# Bioaerosol Detection through Simultaneous Measurement of Particle Intrinsic Fluorescence and Spatial Light Scattering

**John Edward Barton**

A thesis submitted in partial fulfilment of the
requirements of the University of Hertfordshire
for the degree of Doctor of Philosophy

The programme of research was carried out in the
Department of Physics, Astronomy and Mathematics,
University of Hertfordshire

in collaboration with

Defence Science and Technology Laboratory.

December 2005

# Abstract

Interest in the role and detection of airborne biological micro-organisms has increased dramatically in recent years, in part through heightened fears of bioterrorism. Traditional bio-detection methods have generally slow response times and require the use of reagents. Conversely, techniques based on light scattering phenomena are reagent-free and are able to operate in real-time.

Previous research has established that classification of certain types of airborne particles on the basis of shape and size may be achieved through the analysis of the *spatial* light scattering patterns produced by individual particles. Similarly, other research has shown that the intrinsic fluorescence of particles excited by radiation of an appropriate wavelength can be used to establish the presence of biological particles, provided background particles with similar fluorescence properties are not present. This is often not the case.

This thesis, therefore, describes the design, development, and testing of a new type of bio-aerosol detection instrument in which the advantages of both particle spatial light scattering analysis and intrinsic fluorescence are exploited. The instrument, referred to as the Mult-Parameter Aerosol Monitor (MPAM), is unique in simultaneously recording data relating to the size, shape, and fluorescence properties of individual airborne particles at rates up to several thousand particles per second.

The MPAM uses a continuous-wave frequency quadrupled Nd:$YVO_4$ laser to produce both spatial scattering and fluorescence data from particles carried in single-file through the laser beam. This use of a CW laser leads to opto-mechanical simplicity and reduces fluorescence bleaching effects. A custom-designed multi-pixel Hybrid Photodiode (HPD) detector is used to record the spatial scattering data in forward scattering plane whilst particle fluorescence is recorded via a large solid-angle ellipsoidal reflector and single photomultiplier detector.

Calibration tests and experimental trials involving a range of both biological and non-biological aerosols have shown that the MPAM, when supported by appropriate data analysis algorithms, is capable of achieving enhanced levels of discrimination between biological and non-biological particles down to the submicrometre sizes and, in some cases, enhanced discrimination between classes of biological particle.

# Table of contents

# 1 Introduction

In the 1860s the first aerosol impactors were developed. This invention allowed scientists, for the first time, to isolate particles from the air and study them under the microscope. A primary concern for researchers during that period was to investigate the relationship between airborne particles and disease [1]. Since that time both aerosol impactors and other similar technologies have been developed and refined. An array of different instruments are now routinely used in a wide range of aerosol sampling situations. Over more recent years there has been an increasing interest in the role of airborne micro-organisms in the outdoor, indoor and occupational environments. This interest in the presence of airborne biological organisms has arisen for several reasons.

Concern with energy efficiency has meant that the amount of outside air supplied to the interior of buildings has been reduced. This can lead to increasing concentrations of indoor generated air pollutants, including infectious aerosols. The airborne transportation of infectious agents is one of the primary means of disease spread. Tuberculosis, well known to be transmitted by airborne bacteria, has re-emerged as a major public health problem, partly due to the emergence of antibiotic resistant strains.

Recent global events have given rise to an increase in the perceived threat of biological and chemical materials being used as biological warfare agents. The potential impact of the biological component of aerosols on a community could be economically, medically and socially devastating. The ability to detect the biological component of aerosols in *real-time* would therefore, be beneficial to the monitoring and tracking of both ambient and released bioaerosols. Such data could be of use in fields such as agriculture, medicine and public health, the food processing industry and in the early detection of biological warfare agents.

## 1.1 Primary biological aerosol particles

Biological particles normally make up a small but ubiquitous part of both atmospheric aerosols and the indoor air environment. The most commonly identified components of bioaerosols consist of viruses (<~0.5μm), bacteria (~0.5 –10 μm size range) and fungal spores (~10-50 μm size range). These micro-organisms are rarely found as discrete cells but rather bound together by complex polymers. This serves as a protective outer layer conferring a defence against the hostile aerosol environment. Consequently naturally occurring bioaerosols are frequently aggregates of the same species of micro-organism, or mixtures of different sized particles carried on larger organic (plant fragments, skin particles, insect parts etc) or inorganic particles.

Other constituents of bioaerosols include fragments of larger organisms, such as body parts of arthropods, skin scales from mammals and plant debris along with excreta from arthropods such as dust mites and cockroaches.

The above may be described as a naturally occurring biological component of the aerosol. In addition, the air is likely to carry a significant biological component arising from anthropogenic sources. This may be as a result of the accidental or incidental release of bioaerosols from agriculture or industrial processes such as sewage treatment works. The intentional release of airborne pathogens for military or terrorist purposes can also not be ruled out. In this context deliberate attempts to conceal the active biological agent are also likely to be made. This might include steps to use interferent or inhibitory agents in an attempt to thwart any detection technologies. Such activity would be likely to give rise to both a bioaerosol concentration and size distribution that is unlikely to arise naturally.

## 1.2 Significance of the biological component

Bioaerosols play an important role as cloud nuclei. The ability of these particles to act as ice nucleation sources at relatively high temperatures (-4°C) compared to most particles gives them a particular significance in the process of cloud formation [2]. The sources of these nuclei include decaying vegetation, bacteria and marine plankton. Given the large surface area of the oceans such biologically sourced nucleation centres can play a significant role in the process of cloud formation and hence influence world wide climate.

The airborne spread of viruses and bacteria are responsible for both contagious and non-contagious diseases in human, animal and plant populations [3]. This is especially so in indoor environments where common human activities such as sneezing and coughing can result in large numbers of liquid droplets being liberated into the surrounding air. These droplets, which may be contaminated with bacteria, are then free to disperse. Their ability to spread infection through inhalation is one of the challenges to the prevention of the spread of disease in public buildings such as hospitals and the workplace.

Mechanical ventilation and air conditioning systems, which can themselves result in large numbers of liquid droplets being generated into the air, may both assist the distribution of bacteria carrying droplets and, by controlling the relative humidity, ensure they remain viable for long periods. Since the first identified cases of Legionnaires disease *(Legionella pneumophilla)*, poorly maintained air conditioning systems still give rise to periodic outbreaks of this disease by the airborne dissemination of the *Legionella pneumophilla* bacteria in water droplets generated by air conditioning the systems cooling towers [4].

Fungi are responsible for a wide range of allergic reactions and infectious diseases such as blastomycosis and asperigillos. Unlike infectious agents, fungal spores do not have to be viable to retain allergenic properties [5]. It is therefore important to be able to enumerate the total air spora, including the non-culturable components such as hyphal fragments etc [6].

In the military field it is now quite possible for a state of the art biological laboratory to be easily established at low cost, requiring only fairly basic micro-biological knowledge to operate it. Such a laboratory would have the capacity to culture large quantities of biological agents such as anthrax (*Bacillus anthracis*) for later use as a biological weapon.

The application of modern genetic engineering techniques makes it relatively easy to create modified biological agents intended to avoid detection, reduce the effectiveness of medical countermeasures and inflict maximum levels of social and economic damage [7].

It is often assumed that any such biological attack will be directed primarily against human targets. Non-human targets however, such as fisheries, agriculture, forestry, transport networks, etc represent both major economic threats and the potential to create high levels of social disruption while presenting easier targets.

## 1.3 Detection and sampling methods for bioaerosols

The most commonly used detection and identification techniques involve a two step process. Firstly particles are removed from the air suspension by impaction onto a filter membrane, agar or greased surface or impingement into a liquid. Subsequently these collected particles are cultured and stained before being identified and enumerated using microscopy. These methods have a high sensitivity and can be highly specific if required. A large number of studies of bioaerosol distributions using these methods have been published in the literature.

It is now recognised that historical data collected using these techniques on the biological constituents of the ambient aerosol have tended to under represent the biological component. This has arisen for a number of reasons such as: differences in the collection efficiency of the sampler [8] [9], low survival rates during atmospheric transportation, mechanical stress imposed on the organism by the collection method, sampling time, the recovery media used [10] [11], and the relative hardiness of the different micro-organisms.

Further variation in the recovery rate of airborne micro-organisms results from the tendency of some bacterial pathogens (e.g. *Vibrio cholerae*) to respond to environmental stress when released into the environment by entering a viable but non-culturable (VBNC) phase [12]. In this state they are difficult to culture using traditional microbiological culturing

techniques and are therefore likely to be under represented in any sample relying on culturing techniques.

This tendency to under represent the true significance of the biological component may have serious implications for ecological research and estimations of human exposure to infectious agents.

These methods also have the disadvantage of requiring relatively expensive sample collection and preparation techniques and highly trained personnel to carry out the classification of the cultured sample. While in principle it may be possible to automate various stages of this process these methods have a fundamental limitation in the speed with which they can give a result. This arises from the time required for the collected cells to multiply on the culture medium to detectable numbers prior to classification. Culturing times are generally measured in days.

Despite their limitations such sampling methods can provide a useful historical record of the bioaerosol component of the air in a particular region at different times of the year. This information can play an important role in determining if changes in the biological component are part of a normally developing pattern or represent an event outside the normal conditions requiring further monitoring.

## 1.4 Microbiological based techniques

Recently, detection time scales have been reduced by the adoption of methods used in clinical microbiology. These techniques rely on the detection of protein, DNA or RNA material from bacteria or viruses. Immunological based detection methods are designed to detect key proteins found in the specific target organisms. The current generation of this technology requires several processing steps and the use of liquid reagents. By the use of lab-on-a-chip techniques, results can be achieved in minutes although the confidence level in the results is generally low requiring samples to be sent for further laboratory analysis [7].

Molecular methods such as polymerase chain reaction (PCR) [13] have a high degree of sensitivity and specificity. The rapid determination that, for instance, a powder sample has a high concentration of a biological warfare agent such as anthrax is now quite trivial with such methods. However the real-time application of PCR to the detection of bioaerosols still has to overcome several obstacles.

Naturally occurring aerosols are quite complex chemically and may contain high concentrations of natural PCR inhibitors such as humic acids, reducing the efficiency of the

PCR reaction. Current knowledge as to the particular parts of the genetic code to test for are not well established. Degeneracy of the genetic code means there is considerable variability in the DNA. Consequently unless PCR detection is based on strongly conserved regions of the code it is quite conceivable that PCR based on code sequence from one strain of a target organism will not give a sufficiently close match for a different strain to lead to its detection.

It should also be born in mind that as more information on the genome of these agents becomes available it may become possible to reverse engineer biological warfare agents to evade PCR detection while conserving those parts necessary for maintaining the toxicity of the organism [14].

Recent work with these methods has tended to concentrate on developing the micro-biological techniques of the detection process itself. With these methods the issues of how the sample is delivered from the air stream, saturation of the detection chemistry by non-target biological material, the manner in which the output is presented and the potential to operate continuously are areas that remain to be addressed.

Most importantly, the current generation of PCR based techniques still have slow response times, at best ten minutes to several hours rather than real-time responses of a second or less. Methods that allow the rapid detection of biological organisms without first requiring their separation from the sampled air stream followed by the application of difficult to handle reagents would offer many advantages.

### 1.4.1 Light scattering based shape and size measurements

The current generation of light scattering based aerosol monitoring instruments are capable of characterising individual particles on the basis of their size and shape. These instruments are able to classify large numbers (5,000 -10,000) of particles every second and can therefore be considered to be operating in real-time. Such instruments have played an important role in the monitoring of atmospheric aerosols, in detecting contaminants in clean room air and the sizing and characterisation of powders.

Work previously undertaken at the University of Hertfordshire and by others has shown that an analysis of the spatial intensity distribution of light scattered by particles can provide an effective means of determining their shape and size. In this document such spatial intensity distributions will be referred to as scattering profiles, for brevity.

Such light scattering based instruments can play an important role in monitoring both atmospheric aerosols and indoor air quality. The rapid detection of asbestos fibres, for

example, is vital to maintaining safety when handling this type of material. Little published work has been found on the use of these techniques for the identification of the bioaerosol component of aerosols. It is probable that this arises for a variety of reasons, in particular unlike optically based counter/sizers it is only in the last few years that instruments capable of performing particle shape discrimination have made the transition from being laboratory based experimental instruments to being available commercially.

Traditionally the interest in bio-aerosols has been more concerned with the identification of the detailed taxonomy of the sample, rather than the simpler question: Is this particle of biological origin or not?

The levels of shape discrimination that can be achieved by these instruments without recourse to some additional information from, say, staining and visual inspection is not going to answer the taxonomic question. It seems possible however given the characteristic shapes of pollens and bacteria that an instrument discriminating on shape and size may be able to go some way to answering the simpler yes/ no question.

### 1.4.2 Fluorescence based detection

Several groups [15, 16, 17] have demonstrated instruments capable of detecting biological particles directly from within a sample air stream. These instruments exploit the intrinsic fluorescence of biological materials when excited at an appropriate wavelength to read a fluorescence emission signature from individual particles. These instruments also generally record a simultaneous measurement of the particle's size. This may be based on elastically scattered light or aerodynamic particle size (from particle time of flight through illuminating beams).

Unfortunately fluorescence may also be stimulated in a range of non biological materials that are also likely to be found in the aerosol being sampled. The identification of aerosol particles as being of a biological origin consequently requires some further information than a measure of the particle's fluorescence signature (and size) to increase the certainty of the detection method.

## 1.5 Objectives

The current generation of instruments based on either light scattering techniques or the detection of intrinsic fluorescence are already able to provide real-time characterization of an aerosol without first requiring the particle's separation from the sampled air stream. These instruments do not require the use of difficult to handle reagents and can operate

continuously over extended periods. However neither method can provide a sufficient level of discrimination to unambiguously detect the biological component of an aerosol.

The work described in this thesis seeks to combine in a single instrument the advantages that spatial light scattering analysis offers in terms of particle classification according to shape and size with those offered by intrinsic fluorescence measurement.

The primary objective of the project is therefore:

**to develop a real-time multi-parameter instrument capable of achieving enhanced levels of discrimination of particle types down to the submicron size range with the specific aim of differentiating biological from non-biological particles.**

The work to develop this multiparameter aerosol monitor, referred to as MPAM, has been carried out in four stages and these are reflected in the structure of this report:

Stage 1:

A review of the fluorescence of biological materials with particular reference to bioaerosols has been undertaken. The range of instruments employing various measurements of the fluorescence of bioaerosols has been investigated. The review is covered in Chapter 2.

Stage2:

A brief overview of particle shape classification based on scattering profiles has been carried out. Light scattering instruments that have been developed, in particular at the University of Hertfordshire are described. These instruments employ either a wide angle or a low angle scattering geometry for particle characterisation. However prior to this study no systematic experimental investigation of the relative merits of these two different scattering geometries has been undertaken for biological particle characterisation. Stage two of the work was therefore to carry out this comparative evaluation by modifying an existing instrument to record data from both wide and low angle scattering from individual airborne particles. The experimental results are summarised and the conclusions based on this experimental work are presented in Chapter 3.

Stage3:

Based on the outcome of Stages 1 and 2 the MPAM instrument has been designed and built to record the spatial scattering and fluorescence data from individual airborne particles. The instrument has been tested with a range of aerosol materials. This work is described in Chapters 4 and 5.

Stage 4:

Several data processing procedures have been used to analyse the data acquired from the MPAM prototype instrument. An assessment has been made of the relative merits of these various methods. Finally some overall conclusions have been drawn based on the instrument's performance. This stage of the work is described in Chapters 6 and 7.

# 2 Review of the fluorescence of biological materials

## 2.1 Fluorescence

Fluorescence is the emission of electromagnetic radiation that takes place when a molecule in an excited electronic state returns to the ground state. Fluorescence is necessarily a two stage process. The first involves the absorption of a photon whose energy corresponds to an allowed absorption transition followed some time later by emission of radiation and the return to its ground state. The lifetime of the molecule in the excited state depends on competition between the radiative emission and radiationless deactivation processes. Such processes provide an alternative mechanism for a return to the ground state and will result in a quenching of the fluorescence intensity. Often more than one deactivation pathway is present. The minority of molecular species where emission is the predominant deactivation pathway are referred to as fluorophores.

This work is primarily concerned with fluorescence from organic molecules in vivo, consequently the discussion in the main will be limited to processes that influence the excitation and fluorescence of this general class of molecules.

### 2.1.1 Absorption and emission processes

The energy diagram of Figure 2.1 shows the allowed transitions for a polyatomic organic molecule. The diagram shows two excited electronic states $S_1$ and $S_2$, along with the electronic ground state $S_0$. Each of the electronic states is shown divided into a number of vibrational modes[a] $v_i \ldots v_0$ represented by the grey lines in the diagram. At ambient temperatures most organic molecules will be found in the lowest vibrational level of the ground electronic state. Electronic transitions can therefore be regarded as taking place from the molecule's lowest vibrational energy level.

Quantum mechanics does not impose any restriction on changes in vibrational quantum number in going from a ground state to any level of the excited state. Thus the excited energy level can effectively be considered to consist of a series of closely spaced energy

a. These levels can be further divided into discrete molecular rotational bands not shown here. Such bands are normally observed of material in the gaseous phase. In the case of large biological molecules their rotational freedom is normally constrained by their immediate environment and consequently such levels can be ignored.

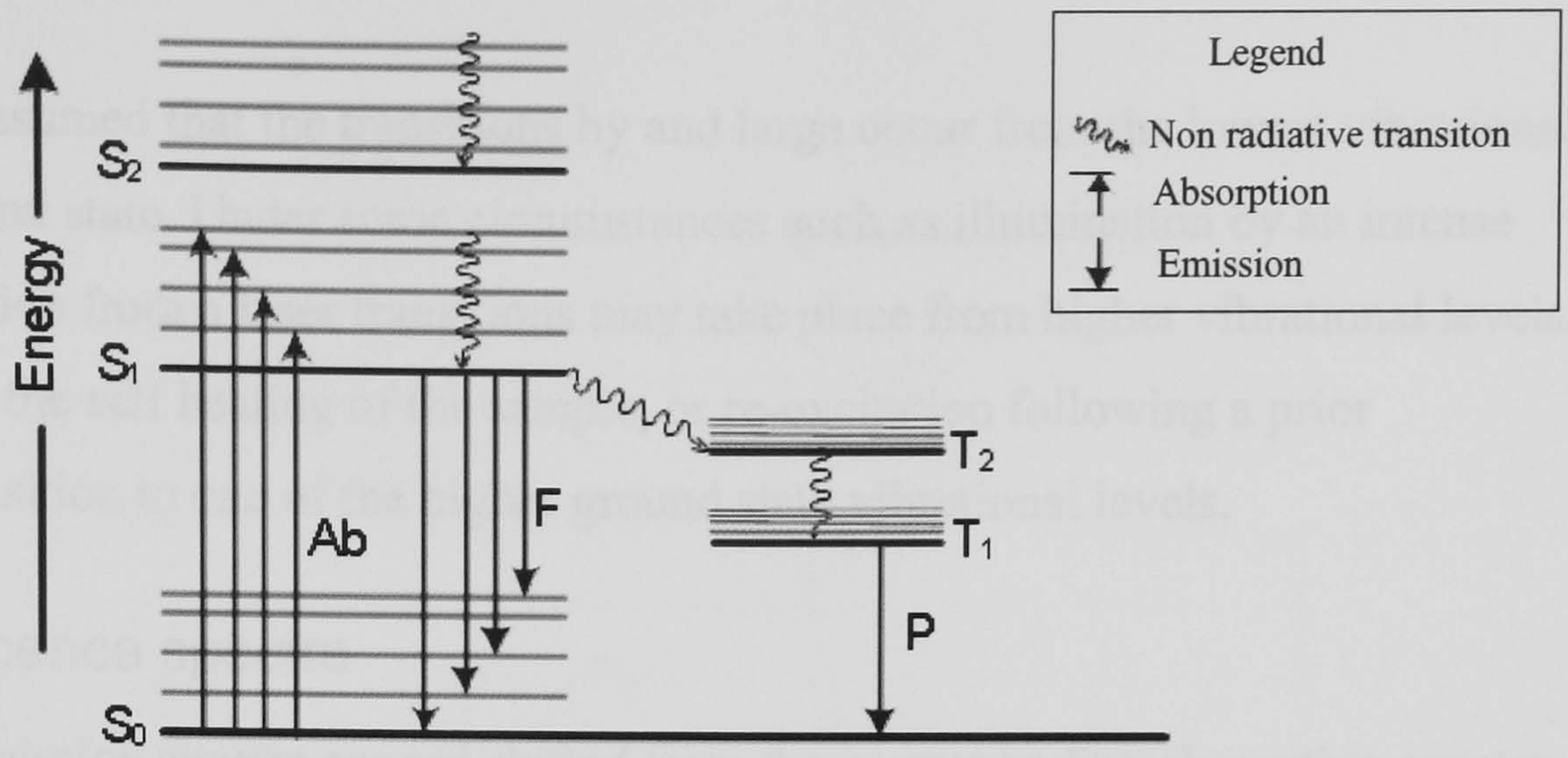



**Figure 2.1 Simplified energy level diagram of a polyatomic molecule having a singlet ground state $S_0$ and two excited singlet electronic levels $S_1$ and $S_2$. Two excited triplet states $T_1$ and $T_2$ are also shown. $A_b$ represents the absorption of a photon. F and P represent fluorescence and phosphorescence respectively.**

bands. Photons of sufficient energy may then be absorbed giving rise to a molecule in an excited state, shown as $S_1$ and $S_2$ in the diagram. This transition is extremely rapid and is generally treated as an instantaneous process.

