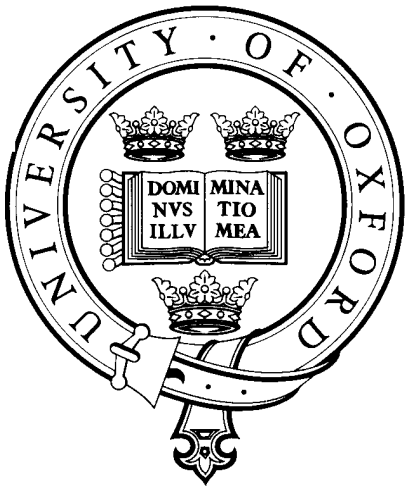
Aerosol Remote Sensing Using AATSR

Andrew Mark Sayer

Supervisors: Don Grainger, Chris Mutlow, Anu Dudhia
Postdoctoral Adviser: Gareth Thomas



First Year Report

Atmospheric, Oceanic and Planetary Physics
Department of Physics
University of Oxford

August 2006

Contents

Abstract

The aim of this document is to describe the work I have been doing over the past year, and show how it will be built upon over the next two years to create a viable doctorate thesis. In this report I provide a brief overview of atmospheric aerosol (Chapter 1) and how we can use probe this aerosol in the atmosphere, focussing on the AATSR instrument in particular (Chapter 2). I then (Chapter 3) discuss the basics of retrieval theory–how we can use measurements of a property of a system to glean information on its state–before describing the results I have obtained to date (Chapter 4). Finally, I outline plans for further work and a rough timetable of how long I anticipate the various areas of study to take (Chapter 5).

I am grateful to Anu Dudhia, Don Grainger, Chris Mutlow and Gareth Thomas for the time they have spent helping me get to grips with the department and the project over the past year. The project studentship is funded by the Natural Environment Research Council (NERC).

Chapter 1

Atmospheric Aerosol

1.1 What is atmospheric aerosol?

Formally defined, aerosols are particles composed of solid or liquid matter, suspended in the atmosphere with dimensions on length scales in the region of 10^{-9} m to 10^{-6} m. This definition encompasses a broad array of different substances with wildly varying shapes, sizes, sources and sinks, and physical and chemical properties. Their large temporal and spatial variability has made global studies difficult, although over the past few decades the importance of aerosols in climate and weather systems has been realised, leading to many attempts–such as the one described in this report–to quantify and classify them.

The study of aerosols began with John Aitken's experiments with dust in the 1880s. He built apparatus to count particles in suspension by expanding and then cooling samples of air. Water condensed on the particles, allowing the resulting drops to be counted with aid of a microscope. Aitken used this apparatus to measure a wide variety of samples, including sea salt produced from air blowing over the ocean. C. T. R. Wilson performed important work from 1897 onwards investigating the growth of water and dust particles, which resulted in the invention of the cloud chamber. From the 1920s to the 1950s work by Kohler, Howell and then Twomey established aerosols as cloud condensation nuclei (CCN) which led to the study of their implications for weather and climate.

It is usual practice to group aerosols by their origins: those produced as a result of naturally-occurring geological processes; those produced by biological processes ('biogenic'); and those produced as a result of human activity ('anthropogenic'). Distinction is also often made between tropospheric and stratospheric aerosol, as these often share different sources and have different radiative effects. A wide variety of literature, including many overviews of particular aspects (such as Andreae and Crutzen [1997], Charlson et al.

[1992], Pueschel [1996] and Schwartz [1996]), surrounds this field. An in-depth treatment of most aspects of atmospheric aerosol–albeit before the era of modern satellites–was given by Twomey [1977].

1.2 Production to removal

1.2.1 Aerosol sources

Definition and quantification of aerosol emission is less clear-cut than that of greenhouse gases or other pollutants. One source may produce particles of a wide variety of sizes or compositions, and so very different properties. Primary aerosols are those substances emitted directly as aerosols, while secondary aerosol forms in the atmosphere from various precursor substances. Combination of aerosol particles may also lead to new particles with different chemical and physical properties from those from which they were made, so it is important to consider not just aerosol sources but how these aerosols interact with each other and the atmosphere during their lifetime. The Intergovernmental Panel for Climate Change (IPCC) have collated and documented estimates for global annual emissions of aerosol and aerosol precusors. This information is summarised in Figure 1.1 and Tables 1.1, 1.2 and 1.3.

Wind-driven sources: mineral dust and sea salt

Mineral dust and sea salt are classified as primary aerosol. As shown in Figure 1.1 (f), mineral dust aerosol is particularly prevalent in tropical and sub-tropical regions where it is kicked up by wind (or human activity) over deserts or vegetation-sparse land. Dust storm frequency has also been linked to climate cycles, such as the El Niño and North Atlantic Oscillation events. Wind blowing over the oceans results in the generation

Aerosol precursor	Northern hemisphere	Southern hemisphere	Global total	Range of estimates
NO_x	32	9	41	-
NH_3	41	13	54	41-70
SO_2	76	12	88	67-130
DMS/H_2S	11.6	13.4	25.0	12-42
VOCs	171	65	236	100-500

Table 1.1: Emissions of aerosol precursors, adapted from IPCC [2001]. Figures indicate the mass of N, C or S in Tg per year as appropriate for the precursor molecule. VOCs are volatile organic compounds. DMS is dimtethyl sulphide, CH_3SCH_3. NO_x, NH_3 and SO_2 are dominated by anthropogenic activities, while DMS/H_2S is primarily biogenic and VOCs are evenly split.

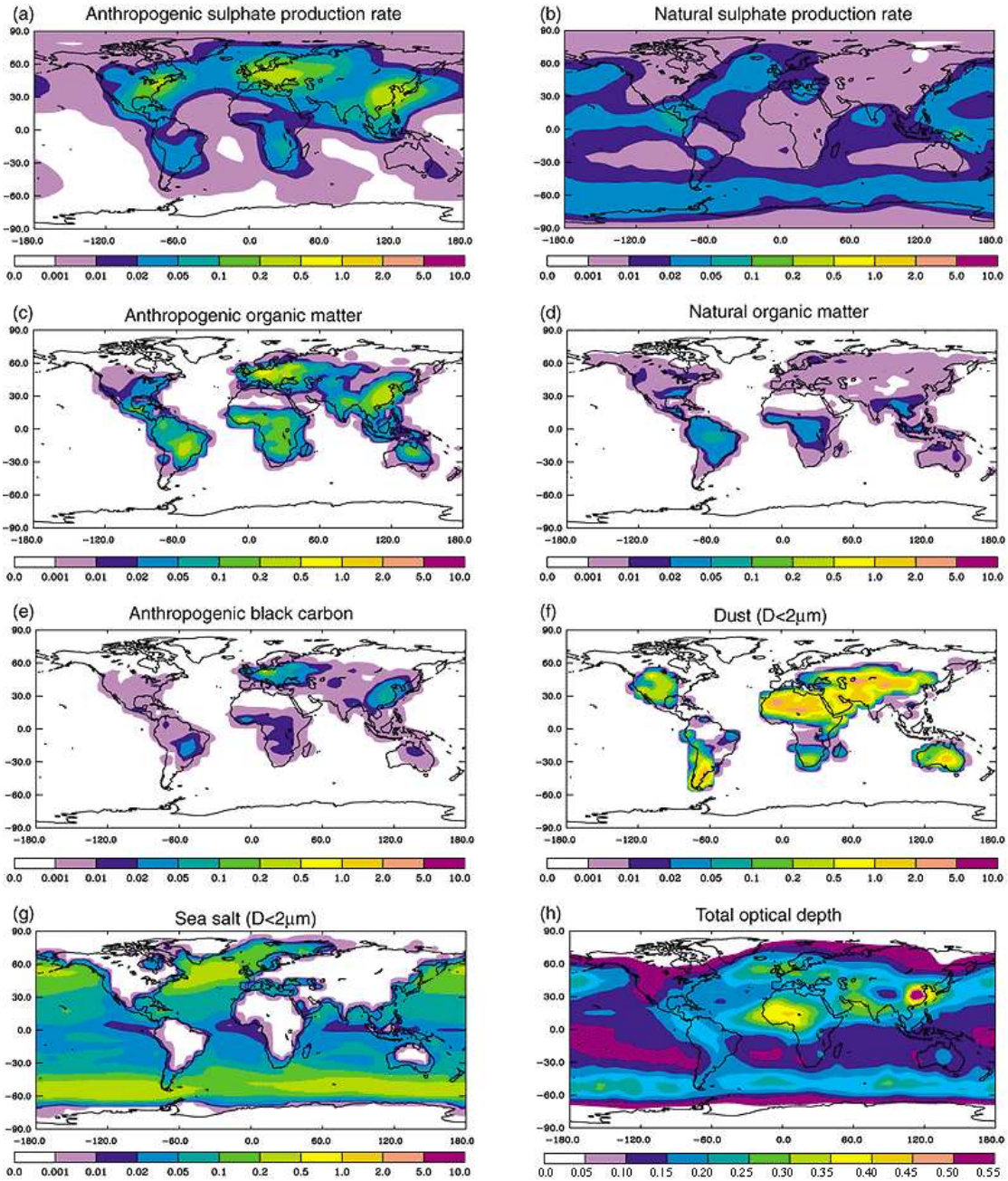


Figure 1.1: Annual average source strength in kg km^{-2} hr^{-1} for the aerosol types considered by the IPCC (a to g) with total aerosol optical depth (h). Taken from IPCC [2001]. Shown are (a) the column average H_2SO_4 production rate from anthropogenic sources, (b) the column average H_2SO_4 production rate from natural sources (DMS and volcanic SO_2), (c) anthropogenic sources of organic matter, (d) natural sources of organic matter, (e) anthropogenic sources of black carbon, (f) dust sources for dust with diameters less than 2 μm, (g) sea salt sources for sea salt with diameters less than 2 μm, and (h) total optical depth according to a model considered by the IPCC.

Aerosol type	Northern hemisphere	Southern hemisphere	Global total	Range of estimates
Carbonaceous aerosols				
Organic matter (0-2 μm):				
Biomass burning	28	26	54	45-80
Fossil fuel	28	0.4	28	10-30
Biogenic (> 1 μm)	-	-	56	0-90
Black carbon (0-2 μm):				
Biomass burning	2.9	2.7	5.7	5-9
Fossil fuel	6.5	0.1	6.6	6-8
Aircraft	0.005	0.0004	0.006	
Industrial dust, etc. > 1 μm	-	-	100	40-130
Sea salt				
< 1 μm	23	31	54	18-100
1-16 μm	1420	1870	3290	1000-6000
Total	1440	1900	3340	1000-6000
Mineral (soil) dust				
< 1 μm	90	70	110	-
1-2 μm	240	50	290	-
2-20 μm	1470	282	1750	-
Total	1800	349	2150	1000-3000

Table 1.2: Emissions of primary aerosol particles for the year 2000, adapted from IPCC [2001]. Figures are Tg per year.

Aerosol source	Northern hemisphere	Southern hemisphere	Global total	Range of estimates
Sulphate (as $NH_4.H_2SO_4$*):*	145	55	200	107-374
Anthropogenic	106	15	122	69-214
Biogenic	25	32	57	28-118
Volcanic	14	7	21	9-48
Nitrate (as NO_3^-*):*				
Anthropogenic	12.4	1.8	14.2	9.6-19.2
Natural	2.2	1.7	3.9	1.9-7.6
Organic compounds				
Anthropogenic VOC	0.15	0.45	0.6	0.3-1.8
Biogenic VOC	8.2	7.4	16	8-40

Table 1.3: Estimates of emissions of secondary aerosol sources, adapted from IPCC [2001]. Figures are Tg per year.

of sea salt aerosol. Global distribution is shown in Figure 1.1 (g). Dependence on wind speed means these aerosol types show strong temporal variability.

Industrial dust and ash

The Industrial Revolution marked the beginning of the time when humans put significant amounts of dust and ash into the atmosphere. This is as the result of emissions from power stations, waste combustion, transportation, and factories. Due to the visible pollution which became commonplace in urban areas, tighter industrial regulations over the years have meant that dust emissions have decreased significantly in developed countries. Industrial dust is a primary anthropogenic aerosol. Volcanic ash, a primary geological aerosol, is ejected into the atmosphere by volcanic eruptions. The most powerful result in this ash reaching the stratosphere.

Organic carbon

Organic carbon aerosol is carbon emitted in volatile organic compounds (VOCs), distributed as shown in Figure 1.1 (c) and (d). Some is directly emitted into the atmosphere, either as products from incomplete combustion from industry or biomass burning or as compounds (e.g. terpenes) emitted by plants, while some secondary organic aerosol (SOA) is produced by modification (typically oxidation) of these VOCs. As a result, organic carbon can be both a primary and a secondary, as well as an anthropogenic or biogenic, aerosol. Many of these compounds contain polar groups (such as hydroxyl or carboxylic acids) meaning they are water soluble, and so readily take part in cloud formation, as discussed in Section 1.2.5.

An important point to note is that although organic carbon is produced both naturally and anthropogenically, its oxidation depends on compounds such as NO_x, O_3 and the hydroxy radical $OH^{\cdot}$. Levels of these, and hence the rate of SOA production, are strongly influenced by human activities creating these pollutants.

Black carbon

Black carbon consists of soot and tarry molecules, and is so called because it is strongly light absorbing. Quantification of carbon aerosol is important because of the very different effects of black carbon (primarily absorbing) and organic carbon (primarily scattering). A primary aerosol, black carbon tends to be insoluble and is produced as the result of incomplete combustion. Its distribution is shown in Figure 1.1 (e).

Other primary biogenic aerosol

Other sources of primary biogenic aerosol include airborne virues and bacteria, as well as pollen and spores and other debris (such as leaf fragments and decaying matter). Little is known about precise spatial or temporal distribution of this class of aerosol, but it would be expected to vary significantly with changing land use patterns.

Sulphate

As a secondary aerosol, sulphate is produced from precursor molecules. These precursors may be produced anthropogenically (by industry or transportation, including SO_2, OCS and CS_2) or naturally (such as H_2S or SO_2 from volcanic eruptions). Over the oceans, DMS (dimethyl sulphide, CH_3SCH_3) emitted by plankton is a major source. In the atmosphere, these precursors are oxidised and combine with water vapour to produce aerosols of mixed H_2SO_4 and H_2O. Sulphate production is shown in Figure 1.1 (a) and (b), although it should be emphasised that there are large uncertainties in emissions (see Table 1.1). DMS emittted by marine plankton is an important source of sulphur although quantifying release is difficult as it is problematic to measure directly, and there have been problems correlating DMS to proxies such as chlorophyll content of a body of water (Andreae and Crutzen [1997]).

The complexity of reaction pathways involved in the generation of sulphate aerosol means that the precise proportion of precursors which end up as H_2SO_4 is uncertain. Being highly soluble, sulphate and reactive intermediates are easily washed out by rain (with estimated lifetimes of 4-7 days in the troposphere) so vertical moisture profiles are important in calculating the sulphate burden accurately. Volcanically-produced SO_2 and OCS from aviation are thought to be amongst the most efficient sulphate precursor sources (in terms of mass of sulphate from mass of precursor). The dry, stable stratosphere in which they are emitted (from the strongest volcanic eruptions in the case of SO_2) means that little rainout or other loss occurs, and the aerosols themselves are able to persist for longer (Pueschel [1996]). Reaction of SO_2 with mineral aerosol particles or oxidation to H_2SO_4 in sea-salt aerosols also reduces the potential radiative forcing of sulphate, as these larger particles have shorter lifetimes and lower scattering efficiencies than pure H_2SO_4 aerosol. Levels of NH_3 are also important, as acid-base reactions forming $NH_4^+.HSO_4^-$ change the refractive index of the aerosol.

Nitrate

Atmospheric nitrate aerosol is linked to NH_3 and H_2SO_4 abundance due to aforementioned neutralisation reactions. The global mean ratio is thought to be close to 1, but in regions where there is surplus ammonia its

oxidisation to nitrate aerosol is possible. Anthropogenic NO_x is also important in polluted areas, although research collated by the IPCC [2001] suggests the total forcing due to anthropogenic nitrate is small.

Like sulphate, nitrate chemistry involves many complex redox cycles. Nitrate aerosols may be effectively removed by formation of the less reactive HNO_3, either by water reacting with N_2O_5 or via the reaction of HCl with $ClONO_2$. Both these reactions are prevalent in the stratosphere, particularly on ice-crystal polar stratospheric clouds (PSCs) in the Antarctic winter, where the redox processes are linked to ozone destruction (Pueschel [1996]).

1.2.2 Particle growth and size distributions

Direct emissions of aerosols and their precursors tend to be of the order of size of molecules, or about 10^{-10} m. These molecular clusters quickly coagulate to give larger particles on length scales of 10^{-9} m, which represent the smallest particles measured by counters. Aerosols of such size are said to be in *nucleation* mode in the classification scheme of Junge [1955].

Aerosols of around 10^{-8} m to 10^{-7} m normally grow less rapidly, and are known as *accumulation* mode aerosol. Particles of this size are still fairly buoyant due to Brownian motion, although the largest may be subject to some slow gravitational settling. They are comparatively long-lived because of their buoyancy and slow growth.

Larger aerosol particles of size scales of 10^{-6} m (1 μm) and greater are often called *coarse* mode particles. Their lifetimes tend to be comparatively short, for two main reasons:

1. Particles of this size have sufficient mass to be affected by gravitational settling. A spherical particle of diameter 1 μm will fall at a speed of 1 mm every 5 seconds or so, with falling speed increasing quadratically with radius.

2. The largest particles act as cloud condensation nuclei (CCN) and so will eventually fall as rain or snow (typically as droplets of size 10^{-3} m, i.e. 1 mm).

The mechanics and statistics of aerosol coagulation have been described in-depth by Twomey [1977]. It suffices to say that when considering samples of aerosol we are most interested in being able to define a size distribution which accurately represents them. Two well-used distributions are outlined below. It is also important to note, however, that size is not the only property we are interested in as a particle's refractive index (which determines its interaction with light) is determined by its chemical composition,

which may alter as the particle grows by both gaining mass from different aerosol sources and through chemical reactions within the particle itself. These changes are known as ageing processes.

The Junge distribution

Junge's original work in 1955 showed that the size distributions of many natural samples of aerosol could be well modelled by inverse power distributions, and as a result this is also sometimes called the Junge distribution. With radius r and for some constants α and c, these distributions take the form:

$$r^{\alpha}\left(\frac{\mathrm{d}N}{\mathrm{dlog}r}\right) = c \tag{1.1}$$

This is solved to give a cumulative distribution for the number density n of particles of size greater than some radius r, which is proportional to $r^{-(\alpha+1)}$. Plots of n against r with logarithmic axes yield straight lines, as in Figure 1.2. Junge's original work placed the value of α at approximately 3, still considered a reasonable estimate for many aerosol samples. When integrating this distribution to obtain total particle number the integral diverges to infinity for all values of α. In practice some minimum radius is chosen as a cutoff (though the particle count becomes sensitive to this minimum value). As a result other measurements, such as effective radius (discussed in Section 3.1.1) or total volume, are more useful when describing the population than total particle number or number density.

Lognormal distributions

More recent studies such as Davies [1974] and Hess et al. [1998] have suggested the lognormal distribution to be most appropriate for atmospheric aerosol. The aerosol models used in the retrieval scheme presented here consider aerosols to be mixtures of lognormally-distributed nucleation, accumulation and coarse mode particles. For an aerosol with mode radius r_m and total number density N_i the number density $n(r)$ of particles larger than r in a distribution with spread σ is given as follows:

$$n(r) = \frac{N_i}{\sqrt{2\pi}\, r \log \sigma_i \ln 10} \exp\left[-\frac{1}{2}\left(\frac{\log r - \log r_m}{\log \sigma_i}\right)^2\right] \tag{1.2}$$

An example of $n(r)$ for a lognormal distribution is shown in Figure 1.3. In practice minimum and maximum radii are specified when these distributions are used to save on computational time. Although the distributions extend to infinity, the number density very far from the modal radius r_m becomes small enough to be neglected.

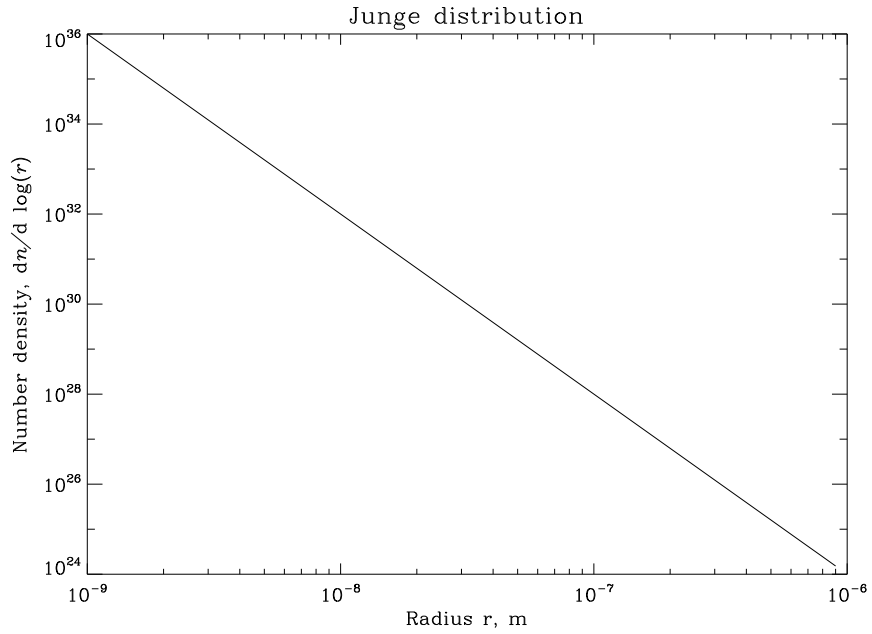


Figure 1.2: Junge distribution number density for radii between 10^{-9} m and 10^{-6} m. For the purposes of this figure α was taken to be 3.0 and c to be 1.0

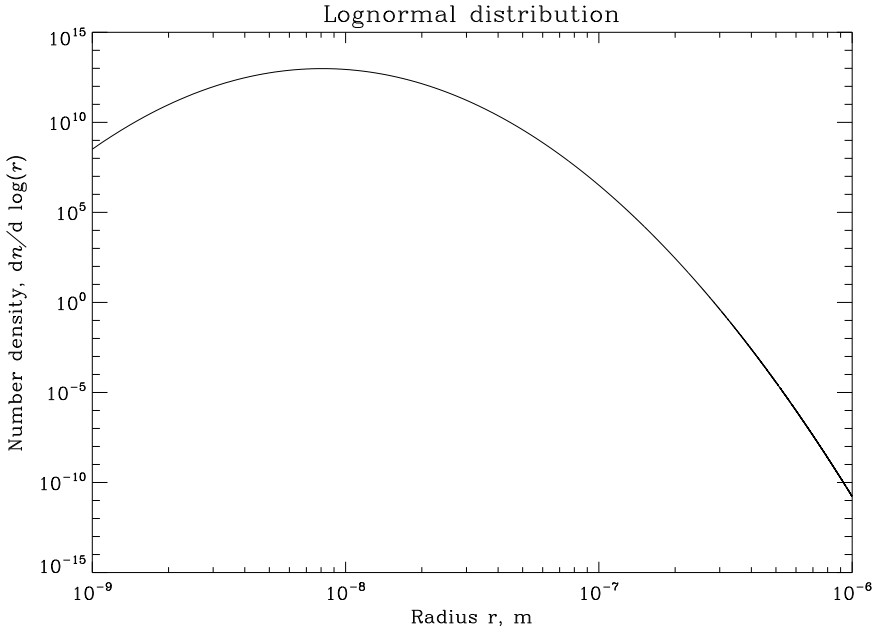


Figure 1.3: Lognormal distribution for radii between 10^{-9} m and 10^{-6} m. This distribution assumes a total number density N_i of 10^6 particles per unit area, a model radius r_m of 10^{-8} m and a standard deviation $\log\sigma_i$ of 0.2.

1.2.3 Vertical transport

Most aerosol is produced at fairly low altitudes and remains confined to the troposphere, due to the lack of large-scale vertical mass exchange between the troposphere and stratosphere. Large particles in particular rarely make their way from troposphere to stratosphere because they are affected by gravitational settling.

The most violent volcanic eruptions are able to inject aerosol and precursor gases directly into the stratosphere (Pueschel [1996]). Although the ash quickly settles, other aerosol (typically sulphate) is able to persist for longer timescales (up to a year) than tropospheric aerosol (days to weeks) because of the high stability of the stratosphere. The Mount Pinatubo eruption of 1991 is estimated to have ejected 14-20 Tg of SO_2 directly, forming a larger quantity of mixed H_2SO_4/H_2O droplets over a timescale of months which persisted for several years and had notable effects on climate during that period. Aircraft exhaust also provides a source of stratospheric aerosol.

Even in periods of volcanic inactivity there is thought to be a compact background stratospheric aerosol later. It is not thought to show strong temporal variability or be dense enough to affect climate: the main climatic perturbations of stratospheric aerosol come from periodic volcanic eruptions.

1.2.4 Horizontal transport

As small particles aerosols are easily transported by winds, and thus the aerosol present in a region may well come from a distant source. This is important in pollutant analysis as the impact of the pollutant will not be confined to the emitting source. Saharan dust is also thought to contribute to fertilisation of the Amazon basin, as transport across the Atlantic Ocean has been observed (Formenti et al. [2001]). Measurement of aerosol transport is difficult as source regions may be large and inhomogeneous, and the distances involved are often great (on the order of 10^3 km). Nevertheless, various methods have been developed to attempt to qualify and quantify the long-range transport of aerosols.

One commonly-used method (Borbély-Kiss et al. [2004], Formenti et al. [2001], Rahn and Lowenthal [1984]) is to calculate the abundances of several trace elements in the aerosol sample, and then attempt to map this data to a particular source or source region. Possible source locations may be narrowed down using wind back-trajectory analysis. Elemental analysis may be performed with proton-induced X-ray emission (PIXE) methods, and commonly-profiled elements include Ti, Ca, Al, Fe and Si for Saharan dust samples and As, Sb, Se, Zn, In, noncrustal Mn and noncrustal V for Arctic aerosol. This method is not considered to be trouble free, one criticism from Thurston et al. [1985] being that as it measures total mass it cannot

account for mass observed in one region coming from different sources, which may alter the apparant relative abundances of the elements analysed.

Transport may also be inferred by monitoring the temporal evolution of large aerosol masses. Light scattering apparatus such as LIDAR and nephelometers are often used in conjunction with satellite data and wind fields to track the movement of aerosol (especially dust clouds). Such methods also often provide some detail about size distribution and vertical profiles and are frequently used in the literature (Dulac et al. [2001], Hamonou and Chazette [1998]), sometimes combined with elemental analysis (Formenti et al. [2001]).

It is also possible to use satellite data to track the movement of aerosol, and quantification of Saharan dust transport will be an eventual goal of this project. Borbély-Kiss et al. [2004] have already made use of Total Ozone Mapping Spectrometer (TOMS) aerosol data for this purpose.

1.2.5 Aerosol sinks

In common with the wide variety of processes leading to the generation of atmospheric aerosol, there exists a wide variety of processes removing of aerosol from the atmosphere. Wet removal processes occur when the aerosol is removed in precipitation (water, fog or ice) while dry removal processes remove aerosol without the involvement of precipitation. Cloud formation (sometimes grouped with general wet removal processes) is another major aerosol removal process and eventually provides the mechanisms for wet removal processes to occur. A mathematical treatment is given by Twomey [1977].

Wet removal processes

Aerosol particles may be rained out by collision with falling raindrops (inertial removal). The efficiency of this process (proportion of particles removed by it) is highest for heavy particles. Removal is also possible through diffusion of aerosols into falling drops, which is most efficient for small raindrops and very small particles (less than 10^{-8} m in size). The strong size dependence of these removal processes means that particles in the region of 10^{-8} m to 10^{-7} m are generally not efficiently removed, although the fact that they are typically among the most numerous particles means that a large total mass is lost. Particles of this size range can also be difficult to collect and count experimentally, hindering studies of their removal. For particles of all sizes the total aerosol amount removed through wet processes does, of course, depend on how much moisture there is.

