectral variance of the single scattering properties with the crystal habit and size distributions assumed, as mentioned in section 7.4.4. For tenuous cirrus in this equatorial region, a better approximation of microphysics may be obtained from aggregate crystals with a small effective radius, between 1 and 5 μm. However, in order to test aggregate crystals in the retrieval, it cannot be assumed that the hexagonal
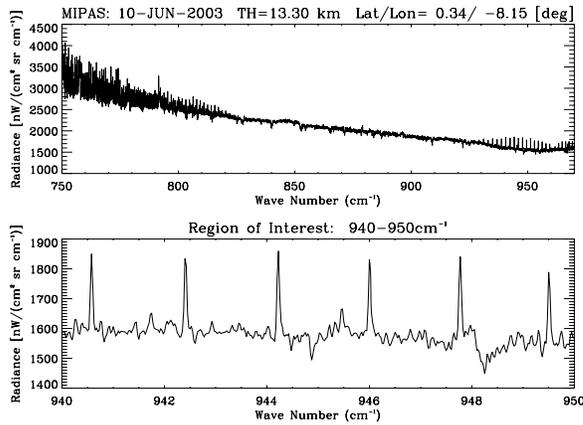
(a) Band A spectra for the 13.30 km limb scan.
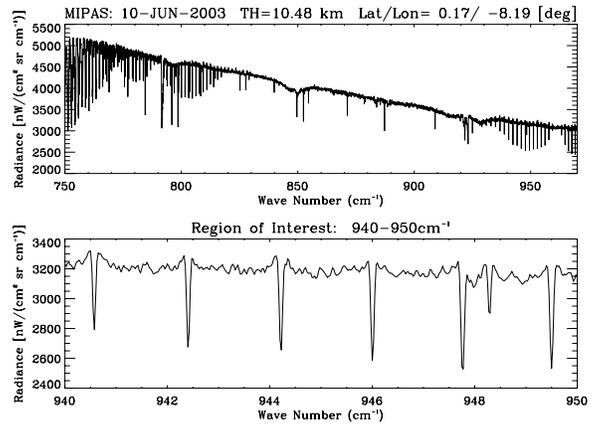
(b) Band A spectra for the 10.48 km limb scan.

Figure 8.10: Apodised MIPAS spectra from profile 21 of orbit 6847. (a) plots the spectra from the first limb scan cloud is detected in and (b) plots the spectra from the subsequent lower limb scan from which the measurement vector for McClouds_RT is derived.
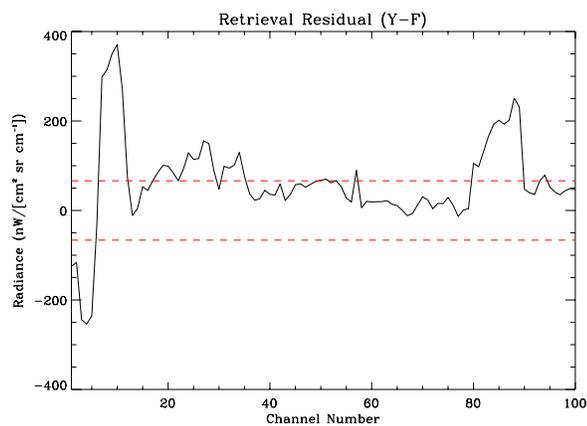
Figure 8.11: Residual (i.e. $\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}})$) from the optimised channel retrieval for the second validation case study.
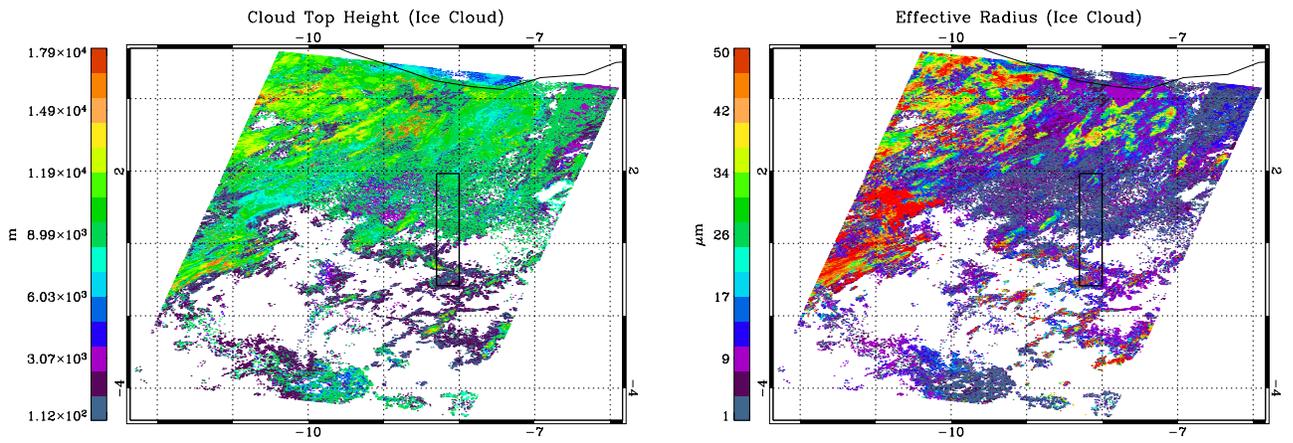
(a) Cloud top height for ice clouds.

(b) Effective radius for ice clouds.
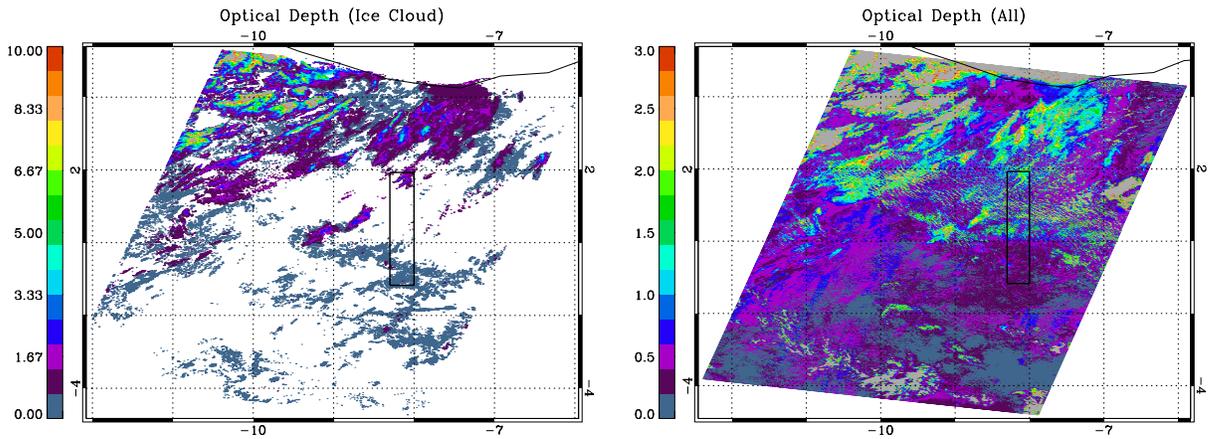
(c) Optical depth for ice clouds. Filtered for cost less than 10.

(d) Optical depth for water and ice clouds.

Figure 8.12: Results from the AATSR cloud retrievals for case study 2.
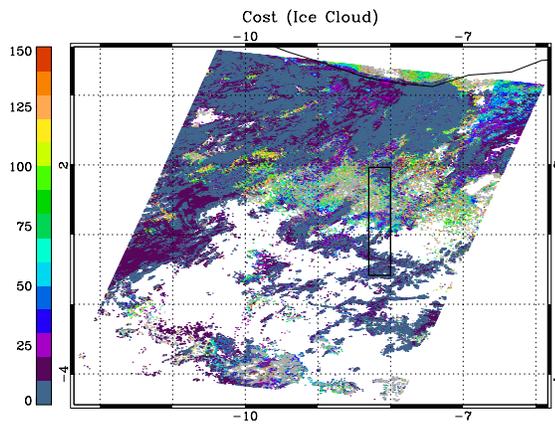
Figure 8.13: Retrieval cost values for ice cloud pixels in the AATSR scene for case study 2.

crystal first guess method and channel optimisation will hold. The modelling study required for the first guess method and the channel optimisation should be repeated in future work for aggregate crystals.

There are many problems with the AATSR retrieval for this scene, and the costs (see figure 8.13) are too substantial over the sample region for any concrete conclusions to be drawn about the comparison to the MIPAS retrieval. Furthermore, the fact that the McClouds_RT solution for the 10.48 km limb scan, is not consistent with the upper limb scan, suggests that more than one limb scan is necessary in the measurement vector. However, the channel optimisation would need to be repeated for the new set of possible measurements and the speed of McClouds_FM would need to be increased for the forward modelling required to be feasible.

8.5   Conclusion

As the first known attempt to compare cloud properties derived from coincident limb and nadir observations, this validation exercise has proved a reasonable success. In particular, there was good agreement in the first case study and the differences in the second case study could be explained but further work is required to understand them fully.

The two case studies show that for optically thick cirrus, i.e. $\tau \gtrsim 0.5$, horizontal inohomogeneities are likely to occur and as such the AATSR cloud retrievals will have an advantage over MIPAS due to the superior horizontal resolution. However, for optically thin cirrus, as evidenced in the second case study, AATSR is either unlikely to detect the cirrus or it will only record a weak signal, whereas MIPAS will record a strong cloud signal. The AATSR retrieval will have difficulty fitting the measurements and will be unreliable for thin cirrus. Thus, to obtain good global cirrus information, the MIPAS retrieval is necessary to obtain information on thin cirrus. In addition, the vertical resolution of MIPAS should provide superior cloud top height accuracy to AATSR.

The two case studies detailed in this chapter can not be considered as a validation of McClouds_RT. However, the results have proved encouraging and further statistics from a much larger sample of coincident measurements are required to fully validate McClouds_RT. Before carrying out further comparisons, it is important to resolve the uncertainties in the McClouds_FM behaviour to broken and inhomogeneous cloud fields. A full model sensitivity study is required to investigate the effect of variable cloud top height and both horizontal and vertical inhomogeneity on the McClouds_FM output. This should improve the interpretation of comparisons of the retrievals for inhomogeneous cloud fields. The retrieval could also be run on these simulated inhomogeneous fields, while still assuming horizontal homogeneity in order to understand the error occurred by this assumption when

it is not valid. Furthermore, due to the large overhead in computing time required to carry out processing on a large dataset, it would be beneficial to first complete a number of the future work areas outlined in section 9.2 to improve both the retrieval method and the understanding of the results.

Chapter 9

Conclusions and Future Work

9.1   Synopsis and Conclusions

This thesis aimed to address the hypothesis that infrared limb sounding of cloud properties can improve current cirrus climatology data. In particular, the MIPAS instrument was the focus of this work, but the methodology developed and the conclusions drawn are equally as valid for similar high resolution infrared limb sounders, e.g. the Tropospheric Emission Spectrometer (TES) which flies on the NASA Aura Satellite. In addressing the hypothesis, the main aim of this thesis was to a develop a cloud retrieval strategy for MIPAS. The research achieved the following:

- It was established in chapter 1 that current cloud climatologies (e.g. ISCCP) underestimate cirrus and in particular, optically thin high-level cirrus. This is due to the use of mostly nadir observations to construct cloud climatology, which in general have a lower detection sensitivity to optically thin cloud than possible with limb observations.

- In chapter 3 the high cloud detection sensitivity of MIPAS was demonstrated and shown to be comparable to state-of-the-art occultation limb sounders. In addition, a powerful new cloud detection method (known as the *principle components cloud test*) was developed in order to be particularly adept at dealing with high thin tropical cirrus which current MIPAS detection methods have a tendency to miss-detect.