Following excitation to any of the higher level electronic states the molecule immediately starts to vibrate at one of the associated vibrational modes. A rapid process of radiationless conversion of energy through the various vibrational levels then takes place. This cascading process can continue until the molecule has reached the lowest vibrational mode of the lowest electronic level, $S_1$ in the Figure. This process is generally referred to as internal conversion. From here the molecule can return to the ground state by means of the emission of a photon giving rise to fluorescence, or by a radiationless deactivation process (referred to as phonon emission), resulting in the fluorescence being quenched. A third deactivation pathway known as intersystem crossing involving a long lifetime metastable state $T_2$ may also take place giving rise to radiative deactivation in the form of phosphorescence. By definition this transition has a low probability for typical fluorescent molecules under normal conditions. Consequently phosphorescence as a predominant deactivation pathway is not considered further here.

The lifetime of the vibrational transitions are in the range 10-1000 femtoseconds compared to fluorescence lifetimes in the order of 10 nanoseconds. As a result of this 1000 fold difference in the relative lifetimes of the two deactivation pathways the probability of the emission of a photon from one of the higher electronic levels is low. Consequently virtually all fluorescence takes place from the lowest vibrational mode of the $S_1$ state giving rise to the characteristic red shift of the fluorescence spectra. (i.e.: emission at longer wavelengths than the excitation wavelength).

The above has assumed that the transitions by and large occur from the lowest vibrational level of the ground state. Under some circumstances such as illumination by an intense source of excitation from a laser transitions may take place from higher vibrational levels. This arises from the self heating of the sample, or re-excitation following a prior deactivation transition to one of the higher ground state vibrational levels.

### 2.1.2 Fluorescence spectra

In general the emission spectra are red shifted from the corresponding absorption spectra. It can be seen from Figure 2.1 that the highest energy transition from the singlet state to the ground electronic state is a 0 – 0 vibrational transition. Since all transitions to the ground state start from this level this determines the maximum energy of the fluorescence. The difference in the energy of the emitted photon differs from the absorbed by an amount equal to the phonon energy losses in both the excited and ground electronic states.

In general molecules will have several other higher energy level singlet and triplet electronic levels compared with the simplified model presented in the diagram. These levels do not normally contribute to the observed fluorescence (or phosphorescence). For a molecule excited to a higher electronic state it will undergo a rapid series of internal conversions to the first excited singlet state. Since this process takes place at a rate of about $10^4$ times that of fluorescence there is a low probability of fluorescence taking place from the higher levels.

The fluorescence spectrum of a particular fluorophore will therefore be largely independent of the wavelength of the excitation radiation. Riboflavin for example has three distinct absorption bands in the UV and visible regions as shown in Figure 2.2, however only a single emission spectrum centered on 525 nm is observed, corresponding with the lowest excited level – ground state transition.

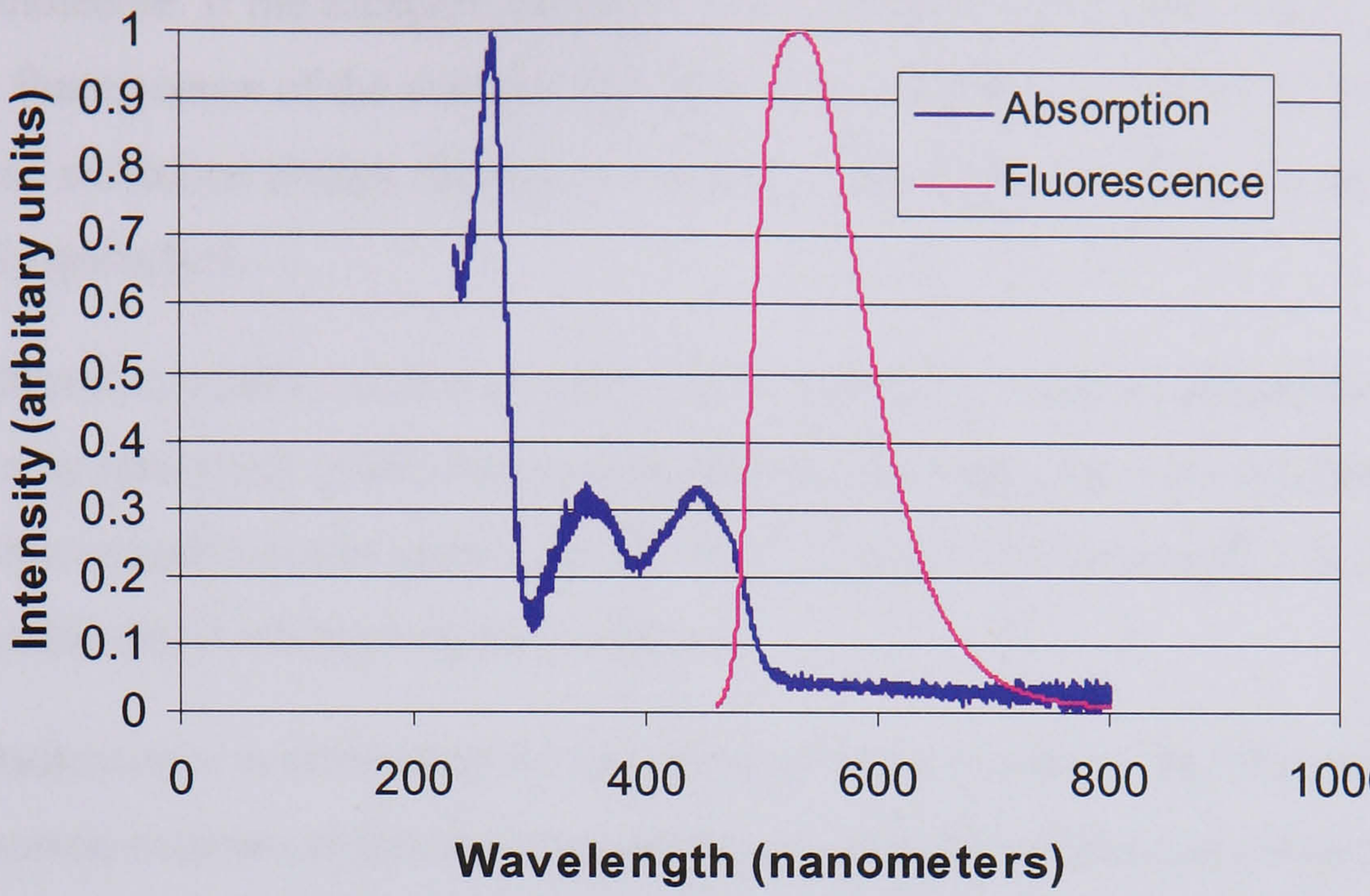



**Figure 2.2 Absorption and emission spectra of riboflavin showing the three absorption bands and a single emission band centred at 525 nm. Both curves have been normalized to their maximum value [18].**

### 2.1.3 Quenching

For the fluorophore molecules, fluorescence is the primary pathway by which the deactivation from a high level excited state takes place. However under some circumstances competing non-radiative mechanisms play a role in the conversion of this excited state energy into thermal energy of the surrounding solvent molecules. Under conditions where this thermal pathway becomes dominant the radiative emissions are severely reduced and the fluorescence is quenched.

A common reason for fluorescence quenching is collisions occurring between excited state molecules and chemical quenchers of the surrounding protein matrix, or from encounters with solvent molecules. One commonly present collisional quencher is oxygen, which because of its small size is able to penetrate the protein structure to some extent. The normally low level of collisional quenching indicates that the fluorescent residues are buried within the protein structure. Conformational changes to the structure of the protein, following intense illumination with a laser for example, may result in these residues being moved near the surface of the protein giving rise to increased collisional quenching of the now exposed fluorophores.

Quenching can also occur when a second (or third) fluorophore is present with an absorption spectrum that overlaps that of the first. If the two types of molecule can undergo transitions of equal energy then the excited state of one molecule, the donor, can transfer energy to an acceptor molecule. In this way fluorescence that may normally be observed from the donor molecule will be quenched by the non-radiative transfer of energy to the

acceptor molecule. If the acceptor molecule itself fluoresces the original energy may now appear as fluorescence of the acceptor molecule [19]. Alternatively if the acceptor molecule passes their excitation energy through a chemical pathway then the fluorescence may be completely quenched.

In large macromolecules, such as proteins or enzymes both acceptor and donor energy transfers may take place within the macromolecule. Confirmation that such energy transfers occur within complex macro molecules has been widely studied especially in the case of carotene chlorophyll energy transfer pathways.

Protein fluorescence is dominated by that of the aromatic amino acids. The precise location of the aromatic residues within the macromolecule, besides influencing chemical quenching processes as already described, is likely to be strongly influenced by donor acceptor transfers along the protein structure.

Collisional quenching, along with resonant energy transfer processes, can explain why the fluorescence of a macromolecule can be expected to be different from that of its component fluorophores. The three aromatic amino acids are the only residues in protein that fluoresce, see Section 2.2. Measured fluorescence is not generally the sum of the emissions of the native amino acids in the proportion found in a particular protein. In some cases fluorescence from the protein is greater than from the free amino acids, possibly because some of the residues are protected from collision quenching for a particular conformation of the protein. This may explain the difference in fluorescence recorded from vegetative and endospore forms of bacteria as a result of the protein conformational changes that take place during sporolation. In other cases the protein fluorescence may be less, or almost entirely quenched. Increases in static quenching may result from static quenchers being brought into close molecular proximity to the fluorophores as the molecular chain folds. In chloroplasts fluorescence may cease entirely, with de-excitation of the fluorophores taking place through resonant energy transfer pathways associated with the chloroplasts photosynthetic processes.

Ultimately the balance struck between such quenching processes and the level of fluorescence observed will depend both on the protein's immediate environment and the conformational form taken up by the folded polypeptide chain.

## 2.2 Fluorescence of biological compounds

Common biological compounds are known to be strong chromophores over a wide range of absorption wavelengths: amino groups absorb in the deep UV (195 nm), β-carotene absorbs strongly in the green and the spectra of the various haemoglobin derivatives have absorption

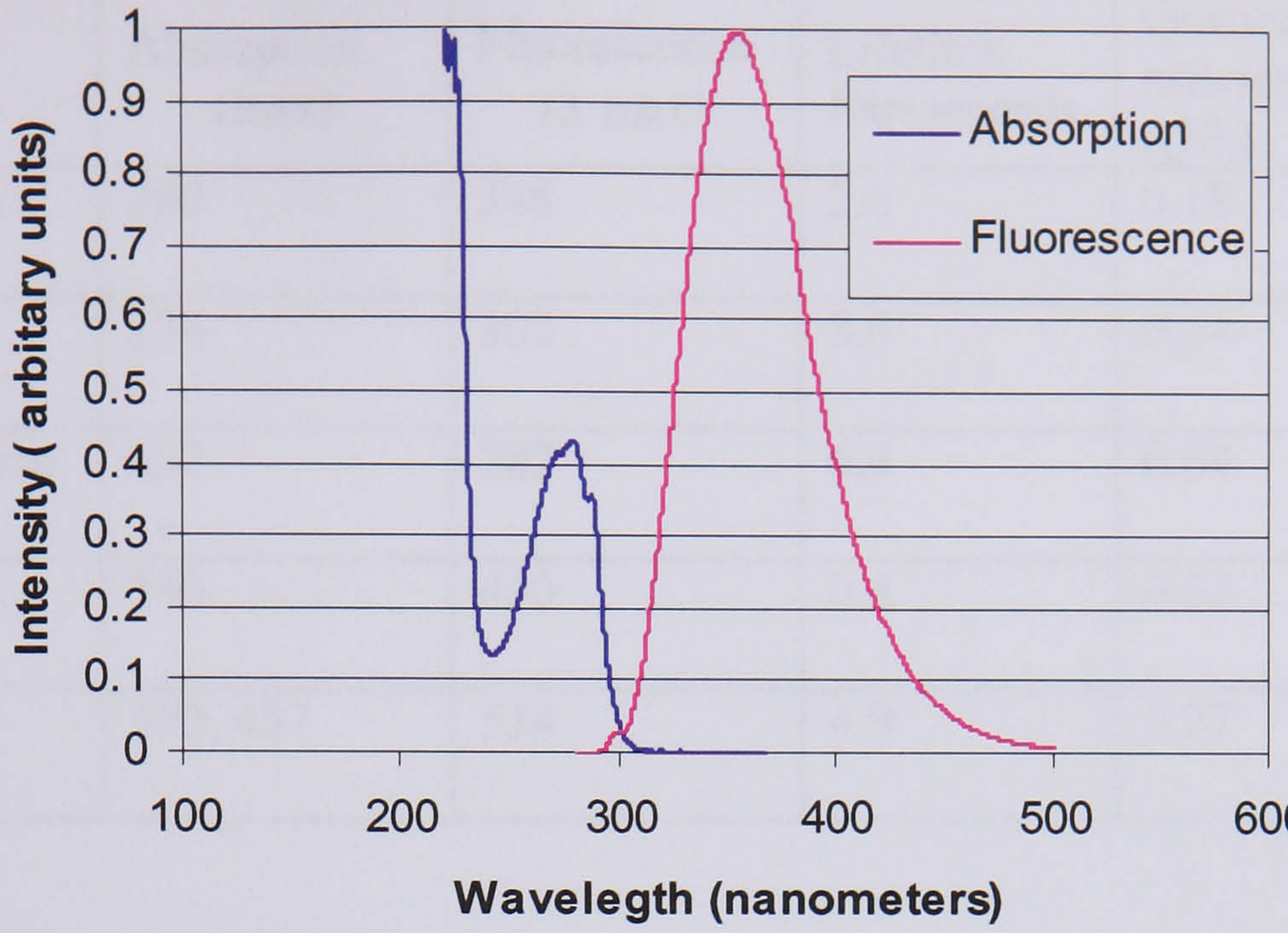



**Figure 2.3 Shows characteristic absorption and fluorescence spectra recorded for tryptophan in solution. Both curves have been normalized to their maximum value [18].**

up to 650 nm. The majority of these groups normally lose there excitation energy via one of the non-radiative mechanisms described above.

For a small group of chromophores fluorescence is the dominant means of relaxation and these are of particular interest in this work. Brown [20] lists the common fluorophores and their respective excitation emission spectra. These include the aromatic amino acids, phenylalanine, tyrosine and tryptophan, the nucleotides NAD(H) and NADP(H), flavins (riboflavin) and chlorophylls. These fluorophores are fundamental to living organisms and their detection within an aerosol can be used as a marker for the presence of biological materials[b].

Figure 2.3 shows the absorption and emission spectra recorded for the amino acid tryptophan. The peak value for both absorption and emission has been normalised to one to make the relative form of the curves clear. The true relative fluorescence quantum efficiency ($QE_f$)[c] can be seen in Table 2.1. The Table summarizes the absorption emission characteristics of the aromatic amino acids along with NADH.

b. These molecules are of a highly complex structure and are not found to exist freely within the environment dissociated from their progenitor organism. Consequently their detection can reasonably be taken as a strong indicator of the presence of biological materials.

c Quantum efficiency is expressed as the ratio of the number of photons absorbed to the number of photons emitted through fluorescence.

| | Absorption (max) | Fluorescence ($\lambda$ max) | Lifetime Nanoseconds | Quantum efficiency ($QE_f$) |
|---|---|---|---|---|
| Tryptophan | 280 | 348 | 2.6 | 0.18 |
| Tyrosine | 274 | 303 | 3.6 | 0.14 |
| Phenylalanine | 257 | 282 | 6.4 | 0.04 |
| NADH | 340 | 450 | 0.4 | 0.02 |
| Riboflavin | 373, 457 | 534 | 4.9 | 0.27 |

**Table 2.1 Summary of the fluorescence characteristics of the aromatic amino acids, NADH and the co-enzyme riboflavin. The wavelength $\lambda$ is in nanometers.**

Many factors, such as the level of exposure of the fluorophore molecules, the degree of sporolation that has occurred and the method of sample preparation, amongst others will influence the $QE_f$ of entire biological organisms compared to that of chemically pure fluorophores in solution. Kunnil *et al* [21] have attempted to quantify the $QE_f$ of dried *B.globigii* spores (washed and unwashed) when fixed to a quartz substrate. They recorded $QE_f$ of ~0.06 per spore which is roughly a third the $QE_f$ of tryptophan as given in the table. The unwashed spores showed a significantly higher $QE_f$ approaching that of tryptophan in solution. This result gives an indication of the significance of the culture medium and other cell debris might have on the level of fluorescence recorded from complete spores.

### 2.2.1 Determination of Excitation Emission Matrix

Li *et al* [22] demonstrated that with multi-wavelength excitation from a mercury lamp it is possible to spectrally resolve the key fluorophores found in bakers yeast cultures. Dalterio *et al* [23, 24, 25] investigated the intrinsic fluorescence of five species of common bacteria. All the bacteria exhibited a broad emission band between 280-580 nm with an intense peak at 330-340 nm. The absorption/excitation band is centred at 290 nm. Li *et al* identified the excitation emission matrix (EEM) peaks as being those from the tryptophan protein L-tryptophan. This EEM agrees well for that of pure tryptophan in solution as reported in the literature and summarised in Table 2.1. Although phenylalanine and tyrosine were known to be present in the bacteria their presence could not be detected from the EEM. These compounds are known to fluoresce strongly in their free form. This lack of measurable fluorescence probably arises from excited energy level transfer effects occurring in the macromolecules. This occurs when these fluorophores transfer most of their excitation energy to an acceptor molecule, such as tryptophan, which may then itself fluoresce.

Kinkennon [26] tested both gram positive and gram negative bacteria and found that for excitation between 250 and 300 nm emissions are dominated by protein fluorescence, especially from tryptophan residues. EEM for excitation between 340 and 460 nm for five different genera of bacteria all resulted in similar spectra showing limited fine detail. The magnitude of fluorescence observed at the longer wavelengths was 5 to 25 times less than that recorded from tryptophan.

Seaver *et al* [27] report that with excitation between 270 and 560 nm (5 nm resolution) there is no spectroscopic difference in the EEM for Escherichia coli (gram negative) and Bacillus subtilis (gram positive). They were unable to observe any emission from the co-enzyme NADH.

### 2.2.2 Fluorescence cross-section

There is little agreement in the published literature on the fluorescence cross-section[d] for any of the commonly investigated bacteria. Values of peak emission cross-section range from $10^{-17}$ to $10^{-12}$cm$^2$/(particle.nm.Sr). An investigation by Seaver *et al* (ibid.) of the fluorescence cross-section of Bacillus subtilis (*B.subtilis*) and Escherichia coli[e] (*E.coli*) showed that total emission stayed constant for excitation wavelengths between 270 and 280 nm and then decreased monotonically with increasing wavelength.

Experiments performed by Cheng *et al* [28] on aerosolised bacterial samples show broadly similar results. They observed an increase of an order of magnitude in the fluorescence cross-section between the vegetative *B.subtilis* and the endospores. The reason for this is not known but may be a consequence of the redistribution of proteins that occurs during the sporolation process.

The only reported comparison between fluorescence cross-sections for bacteria in both aqueous suspension and aerosols has been carried out by Faris *et al* [29]. For *B.subtilis* spores they found the fluorescence cross-section in aqueous solution was four times greater than for the dry aerosolised spores. The magnitude of fluorescence from samples containing single spores was 2.4 times greater than that from clumped spores. The authors attribute this to improved absorption and emission efficiency of the individual spores over spore clumps.

d. The fluorescence cross section is the apparent area of the particle that intercepts the incident radiation resulting in the production of a fluorescence photon per unit solid angle (i.e. within a given cone of observation). The fluorescence cross section is not normally the same as the geometric cross-sectional area. See also section 3.1.1 Mie calculations for scattering cross sections (C*sca*)

e. As with the majority of studies these experiments were carried out using vegetative cultures in solution.

### 2.2.3 Excitation intensity / Photodegradation

Even with high levels of illumination only a small percentage of the molecules are normally excited at any one time. It might therefore be expected that increasing the intensity of the illumination would result in a corresponding increase in fluorescence. Recent studies with individual biological organisms suggest that this may not be true, possibly because of the very small number of fluorescent molecules present.

Faris *et al* found that fluorescence of *B.subtilis* spores did not scale linearly with incident laser radiation power. The loss of linearity occurred at ~ 10 mJ/cm$^2$ during illumination with 10-15 ns pulses at 280 nm. They also observed a 25% reduction in fluorescence for aqueous samples when exposed at low energy densities (<1 mJ/cm$^2$) for long periods (300s). They postulate two, unspecified, photodegradation processes related to integrated UV energy and peak intensity to explain these observations.