Dry removal processes

The largest difference between wet and dry removal is that while wet removal can occur throughout the atmosphere wherever there is precipitating water, dry removal relies on the particles being transported near to the surface by turbulence or gravity. The movement of small particles will be dominated by Brownian motion, while that of heavy particles will feel the effects of gravitational settling. Diffusive removal is again only important for very small particles, and inertial only for larger particles, with low rates of dry removal of particles in the range from 10^{-8} m to 1 μm.

Cloud formation and aerosol removal processes

As previously mentioned, large aerosol particles are able to act as cloud condensation nuclei (CCN). The formation of clouds themselves, then, provides a method to remove aerosol from the atmosphere as well as other removal processes associated with cloud. Cloud removal processes tend to be efficient over all aerosol size ranges due to the high surface area of cloud droplets, as well as the rapid condensation of small particles into larger ones which takes place (especially with soluble particles). Clouds are able to form and rain within an hour, which can lead to rapid removal of aerosol.

Nucleation can only occur on aerosol particles larger than a critical radius of approximately $10^{-8}S_c^{-2/3}$ metres (where S_c is the supersaturation in percent). Following nucleation a period of rapid growth leads to the droplet growing to the order of microns within a few seconds, provided supersaturation is maintained. Further growth is slower, though possible through diffusion of aerosol particles (again only effective for small aerosol) or phoretic forces (due to evaporation or condensation of water, or heat fluxes).

Sufficiently large droplets fall out as rain, removing both directly-incorporated aerosol as well as any caught by wet removal processes. As a result, clouds are considered to be the most important means of removing aerosol from the atmosphere, with small particles being caught up diffusively and larger ones serving as CCN. Water-insoluble aerosol is affected less by wet removal and cloud processes than soluble, although insoluble particles may still act as cloud nucleation sites.

Junge's concept of rainout efficiency E allows us to estimate timescales for aerosol residence in the atmosphere. This relates the amount of contaminant per unit volume of water to the corresponding amount in cloud, and is expressed as follows:

$$E = \frac{cW}{P\chi} \tag{1.3}$$

In the above, c is the concentration of the contaminant in water and χ its concentration in cloud, with W the liquid water content of a cloud and P the volume of rain is the water content of a volume P/W of cloud. Assuming E is close to 1 and taking $W = 1/3$ gm^{-3} we see 1 cm of rain gives a cloud volume of 3 m^3, implying a vertical depth of 300 km can be cleansed. Therefore in principle a few millimetres of rain could effectively clear the atmosphere of aerosol. This suggests that the residence time of aerosol in the atmosphere is determined more by the frequency of rain, and the height at which clouds form, than the total amount of rain that falls.

1.3 Effects on climate

1.3.1 Direct and Indirect effects: aerosol radiative forcing

It is widely acknowledged that by scattering light, aerosols exert a non-negligible influence on weather and climate (Charlson et al. [1992], Pueschel [1996], Schwartz [1996]).

- They have the *direct effect* of increasing planetary albedo by scattering incoming shortwave solar radiation. This exerts a net negative forcing on the Earth.

- They have the *indirect effect* of acting as CCN and so increasing cloud cover (as well as cloud lifetime if droplets are small). This also exerts negative forcing by enhancing shortwave albedo, but has positive forcing effects due to increased absorption of infrared radiation and trapping at night (clouds at night result in the Earth retaining more warmth). Perturbing cloud cover also changes one of the many complex feedbacks into weather and climate so may have other, unforseen effects.

The overall negative forcing of aerosols offsets some of the positive forcing of greenhouse gases, and is thought to be from -0.4 Wm^{-2} to -3.0 Wm^{-2}. While greenhouse gases persist for years, aerosol lifetimes are short and profiles show great spatial variability. We have no past record of global aerosol distribution, and as a result understanding of historical aerosol forcing is poor. The IPCC [2001] has collated a large amount of research into the effects of aerosols on climate, although as shown in Figure 1.4 the aerosol contribution to radiative forcing is not well understood. These unknowns mean it is of great importance to obtain an accurate picture of global aerosol distribution.

In most cases, the absorption and emission of infrared radiation by aerosols is thought to be minor. This is because aerosols are mostly found low in the atmosphere where atmospheric temperature is close to surface temperature. The optical depth of aerosols also decreases at longer wavelengths.

Volcanic eruptions can provide exceptions to this rule: the stratospheric ash aerosol produced absorbs infrared and visible radiation and heats the stratosphere. Over the long term, however, scattering of visible light by the more persistent sulphate aerosol predominates.

Figure 1.5 shows transmission of solar radiation at Mauna Loa observatory with times of major volcanic eruptions labelled. There is a clear and sharp decrease in incoming radiation linked with the eruptions which can persist for several years, due to light scattering by the stratospheric aerosols ejected. Indeed, the 1815 eruption of Tambora in Indonesia is thought (Pueschel [1996]) to be responsible for summer snowstorms and crop failures due to the cooling effects of the aerosol it produced.

Forcing effects are also dependent on the location of the forcing substance. An increase in aerosol optical depth (and resultant increase in planetary albedo) has more impact over a dark surface (such as a cloud-free ocean) than over a bright one (cloudy land, ice or desert).

1.3.2 Rayleigh and Mie scattering

Light scattering from atmospheric particles is described by Mie theory, which reduces to a simpler form known as Rayleigh scattering theory when the particles are spherical and small compared to the wavelength of light being scattered. We can define a size parameter x to set the domains of validity of this approximation:

$$x = \frac{2\pi r}{\lambda} \tag{1.4}$$

Here r is the radius of the particle and λ the wavelength of light considered. When x is small (typically less than 0.1) we can apply the Rayleigh scattering equations without appreciable loss of accuracy. In the case of aerosols, this will be of interest for small particles (less than about 0.1 μm in size) and long wavelengths (around 1 μm and greater, i.e. the infrared region of the spectrum). In these cases, the normalised phase function for scattering intensity of unpolarised light as a function of scattering angle θ is given by:

$$p(\theta) = \frac{3}{16\pi}(1+\cos^2\theta) \tag{1.5}$$

This forward-backward symmetrical distribution is plotted in Figure 1.6. The absolute value of the scattering is proportional to the square of the polarisability, hence radius r to the 6^{th} power. Absorption is proportional to volume (r^3) so for smaller particles it will become relatively more important. Absolute scattering intensity is also proportional to λ^{-4}. This explains why, although effective scatters of shortwave solar radiation, aerosols have little influence on longwave infrared radiation.

When the dimensions of the particle are not small compared to the wavelength of light (i.e. $x > 0.1$), full Mie theory must be used to accurately calculate scattering intensity. This is a computationally expensive exercise. In general, as x increases an increasing proportion of incident light is scattered in the forward direction. This is visible in Figure 1.7. There is strong dependence on both wavelength of light and particle size. While Rayleigh scattering intensity follows a single and almost unruffled curve, the Mie patterns visible show clear oscillations with θ, increasing in frequency with increasing x.

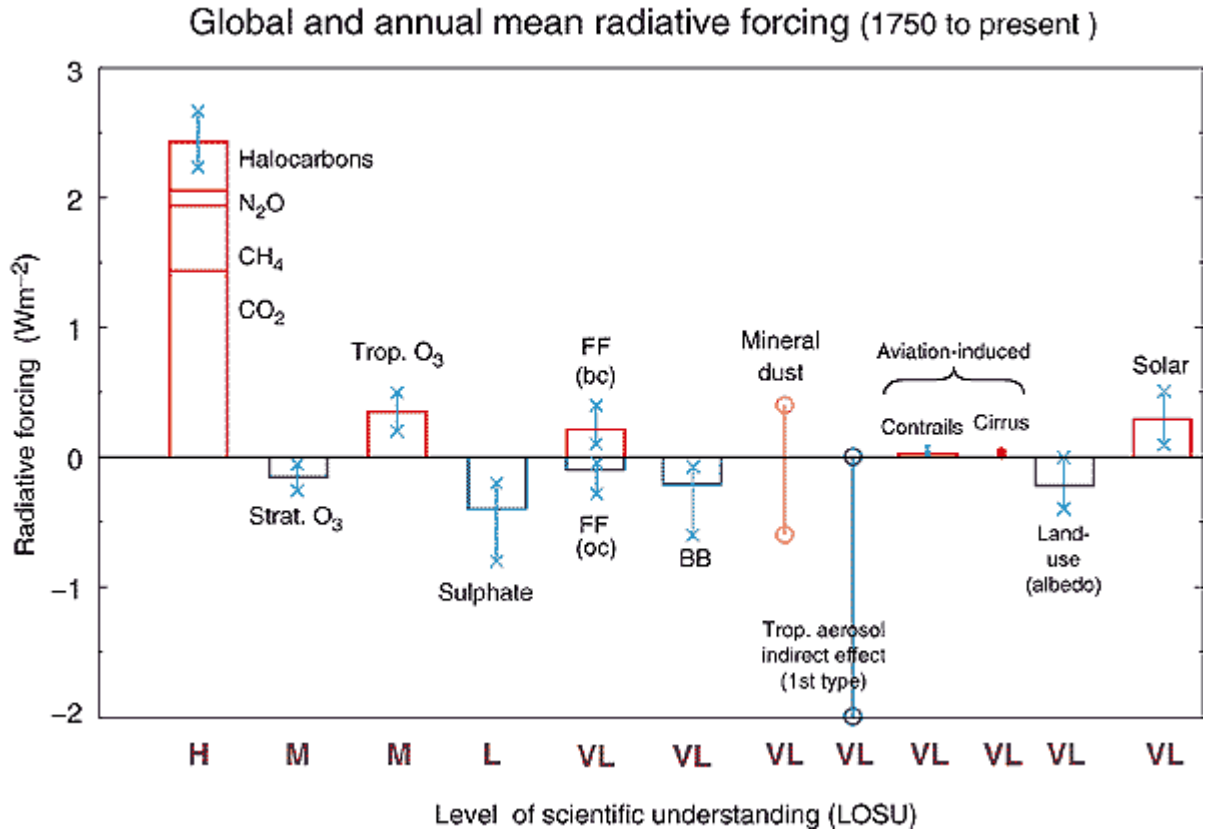


Figure 1.4: Global, annual mean radiative forcings (Wm^{-2}) due to a number of agents for the period from 1750 to the present (about 2000), taken from the third IPCC report IPCC [2001]. The height of the rectangular bar denotes a best estimate value while its absence denotes no best estimate is possible. The vertical line about the rectangular bar with *x* delimiters indicates an estimate of the uncertainty range, guided by the spread in the published values of the forcing and physical understanding. A vertical line without a rectangular bar and with *o* delimiters denotes a forcing for which no central estimate can be given owing to large uncertainties. A level of scientific understanding (LOSU) index is accorded to each forcing, with H, M, L and VL denoting high, medium, low and very low levels, respectively. The well-mixed greenhouse gases are grouped together into a single rectangular bar with the individual mean contributions due to CO_2, CH_4, N_2O, and halocarbons shown. FF denotes fossil fuel burning while BB denotes biomass burning aerosol. Fossil fuel burning is separated into the black carbon (*bc*) and organic carbon (*oc*) components with its separate best estimate and range. It is emphasised that the positive and negative global mean forcings cannot be added up and viewed a priori as providing offsets in terms of the complete global climate impact, and that the forcings show spatial and temporal variability meaning the global annual means described are not a complete picture.

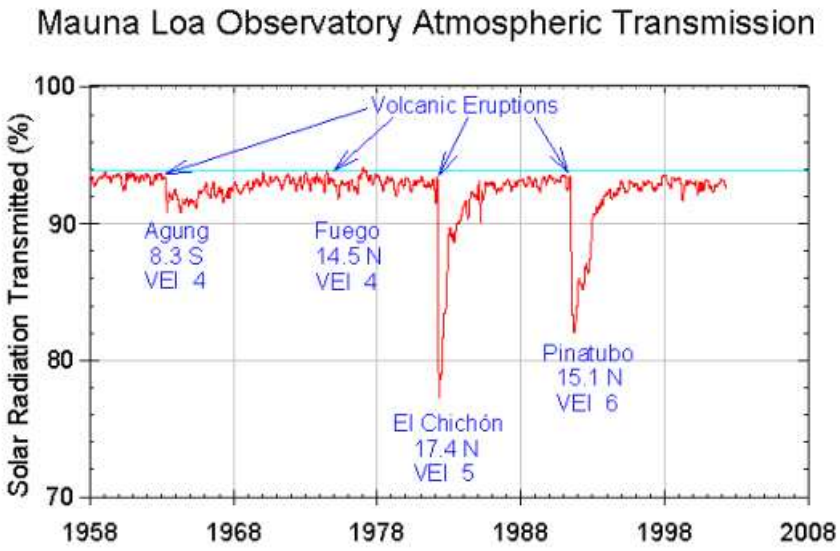


Figure 1.5: Reduction in transmission of solar radiation at Mauna Loa observatory, Hawaii. Incidence of major volcanic eruptions is labelled. Image from the National Oceanic and Atmospheric Administration (NOAA).

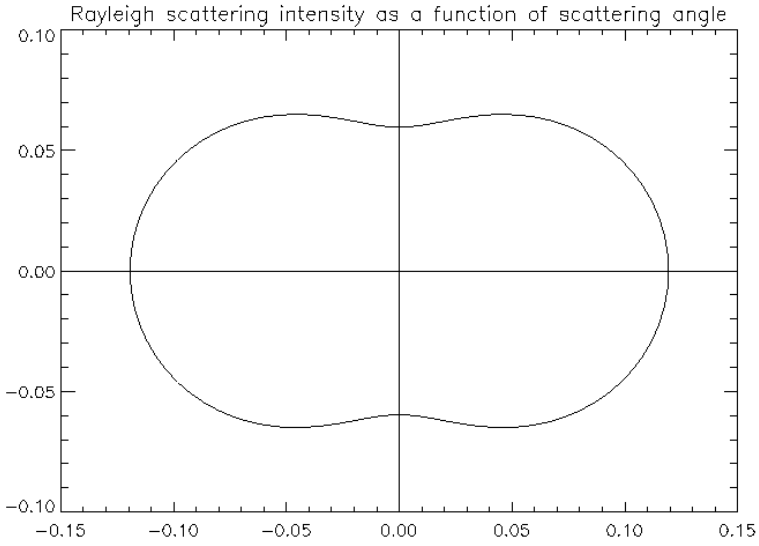


Figure 1.6: Rayleigh scattering intensity for unpolarised light as a function of scattering angle θ. θ increases anticlockwise from the origin, with incident light in the direction of the x axis.

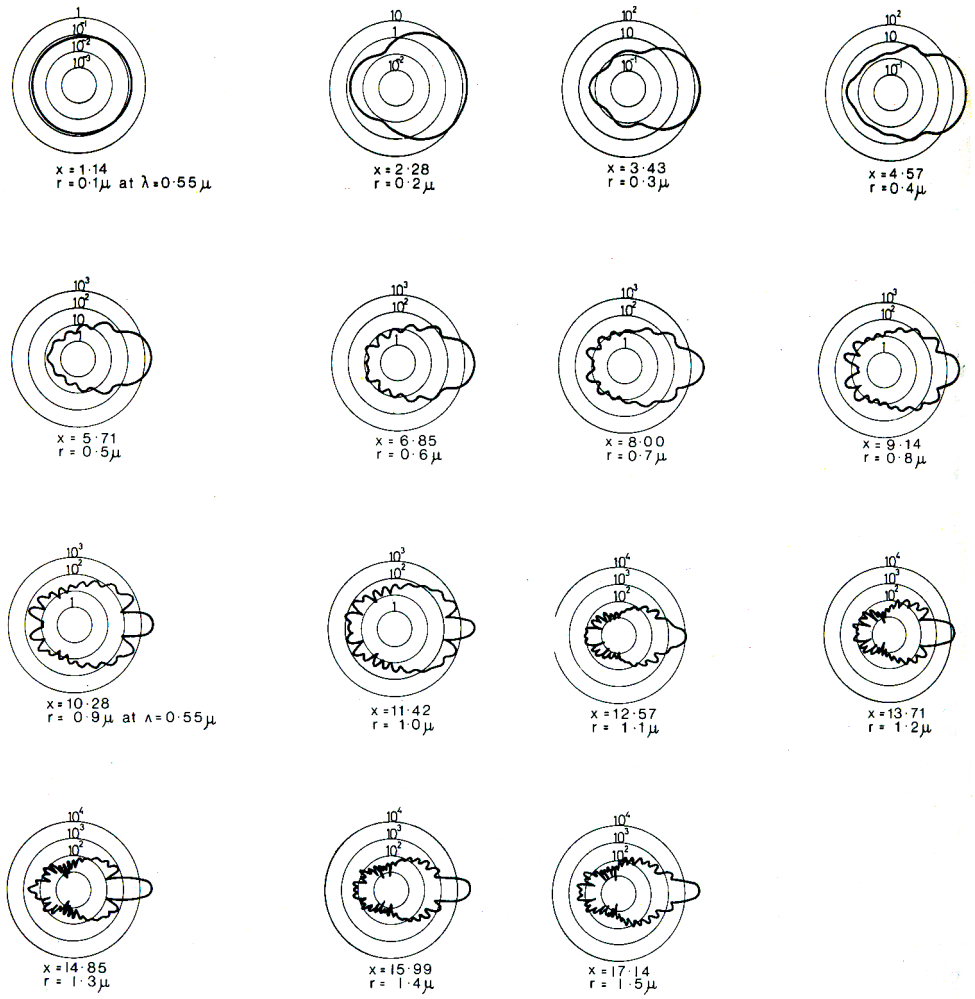


Figure 1.7: Mie scattering intensity as a function of scattering angle θ. θ increases anticlockwise, with incident light from the left. The wavelength of light is 0.55 μm and particle radii r and size parameters x are shown under each figure. Taken from Twomey [1977].

Chapter 2

AATSR and Other Aerosol Instrumentation

The bulk of instrument technical information and figures in this chapter are taken from official AATSR reference documents by ESA [2005].

2.1 The historical context of the AATSR mission

The Advanced Along Track Scanning Radiometer (AATSR) is an instrument aboard the ESA satellite Envisat, launched in March 2002. It is the successor to the earlier instruments ATSR-1, launched on ERS-1 in July 1991 and in operation until March 2000, and ATSR-2, launched on ERS-2 in April 1995 and still in operation (albeit with pointing issues since 2000). They are similar in design and principle, although improvements have been made between instruments. Together the series provides a 15-year data set with many applications in climate studies, aside from their main purpose of monitoring sea surface temperatures.

While ATSR-1 measured radiance at three wavelengths in the infrared part of the spectrum, ATSR-2 and AATSR have an additional four channels in the visible region. It is these visible channels which are key to the instruments' ability to provide data suitable for aerosol retrievals. AATSR is an improvement on ATSR-2 in this regard as ATSR-2 suffers from restrictions on the amount of data which can be downlinked while AATSR is able to provide continous data from all channels.

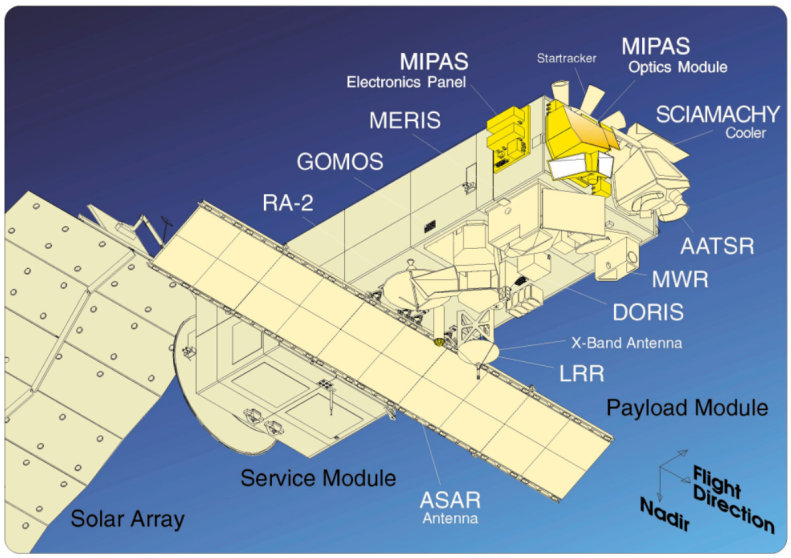


Figure 2.1: Instrumentation aboard Envisat

2.2 Instrument specifications and capabilities

2.2.1 Location

The AATSR instrument is on one end of Envisat, as shown in Figure 2.1. Envisat orbits the Earth in a sun-synchronous polar orbit, with a mean local solar time (MLST) of 10:00 am on for the descending node. It is at an altitude of approximately 800 km and 98.55° inclination to the equator. The reference orbit repeats once every 35 days, and global coverage is achieved once every 3 days as a result of orbital precession. A single Envisat orbit takes about 101 minutes. The orbit is controlled to keep it as precise as possible: within a maximum deviation of ± 1 km from the ground and ± 5 minutes on the MLST at the equator.

2.2.2 Design and data acquisition

The basic structure of the instrument itself can be seen in Figure 2.2. AATSR and its predecessors are unique in that they use two views (near-simultaneous in time) with differing path lengths to discriminate between radiance from the surface and radiance from the atmosphere. The radiometer measures radiance at nadir and along track at seven channels, with positions and widths summarised in Table 2.1 and Figure 2.3.

The instrument has a signal to noise (S:N) ratio of 20:1 at 0.5% spectral albedo. It measures top-of-atmosphere (TOA) radiance over the channels with an absolute accuracy of 5% over its entire range. Light

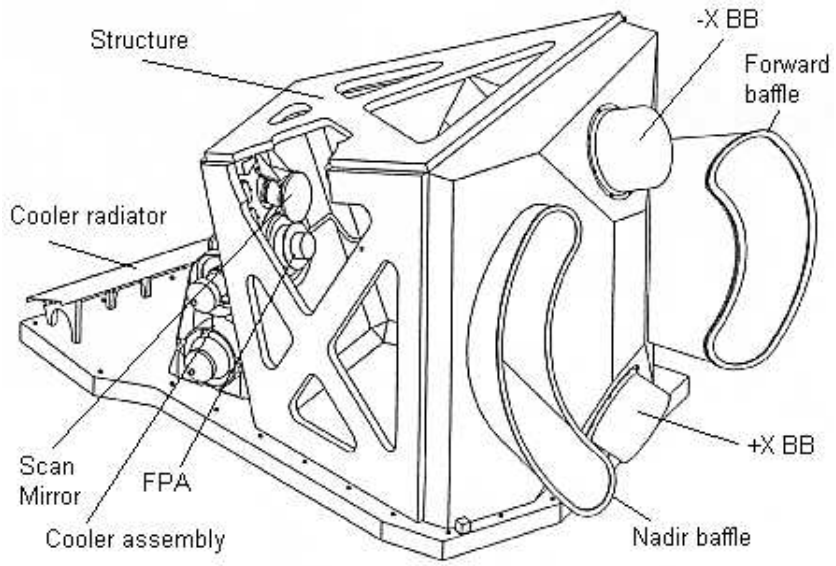


Figure 2.2: AATSR instrument structure. The FPA is the focal plane assembly, and the ±X BB the hot and cold black bodies respectively, used in calibration of the infrared channels.

Channel	Central Wavelength	Bandwidth
0.55 μm	0.555 μm	20 nm
0.66 μm	0.659 μm	20 nm
0.87 μm	0.865 μm	20 nm
1.6 μm	1.61 μm	0.3 μm
3.75 μm	3.70 μm	0.3 μm
11 μm	10.85 μm	1.0 μm
12 μm	12.00 μm	1.0 μm

Table 2.1: AATSR channel locations and bandwidths

is reflected off a scan mirror onto a paraboloid mirror, after which it is focussed and reflected into the visible and infrared focal plane assemblies (FPAs). Here the radiance is converted into electrical signals, which are processed to amplify and digitise them and then sent to data-downlinking systems to be transmitted back to Earth. The baffles shield the nadir and forward windows from direct sunlight, ensuring radiance picked up is only from the region being scanned. The FPA for the infrared channels is cooled to about 80 K whilst that for the visible channels is maintained at ambient temperature.

The viewing geometry of the radiometer is shown on the left of Figure 2.4. The instantaneous field of view (IFOV) consists of two curved swathes around 500 km wide: a nadir view, looking straight down, and

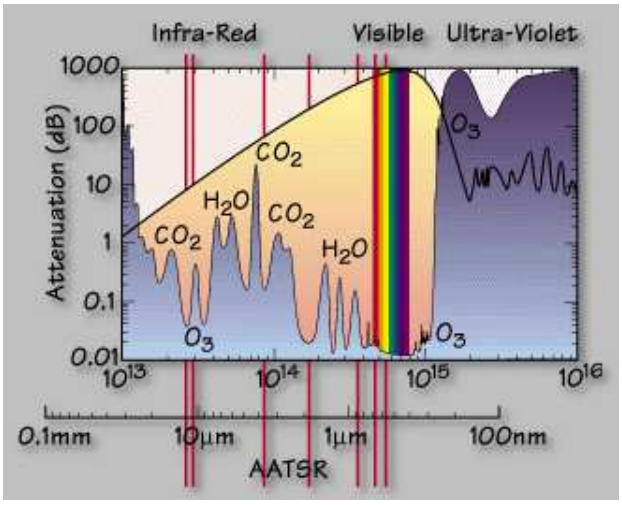


Figure 2.3: Spectral position of AATSR channels

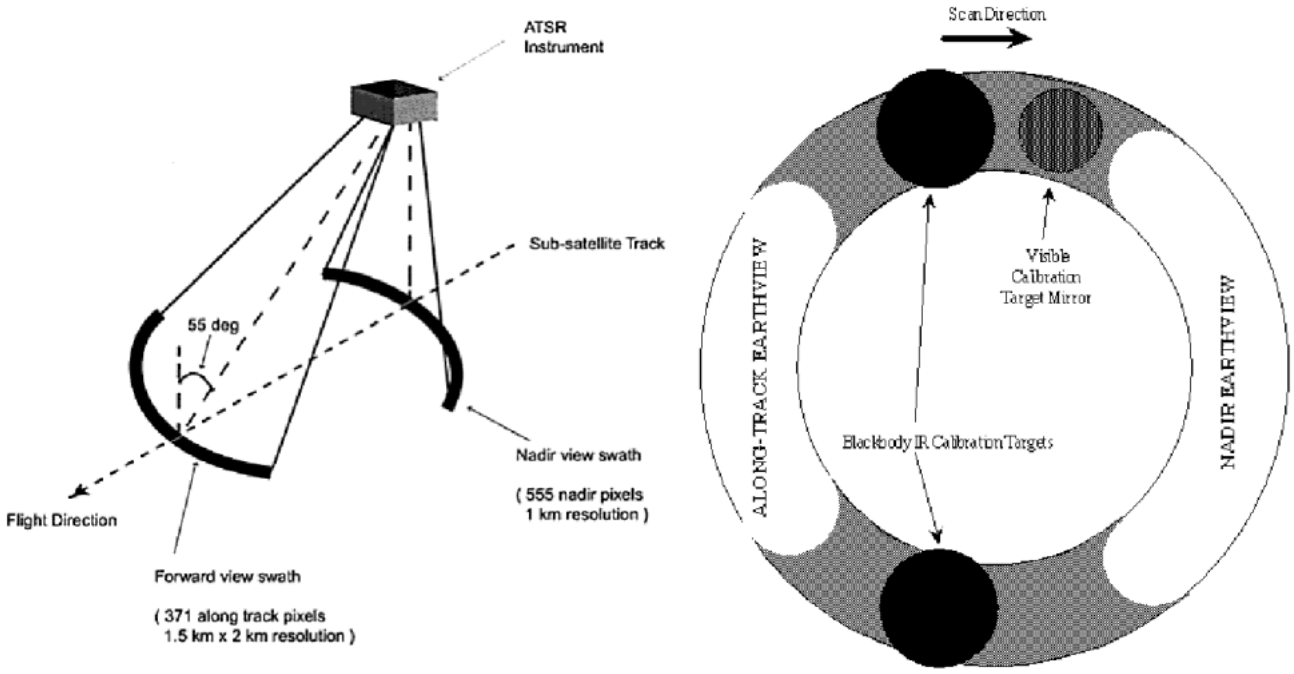


Figure 2.4: (A)ATSR viewing geometry (left) and scan cycle (right)

a forward view inclined at an angle of 55° to the normal to the surface. There are 555 pixels across the nadir swath (with a size of about 1km^2 at the centre) and 371 across the forward (with a size of about 1.5km^2 at the centre). During each scan cycle the satellite moves 1 km (so 1 nadir pixel) onward with respect to the Earth's surface. The scan cycle repeats 6.6 times per second.