- The superior spatial and temporal coverage of MIPAS (typical of infrared limb emission sounders) compared to limb occulation sounders was clearly stated (see figures 3.10 to 3.12 in section 3.4).

- Chapter 4 detailed the development and implementation of a new forward model (i.e. McClouds_FM) to simulate high resolution infrared limb measurements of cirrus.

Spherical geometry and multiple-scattering within clouds is included in the radiative transfer model. Such a forward model is the first requirement in order to determine cloud properties from limb measurements and as such was the major scientific advance achieved by the research in this thesis.

- McClouds_FM was shown to be capable of matching actual MIPAS measurements (within measurement noise) using realistic cloud parameters. The importance of including multiple scattering, as opposed to a single scattering approximation, was demonstrated, showing that single scattering alone could underestimate the radiance arriving at the satellite by up to 25 % (for typical cirrus parameters).

- The dependence of the McClouds_FM output on cirrus optical depth, effective radius and cloud top height was investigated (chapter 5). The focus was the effect of varying the cirrus properties on spectral scattering features and cloud detection measures. It was found that the relationships were non-linear and that there was a peak in the multiple-scattering effect for cirrus with optical thickness between approximately 0.05 and 0.3.

- The simple relationships investigated in chapter 5 were used to develop a method to make a first estimate of the solution to the inverse problem (of inferring cirrus properties from MIPAS measurements). This first estimate is important, due to the non-linearity of McClouds_FM, as it will "point" the retrieval to the neighbourhood of state space that the solution is likely to be found. A method to determine a first guess for an iterative retrieval is a necessity if the retrieval is to be automated and consistent. The first guess was proved to be effective in the retrieval test cases in chapter 7.

- Chapter 7 detailed the development and implementation of a new optimal estimation methodology to determine cirrus properties from MIPAS measurements. As the first known attempt to employ such an approach, the retrieval was highly succesful. It was shown for test cases that considering the non-linearity was necessary to find a solution but it was unlikely to be a problem with respect to the retrieval error analysis.

- The information content and optimisation of the MIPAS measurements was studied. It was found that by selecting spectral measurements (for the retrieval measurement vector) based on maximising the information content, it is theoretically possible to determine the effective radius, number density, cloud top height, and vertical cloud depth to high accuracy, achieving a high retrieval sensitivity with degrees of freedom for signal approaching four, i.e. the optimised retrieval should be able to achieve a degrees of freedom for signal > 3.8 and retrieve (for typical thin cirrus values):

- Effective radius to within 2 μm

- Number density to better than 20 %

- Cloud top height to better than 0.2 km

- Cloud depth to better than 25 %.

- A preliminary validation of the retrieval was carried out in chapter 8 for two case studies comparing coincident MIPAS and AATSR measurements. The results proved encouraging, and without rigorously accounting for the differences in viewing geometry and sensitivity, the first case study showed that cloud retrievals from the two instruments were in reasonable agreement. The second case study demonstrated the lack of sensitivity of nadir radiometers to tenuous cirrus compared to the MIPAS instrument, however, the retrievals for both instruments were problematic.

In conclusion, this thesis has shown that the MIPAS instrument is capable of measuring valuable information on the microphysical and macrophysical properties of cirrus, and in combination with the superior coverage and cloud detection sensitivity over the state-of-the-art satellite observations used in current climatology studies, cirrus climatology data would be greatly enhanced by including cirrus data derived from MIPAS. In order for this improved cirrus climatology to become a reality, there is a great deal of future work to be done, and in addition there is huge potential for improving the cirrus measurements and obtaining a greater understanding of the physics involved. The following section outlines the major areas of future work which should be prioritised.

9.2   Future Work

There is a wide scope for future work to continue the progress made by this thesis in what is an exciting and novel new approach to deriving cloud information from satellite instruments. The aim of the different areas of work falls into three main categories: improving the current methodology, supplementary research, and applying the methodology to an extended data set.

9.2.1   Improvements

There are a number of improvements to the methodolodgy that would enhance the quality of the cirrus information. A short discussion of the major issues follows.

Cloud detection

The principal components cloud test has shown that it is potentially a very powerful tool for cloud detection. The method should be investigated further by repeating the PCA using a wider spectral range, i.e. including the AB band measurements, and also using a seasonal data analysis as opposed to monthly, and experimenting with decreasing the latitude binning resolution, e.g. using hemispheric datasets. It is important to try and generalise the method as much as possible and to extend the region of applicability outside the tropics. Furthermore, the relationship between the contribution coefficients for the first three principle components may also prove to be useful in subtle differentiation between cloud types.

McClouds_FM Improvements

The biggest restriction on the progress of this work is the bottle-neck due to the speed of the forward model calculations. In order for the retrieval to be computationally feasible, it will be necessary to increase the speed. This may be possible by creating 'Look-Up-Tables' of the scattering calculations as they are independent of the atmospheric properties (unlike the radiative transfer), however, a sensitivity study will be required in order to determine the resolution at which the calculations would be required. Alternatively, it may be possible to further optimise the memory use in the code, as simulating on the order of 10^5 photons is heavy both on CPU and memory cost. By converting the McClouds_FM to a more efficient programming language such as FORTRAN or C further gains in speed may also be obtained.

Another avenue is to assess the accuracy to which the radiative transfer calculations are required. Firstly, with respect to spectral resolution, and secondly, with respect to observer position in the atmosphere. Within the retrieval it is not necessary to carry out the radiative transfer calculations for each change in cloud top height and cloud depth, as it is possible to output the radiance at a point along a ray path. By setting a grid of heights to calculate the radiation field at, it is possible to interpolate between these points, and the radiation field can be calculated at the start of the retrieval. In addition, the radiative transfer does not change much with elevation angle for angles greater than 20° and thus the calculations can also be interpolated with respect to elevation angle. The absolute limit of the height and elevation angle grid resolution to provide the required accuracy should be investigated. Furthermore, the RFM may not be the fastest line-by-line code available. It may be possible to pare down the code essential to McClouds_FM, or to find an alternative code to perform the clear-sky radiative transfer.

Other than the issue of speed, McClouds_FM can be improved by adapting it to allow

inhomogeneity within the cloud layer. This can be done by segregating the cloud layer into thinner layers with horizontal boundaries. Each cloud grid box will have its own value of number density and effective radius. Of course this would come at a computational price. Although it is unlikely that it would be possible to retrieve an inhomogeneity parameter, due to the nonlinearity and the multiple possible permutations, it is necessary to be able to simulate inhomogeneous cloud fields in order to improve the methods of comparison for the validation against AATSR retrievals.

There is a minor problem in the current implementation concerning the issue of randomness. McClouds_FM currently uses the *ran2* pseudo-random number generator given in (Press et al., 1992). The generator has a period length of approximately 2.3×10^{18} and is a combination of two Lehmer generators and a Bays-Durham shuffle to break up any serial correlations. However, if a high number of photons are used in the simulation it is possible that the generator will reach period exhaustion. As a result the pseudo-random number generator should be replaced by the MT19937 Mersenne Twister (Matsumoto and Nishimura, 1998). This will give a 623-dimensional equidistribution with a period length of $2^{19937} - 1$.

Validation of McClouds_FM

At the time of beginning this research there were no freely available radiative transfer codes which could simulate multiple scattering and include the necessary spherical geometry for the limb viewing geometry. Recently, the Atmospheric Radiative Transfer System (ARTS, Courcoux and Emde, 2005) has become publicly available, and includes two modules which allow either the discrete ordinates (Emde et al., 2005) or Monte Carlo method (Davis et al., 2004) to simulate the scattering from cloud particles. Although designed primarily for the millimetre and sub-millimetre spectral range, the ARTS model should be applicable to the infrared spectral region and it should be compared against McClouds_FM by way of validation of the radiative transfer.

The Initial Estimate of Cloud Properties

The retrieval sensitivity is lowest for number density and cloud depth as these parameters are most often correlated. As the retrieval sensitivity is often low for these parameters, it is even more important to have a good first estimate as the retrieval will be biased by the first guess. As discussed in section 7.5, because the τ and r_e values are determined by maximising the probability with respect to the *pdf* of IWC, the estimate of N and C_{depth} depends depends on the prior *pdf* used. It may be possible to use different IWC *pdf*s depending on geographics location or in conjunction with a *pdf* of cirrus type. The IWC *pdf* used was representative of thin and subvisible cirrus, but it may be possible to investigate

an ensemble of retrievals by using a range of IWC *pdf*s to determine the first guess.

Also, a study to determine the dependence of the scattering and detection measures on N and C_{depth} directly, rather than τ, should be carried out. Particularly, because C_{depth} is estimated from a height dependent empirical function, which could introduce correlations between C_{top} and C_{depth}. This will require a great deal more computation and will add a degree of complexity to the problem, but it could potentially produce a more accurate estimate of both parameters.

Finally, if the retrieval method is to be used outside of the tropics, the modelling studies for deriving the *ad-hoc* method of making the first guess should be repeated for mid-latitude and polar atmospheres.

The Retrieval Method

There are a number of possible improvements for the retrieval method, primarily concerned with improving the physical understanding of the retrieval results. The main points are:

- The assumption of Gaussian statistics for the prior covariance of number density and cloud depth may be unrealisitic. Using the logarithm of the two parameters in the retrieval should be investigated, as a Gaussian distribution in logarithm space should be an improved assumption.

- The channel optimisation should be repeated including the AB band spectra. This is important as there may be channels in that spectral region which have a higher information content and this may allow an improved retrieval accuracy and degrees of freedom for signal from using less channels in the measurement vector. In addition, there may be channels which have improved information on low optical depths.

- The weighting function with respect to the atmospheric parameters, $\mathbf{K}_b$ should be investigated, by carrying out a sensitivity study. As $\mathbf{K}_b$ will also depend on $\mathbf{x}$, it is important to understand the variance with respect to the cloud state. For example, the weighting function with respect to temperature has a reasonably strong peak at the cloud top height for optically thick clouds. If the forward model speed is improved then the model parameter error can be propogated through the retrieval.

- The channel optimisation and information content study should also be repeated, using two limb scans in the measurement vector: the first limb scan cloud is detected in and the subsequent lower scan. By using two limb scan measurements it should avoid retrievals where the retrieved cloud parameters (in particular the cloud top height) is not consistent with the cloud top event limb scan. In order to implement this approach McClouds_FM will need to use a high number of pencil beams to

account for the FOV adequately in the cloud top event limb scan. For this to be used in a practical retrieval method, the speed improvements of McClouds_FM will first have to be achieved as each model calculation will be a factor of n times slower for an n beam FOV.

- The channel optimisation and information content study should be repeated for mid-latitude and polar atmospheres, so the retrieval method will be valid for cirrus outside the tropics.

9.2.2 Supplementary Research

There are several key areas which can build on the methodology developed in this thesis. These areas are outlined below and are intended as pre-cursor studies to improve the results of applying the retrieval methodology with respect to obtaining global cirrus data.

Crystal Habit

The crystal habit is an important parameter to investigate. As the single scattering properties are dependent on the habit, it is important that the habit assumed in the model is as close to the true crystal shape of the cloud in the FOV. The *ad-hoc* method for determining the first guess must be re-derived for additional crystal shapes. In addition, the channel optimisation and information content study must also be repeated. Aggregates, poly-crystals, hexagonal plates, bullet-rosettes and mixtures of the various shapes with size distributions determined by *in situ* field studies (e.g. Baum et al., 2005) and accurately calculated single scattering properties (e.g. Baran, 2003) should all be investigated.