Seaver [27] [30] found that to maintain a reproducible fluorescent spectrum, pulse energies (40 ns/pulse) had to be limited to ~0.5 mJ/pulse. At this pulse energy the photon density (~1x10$^{17}$ photons/cm$^2$) approaches saturation of the tryptophan absorption.

Pan *et al* [15] observed additional lines in the spectrum of *B.subtilis* spore conglomerates excited at 266 nm at energies above 0.1mJ/pulses. They suggest that these additional lines may result from laser induced plasma. They conclude that increasing the laser intensity does not improve their S/N ratio but degrades the fluorescence spectrum.

### 2.2.4 Interference

The process of fluorescence is not restricted to the biological compounds. For example cigarette smoke has been shown to have a broad fluorescence spectrum when excited at 363 nm. Other materials commonly found in aerosols that are known to fluoresce include clays[f], aviation fuels and the polycyclic aromatic hydrocarbons (PAH) from both anthropogenic and natural origin. The excitation spectrum for PAHs falls between 240 and 310 nm with emission from 350 nm upwards for $\lambda_{exc}$ of 310 nm [31]. Some spectra of pure aromatic hydrocarbons (only carbon and hydrogen, without non-hydrocarbon substituted compounds) are shown in Figure 2.4, along with that of tryptophan. All the spectra have again been normalized to a maximum value of one, making a comparison possible. The PAH spectra are recorded from solutions of pure aromatic compounds in a non-polar solvent [32]. One of the ways in which PAHs are most likely to be encountered as aerosol

f. No published spectra for the common constituents of clays have been found with an EEM at relevant wavelengths.

particles is as a result of the burning of fossil fuels in the internal combustion engine. These are likely to be non pure aromatic hydrocarbons that have condensed onto soot particles. The fluorescence of these more complex aromatic compounds can exhibit significant shifts and modifications in their fluorescence spectra. Nevertheless the pure PAH spectra of Figure 2.4 give an indication of the interference caused by the presence of PAHs, potentially masking the fluorescence from biological particles.

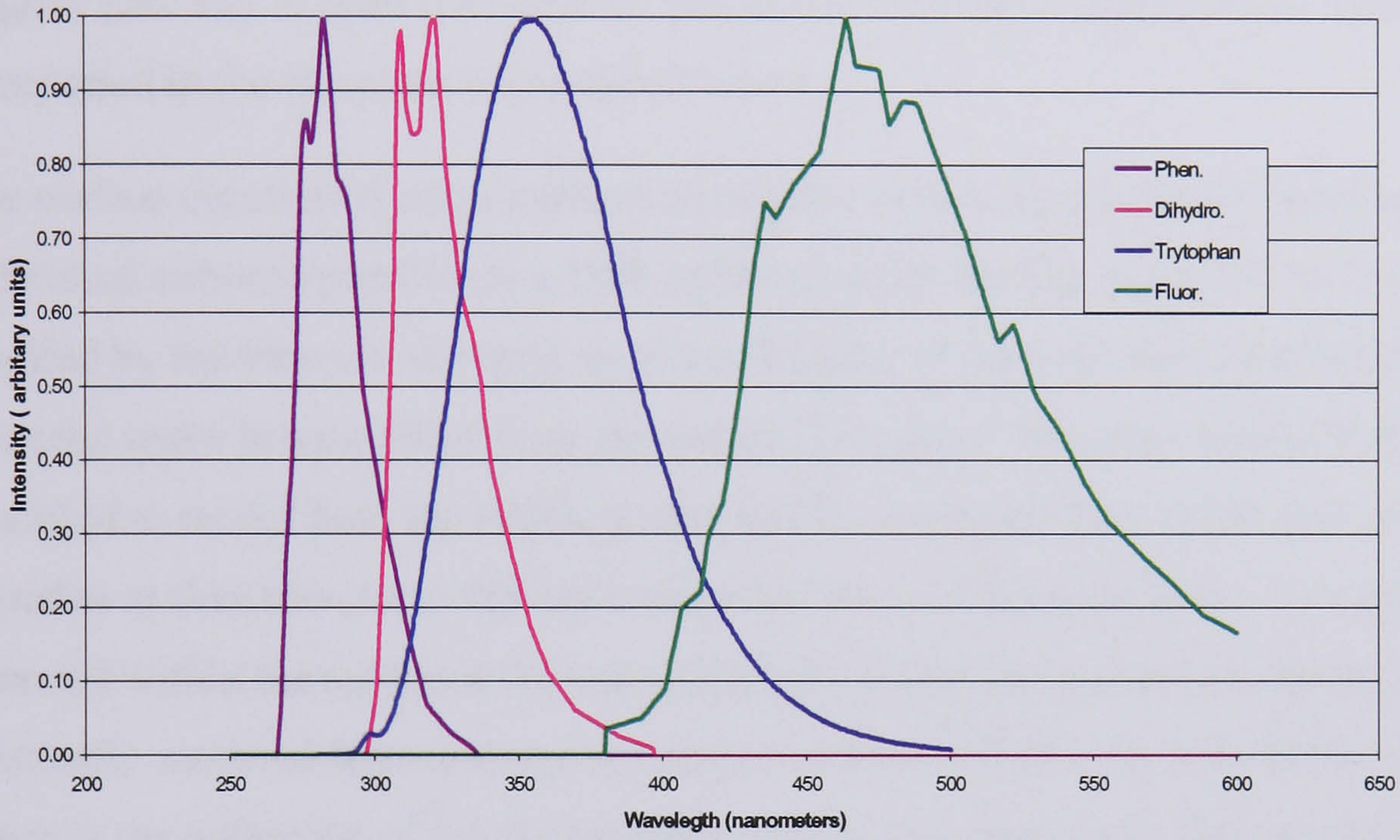



**Figure 2.4 Fluorescence spectra from the pure PAHs Phenylcyclohexane (λexc.265 nm), 9,10 – Dihydrophenanthrene (λexc.265 nm), Fluoranthene (λexc.288 nm) and the aromatic amino acid tryptophan.**

## 2.3 Intrinsic fluorescence detection Instruments

Over the past ten years, several systems designed to detect intrinsic fluorescence from aerosols have been described in the literature. The ultimate aim in the development of these instruments has been to make use of the intrinsic fluorescence of bioaerosols, along with other parameters of the aerosol particle to discriminate between biological and non-biological aerosols.

These instruments have been developed as research instruments with one later being developed into a commercially available instrument. Results have been described for laboratory tests using both standard and fluorescent dyed polystyrene latex (PSL) spheres along with various bioaerosols. For some instruments the results of field trials have been reported. Although some field trials have been undertaken in the UK currently the only published data comes from trials undertaken by the US and Canadian Army. The main test site used by the US Army is a large flat high desert environment with low background

aerosol concentration, typically 20 counts/l [33] in the size range 0.5-20 μm. In these trials, large numbers of known bacteria were aerosolised by course techniques, e.g. agricultural spray equipment generating clusters of the released organisms. This enabled back-to-back sampling to be carried out from both test instruments and conventional impactor measurements.

Currently there are no published results from background studies of bioaerosols for these types of instruments, although Agranovski *et al* have published brief results from trials at a piggery [34, 35]. A brief overview of the instruments and a discussion of their performance as reported in the literature is presented below.

The earliest description of an instrument capable of detecting intrinsic fluorescence from individual airborne particles is a 1995 publication by Pinnick *et al* [36]. This work was inspired by the then recent report by Ho and Fisher of fluorescence detection from a single bacterial spore in a modified flow cytometer. This paper describes a particle counter modified to record both the elastic scatter and fluorescence from individual airborne particles as they traverse a 488 nm continuous wave (CW) laser beam. The sample flow was operated within the cavity of the argon ion laser to provide a greater intensity. Both elastically scattered light and any fluorescence were collected by a parabolic mirror to enhance the collection of the fluorescence signal. The instrument demonstrated the ability to discriminate between standard (non-fluorescent at 488 nm) and fluorescent dyed polystyrene spheres.

This argon ion laser has a wavelength of 488 nm which is near the excitation maximum for the flavin compounds and would also excite the chlorophylls. The instrument was able to detect fluorescence from *B.subtilis* and Bacillus anthracis (*B.anthracis*) spores along with pollens (Puff ball spores). The size range reported for spores exhibiting measurable fluorescence in this instrument was 2 – 4 μm, indicating the fluorescence was from spore clusters. Low levels of fluorescence were observed from non-biological, but naturally occurring, particles such as kaolin and hematite. These non-biological particles are not strongly excited at this wavelength. Their results show that this type of instrument is capable of discriminating single biological from non-biological aerosol particles on the basis of their intrinsic fluorescence.

The same group has built an experimental instrument to measure the fluorescence spectra from a range of different biological particles excited by a 266 nm pulsed laser [37, 38]. In these instruments the sample flow first intersects an arrangement of a crossed continuous wave laser diode beams which are employed to define the sample volume. The elastically scattered light from particles traversing the beams is used to generate a trigger signal to

pulse the 266 nm laser. This beam is displaced from the trigger beam to ensure the particles are irradiated.

This instrument demonstrated the feasibility of recording fluorescence spectra from individual airborne particles in real time. The spectra recorded from Bacillus thuringensis (*B.thuriningensis*) and *E.coli* were of a broad nature, exhibiting no fine structure. In both cases the tryptophan emission dominated the spectra. More recently Pinnick's work has been further developed to allow the use of either a 266 nm or 351 nm Q switched laser to record individual spectra from 2-5 μm biological particles [15]. The published results suggest that multiple wavelength excitation will be necessary to achieve spectral discrimination between different biological materials.

Only a single reference in the literature has been found to the effect of common interferents on the instrument's ability to detect bacteria [39]. In this case the instrument was challenged with side stream tobacco smoke and the fluorescence spectrum compared with that for *B.atrophaeus* [g] clusters. The authors conclude that because both the peak position and the line shape of the fluorescence are similar it would not be possible to distinguish the smoke particles from similarly sized bacteria. This particular bacterium has been widely used in research in this field and this particular variety is generally referred to as BG. This naming convention has also been adopted here.

Hairston *et al* [16] describe a prototype instrument for real time detection of bioaerosols by the simultaneous measurement of aerodynamic particle size[h] and intrinsic fluorescence. In this method the aerosol to be sampled is drawn through a nozzle designed to produce a sharp acceleration in the airflow. Small particles in the flow follow the change in velocity of the air stream, while due to their greater inertia, larger particles accelerate more slowly. This difference is detected by time of flight measurements as the particles cross two closely spaced beams (633 nm HeNe or 680 nm diode) near the nozzle exit. The time of flight data is then converted to particle size by reference to a stored calibration table. A helium-cadmium laser operating at 325 nm is used to excite NAD(P)H (excitation max at

g . *B.globigii* (BG) also known as *B.Subtilis* var. *niger* has been renamed by the International Committee on Systematics of Prokaryotes (ICSP) and is now called *Bacillus Atrophaeus*. However both the previous names are still in widespread use.

h. Aerodynamic size for a particular particle is defined as the diameter of a unit density (1g/cm3) sphere that has the same settling velocity as the particle.

~340 nm)[i], with the fluorescence detected in the range of 420 to 580 nm. The UV beam is aligned to intersect the particle flow just below the red beams. The instrument demonstrated very good discrimination between appropriately dyed PSL spheres and non fluorescent PSL spheres. When tested with BG spores 16.7% of the total particle count exhibited fluorescence, after correcting for spurious noise resulting primarily from the system lasers not being fully attenuated by the optical filters. This compares with 3% of the non-fluorescent PSL spheres being classified as fluorescent[j]. The authors conclude that the test results demonstrate the detection of intrinsic fluorescence from individual spores.

The authors suggest that the presence of NAD(P)H may be an indicator of bacterial viability given the role of NAD in the energy cycle of the cell although they recognise that the determination of viability with any degree of precision is difficult [10]. In their investigation, the target organisms were endospores which by their nature are dormant entities containing very low levels of biological material. In such organisms, metabolic processes are reduced to an absolute minimum. Under such circumstances the detection of fluorescence from individuals may be a poor indicator of future viability. It should be noted that, unlike the infectious agents, many spores do not have to be active to induce allergenic respiratory diseases. Consequently the total concentration and composition of the aerosol including non-viable spores, hyphen fragments etc. is of interest [6].

This instrument has since been developed into a commercially available system the UV-APS (Ultraviolet Aerodynamic Particle Sizer Spectrometer Model 3312. TSI, St Paul, MN, U.S.A.) providing both aerodynamic particle sizing and a fluorescence magnitude signal derived from a pulsed Nd:YAG laser operating at 349/355 nm with a pulse width of 6 ns at 30 mJ/pulse (max). The time of flight signal is used to trigger the pulsed laser at the correct time to irradiate the particle.

Brosseau *et al* [40] have carried out laboratory tests on two UV-APS instruments with both standard and dyed PSL, gram negative and gram positive bacteria in both vegetative and spore forms where appropriate. Both instruments recorded the highest percentage of fluorescence counts (≈63% fluorescent) for BG spores that had been heat treated and where therefore non-viable. The metabolically active vegetative state gave a lower mean

i. They suggest that the presence of NAD(P)H may be an indicator of bacterial viability, although the determination of viability with any degree of precision is difficult.

j. The authors attribute the high percentage of non fluorescent PSL due to cross contamination within the atomisation system which had previously been used for disbursing the fluorescent PSL.

percentage fluorescence count of ≈33%. Overall for the organisms tested they reported no correlation between the culturability of the samples and the percentage count of fluorescent particles.

Currently the origin of the increased fluorescence observed by Brosseau *et al* remains unclear. Although the commercial instruments use a slightly different wavelength from that of the original instrument as described by Hairston, it is still surprising that such large differences in percentage fluorescence counts should be reported from nominally the same bacterial organism. The fact that the highest level of fluorescence was reported from the inactivated spores suggest that the use of fluorescence measurements from NAD compounds as an indicator of viability should be treated with some caution.

This phenomenon has been further investigated by Ristovski *et al* with the UV-APS [41]. In this experiment relative levels of fluorescence from Micrococcus Luteus *(M.luteus)* and *B.atrophaeus* were recorded from both viable and dead (sterilised by autoclaving) samples. They report that the fluorescent signals from the sterilised samples were weaker than for the corresponding aerosols generated from the same suspensions prior to sterilisation. However when they compared particles of the same size they found that those sampled after autoclaving had a larger fluorescence signal. They attribute this to water evaporating from the lysed bacterial cells during aerosolisation since there is no longer a cell membrane preventing loss of liquid from the cells. As a result, the fluorophores become more concentrated within the cell, giving rise to the same level of fluorescence from a smaller particle.

One study has been reported by Agranovski *et al* [27]of the use of the UV-APS instrument to record the concentration of bioaerosols at a piggery. In this experiment the instrument was positioned at various locations both within and around the piggery. The aerosol was subsequently sampled with an AGI-30 impinger and an Anderson impactor from which conventional culturing techniques were used to establish the bioaerosol concentration. High particle counts were recorded with up to $10^6$ bacteria per $m^3$ being reported. The concentration of bioaerosols reported was higher than the viable counts found for the impinger or Anderson impactor.

Seaver *et al* [17] describe an instrument that records both a size parameter based on elastic scattering and intrinsic fluorescence from a pulsed 266 nm frequency quadrupled Nd:YAG laser. In this instrument the elastic scatter signal is generated with a 780 nm CW laser. The elastically scattered signal from single particles delivered by the nozzle system is used to trigger the UV laser. Both the elastic scatter and fluorescence are collected over a large solid angle by a parabolic reflector to achieve high S/N ratios. In field trials [42] the

instrument's ability to detect a low concentration of released known biological organisms from within a low level ambient background was demonstrated. In field trials in which both dust particles (Kaolin) and a bioaerosol, Erwinia herbicola (*Erwinia.h*), were sequentially released, the instrument discriminated between the two materials on the basis of the increased fluorescence recorded from the bioaerosol.

In an attempt to improve the discrimination that can be achieved between different biological organisms Sivaprakasam *et al* [43] have developed a laboratory based instrument using multiple wavelengths to excite intrinsic fluorescence. The instrument takes advantage of the newly available laser sources to interrogate particles with both 266 nm and 355 nm laser pulses. In this instrument particles are directed by a nozzle through the focal point of an elliptical mirror inside an aerosol chamber. A 810 nm continuous wave diode laser is focused to intercept the aerosol particles at the focal point of the mirror. Light scattered by individual particles passing through this beam is collected and focused onto a photodiode by the elliptical mirror. This light scattering signal is used to estimate the size of the particle and trigger the pulsed UV lasers. These pulses are separated in time by 400 ns. Any fluorescence signal excited by these pulses is then detected in three broad spectral bands centered at 350 nm, 450 nm and 550 nm.

Although only initial results have been reported to date, the results indicate that significantly improved discrimination between bacterial spores, vegetative bacterial cells and proteins can be achieved based on the two wavelength excitation approach.

## 2.4 Summary

The current generation of instruments as described fall into two broad classes, those that measure the spectra of the intrinsic fluorescence and those that record the total undispersed fluorescence.

To date, the former group of instruments has only been operated under laboratory conditions. To generate noise free spectra it is necessary to integrate the spectra from a number of particles. As expected these spectra are dominated by the fluorescence from the aromatic amino acids and it is not possible to discriminate between the different species of (washed) bacteria from their spectra. The spectra do however indicate an ability to discriminate between the same bacteria prepared in different ways. The spectra from washed and unwashed *E.herbicola* showed some differences in the longer wavelengths not dominated by the tryptophan peak. It may therefore be possible to discriminate between the same or similar organisms on the basis of the culture media. This may be useful in attributing a possible source to an intentional large scale bioaerosol release.

With the current generation of technology using single wavelength excitation the collection of spectra from single particles is unable to provide greater species to species discrimination than the technically simpler measurement of total undispersed fluorescence intensity.

The second class of instrument mentioned above measures the total undispersed fluorescence over a band of wavelengths determined by the excitation wavelength. This is used with a size parameter, based on elastic scatter from a particle or, in the case of the UV-APS, the aerodynamic size.

Both classes of instrument have a red beam(s) to provide the sizing signal and to trigger the UV excitation source. This introduces the possibility that the particle will not be optimally positioned for the collection optics at the time the UV source is triggered. This arises from the inevitable uncertainty as to the position of the particle when detected by the triggering beam/electronics, slight variations in the particles trajectory within the aerosol stream and a small amount of jitter in the firing of the UV laser.

This is of greatest concern with the UV-APS where the particle travels at a different velocity according to its mass. An additional disadvantage of the pulsed excitation source is that the high energy delivered by the pulse, necessary to induce sufficient excited states to produce measurable levels of fluorescence, can cause photo bleaching [15]. Therefore increasing the excitation intensity in an attempt to improve S/N ratio results in a degraded fluorescence signal.

## 2.5 Conclusion

The instruments described have demonstrated the feasibility of detecting intrinsic fluorescence from single aerosol particles in real time. Although spectra have been recorded, the domination of the tryptophan peak at the excitation wavelengths used has meant that these have not been useful in providing enhanced discrimination between different species of bacteria. These results indicate that total undispersed fluorescence intensity may at the very least be sufficient to discriminate between biological particles and some of the nonbiological particles commonly found in aerosols.

Those instruments where the target fluorophore is the nicotinamide compounds or riboflavin have reported significant differences in the number of particles being reported as fluorescent when challenged with the same organism. Although the spectral studies reported show up differences in the fluorescent response due to variation in bacterial strains, growth states, culture media and preparation techniques, this would not seem to fully account for the reported differences. For those instruments where the target fluorophore is one of the aromatic amino acids, there is a greater agreement between the different reports as to likely

levels of fluorescence even though the same general considerations will influence the results.

Without further fundamental work to investigate the cause of these variations the experimental results reported would suggest that the aromatic amino acids should be the preferred target molecule, for which fluorescence excitation at approximately 280 nm would be optimal.

Also, with pulsed excitation lasers there is a trade off between the desire to increase the level of irradiance to improve fluorescence S/N ratios and the requirement to limit the energy delivered to avoid photo-bleaching and the consequent loss of fluorescence signal. This limitation can be reduced by the use of very long pulse durations (difficult to achieve in practice) or by CW laser excitation at reduced power (where the duration of illumination is determined by the transit time of the particle through the beam) . Excitation energies below those giving rise to photo-bleaching can then be used with the fluorescence recorded from multiple excitation emission cycles.

Overall, therefore, it was concluded that the optimal excitation source for the new MPAM multiparameter instrument would be a CW laser operating as close as possible to 280 nm wavelength. In reality, available laser technologies restricted the choice to a narrow band in the mid-260 nm wavelength range (see Section 4.1.1)

# 3 Particle characterisation by light scattering

A variety of techniques are available for the physical measurement of aerosol particles. The precise technique to be applied in any situation will depend on which attributes of the aerosol are of interest to the observer and often on the limits of the technology currently available. This report describes rapid methods of aerosol characterisation based on light scattering techniques. For a more general overview reference can be made to one of the texts on aerosol measurement techniques, for example Willeke and Baron [44].