With an altitude of around 800 km, the forward view samples approximately 1,000 km in front of the nadir at any given time. Thus after around 150 seconds (at 6.6 scan cycles per second) the satellite has moved such that nadir view samples the same region, giving two views of the scene with differing path lengths. While this time delay means the views are not *strictly* simultaneous, it is still valid to treat them as such because this timescale is short compared with that of atmospheric processes. Were it not, then the dual-view approach of the (A)ATSR instruments would not be able to be exploited in this way.

The FPA is occassionally warmed up to outgass. This is because contaminants condense onto the cold surfaces of the FPA and detectors, degrading instrument operation due to signal attenuation, as well as increasing the burden on the FPA coolers. These contaminants do not affect signal calibration because they affect both views of the Earth and of calibration targets. During outgassing, which occurs for around 2 days at a time each 3 months, no data is collected from the IR or 1.6 μm channels. As the aerosol retrieval uses the 1.6 μm channel, aerosol retrievals cannot be performed during outgassing. Data users are notified in advance of scheduled outgassings.

2.2.3 Instrumental calibration and long-term performance

AATSR is calibrated during each instrument scan cycle, the pattern of which is shown on the right of Figure 2.4. The baffles covering the viewing windows and the black bodies themselves are also visible on the exterior of the instrument in Figure 2.2. Between taking nadir and forward views, a black-body target is viewed to calibrate the IR channels. One is maintained at 305 K while the other is at the ambient instrument chamber temperature (256 K). This range encompasses all the expected brightness temperatures that the instrument measures. The black bodies are cylindrical, insultating and non-reflective with a temperature monitoring system. There is also a visible-channel calibration system (VISCAL) which views the sun once per orbit near sunrise, and uses an opal tile to scatter light onto the detector at the appropriate time in the scan.

The performance of the instrument over its life to date was discussed at the recent MERIS/(A)ATSR workshop by Mutlow et al. [2005]. The instrument was found to be fulfilling its original specifications. Some corrections may be needed to account for instrumental drift (Smith [2005]) although these have not been finalised.

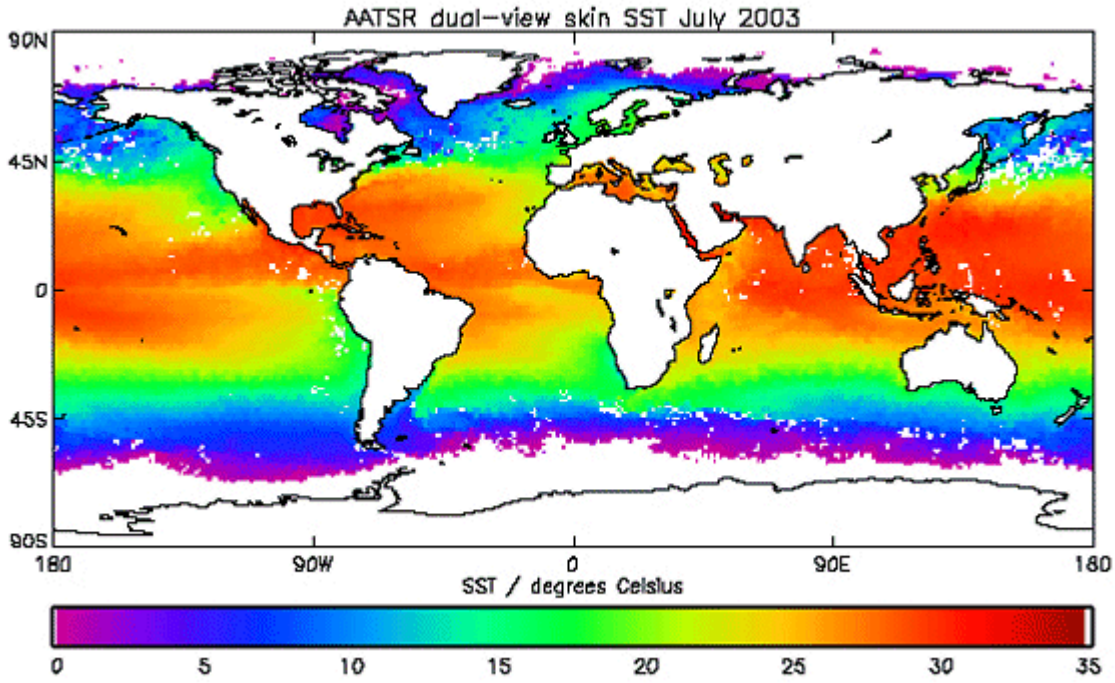


Figure 2.5: Global mean SST for July 2003, courtesy of Anne O'Carroll (Met Office)

2.3 Non-aerosol uses for AATSR data

2.3.1 Sea surface temperature retrievals

As previously mentioned, the primary objective of AATSR and predecessors was to provide a continuous, global record of sea surface temperature (SST) over a long period of time. Satellites make ideal platforms for such instruments, and offer several advantages over other data sources such as ships and buoys. These include a much greater coverage, and also the fact that ships and buoys give measurements which are not truly of the sea *surface skin* temperature but instead that of the near-surface bulk sea. The difference between bulk and skin temperatures can be of the order of tenths of a degree. As it provides a single data set AATSR is self-consistent whereas ships and buoys, being many data sets collected in different ways, are not. It should, however, be noted that ship and buoy measurements are still useful (for example for validation of results from satellite data).

Measuring SST accurately is very important in climate research. This is particularly true for the detection of global warming, where high precision is required so any long-term trend is not hidden by noise. As a result, the AATSR mission requirements state that the instrument should be able to retrieve SST with an

accuracy of better than 0.3 K (1σ limit), and up to 0.1 K under favourable conditions. In addition, there are strict requirements on instrumental calibration, to assess any long-term drift and provide continuity with the ATSR-1 and ATSR-2 data sets.

Temperature is retrieved using the infrared channels. The visible channels are used for cloud flagging (1.6 μm) and atmospheric correction for aerosol scattering (others), and it is for this reason that AATSR is suitable for performing aerosol retrievals. At nighttime the visible channels cannot be used so differences in correlations between thermal channels are used for these purposes instead. Similarly, aerosol retrieval can only be performed during the day. Retrievals are often presented as monthly means (as in Figure 2.5). These have the advantage that the entire globe will have been sampled several times by the satellite during this period, so the likelihood of at least one cloud-free retrieval for each pixel is increased.

Some work (such as Good et al. [2005]) has been done into trends in SST from ATSR-2 and AATSR data. For longer-term trends, however, (A)ATSR data is typically used in conjunction with other data sets to extend the range of time spanned. SST measurements are also useful in other modelling studies, such as coupled ocean-atmosphere circulation models or studying periodic oceanic features such as the El Niño/La Niña Southern Oscillation (Cardon et al. [1998]).

2.3.2 Land remote sensing

A secondary objective of AATSR was to continue ATSR-2's records of data for land remote sensing. Both land surface temperature (LST) and vegetation properties such as Normalized Difference Vegetation Index (NDVI), which provides information about surface type, are routinely retrieved. LST retrievals are similar in principle to SST retrievals.

A difficulty associated with remote sensing over land as opposed to the sea is that the land surface albedo often shows a strong dependence on viewing geometry. In other words, it has a steep bi-directional reflectance distribution function (BRDF), which makes modelling more difficult. In addition the land surface can be highly heterogeneous over the scale of the instrument's resolution–adjacent fields of different crops can have very different albedos. Despite these difficulties, progress in land remote sensing including retrieval of LST (Sória and Sobrino [2005]) and BRDF has been made with (A)ATSR data, with retrieval of BRDF often performed simultaneously with aerosol retrieval (North et al. [1999], North [2002]).

Burning vegetation can also be monitored using the same method as retrievals of LST, although the fact that fire temperatures exceed that of the hot calibration body (305 K) means that the temperatures retrieved

are not as precise as those for SST or 'normal' LST. Still, AATSR has become an important tool in the monitoring of biomass burning (Huang and Siegert [2004]). The instrument was designed such that strong signals from fire do not swamp the detector so useful information can still be gleaned from the data.

2.3.3 Retrieval of cloud properties

AOPP and the Rutherford Appleton Laboratory (RAL) at Didcot together have a scheme for aerosol and cloud retrievals called ORAC (Oxford-RAL Aerosol and Cloud) which has been used with ATSR-2 and AATSR data and will be discussed in detail in the next chapter. The principles behind aerosol and cloud retrievals are similar in that aerosol may be considered to be a very optically thin layer of cloud. One important difference is that the aerosol retrieval does not currently use AATSR's infrared channels, while the cloud retrieval does.

Cloud parameters retrieved by ORAC include optical depth, effective radius, cloud-top temperature, the fraction of cloud within a pixel or superpixel, and cloud type. Cloud-top temperature may be converted to height and pressure.

2.4 Other instruments in use for aerosol characterisation

A wide variety of instrumentation is in current or recent use to characterise atmospheric aerosols. These can be broadly divided into two categories: satellite instruments similar to AATSR for remote sensing, and terrestrial instruments (either on the ground or aboard ships and aircraft).

2.4.1 Satellite-borne

MERIS

Like AATSR, the MEdium Resolution Imaging Spectrometer MERIS is an instrument aboard Envisat and the two have several common features and goals. Both are radiometers, with MERIS measuring at 15 wavelengths in the visible region. MERIS also has a higher resolution (approximately 300 m) and wider field of view (68.5°). MERIS only has one viewing geometry, which gives AATSR an advantage in terms of making corrections for the aerosol contribution to signal. MERIS is primarily used for ocean colour measurements and vegetation remote sensing.

Being jointly aboard ENVISAT offers opportunities for synergy between MERIS and AATSR data and derived products, which has been discussed recently by Silió and Bricaud [2005]. MERIS is used along with (A)ATSR and SEVIRI in the GlobAerosol project (discussed by Thomas [2006]), although the retrieval algorithms differ.

SCIAMACHY and GOME

The SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) is also aboard Envisat and is the successor of the Global Ozone Monitoring Experiment (GOME), which flew on ERS-2 alongside ATSR-2. These instruments are primarily used for retrieval of vertical profiles of various atmospheric trace gases (such as work by Dils [2006] and Richter [2006]), although they have also been used to characterise aerosol distribution.

GOME measured in the visible at a nadir geometry, while SCIAMACHY extends the range of wavelengths samples into the infrared and is capable of several viewing geometries. Both instruments make use of the differential optical absorption spectroscopy (DOAS) technique. The horizontal resolution of the instruments (60 km $\times$ 30 km at best) is markedly poorer than that of the ATSR series. SCIAMACHY's variable viewing geometries allow for good retrieval of vertical profiles.

MISR

The Multi-angle Imaging SpectroRadiometer (MISR) is located on NASA's TERRA satellite and run by the Jet Propulsion Laboratory (JPL) at CalTech. Designed and widely used (Abdou et al. [2005]) for aerosol retrieval, like AATSR it measures in the visible and infrared (although only over 4 channels) at 9 widely-spaced angles. It is capable of a very high spatial resolution of 275m (250m at nadir) although in practice for a more reasonable data transmission rate it tends to work with a resolution of between 1 $\times$ 1 km and 4 $\times$ 4 km, adjustable for each detector. MISR is currently the only instrument which is routinely used to produce multi-angle retrievals of aerosol optical depth.

MODIS

An instrument carried on two satellites (NASA's TERRA and AQUA), the MODerate resolution Imaging Spectroradiometer (MODIS) measures at 36 bands in the visible and infrared. Resolution varies between 250 $\times$ 250 m and 1 $\times$ 1 km dependent on band, and it has a wide swath (55.5°, giving a 2,330 km overpass). As well as retrieving aerosol (Abdou et al. [2005]), a primary use of MODIS is land surface classification

and this land product is used to set the *a priori* surface albedo used in the AATSR retrieval (described in Section 3.3.2). Over the ocean MODIS is used for ocean colour and SST measurements.

SEVIRI

Unlike the other satellite instruments mentioned here, the Spinning Enhanced Visible and Infra-Red Imager is in a geostationary orbit. Located above equatorial Africa aboard the Meteosat 8 satellite, it images the visible portion of the globe every 15 minutes using 4 channels in the visible/near-infrared region at 1 km $\times$ 1 km resolution and a further 8 channels in the infrared at 3 km by 3 km resolution. Launched on 29 January 2004, it is expected to have a lifetime of 7 years. As well as aerosol detection, it is used in remote sensing of cloud and surface properties.

SEVIRI deserves special mention in that it is used in the GlobAEROSOL project alongside (A)ATSR and MERIS (Thomas [2006]). The same principles apply to both (A)ATSR and SEVIRI aerosol retrievals, although as previously mentioned the SEVIRI viewing geometry is different, as are the precise channel wavelengths and bandwidths.

AVHRR

The Advanced Very High Resolution Radiometers (AVHRR) are a series of instruments which have flown on the American satellites TIROS-N and NOAA-6 onwards (with improvements made since NOAA-15). They are primarily used for surface temperature and cloud measurements, but have also been used to obtain aerosol properties (Wetzel and Stowe [1999]). AVHRR measures at 5 bands in the visible and infrared regions at a resolution adjustable to either 1.1 $\times$ 1.1 km or 4 $\times$ 4 km (which enables global coverage).

OMI

A joint venture between the American NASA, the Netherlands' Agency for Aerospace Programs (NIVR) and Finnish Meteorological Institute (FMI), the Ozone Monitoring Instrument (OMI) aboard AURA was designed as a successor to the Total Ozone Mapping Spectrometer (TOMS).

OMI is a nadir viewing radiometer which measures backscattered solar radiation from 0.270 to 0.500 μm, with a spectral resolution of about 0.5 nm. It has a 2600 km swath, resulting in daily global coverage, and a resolution of 13 $\times$ 24 km in normal operational mode. In addition to O_3, OMI has been used to retrieve aerosol properties (Braak [2006]).

CALIPSO

Another joint venture (between NASA and the French CNES), the recently-launched Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite was designed specifically to monitor cloud and aerosol. Launched on April 28 2006, the mission's initial data is currently being released.

CALIPSO carries 3 co-aligned instruments measuring at nadir. The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) measures backscattered light intensity at 532 nm and 1064 nm. The Imaging Infrared Radiometer (IIR) measures at 8.7 μm, 10.5 μm, and 12.0 μm and has a 64 $\times$ 64 km at 1 $\times$ 1 km resolution. Finally, the Wide Field Camera (WFC) measures intensity over the 270-620 nm region, designed to overlap with a MODIS band. Together, these instruments will be able to probe cloud and aerosol properties and vertical structure in the atmosphere.

2.4.2 Terrestrial

AERONET

The AERONET (AErosol RObotic NETwork) is a global system of sun photometers coordinated by NASA [2006], providing data on aerosol properties and precipitable water. Photometers work by measuring light intensity, in a similar way to the (A)ATSR instruments. A diverse range of environments are covered, as shown in Figure 2.6. Data is available from 1993 onwards (although not all sites have been active since then). It has recently been extended to provide ocean colour products at several coastal sites.

Because of the long spatial and temporal coverage, AERONET sites are often chosen to validate satellite retrievals (for example by Abdou et al. [2005]). Ground-based photometers have an advantage over satellites in that measurement can be taken continually at a specific location if continuity is required, although as they are point measurements they are less useful than satellites for estimating global coverage. Being physically located on the ground also makes maintenance of such instrumentation simpler than that of a satellite, and their calibration is aided by the presence of the Sun as a well-defined light source to sample.

Other land, sea and air campaigns

The literature shows a wealth of one-off or ongoing campaigns of aerosol characterisation based upon data collected on land (Lu et al. [2006]), at sea (Ramachandran [2005]) and increasingly by aircraft (such as SAMUM 2005, Inomata et al. [2006], Targino and Noone [2006] and Yoon et al. [2006]), although most of these are local as opposed to worldwide in scope.

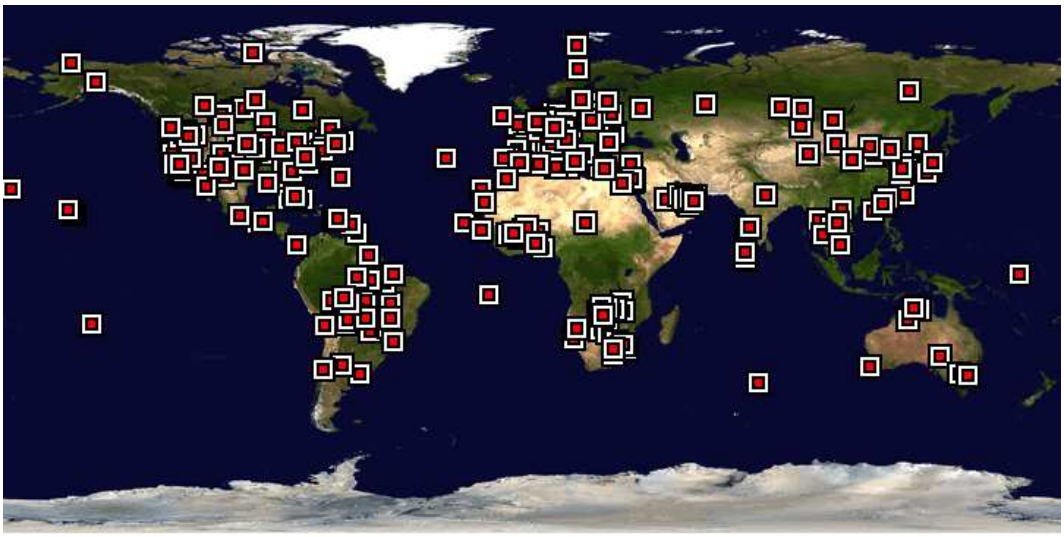


Figure 2.6: AERONET global coverage. Red squares indicate the presence of an AERONET station.

A variety of techniques are employed ranging from photometry and LIDAR to mass spectrometry. In one example, the SAharan Mineral DUst ExperiMent (SAMUM 2005) was an international project carried out largely over Morocco in 2005, and involved several flights in the custom aircraft Falcon. This had a payload of a nadir-viewing LIDAR system, a photometer and particle counters able to characterise aerosol microphysical and chemical properties. The goal of the project was to observe the radiative effects of Saharan dust as it is blown off the desert. Validation studies are being carried out with ground-based LIDAR and *in situ* measurements, as well as data from the MERIS and MISR satellite instruments.

Chapter 3

Theory and Retrieval Algorithm

3.1 The inverse problem

The field of retrieval theory is dedicated to the solution of the inverse problem, defined and investigated thoroughly by Rodgers [2000]. There are systems in which we are unable to measure directly the quantity we are interested in, but are able to measure some other quantity somehow related to or resulting from it. From this measurement y we wish to find out something about the state x the system is in. Given $y = F(x)$ for some function F known as the Forward Function, approximated by the Forward Model (FM) $f(x)$, solving the inverse problem means being able to find the inverse function $x = f^{-1}(y)$ and so obtain state information x from measurement y.

The premise behind performing an aerosol retrieval is that aerosol in the atmosphere will affect the light received at the satellite through a combination of absorption and scattering. The radiances AATSR measures form our measurement vector y and the state information vector x we wish to retrieve consists of various atmospheric aerosol and surface properties (namely aerosol optical depth, effective radius, and surface albedo at 0.55 μm for both forward and nadir viewing geometries). Aerosol radiative properties depend on factors such as aerosol particle size, shape, chemical composition and vertical profile so a forward model must take such factors into account. The use of satellite data in this way is known as remote sensing (or remote sounding), as the measurements are taken at a distance.

The algorithm developed and used through this project is based on the GlobAEROSOL algorithm for retrieval of aerosol properties (Thomas et al. [2006]), developed at Oxford and the Rutherford Appleton Laboratory (RAL), which has itself been developed from an earlier algorithm designed for ATSR-2 (Marsh et al. [2004]). The scheme is known as the Oxford-RAL retrieval of Aerosol and Clouds (ORAC) and is

applicable to both ATSR-2 and AATSR date. Continuing the chain further back, ORAC was built from the Enhanced Cloud Processor (ECP) described by Watts et al. [1998], which was designed to retrieve cloud properties from ATSR measurements.

The main development from the GlobAERSOSOL scheme to the one described here is that the former only uses the instrument's nadir view channels, while this scheme makes use of both nadir and forward views (and hence is known as a dual-view retrieval). This doubles the number of measurements in use, and returns one more piece of state information (the forward-view surface albedo at 0.55 μm).

3.1.1 Retrieved aerosol parameters

A sample of atmostpheric aerosol can be described by its number density, size and shape distribution, vertical profile and chemical make-up (which determines refractive index). Assuming the particles are spherical allows computation of radiative properties from Mie theory.

Aerosol optical depth τ is the first quantity retrieved in the ORAC scheme, and this is the parameter which has the largest impact on observed radiance. Optical depth at 0.55 μm is retrieved directly during the iteration, and later related to optical depth at 0.87 μm using lookup tables. Optical depth is defined as:

$$\tau(\lambda) = \int_0^\infty \beta_e(z,\lambda)\, dz = \int_0^\infty (\beta_s + \beta_a)(z,\lambda)\, dz \tag{3.1}$$

The total extinction coefficient β_e, is defined as the sum of the extinction due to absorption (β_a) and scattering (β_s). Radiative characteristics are described by the vertical profiles of β_a and β_s along with the scattering phase function $P(\theta)$, which determines the spatial distribution of the scattered radiation. Other useful aerosol optical properties include the single scattering albedo ω_o, which is the ratio β_s/β_e, and the asymmetry parameter g which is the weighted integral of $P(\theta)$ from 0 to 180°. For a given aerosol shape, size, and refractive index β_e is proportional to the aerosol concentration while $P(\theta)$ is not.

Mie theory shows that the extinction coefficient β is given by:

$$\beta(z,\lambda) = \int_0^\infty Q_e(z,m,x)\, \pi r^2\, n(z,r)\, dr \tag{3.2}$$

Q_e is the Mie extinction efficiency factor, dependent on the Mie size parameter $x = 2\pi r/\lambda$ and the refractive index of the particles ($m = m_r + im_i$). $n(r)$ is the aerosol number size distribution, dependent on radius r.

The most appropriate distribution to use to characterise atmospheric aerosol is the log-normal one (Davies [1974]). This is described by its modal radius (r_i), standard deviation of $\ln(r)$ (σ) and total number density (N). The second parameter retrieved, then, tells us about aerosol particle size. This is the effective radius r_e, defined as the ratio of the 3rd and 2nd moments of the size distribution:

$$r_e = \frac{\int_0^\infty r^3 n(r) dr}{\int_0^\infty r^2 n(r) dr} \tag{3.3}$$

Vertical distribution also affects radiative properties, although this is currently not retrieved as the visible channels do not provide information on it (shown in more detail in Section 3.4.3). As a result the aerosol height distribution at height h and for scale height Z (dependent on aerosol type) is fixed and described as follows:

$$N(h) = N(0)e^{-h/Z} \tag{3.4}$$

No information on aerosol shape distribution is currently retrieved and aerosol is assumed to be spherical. This is largely due to difficulties in the modelling of radiative characteristics of non-spherical particles, although some progress has recently been made in this field (such as Dubovik et al. [2005]). Work on this project to date has been focussed on oceanic scenes, for which aerosol particle sphericity is a reasonably valid assumption. However, other classes of aerosol (particularly desert dust) tend not to be spherical so the assumption of sphericity must be taken as a (currently) necessary approximation.

3.1.2 Retrieved surface parameters

A large proportion of the radiance arriving at the satellite is due to direct reflection of sunlight off the surface, particularly over land. The third and fourth retrieved parameters are the surface albedo at 0.55 μm for the nadir and forward viewing geometries respectively. These are retrieved as independent values, although a high degree of correlation between them would be expected for many surfaces.

The initial guess at surface albedo (taken from modelled *a priori* values, discussed in Section 3.3.2) for each view is used to form a picture of the spectral shape of the surface by calculating the ratio of the albedo at wavelength λ to that at 0.55 μm for all four channels. During the retrieval process, the spectral shape (i.e. ratio between channels) is constrained and only the brightness (i.e. absolute magnitude) is allowed to vary. This effectively means that instead of needing eight variables to describe the surface (one per measurement channel used per view) only two are required (one per view for the 0.55 μm channels) provided the correct surface spectral shape is known.

3.2 The forward model

The aerosol forward model to compute radiance given the atmospheric state uses a series of lookup tables rather than perform radiative transfer calculations on-the-fly, as using lookup tables is a computationally cheaper method. The lookup tables are generated using assumed values of aerosol refractive index, size distribution, phase function and vertical distribution, as well as assumptions of the absorption of atmospheric gases. They are generated for a 32 layer plane parallel atmosphere stretching from 0 - 100 km in height, with a 1 km spacing up to 25 km, stretching to 30 km for the top layer.

The Optical Properties of Aerosols and Clouds (OPAC) database of Hess et al. [1998] provides information on physical (e.g. size and vertical distribution) and optical (e.g. wavelength-dependent complex refractive index) properties, which are used to build up profiles for the characteristic aerosol types. Information on various aerosol components is used to generate four aerosol classes: continental clean, desert, maritime clean and Antarctic. Information on the components can be found in Table 3.1 and details of the assumed composition of the classes in Table 3.2. Increased relative humidity provides an increase in the amount of atmospheric water vapour which can condense onto the particles, altering their size and refractive index. This effect is taken into consideration utilising the method of Hănel [1976].

- The water-insoluble component of aerosol particles consist mostly of soil particles containing organic material. The water-soluble substances are a mixture of sulphate, nitrate, and organic compounds, the mixing ratio of which varies with the type of sources (such as biogenic, fossil fuel combustion, volcanic, biomass burning).
- The term ‘soot’ represents absorbing black carbon that is either a product of direct particle emission into the atmosphere by combustion processes that form highly absorbing graphitic along with weakly absorbing organic particles, or a transformation within the atmosphere of combustion related gases to particles.
- Sea salt particles consist of the various kinds of salt contained in seawater. Two modes are given to allow for a different wind-speed-dependent increase of particle number for particles of different sizes.
- Desert dust is produced in arid regions and consists of a mixture of quartz and clay minerals. It is modelled with three modes to allow for increasing relative amounts of large particles with increasing turbidity. The size of mineral aerosol particles is assumed to be insensitive to relative humidity effects.
- Sulphate derivates, whether from natural or anthropogenic sources, are simply termed ‘sulphate’, except where they occur in a mixture of different water-soluble substances. Sulphate is found in Antarctic aerosol and in the stratospheric background aerosol.

Component	r_i (μm)	σ	Wet Mode Radius
Insoluble	0.47	2.51	indep.
Water-soluble	0.0212	2.24	*
Soot	0.0118	2.00	*
Sea Salt (Accumulation)	0.209	2.03	*
Sea Salt (Coarse)	1.75	2.03	
Mineral (Nucleation)	0.07	1.95	
Mineral (Accumulation)	0.39	2.00	indep.
Mineral (Coarse)	1.90	2.15	
Sulphate	0.0695	2.03	*

Table 3.1: Aerosol components, modal radius, and standard deviation for a log-normal size distribution and an assumed relative humidity of 0 %. Entries in the fourth column refer to the publications where the values of the wet mode radii can be obtained: * refers to d'Almeida et al. [1991] or Hess et al. [1998]. The label regarding the wet mode radius is set to "indep." if the aerosol component is not hygroscopic.

Mie theory is used to calculate scattering phase function for the central wavelength of each channel over a range of 20 effective radii r_e (from 0.02 μm to 20 μm). The exponentially-decreasing vertical profile described previously is then used to give a vertical profile of scattering. Phase function generation requires the following pair of assumptions:

- The radiative properties of the aerosol are constant across the width of each instrument channel. This is a reasonable assumption because the features of aerosol extinction spectra are very broad in comparison with gas absorption or emission features.
- The aerosol size distribution is, as previously mentioned, assumed log-normal. Changes in effective radius are achieved by adjusting the proportions of the differently-sized aerosol modes. If effective radius for an aerosol type equals that of the smallest or largest component in it, then the type becomes a single-component aerosol. If the size lies outside this range then the modal radius is shifted (keeping the distribution width constant)–in such situations the validity of the assumed size distribution becomes poorer. However, for very small aerosol particles optical properties become less important as effects become dominated by Rayleigh scattering.