By running McClouds_RT in different crystal shape 'modes' it may be possible to infer the crystal shape of the ice clouds as the shape which has the lowest cost, i.e. the shape for which the simulated measurements has the best fit to the actual measurements.

Crystal Size Distribution

As evidenced in this thesis, one of the main problems of trying to retrieve cirrus properties from MIPAS is that there can be multiple solutions, which in effect means that the cost function surface is "bumpy" and pocked with multiple minima.

One possible reason for this is evidenced in chapter 5 when modelling the scattering effect index (SEI-1) as a function of effective radius and optical thickness in section 5.2. The SEI-1 surface (see figure 5.9) behaved in a somewhat unphysical and unexpected way, in that there were numerous peaks and troughs for a given optical thickness. It was mentioned that this is due to the size distributions assumed in this thesis. The size distributions were taken from *in situ* studies (Fu, 1996) and as they are technically only a small sample of the

underlying distributions typical of cirrus, they will have a high variance associated. This is shown in the relationship between the total number density of the size distributions and the effective radius. The total number density does not vary smoothly with effective radius and the peaks and troughs are correlated to those in the modelled function of SEI-1 with respect to effective radius. The scattering effect index can be assumed to be analogous to the modelled measurements, and therefore the complex and non-linear behaviour of McClouds_FM with respect to the cloud properties is partly due to the assumption of the *in situ* derived size distributions. For future McClouds_RT studies the ice crystal scattering properties should be re-calculated with a new set of size distributions (e.g. a log-normal distribution) in which the total number density is either constant or varies smoothly with respect to the effective radius. By making this new assumption the scattering effect index should become a smoother function with respect to r_e and τ, and by extension, the new size distributions should smooth the cost function of the cirrus retrieval. The initial estimate method will need to be re-derived, but it should also be less important to the retrieval if the cost function surface is smoother as there will be less local minima.

Investigating the Effect of Horizontal Inhomogeneity

The results of the comparison between MIPAS and AATSR cirrus retrievals in chapter 8 imply that the assumption of horizontal homogeneity in the McClouds forward model and retrieval may not hold that often in reality. It is important to improve the estimate of the error incurred by this assumption of homogeneity. As suggested in section 8.5, a study should be carried out in which inhomogeneous cirrus fields of varying properties should be input to McClouds_FM in order to produce a set of spectra representative of broken and horizontally varying cirrus. McClouds_RT can then be run on these spectra, under the assumption that the cirrus is homogeneous. Studying the results should yield a better understanding of how the McClouds_RT results should be interpreted when the true cloud field is known to be inhomogeneous. If the relationship is well-behaved then it may be possible to quantify the error on the retrieval incurred by the assumption of homogeneity.

Adaptation to other Instruments

The limb cloud retrieval methodology could be adapted for use with other infrared limb instruments to increase the volume of data on thin cirrus. One instruments in particular is the Tropospheric Emission Spectrometer (TES). TES is an FTS instrument and thus McClouds_RT is easily adaptable to TES, simply accounting for the minor differences in the viewing and re-calculating the channel optimisation for the TES spectral range and resolution. TES measures in the spectral range from 650–2250 cm^{-1} at either 0.08 cm^{-1} or 0.02 cm^{-1}, which is comparable to MIPAS. TES flies on the NASA Aura Spacecraft which

also carries the Microwave Limb Souder (MLS) which would provide valuable validation data.

9.2.3 Applications

The most exciting aspect of the developments of this thesis is the potential to apply the methods to large datasets.

Qualitative Studies

The qualitative studies mentioned in section 5.4.1 should be a high priorirty for an immediate application of the progress achieved by this thesis. Because the initial estimate methods are fast and simple it will allow large amounts of data to be processed with a relatively low overhead in terms of computation time.

A particularly interesting application of the initial estimate method woth be to investigate the distribution of upper tropospheric cirrus and the relative differences in cloud top height, optical thickness and effective radius over the entire MIPAS dataset to see if there are any noticeable correlations with the moderate ENSO (El Niño/Southern Oscillation) of 2002/2003.

Further Validation

The comparison between McClouds_RT and the AATSR retrieval was relatively succesful and the method should be pursued to obtain more statistics. In addition, a comparison to *in situ* aircraft measurement campaigns (e.g. CRYSTAL-FACE) would be the most effective validation if any coincident measurements can be found. Furthermore, if McClouds_RT is to be used outside of the tropics, it should be validated against ground based lidar and radar. However, these ground based instruments tend to attenuate at low altitudes and it may be that there are no suitable measurements for comparison.

Processing the MIPAS Data Catalogue

As stated above, if the entire MIPAS data set is to be processed, there must either be a vast improvement in the McClouds_RT processing time or there must be a concerted effort to pool computational resources. If either of these problems can be overcome, then the potential data from such an exercise would provide unprecedented spatial and temporal sampling of thin upper tropospheric cirrus on a global-scale. This would augment the state-of-the-art cirrus climatologies. Further, one of the most interesting uses of the data would be to perform a radiation budget study to determine the radiative impact of thin cirrus globally. A comparison to current radiation budget studies which use climatologies

which distinctly underestimate thin cirrus, would provide a distinct advance in the current understanding of the role of thin cirrus in global climate change.

Appendix A

Principal Components Analysis: Application to the MIPAS Data

This appendix provides the formal description of the PCA applied to the MIPAS data from August 2003, and discussed in section 3.5.

The MIPAS data was binned by latitude and tangent height, as stated in section 3.5.1, such that for each latitude bin, and then for each tangent height bin in turn, the PCA considers the $m \times n$ matrix of values denoted by $\mathbf{R} = \{r_{ij}\}$ where i is the ith limb scan in the bin and j is the radiance at the jth spectral point. The limb scan measurements can be considered row vectors in $\mathbb{R}^n$ and the radiances at a given spectral point as column vectors in $\mathbb{R}^m$. PCA is used to search for the best-fitting set of orthogonal axes in $\mathbb{R}^n$ to replace the initial n axes in this space.

A.1   The Eigenvalue Problem

Following the derivation in (Murtagh and Heck, 1987), the axis which best fits the points in $\mathbb{R}^n$ is determined by minimising a goodness of fit criterion. In this case, the goodness of fit criterion is defined as the squared deviation of points from the axis. If the frame is first centred on the data points, then minimising the squared deviation of points from the axis is equivalent to maximising the sum of squared projections onto the axis (by Pythagaros Theorem - see figure A.1). Denote the $m \times n$ matrix of centred data points as $\mathbf{X} = \{x_{ij}\}$ where i is the ith limb scan in the bin and j is the centred radiance at the jth spectral point. If $\mathbf{u}$ is the $n \times 1$ vector corresponding to the best fit axis, then $\mathbf{Xu}$ gives the projection of the m limb scans onto this axis. The squared projections of points on this new axis, for all points, is

$$(\mathbf{Xu})^T(\mathbf{Xu}), \tag{A.1}$$

where u is abritarily (but reasonably) chosen to be of unit length, i.e. $\mathbf{u}^T\mathbf{u} = 1$. A Lagrange
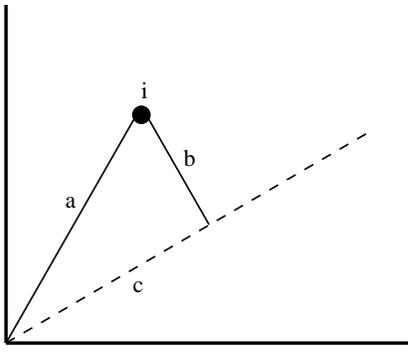
Figure A.1: Projection onto an axis. For each point i, by Pythagaros Theorem, $a^2 = b^2 + c^2$; since a is constant, the choice of the new axis which minimises b simultaneously maximises c.

multiplier, λ, can then be used to find the maximum of the multivariate quadratic form $\mathbf{u}^T \mathbf{Y} \mathbf{u}$ (where $\mathbf{Y} = \mathbf{X}^T \mathbf{X}$) subject to the constraint $\mathbf{u}^T \mathbf{u} = 1$. Thus, the optimal value of $\mathbf{u}$ is the solution of

$$\mathbf{Y} \mathbf{u} = \lambda \mathbf{u}, \tag{A.2}$$

which is the simple eigenvalue and eigenvector problem for the matrix $\mathbf{Y}$. By finding all eigenvectors, $\mathbf{u}_i$ and corresponding eigenvalues λ_i which satisfy equation A.2, the principle components are obtained such that $\mathbf{u}_i^T \mathbf{u}_j = 0$ when $i \neq j$. The eigenvalue, λ_i can be considered as a figure of merit for the corresponding axis, indicating the amount of variance explained by that axis. The principle components (eigenvectors) are ordered by decreasing order of corresponding eigenvalue. The set of principle components are often referred to in the literature as empirical orthogonal functions (EOFs) of variability.

A.2 Defining the Data Set

The following definitions apply to each tangent height bin, of a given latitude bin containing m limb scans.

The original limb scan measurements $\mathbf{R}$ are first centred by subtracting the mean spectrum from each data point to get the deviation from the mean. The centred matrix $\triangle \mathbf{R} \equiv \{ \triangle r_{ij} \}$ then has the $(i, j)^{\text{th}}$ value

$$\triangle r_{ij} = r_{ij} - \bar{r}_j \tag{A.3}$$

where the average spectrum obtained is given by the $n \times 1$ column vector, $\bar{\mathbf{r}}$, with j^{th} value

$$\bar{r}_j = \frac{1}{n} \sum_{i=1}^{m} r_{ij} . \tag{A.4}$$

To treat each of the radiance measurements equally, the deviations from the mean at each

wavenumber are then normalised by the standard deviation σ_j over all limb scans at this wavenumber. The resultant matrix is known as the standardised matrix, $\mathbf{X}$ and the $(i,j)^{\text{th}}$ value is given by:

$$x_{ij} = \frac{r_{ij} - \bar{r}_j}{\sigma_j \sqrt{n}} \tag{A.5}$$

where

$$\sigma_j^2 = \frac{1}{n} \sum_{i=1}^{m} (r_{ij} - \bar{r}_j)^2 \; . \tag{A.6}$$

The $1/\sqrt{n}$ term is introduced into the standardised matrix for the convenient definition of correlations. It is thus the matrix, $\mathbf{Y} \equiv \mathbf{X}^T\mathbf{X}$ for which the eigenvector and eigenvalue soultions should be found (as in section 3.5.1). This is the matrix of the $(j,k)^{\text{th}}$ term:

$$y_{jk} = \sum_{i=1}^{m} x_{ij}x_{ik} = \frac{1}{n} \sum_{i=1}^{m} \frac{(r_{ij} - \bar{r}_j)(r_{ik} - \bar{r}_k)}{\sigma_j \sigma_k} \tag{A.7}$$

which is the correlation coefficient between variables j and k.

The limb scan measurements can then be approximated using the resultant eigenvectors. The jth value of an approximated limb scan will be given by:

$$r_j \cong \tilde{r}_j = \bar{r}_j + \sum_{i=1}^{p} c_i u_{i,j} \sigma_j \tag{A.8}$$

where $\bar{r}_j$ is the mean spectrum at the jth wavenumber, σ_j is the standard deviation at the jth wavenumber, $u_{i,j}$ denotes the value at the jth wavenumber of the ith eigenvector $\mathbf{u}_i$ (i.e. principle component) and p is the number of principle components/eigenvectors included in the approximation. As p increases, so will the accuracy of the approximation. As the eigenvectors are orthogonal and normalised, the coefficient c_i for each limb scan in the bin is given by

$$c_i = \sum_{j=1}^{n} \frac{(r_j - \bar{r}_j)u_{i,j}}{\sigma_j} \tag{A.9}$$

where j sums over the individual wavenumbers.