Optical scattering techniques can be broadly grouped into two methods, ensemble scattering and single particle scattering. Ensemble scattering records the scattering from a multitude of randomly orientated particles and is able to provide particle size distribution information. However because of the multiple scattering that occurs within the sample this technique is generally unable to provide shape characterisation data. The alternative technique of relying on the detection of scattering from a single particle to a suitably configured array of detectors enables both a size and shape profile to be determined.

Under circumstances where the typical background aerosol is generally known, the characteristic shape of certain types of particles may be used to facilitate their detection. Respirable airborne asbestos fibres are a well known example where the particles characteristic shape is employed in its real-time detection, although further detailed chemical analysis may be necessary to determine exact fibre mineralogy.

In this work, it was intended to use real-time measurement of both the particle's intrinsic fluorescence as well as its size and shape to achieve particle discrimination. The former parameter was dealt with in the previous Chapter; the latter would be achieved through *spatial light scattering*, as described below.

## 3.1 Theoretical background

The spatial intensity distribution of light scattered by an individual particle, its *scattering profile*, is a complex function of the particle's size, shape and orientation with respect to the illumination. Other factors of equal importance are the wavelength and polarisation state of the incident radiation. Commonly the particle's size is expressed by a dimensionless size parameter $\alpha$. The is defined as $\alpha = 2\pi a/\lambda$, where a is the effective radius of the particle and $\lambda$ is the wavelength of the incident light.

If the particle is small compared with the wavelength ($\alpha \ll 1$) and provided its relative refractive index is near to unity and only itself weakly dependent on wavelength then the particle can be treated as an group of oscillating dipoles radiating in phase in all directions.

The intensity of scattering to any particular angle is primarily a function of volume and hence is insensitive to particle shape [45]. Such scattering is historically defined as Rayleigh scattering and the scattered irradiance is simply proportional to $1/\lambda^4$. Consequently, for a radiation source of any particular wavelength a lower limit is set on the size of particles whose shapes can be investigated by light scattering methods.

As the size of a particle increases, such that $\alpha > 1$, full account must be taken of the phase relationships of the scattering from the dipole elements. The increasing possibilities of reinforcement or cancellation of the scattering from the individual dipoles means the scattering becomes increasingly complex, with considerable fine detail emerging in the scattering profile.

Ultimately, for those particles very much larger than the wavelength the much simpler techniques from geometrical optics such as refraction and reflection can be applied to determine the spatial scattering profile and the particle's shape.

The scattering and absorption of radiation observed from any particular particle is of course also influenced by its chemical composition. This is expressed in terms of a material's complex refractive index $N = n + ik$, where the real part represents scattering and the imaginary part represents absorption. Both n and k also depend on the wavelength. If k is equal to 0 at a given wavelength the particle does not absorb radiation at this wavelength. Some typical values of N are shown in Table 3.1. As can be seen the naturally occurring iron oxide haematite absorbs quite strongly, compared to the other materials listed. Values of N for *Erwinia herbicola* are also shown, being typical of the bacteria used in this work. In general, biological materials behave as linear, homogeneous and isotropic media; despite their complex chemical makeup [46].

| **Substance** | **λ** | **n** | **k** |
|---|---|---|---|
| Water | 500nm | 1.339 | 0 |
| Water | 200nm | 1.327 | 9.137E-8 |
| NaCl | 500nm | 1.544 | 0 |
| Haematite | 500nm | 2.6 | 1.0 |
| oleic acid | 632nm | 1.445 | 0 |
| Erwinia herbicola | 500nm | 1.526 | 0.0002 |
| Erwinia herbicola | 266nm | 1.593 | 0.00373 |

**Table 3.1 Refractive indices of some of the materials used in this experimental work.**

### 3.1.1 Scattering from spheres

Homogenous spherical particles can be modelled using Mie theory. The Mie calculation gives the particles scattering cross section (C*sca*)[a] and an exact solution to the scattering profile [45]. Currently, many software applications are available to carry out such modelling. Figure 3.1 is a typical plot of the Mie scattering from a 3μm diameter oleic acid sphere illuminated by radiation of wavelength 266 nm [47]. Perhaps the most important point to note is that the scattering is highly asymmetric in contrast to Rayleigh scattering. This can be seen most strikingly in the linear plot where it is difficult to see the backscattering at all. The scattering into the forward lobes is more than a 1000 times greater than in the back direction. Many practical counting/sizing instruments operate in this forward lobe to take advantage of this preponderance of the scattering in the forward direction. Considering the polar plot one other feature of the scattering becomes immediately apparent. Apart from the very narrow angles of the central forward lobe the scattering intensity has considerable fine structure (in the forward direction) for scattering angles immediately outside the central lobe.

Clearly many particles are far from spherical and while Mie theory or an extended variant of Mie can be used at least for a first order approximation to many scattering problems it is essentially limited to the class of near spherical particles. For the non-spherical particles, even those of a regular form, generally a simplified approach is required in order to arrive at a theoretical scattering profile within a reasonable computational time.

### 3.1.2 Scattering from fibres

Under certain conditions the scattered field can be approximated by treating the particle as a set of independent dipole scatterers and subsequently superimposing the resulting waveforms. In this approximation, known as Rayleigh-Gans theory, the interactions between the individual dipoles are ignored. In this manner it is possible to obtain relatively simple expressions to determine the scattering from fibre-like particles.

Theory predicts that for a vertically aligned finite cylinder illuminated by a horizontal beam the scattering will be confined to the horizontal plane of the beam axis. However if the cylinder is inclined from the vertical the scattering assumes a conic section. As the angle of

a The scattering cross section C*sca* is the area of the particle that interacts with the incoming light resulting in scattering. When the wavelength is similar to the particle's diameter the C*sca* is larger than the geometric cross section of the particle.

inclination is increased the scattering becomes increasingly circular, eventually forming a complete circle.

Hirst *et al* [48] have used Rayleigh-Gans theory to model the scattering profiles from ideal fibres. Their results showed that the scattering profiles are highly sensitive to the particle's orientation to the illuminating beam. For relatively high scattering angles small changes in orientation of the fibre resulted in significant changes in both the shape and overall complexity of the scattering profiles. In contrast the scattering profiles at lower angles were shown to be less sensitive to small changes in the orientation of the fibre.

### 3.1.3 Scattering from particles of more complex or irregular shape

In general, the majority of real world particles are neither spherical or simple fibres. The majority of particles are likely to be heterogeneous in structure and of irregular shape. Unsurprisingly, the theoretical modelling of scattering profiles from these types of particles has proven to be computationally more difficult. However, methods such as Finite Difference Time Domain, Finite Element techniques or the Extended Boundary Condition Method can be used to model such complex particles. These techniques and their field of application is extensively covered in the literature, see for example Wriedt [49] or Jones [50].

Where the primary concern is the identification and characterisation of particles on the basis of a recorded scattering profile, without a priori knowledge of at least some of the parameters that determine the nature of the scattering profile, the theoretical models do not easily provide an inversion of scattering profiles (i.e. the determination of a particle's characteristics from the scattering data) [51].

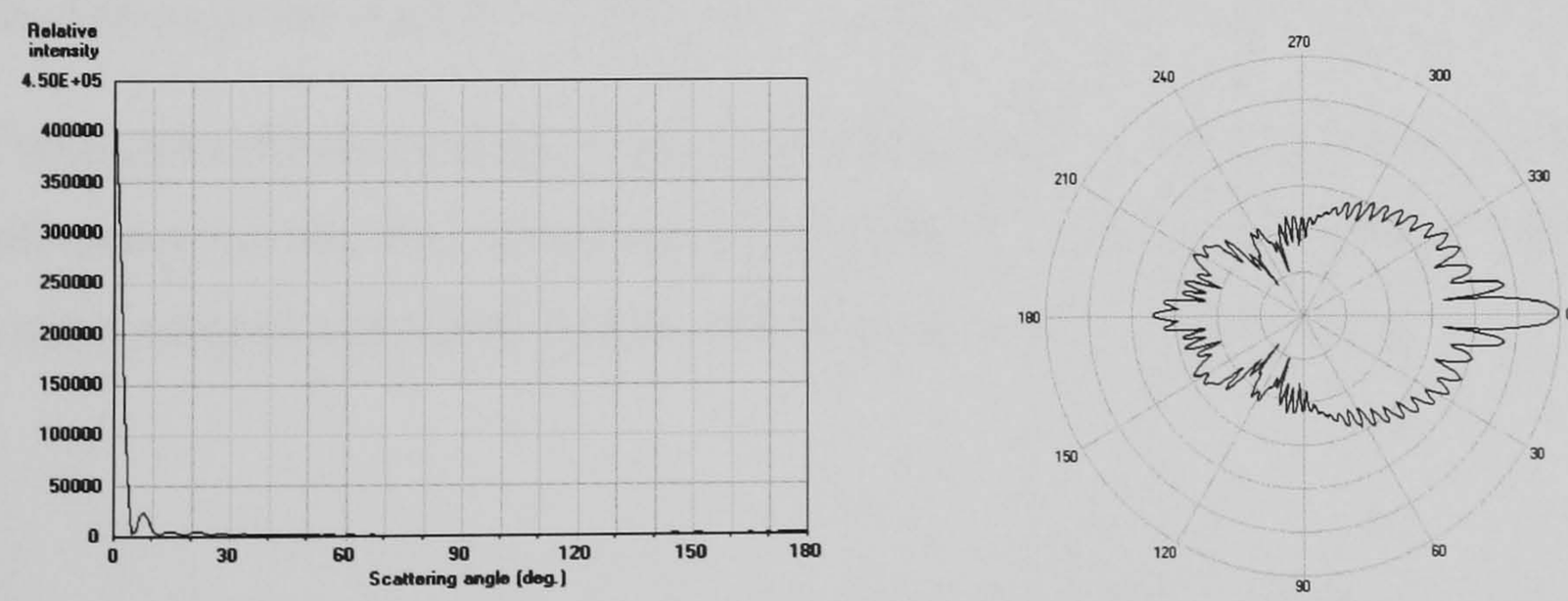



**Figure 3.1 Mie scattering from a spherical 3μm diameter oleic acid droplet illuminated at 266nm. The scatter intensity is in arbitrary units on a linear scale. The right hand figure shows the same data on a polar plot to a log scale.**

Because the theory is complex for particles of widely varying shape, size and orientation to the illumination beam, the problem can also be approached experimentally as outlined below.

## 3.2 Light scattering instruments

The orientation of the scattering system is shown in Figure 3.2. An incident plane wave propagates in the direction of the x-axis, striking the particle centered on the xyz co-ordinate system. The variation in the light scattered may be described in terms of the polar scattering angle $\theta$ (the 0° - 180° scattering in a plane containing the axis of the incident illumination) and the azimuthal scattering $\phi$. For any scattering angle $\theta$, $\phi$ may be 0° to 360° in a plane perpendicular to the axis of the incident illumination.

### 3.2.1 Particle counter/sizers

Optical particle counters measure the scattered light intensity from particles as they transit a laser beam, the intensity of the light scattered by a particle being determined to some extent by its size. Generally counter/sizers have a single detector element. Often this is arranged to collect light from the forward scattering lobe to maximize the signal from smaller particles, as illustrated in Figure 3.1. The magnitude of the detector's output is then matched to the instrument's response via a set of calibration tables derived from known standard particles, usually standard sizes of latex spheres.

Most real life particles are non-uniform in structure and often non-spherical leading to erroneous sizing. Considerable effort has therefore been expended in characterising their response to non-spherical particles. For example Umhauer and Bottlinger [52] investigated the influence of the shape of irregular non-spherical particles on the scattering intensity. This was done by repeatedly moving the same particle, suspended in an oscillating electric field through the sample volume with a continually changing random orientation.

These instruments measure a single parameter and are not intended to furnish further information about the characteristics of particles. Additional attributes must be evaluated if a more complete characterisation of the aerosol is to be determined.

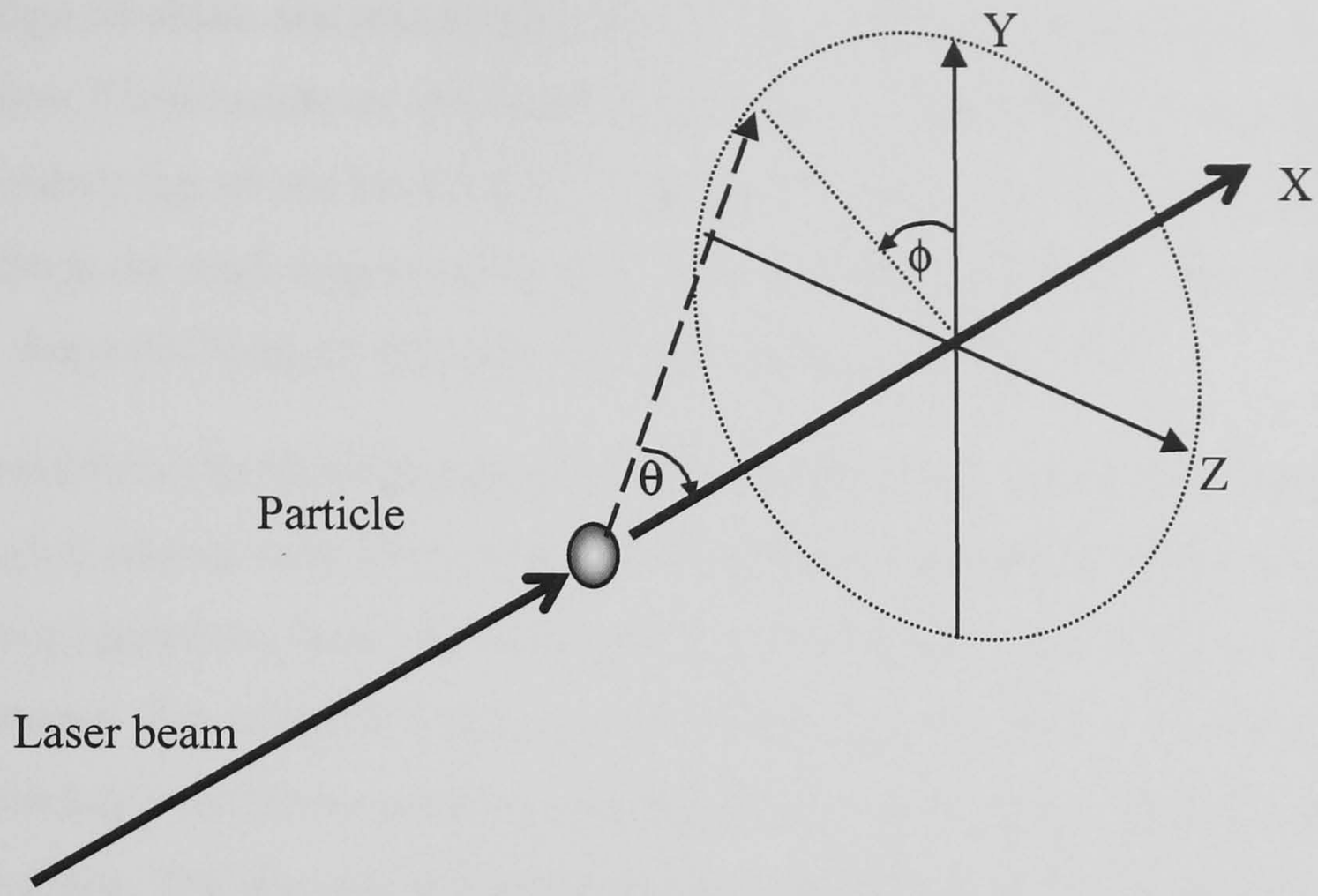



**Figure 3.2 A representation of the polar θ and azimuthal ϕ scattering angles shown relative to a set of arbitrary instrument axis**

### 3.2.2 Instruments intended to investigate both particle shape and size

In the 1980's several groups demonstrated the possibility of discriminating between different types of human blood cells by the simultaneous measurement of various combinations of forward scatter, side scatter and back scatter from individual cells within a liquid flow. For example Sloot *et al* [53] modified a flow cytometer to record the cross polarization state for forward, side and back scatter. They demonstrated that this method had the potential to discriminate between arbitrarily shaped particles.

Spinrad *et al* [54] have demonstrated that it is possible to discriminate between spherical test particles and marine algae, which are variously ovoid, encrusted sphere, cylindrical or irregular pear in shape depending on species. A flow cytometer was configured to simultaneously measure chlorophyll fluorescence, 90°side scatter and forward angle light scatter for both unpolarized and cross polarized scattering. In this instrument the fluorescence was solely used to gate the instrument on the algal and fluorescent PSL test particles within the liquid flow. Their results show that there is a high degree of correlation between the unpolarized scatter and the cross polarized scatter from the more highly aspherical particles.

The majority of experimental work however, has concentrated on investigating the variation in the spatial intensity of the scattered light. Latimer *et al* [55]has modelled the scattering

from a range of oblate and prolate spheroids lying at various inclinations to the source of illumination. Their results showed significant levels of asymmetry in the scattering profiles for spheroids lying off the beam axis. The same group has also calculated the effects of particle shape on small angle scattering ($0° \leq \theta \leq 20°$) for spheroids. They concluded that a particles shape can strongly influence its low angle scattering profile.

This sensitivity in the scattering profiles to a particles shape and orientation has meant that most detailed studies have been restricted to particles that are spherically symmetrical. Some groups however, have exploited this shape sensitivity to discriminate between various particle shapes. For example, Diehl *et al* developed an instrument to distinguish fibrous from individual non fibrous particles suspended in water as they drifted slowly though the sample volume. The instrument measured light scattered to pairs of detectors arranged at scattering angles of ± 45°, ±90°, +45° with -135°, and 45° with 135° as particles passed through the sample volume. By comparing the ratio of detector outputs from pairs of detectors the instrument was able to distinguish fibrous from non-fibrous particles in the sub-2 μm size range.

A major limitation of this type of instrument, where the particle simply drifts through the beam for single particle characterization, is that the recorded scattering intensity distribution is strongly influenced by both the particle's shape and orientation to the illumination. Mungai and Warren [56] have used EBCM to model scattering from non-spherical Chebyshev particles. They conclude that the particle's orientation is the most significant determinant of the scattering profile. Consequently, without some degree of control over the particle's orientation within the sample space, particles of identical shape will have substantially different scattering profiles, making the efficient classification of shape virtually impossible.

Flow-through instruments intended to investigate individual particle shape employ a system of sample delivery that provides some degree of control over the orientation of particles within the measurement volume. In the majority of cases this is achieved by a nozzle system that causes non spherical particles to adopt a preferred orientation with respect to the flow lines.

Bartholdi *et al* [57] describe an instrument that uses "hydrodynamic control" to preferentially orientate particles within a liquid flow. They were able to capture near 360° polar scattering from single particles with an annular section of an ellipsoidal reflector which directed scattered radiation onto an array of photodiodes. Their experimental results closely matched the modelled symmetrical scattering from test spherical particles. They

conclude that any asymmetry detected in the scattering could be used in discriminating among different biological cells.

Various instruments have been developed to investigate the shape of particles suspended in an air stream. Typically the particles are delivered into the sample volume entrained in a fine laminar flow such that the more elongated particles tend to align axially with the flow.

An instrument intended to classify individual airborne particles rapidly was developed by Ludlow [58]. This consisted of a spherical scattering chamber the centre of which was coincident with the intersection of a laser beam and an orthogonally aligned sample flow. Three photomultiplier detectors were mounted symmetrically around the beam axis at the single scattering angle of 60°. By comparing the output from the detectors the instrument was able to differentiate between spherical droplets and non spherical particles.

This concept has been further developed by, Kaye *et al* [59] by incorporating an ellipsoidal reflector that allows a much greater proportion of the scattered light to be collected, as illustrated in Figure 3.3. Three PMT detectors are arranged symmetrically around the beam axis and record azimuthal variations in the scattering profiles. Light scattered between the polar angles of 27° and 140° is incident on this mirror and is focused onto the three PMTs. The signal falling on each detector is therefore essentially the integral of the scattering over a wide polar range. By comparing the output of the three detectors, spheroidal particles with aspect ratios as low as 1.08:1 could be discriminated from true spheres. Any light scattered in the forward direction between the angles of 4° and 27° is also focused onto a single PMT and this is used to ascribe a size parameter based upon the total equivalent scattering from a perfect sphere.