For each layer at which the aerosol distribution is defined, the extinction coefficient β_e, single scattering albedo ω_a and the coefficients of a Legendre expansion of the scattering phase function are calculated for each channel and effective radius.

Type	Components	Mix. Ratios ($1/cm^{-3}$)
Continental clean	total	2,600
	water-soluble	2,600
	insoluble	0.15
Desert	total	2,300
	water-soluble	2,000
	mineral (nucleation)	269.5
	mineral (accumulation)	30.5
	mineral (coarse)	0.142
Maritime clean	total	1,520
	water-soluble	1,500
	sea salt (accumulation)	20
	sea salt (coarse)	0.0032
Antarctic	total	43
	sulphate	42.9
	sea salt (accumulation)	0.047
	soot	0.0053

Table 3.2: Composition of aerosol classes used in the retrieval. A 50% relative humidity is be considered for the hygroscopic components. Mixing ratio values are found in Table 4 of Hess et al. [1998].

3.2.1 Modelling atmospheric gas absorption

Following on from aerosol scattering properties, gas absorption over the instrument band passes is calculated and convolved with the instrument filter transmission functions using the MODTRAN database (version 3.5-v1.1) outlined by Brown et al. [2004]. This database provides tropical, mid-latitude summer and winter, subarctic summer and winter, and US Standard Atmosphere climatological atmospheres for H_2O, CO_2, O_3, N_2O, CO, CH_4, plus single profiles for HNO_3, NO, NO_2, SO_2, O_2, N_2, NH_3 and CFCs.

ORAC lookup tables are generated using the mid-latitude summer atmosphere only. This simplification can be made as gas absorption is weak compared to aerosol extinction in the visible and the instrumental channels are free from strong absorption features of gases which show large spatial and temporal variability (most notably H_2O).

3.2.2 Modelling atmospheric transmission and reflectance

To conclude lookup table generation, atmospheric transmission and bidirectional reflectance values are computed from aerosol phase function and gas absorption previously calculated. The DIscrete Ordinates Radiative Transfer (DISORT) software package (Olesen [2005]) is used for this purpose. DISORT is a

well-documented and often used general purpose algorithm for time-independent multiple-scatter transfer calculations. It solves the equation for the transfer of monochromatic light at wavelength λ as described by:

$$\mu \frac{dL_\lambda(\tau_\lambda, \mu, \phi)}{d\tau} = L_\lambda(\tau_\lambda, \mu, \phi) - L^S_\lambda(\tau_\lambda, \mu, \phi) \tag{3.5}$$

In equation 3.5, $L_\lambda(\tau_\lambda, \mu, \phi)$ is the intensity along direction (μ, ϕ), where μ is the cosine of the zenith angle and ϕ is azimuth angle, at optical depth τ_λ, measured perpendicular to the surface of the medium. $L^S_\lambda(\tau_\lambda, \mu, \phi)$ is the source function.

The surface albedo is set to zero for lookup table creation. The equation is solved for L to give tabulated transmission or reflectance (depending on the file) for optical depths τ at 0.55 μm for each of the 20 effective radii, 9 logarithmically-spaced 0.55 μm optical depths and sun/satellite geometry (specified by 20 equally spaced zenith angles and 11 equally spaced azimuth angles). Linear interpolation is used for values falling in between the points. Five lookup tables, known as Static Application Data (SAD) files, for each aerosol type/instrument channel combination are produced:

- Bidirectional reflectance of the aerosol layer, R_{BD}
- Reflectance of the aerosol layer to diffuse radiance, R_{FD}
- Diffuse transmission of the incident beam, T_{BD}
- Direct transmission of the beam, T_{DB}
- Transmission of diffuse incident radiance, T_D

Molecular absorption and Rayleigh scattering in the aerosol layer

The effects of molecular absorption and Rayleigh scattering are included in DISORT by an adjustment of the aerosol layer's optical depth, the particle's single scattering albedo and phase function in the following manner:

$$\tau = \tau_a + \tau_R + \tau_g \tag{3.6}$$

$$\omega = \frac{\tau_R + \omega_a \tau_a}{\tau_g + \tau_R + \tau_a} \tag{3.7}$$

$$P(\theta) = \frac{\tau_a \omega_a P_a(\theta) + \tau_R P_R(\theta)}{\tau_a \omega_a + \tau_R} \tag{3.8}$$

Here τ_a, τ_R and τ_g are the contributions to the total optical depth τ from aerosol scattering, Rayleigh scattering and gaseous absorption within each layer respectively. The aerosol single scattering albedo is denoted ω_a. For each layer bounded by lower and upper pressure levels p_l and p_u respectively and ground level pressure p_0, τ_R is calculated as:

$$\tau_R = \frac{\tau_{RT}[p_u - p_l]}{p_0} \tag{3.9}$$

τ_{RT}, the wavelength-dependent Rayleigh scattering optical depth for a column of atmosphere extending from the ground surface to the top of the atmosphere (TOA) is obtained fromJustus and Paris [1985]:

$$\tau_{RT}(\lambda) = \frac{p_0}{p_s} \times \frac{1}{117.03\lambda^4 - 1.316\lambda^2} \tag{3.10}$$

This takes the value of p_s (standard pressure) as 1013.25 mbar, and requires the ground pressure p_0 in mbar and wavelength λ in μm.

3.2.3 Computing top of atmosphere radiances

The forward model uses the lookup tables to predict a top of atmosphere (TOA) radiance using a three layer radiative transfer model, shown diagrammatically in Figure 3.1. The solar beam is incident on the aerosol layer (described by many layers in DISORT).

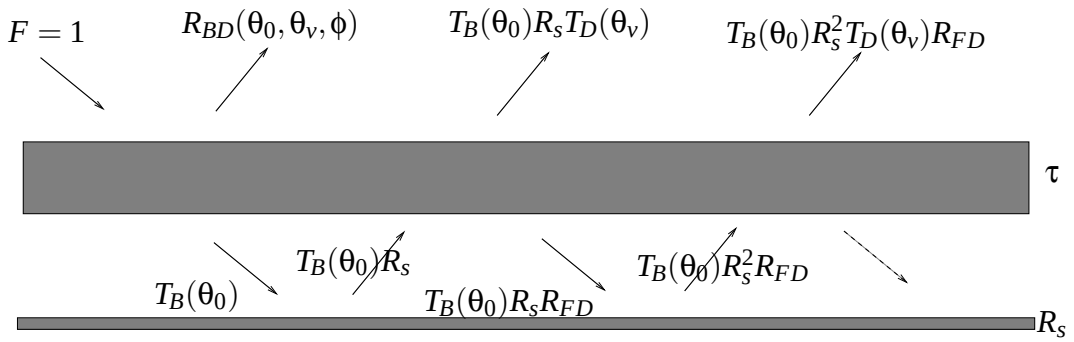


Figure 3.1: Aerosol layer - surface interactions.

The first contribution to the TOA radiance received at the satellite is the direct bidirectional reflectance (R_{BD}) of the aerosol layer. Transmission through the layer is partly by direct transmission of the beam (T_{DB}) and partly by diffuse transmission of scattered radiance (T_{BD}). These can be combined to give the total downward transmission through the aerosol layer $T_B = T_{BD} + T_{DB}$, as a Lambertian surface is assumed (meaning any directionality in the radiance is lost on reflection). The reflected radiance is assumed to be

diffuse and is partially transmitted (T_D) by the aerosol layer into the view direction, thereby giving the second term in the solar component. The aerosol layer also reflects downwards, giving rise to a series of multiple reflections and transmissions giving with rapidly decreasing contributions to the TOA reflectance. This multiple-scattering process is represented as follows:

$$\begin{aligned} R = & R_{BD}(\theta_0, \theta_v, \phi) \\ & + T_B(\theta_0) R_s T_D(\theta_v) \\ & + T_B(\theta_0) R_s^2 T_D(\theta_v) R_{FD} \\ & + T_B(\theta_0) R_s^3 T_D(\theta_v) R_{FD}^2 \\ & + \cdots \end{aligned} \tag{3.11}$$

Here θ_0 is the solar zenith angle, θ_v is the satellite (instrumental view) zenith angle and ϕ is the relative (solar to satellite) azimuth angle. This expression can be simplified to give:

$$R = R_{BD}(\theta_0, \theta_v, \phi) + T_B(\theta_0) T_D(\theta_v) R_s (1 + R_s R_{FD} + R_S^2 R_{FD}^2 + \cdots) \tag{3.12}$$

Further simplification as a geometric series limit gives the equation used to calculate TOA radiance seen by the satellite:

$$R = R_{BD}(\theta_0, \theta_v, \phi) + \frac{T_B(\theta_0) T_D(\theta_v) R_s}{(1 - R_s R_{FD})} \tag{3.13}$$

An example of TOA radiance predicted by the forward model as a function of τ and r_e is shown for $\theta_0 = 60^\circ$ and $\theta_v = 55^\circ$ as Figure 3.2. TOA radiance is given in normalised units to the satellite's detectors. The range of optical depth and effective radii scanned was chosen to encompass the most commonly encountered values for this aerosol type (maritime clean). The measurements give us the most information when contours on different channels are orthogonal, and as Figure 3.2 shows, over this range of τ and r_e the contours tend not to be parallel. Similar patterns are encountered for other geometries.

Forward model radiance gradient

The gradient version of the forward model $(\partial y / \partial x)$ is required for two purposes:

1. The gradient with respect to the retrieved state parameters x is needed for the inversion of the non-linear reflectance model by the Levenberg – Marquardt algorithm, which will be discussed later.

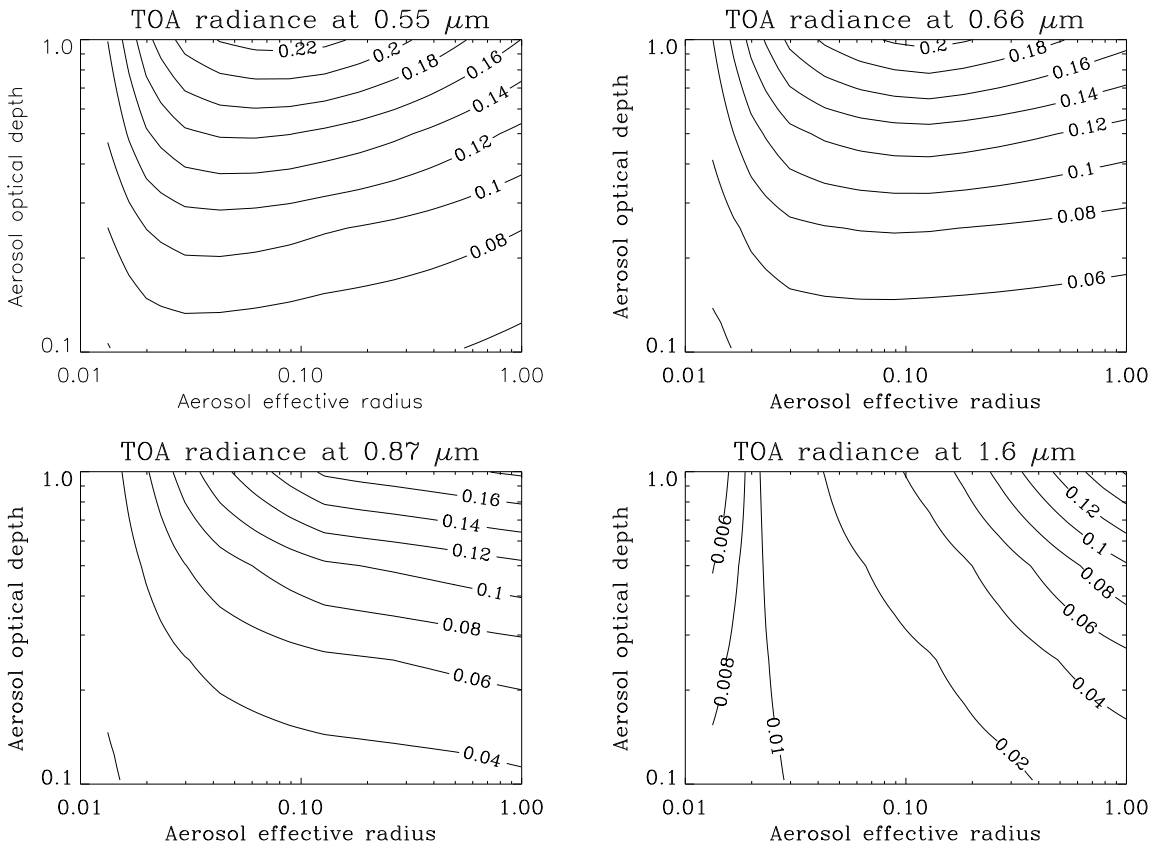


Figure 3.2: TOA radiance arriving at satellite nadir view channels as predicted by the forward model for $\theta_0 = 60°$, $\theta_v = 55°$, a surface albedo of 0.02 at all wavelengths and the maritime aerosol class.

2. The gradient with respect to non-retrieved model parameters (such as surface reflectance spectral shape) is used to judge the sensitivity to these parameters, and thus estimate their contribution to the retrieval error.

Equation 3.13 can be rewritten as follows:

$$R = R_{BD} + S \tag{3.14}$$

From this it can be stated that the gradient of the model with respect to aerosol optical depth τ is given by:

$$\frac{\partial R}{\partial \tau} = R'_{BD} + S\left[\frac{T_D T'_B + T'_D T_B}{T_B T_D} + \frac{R_S R'_{FD}}{1 - R_S R_{FD}}\right] \tag{3.15}$$

Here all $'$ indicate $\partial/\partial\tau$. The gradient of the model with respect to aerosol particle radius r_e is given by an analogous expression where all $'$ indicate $\partial/\partial r_e$. The gradient of the model with respect to the surface reflectance also follows directly from equation 3.13:

$$\frac{\partial R}{\partial R_s} = \frac{S}{R_s(1 - R_s R_{FD})} \tag{3.16}$$

3.3 *A priori* information

The retrieval algorithm makes use of *a priori* information on the retrieved state parameters x. This serves a dual purpose: firstly, to provide an initial guess at the values of x for the iteration, and secondly to act as a constraint on the retrieval and so make the process less likely to retrieve unphysical values for x. The precise treatment of the *a priori* information by the retrieval algorithm is discussed in Section 3.6.

3.3.1 Aerosol properties

Each different aerosol class in the retrieval has its own associated *a priori* optical depth τ and effective radius r_e, with associated errors, dependent on its composition. For the aerosol classes used in the GlobAEROSOL project these are listed in Table 3.3.

3.3.2 Surface albedo

Creating an *a priori* model of the surface for a pixel depends on whether it is a land or sea pixel. In both cases an effective Lambertian reflectance is calculated for compatability with the ORAC forward model. Over land, the MODIS land surface bidirectional reflectance product described by Scatterfield et al. [2004] is used (for both viewing geometries). If a pixel is missing it is filled by nearest neighbour sampling; if a large chunk is missing it is set to a default value. The product is interpolated onto a sinusoidal grid and

Aerosol Class	ORAC label	Log optical depth $\log(\tau)$	Error on $\log(\tau)$	Log effective radius $\log(r_e)$, μm	Error on $\log(r_e)$
Continental clean	A00	-1.2	1.0	-0.64	0.5
Desert	A03	-0.54	1.5	0.14	0.5
Maritime clean	A05	-1.0	1.0	-0.08	0.5
Antarctic	A09	-0.19	2.0	-0.82	0.5

Table 3.3: *A priori* information for the aerosol classes used in the retrieval scheme. Quantities are logarithmic as this is the form in which the retrieval code deals with them.

averaged to the retrieval resolution.

Over water the situation is slightly more complicated. Surface albedo is modelled using the method of Sidran [1981], which involves specular reflections from the wind-roughened surface (as modelled using the Fresnel equations and Cox and Munk [1954*a*] statistics. The wind speed at the pixel is obtained from data from the European Centre for Medium-Range Weather Forecasts (ECMWF [2006]). The formulation of a more accurate sea surface model is one of the goals for the next two years of work.

Sun-glint

The specular reflection of sunlight from the ocean surface (sun-glint) poses a problem for nadir measurements over the ocean. In some areas the surface contribution to TOA-observed radiance becomes completely dominant and masks any aerosol signal past the limits of detection of the retrieval scheme. As a result, no retrieval is attempted on pixels in the sun-glint region of the scene. The criterion for determining whether a pixel lies in the sun-glint region is a simple threshold based on the Cox and Munk-predicted radiance. This flagging of pixels for sun-glint contamination occurs during the creation of the *a priori* surface for the pixel.

Sun-glint is less of an issue for the forward viewing geometry, for two reasons:

- The forward geometry has a longer atmospheric path length, so the relative aerosol contribution to the signal is magnified.
- Because the instrument orbits at a local time of around 10am, the sun's position in relation to the instrument is always less favourable for the nadir view than the forward view. In other words, the forward view is less likely to be affected by sun-glint because it is not pointing at the sun-glint region.

The above means that while nadir measurements for a pixel may be unusable the forward measurements might still be. Nevertheless no retrieval is carried out in these regions, as using just the forward view's data in the retrieval might introduce a bias into the results due to imperfections in the forward model.

3.4 Discussion of forward model errors

It is not feasible to accurately model the full range of variation in aerosol composition, size, shape and vertical distributions, atmospheric composition and surface properties which affect the TOA-observed radiance. Even if all factors could be modelled the measurements do not provide enough information to retrieve them all: mathematically, the problem is ill-posed. The resulting approximations and assumptions in the forward

model will result in errors in the retrieved state parameters and it is important that the most important of these errors be quantified.

Discussion in this section is based upon the results found in the GlobAEROSOL single-view retrieval by Thomas et al. [2006]. The conclusions reached, however, should hold for the dual-view retrieval.

3.4.1 Surface albedo

Except in circumstances of extremely high aerosol loading (or cloud), a large portion of the instrumental signal will be dominated by surface features, with the aerosol signature essentially adding a perturbation. The forward view has a longer path length (hence aerosol contribution) than nadir although the surface is still a large effect. As a result the accuracy of the surface albedo is of key importance in determining the overall accuracy of the retrieval. Although surface albedo is a retrieved quantity the spectral shape is constrained by the *a priori* data for each pixel, so it is this spectral shape which is of greatest importance to model correctly. Several points can be made about surface modelling:

- Snow and ice surfaces have a very high albedo, making the surface contribution to radiance even greater. As a result retrieval of aerosol properties becomes more difficult as the aerosol signal becomes very small.
- Although the MODIS surface reflectance product is used over all land surfaces, it in turn depends on the MODIS aerosol products, which are considered inaccurate over desert regions. This can result in problems with the ORAC aerosol retrieval over desert regions (particularly the Sahara, Arabian peninsular, Tibetan high plateau and Gobi desert).
- Ocean colour varies across the world, although in this retrieval the same optical properties are used globally to set the *a priori*. This introduces some error in regions where high surface roughness produces a large amount of foam, or where there are large plankton blooms (both of which will change the spectral shape of the ocean).

3.4.2 Aerosol size distribution

Figure 3.4.2 gives an example of the difference between the scattering phase functions that correspond to two grossly different size distributions, which have the same effective radius r_e. Even in this extreme case the scattering phase function for the two distributions is remarkably similar and the single scatter albedo ω_a (which is the other aerosol parameter used by DISORT to calculate TOA radiances) for each distribution

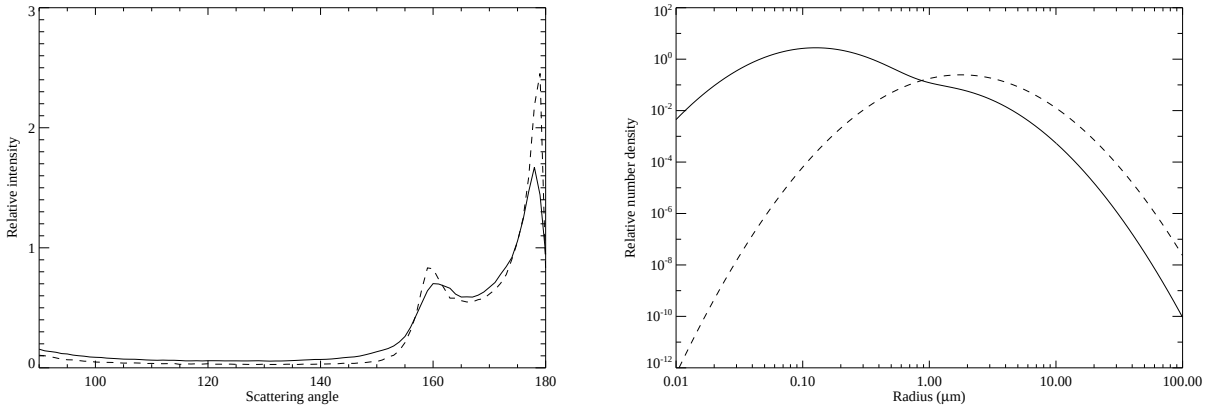


Figure 3.3: Comparison of the phase functions at 0.55 μm (*left*) of two size distributions (*right*), both with an effective radius of 10.3 μm. The first distribution (solid line) consists of the accumulation mode and coarse mode sea salt components from the OPAC database (Hess et al. [1998]) at 0% relative humidity, while the second distribution (dashed line) is the accumulation mode alone, with modal radius increased by a factor of 14 to give it the same effective radius as the combined distribution. Taken from Thomas et al. [2006].

is almost identical. It is likely, therefore, that this method of setting the effective radius does not introduce significant errors in the modelled radiances.

3.4.3 Sensitivity to vertical profiles of aerosol parameters

A simple 5-layer forward model from the ground to 2 km has been used to test whether the TOA radiance varies as a function of the aerosol number distribution, ω_0 and g. Note that this analysis was done using a previous version of the lookup table generation that used the asymmetry parameter g to describe the angular distribution of the aerosol scattering, rather than a Legendre expansion of the phase function. The TOA radiance was investigated for a model with constant $\omega_0 = 0.8$ and $g = 0.61$ for all layers. Surface albedo was set to 0.0. The following four vertical profiles were tested:

- Exponentially increasing with height
- Exponentially decreasing with height
- Linearly increasing with height
- Linearly decreasing with height

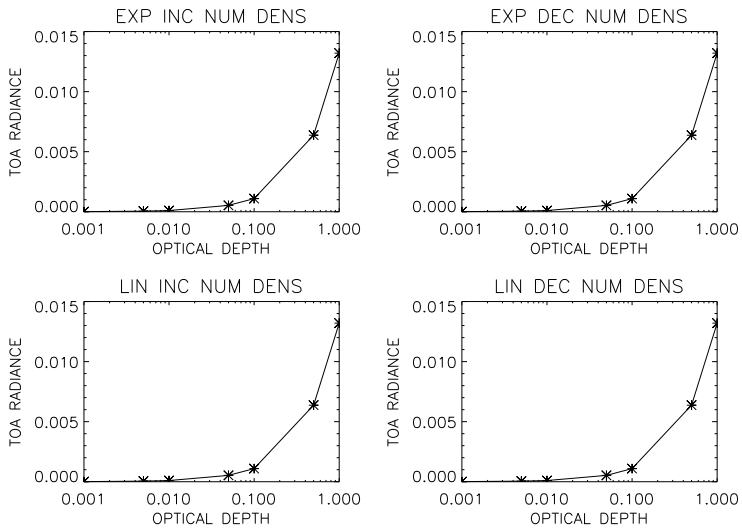


Figure 3.4: TOA radiance values for each of the four vertical aerosol N distributions at $\lambda = 0.67\,\mu$m. Taken from Thomas et al. [2006].

The TOA radiances for the vertical profiles are shown in Figure 3.4. This figure shows that the radiance values for all profiles are identical–if ω_0 and g were kept constant with height, the forward model would not be sensitive to the aerosol profile. Keeping all parameters constant and varying ω_0 produces the set of graphs in Figure 3.5. It can be seen that there is very little difference between the four plots when $\tau < 0.1$; however, for $\tau > 0.1$ the TOA radiance is dependent on whether ω_0 increases or decreases with height. This can be explained by more of the incoming radiation being absorbed high in the model atmosphere when ω_0 decreases with height. For the set of graphs in Figure 3.6 the only parameter which is permitted to vary with height is the aerosol asymmetry parameter g. Again it can be seen that the TOA radiance varies with its vertical distribution.

The conclusion to be drawn from these tests is that the forward model is sensitive to the vertical distribution of aerosol–although it is not directly sensitive to the vertical number distribution N it *is* sensitive to the vertical profiles of ω_0 and g. Equations 3.6, 3.7 and 3.8 show that values for τ, ω_0 and g within the layers of the forward model are dependent on τ_a and hence on N.

Due to an absence of information about vertical aerosol distribution at a given place and time this dependence cannot be accounted for in some *a priori*, and the measurements do not provide enough information to allow the retrieval of the vertical distribution and the other retrieved parameters. As a result this dependence adds a forward model error into the retrieval because the model does not take it into account.

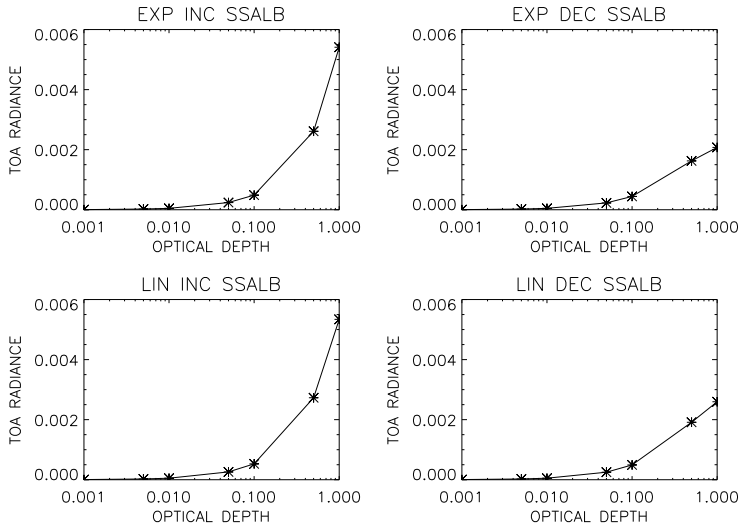


Figure 3.5: TOA radiance values for each of the four vertical aerosol ω_0 distributions at $\lambda = 0.67\,\mu$m. Taken from Thomas et al. [2006].

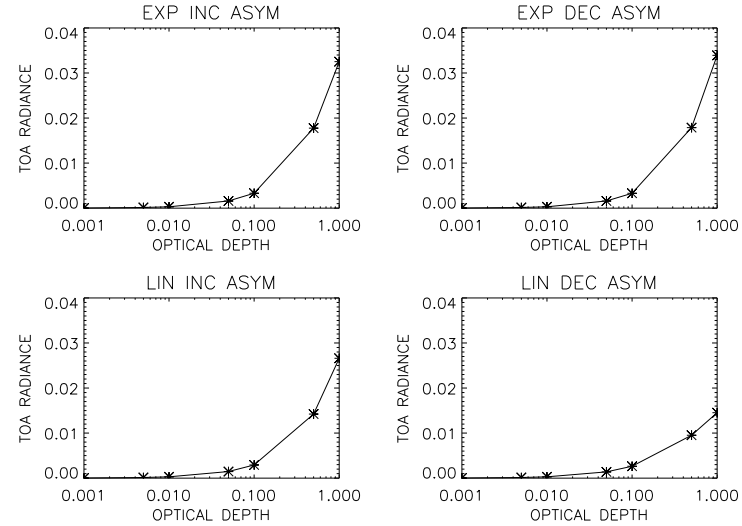


Figure 3.6: TOA radiance values for each of the four vertical aerosol g distributions at $\lambda = 0.67\,\mu$m. Taken from Thomas et al. [2006].