The PCA considers the statistical quantities defined above (i.e. standard deviation, mean, etc.) in terms of *population* rather than *sample*. That is to say, the PCA is carried out on each height bin of each latitude bin, under the assumption that the MIPAS data being examined is taken as all that can be immediately studied, rather than being representative of an underlying distribution of measurements.

A.3 Aspects of Practical Implementation

This section details the practical considerations for implementing the PCA on a $m \times n$ matrix of m limb scan measurements (with $m < n$) for a single tangent height bin, of a given latitude bin.

In order to practically obtain the eigenvectors and eigenvalues of the correlation matrix $\mathbf{Y}$, first note that it is an $n \times n$ square-matrix. Next, assume that $\mathbf{Y}$ has eigenvalues λ_1, λ_2, ..., λ_k which are nondegenerate and have corresponding linearly independent eigenvectors $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k$. Define the matrix composed of the eigenvectors

$$\mathbf{P} = [\mathbf{u}_1 \, \mathbf{u}_2 \, \cdots \, \mathbf{u}_k] \tag{A.10}$$

and the diagonal matrix composed of eigenvalues

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix}. \tag{A.11}$$

Then

$$\mathbf{Y} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1} \tag{A.12}$$

(which is known as the eigenvalue decomposition) and the problem of finding the eigenvalues can be reduced to finding the diagonalisation of the correlation matrix. As $\mathbf{Y}$ is both real and symmetric, Householder reductions[1] can be used to reduce the correlation matrix into tridiagonal[2] form, which is an $O(n^3)$ problem (if eigenvectors are required[3]) and then using the iterative QR method[4] to diagonalise the resultant matrix. However, as the limb spectra are measured at high resolution, there are a large number of spectral points (n=8001 in this case), and as a result the Householder reductions become computationally onerous. Therefore, a more efficient computational method should be employed.

In order to understand the reasoning for selecting a method to solve the eigenvalue value problem, it is useful to first study the analogous problem for non-square matrices, i.e. the

---

[1]The Householder algorithm reduces an $n \times n$ symmetric matrix to tridiagional form by $n - 2$ orthogonal transformations. Each transformation annihilates the required part of a whole column and whole corresponding row. For more details see pp. 470-475 of Press et al. (1992)

[2]A square matrix is tridiagional if it has nonzero elements only on the diagonal and the slots horizontally or vertically adjacent the diagonal (i.e. along the subdiagonal and superdiagonal).

[3]If the eigenvectors of the final tridiagonal matrix are found then the eigenvectors of $\mathbf{Y}$ can be obtained by applying the accumulated Householder matrices to those eigenvectors.

[4]The QR algorithm consists of a sequence of orthogonal transformations which uses the fact that any real matrix can be decomposed into an orthogonal and upper tridiagonal matrix. The diagonal elements will converge to the eigenvalues and the super- and sub- diagonal elements will converge to zero. Full details can be found in Press et al. (1992) pp. 476-478.

singular value problem. The singular value decomposition of a non-square $m \times n$ real matrix $\mathbf{A}$ (with $m < n$) is of the form

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T \tag{A.13}$$

where $\mathbf{U}$ is an $m \times p$ matrix, $\mathbf{V}$ is a square $n \times p$ matrix, and $\mathbf{D}$ is an $p \times p$ diagonal matrix containing the p singular values of $\mathbf{A}$ (where $p \leqslant m$ is the rank of $\mathbf{A}$). The columns of $\mathbf{U}$ are described as the left singular vectors and the columns of $\mathbf{V}$ are described as the right singular vectors. Notice the similarity between the form of equation A.13 and its square matrix analogue, equation A.12.

The standard eigenvalue problem is meaningless for non-square matrices as $\mathbf{A}\mathbf{u}$ will have a different dimension to $\mathbf{u}$. However, an eigenvalue problem can be constructed by considering the symmetric problem:

$$\begin{pmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \tag{A.14}$$

where $\mathbf{v}$ is of dimension n and $\mathbf{u}$ is of dimension m. The vectors $\mathbf{u}$ and $\mathbf{v}$ are singular vectors of $\mathbf{A}$ and λ is a singular value.

To apply the symmetric problem to the MIPAS data set, replace $\mathbf{A}$ in equation A.14 with $\mathbf{X}$ (as defined in equation A.5). Notice that the symmetric eigenvalue problem is equivalent to the 'shifted' eigenvalue problem:

$$\mathbf{X}\mathbf{v} = \lambda\mathbf{u} \tag{A.15}$$
$$\mathbf{X}^T\mathbf{u} = \lambda\mathbf{v} . \tag{A.16}$$

By substitution:

$$\mathbf{X}^T\mathbf{X}\mathbf{v} = \lambda\mathbf{X}^T\mathbf{u} = \lambda^2\mathbf{v} \tag{A.17}$$
$$\mathbf{X}\mathbf{X}^T\mathbf{u} = \lambda\mathbf{X}\mathbf{v} = \lambda^2\mathbf{u} \tag{A.18}$$

showing that $\mathbf{u}$ and $\mathbf{v}$ are the eigenvectors of $\mathbf{X}\mathbf{X}^T$ and $\mathbf{X}^T\mathbf{X}$ respectively. Therefore, as the correlation matrix $\mathbf{Y} = \mathbf{X}^T\mathbf{X}$ (as defined in equation A.7), then the eigenvectors of $\mathbf{Y}$ are equivalent to the right singular vectors of $\mathbf{X}$ and the eigenvalues of $\mathbf{Y}$ are equivalent to the square of the singular values of $\mathbf{X}$.

From equation A.13 it can be seen that finding the singular vectors of $\mathbf{X}$ can be simplified to a diagonalisation problem. As $\mathbf{X}$ is rectangular (with $m < n$), the diagonalisation problem also implies that the form of the diagonal matrix will be square, i.e. of different dimension to

X. It is solved by first reducing $\mathbf{X}$ to an $m \times m$ lower bidiagonal matrix[5] using householder reductions, then diagonalising this matrix by using the iterative QR method. This method will also give $m - p$ surplus singular vectors.

By making these observations, it is intuitively clear that the eigenvectors of $\mathbf{Y}$ are more efficiently found by searching for the singular vectors of the smaller matrix $\mathbf{X}$ as there are far less operations required i.e. $O(m^3)$ householder reductions as opposed to $O(n^3)$, as $m \ll n$. In additon, the workload of the QR method is $O(N)$ per iteration for an $N \times N$ tridiagonal matrix, so only employing the QR method to the $m \times m$ bidiagonal form of $\mathbf{X}$ as opposed to the $n \times n$ tridiagonal form of $\mathbf{Y}$ will further reduce the number of operations.

---

[5]A square matrix is lower bidiagonal if it has nonzero elements only on the diagonal and subdiagonal. Therefore, a bidiagonal matrix is also tridiagonal.

Appendix B

Monte Carlo Scattering Algorithm

In order to use the Monte Carlo method to simulate the path of a photon through a cloudy medium, the radiative transfer process is considered as the stationary Markov chain whose states are photon interactions with atoms (or molecules) of the medium. Each interaction can have one of two outcomes; absorption or scattering of the photon, i.e. the photon path terminates, or the photon leaves the collision travelling in a new direction.

Each photon is traced through the cloud by determining the distance between scattering events and the direction after scattering by sampling the appropriate probability distributions, until the photon is either absorbed or exits the cloud.

B.1 Describing Multiple Scattering

The straight path between two successive collisions is called a *free path*, and is an important quantity in the simulation. The angle between the previous and the new direction is called the scattering angle.

Now consider further the notion of the free path. For any pencil of monochromatic light travelling in a direction $\vec{\psi}$ after traversing a path of length L, and in the absence of multiple scattering, the radiance is attenuated according to Beer's law

$$I(\vec{r}, \omega) = I(\vec{r_0}, \omega) \exp(-\tau) \tag{B.1}$$

where $\tau = \beta_{\text{ext}} L$ is the optical depth of the path from initial position $\vec{r_0}$ to final position $\vec{r} = \vec{r_0} + \vec{\psi} L$. The quantity β_{ext} is the volume extinction coefficient, which is equal to the sum of the scattering coefficient and the absorption coefficient, i.e. $\beta_{\text{ext}} = \beta_s + \beta_a$.

However, in a cloudy medium there is multiple scattering, but the length of the free path between collisions is uncertain. The *free path length l* of photons travelling in a direction

$\vec{\psi}$ can be expressed in terms of a probability density function

$$f_l(x) = \beta_{\text{ext}}[\vec{r}(x)] \exp\left\{-\int_0^x \beta_{\text{ext}}[\vec{r}(x')] \, dx'\right\} \tag{B.2}$$

where $\vec{r}(x) = \vec{r}(0) + x\vec{\psi}$ (such that x has dimensions of length), $\vec{r}(0)$ is the initial position, and $\beta_{\text{ext}}[\vec{r}(x)]$ is the volume extinction coefficient for the path from position $\vec{r}(0)$ to position $\vec{r}(x)$. The quantity $\tau(x) = -\int_0^x \beta_{\text{ext}}[\vec{r}(x')] \, dx'$ is the optical depth of $[\vec{r}(0), \vec{r}(x)]$. The free path length pdf can be interpreted such that $f_l(x)dx$ is the probability that $l \in (x, x + dx)$ where $|\vec{r}(x) - \vec{r}(0)| = x$.

The cosine of the scattering angle, μ for the collision at the point $\vec{r}$ has the pdf $\mathbf{P}(\mu, \vec{r})$, which is known as the phase function of the particle at point $\vec{r}$. The phase function is normalised by

$$\int_{-1}^{+1} \mathbf{P}(\mu, \vec{r}) \, d\mu = 1 \tag{B.3}$$

The probabilities of absorption and scattering at a point $\vec{r}$ are defined by

$$1 - \tilde{\omega}(\vec{r}) = \frac{\beta_a(\vec{r})}{\beta_{\text{ext}}(\vec{r})}, \qquad \tilde{\omega}(\vec{r}) = \frac{\beta_s(\vec{r})}{\beta_{\text{ext}}(\vec{r})} \tag{B.4}$$

where $\tilde{\omega}(\vec{r})$ is known as the *single scatter albedo* (of the particle) at point $\vec{r}$.

B.2   The Algorithm

The following pseudocode is used to simulate the photon paths (for a pencil beam of light) through a cloud:

1. The original photon position and direction of travel are selected:

$$\vec{r}_0 = (x_0, y_0, z_0), \qquad \vec{\psi}_0 = (a_0, b_0, c_0) \tag{B.5}$$

2. The current coordinates of the photon (x', y', z') are set i.e. $x' = x_0$, $y' = y_0$, $z' = z_0$, and the current direction of the photon (a', b', c') is set i.e. $a' = a_0$, $b' = b_0$, $c' = c_0$.

3. The free path length l is simulated.

4. The escape from the cloud is examined.

   - If the photon has escaped from the cloud:

     (a) Record the exit point from the cloud, the exit direction, and the number of scatters undergone.

(b) If there are no more photons to enter the cloud then the algorithm ends here, or else go to **1**.

• If the photon has not escaped from the cloud then continue to **5**.

5. The coordinates of the new collision are calculated :

$$x = x' + a'l, \qquad y = y' + b'l, \qquad z = z' + c'l \qquad \text{(B.6)}$$

6. The type of collision (absorption or scattering) is simulated.

• If the photon is absorbed then:

(a) The path terminates and the absorption point and number of scatters undergone are recorded.