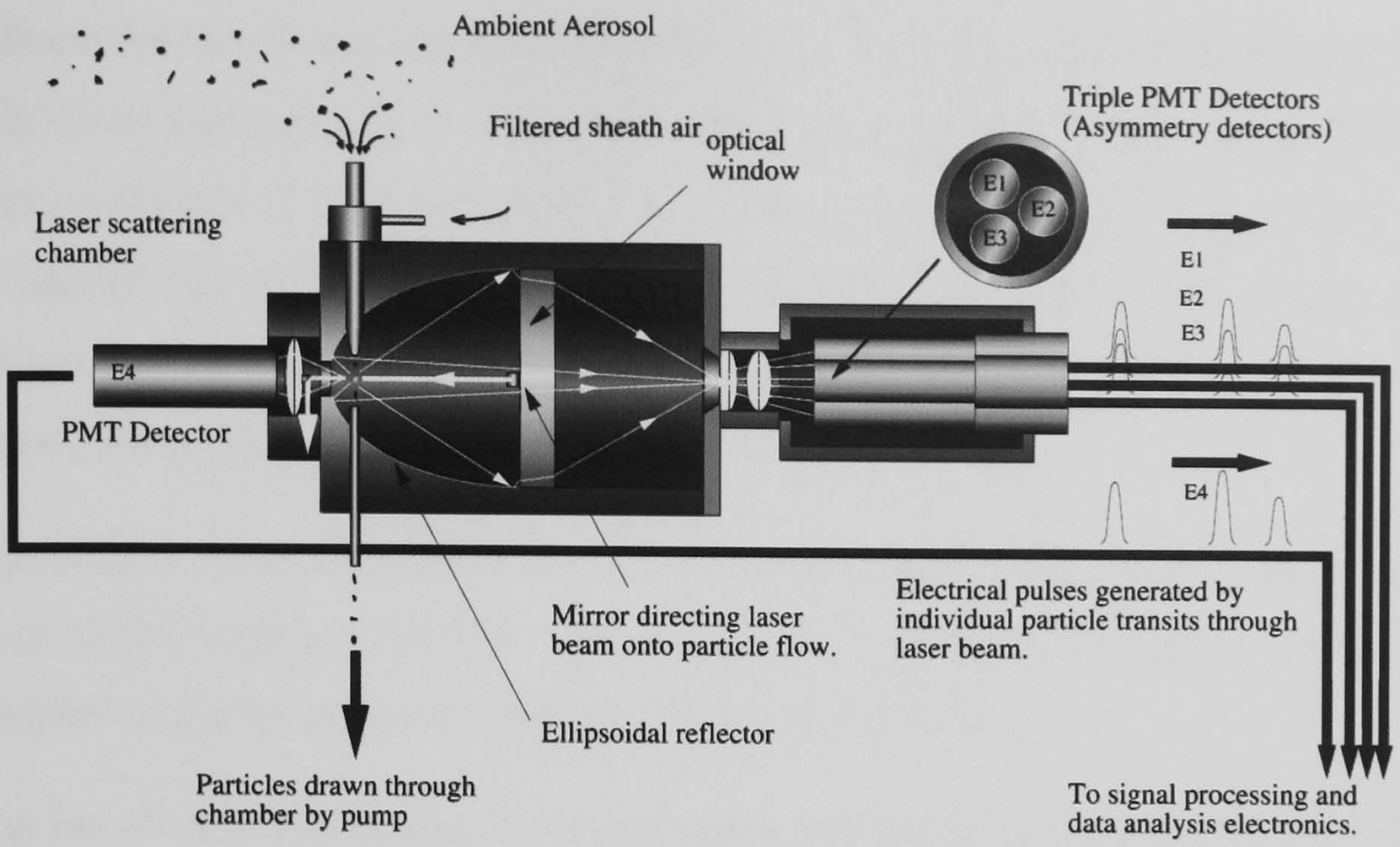



**Figure 3.3 Schematic diagram of a scattering chamber incorporating an ellipsoidal reflector that allows light scattered between the polar angles of 27° and 140° to be collected .**

Each particle thus generates four signal pulses, which are fed to the acquisition electronics. Data processing is then used to drive a graphical interface screen providing both shape and size information. The use of just four detectors allows fast operation (~10,000 particles/s) whilst giving an indication of scattering asymmetry, related to the particle's shape. However, although the instrument was able to provide a basic particle shape metric the combination of a small number of detectors (i.e. limited spatial resolution) meant it was unable to realize the full potential of spatial scattering profiles in the characterization of non-spherical particles.

This limitation was overcome to some extent by the development of the Dual Amplitude Weighted Nephelometer-Aerosol (DAWN-A) scattering chamber [60] (Wyatt Technology Corporation, Santa Barbara, CA, USA). This instrument consists of a sealed spherical chamber, the centre of which is coincident with a laser beam and an orthogonally aligned sample flow. The chamber has 72 small and two large detector ports. These small ports are arranged on four azimuthal planes orientated at 45° intervals. Each small aperture has a plug to seal the chamber or alternatively it can accept an optical fibre probe to convey scattered light to a discrete PMT and its associated signal processing electronics. The two larger apertures subtend a large azimuthal angle and are centred on the polar scattering angles of 25° and 155° respectively. The instrument thus provides a flexible test bed for the investigation of both polar and azimuthal scattering. Provision is made to calibrate the response of the different ports and once this is done the instrument has shown excellent

agreement for the +θ and the -θ values indicative of spherically symmetric scattering from spherical test particles (e.g. 0 .5 μm PSL). In contrast, the scattering from cells of the bacterium *Bacillus Atrophaeus* gave rise to significant differences in the scattering recorded between individual cells resulting from differences in both size and orientation with respect to the beam direction. The bacteria's inhomogeneous structures also produced dramatic asymmetries in the scattering.

The flexibility of the instrument has led to it being used by many researchers. For example Dick *et al* [61] used the DAWN-A instrument to investigate the refractive index of atmospheric organic carbon at different relative humidities.

Whilst the DAWN instrument has undoubtedly provided a powerful platform with which to study spatial scattering, the quantity of data it can capture is still limited by its use of a relatively small number of discrete detectors located at preset angles.

To further extend the fundamental understanding of the spatial light scattering, Hirst and Kaye [62],[63 ] have developed an instrument that uses an intensified charge-coupled device (CCD) camera to record high resolution scattering profiles from individual particles. The scattering chamber design is similar to that described in Figure 3.3, but with the triple PMT detectors replaced by the CCD camera. The so called 'wide-angle' scattering image captured by the camera was therefore a two-dimensional transformation of the three-dimensional spatial intensity distribution scattered to the ellipsoidal reflector. This represent approximately 83% of the of the 4π sr of scattering about the particle.

The examples shown in Figure 3.4, recorded using this CCD system, illustrate the variations these wide-angle scattering patterns can assume for different particle types and gives an indication of the potential for particle classification from high resolution scattering profiles. In each image, the inner dark disc corresponds to the hole in the rear of the ellipsoidal reflector in Figure 3.3 (27° scattering angle), and the outer perimeter of the image corresponds to 140° scattering angle. The vertical dark dumb-bell is the shadow of the aerosol inlet and exhaust tubes, the circles being at 90° above and below the particle respectively.

Young-Le Pan *et al* [64] have employed a CCD camera system also based on an ellipsoidal reflector arranged to capture scattering profiles over the range of $\theta = 48°$ to 164°, $\phi = 0$ to 360° covering 63% of 4π sr of the total sphere. The use of a high speed pulsed laser (70 ns) to illuminate the particle made it possible to record high resolution camera images from individual aerosol particles. They compared the scattering profiles from PSL spheres compared to those from single vegetative BG spores. The highly detailed scattering profile

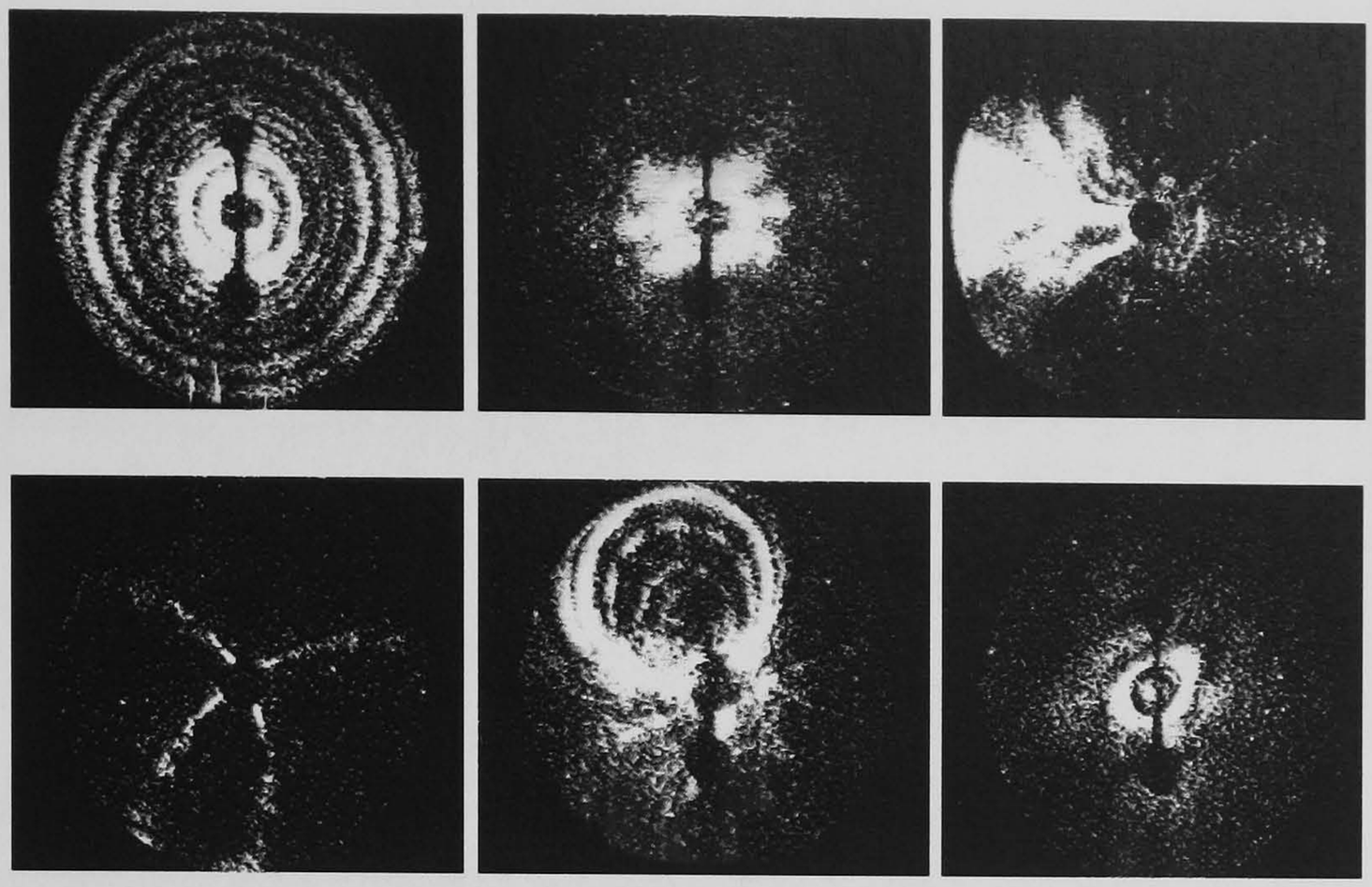


**Figure 3.4. Examples of wide-angle spatial light scattering patterns recorded on CCD camera from various particles. Clockwise from top-left: 10 μm water droplet; doublet of 1μm PSL spheres; copper flake (~5 μm); salt crystal corner-on (~3 μm); fibre tilted from vertical; 2 μm ellipsoidal particle.**
(Particle instruments research group library images U.H)

from the spores reflected more complex surface detail and structural homogeneity of the spores compared with the typical scattering profile from the spheres. Scattering profiles from clusters of both (PSL) spheres and spores were also captured. The high resolution fine detail in these images makes distinguishing between them by eye readily possible.

Camera systems based on the capture of such detailed images provide important information on the detail scattering from different materials and particle types. They allow the identification of salient features of scattering from a wide range of particle morphologies. However, for instruments intended to run in real time they have some limitations. Firstly, it is not currently possible to have a suitable camera system running at the speed required for real time studies. Secondly, even if the cameras could operate at such a rate the quantity of data produced would overwhelm the present generation of data processing algorithms and processing hardware.

### 3.2.3 Low angle scattering instruments

Hirst *et al* [65 ],[66] have developed a prototype asbestos monitoring instrument, SAFIRE (System for Airborne Fibre Recognition). In this instrument light scattered between the scattering angles of $\theta = 5°$ to $30°$ is collected to form the scattering profile. Figure 3.5 shows a schematic of this, so called 'low angle' scattering chamber used in the SAFIRE instrument. They have shown that in the specific case of asbestos fibres, even at low

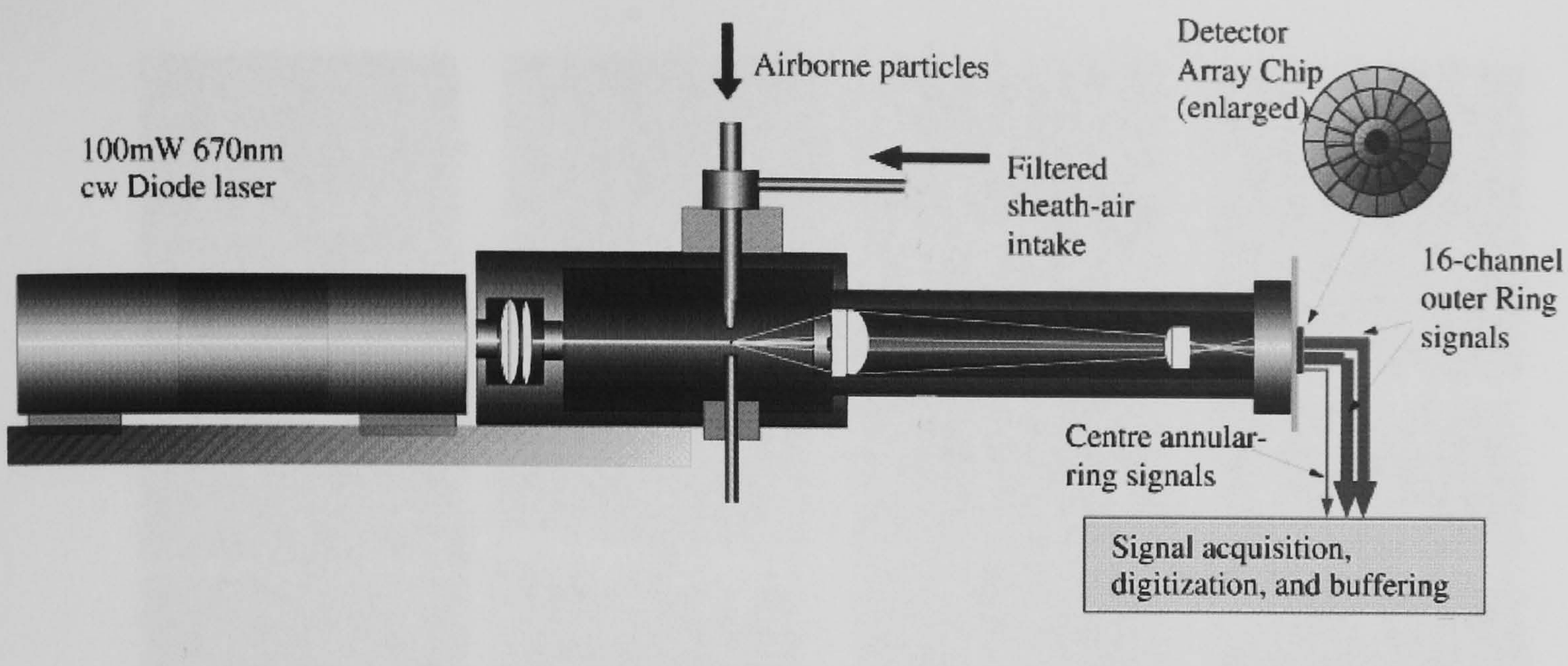



**Figure 3.5 Schematic representation of the low angle prototype of the SAFIRE fibre monitor scattering instrument.**

scattering angles (i.e. less than 30°) the scattering profiles showed characteristic features sufficient to allow the differentiation of fibres from other particle types.

This development followed an initial investigation of scattering from asbestos fibres with the camera system previously described in section 3.2.2. For example, the top row of Figure 3.6 shows wide-angle scattering images from individual airborne crocidolite asbestos fibres recorded using this system. The left-most image shows scattering in the form of a horizontal disc containing the beam axis, as expected from vertically aligned fibres. However, if the fibre is inclined from the vertical by only a few degrees the scattering assumes a conic section, as shown in the central images. Ultimately as the inclination increases the scattering forms a complete circle as in the right most image.

In contrast the low-angle camera images recorded on a prototype of the SAFIRE instrument, shown in the second row of Figure 3.6, illustrate the predominantly horizontal scattering exhibited by the crocidolite fibres and are far less susceptible to the pronounced conic section variations seen in the wide-angle profiles.

Since, in any instrument employing a laminar-flow aerosol delivery system of the type used in the above examples, elongated fibres are always likely to be inclined by a few degrees to the vertical (the beam lying in the horizontal plane as shown in Figure 3.2) the use of wide-angle collection optics can lead to the wide variation in scattering profiles. However, by restricting the scattering profiles to the low-angle range, the presence of pronounced conic sections in the profiles is avoided, thus reducing the variation in the scatter profiles and making the task of classifying scattering patterns significantly less difficult and computationally more efficient.

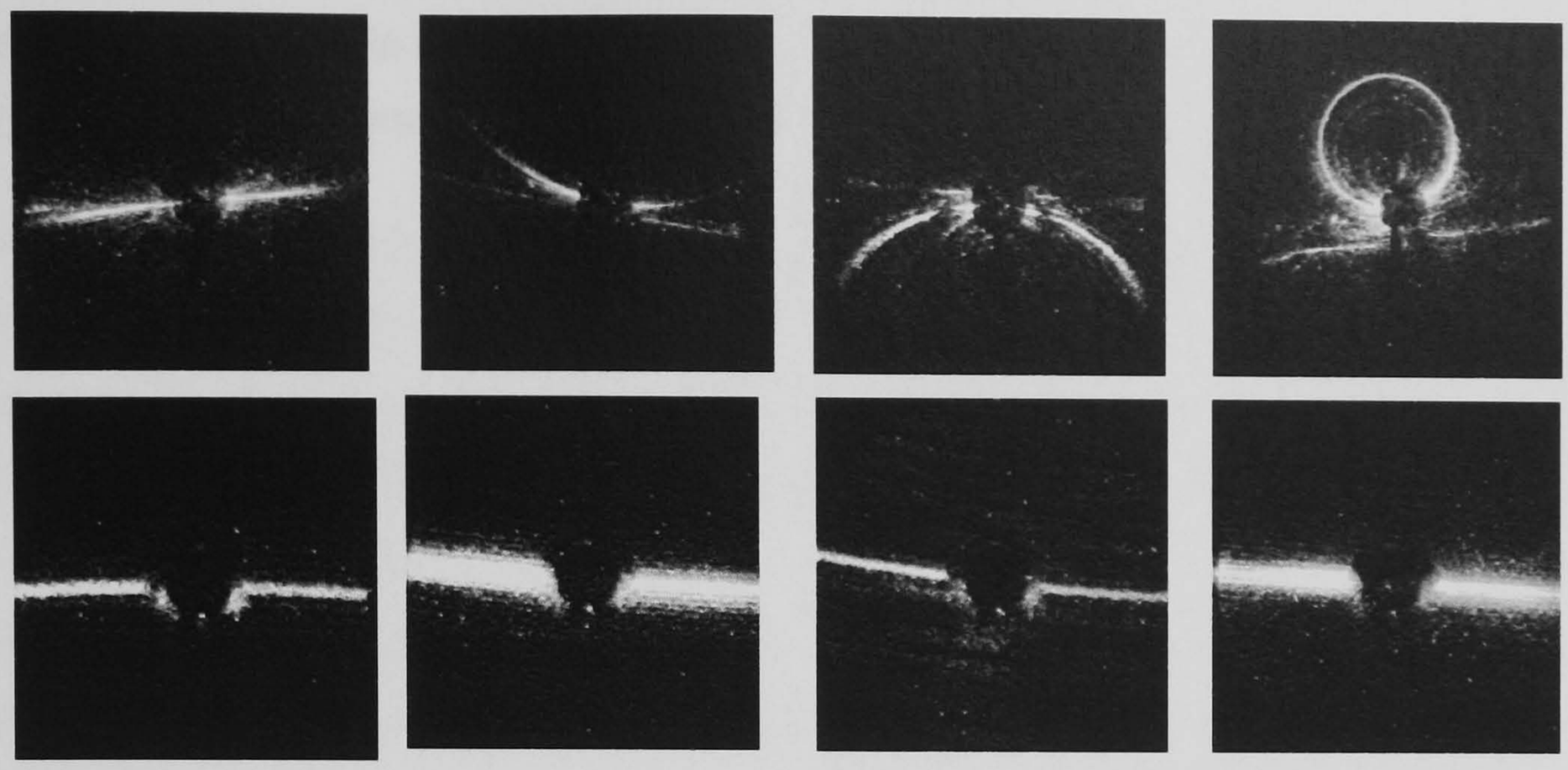


**Figure 3.6 The top row shows examples of scattering profiles from individual airborne crocidolite asbestos particles captured using a wide-angle camera system.**
**Bottom: Similar scattering profiles recorded using a low-angle camera instrument.**
(Particle instruments research group library images U.H)

Hirst (*ibid*)was also able to show that by operating at low angles in the near forward direction, the SAFIRE instrument had two further advantages:

- The simplified low-angle scattering profiles meant that a high-speed low-resolution multi-element detector could be used; a prerequisite for any instrument classifying in real-time, such as might be required to monitor for potentially pathogenic airborne biological particles.