3.5 Data preprocessing

The ORAC processor expects all input data to be presented in unformatted binary files. Since such files are an exact reproduction of a computer's internal binary representation of data, the format of these files is

architecture dependant and thus ORAC input files are not portable between different systems. The primary role of the preprocessing is to read in the required data in it's own format (generally more portable standards), manipulate it and output it in unformatted form. The use of preprocessing code in this way enables the FORTRAN retrieval code to be treated almost as a 'black box', while the preprocessing code can be written in a higher level language (here Interactive Data Language, IDL, by RSI [2006]) which is much better suited to the quick development of new routines for reading different data. This provides a stable, highly optimised retrieval scheme, while maintaining the flexibility needed to allow the easy application of the scheme to new input and ancillary data.

IDL was selected as the language for the majority of the preprocessing code as there are readily available I/O routines for most scientific data formats and because of its inbuilt routines for data manipulation. The interactive nature of IDL is also a considerable advantage in the development and debugging of code, while the freely available run-time only version of IDL (IDL Virtual Machine) makes it economical to implement the ORAC processing chain on multi-processor computational clusters. Analysis of output data is also typically performed using IDL routines. The preprocessing additionally performs some important functions required by the retrieval:

Land flagging

The ORAC land flagging scheme uses the CIA coastline database. It consists of vector descriptions of land outlines rivers and political boundaries at approximately 1km resolution. The database was modified at RAL for use with ERS-2 and Envisat satellite instruments. The modified database is coded in hierarchical format which makes searching for the correct land surface flag fast and efficient.

Setting surface albedo

As mentioned in Section 3.3.2, *a priori* land surface reflectance is set using the MODIS surface albedo product for the appropriate date. Sea surface albedo and any resulting sun-glint are also calculated using the Cox and Munk algorithm.

Cloud flagging

Retrievals are only performed on cloud-free pixels. The ESA cloud flag product (produced in the ESA SST retrieval) is used to flag cloud over the oceans. Over land, the Stevens cloud flag developed at RAL is used. This is based on a NDVI technique looking at ratios between the visible channels, as shown in Equation 3.17. It is a simple threshold technique, with different thresholds for different surfaces (desert, clean water,

dirty water, ice, green vegetation, other vegetation, cloud, unclassified) based on test data sets. Thresholds are set conservatively to identify the maximum possible cloud contamination.

$$\frac{V.87\mu m - V.67\mu m}{V.87\mu m + V.67\mu m} \, (vs) \, \frac{V.67\mu m - V.55\mu m}{V.67\mu m + V.55\mu m} \tag{3.17}$$

Cloud flagging is performed at the full instrumental resolution regardless of the spatial resolution required by the retrieval, which minimises errors in cloud flagging due to sub-pixel clouds. Cloud flags in forward and nadir views may differ (due to both cloud movement in the time between views and differences in apparent height due to viewing geometry) and a pixel is only flagged as cloud-free if both forward and nadir views register as cloud-free.

Superpixelling

As well as running at full instrumental resolution of 1 km $\times$1 km the retrieval allows spatial averaging of data ('superpixelling'). If run at reduced resolution, radiances are averaged to give a mean and variance for use during the retrieval. The fraction of pixels within each superpixel containing cloud is also calculated, to give a cloud fraction product. An advantage of running at reduced resolution is that it cuts computational time by a factor of about n^2 for an n km $\times$ n km superpixelling as compared to running at full resolution. It also decreases the variance on the measurement by a factor of the number of cloud-free pixels within the superpixel, hence providing a way to compensate for measurement noise. However, very large superpixelling may impact the accuracy of the retrieval, given factors such as geometry and surface conditions varying between individual pixels.

Driver file generation

The behaviour of the ORAC retrieval code is controlled by a driver, generated in IDL preprocessing. It is a text file which passes various pieces of information to the FORTRAN code such as file paths, *a priori* and first guess settings, convergence criteria, and active instrumental channels.

3.6 The retrieval algorithm

The ORAC retrieval scheme is an optimal estimation (OE) scheme which uses the Levenburg-Marquardt algorithm. The basic principle of optimal estimation is to maximise the probability of the retrieved state based on the value of the measurements and *a priori* information. Formally, it is required to maximise the conditional probability $P = P(x|y, x_b, b)$ with respect to the values of the state vector x, where x_b is the *a*

priori value of the state vector and b are all other parameters not modelled by the forward model. The full state vector x is formally defined as follows:

$$x = [\tau_{0.55}, r_e, R_{s(0.55,FORWARD)}, R_{s(0.55,NADIR)}]$$

The terms, as previously mentioned, are the optical depth at 0.55 μm, effective radius, and surface albedo at 0.55 μm for forward and nadir viewing geometries. The GlobAEROSOL retrieval lacked the forward view, and so the term involving the forward view surface albedo was not present in x.

3.6.1 Definition of retrieval cost J and method of minimisation

The assumption is made that errors in the measurements, *a priori* and model parameters are Normally distributed with zero mean and variance given by S_y, S_x and S_b respectively. Then, the conditional probability takes on the quadratic form:

$$\begin{aligned} P(x) \propto & \exp[-(y(x)-y_m)S_y^{-1}(y(x)-y_m)^T] \\ & \times \exp[-(x-x_b)S_x^{-1}(x-x_b)^T] \\ & \times \exp[-(b_t-b)S_y^{-1}(b_t-b)^T] \end{aligned} \tag{3.18}$$

The three terms present in the equation represent weighted deviations from measurements, the *a priori* state and the model parameters respectively. b_t has been written for the (unknown) true values of the model parameters. Maximising probability is equivalent to minimising the negative logarithm, so that the quantity J is equivalently minimised where:

$$\begin{aligned} J(x) = & (y(x)-y_m)S_y^{-1}(y(x)-y_m)^T \\ & + (x-x_b)S_x^{-1}(x-x_b)^T \\ & + (b_t-b)S_b^{-1}(b_t-b)^T \end{aligned} \tag{3.19}$$

The minimisation is done with respect to the state variable x, so that the derivative of J is independent of the third term and b therefore cannot be part of the solution. There are many techniques used in the minimisation of J. Essentially any method of finding the minimum is acceptable, with the caveat that in an operational context it must be robust and fast. The particular characteristics of this problem are that:

1. First and second derivatives of J (with respect to x) exist and are continuous.

2. Multiple minima are unlikely.

3. J is likely to be approximately quadratic in the region of the solution and far from quadratic elsewhere.

Characteristic 1 implies descent algorithms that make use of the local gradient are possible, and these are generally faster than methods that do not. 2 implies that excessive domain searches to avoid minor minima are probably not required. 3 is a result of the non-linear nature of the forward model, and means that quick convergence from a poor starting position is unlikely.

The cost function to be minimised is a reduced form of equation 3.19 since there is no explicit dependence on b:

$$\begin{aligned} J(x) &= (y(x)-y_m)S_y^{-1}(y(x)-y_m)^T \\ &\quad + (x-x_b)S_x^{-1}(x-x_b)^T \end{aligned} \tag{3.20}$$

The first and second derivatives of J with respect to x are given by:

$$\begin{aligned} J' &= \frac{\partial J}{\partial x} = K_x^T S_y^{-1}(y(x)-y_m) + S_x^{-1}(x-x_b) \\ J'' &= \frac{\partial^2 J}{\partial x^2} = K_x^T S_y^{-1} K_x + S_x^{-1} \end{aligned} \tag{3.21}$$

The expression for J'' is a commonly-used approximation in that K_x is assumed to be independent of x, i.e. the radiative transfer is linear in x. This is only strictly true near the solution (in the region where J is quadratic) but since J'' is only employed near the solution (as will be discussed in the next section) the approximation is acceptable.

Levenburg-Marquardt algorithm

The minimum is found by starting at some 'first guess' state x_0 (here set equal to the *a priori* value) and then proceeding to make steps δx_n, based on some algorithm. Assuming the value of J decreases at each step, the updated x vector is taking the process towards the cost function minimum. To be consistent with the three points made above, ORAC uses the Levenburg-Marquardt algorithm to define the value of δx_n. This algorithm uses a weighted combination of steepest descent and Newtonian descent according to the characteristic of the cost function.

When the cost function is near quadratic (generally near the solution) the efficiency of the Newtonian scheme is employed, while when the cost function is far from quadratic (generally when far from the solution) the robustness of the steepest descent algorithm is favoured.

The steepest descent algorithm is intuitively the simplest. The vector $-J'$ defines the 'downward' direction of the local steepest gradient. A move $\delta x = -J'$ is almost certainly at least approximately in the direction of the minimum (although it may be too far or hardly far enough). The step is therefore usually scaled, $\delta x = -\alpha J'$ for some variable α. If J is found to be decreasing (so the step is not far enough) α can be increased to move faster; if J increases (meaning the step was too far in this direction) then α is reduced until J decreases. J must eventually decrease with this method otherwise there is an error in the calculation of $\partial J/\partial x$.

Steepest descent can be very slow to converge, especially near the solution where the gradient necessarily becomes small. It is, however, very robust. In contrast, Newtonian descent is very fast near the solution because it will find it in one iteration if J is quadratic. Newton's method finds the root of an equation and is therefore applied here in the form to find the root of $J' = 0$. the Newton step is therefore defined as $\delta x = -J'/J''$. The problem with Newtonian descent is that away from the solution J can be very non-quadratic; the J'' can easily be the 'wrong' sign leading the step taken be away from the solution. This is not a problem scaling can solve.

A combination of both methods as used in this retrieval achieves the best results. Before each step is taken, the resulting cost is checked to see whether it decreases. If so it is taken and the next step is adjusted to be more Newtonian. If instead J increased, the step is not made and an adjustment is made towards more steepest descent. The resulting Levenburg-Marquadt procedure becomes steepest descent far from the solution, and Newtonian iteration close to it. It can be described formally be equation 3.22. Here I is the unit matrix (size $x \times x$) and α the control variable described. When α is large (compared to J'') the step tends to steepest descent; when small the step approximates Newtonian.

$$\delta x = -(J'' + \alpha I)^{-1} J' \tag{3.22}$$

For the initial step, α is set proportional to the average of the diagonals of J'' to obtain a reasonable value:

$$\alpha_0 = MQ_{start} \times \overline{Trace(J'')} \tag{3.23}$$

After a successful step with decreasing J, α is decreased as follows:

$$\alpha_{n+1} = \alpha_n \div MQ_{step} \tag{3.24}$$

Following an unsuccessful step attempt α is increased:

$$\alpha_{n+1} = \alpha_n \times MQ_{step} \tag{3.25}$$

In the ORAC retrieval the control parameters MQ_{start} and MQ_{step} have default values of 0.001 and 10 respectively. These are defined by the driver file and easily changed in the IDL preprocessing code.

Boundaries

The state space in which the solution is to be found is typically bounded by physical constraints and the retrieval needs to ensure that values are not retrieved outside of these boundaries. These bounds are optical depths τ of between 0.01 and 2, effective radii r_e between 0.01 μm and 10 μm, and surface albedoes of between 0 and 1 (for both views).

Convergence criteria

The iteration process proceeds until the decrease in J between consecutive iterations (δJ_n) is so small as to be considered negligible. This is determined as being smaller than a preset value (the default is 0.05). This prevents a parameter which is not well constrained from oscillating or becoming unstable (because the cost function is 'flat' in that direction). Such an instability makes it appear that the solution is not yet found whereas, in fact, the cost is minimal and cannot decrease further. The use of a convergence criterion on δJ_n avoids this problem. If convergence has not occurred after 25 iterative steps, the retrieval is stopped.

3.6.2 Quality control information

This is a very important aspect of any retrieval as there needs to be an effective way of checking the validity of the result. Fortunately, OE provides diagnostics which allow a reasonably strict quality control to be applied. There are three principal diagnostics directly available from the inversion method:

- A check of the 'goodness of fit' of the solution to its constraints. This should identify situations where the scene is not well-modelled by the radiative transfer mixed aerosol layer.
- The error field for the retrieval. Scenes that are modelled successfully but have larger than expected errors can indicate problems with the scene, e.g. missing channels.

- The number of iterations taken to reach the solution. A very high number of iterations may indicate a 'false' convergence (to unrealistic values), or no convergence at all (in the case of over 25 steps).

Model fit

The value of J at the solution indicates whether the solution is good to within the statistical accuracies assumed for measurements and *a priori* data. If, at the solution, none of the measurements deviate from the calculated values (i.e. $y(x) - y_m$) by significantly more than their expected noise values (S_y) and no state variables deviate from their *a priori* values (i.e. $x - x_b$) by significantly more than the *a priori* error (S_x), then J will be of order $n_y + n_x$ (where n_y is the number of measurements used and n_x is the number of state variables retrieved).

Because of the general lack of *a priori* and the bounded state variables there are actually fewer degrees of freedom than $n_y + n_x$. A bounded fraction $(0-1)$ cannot contribute anywhere near its expected S_x value $(\sim \infty)$ to J. Therefore it can be stated that if there are n_b state variables that are bounded but without significant *a priori* then an acceptable solution will have J of order $n_y + n_x - n_b$.

The largest contribution to J is likely to be from measurement fit, meaning it is the final $y(x) - y_m$ that will most often indicate a problem with the solution. The retrieval keeps track of J_y and J_x separately (where the total cost is, from equation 3.20 $J = J_x + J_y$). The general procedure to track down the reason for a high cost would be to check J_x and J_y for measurement or *a priori* misfit and then either $y(x) - y_m$ or $x - x_b$ for the offending measurement or state variable.

Because the value of J depends on the estimation of values for S_y and S_x it is likely that J will not initially be of order $n_y + n_x - n_b$ as expected. Values too low imply an overestimation of (probably) measurement error; values too high imply either underestimation of noise levels or convergence criteria that are too loose.

3.6.3 Linear error analysis

OE allows rigorous error analysis of the retrieved state vector $\hat{x}$. It is based on the assumption that in the vicinity of the solution the radiative transfer model is linear. In this case, it can be shown that the solution for the minimum of the cost function (equation 3.19) is found for:

$$\hat{x} = (S_x^{-1} + K_x^T S_y^{-1} K_x)^{-1} (S_x^{-1} x_b + K_x^T S_y^{-1} y_m) \tag{3.26}$$

We can define an 'inversion operator' $D_y = \partial / \partial y_m$ at the solution, which will be given by:

$$D_y = (S_x^{-1} + K_x^T S_y^{-1} K_x)^{-1} K_x^T S_y^{-1} \tag{3.27}$$

From Rodgers [2000], three error sources can be identified. These are known as 'null space' (or *a priori smoothing*) error, measurement error and model parameter error and are expressed as covariances S_N, S_M, and S_S respectively.

Null space error

This expresses error arising from a lack of information in the measurement system (hence the term 'null space'). It is calculable as follows:

$$S_N = (D_y K_x - I) S_x (D_y K_x - I)^T \tag{3.28}$$

Measurement error

This expresses the mapping of the errors in the measurements onto the solution and is defined by:

$$S_M = D_y S_y D_y^T . \tag{3.29}$$

Model parameter error

This expresses the mapping of the errors in the forward model parameters onto the solution. The term is frequently neglected from error analyses on the grounds that S_b is either small (all parameters are well-known) or that all uncertain parameters are included in the state vector x, and errors are therefore manifest through S_N and S_M. It is also common to include the forward model error into S_y.

$$S_S = (D_y K_b) S_b (D_y K_b)^T \tag{3.30}$$

The total of the null space and measurement errors gives the familiar total error:

$$\hat{S}_T = (S_x^{-1} + K_x^T S_y^{-1} K_x)^{-1} \tag{3.31}$$

All error terms are covariances and the correlations implied by the off-diagonal terms are important for interpreting the information content of the measurements. However, it is unrealistic to output the full covariances for all three error sources in an operational code environment and these are only available as special breakpoint variables for diagnosis. The ORAC nominal output contains only the square root of the

diagonals for the total error, S_T (or optionally $S_T + S_S$). The folllowing points are worth noting about the error estimate:

1. It applies to each state parameter separately–a high expected error in one parameter does not necessarily imply a high error in another.
2. It is conditional upon the correct solution being found.
3. It is conditional upon atmospheric and aerosol conditions corresponding to the model assumed.

3.6.4 Defining the statistical constraints

The statistical constraints on the solution $\hat{x}$ are the error covariance matrices S_x and S_y, i.e the quality of the measurements and *a priori* values of the state. The error covariance matrix S_b defines the contribution of the model parameter error to the retrieval error and does not affect the behaviour of the retrieval.

3.7 Speciation and calculation of the Ångstrøm coefficient

As described the ORAC retrieval scheme determines the optical depth at two wavelengths and effective radius for a given aerosol model (plus surface albedo information) which best reproduce the observed radiance. The GlobAEROSOL project requires two additional products, which are also calculated by this retrieval: the Ångstrøm coefficient and speciation of the aerosol. The Ångstrøm coefficient is defined by:

$$\frac{\mathrm{d}\log[\tau(\lambda)]}{\mathrm{d}\log[\lambda]} \tag{3.32}$$

By assuming that τ varies linearly with λ in the visible, it can be directly calculated from the two retrieved optical depths:

$$\text{Å}_{0.55,0.865} = \frac{\log[\tau_{0.865}/\tau_{0.55}]}{\log[0.865/0.55]} \tag{3.33}$$

For GlobAEROSOL a retrieval is performed for each of the aerosol classes, and speciation is determined by the class which provides the best fit to the observed radiances. For single-view retrievals there has been concern (Marsh et al. [2004]) about large errors arising from insufficient information content from the measurements, particularly due to uncertainties in the surface spectral shape. It may be possible to provide some constraint to the variation of speciation across adjacent pixels, especially over the ocean (where both aerosol properties and the surface reflectance can be expected to be reasonably homogeneous over a few tens of kilometers). This has not currently been investigated.

Chapter 4

Progress and Results To Date

4.1 Development of the forward-view retrieval

As has been previously mentioned, the aerosol retrieval present in the GlobAEROSOL scheme makes use of only the nadir view channels on the ATSR-2 and AATSR instruments. The initial work performed in the research project involved adapting the code to allow it to make use of data from the forward view of the dual-viewing geometry, which involved making changes to the IDL preprocessing steps of the scheme.

A pair of scenes were analysed to reveal similarities and differences in the retrieved data. The first uses data from the orbit on September 5^{th} 2004 between $28^o - 64^o$ north and $16^o - 42^o$ west, and the second August 28^{th} 2004 from $8^o - 40^o$ south and $108^o - 120^o$ east. Figure 4.1 shows the geographical location of these two scenes, as well as radiances received at 0.55 μm at the nadir view.

Due to these locations being mostly at sea, the marine aerosol class lookup tables were used during the retrievals. Retrievals were run with pixel latitude and longitude dimensions of 10 km to provide a reasonable number of data points and decrease the measurement noise through superpixelling, but without requiring excessive computational time to run. The scenes were chosen for the following reasons:

- They are comparatively cloud-free, yielding a high number of pixels to perform the retrieval on.
- They are largely over the ocean. Being comparatively homogeneous, the ocean surface is in principle easier to describe than the land. Therefore by attempting retrievals above the ocean we know our description of each pixel is of a similar quality.
- One is in the northern hemisphere (the Atlantic Ocean, to the northwest of the United Kingdom) while the other is in the southern (clipping over western Australia and Indonesia). This provides two

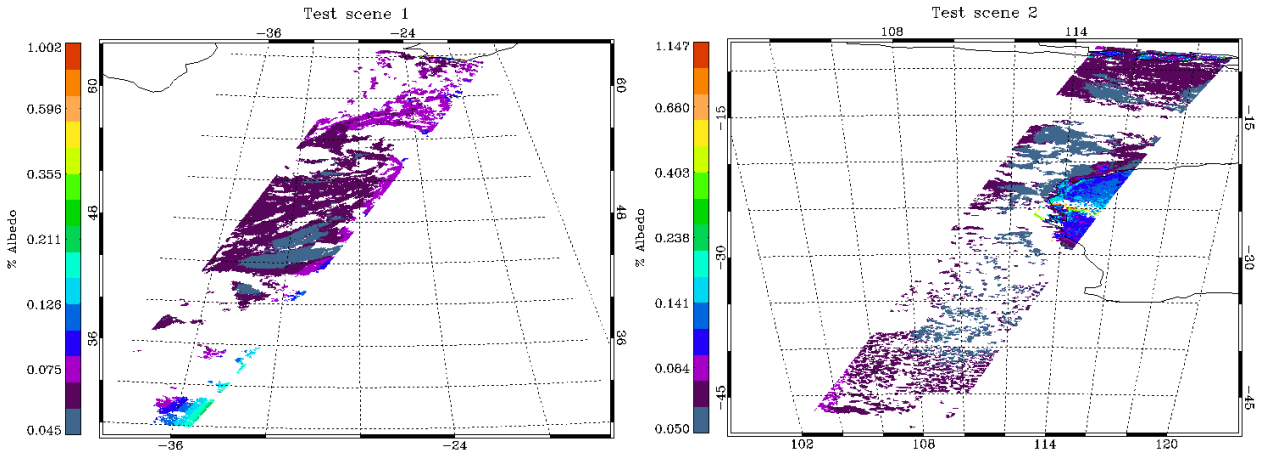


Figure 4.1: Radiances received at 0.55 μm at nadir view for the two scenes compared. The satellite's native units are of apparant albedo. Missing data points from the swathe are due to cloud-flagging. Note the bright streak in the southwest of test scene 1; this is the sun-glint region.

geometrically different locations to probe.

Retrieved aerosol optical properties for forward and nadir views would be expected to match within error bars, assuming the state of the atmosphere does not change much in the few minutes between recording of forward and nadir view data. Some discrepancy would be encountered because the aerosol is in the atmosphere and not flat on the surface (so as a result will appear in slightly different locations in the different views), though if the bulk of the aerosol is low in the atmosphere, and variability on horizontal scales of several kilometres is not too great, then this discrepancy will be minimal. The forward viewing geometry would be expected to give smaller error bars on retrieved aerosol properties, as its longer atmospheric path length means that it has a greater signal from aerosol than the nadir view.

Retrieved surface albedo would be expected to differ because true surfaces tend not to be Lambertian. One would, however, expect strong correlations in retrieved albedo: a comparatively bright surface is likely to be bright at both viewing geometries.

Data points were filtered to consider for the analysis only those meeting the following criteria:

- The retrieval converged, with a cost less than 20 (so represents a good fit).
- The pixel is a sea pixel.
- Retrieved quantities do not lie exactly on state boundaries.

A note on cost

In the GlobAEROSOL retrieval scheme, retrieval cost is divided by the number of active channels (4 for their single-view nadir retrievals) to provide an effective cost-per-channel measure of the success of the retrieval. Convergence criteria are scaled by a corresponding amount. That convention has been adhered to here so should be borne in mind when discussing results. It does not impact much when comparing single-view retrievals, as both have the same number of data points and retrieved parameters, but will be an issue when comparing single-view with dual-view retrievals.

4.1.1 Retrieved aerosol properties

Comparative optical depths and effective radii for the two test scenes are shown in Figures 4.2 and 4.3 respectively. The correlations between values, especially for optical depth in the second scene, are unfavourable although this is not necessarily indicative of poor agreement. The ranges of τ and r_e retrieved are fairly small, and error bars significant within this range, so a low correlation on these scales is not too surprising.

The bottom pair of each set of four graphs considers the ratio of the difference in the retrieved parameters to the total uncertainty on them (i.e. the ratio of $\tau_f - \tau_n$ to $S_{\tau,f} + S_{\tau,n}$ for retrieved τ with error S_f for the forward view and S_n for the nadir). This provides a useful measure of how frequently the retrieved quantities agree with each other: this error-normalised difference being less than 1 implies that, within error bounds, the retrievals agree. This simple measure makes the approximation that any errors in retrieved optical depth on the forward and nadir views will not be correlated.

For both scenes and parameters (though more so in the second test scene), a large proportion of pixels have a normalised difference score of less than 1 as shown by the red lines in the figures. We can say, then, that there is reasonable agreement between the forward and nadir view retrieved aerosol properties. On the flip side, this is perhaps due more to the size of the error bars than any to strong correlation in the results (particularly noticeable with effective radius), and there are some clear discrepancies. The salient points are that the range of retrieved values is not large and, although the agreement between the views is questionable on an individual pixel basis, both views retrieve aerosol properties in similar ranges and with similar precision.

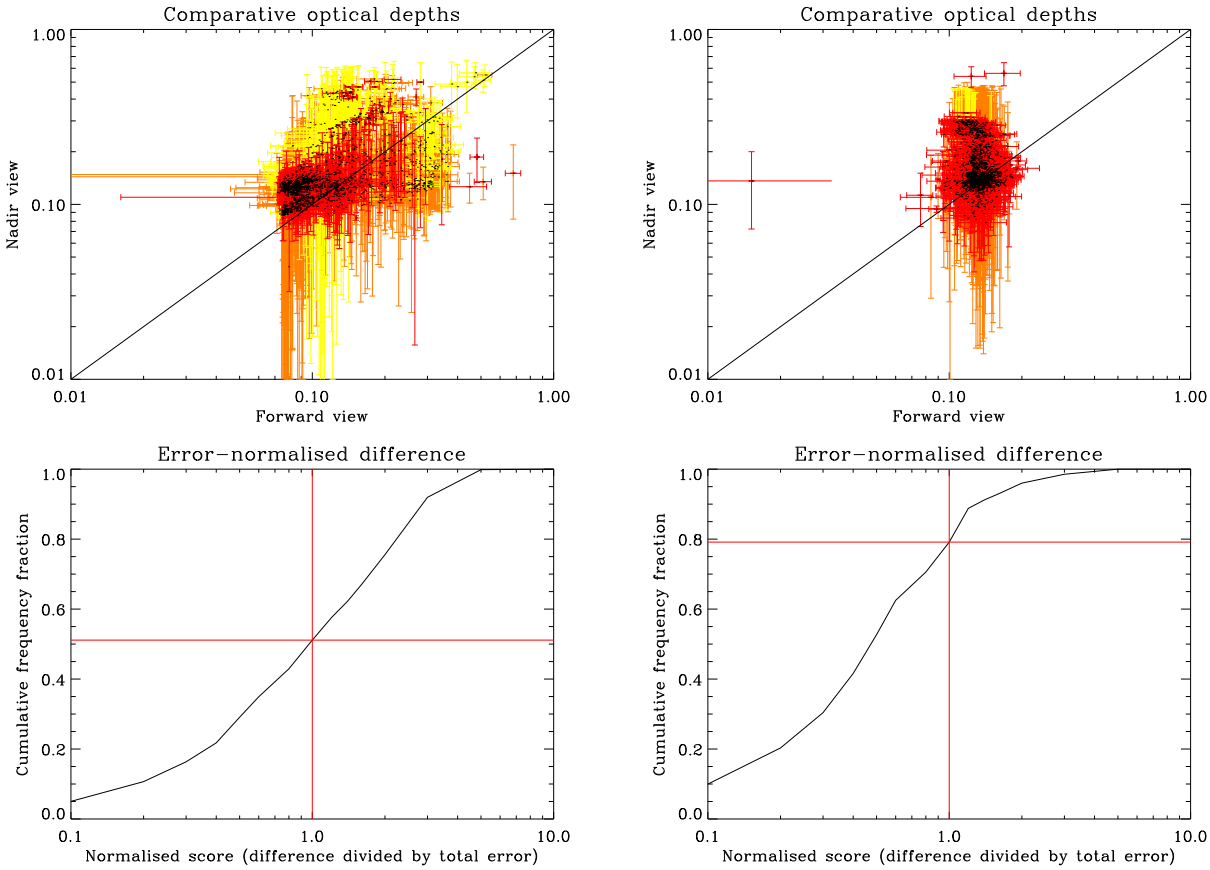


Figure 4.2: Retrieved optical depths τ at 0.55 μm for forward and nadir views for the first (left two graphs) and second (right two) test scenes. *Top:* τ compared pixel-by-pixel. Coloured lines are error bars. Yellow indicates cost $J < 1$, orange $1 < J < 5$ and red $5 < J < 10$. The Spearman correlation coefficient r is 0.524699 for the left and -0.085780 for the right. *Bottom:* Difference between forward and nadir retrieved τ divided by the sum of their estimated errors.

4.1.2 Retrieved surface albedo

Retrieved surface albedo at 0.55 μm is compared in Figure 4.4. The normalised difference score described for aerosol optical properties was not calculated as one would not expect the surface albedo to match between the views' differing geometries. The correlation between the views is small but again reasonable considering the magnitude of the error bars. Most retrieved values for sea albedo are from around 0.1% to 1-2%, which seems fairly low compared to model estimates of sea surface albedo which will be discussed in Section 4.4.