(b) If there are no more photons to enter the cloud then the algorithm ends here, or else go to **1**.

• If the photon is scattered then continue to **7**.

7. The cosine of the scattering angle μ is simulated.

8. The coordinates of a new direction (a, b, c) are calculated.

9. Now set the current photon location to the new coordinates $(x', y', z') = (x, y, z)$, and set the current direction of travel to the new direction $(a', b', c') = (a, b, c)$.

10. Increment the counter for the number of scatters the photon has undergone.

11. Go to **3**.

B.3   Random Simulation of Free Path Length

When, simulating the transfer of photons through a cloudy medium, the difficulty comes when simulating the free path length if β_{ext} is not constant for the cloud. By integrating

equation B.2, the cumulative distribution function[1] (cdf) of the free path length is obtained

$$
\begin{aligned}
F_l(x) &= \int_0^x f_l(x)\, dx \\
&= \int_0^x \beta_{\text{ext}}[\vec{r}(x)] \exp[-\tau(x)]\, dx \\
&= \int_0^{\tau(x)} \exp[-\tau(x)]\, d\tau(x) \qquad\qquad\qquad (\text{B.7})\\
&= -\exp[-\tau(x)]\big|_0^{\tau(x)} \\
F_l(x) &= 1 - \exp[-\tau(x)]
\end{aligned}
$$

for $x \geq 0$. This can also be expressed in terms of optical depth

$$
\tau(x) = -\ln[1 - F_l(x)]. \qquad\qquad\qquad (\text{B.8})
$$

This is easy to solve if $\beta_{\text{ext}}[\vec{r}(x)]$ is a step function, or slowly varying. However, in the context of remote sounding cloud properties, a bulk (or average) volume extinction coefficient for the (whole) cloud is more appropriate. Therefore, in McClouds_FM it is assumed that β_{ext} is constant through the cloud. This simplifies the problem of simulating the free path length $l = \tau(l)/\beta_{\text{ext}}$.

Note that as $\int_0^\infty F_l(x)\, dx = 1$ must hold (as $F_l(x)$ is the cdf of l), then for any x, $F_l(x)$ is uniformly distributed over the interval $[0, 1]$, i.e. any value for the cdf is equally likely on the interval $[0, 1]$. Now, let $F_l(x) = Y$ where Y is an uniformly distributed random number on the set $[0, 1]$, then the free path length can be simulated by,

$$
l = \frac{-\ln[1 - Y]}{\beta_{\text{ext}}} \qquad\qquad\qquad (\text{B.9})
$$

where Y is generated by the computer.

B.4   The Scattering Angle

The phase function measures the angular distribution of radiation after interaction with a particle. The convention used in McClouds_FM is that the phase function describes the angular distribution of radiation with respect to the direction of radiation before the collision by two angles; the azimuthual and polar scattering angles ϕ and θ, respectively. These scattering angles are illustrated in figure B.1. It is assumed in McClouds_FM that for any collision, for a given polar angle, scattering of radiation is azimuthually isotropic.

---

[1]The *cumulative distribution function* (cdf) of a real-valued random variable X is defined as

$$
P(x) = Pr\{X \leq x\} \quad \forall \in \mathbf{R}.
$$
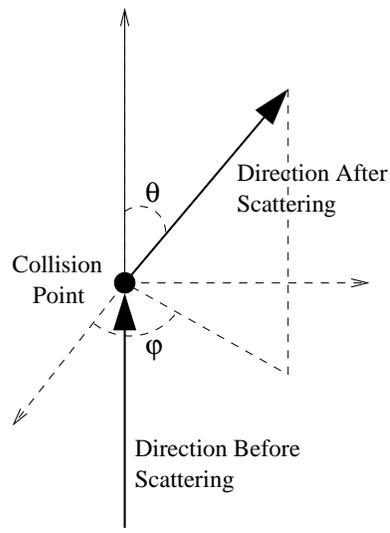
Figure B.1: Scattering Angles.

Thus, for scattering from any collision the new azimuthual angle, can be simulated from

$$\phi = 2\pi\alpha \qquad (B.10)$$

where α is a random uniform deviate.

However, radiation is scattered anisotropically over the polar angle. The cosine of the polar angle has a pdf given by the polar phase function, $\mathbf{P}(\cos\theta)$ and is normalised as in equation B.3.

The first step is to form the cdf of the phase function, which shall be denoted by $\mathbf{D}(\cos\theta)$. As $\cos\theta$ and the cdf $\mathbf{D}(\cos\theta)$ are one-to-one, $\cos\theta$ can be sampled by first sampling $y = \mathbf{D}(\cos\theta)$ and then solving for $\cos\theta$ by inverting $\mathbf{D}(\cos\theta)$, i.e. $\cos\theta = \mathbf{D}^{-1}(y)$. This method is known as sampling via inversion of the cdf, and is a general method for generating random variates from univariate distributions. As the cdf is uniformly distributed on $[0, 1]$, a random number generator is used to obtain random uniform deviates to generate a sample R from the cdf $\mathbf{D}(\cos\theta)$. Then the value of $\cos\theta$ is determined by inversion $\cos\theta = \mathbf{D}^{-1}(R)$. This can be shown to have the required distribution. Under the stated transformation, and for a particular sampled value from this method, e.g. $\cos\Theta$,

$$\begin{aligned} Pr\{\cos\Theta \le \cos\theta\} &= Pr\{\mathbf{D}^{-1}(R) \le \cos\theta\} \\ &= Pr\{R \le \mathbf{D}(\cos\theta)\} \\ &= \mathbf{D}(\cos\theta). \end{aligned} \qquad (B.11)$$

Bibliography

Ackerman, S. A., Smith, W. L., Spinhirne, J. D. and Revercomb, H. E. (1990). Tue 27-28 October 1986 FIRE IFO Cirrus Case Study: Spectral Properties of Cirrus Clouds in the 8-12 μm Window, *Mon. Wea. Rev.* **118**: 2377–2388.

Albrecht, B. A. (1992). Clouds and their effect on climate, *in* W. A. Nierenberg (ed.), *Encyclopedia of Earth System Science*, Vol. 1, Academic Press, Inc., San Diego, CA, pp. 545–556.

Arking, A. (1991). The radiative effects of clouds and their impact on climate, *Bulletin of the Am. Meteor. Soc.* **71**: 795–813.

Arnott, W. P., Dong, Y. Y., Hallett, J. and Poellot, M. R. (1994). Role of small ice crystals in radiative properties of cirrus: A case study, FIRE II, november 22, 1991, *J. Geophys. Res.* **99**: 1371–1381.

Baker, M. B. (1997). Cloud microphysics and climate, *Science* **276**: 1072–1078.

Baran, A. J. (2003). Simulation of infrared scattering from ice aggregates by use of a size-shape distribution of circular ice cylinders, *Appl. Opt.* **42**: 2811–2818.

Baran, A. J. (2004). On the scattering and absorption properties of cirrus cloud, *J. Quant. Spectrosc. Radiat. Transfer.* **89**: 17–63.

Baran, A. J. and Francis, P. N. (2004). On the radiative properties of cirrus cloud at solar and thermal wavelengths: A test of model consistency using high-resolution airborne radiance measurements, *Q. J. R. Meteorol. Soc.* **130**: 763–778.

Baran, A. J., Francis, P. N., Labonnote, L.-C. and Doutriaux-Boucher, M. (2001). A scattering phase function for ice cloud: Tests of applicability using aircraft and satellite multi-wavelength radiance measurements of cirrus, *Quart. J. Roy. Meteor. Soc.* **127**: 2395–2416.

Baran, A. J., Watts, P. D. and Francis, P. N. (1999). Testing the coherence of cirrus microphysical and bulk properties retrieved from dual-viewing multispectral satellite radiance measurements, *J. Geophys. Res.* **104**: 31673–31683.

Barnes, A. A. (1980). Observations of ice particles in thin air, *J. Rech. Atmos.* **14**: 311–315.

Baum, B. A., Heymsfield, A. J., Yang, P. and Bedka, S. T. (2005). Bulk scattering models for the remote sensing of ice clouds. 1: Microphysical data and models, *J. Appl. Meteor.* **XX**: (in press).

Baumgardner, D., Chepfer, H., Raga, G. B. and Kok, G. L. (2005). The shapes of very small cirrus particles derived from in situ measurements, *Geophys. Res. Lett.* **32**: L01806. doi:10.1029/2004GL021300.

Birk, M. and Wagner, G. (2003). Radiometric characterisation of mipas/envisat, *Atmospheric Science from Space using Fourier Transform Spectrometry Workshop*, IMK, Germany.

Buschmann, N. and Albers, F. (2000). Spatial inhomogeneities of cirrus clouds as derived from in situ ice particle measurements, *Phys. Chem. Earth Pt. B* **25**: 77–81.

Carli, B. et al. (2004). First results of mipas/envisat with operational level 2 code, *Advances in Space Research* **33**: 1012–1019.

Carlotti, M. (1988). Global-fit approach to the analysis of limb-scanning atmospheric measurements, *Appl. Opt.* **27**: 3250–3254.

Cess, R. D. et al. (1990). Intercomparsion and interpretation of cloud-climate feedback processes in nineteen atmospheric general circulation models, *J. Geophys. Res.* **95**: 16601–16615.

Cess, R. D. et al. (1996). Cloud feedback in atmospheric general circulation models: an update, *J. Geophys. Res.* **101**: 12791–12794.

Chen, J. P., McFarquhar, G. M., Heymsfield, A. J. and Ramanathan, V. (1997). A modeling and observational study of the detailed microphysical structure of tropical cirrus anvils, *J. Geophys. Res.* **102(D6)**: 6637–6653.

Chen, T., Rossow, W. B. and Zhang, Y. (2000). Radiative effects of cloud-type variations, *J. Climate* **13**: 264–286.

Chepfer, H. et al. (2005). Particle habit in tropical ice clouds during CRYSTAL-FACE: Comparison of two remote sensing techniques with insitu observations, *JGR* **XX**: X–X. Submitted September 2004.

Chung, S., Van Delst, S. A. A. P. F. and Menzel, W. P. (2000). Model calculations and interferometer measurements of ice-clouds characteristics, *J. Appl. Meteor.* **39**: 634–644.

Collins, D. G., Blattner, W. G., Wells, M. B. and Horak, H. G. (1972). Backward monte carlo calculations of the polarization characteristics of the radiation emerging from spherical-shell atmospheres, *Appl. Opt.* **11**: 2684–2705.

Colman, R. A. (2003). A comparison of climate feedbacks in general circulation models, *Clim Dyn* **20**: 865–873.

Colman, R. A., Power, S. B. and McAvaney, B. J. (1997). Non-linear climate feedback analysis in an atmospheric gcm, *Clim Dyn* **13**: 717–731.

COMET Program (2005). Winter microphysics topics: Winter precipitation processes [online], *Available from*: http://meted.ucar.edu/norlat/snow/ [Accessed 26 January 2005]. UCAR Cooperative Program for Operational Meteorology, Education and Training (COMET).

Courcoux, N. and Emde, C. (2005). The atmospheric radiative transfer system [online], *Available from*: http://www.sat.uni-bremen.de/arts/ [Accessed 1 April 2005]. The Satellite Microwave Atmospheric Sounding Group, Institute für Umweltphysik (IUP) Unversität Bremen.

Davis, C., Emde, C. and Harwood, R. (2004). A 3D polarized reversed monte carlo radiative transfer model for mm and sub-mm passive remote sensing in cloudy atmospheres, *IEEE TGARS* **43**: 1096–1101.