- A highly simplified opto-mechanical design suitable for use in a practical fieldable instrument could be used since the complexities associated with the use of an ellipsoidal reflector could be avoided.

## 3.3 Experimental comparison of wide-angle & low-angle scattering.

As outlined above, there were clear mechanical and speed advantages to be gained by using low-angle forward scattering pattern analysis in the assessment of particle shape. However, historically it had been thought that the relatively narrow forward angle scattering would be largely insensitive to particle shape.

For example, Zerull *et al* [67] used microwave radiation to investigate the scattering from a range of spherical and non spherical particles having size parameters $1.9 < \alpha < 17.8$.

Considerable variations in scattering intensities were observed for side ($\theta = 55° - 160°$) and back scatter ($\theta = 160° - 180°$) but not in the forward direction ($\theta = 5° - 55°$).

Sachweh *et al* [68] also investigated variations in the azimuthal scattering from spherical and non spherical particles using the DAWN-A instrument. As part of this investigation measurements were carried out at scattering angles of 40°, 50°, 90°, and 140° to identify the most appropriate scattering angle at which to record any azimuthal variation. They showed that with the DAWN-A instrument, scattering in the forward direction has the greater possibility of confusing spherical and non-spherical particles. They concluded that the probability of non-spherical particles being incorrectly identified as spherical was greater at $\theta = 40°$ than at $\theta = 55°$, indicating that, with the DAWN-A instrument, side scattering is better for shape discrimination.

Mugnai and Wiscombe [56] [69] have used extended boundary condition method (ECBM) to model scattering from non spherical Chebyshev particles. They found that the near and mid-forward scattering region is the least sensitive to non-sphericity, with the difference between spherical and non-spherical scatter intensities generally small (~ 7%). However as the authors point out

> "on an absolute scale, however, these differences are rarely negligible. A few percent difference in a large forward-scattered intensity is comparable with a 100% or more difference in a small side-scattered intensity."

Furthermore, as the relative deformation of the particle is increased and or the size parameter approaches higher values ($\alpha > 40$), larger differences in the scattering from spherical and non-spherical particles occurs.

Evidently, the combination of differences in the field of view, the number and sensitivity of the detector elements used (small number of discrete or high resolution camera systems, etc.), the types of particle under investigation and the data analysis methods applied, made it difficult to draw any definitive conclusion about the scattering angle and detector combination to be used to achieve optimal particle shape classification. **Stage 2** of this work therefore involved designing and constructing an instrument to record comparatively low resolution data[b] from both wide and low-angle scattering from individual airborne particles of both biological and non-biological origin. The results of this comparative study would

b Because of the requirement to operate in real-time, high resolution camera systems can be rejected not least because of the low frame rate of these systems but also because of the difficulties of data processing requirements needed to handle high resolution images.

inform the design of the subsequent multiparameter shape and fluorescence measurement instrument.

### 3.3.1 HiLo system overview

An existing scattering chamber based on an ellipsoidal reflector was modified by the author, as described below, to allow the recording of the scattering profile in both wide-angle and low-angle ranges from the same aerosol sample (although not the same particle – see below). A detailed description of the system, known as the HiLo system, can be found in the MPAM Final Report [70]. Only the briefest description of those features salient to the current discussion are presented here.

The HiLo instrument is shown in Figure 3.7. The electronic data acquisition assembly, including a multi-element Hybrid Photo Diode (HPD) detector, (see Section 4.1.2) could be mounted at either end of the chamber.

The detector comprised three concentric annuli: the innermost annulus was complete and could be used in the assessment of particle size whilst the middle-annulus was divided into 6 detector elements and the outer annulus into 24, so as to allow assessment of azimuthal scattering. Particle spatial light scattering profiles seen by the detector were digitised and passed to a host PC for storage. The PC displayed in real-time a proportion of the data, allowing the user to observe the nature of the scattering from the aerosol under test. The raw values from the detector were logged to disk without any data processing taking place. These were subsequently converted to a format suitable for processing in a spreadsheet package.

Of significance here is the relative scattering angles recorded by the instrument for both wide- and low-angle regimes, as given in Table 3.2. The scattering angle range used in the low-angle forward direction was designed to be close to that used in the SAFIRE instrument. This would provide the possibility of examining the full particle classification potential of low-angle scattering when confronted with a greater range of particle

| | Central beam stop | | Inner annulus (continuous) | | 2nd annulus (6 detectors) | | Outer annulus (24 detectors) | |
|---|---|---|---|---|---|---|---|---|
| | Min | Max | Min | Max | Min | Max | Min | Max |
| Low-angle Regime | 0° | 4.2° | 4.2° | 8.3° | 8.3° | 14.5° | 14.5° | 29.6° |
| Wide-angle Regime | 0° | 39.6° | 39.6° | 76° | 76° | 105.7° | 105.7° | 140.4° |

**Table 3.2 Scattering angles θ subtended by each ring of the HPD detector.**

morphologies than simply fibres; in particular biological particle shape classification could be assessed.

The wide-angle range used was largely predetermined by the configuration of the existing ellipsoidal reflector. As this would be similar to the wide-angle scattering instruments alluded to previously (as in Fig. 3.3) it would be possible to make an approximate comparison between the new HiLo instrument's performance and these earlier instruments.

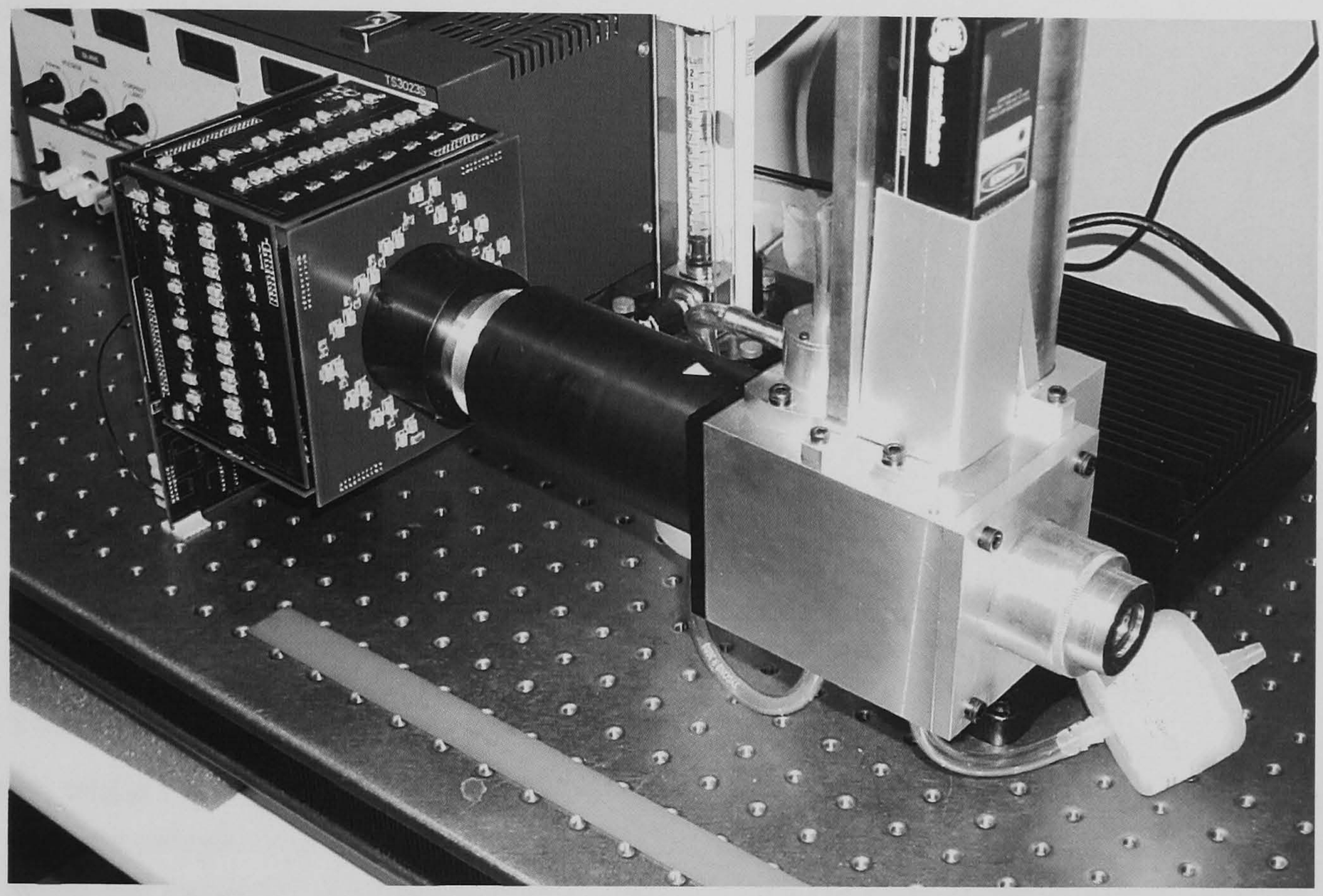


**Figure 3.7 Photograph of the HiLo spatial scattering instrument. The detector and associated electronics are shown in the low angle position. The wide angle position is on the right.**

## 3.3.2 Experimental work

Experimental work was carried out with a range of aerosol particle types including some biological materials. These varied in morphologies from spherical (water droplets, PSL), cubic (NaCl crystals), low aspect ratio ellipsoidal particles (haematite, bacterial cells) to fibres (caffeine).

Each aerosol was generated independently by the most appropriate method for the particle type. They were generated into a ballast chamber, except for the water droplets that were sprayed directly over the inlet nozzle. Data was logged for both the wide and low angle as

near to contemporaneously as repositioning of the detector allowed. Typically, data from two to five thousand particles were recorded at each position.

These results are summarised below.

### 3.3.3 Comparative results from high and low angle scattering

The HiLo results showed that, in terms of particle type discrimination, the low-angle scattering regime was at least as effective as the wide-angle regime for the particle types examined. This conclusion may be surprising given the generally held notion that shape information contained within scattering profiles is predominant at the higher scattering angles and decreases as the low forward angles are approached. However, there are three factors which have been alluded to above which help to explain this result.

The low-angle regime investigated extends out to the order of ~30°scattering angle and is therefore outside the very low angle shape insensitive forward scattering lobe which many particle sizing instruments operate within.

There is a predominance of scattering in the forward direction over backscatter, resulting in a higher signal-to-noise ratio for the same particle in the low-angle regime over the wide-angle regime. Consequently the smaller solid angle of collection is more than offset by the improved signal level. That this better S/N ratio aids inversion of scattering data to extract particle parameters has previously been demonstrated by Ulanowski [71] [72].

Lastly, while the higher scattering angles have a greater information content with respect to particle morphology and orientation, this information does not aid particle classification when the particle's orientation is only loosely constrained, as it inevitably is with instruments that use laminar air-flow delivery systems. For classification carried out at high speed by statistical methods, the additional complexity of the (wide-angle) scattering profiles gives rise to greater confusion between particle types. The lower sensitivity to particle orientation present in the low-angle scattering regime is a positive advantage when it comes to attempting to classify particles by automatic methods on the basis of their shape.

## 3.4 Summary and conclusion

Spatial light scattering analysis has been shown to be capable of classifying aerosols into distinct particle types.

Both experimental and commercial light scattering instruments have been developed for sizing and, less commonly, shape characterisation of aerosols. The instruments intended to classify by shape have generally collected scattering profiles over wide angles.

While it has generally been held that high scattering angles are required for the determination of particle shape from scattering profiles, both the review of the theory presented and the low-angle instruments described have shown that good levels of size and shape discrimination can in fact be achieved with relatively low collection angles in the forward direction. In particular, when combined with low resolution detectors such a collection regime can provide efficient high-speed particle shape classification.

This conclusion was put into practice in the design of the MPAM multiparameter aerosol classification instrument, as described in the following Chapter.

# 4 MPAM instrument design

## 4.1 Overall design considerations.

**Spatial Scattering Geometry:** The conclusions from the experimental work comparing the wide- and low-angle scattering geometries provided by the HiLo instrument and the low angle scattering geometry used in the SAFIRE fibre monitor as described in Chapter 3.2 has shown that low scattering angles in the forward direction can successfully be used to classify aerosols according to both size and shape.

In particular, when machine classification methods are used (necessary to satisfy the requirement for real time operation of the new instrument), better levels of discrimination can be achieved from low-angle scattering than from the more ambiguous wide-angle scattering profiles.

Consequently, a design based on the low-angle scattering geometry used in the HiLo chamber was developed for capturing size and shape data in MPAM.

**Particle Fluorescence Measurement:** The review of fluorescence reported in Chapter 2 indicated that the optimum bio-fluorophores are the aromatic amino acids and, in particular, tryptophan. This determined that the excitation source wavelength would ideally be within the range 220 –290 nm. Given the potential difficulties of detecting florescence from the smaller bioaerosol particles and the uncertainty as to the magnitude of any fluorescence produced it was decided to develop a design that would maximise this signal:

Firstly, it was decided to use undispersed fluorescence over a broad band of wavelengths, as the reported studies showed that fluorescence spectra did not provide significantly better inter-species discrimination. Secondly, it was also decided to use an ellipsoidal reflector to collect the fluorescence signal over a large solid angle since the fluorescence was expected to be broadly isotropic [73]. Such a collector arrangement would maximise the amount of the fluorescence that could be focused on to the detector. As a result the maximum fluorescence signal could be recorded by a single high gain detector.

Two independent recent technological developments meant that an instrument satisfying these design criteria could be realised. Firstly, a CW laser source operating at 266 nm would shortly be available. This was a prototype product from Coherent Inc (Santa Clara, CA, U.S.A.) in which continuous-wave 532 nm laser radiation from their established high-power VERDI™ laser was frequency-doubled to produce CW 266 nm radiation. This wavelength was near the maximum excitation wavelength for tryptophan and was therefore ideal. Secondly, the availability of a solid state detector with high gain in the UV (see below)

meant that the 266 nm laser line could also be used to provide the spatial scattering profiles. These two developments are described more fully in following sections.

### 4.1.1 CW 266 nm Laser from Coherent Inc.

The 266 nm laser output is achieved by the use of a non-linear BBO (Beta-Barium Borate $\beta$-$BaB_2O_4$) crystal in a resonant cavity. The Monolithic block doubler[a] (MBD), as shown in Figure 4.1, makes use of the crystal to provide frequency doubling from the narrow line-width 532 nm pump laser to the required 266 nm UV output. The monolithic block construction ensures all the optical components are sufficiently rigidly mounted to maintain the precise optical coupling required within the cavity.

The doubling efficiency of the crystal is a function of the square of the fundamental pump intensity. By arranging the non-linear BBO crystal in an active Fabry-Perot resonant cavity, intra-cavity intensity levels up to 60 times that of the pump can be achieved. Consequently, the second harmonic power level can be a few thousand times greater than the case when no cavity is used.

The butterfly configuration of the cavity has the advantage that two of the mirrors can be used to actively control the resonance within the cavity. One mirror is mounted on a fast acting piezo electric actuator, with the other mounted on a piezo electric stack. This has a slower response but a much greater travel. A saw tooth voltage is applied to this piezo which has the effect of ramping the enhancement cavity length. As the optical path length changes the intra-cavity intensity passes through a series of resonant fringes. The electronic control system is designed to lock the cavity to the peak of one of these fringes.

The enhancement achieved by the cavity meant it was possible to provide a CW doubled output from a relatively low power Nd:YVO4 CW laser producing 2 watts at 532 nm that, after passing through the MDB, produced ~200 mW CW UV power at 266 nm.

a Supplied by Microlase Optical Systems Glasgow UK.

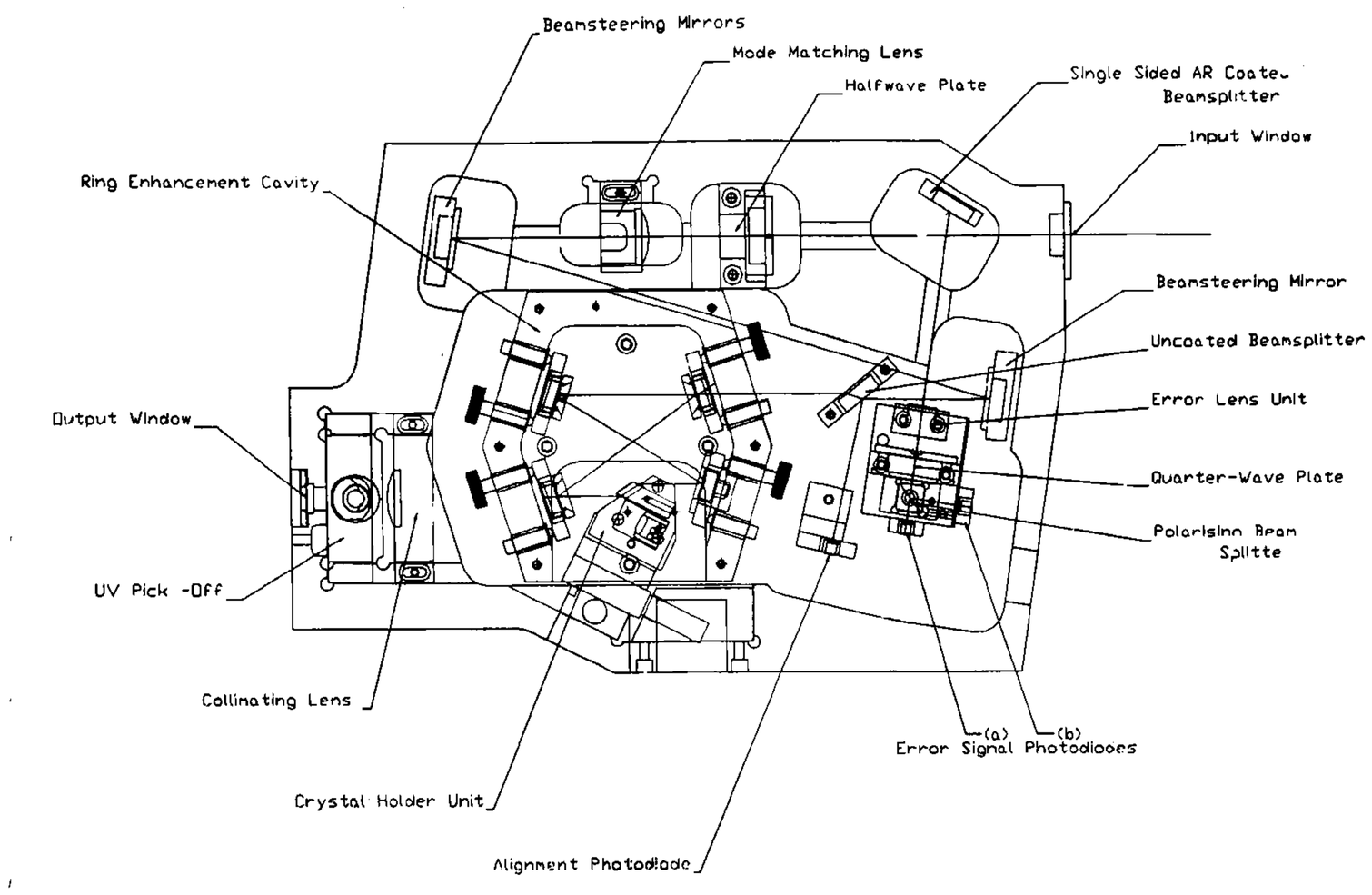



**Figure 4.1.Simplified schematic of the mono block chamber construction showing the butterfly configuration of the Fabry-Perot resonant chamber.**

## 4.1.2 Choice of detector for spatial data

To record spatial scattering from particles down to the sub-micron sizes, a multi-element detector with inherent signal gain was required. To operate in real-time the detector had to have sufficiently fast response to allow a maximum throughput of ~10,000 particles/s. While CCD camera systems had sufficient gain they could only operate at low speed as a result of their inherently low frame rate. The high resolution of camera images would also create high data processing overheads in classifying the detailed images. Even with efficient data processing algorithms and high speed processors these systems could not operate in real-time.

Multi-element photomultiplier (PMT) arrays have been developed that have both high gain and can be operated at high speeds. By selection of an appropriate photocathode material they can also have a high sensitivity in the UV. However PMT arrays have several disadvantages. They have high crosstalk between adjacent channels; the response is non-linear; there is a significant variation in the gain between the individual elements; and the pixels are arranged in a rectilinear array. While the latter limitation could be overcome by use of a optical fibre coupling to map the desired detector geometry (i.e.: a radial array to interrogate the azimuthal scattered intensity from the particle, as in the SAFIRE instrument – Section 3.2.3) to a rectilinear array, this solution would present other problems. In particular, standard fibre optics would not transmit in the UV.

Photodiode systems are able to operate at high speed and do not suffer from the problems of high crosstalk and the other disadvantages of PMTs. Whilst these devices have the further

advantage of being available as radial arrays, they have the major drawback of very low sensitivity in the UV and no inherent signal gain.

However, a new type of detector had recently become available, described as hybrid photodiode detector (HPD), manufactured by Delft Electronic Products (Roden, the Netherlands).