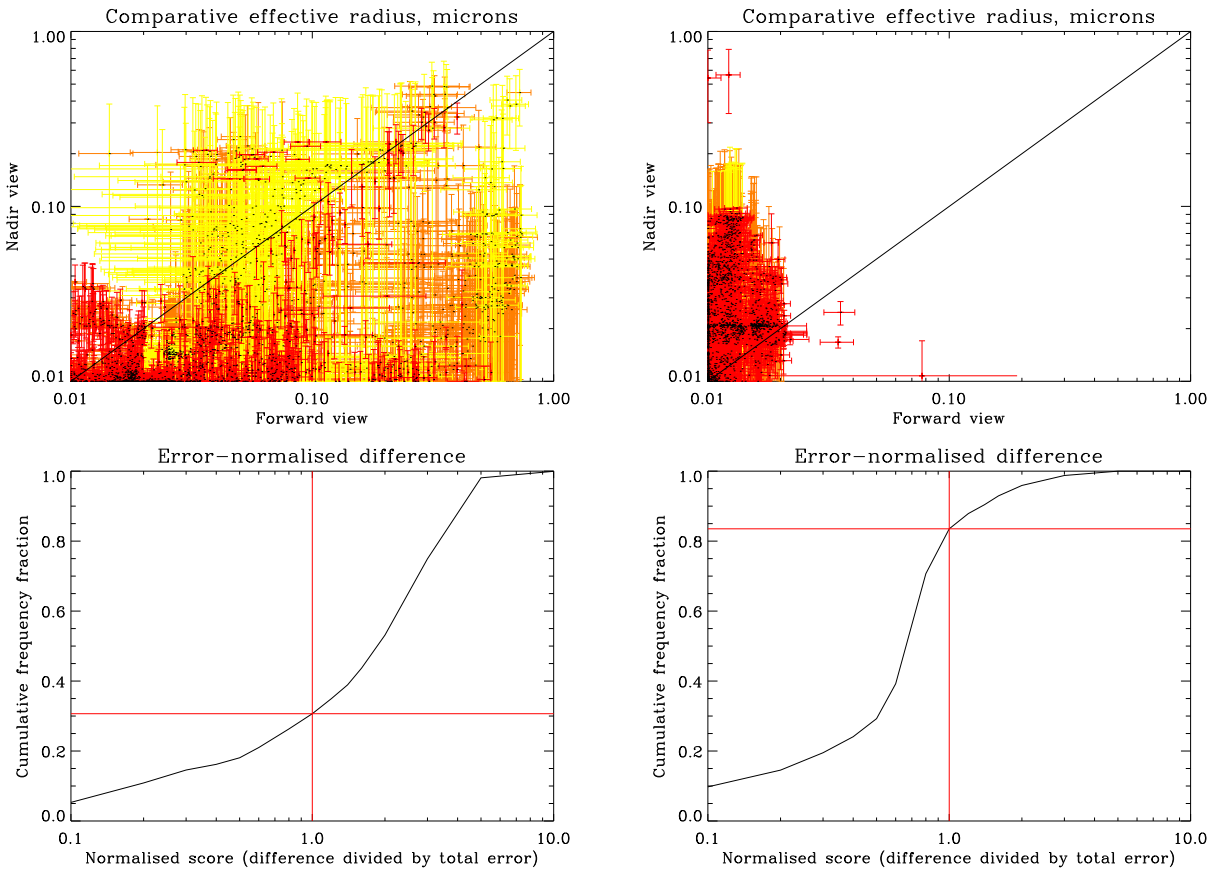


Figure 4.3: Retrieved effective radius r_e for forward and nadir views for the first (left two graphs) and second (right two) test scenes. *Top:* r_e compared pixel-by-pixel. Coloured lines are error bars. Yellow indicates cost $J < 1$, orange $1 < J < 5$ and red $5 < J < 10$. The Spearman correlation coefficient r is 0.546713 for the left and 0.257511 for the right. *Bottom:* Difference between forward and nadir retrieved r_e divided by the sum of their estimated errors.

4.1.3 Retrieval costs

Figure 4.5 shows that while the forward view performs better in the first test scene, the nadir performs better in the second. As previously mentioned, the forward view's longer path means the forward view would be expected to perform better. This suggests perhaps some problem with the forward model under certain geometric conditions. It is interesting to note that the *a priori* contribution to cost remains in the region of 1-1.5 where the total cost is greater than this, for both views in both scenes. This suggests that future improvements to the *a priori* information, particularly the surface albedo (which is treated as having the smallest *a priori* error) could mean a lower overall cost.

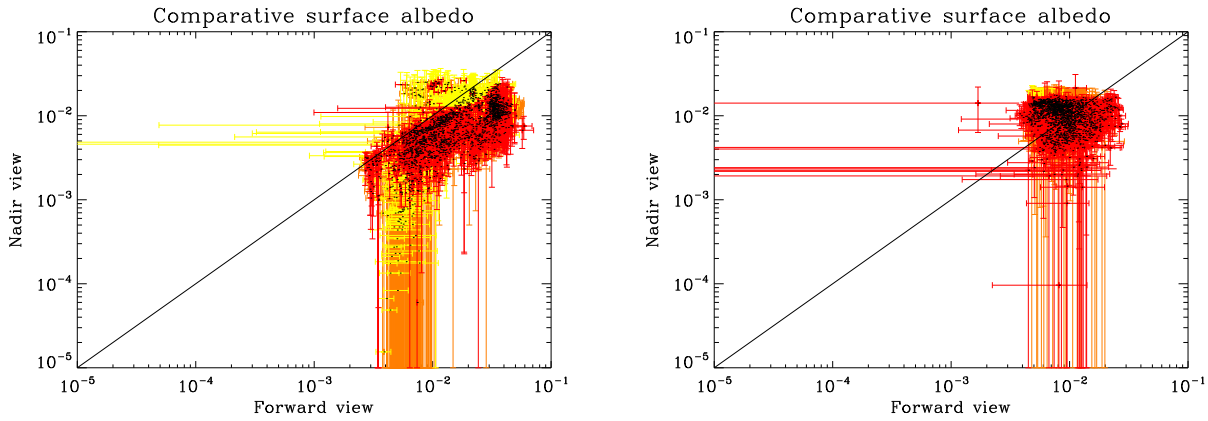


Figure 4.4: Retrieved surface albedo R_s at 0.55 μm for forward and nadir views for the first (left) and second (right) test scenes. Coloured lines are error bars. Yellow indicates cost $J < 1$, orange $1 < J < 5$ and red $5 < J < 10$. The Spearman correlation coefficient r_s is 0.709167 for the left and -0.190553 for the right.

4.1.4 Constraining surface albedo

The surface albedo can be effectively constrained in a retrieval, by simply setting the uncertainty on the *a priori* value to be very small. This essentially forces the retrieval to use the *a priori* information, as iterating away from this value for surface albedo will yield a step of prohibitively high cost which will then be rejected.

If the surface model were correct then decreasing the *a priori* error in this way would enable a more accurate elucidation of the aerosol contribution to observed radiance, and decrease the uncertainty on retrieved parameters. With an imperfect surface model, however, we might expect an increase in cost as the retrieval becomes unable to find a solution matching the observed radiance by merely changing optical depth and effective radius.

Retrievals have been performed for both test scenes in which surface albedo was tightly constrained to the *a priori* value (with an error of 10^{-4} instead of the standard 10^{-2}). For the majority of pixels the cost was found to increase, and the agreement between nadir and forward-view retrievals worsened. These results are not presented here for brevity. This is an indication that a better model of the sea surface is required to improve the quality of the retrieval.

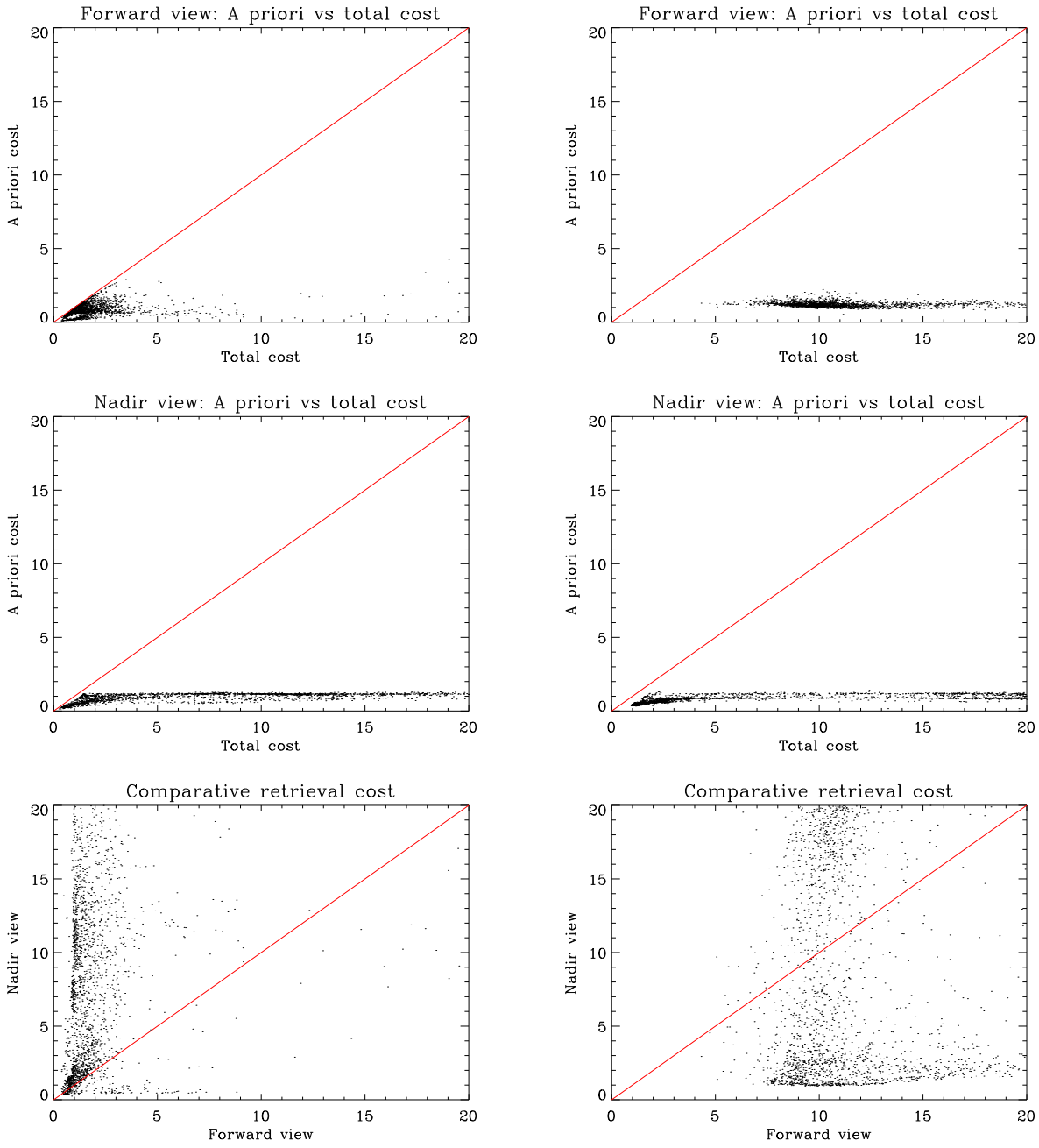


Figure 4.5: Retrieval costs J for the first (left three graphs) and second (right three graphs) test scenes. *Top:* The *a priori* contribution to total cost for forward view retrievals. *Middle:* The *a priori* contribution to total cost for nadir view retrievals. *Bottom:* Comparative forward and nadir retrieval costs.

4.2 Development of the dual-view retrieval

The second stage of progress so far has consisted of developing the retrieval code so that data from both the forward and nadir views may be utilised, involving modifications to both the IDL preprocessing code and the ORAC code written in FORTRAN. This means that instead of using 4 measurements and retrieving 3 parameters, the retrieval now works as described in the preceding chapter and uses 8 measurements to retrieve 4 parameters. In principle, this increased number of measurements should mean smaller errors on retrieved quantities. One would also expect the retrieved quantities themselves to lie somewhere in between those values returned from single-view retrievals, with a bias toward whichever single-view retrieval performed better. Both test scenes previously compared were run with the dual-view retrieval code and analysed with the same constraints on pixel quality.

4.2.1 Retrieved aerosol properties

Figure 4.6 compares retrieved optical depth τ and effective radius r_e between single and dual-view retrievals. The left-hand portions of the figure compare the differences between dual-view and both single-view retrievals: so if the retrieved parameter for the dual view lies in between that of the single-view retrieved values, points in the graph will fall in the top-left and bottom-right quadrants.

Although there are points in these regions there are a considerable number in other portions of the graphs. This is not necessarily indicative of a problem in the dual-view retrieval as many points lie close to the lines $x = 0$ or $y = 0$ and so, within error bars, could lie in the expected quadrants. There are very few points where dual-view retrieved properties differ considerably from both single-view retrieved estimates. The very different distribution of points among the graphs from the two test scenes further indicate possible geometric effects leading to particular biases either in the forward model or surface spectral shape.

Error estimates are also shown on the right of Figure 4.6. In all cases, the dual-view retrieval tends to show smaller error bars than the nadir retrieval. In many cases (although not with such a majority) the same is true when comparing dual-view and forward retrievals. Smaller error bars would be expected due to the larger number of measurements used in the dual-view retrieval, so larger errors on the dual-view retrieval indicate a worse fitting of the data. This is not entirely unexpected: Section 4.1 showed that forward and nadir retrievals did not always retrieve values which agreed within error bars, so in these cases the dual-view retrieval might be expected to find some compromise solution fitting both views' sets of measurements reasonably, as opposed to one very well and the other poorly.

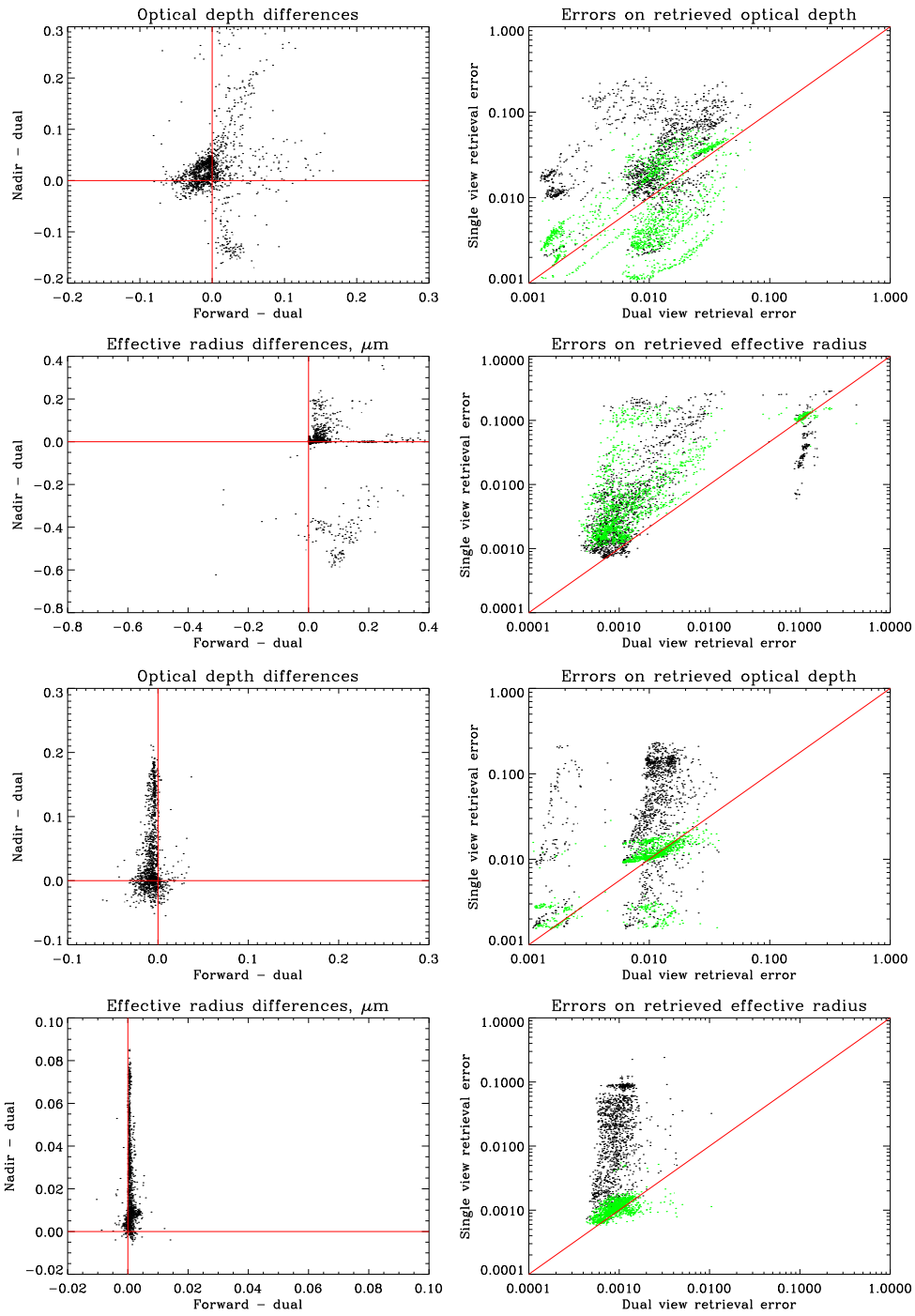


Figure 4.6: Comparative aerosol properties for the first (top four) and second (bottom four) test scenes. *Left:* Differences between dual-view and single-view τ (top) and r_e (bottom). *Right:* errors on retrieved τ (top) and r_e (bottom), with dual/nadir comparisons in black and dual/forward in green.

4.2.2 Retrieved surface albedo

In contrast to aerosol properties, the surface albedo is considered to be independent betweem viewing geometries so one would expect a strong correlation between estimates from single-view and dual-view retrievals. Differences in retrieved values would be due to the retrievals' differing estimates of aerosol properties. Figure 4.7 shows that such strong correlations are indeed observed. There do not appear to be any particularly strong biases evident in the retrievals (save perhaps the dual-view forward albedo being higher than the single-view for the first test scene), although this does not necessarily tell us much as the data were expected to be in good agreement.

Comparative error estimates are also favourable, with the dual-view retrieval performing considerably better than the nadir-only retrieval and slightly better than the forward-only. Errors are still not always small compared to the magnitude of the albedo, being of the order of 10^{-4} to 10^{-2} for albedos on the order of 10^{-3} to 10^{-1}.

4.2.3 Retrieval costs

As previously mentioned, the GlobAEROSOL retrieval divides total cost by the number of active channels to provide a cost-per-channel measurement. When comparing single and dual-view retrievals this is no longer appropriate, as there both the number of measurements being used and number of retrieved parameters differ. In this situation χ^2 provides a more useful comparative measure of cost. This is found by dividing the cost J by the number of degrees of freedom (number of measurements - number of state parameters retrieved) instead of the number of measurements. To convert from the per-measurement cost to χ^2, then, single-view costs increase by a factor of 4 (multiply by 4 measurements and divide by 1 degree of freedom) while dual-view retrieval costs increase by a factor of 2 (multiply by 8 measurements and divide by $8-4=4$ degrees of freedom). The χ^2 measure of cost will be used in discussion here.

Figure 4.8 shows that the performance of the dual-view retrieval is comparable to that of the single-view retrievals in terms of how well it fits the data. In the first test scene it tends to perform better than the nadir retrieval and comparably to the forward, while in the second this trend is reversed. It is worth noting that if the cost-per-measurement value J were applied, the dual-view retrieval would in fact seem to be performing poorly compared to the single-view retrievals, meaning it fits the data less well, which might be expected due to the disagreement between forward and nadir view retrievals.

The comparable performance of the dual and single-view retrievals in terms of χ^2 is encouraging, in

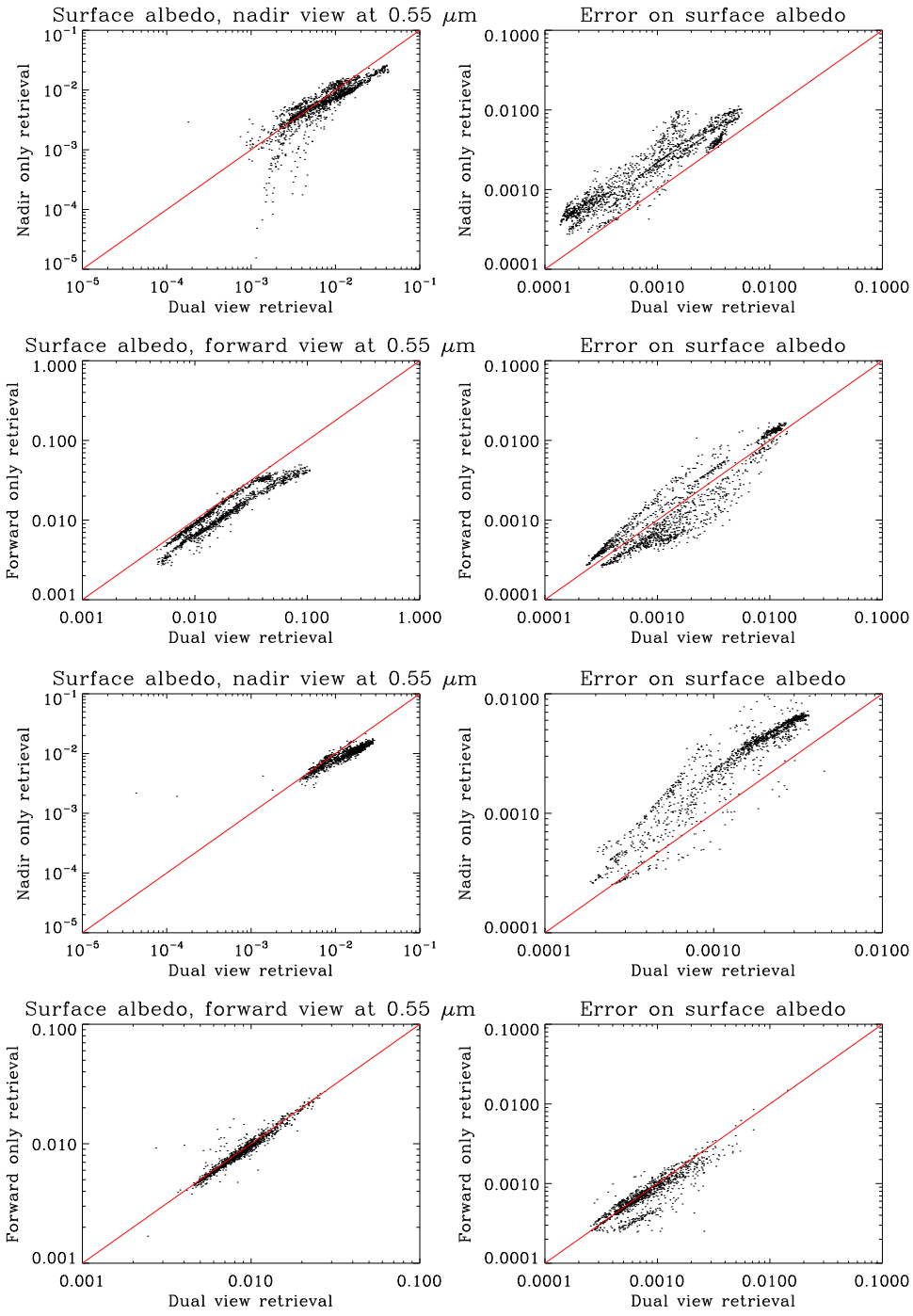


Figure 4.7: Comparative retrieved 0.55 μm surface albedo for the first (top four) and second (bottom four) test scenes. *Left:* Dual view retrieval against single view retrieval for nadir (top) and forward (bottom) geometries. For the first scene, Spearman correlation coefficients are 0.879977 (nadir/dual) and 0.969733 (forward/dual). For the second they are 0.941172 (nadir/dual) and 0.960581 (forward/dual). *Right:* Comparative error estimates for these retrievals.

that it means that if sources of error can be identified and eliminated which lead to single-view forward and nadir retrievals being in better agreement, the dual-view retrieval will benefit comparably more because of its larger number of measurements. Finally, it is worth noting that the *a priori* contribution to cost is comparable to the proportion found for single-view retrievals (Figure 4.5).

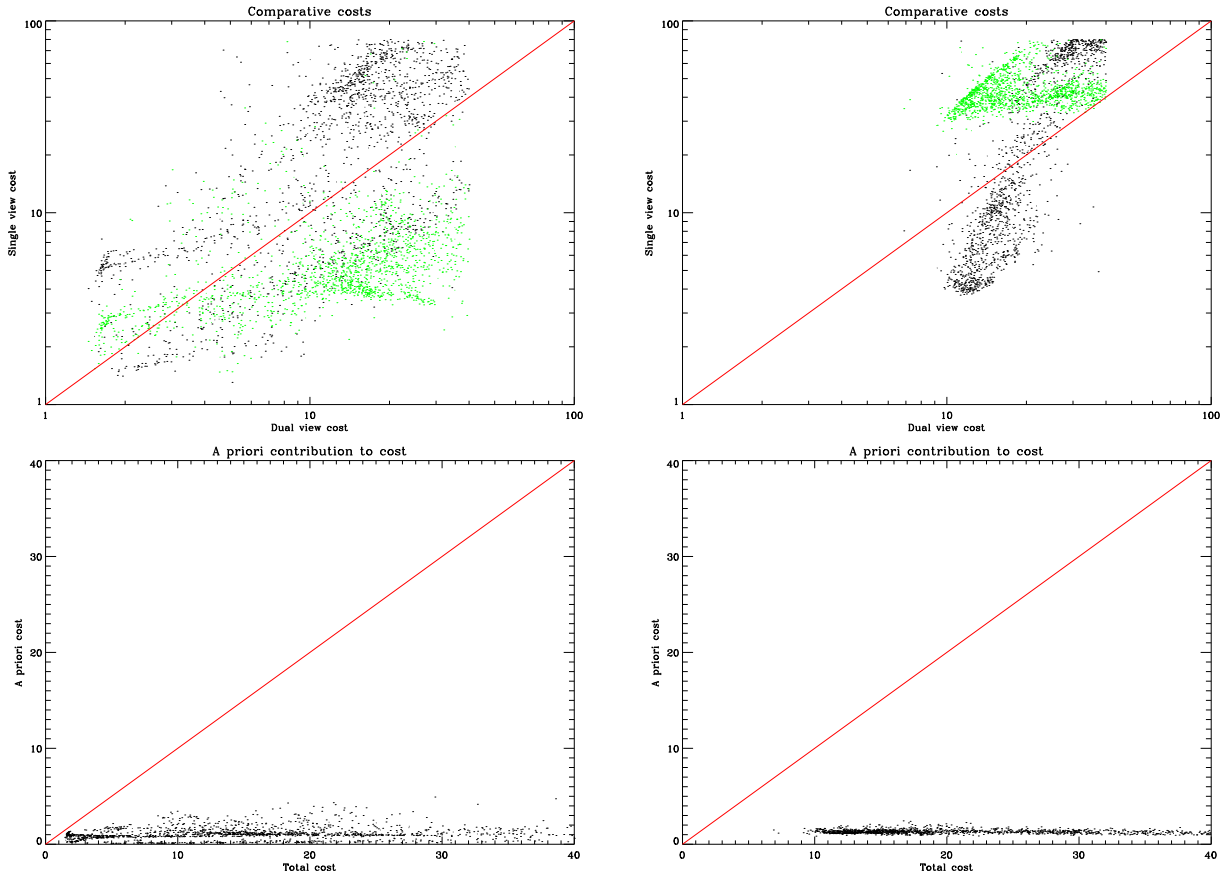


Figure 4.8: Comparative χ^2 costs (see text) for the first (left) and second (right) test scenes. *Top:* Dual-view cost against single-view cost for nadir (black) and forward-view (green) retrievals. *Bottom:* The *a priori* contribution to cost for the dual-view retrieval.