Davis, J. M., Cox, S. K. and McKee, T. (1979). Vertical and horizontal distributions of solar absorption in finite clouds, *J. Atmos. Sci.* **36**: 1976–1984.

Dean, S., Lawrence, B., Grainger, R. G. and Heuff, D. (2005). Orographic cloud in a GCM: The missing cirrus, *Climate Dynamics* **24**: 771–780.

Del Genio, A. D. (2002). GCM simulations of cirrus for climate studies, pp. 310–326.

Del Genio, A. D. and Kovari, W. (2002). Climatic properties of tropical precipitating convection under varying environmental conditions, *J. Climate* **15**: 2597–2615.

DeMott, P. J. (2002). Laboratory studies of cirrus cloud processes, pp. 102–135.

DeMott, P. J. et al. (1998). The role of heterogeneous freezing nucleation in upper tropospheric clouds: Inferences from success, *Geophys. Res. Lett.* **25**: 1387–1390.

Dowling, D. R. and Radke, L. F. (1990). A summary of the physical properties of cirrus clouds, *J. Appl. Meteor.* **29**: 970–978.

Dudhia, A. (2004). The Reference Forward Model (RFM): Software User's Manual (SUM) [online], *Available from:* http://www-atm.physics.ox.ac.uk/RFM/sum.html. Updated Version of ESA Technical Document PO-MA-OXF-GS-0003.

Dudhia, A., Jay, V. L. and Rodgers, C. D. (2002). Microwindow selection for high spectral resolution sounders, *Appl. Opt.* **41**: 3665–3673.

Edwards, D. P. (1992). *GENLN2: a general line-by-line atmospheric transmittance and radiance model/ Version 3.0 description and users guide*, National Center for Atmospheric Research, Boulder, Colorado. NCAR/TN-367+STR.

Emde, C. et al. (2005). A polarized discrete ordinate scattering model for simulations of limb and nadir longwave measurements in 1D/3D spherical atmospheres, *J. Geophys. Res.* **109(D24)**: D24207. doi:10.1029/2004JD005140.

ESA (2000). *Envisat-MIPAS: An Instrument for Atmospheric Chemistry and Climate Research, SP-1229*, ESA Publications Division, ESTEC, Noordwijk, The Netherlands. Technical Coordinator: C. Readings. Editor: R. A. Harris.

ESA (2005). Envisat instruments: MIPAS [online], *Available from*: http://envisat.esa.int/instruments/mipas/ [Accessed 20 May 2004].

Ewen, G. B. L. (2002). Infrared limb observations of cloud: First year DPhil report, Department of Atmospheric, Oceanic and Planetary Physics, University of Oxford.

Ewen, G. B. L., Grainger, R. G., Lambert, A. and Baran, A. J. (2005). Infrared radiative transfer modelling in a 3D scattering cloudy atmosphere: Application to limb sounding measurements of cirrus, *J. Quant. Spectrosc. Radiat. Transfer.* **96**: 45–74.

Fletcher, R. (1971). A modified Marquardt subroutine for non-linear least squares, *Technical Report HL71/2812(C13)*, Theoretical Physics Division, Atomic Energy Research Establishment (A.E.R.E), Harwell, U.K. Report R.6799.

Fowler, L. D. and Randall, D. A. (1999). Simulation of upper tropospheric clouds with the Colorado State University general circulation model, *J. Geophys. Res.* **104(D6)**: 6101–6122. 10.1029/1998JD200074.

Francis, P. N., Jones, A., Saunders, R. W., Shine, K. P., Slingo, A. and Sun, Z. (1999). An observational and theoretical study of the radiative properties of cirrus: Some results from ICE'89, *Quart. J. Roy. Meteor. Soc.* **120**: 809–848.

Fridland, A. M. et al. (2004). Evidence for the predominance of mid-tropospheric aerosols as subtropical anvil cloud nuclei, *Science* **304**: 718–722. doi: 10.1126/science.1094947.

Frieden, B. R. (2004). *Science from Fisher information : a unification*, 2nd edn, Cambridge University Press, Cambridge, UK.

Fu, Q. (1996). An accurate parameterization of the solar radiative properties of cirrus clouds for climate models, *J. Climate* **9**: 2058–2082.

Furakawa, Y. and Kobayashi, T. (1978). On the growth mechanisms of polycrystalline snow crystals with a specific grain boundary, *J. Cryst. Growth* **49**: 259–264.

Garrett, T. J. et al. (2004). Convective generation of cirrus near the tropopause, *J. Geophys. Res.* doi:10.1029/2004JD004952.

Garrett, T. J. et al. (2005). Evolution of a florida cirrus anvil, *J. Atmos. Sci.* submitted.

Garrett, T. J., Gerber, H., Baumgardner, D. G., Twohy, C. H. and Weinstock, E. M. (2003). Small, highly reflective ice crystals in low-lattitude cirrus, *Geophys. Res. Lett.* **30(21)**: 2132. doi:10.1029/2003GL018153.

Gerber, H., Arends, B. G. and Ackermann, A. S. (1994). New microphysics sensor for aircraft use, *J. Atmos. Ocean. Tech.* **1**: 22–27.

Gierens, K. (2003). On the transition between heterogeneous and homogeneous freezing, *Atmos. Chem. Phys.* **3**: 437–446.

Glatthor, N. et al. (1999). Intercomparison of the KOPRA and the RFM radiative transfer codes, *Proceedings of the European Symposium on Atmospheric Measurements for Space*, pp. pp. 757–764.

Goodman, J. et al. (1998). Shape and size of contrails ice particles, *Geophys. Res. Lett.* **25(9)**: 1327–1330.

Gourdeau, J. (2004). Clouds and particles: Characteristics of the different cloud types [online], *Available from*: http://www.espere.net [Accessed 14 January 2005]. Environmental Science Published for Everybody Round the Earth (ESPERE).

Haag, W. et al. (2003). Freezing thresholds and cirrus cloud formation mechanisms inferred from in situ measurements of relative humidity, *Atmos. Chem. Phys.* **3**: 1791–1806.

Hallett, J., Arnott, W. P., Bailey, M. P. and Hallett, J. T. (2002). Ice crystals in cirrus, *in* D. K. Lynch, K. Sassen, D. O. Starr and G. Stephens (eds), *Cirrus*, Oxford University Press, Inc., New York, pp. 147–167.

Hamilton, C. J. (1997). Clouds from space [online], *Available from*: http://www.solarviews.com [Accessed 18 January 2005].

Harrison, E. F. et al. (1990). Seasonal variation of cloud radiative forcing derived from the Earth Radiation Budget Experiment, *J. Geophys. Res.* **95**: 18687–18703.

Harshvardhan, Weinman, J. A. and Davies, R. (1981). Transport of infrared radiation in cuboidal clouds, *J. Atmos. Sci.* **38**: 2500–2513.

Hartmann, D. L., Holton, J. R. and Fu, Q. (2001). The heat balance of the tropical tropopause, cirrus, and stratospheric dehydration, *Geophys. Res. Lett.* **28(10)**: 1969–1972.

Havemann, S. and Baran, A. J. (2001). Extension of T-matrix to scattering of electromagnetic plane waves by non-axisymmetric dielectric particles: application to hexagonal ice cylinders, *J. Quant. Spectrosc. Radiat. Transfer.* **70**: 139–158.

Herring, D. and Kaufman, Y. (2005). NASA Earth observatory: Cloud forcing [online], *Available from*: http://earthobservatory.nasa.gov/Observatory/Datasets/cldforc.erbe.html [Accessed 28 January 2005]. NASA Earth Science Enterprise.

Hervig, M. and McHugh, M. (1999). Cirrus detection using HALOE measurements, *Geophys. Res. Lett.* **26**: 719–722.

Heymsfield, A. J. (1975). Cirrus uncinus generating cells and the evolution of cirriform clouds. parts I-III., *J. Atmos. Sci.* **32**: 799–830.

Heymsfield, A. J. (1986). Ice particles observed in cirriform cloud at -83 c and implications for polar stratoshperic clouds, *J. Atmos. Sci.* **43**: 851–855.

Heymsfield, A. J. and McFarquhar, G. M. (1996). High albedo of cirrus in the tropical pacific warm pool: Microphysical interpretations from CEPEX and from Kwajalein, Marshall Islands, *J. Atmos. Sci.* **53**: 2424–2451.

Heymsfield, A. J. and McFarquhar, G. M. (2002). Mid-latitude and tropical cirrus: Microphysical properties, *in* D. K. Lynch, K. Sassen, D. O. Starr and G. Stephens (eds), *Cirrus*, Oxford University Press, Inc., New York, pp. 78–101.

Heymsfield, A. J. and Miloshevich, L. M. (1993). Homogeneous ice nucleation and supercooled liquid water in orographic wave clouds, *J. Atmos. Sci.* **50**: 2335–2353.

Heymsfield, A. J. and Miloshevich, L. M. (1995). Relative humidity and temperature influences on cirrus formation and evolution: Observations from wave clouds and FIRE II, *J. Atmos. Sci.* **52**: 4302–4326.

Heymsfield, A. J. and Miloshevich, L. M. (2003). Paramterizations for the cross-sectional area and extinction of cirrus and stratiform ice cloud particles, *J. Atmos. Sci.* **60**: 936–956.

Heymsfield, A. J. and Platt, C. M. R. (1984). A parameterization of the particle size spectrum of the ice clouds in terms of the ambient temperature and the ice water content, *J. Atmos. Sci.* **41**: 846–855.

Heymsfield, A. J. et al. (2004). Effective ice particle densities for cold anvil cirrus, *Geophys. Res. Lett.* doi:10.1029/2003GL018311.

Heymsfield, A. J., Lawson, R. P. and Sachse, G. W. (1998). Growth of ice crystals in a precipitating contrail, *Geophys. Res. Lett.* **25(9)**: 1335–1338.

Heymsfield, A. J., Miller, K. M. and Spinhirne, J. D. (1990). The 27-28 october 1986 fire ifo cirrus case study: Cloud microstructure, *Mon. Wea. Rev.* **118**: 2313–2328.

Höpfner, M. (2004). Study of polar stratospheric clouds on high resolution mid-ir limb emission spectra, *J. Quant. Spectrosc. Radiat. Transfer.* **83**: 93–107.

Höpfner, M. et al. (2002). Evidence of scattering of tropospheric radiation by PSCs in mid-IR limb emission spectra: MIPAS-B observations and KOPRA simulations, *Geophys. Res. Lett.* 10.129/2001GL014443.

Hu, Y. X., Winker, D., Yang, P., Baum, B., Poole, L. and Vann, L. (2001). Identification of cloud phase from PICASSO-CENA lidar depolarization: a multiple scattering sensitivity study, *J. Quant. Spec. Rad. Tran.* **70**: 69–579.

Huang, H. L. et al. (2004). Inference of ice cloud properties from high spectral resolution infrared observations, *IEEE Trans. Geosci. and Remote Sens.* **42**: 842–853.

IPCC (2001). *Climate Change 2001: The Scientific Basis*, Cambridge Univ. Press, Cambridge, UK. J. T. Houghton, L.G. Meira Filho, B. A. Callander, N. Harris, A. Kattenberg, and K. Maskell (eds).

Irwin, P. G. J. and Dyudina, U. (2002). The retrieval of cloud structure maps in the equatorial region of jupiter using a principle component analysis of Galileo/NIMS data, *Icarus* **156**: 52–63.

Jensen, E. J., Toon, O. B., Selkirk, H. B., Spinhime, J. D. and Schoerbel, M. R. (1996). On the formation and persistence of subvisible cirrus clouds near the tropical tropopause, *J. Geophys. Res.* **101**: 21316–21375.