A schematic view of an HPD detector is shown in Figure 4.2. A photocathode and a large area semiconductor diode are separated by a small evacuated gap. Photoelectrons produced by the photocathode are accelerated across this gap by the applied high voltage. Above a device dependent threshold voltage, every 3.6eV energy increase creates one additional electron hole pair in the diode. This gives a linear signal amplitude growth with applied voltage. Unlike the operation of conventional photodiodes, the electron hole production is no longer a function of the energy of the incident radiation and hence the sensitivity of the device is essentially decoupled from the energy of the incident radiation. The gain is simply determined by the accelerating voltage applied across the device. Typically, current gain could be increased to 2200 at the maximum working voltage of 12kV. The photocathode

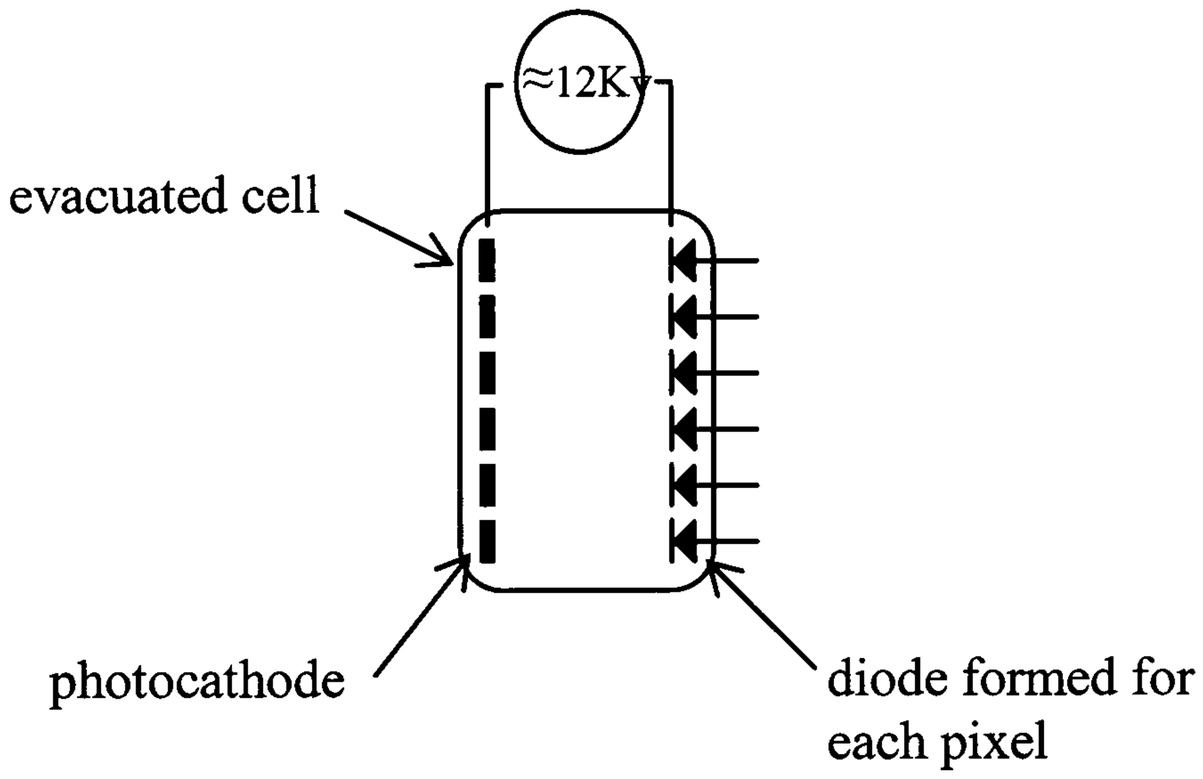



**Figure 4.2 Structure of a hybrid photo diode (HPD).**

materials can then be selected to provide the maximum sensitivity at the desired detection wavelength. A device with a photocathode having a maximum sensitivity at 270 nm (65mA/W, 25% QE) was specified as this provided the best match with the 266 nm laser line that MPAM would employ.

The short electron trajectory and the uniform linear field across the device means there is very little crosstalk between adjacent elements. The overall gain stability of the device is primarily determined by the stability of the high voltage supply. As the device draws very low current it is relatively easy to maintain a high level of gain stability with a high voltage

power supply. Photo-conversion in metal photocathodes occurs on a time scale of picoseconds and the photoelectron acceleration in a uniform field of 13kV has a negligible time spread. The speed of the device is ultimately determined by the characteristics of the semiconductor diode. The combined rise and fall time for the device was given as 25 ns. Since the scattered light pulse duration in the MPAM instrument (determined by the particle transit time through the laser beam) was expected to be greater than 1 μs, the response time of the HPD would clearly be faster than required.

The overall characteristics of the HPD device matched the specification for gain, sensitivity, linearity, and low cross talk required by MPAM. However, it was necessary to develop a custom design that would meet the additional requirement of pixels (detector elements) arranged in radial arrays, allowing direct assessment of the azimuthal variation in scattered light patterns generated by individual particles.

#### 4.1.2.1 Detector geometry

A previous investigation [74] involving the mapping of particle spatial scattering profiles captured with a high-resolution camera onto various low-resolution multi-element detector array geometries had demonstrated that simple radial arrays could provide good particle shape discrimination. The designs investigated varied from an 8-pixel radial array to 64 pixels configured in concentric rings. This simulation process demonstrated that a detector geometry consisting of typically ~30 radial pixels gave optimal levels of particle classification[b].

Based on these results, it was decided to employ a similar detector geometry. For cost reasons it was necessary to design a detector configuration which could map onto one of DEP's existing diode chip carriers. The layout of connectors on the chip carrier meant it was possible to implement a design with a greater azimuthal resolution than had previously been possible. The final design selected, shown in Figure 4.3, comprises three annular rings and a central aluminised beam stop. The innermost ring, used for particle detection, is continuous and is therefore least sensitive to particle orientation. The next two rings are divided into 6 and 24 segments respectively. The six pixel ring could be used to acquire scattering data over reasonably large solid angles, potentially offering enhanced performance in the detection of low signal levels from the smallest particles. The higher azimuthal resolution of

b The designs considered in the modelling were restricted to those that could be developed to fit the chip carrier used in the SAFIRE instrument. However the general conclusions were considered sufficiently general to be applicable to this work.

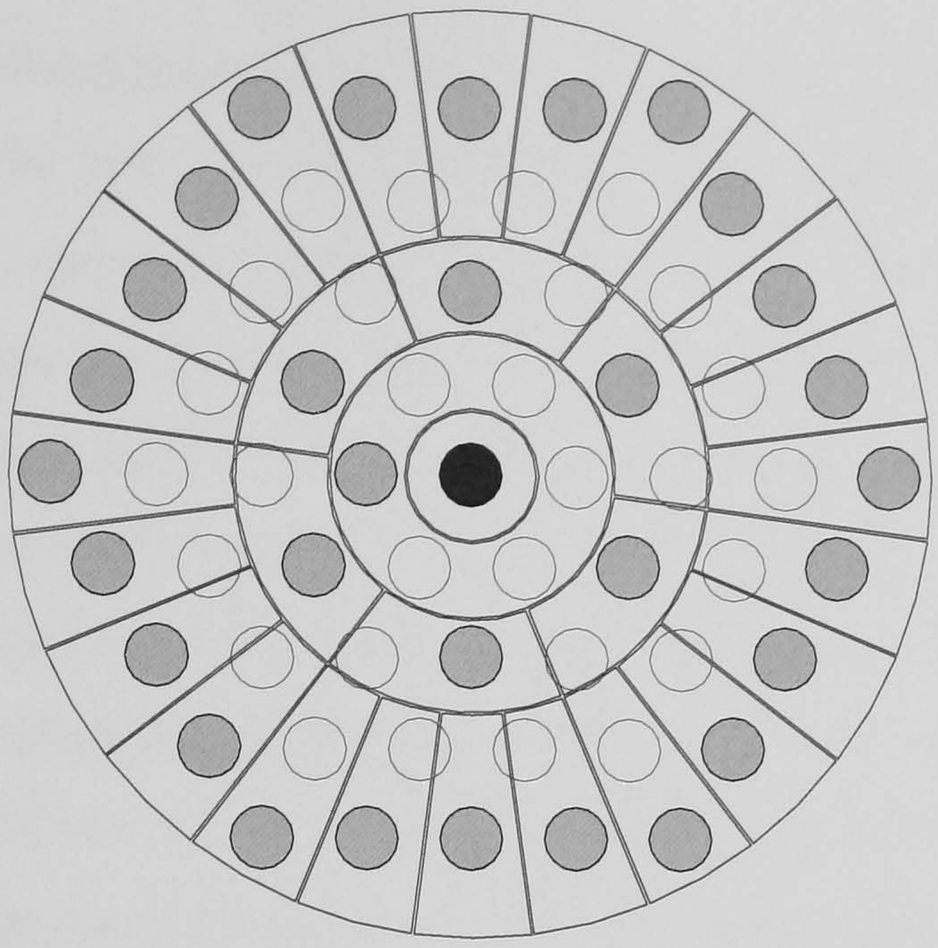


**Figure 4.3. Schematic diagram of the layout of the HPD detector. The small circles represent the solder points on the 61 pin DEP carrier used. Only the solid-grey connections have been implemented giving rise to 31 active detector pixels. The black central circle is an aluminium beam stop.**

the 24-pixel ring has the potential to record some of the finer structure in the scattering patterns.

### 4.1.3 Choice of detector for particle fluorescence signal

As undispersed fluorescence was to be recorded, only a single element detector was required. An estimation as to the likely level of fluorescence signal was carried out, intended to determine the required level of sensitivity for the detector. From the published literature it was clear that there was little agreement on the fluorescence cross-section for any of the commonly investigated bacteria. Values of peak emission cross-section ranged from $10^{-17}$ to $10^{-12}$cm$^2$/particle.nm.sr. Also for reasons discussed in section 2.2, the *in vivo* fluorescent quantum efficiency was likely to be other than that quoted for purified fluorphore in solution. Consequently, there is a wide margin of uncertainty over any estimates as to the likely fluorescent signal strength. To allow an estimate to be made, the following assumptions were made:

- the excitation intensity would be sufficiently high for all fluorophore molecules to be excited;
- the particle transit through the UV laser beam would be 3 μs;
- the ellipsoidal reflector would collect 80% of the solid angle around the particle;
- the collection optics would be loss-free;
- the fluorescence emission would be at a single wavelength of 350 nm, the peak of tryptophan emission.

A typical bioaerosol organism has a density of 1.45g/cm$^3$. The published dry mass for B.cerus, for example, is 0.67 x $10^{-12}$g including exosporium, while that for B.atrophaeus is 0.53 x $10^{-12}$g. For a spore of approximately 0.75 x 1.25 μm and assuming a prolate spheroidal cell shape, this gives a mass of 0.5 x $10^{-12}$g/spore. Tryptophan makes up 5% of the total mass of the spore [75] and has a molecular mass 204.23. The number of tryptophan molecules per spore would therefore be ~0.7 x$10^7$ molecules. If each molecule emits one photon during the period of illumination of the cell (a conservative estimate), and with an optical collection efficiency of 80%, the number of photons arriving at the PMT photocathode would be ~5 x $10^6$. A typical PMT photocathode quantum efficiency of 5% at 350 nm wavelength would result in ~ 3 x $10^5$ photoelectrons being generated. With a typical PMT gain of ~$10^5$, the anode charge would be ~ 3 x $10^{10}$ electrons, which would arrive at the anode over the 3 μs particle time-of-flight (TOF) through the laser beam giving rise to a detectable output from a standard PMT-amplifier configuration.

### 4.1.4 Selection of 266 nm blocking filter

When the particle is irradiated with UV radiation, both fluorescent emission and elastically scattered UV will be produced. Since originating from the same point in space (i.e.: the particle), these radiations will follow the same optical path. To detect the fluorescence signal alone from a particle therefore, the UV excitation radiation elastically scattered from the particle had to be prevented from reaching the fluorescence detector.

An investigation into the use of 'broadband' interference filters with pass-band in the desired fluorescence detection range showed them to be unsuitable in this application for several reasons. Firstly, they achieve only low transmission levels of around 20% in the pass band; consequently, the relatively low level signals of interest would be significantly attenuated by the filter. Secondly, the transmission bandwidth is both narrow (+/-30 nm FWHM) and symmetrical about the centre wavelength, potentially resulting in the loss of any longer wavelength component of the fluorescence.

Thirdly, typical out of band transmission for interference filters at 266 nm was 0.1% of the maximum transmission. Since the magnitude of the elastically scattered UV radiation could be at least three orders of magnitude greater than that of the fluorescence, the magnitude of the elastically scattered 266 nm signal passed by the filter could be greater than the fluorescence signal itself, making exact assessment of the fluorescence magnitude impossible. Fourthly, another problem with this type of filter was the shift in centre transmission wavelength with the angle of incidence. The use of an ellipsoidal reflector to collect the fluorescent light meant that the angle of incidence of the light onto the filter surface could lie between the near normal and typically incident angles up to 55°. The

effects of off-normal angles of incidence were to decrease in the centre wavelength, decrease transmittance, and increase the transmitted bandwidth. Consequently for off-normal radiation at higher angles there would be an even greater loss of the signal.

### 4.1.5 Coloured Glass filters

As an alternative to interference filters, coloured glass transmission filters were investigated. Coloured glass filters attenuate by both ionic absorption and colloidal scattering. The effect of transmission differences with angle of incidence on this class of filters arises from the increased path-length through the filter as the angle of incidence moves away from the normal. Furthermore, losses by surface reflection also varies with angle of incidence. With this type of filter, these effects result in a small shift in the cut-off point towards longer wavelengths. However, the filters are available as long wave pass with transmission levels above 90% compared with 20% for interference filters. The filters have the disadvantage of exhibiting fluorescence, probably from the glass substrate at high levels of UV irradiance.

A low fluorescence long pass filter was available from one manufacturer (Oriel Corporation) with filter fluorescence 7 times lower than that of a standard Schott Glass filter. The shortest cut-on wavelength of these filters was 370 and it was considered that this could result in too great a loss of the 300-400 nm wavelength fluorescence signal, sufficient to outweigh the improved performance of the filter.

In addition, the manufacturer was unable to provide further technical detail on the precise cause of the filter autofluorescence but did say that this type of filter was regularly used in fluorescence studies and should therefore be acceptable in the proposed application. The option of the manufacture of a bespoke filter was investigated with several companies. None were able to provide improved performance over their stock components at a reasonable cost or time scale.

It was therefore determined to use a standard Schott Glass WG295 filter[c] as this best met the design criteria. This filter has an attenuation of $1 \times 10^{-5}$ at 266 nm, thus providing excellent blocking of the elastically scattered UV, and a 50% cut point at 295 nm, allowing fluorescence collection at all wavelengths greater than this (up to the limit of the PMT photocathode response curve).

[c] Schott Glass Technologies WG295 x 2mm thickness.

## 4.2 Summary of overall design specification

By taking advantage of both a newly available CW laser source and high gain hybrid photodiode technology as described above, an instrument based on a single beam configuration used to both excite intrinsic fluorescence and generate light scattering profiles could be constructed. This configuration (potentially) offered several advantages over currently existing instruments:

- A simple optical design could be used that did not require multiple beams to trigger the instrument;
- the problem of jitter in the timing of pulsed UV laser systems would be completely eliminated;
- the UV beam could be operated at a low instantaneous power than common in pulsed lasers, thus avoiding the possibility of optical bleaching;
- the instrument could be operated continuously at high speed.

# 5 The MPAM instrument

Based on the design specification outlined above, the MPAM instrument was constructed over a period of approximately 15 months. A detailed description of the instrument design and operation is given in the following sections.

The MPAM system comprises three major components: the scattering chamber, the monoblock frequency doubler (MBD) / laser assembly, and the detectors and their associated electronics. These are all mounted on a single optical bench as shown in Figure 5.1 below.

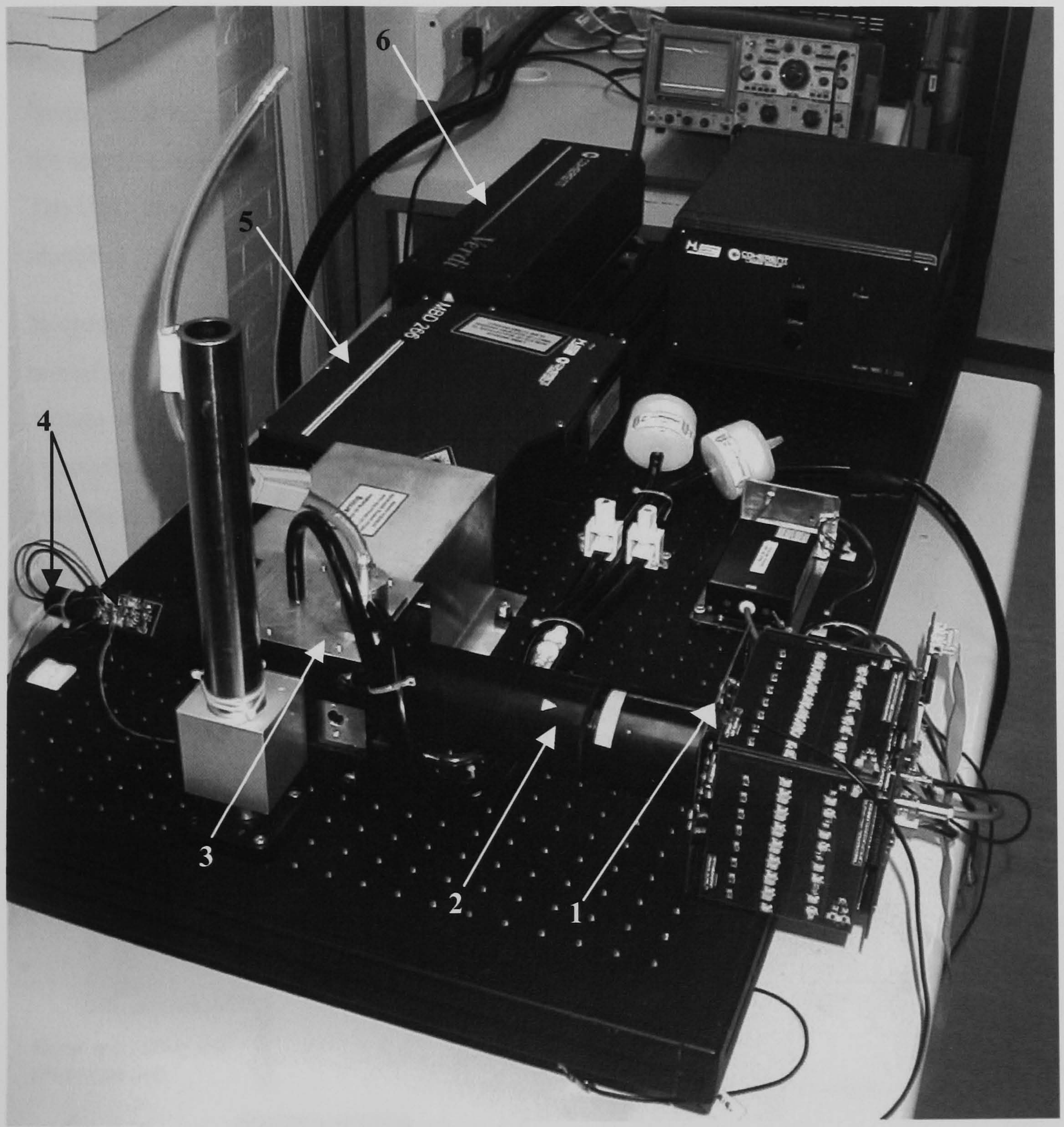



**Figure 5.1 A photograph showing the configuration of the main system components mounted on a optical bench. The MPAM chamber is in the foreground with the MBD and Verdi pump laser behind. Key: (1) HPD mounting & detector electronics. (2) Forward lens mounting tube. (3) Chamber. (4) PMT fluorescence pre amplifier. (5) MBD doubler. (6) Verdi pump laser**

Both sets of detector electronics (i.e. scattering and fluorescence) are mounted directly onto the scattering chamber. The associated electronic control systems, power supplies, pumps and cooling system are housed independently from the table. This ensures there is no mechanical modulation of the MBD cavity from sources of mechanical vibration (PC cooling fans, pumps etc). A standard computer is used to provide both monitoring and control of the system and for displaying and logging data.

## 5.1 Overview of Operation

A simplified schematic of MPAM scattering chamber is shown in Figure 5.2. The chamber is based on an ellipsoidal reflector intended to capture the maximum fluorescence photon flux and direct this onto a single PMT detector, any elastically scattered 266 nm radiation having been blocked by the long-pass filter. The spatial scattering data is acquired through the aperture in the ellipsoidal mirror, and focused onto the HPD spatial scattering detector. The HPD detector and the data acquisition electronics assembly are mounted directly on the chamber.