4.3 Discussion of residuals in forward model-predicted radiance

Analysis of the distribution of residuals can provide important insights into the performance of the forward model and retrieval algorithm. In the case of no forward model error (so all error coming from the instrumental noise), one would expect the residuals in the data to have Normal distributions with a mean of 0 and a standard deviation equal to the instrument's noise. A non-Normal distribution would indicate the retrieval

is not finding solutions (or the convergence criteria are too loose); an excessive width, that there is large error in the modelling; and a persistent bias, that there is some persistent bias in the modelling.

Residuals for the two test scenes for single-view and dual-view retrievals for both viewing geometries are shown in Figures 4.9, 4.10, 4.11 and 4.12. Several points are worth noting about these distributions:

- Distributions are, by and large, Normal with only one modal value. This is encouraging and gives confidence that the retrieval algorithm is working and finding a solution.
- In all cases the standard deviation of the distribution of residuals is considerably smaller than the noise on the instrumental channel, indicated by red lines. This is due to superpixelling reducing the measurement error as described in Section 3.5, up to tenfold for a completely cloud-free pixel.
- Although less evident for the forward view in Figure 4.9, residuals tend to show biases consistent between views and scenes. Specifically, the 0.55 μm and 0.67 μm channels tend to have noticeable negative and positive residuals respectfully. This is indicative of some problem in either the forward model or sea surface spectral shape.
- The dual-view residuals look worse than single-view residuals (in terms of distribution shape and spread): this is to be expected due to disagreements between the results of forward and nadir single-view retrievals.

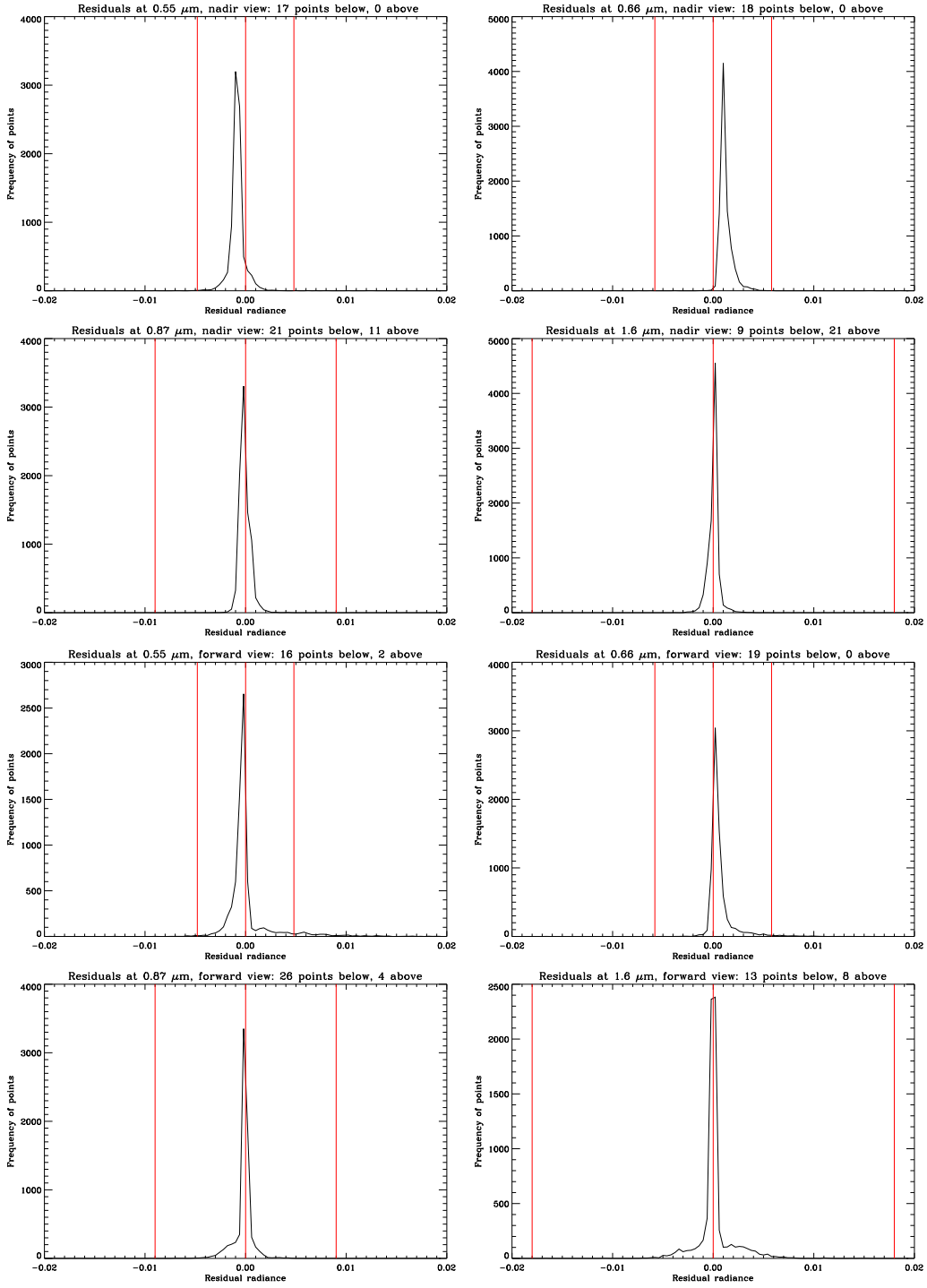


Figure 4.9: Residuals on radiance for the instrumental channels used, for single-view nadir (top four graphs) and forward (bottom four graphs) retrievals for the first test scene. Red lines indicate the channel's noise sensitivity. The number of points falling above and below these ranges is noted in each graph's title.

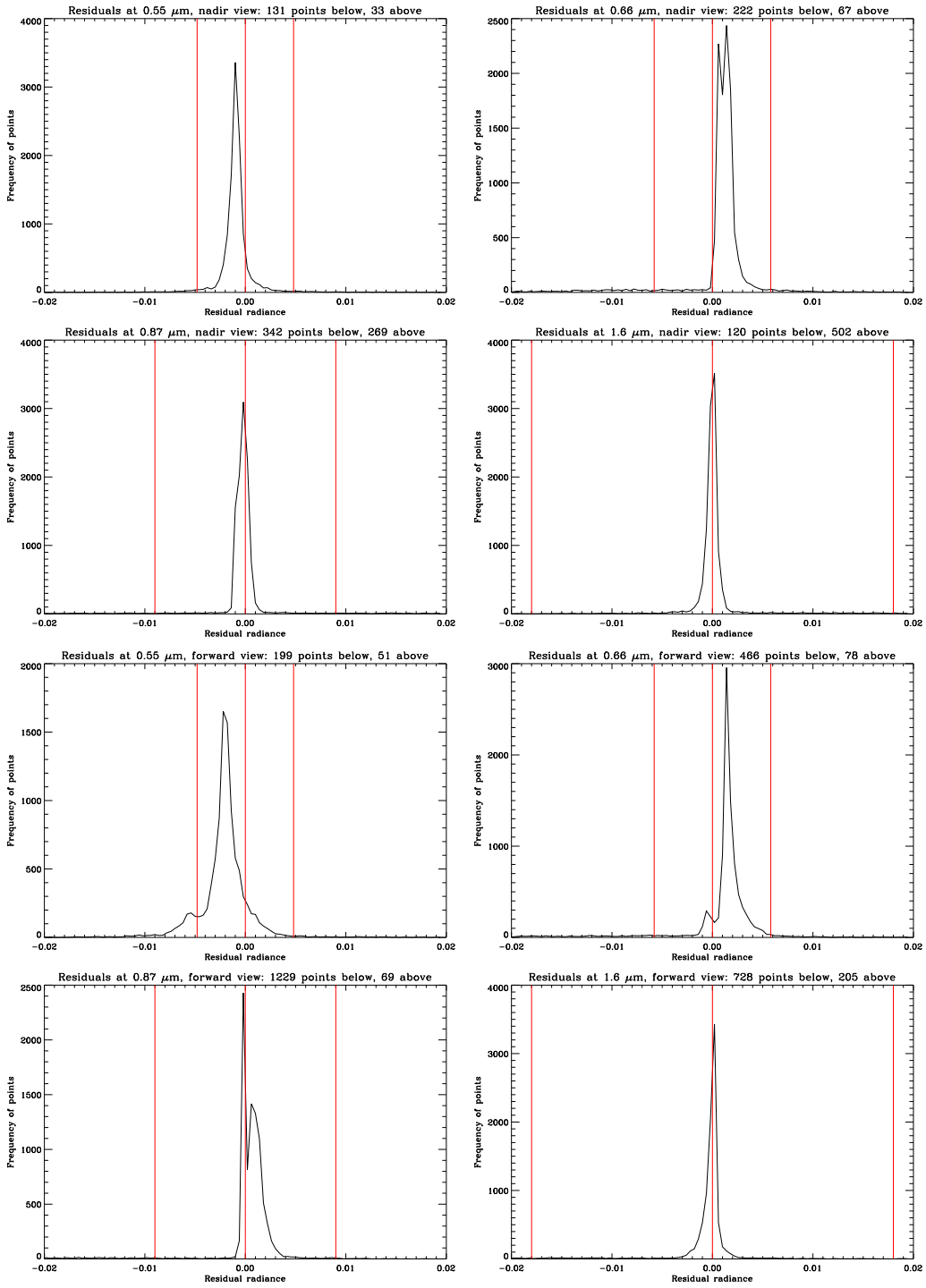


Figure 4.10: Residuals on radiance for the instrumental channels used, for single-view nadir (top four graphs) and forward (bottom four graphs) retrievals for the second test scene. Red lines indicate the channel's noise sensitivity. The number of points falling above and below these ranges is noted in each graph's title.

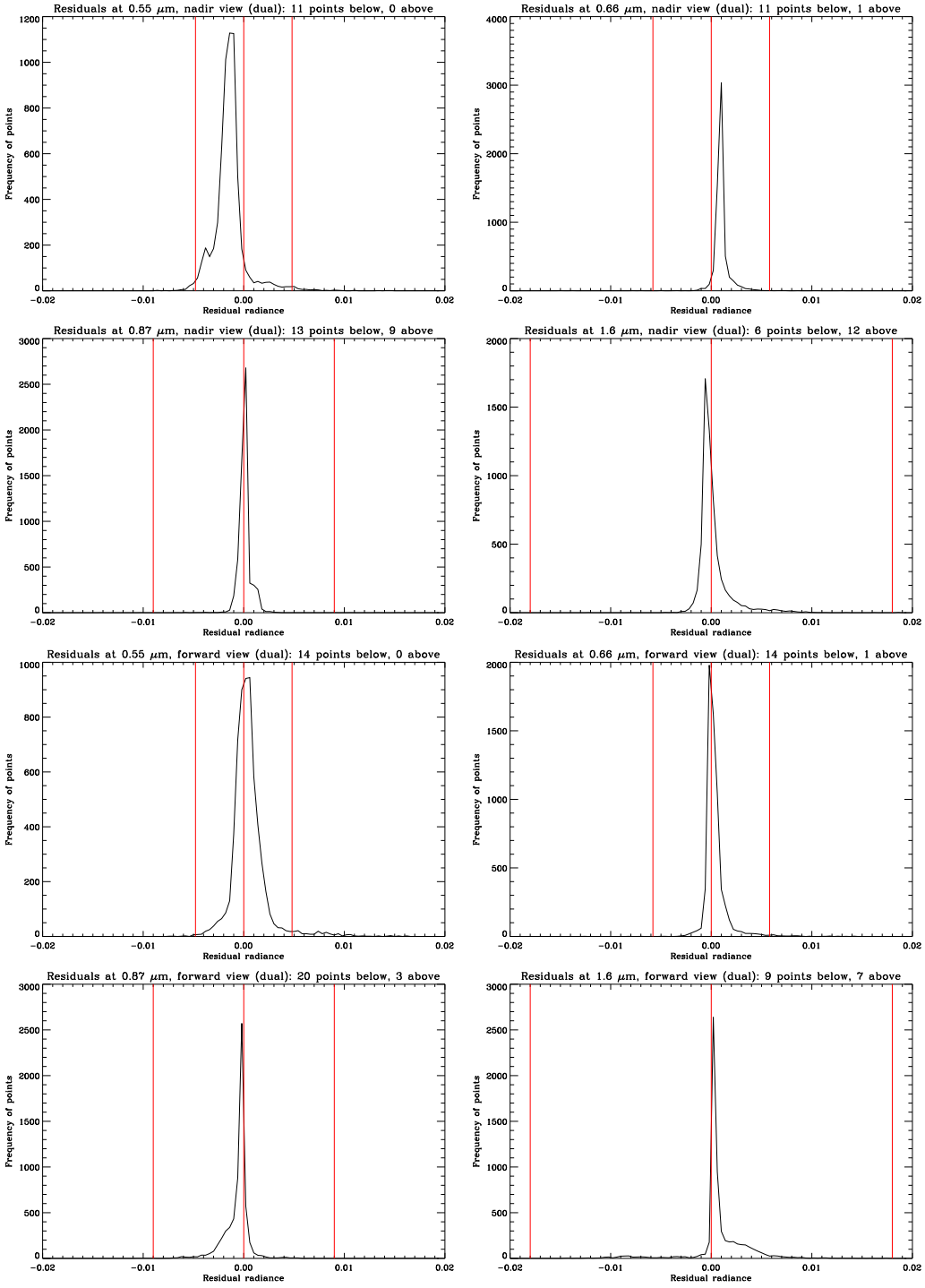


Figure 4.11: Residuals on radiance for the instrumental channels used, for dual-view nadir (top four graphs) and forward (bottom four graphs) retrievals for the first test scene. Red lines indicate the channel's noise sensitivity. The number of points falling above and below these ranges is noted in each graph's title.

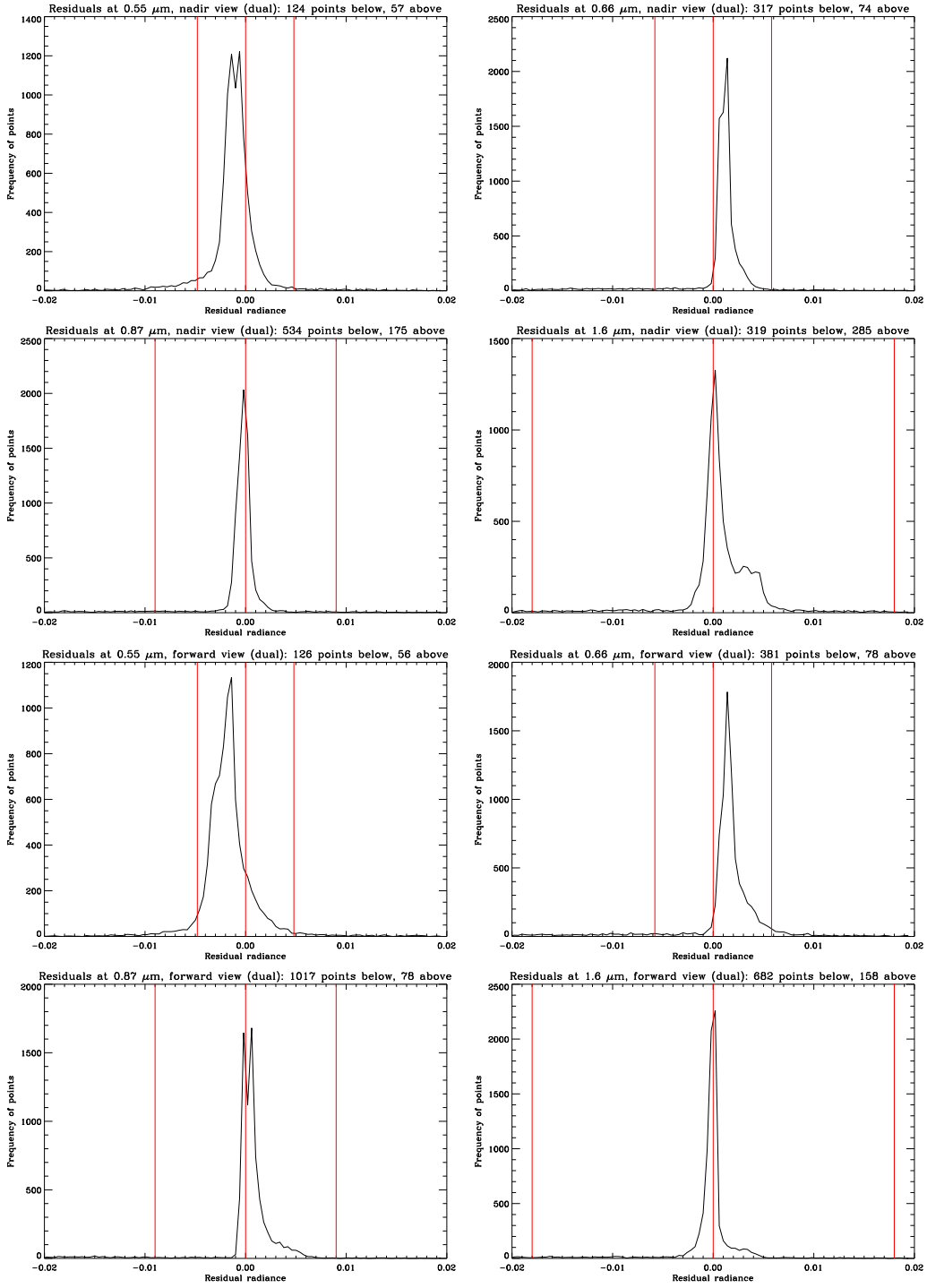


Figure 4.12: Residuals on radiance for the instrumental channels used, for dual-view nadir (top four graphs) and forward (bottom four graphs) retrievals for the second test scene. Red lines indicate the channel's noise sensitivity. The number of points falling above and below these ranges is noted in each graph's title.

4.4 Sea surface models

There is scope for development of new models of the sea surface to try to obtain a more accurate spectral shape, and thus hopefully remove one source of bias affecting the residuals. There is some concern over the fact that some very low values for the sea albedo at 0.55 μm are being retrieved (of the order of 10^{-4} or lower), when it has been well documented (e.g. from Cox and Munk [1954*a*] onwards) that typical values are of the order of a few percent or higher.

4.4.1 Reflection from a smooth surface: the Fresnel equation

Perhaps the simplest model of the spectral shape of the ocean surface (aside from assuming it reflects in equal proportions at all wavelengths) can be obtained from the Fresnel equations. These describe the intensity of light seen by an observer resulting from reflection at the boundary between two uniform and flat surfaces receiving diffuse illumination. This depends on the angle of incident light θ_i (equal and opposite to the observer's viewing angle θ_r), the angle θ_t which the incident wave is refracted to in the water and the real parts of the refractive indices n_{air} and n_{water} of our two mediums. Assuming the incident light is unpolarised, the fraction of reflected light (i.e. the albedo of the surface) is given as follows:

$$R = \frac{1}{2}\left(\left[\frac{\sin(\theta_i - \theta_t)}{\sin(\theta_i + \theta_t)}\right]^2 + \left[\frac{\tan(\theta_i - \theta_t)}{\tan(\theta_i + \theta_t)}\right]^2\right) \tag{4.1}$$

The angle θ_t of the refracted beam transmitted through the water can be obtained using Snell's Law, that $n_{air}\sin\theta_t = n_{water}\sin\theta_i$. As refractive index depends on wavelength, we obtain different variation of albedo with geometry at different wavelengths (and hence a spectral shape changing with viewing geometry) as shown in the top half of Figure 4.13.

According to the Fresnel model, there is little change in albedo or spectral shape at nadir satellite zenith angles (up to about 22^o). Albedo here is of the order of 2%, which provides a rough estimate of the magnitude of the value we would expect to retrieve. At the forward viewing geometry (near $\theta_r = 55^o$) the albedo is slightly higher and the spectral shape different. This result shows that, even in a simple model, it is important to consider the effects of viewing geometry on spectral shape. Predictions of a higher albedo in the forward view tally with observed retrieved quantities (for example in Figures 4.4 and 4.7).

Sidran [1981] used a model based on the Fresnel equations which makes use of the imaginary parts of the complex refractive index. For a flat surface, this gives the albedos and spectral shape shown in the bottom half of Figure 4.13. These show the same overall pattern of variation of albedo with angle, although

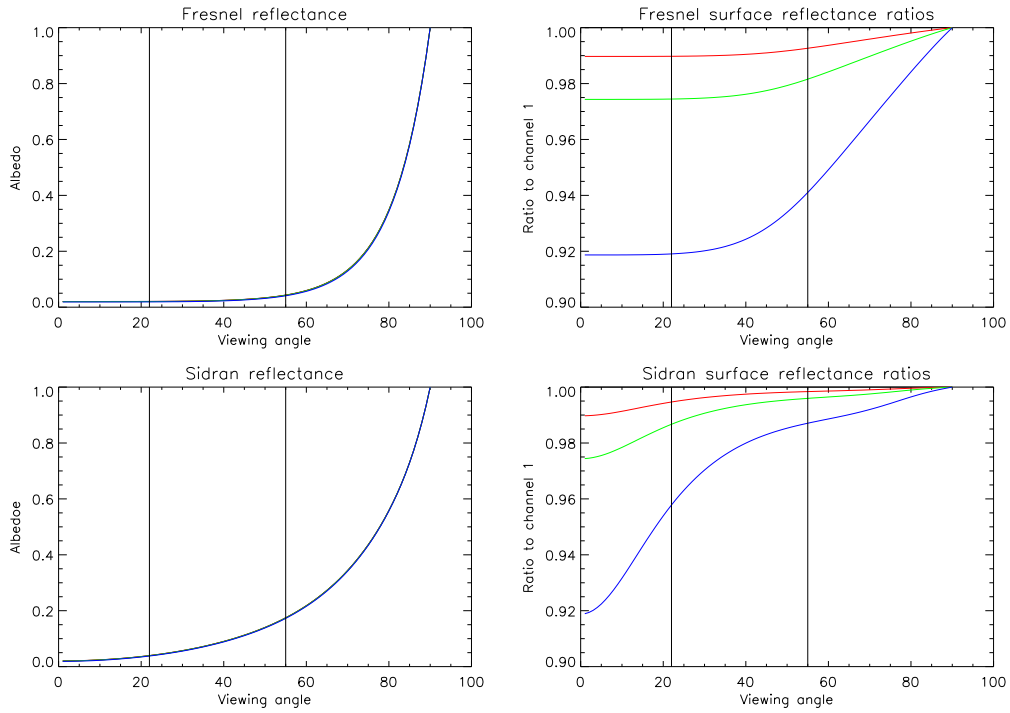


Figure 4.13: *Left:* Predicted sea surface albedo according to the Fresnel (top) and Sidran (bottom) models for a flat surface. *Right:* Ratios of channel albedo to 0.55 μm albedo for Fresnel (top) and Sidran (bottom) models. In all graphs, 0.55 μm values are in black, 0.66 μm in red, 0.87 μm in green and 1.6 μm in blue.

the increase is noticeably steeper and spectral shape varies more dramatically. Retrieved forward surface albedo is not as high as predicted by this model.

Retrievals have been performed using both the flat Fresnel and Sidran models to set the *a priori* surface albedo and constrain spectral shape. In terms of the overall cost J and the distribution of residuals, both of these models performed less well than the model currently applied in the retrieval scheme, described in Section 3.3.2. A more complex model, then, is needed to describe the sea surface accurately.

4.4.2 Wind-ruffled surfaces

In their 1994 paper Sathe and Vaithiyanathan [1994] used the statistics derived by Cox and Munk to relate wind speed to sea slopes (1954*a*, 1954*b*) in combination with the Fresnel equations to develop a model for surface albedo, taking into account both reflection from slopes (which would make a different angle from their normal to an observer than a flat sea, effectively changing θ_i and θ_r) and foam generated as a result of that wind.

Fresnel albedo is calculated for a range of effective viewing angles, and related to the overall albedo seen by considering the probability of a certain slope angle given the wind speed at the location. For an area, as opposed to a single point on the surface, this probability is related to the fractional cover of that slope angle in the area. A small additional term due to foam reflectivity is added, though it was noted that estimates both of foam reflectivity and dependence of foam cover with wind speed were still uncertain. Albedos and spectral shapes for this model over a range of representative wind speeds are shown in Figure 4.14.

Note that the graph in Figure 4.14 corresponding to a wind speed of 0 ms^{-1} shows subtle differences to the flat surface in the top portion of Figure 4.13. This is because even with no local wind there is assumed to be some small residual wave motion due to winds at other locations.

The addition of wind effects has changed the predicted albedo and spectral shape in several important ways. Firstly, the absolute albedo at small viewing angles increases slightly, more noticeably at higher wind speeds. Secondly, the albedo at larger angles decreases significantly. As wind speed increases, the spectral shape varies less with viewing angle, as does the albedo itself.

Figure 4.14 also indicates we would not expect to see very low surface albedoes: they are of the order of several percent for nadir viewing angles. When applied to the retrieval scheme, however, this model has several problems. Firstly, in the nadir geometry the sunglint region is not flagged (as the model predicts a below-threshold surface albedo) although the effects of sun-glint are clearly visible in the results, shown in Figure 4.15, in the form of streaks of high optical depth and effective radius on the right-hand side. In contrast the entirety of the forward view becomes flagged as sun-glint and hence unusable. This behaviour may be due to the fact that the model assumes diffuse illumination, which is not the case for a cloudless sky.

4.4.3 Use of the sun-glint region as a guide to surface spectral shape

Retrievals are not performed in the sun-glint region of a scene because the signal is dominated by the surface term, which essentially masks any aerosol perturbation. Although it is not good practice to obtain one's *a priori* from data (as this stops data and *a priori* being independent), consideration of the spectral shape of radiance received in the sun-glint region could provide a ballpark for approximate values and ranges one might expect from a model of the surface.

Figure 4.16 shows received radiance and its spectral shape in the sun-glint region of the first test scene (see Figure 4.1) for both viewing geometries. As the nadir view has a shorter path length, we would expect

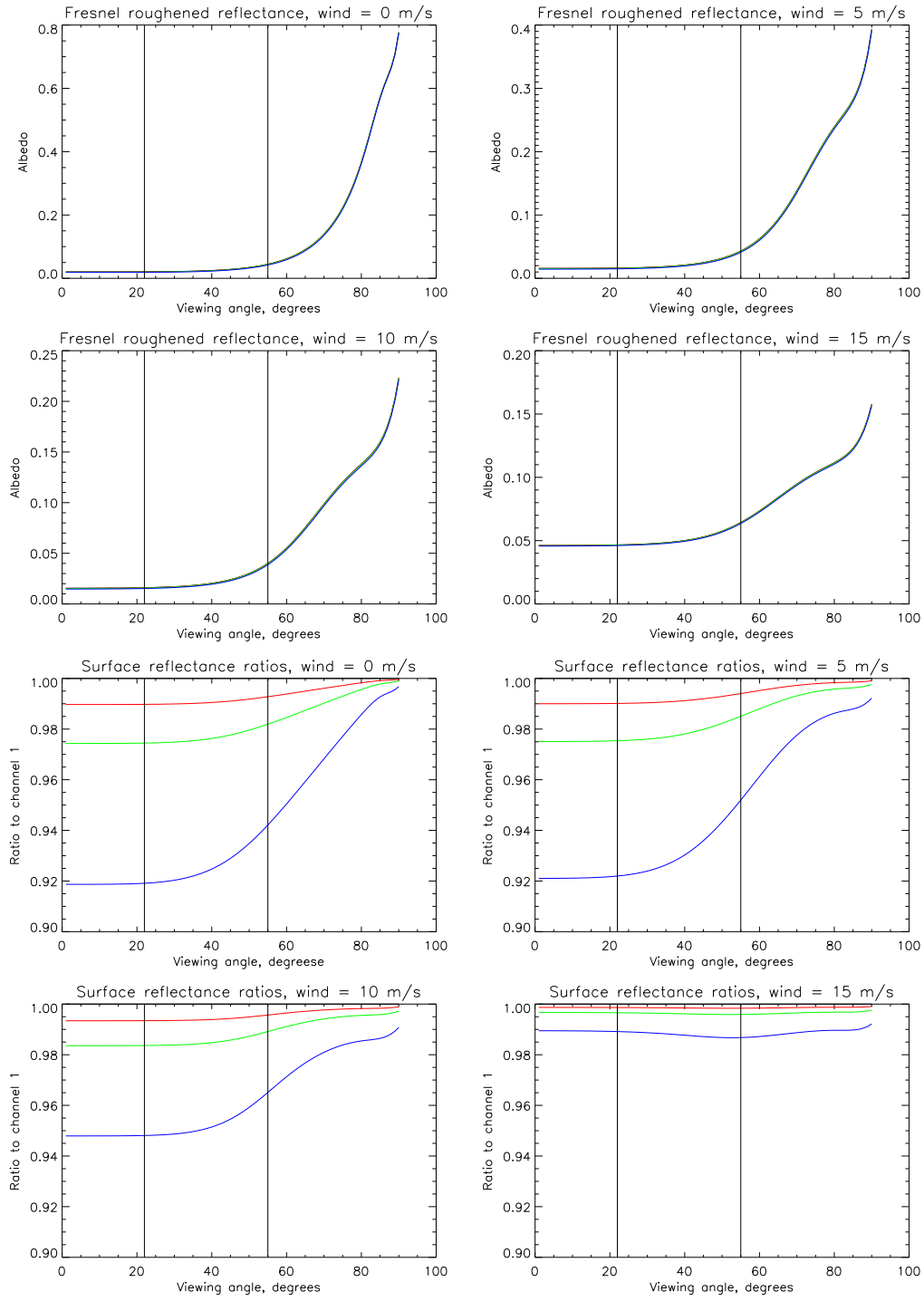


Figure 4.14: *Top four graphs:* Sea surface albedo for a wind-ruffled surface according to the model of Sathe and Vaithiyanathan [1994], for a range of typical wind speeds. *Bottom four graphs:* Ratios of channel albedo to 0.55 μm albedo for this model at these same wind speeds. In all graphs, 0.55 μm values are in black, 0.66 μm in red, 0.87 μm in green and 1.6 μm in blue.