Jin, Y., Rossow, W. B. and Wylie, D. P. (1996). Comparison of the climatologies of high-level clouds from HIRS and ISCCP, *J. Climate* **9**: 2850–2879.

Kahn, B. H., Eldering, A., Ghil, M., Bordoni, S. and Clough, S. A. (2004). Sensitivity analysis of cirrus cloud properties from high-resolution infrared spectra. Part I: Methodology and synthetic cirrus, *J. Climate* **17**: 4856–4870.

Kahn, B. H., Eldering, A., Irion, W., Mills, F. P., Sen, B. and Gunson, M. R. (2002). Cloud identification in Atmospheric Trace Molecule Spectroscopy infrared occultation measurements, *Appl. Opt.* **41**: 2768–2780.

Kahn, B. H. et al. (2003). Near micron-sized cirrus cloud particles in high-resolution infrared spectra: An orographic case study, *Geophys. Res. Lett.* **30(8)**: 1441. doi:10.1029/2003GL016909.

Kent, G. S., Wang, P. H. and Skeens, K. M. (1997). Discrimination of cloud and aerosol in the Stratospheric Aerosol and Gas Experiment III occultation data, *Appl. Opt.* **36**: 8639–8649.

Kiehl, J. T. (1994). On the observed near cancellation between longwave and shortwave cloud forcing in tropical regions, *J. Climate* **7**: 559–565.

Kiehl, J. T., Hack, J. J. and Briegleb, B. P. (1994). The simulated earth radiation budget of the National Center for Atmospheric Research community climate model CCM2 and comparisons with the earth radiation budget experiment (ERBE), *J. Geophys. Res.* **99**: 20815–20827.

Kinne, S., Akerman, T. P., Heymsfield, A. J., Valero, F. P. J., Sassen, K. and Spinhirne, J. D. (1992). Cirrus microphysics and radiative-transfer - cloud field-study on 28 october 1986, *Mon. Wea. Rev.* **120**: 661–684.

Klinedinst, G. (2003). CRYSTAL-FACE space station images [online], *Available from*: http://cloud1.arc.nasa.gov/crystalface/ [Accessed 18 January 2005]. NASA Earth Science Mission: The Cirrus Regional Study of Tropical Anvils and Cirrus Layers - Florida Area Cirrus Experiment.

Koop, T., Luo, B., Tsias, A. and Peter, T. (2000). Water activity as the determinant for homogeneous ice nucleation in aqueous solutions, *Nature* **406**: 611–614.

Korolev, A., Isaac, G. A. and Hallett, J. (2000). Ice particle habits in stratiform clouds, *Quart. J. Roy. Meteor. Soc.* **126**: 2873–2902.

Krueger, A. F. and Fritz, S. (1961). Cellular cloud patterns revealed by TIROS I, *Tellus* **13**: 1–7.

Langley ASDC (2003). SAGE II version 6.20 processing species readme file [online], *Available from*: http://www-sage2.larc.nasa.gov [Accessed 26 March 2005]. Langley Atmospheric Sciences Data Center, Hampton, Virginia, USA.

Lawson, R. P., Heymsfield, A. J., Aulenbach, S. M. and Jensen, E. J. (1998). Shapes, sizes and light scattering properties of ice crystals in cirrus and a persistent contrail during success, *Geophys. Res. Lett.* **25(9)**: 1331–1334.

Lee, K.-M. et al. (2003). Properties of polar stratospheric clouds observed by ilas in early 1997, *J. Geophys. Res.* **108(D7)**: 4228–4244. doi:10.11029/2002JD002854.

Liao, X., Rossow, W. B. and Rind, D. (1995 (1)). Comparison between SAGE II and ISCCP high-level clouds 1. global and zonal mean cloud amounts, *J. Geophys. Res.* **100**: 1121–1135.

Liao, X., Rossow, W. B. and Rind, D. (1995 (2)). Comparison between SAGE II and ISCCP high-level clouds 2. locating cloud tops, *J. Geophys. Res.* **100**: 1137–1147.

Liou, K. N. (1986). Influence of cirrus clouds on weather and climate processes:a global perspective, *Mon. Wea. Rev.* **114**: 1167–1199.

Liu, Q., Simmer, C. and Ruprecht, E. (1996). Three-dimensional radiative transfer effects of clouds in the microwave spectral range, *J. Geophys. Res.* **101(D2)**: 4289–4298.

Luo, B. P. et al. (2003). Dehydration potential of ultrathin clouds at the tropical tropopause, *Geophys. Res. Lett.* **30(11)**: 1557. doi:10.1029:2002GL016737.

Lynch, D. K. (1996). Cirrus clouds: Their role in climate and global change, *Acta Astronautica* **38**: 859–863.

Lynch, D. K. and Sassen, K. (2002). Subvisual cirrus, *in* D. K. Lynch, K. Sassen, D. O. Starr and G. Stephens (eds), *Cirrus*, Oxford University Press, Inc., New York, pp. 256–264.

Marks, C. J. and Rodgers, C. D. (1993). A retrieval method for atmospheric composition from limb emission measurements, *J. Geophys. Res.* **98**: 14939–14953.

Massie, S., Gettelman, A. and Randel, W. (2002). Distribution of tropical cirrus in relation to convection, *J. Geophys. Res.* **107(D21)**: 4591. doi:10.1029/2001JD001293.

Massie, S., Lowe, P., Tie, X., Hervig, M., Thomas, G. and III, J. R. (2000). Effect of the 1997 el nino on the distribution of upper tropospheric cirrus, *J. Geophys. Res.* **105**: 22,725–22,741.

Massie, S., Randel, W. and Wu, F. (2003). Halogen Occultation Experiment and Stratoshperic Aerosol and Gas Experiment II observations of tropopause cirrus and aerosol during the 1990s, *J. Geophys. Res.* **108**: 4222–4233.

Matsumoto, M. and Nishimura, T. (1998). Mersenne Twister. a 623-dimensionally equidistributed uniform pseudo-random number generator, *ACM Transactions on Modeling and Computer Simulation* **8**: 3–30.

McFarquhar, G. M. and Heymsfield, A. J. (1996). Microphysical characteristics of three anvils sampled during the Central Equatorial Pacific Experiment, *J. Atmos. Sci.* **53**: 2401–2423.

McFarquhar, G. M. and Heymsfield, A. J. (1997). Parameterization of tropical cirrus ice crystal size distributions and implications for radiative transfer: Results from CEPEX, *J. Atmos. Sci.* **54**: 2187–2200.

McFarquhar, G. M., Heymsfield, A. J., Spinhirne, J. D. and Hart, B. (2000). Thin and subvisual tropopause tropical cirrus: observations and radiative impacts, *J. Atmos. Sci.* **57**: 1841–1853.

Mergenthaler, J. L., Roche, A. E., Kumer, J. B. and Ely, G. A. (1999). Cryogenic Limb Array Etalon Spectrometer observations of tropical cirrus, *J. Geophys. Res.* **104**: 22,183–22,194.

Michael, S. (2005). Three dimensional imagery's hologram production lab [online], *Available from*: http://www.3dimagery.com [Accessed 10 January 2005].

Miller, S. D. and Stephens, G. L. (2001). Cloudsat instrument requirements as determined from ECMWF forecasts of global cloudiness, *J. Geophys. Res.* **106**: 713–733.

Milz, M. et al. (2004). Feasibility of measurements of water vapor and ice clouds in the tropical UT/LS region with MIPAS/Envisat, *Advances in Space Research* **34**: 815–819.

Minnis, P. (2002). Satellite remote sensing of cirrus, *in* D. K. Lynch, K. Sassen, D. O. Starr and G. Stephens (eds), *Cirrus*, Oxford University Press, Inc., New York, pp. 147–167.

Mishchenko, M. I. (1991). Light scattering by randomly oriented axially symmetric particles, *J. Opt. Soc. Am.* **8**: 871–882.

Mitchell, D. L. and Arnott, W. P. (1994). A model predicting the evolution of ice particle size spectra and radiative properties of cirrus clouds. part II: Dependence of absorption and extinction on ice crystal morphology, *J. Atmos. Sci.* **51**: 817–832.

Murphy, D. M., Thomson, D. S. and Mahoney, N. J. (1998). In situ measurements of organics, meteoric material, mercury, and other elements in aerosols at 5 to 19 kilometers, *Science* **282**: 1664–1669.

Murtagh, F. and Heck, A. (1987). *Multivariate Data Analysis*, D. Reidel Publishing Company, Dordrecht, Holland.

Nakajima, T. and King, M. D. (1990). Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. part I: Theory., *J. Atmos. Sci* **47**(15): 1878–1893.

Norton, R. H. and Beer, R. (1976). New apodizing functions for Fourier spectrometry, *J. Opt. Soc. Am.* **66**: 259–264.

Oikarinen, L., Sihvola, E. and Kyrölä, E. (1999). Multiple scattering radiance in limb-viewing geometry, *J. Geophys. Res.* **104(D24)**: 31,261–31,274.

Plass, G. N. and Kattawar, G. W. (1968). Monte carlo calculations of light scattering from clouds, *Appl. Opt.* **7**: 415–419.

Platnick, S. et al. (2003). The MODIS cloud products: Algorithms and examples from Terra, *IEEE Trans. Geosci. Remote Sens.* **41**: 459–473.

Platt, C. M. R. and Dilley, A. C. (1984). Determination of cirrus particle single-scattering phase function from lidar and solar radiometric data, *Appl. Opt.* **23**: 380–386.

Potter, B. E. and Holton, J. R. (1995). The role of monsoon convection in the dehydration of the lower tropical stratosphere, *J. Atmos. Sci.* **52**: 1034–1050.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1992). *NUMERICAL RECIPES in C: The Art of Scientific Computing*, second edn, Cambridge University Press.

Pruppacher, H. R. and Klett, J. D. (1978). *Microphysics of clouds and precipitation*, Reidel, Dordrecht, Holland.

Queney, P. (1948). The problem of airflow over mountains: A summary of theoretical studies, *Bulletin of the Am. Meteor. Soc.* **29**: 16–26.

Rajeevan, M. and Srinivasan, J. (2000). Net cloud radiative forcing at the top of the atmosphere in the asian monsoon region, *J. Climate* **13**: 650–657.

Ramanathan, V. and Collins, W. (1991). Thermodynamic regulations of ocean warming by cirrus clouds deduced from observations of the 1987 El Niño, *Nature* **351**: 27–32.

Ramanathan, V. et al. (1989). Cloud-radiative forcing and climate: Results from the Earth Radiation Budget Experiment, *Science* **243**: 57–63.

Randall, D., Khairoutdinov, M., Arakawa, A. and Grabowski, W. (2003). Breaking the cloud parameterization deadlock, *Bulletin of the Am. Meteor. Soc.* **84**: 1547–1564.

Raschke, E. et al. (1998). Cloud-radiation studies during the European Cloud and Radiation Experiment (EUCREX), *SURV. GEOPHYS.* **19**: 89–138.

Remedios, J. J. (1999). Extreme atmospheric constituent profiles for mipas, *Proceedings of the European Symposium on Atmospheric Measurements from Space*, Vol. 2, ESTEC, Netherlands, pp. 779–783.

Remsberg, E. E. and Deaver, L. E. (2005). Nasa HALogen occultation experiment project [online], *Available from*: http://haloedata.larc.nasa.gov/home/ [Accessed 2 April 2005].

Ridolfi, M. et al. (2000). Optimised forward model and retrieval scheme for MIPAS near–real–time data processing, *Appl. Opt.* **39**: 1323–1340.

Rodgers, C. D. (1976). Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation, *Geophys. Space. Phys.* **14**: 609–624.