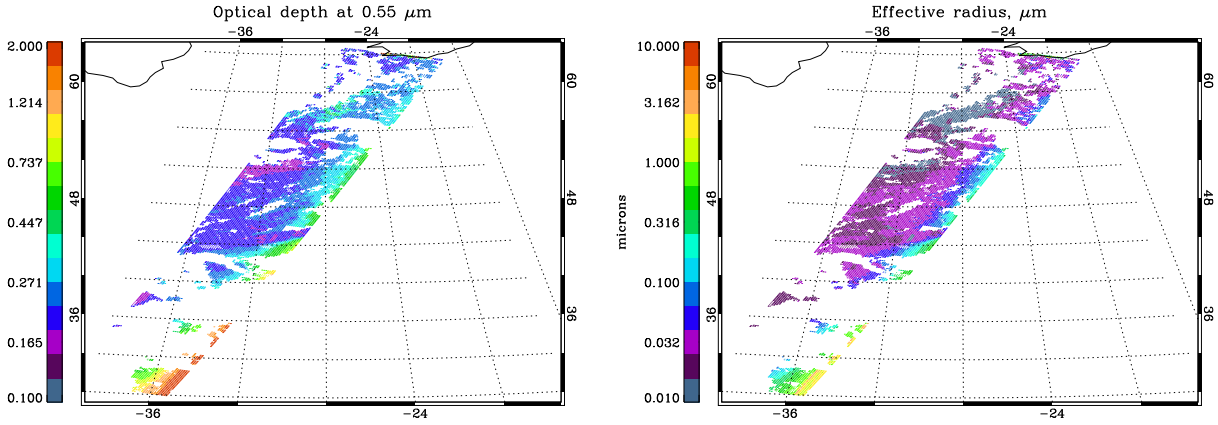


Figure 4.15: Nadir view retrieved optical depth at 0.55 μm and effective radius for the first test scene, using the model of Sathe and Vaithiyanathan [1994] for *a priori* sea surface albedo.

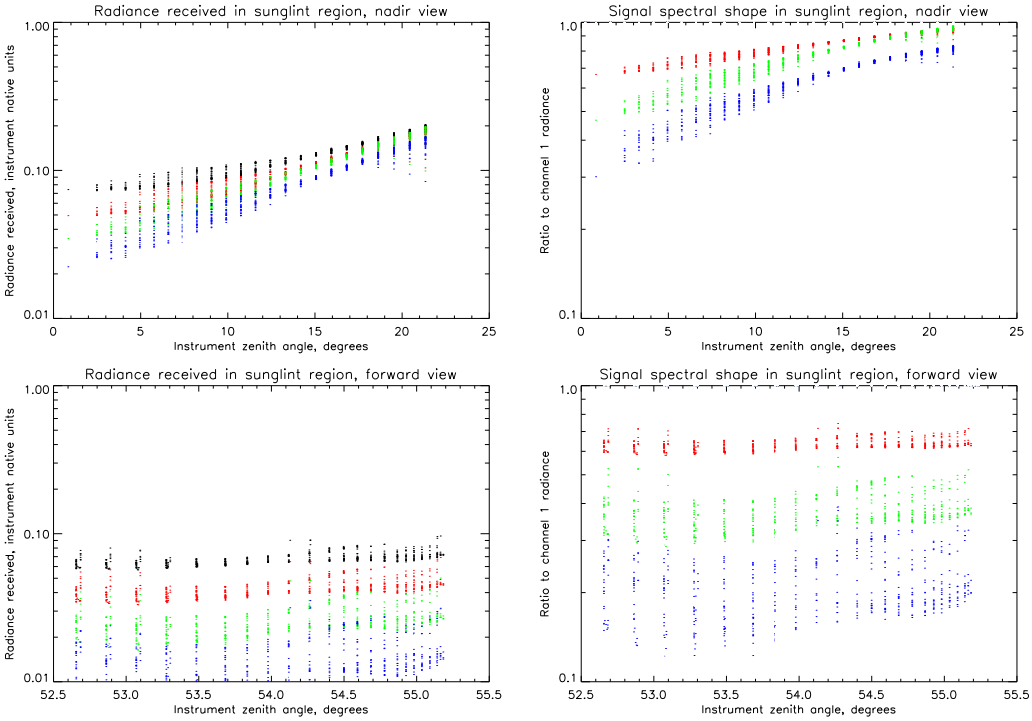


Figure 4.16: *Left:* Radiance received in the sun-glint region at nadir (top) and forward (bottom) viewing geometries, plotted against viewing angle for consistency with previous figures. *Right:* Ratios of channel received radiance to 0.55 μm radiance for nadir (top) and forward (bottom) viewing geometries in the sun-glint region. In all graphs, 0.55 μm values are in black, 0.66 μm in red, 0.87 μm in green and 1.6 μm in blue.

this spectral shape to be the best match for the true surface spectral shape. The following observations can be made from these graphs:

- Higher radiance is, in general, observed at the nadir view than the forward. This can be explained by the fact that the nadir view is looking at the sun-glint region, while the forward is less affected by glint.
- The same wavelength dependence is observed in radiance as in the sea surface models: longer wavelengths are duller.
- The relative flatness of spectral shape throughout viewing angles sampled by the nadir view predicted by the smooth Fresnel surface and wind-roughened Sathe and Vaithiyanathan [1994] models is also seen in the radiance.
- Although the overall shape is similar, the magnitudes of the ratios differ considerably. Radiance observed is much less spectrally flat than the surface spectral shape predicted by the models. This may indicate that there is a stronger dependence on refractive index, or some other wavelength-dependent factor, than is accounted for by the models. However, it must not be forgotten that because of the presence of aerosol (particularly with the forward view's longer path and lower susceptibility to sun-glint) we are not truly looking at the surface, thus we do not know how much of the discrepancy is due to the aerosol and how much the surface spectral shape.

4.4.4 Temporal and spatial variation of the refractive index of water

The refractive index of water is dependent on its composition, which may be non-uniform in space or time because of factors including pollutants and the levels of organic matter from phytoplankton, zooplankton and products of their life cycles. Currently, one value is used for refractive index at each wavelength. This is a problem because bodies of water the world over are not uniform. Indeed, there is a wide field devoted to the analysis of diverse ocean colour (Morel [1988], Lee and Hu [2006]).

Spatial variations in refractive index affecting how well a body of water is described by the calculated spectral shape could explain the differing qualities of fit and trends observed in the test scenes. Another possibility would be to perform retrievals over the same expanse of ocean and see how the quality of fit or distribution of residuals varies annually, as this could be an indicator of a changing sea spectral shape. Alternatively, both spatial and temporal differences in quality of retrieval fit could simply be due to different aerosol being present, as well as the changing solar and observation angles, and described by the forward

model to a different degree of success. Because of the known non-uniformity in ocean colour, however, it seems that changing refractive indices must play at least some role.

4.5 Consequences of a low *a priori* sea surface albedo

For both single and dual-view retrievals, retrieved sea surface albedo at 0.55 μm is in many cases lower than expected, despite the retrieval cost being low and the residuals small (thus indicating a good fit). One cause for a low retrieved albedo is a low *a priori* albedo. A lower than expected retrieved albedo means that less radiance is being attributed to the sea surface, meaning more in turn must come from the aerosol. This may have the effect of decreasing the apparant effective radius: as discussed in Section 1.3.2, small particles scatter more light backward than larger ones, so would increase the radiance seen by the satellite. Therefore, with an apparantly dull surface the retrieval may consequently produce a small value for r_e. In many cases in the test scenes described, the effective radius lies close to the lower limit of the state. Conversely, with a low surface albedo we might expect optical depth to be overestimated compared to the true case.

4.6 Transferable skills developed

In the course of my work so far in AOPP I have had the opportunity to acquire and advance some transferable skills which are applicable to for my doctoral studies and beyond. These include most notably learning how to program with the computer languages IDL and FORTRAN 90, as well as use of the LaTeX typesetting language and BibTeX bibliography format. Attendance of numerous seminars and group meetings, as well as several conferences and workshops, has improved my ability to assimilate information while working as part of a larger group.

I have also attended several IT short courses at OUCS (Oxford University Computing Service): Powerpoint as a tool for presentations and posters; CSS for web development; and XHTML. The latter two courses have led to my assuming responsibilities for maintenance of the webpage of the Earth Observation research group, as well as the Trinity College graduate website.

Chapter 5

Future Work

Although the final state of the scope of the research project has not been finalised (as it is difficult to predict precisely how much will be achieved in the available time) several areas have been highlighted for particular attention. These are discussed below.

5.1 Algorithm improvements

5.1.1 *A priori* surface albedo improvements

Sea surface albedo

Further work into sea surface models will provide a more accurate *a priori* model, as well as information as to whether the low values currently retrieved are unphysical. As well as modelling the effects of wind, a possible approach to the problem would include information on the spatial and temporal variation of pigment concentrations and their effects on ocean colour. Pigments are thought to be most important for the 0.55 μm channel, as this is an absorption wavelength of chlorophyll. Models such as Gordon et al. [1988] have been developed to take pigments into account, and the classification scheme of Morel [1988] has been widely adopted (used for example recently by Lee and Hu [2006]) to describe global variation. The SeaWIFS instrument often used for such measurements does not measure as far as the 1.6 μm channel used by the (A)ATSR instruments and so the ocean colour product from it may not be directly used.

It is hoped that the current model's basic suitability for wind speed effects can be assessed and, if necessary, improved within a few months. The time required to investigate the effects of ocean colour on the instrument channels used and make any adjustments to the algorithm is uncertain.

Land surface albedo

As described in Section 3.3.2, the same MODIS land data is used for the *a priori* surface albedo for both instrument viewing geometries. In principle as surface albedo is retrieved this is not a large source of error, although it carries with it the inbuilt assumption that the spectral shape of the surface is the same for both nadir and forward views. This may not be the case, and so a full BRDF will be developed from this product to more accurately describe the surface albedo.

It is also possible that constraining the forward and nadir view surface albedo to some fixed ratio, following the method of North et al. [1999] (validated in North [2002]) over land, would improve results.

This area of work is prioritised lower than improving *a priori* sea surface albedo, as the focus of this project (as described in Section 5.2.1) will be mostly over the ocean. As the Grape and GlobAEROSOL projects are global in scope, however, the improvements will eventually be made to the ORAC algorithm.

5.1.2 Introduction of infrared channels

Current work in the research group involves the production of a forward model for the infrared channels of (A)ATSR and SEVIRI (located at 3.6, 11 and 12 μm as shown in Table 2.1) for use with the GlobAEROSOL project. Information from these channels should allow the retrieval of aerosol layer effective height, as well as something about surface emissivity (at the infrared wavelengths). Infrared wavelengths should also be more sensitive to effective radius than the visible, improving the retrieval of this parameter. When this forward model is complete, lookup tables will be generated allowing its incorporation into the retrieval schemes. The forward model is tentatively hoped to be finished around December 2006, after which implementation into the retrieval scheme and testing should take an additional few months.

5.1.3 Instrument channel drift correction

One reason for the need for calibration is that instruments decay over time. Over the course of its lifetime, the degradation of the ATSR-2 visible channels was monitored using stable terrestrial sites by Smith et al. [2002], and corrections for the long-term drift in gain applied. Current work by Smith [2005] is investigating the need for similar corrections for AATSR data, although as yet no finalised algorithm has been published. The 0.55 μm channel is thought to be the one most affected by any possible correction, as in both ATSR-2 and current AATSR data it drifts visibly faster than the others.

If not applied to the raw data, any correction algorithm published will be applied during the preprocessing stage. The timing of the publication of a correction method is uncertain, but when available it should not take long to apply. If progress appears slow on this front then one possibility is to treat the measurement errors as larger than the official values (particularly for the 0.55 μm channel) and see how this affects results in the interim.

5.2 Areas of focus of study

5.2.1 Transport of Saharan dust

Saharan dust is known to be blown long distances from the desert by wind (see, for example, Dulac et al. [2001], Hamonou and Chazette [1998] or Borbély-Kiss et al. [2004]), sometimes even reaching South America (Formenti et al. [2001]) where it may fertilise the Amazon basin. It is hoped that this desert dust can be observed blowing from Africa, and the mass flux, as well as the final location of the dust can be estimated. This would be an important step in the quantification of global aerosol transport.

Focussing on this part of the world means the ocean surface is particularly important to model, as this will be the background surface for most of the scenes of interest. Also important is that the aerosol model used for desert dust is accurate, which is something that could be addressed by *in situ* measurements. In particular, the model could be extended to account for non-spherical particles following the method of Dubovik et al. [2005]. Addition of the infrared channels to the retrieval will improve the performance of the retrieval over desert, as the contrast between blown desert dust and the desert surface will be greater in the infrared (where airborne dust will be cooler) then the visible. The cloud detection algorithm will also be investigated, as it is suspected that large dust particles may be sometimes incorrectly flagged as cloud.

It should be possible to have preliminary estimates of mass flux by the end of 2006, which can be refined as the algorithm is improved.

5.2.2 Validation exercises

It is important that the retrieval algorithm developed during the project be validated against independent data to test its performance and make clear any shortcomings. Three distinct avenues for validation exist using the aerosol instrumentation described in Section 2.4:

- Particular targets for validation with satellite instruments might be MERIS (being jointly aboard Envisat should offer a lot of data for comparison) and SEVIRI (as it is located above the area of particular

interest). As SEVIRI also uses the ORAC retrieval algorithm the data retrieved will not be completely independent. It will provide a useful check on the retrieved values, and monitor discrepancies arising from differences in raw measurements, but would not bring to light any problems (such as a bias or offset in retrieved quantities) with the algorithm specifically.

- AERONET stations have been widely used in satellite validation and it is anticipated that they will be used for validation of the dual-view algorithm.

- Campaigns have been carried out specifically monitoring dust over the Sahara, such as aircraft measurements taken as part of SAMUM (2005). Having similar goals to this project, they would provide ideal validation opportunities.

Validation is expected to be an ongoing exercise over the next two years. As better descriptions of aerosol properties are obtained using the retrieval algorithm they can be validated, and any issues with the algorithm coming to light as a result of this validation can be described and, hopefully, minimised.

Bibliography

Abdou, W. A., Diner, D. J., Martonchik, J. V., Bruegge, C. J., Kahn, R. A., Gaitley, B. J. and Crean, K. A. [2005], 'Comparison of coincident Multiangle Imaging Spectrometer and Moderate Resolution Imaging Spectrometer aerosol optical depths over land and ocean scenes containing Aerosol Robotic Network sites', *J. Geophys. Res.* **110**.

Andreae, M. O. and Crutzen, P. J. [1997], 'Atmospheric aerosols: Biogeochemical sources and role in atmospheric chemistry', *Science* **276**(5315), 1052–1058.

Borbély-Kiss, I., Kiss, A. Z., Koltay, E., Szabó, G. and Bozó, L. [2004], 'Saharan dust episodes in Hungarian aerosol: elemental signatures and transport trajectories', *J. Aerosol Sci.* **In press**.

Braak, R. [2006], Ozone Monitoring Instrument aerosol products and validation, *in* 'ACCENT Fifth AT-2 Workshop'.

Brown, J., Hoke, M., Doherty, K., Anderson, G. and Berk, A. [2004], *MODTRAN 4 Software*. http://www.vs.afrl.af.mil/ProductLines/IR-Clutter/modtran4.aspx [Accessed 19 July 2006].

Cardon, K., Goryl, P., Scharoo, R. and Benveniste, J. [1998], '1997/1998 El Niño observed by ERS', *ESA publications* .

Charlson, R. J., Schwartz, S. E., Hales, J. M., Cess, R. D., Coakley, J. A., Hansen, J. E. and Hofmann, D. J. [1992], 'Climate forcing by anthropogenic aerosols', *Science* **255**(5043), 423–430.

Cox, C. and Munk, W. [1954*a*], 'Measurement of the roughness of the sea surface from photographs of the Sun's glitter', *J. Opt. Soc. Am.* **44**, 838–850.

Cox, C. and Munk, W. [1954*b*], 'Statistics of the sea surface derived from Sun glitter', *J. Mar. Res.* **13**, 198–227.

d'Almeida, G., Koepke, P. and Shettle, E. [1991], *Atmospheric aerosols: global climatology and radiative characteristics*, A. Deepak Publishing.

Davies, C. [1974], 'Size distribution of atmospheric particles', *J. Aerosol Sci.* **5**, 293–300.

Dils, B. [2006], The evaluation of SCIAMACHY CO scientific data products, using ground-based ftir measurements, *in* 'ACCENT Fifth AT-2 Workshop'.

Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B., Mischenko, M., Yang, P., Eck, T., Volten, H., Muñoz, O., Veihelmann, B., van der Zande, W. J., Leon, J.-F., Sorokin, M. and Slutsker, I. [2005], 'The application of spheroid models to account for aerosol particle non-sphericity in remote sensing of desert dust', *J. Geophys. Res.* .

Dulac, F., Chazette, P., Gomes, L., Chatenet, B., Berger, H. and Vinicula Dos Santos, J. M. [2001], 'A method for aerosol profiling in the lower troposphere with coupled scatter and meteorological rawindsondes and first data from the tropical Atlantic off Sahara', *J. Aerosol. Sci.* **32**, 1069–1086.

ECMWF [2006], *European Centre for Medium-Range Weather Forecasts Home Page*. http://www.ecmwf.int/ [Accessed 19 July 2006].

ESA [2005], *AATSR Frequently Asked Questions*, 1st edn. See also http://envisat.esa.int/instruments/aatsr/ [Accessed 13 March 2006].

Formenti, P., Andreae, M. O., Lange, L., Roberts, G., Cafmeyer, J., Ratja, I., Maenhaut, W., Holben, B., Artaxo, P. and Lelieveld, J. [2001], 'Saharan dust in Brazil and Suriname during the Large-Scale Biosphere-Atmosphere Experiment in Amazonia (LBA) - Cooperative LBA Regional Experiment (CLAIRE) in March 1998', *J. Geophys. Res.* **106**(D14), 14919–14934.

Good, S., Lawrence, S. and Llewellyn-Jones, D. [2005], The measurement of climate change using data from the along-track scanning and advanced very-high resolution radiometers, *in* '2005 ESA MERIS/(A)ATSR Workshop'.

Gordon, H. R., Brown, O. B., Evans, R. H., Brown, J. W., Smith, R. C., Baker, K. S. and Clark, D. K. [1988], 'A semianalytic radiance model of ocean colour', *J. Geophys. Res.* **93**(D9), 10909–10924.

Hamonou, E. and Chazette, P. [1998], 'Evidence of saharan mineral aerosols transport to the mediterranean inside well-defined layers', *J. Aerosol Sci.* **29**, S1263–S1264.

Hess, M., Koepke, P. and Schult, I. [1998], 'Optical properties of aerosols and clouds: The software package OPAC', *Bull. Am. Met. Soc.* **79**(5), 831–944.

Hűnel, G. [1976], 'The properties of atmospheric aerosol particles as functions of the relative humidity at thermodynamic equilibrium with the surrounding moist air', *Adv. Geophys.* **19**, 73–1988.

Huang, S. and Siegert, F. [2004], 'ENVISAT multisensor data for fire monitoring and impact assessment', *Int. J. Rem. Sens.* **25**(20).

Inomata, Y., Iwasaka, Y., Osada, K., Hayashi, M., Mori, I., Kido, M., Hara, K. and Sakai, T. [2006], 'Vertical distributions of particles and sulfur gases (volatile sulfur compounds and SO_2 over East Asia: Comparison with tho aircraft-borne measurements under the Asian continental outflow in spring and winter', *Atmos. Env.* **40**, 430–444.

IPCC [2001], *Climate Change 2001: The Scientific Basis*, Intergovernmental Panel for Climate Change (IPCC).

Junge, C. E. [1955], 'The size distribution and aging of natural aerosols as determined from electrical and optical data on the atmosphere', *J. Atmos. Sci.* **12**(1), 13–25.

Justus, C. G. and Paris, M. V. [1985], 'Modelling solar spectral irradiance and radiance at the bottom and top of a cloudless atmosphere', *J. Appl. Meteorol.* **24**, 193–205.

Lee, Z. and Hu, C. [2006], 'Global distribution of Case-1 waters: An analysis from SeaWiFS measurements', *Remote Sens. Environ.* **101**, 270–276.

Lu, G., Brook, J. R., Alfarra, M. R., Anlauf, K., Leaitch, W. R., Sharma, S., Wang, D., Worsnop, D. R. and Phinney, L. [2006], 'Identification and characterization of inland ship plumes over Vancouver, BC', *Atmos. Env.* **40**, 2767–2782.

Marsh, S., Dean, S., Grainger, R., Quijano, A. and Thomas, G. [2004], *An Optimal Estimation Aerosol Retrieval Scheme For ATSR-2*, Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford.

Morel, A. [1988], 'Optical modeling of the upper ocean in relation to its biogenous matter content (Case I waters)', *J. Geophys. Res.* **93**(C9), 10749–10768.

Mutlow, C., Corlett, G. and Smith, D. [2005], The AATSR sensor and its in-flight performance, *in* '2005 ESA MERIS/(A)ATSR Workshop'.

NASA [2006], *AERONET: The AERosol RObotic NETwork home page*. http://aeronet.gsfc.nasa.gov/ [Accessed 19 July 2006].

North, P. R. [2002], 'Estimation of aerosol opacity and land surface bidirectional reflectance from ATSR-2 dual-angle imagery: Operational method and validation', *J. Geophys. Res.* **107**(0).

North, P. R., Briggs, S. A., Plummer, S. E. and Settle, J. J. [1999], 'Retrieval of land surface bidirectional reflectance and aerosol opacity from ATSR-2 multiangle imagery', *IEEE Trans. Geosci. Remote Sens.* **37**(1), 526–537.

Olesen, F. S. [2005], *DISORT Algorithm*.
http://imk-msa.fzk.de/msa-public/Software-tools/Modtran/science/disort.htm [Accessed 19 July 2006].

Pueschel, R. F. [1996], 'Stratospheric aerosols: Formation, properties, effects', *J. Aerosol Sci* **27**(3), 383–402.

Rahn, K. A. and Lowenthal, D. H. [1984], 'Elemental tracers of distant regional pollution aerosols', *Science* **223**(4632), 132–139.

Ramachandran, S. [2005], '$PM_{2.5}$ mass concentrations in comparison with aerosol optical depths over the Arabian Sea and Indian Ocean during winter monsoon', *Atmos. Env.* **39**, 1879–1890.

Richter, A. [2006], Glyoxal measurements with SCIAMACHY - a new tropospheric species from satellite measurements, *in* 'ACCENT Fifth AT-2 Workshop'.

Rodgers, C. D. [2000], *Inverse methods for atmospheric sounding: Theory and Practice*, Series on Atmospheric, Oceanic and Planetary Physics–Vol. 2, World Scientific.

RSI [2006], *IDL Home Page, Research Systems Inc.*
http://www.ittvis.com/idl/index.asp [Accessed 19 July 2006].

SAMUM: Saharan mineral dust experiment home page [2005].
http://www.tropos.de/samum/index.html [Accessed 19 July 2006].

Sathe, P. V. and Vaithiyanathan, R. [1994], 'Total reflectance of wind-roughened sea surfaces: directional properties', *Comp. & Geosci.* **20**(10), 1447–1453.

Scatterfield, L., O'Bannon, J. and Vermote, E. [2004], *MODIS Land Surface Reflectance Home Page*.
http://modis-land.gsfc.nasa.gov/ [Accessed 19 July 2006].

Schwartz, S. E. [1996], 'The Whitehouse Effect - Shortwave Radiative Forcing Of Climate By Anthropogenic Aerosols: An Overview', *J. Aerosol. Sci.* **27**(3), 359–382.

Sidran, M. [1981], 'Broadband reflectance and emissivity of specular and rough water surfaces', *Appl. Opt.* **20**(18), 3176–3183.

Silió, A. and Bricaud, A. [2005], Synergetic use of ocean colour (MERIS) and sea surface temperature (AATSR) for a study of variations in total and new primary production in upwelling areas: Application to the benguela ecosystem, *in* '2005 ESA MERIS/(A)ATSR Workshop'.

Smith, D. L. [2005], *Report for AATSR VisMon contract*, 1st edn. Prepared by the Rutherford Appleton Laboratory.

Smith, D. L., Mutlow, C. T. and Rao, C. R. N. [2002], 'Calibration monitoring of the visible and near-infrared channels of the Along-Track Scanning Radiometer-2 by use of stable terrestrial sites', *Appl. Opt.* **41**(3), 515–523.

Sória, G. and Sobrino, J. A. [2005], AATSR derived land surface temperature from heterogeneous areas, *in* '2005 ESA MERIS/(A)ATSR Workshop'.

Targino, A. C. and Noone, K. J. [2006], 'Airborne in situ characterization of dry urban aerosol optical properties around complex topography', *Atmos. Res.* **79**, 148–174.

Thomas, G. [2006], A unified, global aerosol dataset from MERIS, (A)ATSR and SEVIRI, *in* 'ACCENT Fifth AT-2 Workshop'.

Thomas, G., Dean, S., Carboni, E., Grainger, R., Poulsen, C., Siddans, R. and Kerridge, B. [2006], *GlobAEROSOL Data User Element–Aerosol Extraction Algorithm Definition, ATSR-2/AATSR Algorithm Theoretical Bais Document*, 2nd edn, Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford.

Thurston, G. D., Laird, N. M., Rahn, K. A. and Lowenthal, D. H. [1985], 'Tracing aerosol pollution', *Science* **227**(4693), 1406, 1408, 1412. Letters discussing Rahn and Lowenthal's Elemental Tracers paper (Science, 1984).

Twomey, S. [1977], *Atmosperic Aerosols*, Elsevier.

Watts, P., Mutlow, C., Baran, A. and Zavody, A. [1998], 'Study on cloud properties derived from meteosat second generation observations', *EUMETSAT ITT* .

Wetzel, M. A. and Stowe, L. L. [1999], 'Satellite-observed patterns in stratus microphysics, aerosol optical thickness, and shortwave radiative forcing', *J. Geophys. Res.* **104**, 31287–31300.

Yoon, S.-C., Kim, S.-W., Kim, J., Sohn, B.-J., Jefferson, A., Choi, S.-J., Cha, D.-H., Lee, D.-K., Anderson, T. L., Doherty, S. J. and Weber, R. J. [2006], 'Enhanced water vapor in Asian dust layer: Entrainment processes and implication for aerosol optical properties', *Atmos. Env.* **40**, 2409–2421.