Rodgers, C. D. (1996). Information content and optimisation of high spectral resolution measurements, *in* P. B. Hays and J. Wang (eds), *SPIE*, Vol. 2830, pp. 136–147. *Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research II.*

Rodgers, C. D. (2000). *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific Publishing.

Rodgers, C. D. and Connor, B. J. (2003). Intercomparison of remote sounding instruments, *J. Geophys. Res.* **108**(D3): 609–624. doi:10.1029/2002JD002299.

Rogers, R. R. and Yau, M. K. (1989). *A Short Course in Cloud Physics*, third edn, Pergamon Press, Oxford, UK.

Rossow, W. B. (1989). Measuring cloud properties from space: A review, *J. Climate* **2**: 201–213.

Rossow, W. B. and Schiffer, R. A. (1991). ISCCP cloud data products, *Bulletin of the Am. Meteor. Soc.* **72**: 2–20.

Rossow, W. B. and Schiffer, R. A. (1999). Advances in understanding clouds from ISCCP, *Bulletin of the Am. Meteor. Soc.* **80**: 2261–2287.

Rossow, W. B. et al. (1985). ISCCP cloud algorithm intercomparison, *J. Climate Appl. Meteor.* **24**: 877–903.

Rossow, W. B., Garder, L. C., Lu, P.-J. and Walker, A. (1993). Cloud detection using satellite measurements of infrared and visible radiances for ISCCP, *J. Climate* **6**: 2341–2369.

Rothman, L. S. et al. (2003). The hitran molecular spectroscopic database:edition of 2000 including updates through 2001, *J. Quant. Spectrosc. Radiat. Transfer.* **82**: 5–44.

Sandor, B. J., Jensen, E. J., Stone, E. M., Read, W. G., Waters, J. W. and mergenthaler, J. L. (2000). Upper tropospheric humidity and thin cirrus, *Geophys. Res. Lett.* **27(17)**: 2645–2648.

Santacesaria, V. et al. (2003). Clouds at the tropical tropopause: A case study during the ape-theseo campaign over the western indian ocean, *J. Geophys. Res.* **108(D2)**: 4444. doi:10.1029:2002JD002166.

Sassen, K. (1991). Corona-producing cirrus cloud properties derived from polarization lidar and photographic analyses, *Appl. Opt.* **30**: 3421–3428.

Sassen, K. (2002). Cirrus clouds: A modern persepective, *in* D. K. Lynch, K. Sassen, D. O. Starr and G. Stephens (eds), *Cirrus*, Oxford University Press, Inc., New York, pp. 147–167.

Sassen, K. and Cho, B. S. (1992). Subvisual-thin cirrus lidar dataset for satellite verification and and climatoligical research, *J. Appl. Meteor.* **31**: 1275–1285.

Saunders, R. W. and Kriebel, K. T. (1988). An improved method for detecting clear sky and cloudy radiances from AVHRR data, *Int. J. Remote Sensing* **9**(1): 123–150.

Schneider, S. H. (ed.) (1996). *Encyclopedia of Climate and Weather*, Vol. 2, Oxford University Press, New York, pp. 152–154.

Schumann, U. (1996). On conditions for contrail formation from aircraft exhausts, *Meteor. Z.* **5**: 4–23.

Scorer, R. S. (1978). *Environmental Aerodynamics*, Ellis Horwood, Chichester, UK.

Sherwood, S. C. and Dessler, A. E. (2000). On the control of stratospheric humidity, *Geophys. Res. Lett.* **27(16)**: 2513–2516.

Sherwood, S. C., Chae, J.-H., Minnis, P. and McGill, M. (2004). Underestimation of deep convective cloud tops by thermal imagery, *Geophys. Res. Lett.* **31**: L11102. doi:10.1029/2004GL019699.

Shupe, M. D. and Intrieri, J. M. (2004). Cloud radiative forcing of the arctic surface: The influence of cloud properties, surface albedo, and solar zenith angle, *J. Climate* **17**: 616–628.

Spang, R., Eidmann, G., Riese, M., Offerman, D. and Preusse, P. (2002). CRISTA observations of cirrus clouds around the tropopause, *J. Geophys. Res.* **107(D23)**: 8174–8184.

Spang, R., Remedios, J. J. and Barkley, M. P. (2004). Colour indices for the detection and differentiation of cloud types in infra-red limb emission spectra, *Advances in Space Research* **33**: 1041–1047.

Stamnes, K., Tsay, S. C., Wiscombe, W. and Jayaweera, K. (1988). Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media, *Applied Optics* **27**: 2502–2509.

Stephens, G. L. (1980). Radiative properties of cirrus clouds in the infrared region, *J. Atmos. Sci.* **37**: 435–446.

Stephens, G. L., Tsay, S.-C., Stackhouse Jr., P. W. and Flatau, P. J. (1990). The relevance of the microphysical and radiative properties of cirrus clouds to climate and climate feedback, *J. Atmos. Sci.* **47**: 1742–1753.

Sunil Kumar, S. V., Parameswaran, K. and Krisha Murthy, B. V. (2003). Lidar observations of cirrus cloud near the tropical tropopause: General features, *Atmos. Res.* **66**: 203–227.

Takano, Y. and Liou, K. N. (1989). Solar radiative-transfer in cirrus clouds .1. single-scattering and optical-properties of hexagonal ice crystals, *J. Atmos. Sci.* **46**: 3–19.

Thomason, L. W. (2003). SAGE II data distribution: cloud occurrence [online], *Available from*: http://www-sage2.larc.nasa.gov [Accessed 26 January 2005]. NASA Langley Research Center - The Stratospheric Aerosol and Gas Experiment II.

Thomason, L. W. (2004). SAGE II Version 6.20 data distribution [online], *Available from*: ftp://ftp-rab.larc.nasa.gov/pub/sage2/v6.20 [Accessed 28 March 2005]. NASA Langley Research Center.

Thomason, L. W. (2005). SAGE III data distribution (l2 monthly cloud presence data) [online], *Available from*: http://www-sage3.larc.nasa.gov/data/ [Accessed 28 March 2005]. NASA Langley Research Center.

von Clarmann, T. et al. (2002). Retrieval of temperature and tangent altitude pointing from limb emission spectra recorded from space by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), *J. Geophys. Res.* art. no. 4746.

Wang, P.-H., Minnis, P., McCormick, M. P., Kent, G. S. and Skeens, K. M. (1996). A 6-year climatology of cloud occurrence frequency from Stratospheric Aerosol and Gas Experiment II observations (1985-1990), *J. Geophys. Res.* **101**: 407–429.

Watts, P. D. (1995). Potential use of along track scanning radiometer data for cloud parameter retrieval, *Proc. Spie. Int. Soc. Opt. Eng.* (2578): 30–45.

Watts, P. D., Mutlow, C. T., Baran, A. J. and Zavody, A. M. (1998). Study on cloud properties derived from METEOSAT second generation observations, *Technical Report 97/181*, EUMETSAT ITT. Final Report.

Weickmann, H. (1947). *Die Eishpase in der Atmosphare*, Royal Aircraft Establishment, Farnborough, UK.

Weinman, J. A. and Davies, R. (1978). Thermal microwave radiances from horizontally finite clouds of hydrometeors, *J. Geophys. Res.* **83**: 3099–3107.

Weinstock, E. M., Hinstock, E. J. et al. (1994). New fast-response photofragment fluorescence hygrometer for use on the NASA ER-2 and the perseus remotely piloted aircraft, *Rev. SCi. Instrum.* **65**: 3544–3554.

Whitelaw, T. A. (1996). *Introduction to Linear Algebra*, 2nd edn, Blackie Academic and Professional (Chapman and Hall), Glasgow, UK.

Whiteway, J. et al. (2004). Anatomy of cirrus clouds: Results from the Emerald airborne campaigns, *Geophys. Res. Lett.* **31**: L24102. doi:10.1029/2004GL021201.

Wielicki, B. A. et al. (2002). Evidence for large decadal variability in the tropical mean radiative energy budget, *Science* **295**: 841–844.

Winker, D. M. and Trepte, C. R. (1998). Laminar cirrus observed near the tropical tropopause by LITE, *Geophys. Res. Lett.* **25**: 3351–3354.

WMO (1975). *International Cloud Atlas: Manual on the Observation of clouds and other Meteors*, Vol. 1 of *WMO - No. 407*, World Meteorlgical Organization, Geneva, Switzerland.

Wofsy, S., Michelsen, H. and McCormick, M. (2003). SAGE III cloud product algorithm theoretical basis document, *Available from*: http://eospso.gsfc.nasa.gov/eos_homepage/for_scientists/atbd/ [Accessed 26 March 2005]. The SAGE III ATBD Team, Langley Atmospheric Sciences Data Center, Hampton, Virginia, USA.

Wong, T. (2002). The Earth Radiation Budget Experiment (ERBE) [online], *Available from*: http://asd-www.larc.nasa.gov/erbe/ [Accessed 10 January 2005]. NASA Homepage.

Woodbury, G. E. and McCormick, M. P. (1983). Global distributions of cirrus clouds determined from SAGE data, *Geophys. Res. Let.* **10**: 1180–1183.

Woodbury, G. E. and McCormick, M. P. (1986). Zonal and geographical distributions of cirrus clouds determined from SAGE data, *J. Geophys. Res.* **91**: 2775–2785.

Wylie, D. (2004). 11 years of HIRS data from 1989-2000 [online], *Available from*: ftp://ftp.ssec.wisc.edu/wylie/ [Accessed 25 January 2005]. University of Wisconsin, USA.

Wylie, D. P. and Menzel, W. P. (1989). Two years of cloud cover statistics using VAS, *J. Climate* **2**: 380–392.

Wylie, D. P. and Menzel, W. P. (1999). Eight years of global cirrus cloud statistics using HIRS, *J. Climate* **12**: 170–184.

Wylie, D. P. and Wang, P.-H. (1997). Comparsion of cloud frequency data from the high-resolution infrared radiometer sounder and the Stratospheric Aerosol and Gas Experiment II, *J. Geophys. Res.* **102**: 29,893–29,900.

Wylie, D. P., Menzel, W. P., Woolf, H. M. and Strabala, K. L. (1994). Four years of global cirrus cloud statistics using HIRS, *J. Climate* **9**: 1972–1986.

Xu, K. M., Wielicki, B. A. and Wong, T. (2005). Statistical analyses of satellite cloud object data to study climate sensitivities, *16th Symposium on Global Change and Climate Variations*, American Meteorological Society, San Diego, California.

Yang, P. and Baum, B. A. (2003). Satellite remote sensing; cloud properties, *in* J. R. Holton, J. A. Curry and J. A. Pyle (eds), *Encyclopedia of Atmospheric Sciences*, Vol. 5, Elsevier Science Ltd., pp. 1956–1965.

Yang, P. et al. (2001). Radiative properties of cirrus clouds in the infrared (8–13 μm) spectral region, *J. Quant. Spectrosc. Radiat. Transfer.* **70**: 473–504.

Yang, P., Liou, K. N., Wyser, K. and Mitchell, D. (2000). Parameterization of the scattering and absorption properties of indivi dual ice crystals, *J. Geophys. Res.* **105**(D4): 4699–4718.

Zhang, M. (2004). Cloud-climate feedback: How much do we know?, *in* X. Zhu et al. (eds), *World Scientific Series on Meteorology of East Asia*, Vol. 3, World Scientific Publishing.

Zhang, Y., Rossow, W. B., Lacis, A. A., Oinas, V. and Mischenko, M. I. (2004). Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data, *J. Geophys. Res.* **109**: D19105. DOI:10.1029/2003JD004457.