In operation, particle laden air from a test aerosol is drawn through a scattering chamber at a rate of approximately 1 litre per minute. A filtered flow of sheath air of ~ 4 litres per minute aerodynamically focuses the sample airstream and constrains it to pass through the circularly polarized 266 nm beam from the UV laser. The airflow also has the effect of preferentially aligning elongated particles with their long axis parallel to the flow. (As

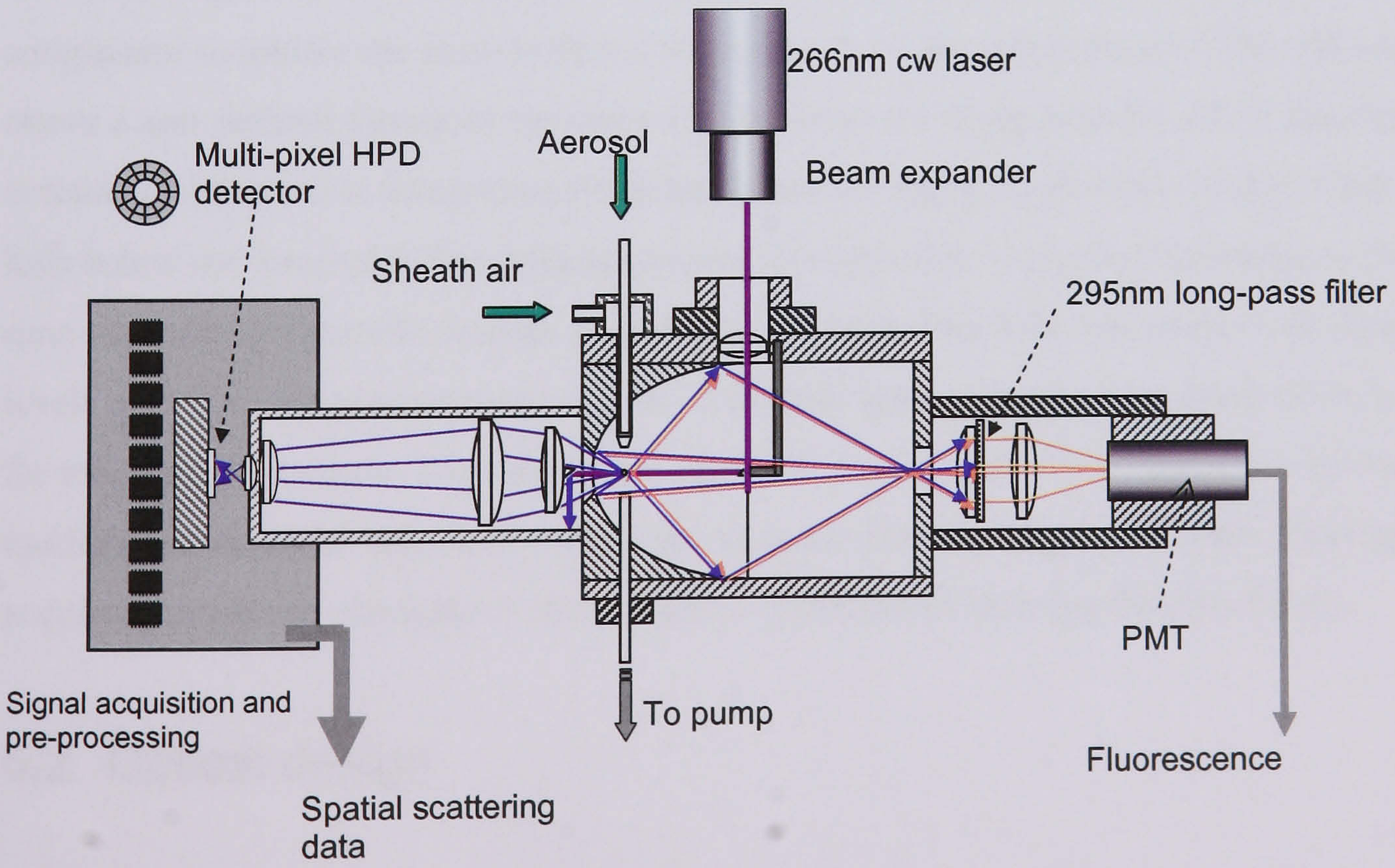



**Figure 5.2. Schematic layout of the MPAM chamber.**

discussed in Chapter 3, this is an important attribute in classifying particle shape from the variation in azimuthal scattering). At the point of intersection with the sample airflow, the beam dimensions are ~2.4 mm width by ~100 μm depth. The intersection of the beam with the sample airflow column defines the *scattering volume* as a horizontal disc approximately 1 mm diameter and 100 μm in depth. This volume is sufficiently small that for particle number concentrations of up to ~$10^5$/l, particle coincidences in the beam are rare. Light scattered both elastically and inelastically by the particle between angles of 30° to 140° to the beam axis is reflected from the ellipsoidal reflector and onto the photomultiplier detector via the blocking filter. The filter thus allows only the particle fluorescence signal to reach the PMT.

Light scattered by the particle in the forward direction (between angles of 4° and 28°) is imaged onto the hybrid photodiode detector, HPD, giving spatial scattering information that ultimately is used to both achieve particle shape classification and provide an estimate of particle size.

The data acquisition electronics comprise 32 processing channels, 31 of which process data from the HPD whilst the 32nd processes data from the fluorescence PMT detector at the opposite end of the chamber. Each channel comprises a transimpedance amplifier, dc-restoration circuitry (which removes any dc levels from the signals caused by background chamber illumination, etc.) followed by signal integration and digitization.

The inner ring of the HPD acts also as a particle-detect trigger, its output being passed to a comparator to initiate the particle detect sequence. Once the output level of this channel is above a user defined threshold the control logic forces the integrators for all 32 channels to commence integration. Integration terminates once the signal level on the trigger channel falls below the threshold. The integration period is typically ~3 μs, corresponding to the time-of-flight of a particle through the beam. The integrators hold their respective signal levels until they are passed to one of four 12-bit analogue-to-digital converters (ADC), one for every eight channels. The subsequent digitized scattering and fluorescence values are multiplexed via serial data lines to the host computer for buffering and storage. Data may be acquired for particle throughput rates of up to ~5,600 particles/s (see Section 5.6.6).

## 5.2 Optical design

### 5.2.1 Input Optics

The 266 nm output of the MBD is passed through a beam expander giving a beam diameter of 2.4 mm (between the $1/e^2$ points). The beam then passes through a cylindrical lens before

being turned through 90° by a front-silvered 45° mirror held on the optical axis by a mechanical assembly, allowing for small adjustments to the beam alignment to be made. This configures the beam at the scattering volume to be the required 100 μm depth and sufficiently wide that the intensity variation across the 1 mm wide scattering volume is less than ~ +/- 10% (necessary to ensure particles experience similar irradiance regardless of their lateral trajectory through the scattering volume).

### 5.2.2 Scattering Optics

The range of angles seen by the HPD is determined essentially by the geometry of the ellipsoidal reflector used and the size of the scattering volume. The ellipsoidal reflector has a 12.7 mm aperture located 12.1 mm from its primary focal point. An output image diameter of 18.0 mm is required to match the input window of the HPD. The scattering particle's position is constrained by the inlet nozzle design to lie within the limits of the sample flow diameter of 1mm, thus limiting the maximum scatter angle to 30° from the optical axis.

Particles can, of course, traverse the 1mm diameter scattering volume at any lateral point. It is essential, therefore, that the optical collection system was designed so that the particle scattering pattern as seen by the HPD was insensitive to the exact position of the particle within the scattering volume. These constraints were used to set up an optical model using the design software Zeemax®. The scattering volume was modelled by a 1mm diameter disc with rays traced from points on both the circumference and the centre of the disc at 5° intervals. This disc is centered on the intersection of the sample flow and the focus of the ellipsoidal reflector. The optical design is shown in Figure 5.3. The disc representing the sample volume is shown to the right with the fused silica window of the HPD forming the left most optical element. For each of the scattering angles shown, rays are traced from three points on the disc representing particles following trajectories at the edge and centre of the sample flow. This can be most easily seen in the enlarged view below. These limits also correspond with a particle passing ±0.5 mm from the primary focus of the ellipsoidal reflector. From this figure it can be seen that, irrespective of the particle's position, the scattering angle ray order is always preserved at the output, i.e. rays further from the axis always have a greater scattering angle than rays closer to the axis.

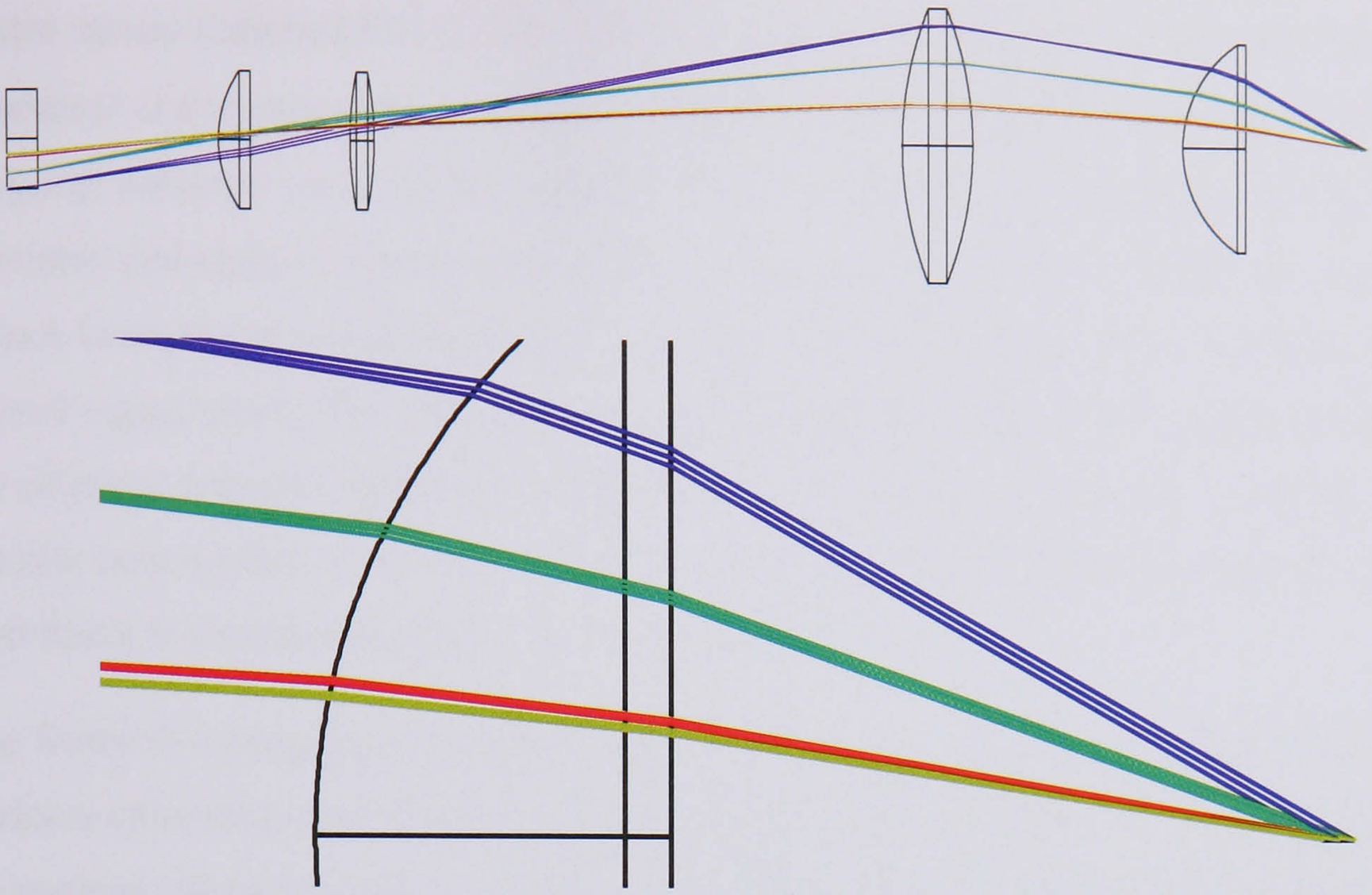


**Figure 5.3 Top: Zemax™ drawing showing the optical collection components with the scattering volume located to the right and the HPD optical window on the left. The lower detail shows rays traced from the nearest, centre and furthest limit of the scattering volume to illustrate the effect of particle position changes.**

### 5.2.3 Fluorescence collection optics

As with the scattering optics design, the scattering volume is again modelled by a 1 mm diameter disc, this time located at the primary focus of the ellipse. As only the fluorescence intensity is recorded, there is no longer the necessity to preserve ray order. The design simply has to ensure that all rays reach the PMT photocathode and consequently the design criteria could be considerably relaxed. The ellipsoidal reflector enables light to be collected over a large solid angle (collection angle of 30° to 140°) representing approximately 80% of the total sphere centred at the focus of the ellipse. The reflector used in MPAM (Melles Griot model 02-REM-001) has a rhodium coating that has a high reflectance in the visible ( >94%) but also has a relatively high reflectance at 266 nm (~60%); thus both scattered and fluorescence wavelengths are reflected towards the PMT. A UV blocking filter as described in section 4.1.5 is positioned in the optical path to prevent the 266 nm radiation reaching the PMT. Light passing through the filter is focused onto the PMT by an additional lens system.

## 5.3 MPAM Chamber Mechanical Design & Construction

Figure 5.4 shows the mechanical assembly drawing of the chamber. The chamber is mounted on the table by a single cylindrical precision column. The 266 nm beam enters the chamber via the beam input lens sub-assembly. This sub-assembly also supports the 45°

mirror mount (labelled FL11). This allows the mirror to be rotated to fine tune beam alignment and ensure that the beam is turned by precisely 90° along the optical axis of the elliptical reflector. An elliptical cross section beam shroud, 2 cm long (not shown in the assembly drawing) is glued to the mirror mount so as to surround the reflected beam and reduce background scattering from the mirror entering the chamber and contributing to overall signal noise. The reflected beam passes along the shroud and through the focus of the elliptical reflector before impinging on a 4 mm diameter black felt beam stop glued to the first lens surface. All of the non-optical surfaces within the chamber body are painted matt black to increase absorption of any stray light.

The forward-scattering lens mounting tube is arranged so that the focus of the scattering optics is coincident with the primary focus of the ellipsoidal reflector. Some adjustment of the position along the optical axis of the chamber can be made to enable both the lens system and reflector to be positioned accurately.[a]

Since any light scattered in the forward direction between 4°- 28° is imaged onto the HPD detector by the lens system, fluorescence excited in a particle that is emitted between these angles will also be imaged onto the detector. As the fluorescence will generally be several orders of magnitude below the scatter intensity it will have (effectively) no influence on the scattering profile.

The majority of the mechanical components of the instrument were machined in aluminium alloy which was then anodised black. Aluminium was chosen because of its lightness and strength. This meant the chamber could be rigidly mounted on a single mounting post, simplifying the ease with which alignment of the chamber could be carried out. The anodised finish gives the surfaces an additional hardness preventing corrosion and fusion between the sliding surfaces. An investigation of reflections from matt black paint, blank anodised aluminium and black felt (Coherent Ealing Beam Stop Black felt) showed that the black felt was in fact the best anti-reflection surface, followed by the matt black paint, though neither was suitable for surfaces subject to movement or abrasion. All three materials showed low-level fluorescence under 266 nm irradiation although it was not possible to quantify the amount.

a This is necessary as the manufacturer does not specify the relationship between the mechanical and optical dimensions of the ellipse.

| | |
|---|---|
| 1 | MAIN CHAMBER |
| 2 | ELIPTICAL REFLECTOR HOLDER |
| 3 | LOCKING RING |
| 4 | SPACER RING |
| 5 | PLATE |
| 6 | LENS HOLDER |
| 7 | LENS SPACER |
| 8 | LENS LOCKING RING |
| FL8 | LASER LENS HOLDER |
| FL9 | CYLINDRICAL LENS SEAT |
| FL10 | LENS LOCKING RING |
| FL11 | PRISM MOUNTING HEAD |
| FL11A | PRISM MOUNT |
| FL11B | MOUNTING TUBE |
| FL12 | FORWARD PLATE |
| FL13 | FORWARD LENS MOUNTING TUBE |
| FL14 | LENS HOLDER |
| FL15 | LENS HOLDER |
| FL16 | PHOTODIODE HOLDER |
| FL17 | PMT MOUNTING |
| FL18 | EXHAUST |
| FL19 | MOUNTING PLATE |
| FL20 | SAMPLE INLET ASSEMBLY |
| FL21 | HPD & MOUNTING |

| | |
|---|---|
| A | CYLINDRICAL LENS 01-LQC-006/074 |
| B | MIRRORED PRISM 01-PRB-001/028 |
| C | EALING LP FILTER A45-062 |
| D | ELLIPTICAL REFLECTOR MG 02-REM-001 |
| E | LENS COHERENT 34-6767 |
| F | LENS MG 01-LQD-009 |
| G | LENS MG 01-LQD-007 |
| H | LENS MG 01-LQP-001 |

optics design in: FWD_HPDuv_BUILT

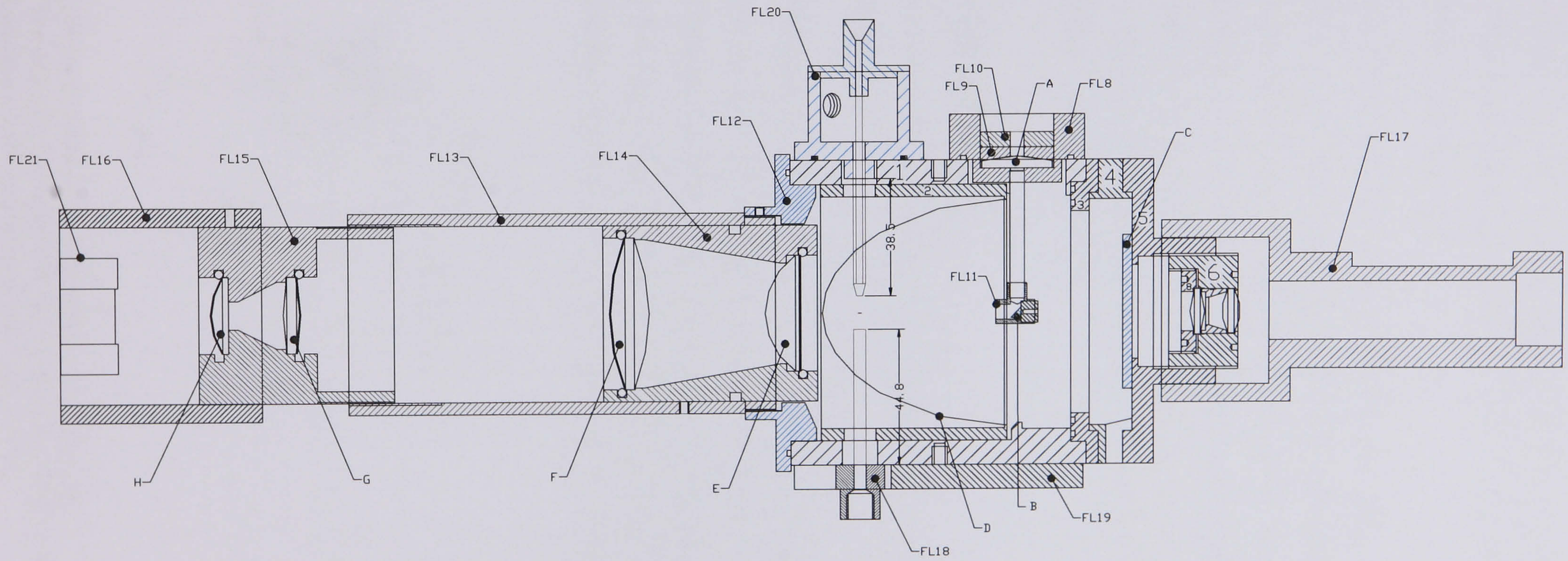



**Figure 5.4. Mechanical assembly drawing of the chamber. For clarity the mechanical mounting components have not been shown.**

### 5.3.1 MBD & Verdi Mounting

Both the MBD and Verdi™ pump laser are mounted on closed-cycle heat sinks to keep the laser base plate temperature to within 0.1°of its set point. Although the Verdi can be operated with convective air cooling, more precise temperature control is required to achieve the maximum pointing stability of its output.

This ensures the most stable optical coupling between the MBD and also helps to minimise the effects of any pointing variation of the Verdi output. Unlike the optical input coupling of the MBD, there is no beam steering adjustment on the output. Although the optical output is nominally aligned with the mechanical axis, the precise alignment of the output varies from crystal to crystal and to a lesser extent with the active spot in use on any particular crystal. As BBO crystals have a limited number of operating hours it would be necessary to replace the crystal during the lifetime of the instrument.

### 5.3.2 Particle inlet and exhaust assembly

The particle nozzle consists of two concentrically aligned tubes mounted in the scattering chamber assembly as shown in Figure 5.5.

The particle delivery system is designed to operate with a laminar flow. It consists of an inner nozzle delivering the sample aerosol, surrounded by a sheath of filtered air that helps to constrain the sample flow and prevent contamination of mechanical surfaces by aerosol particles. An additional port in the side of the scattering chamber allows a flow (~100 ml/min) of clean bleed-air to enter the chamber via a 0.1µ HEPA filter. This helps to

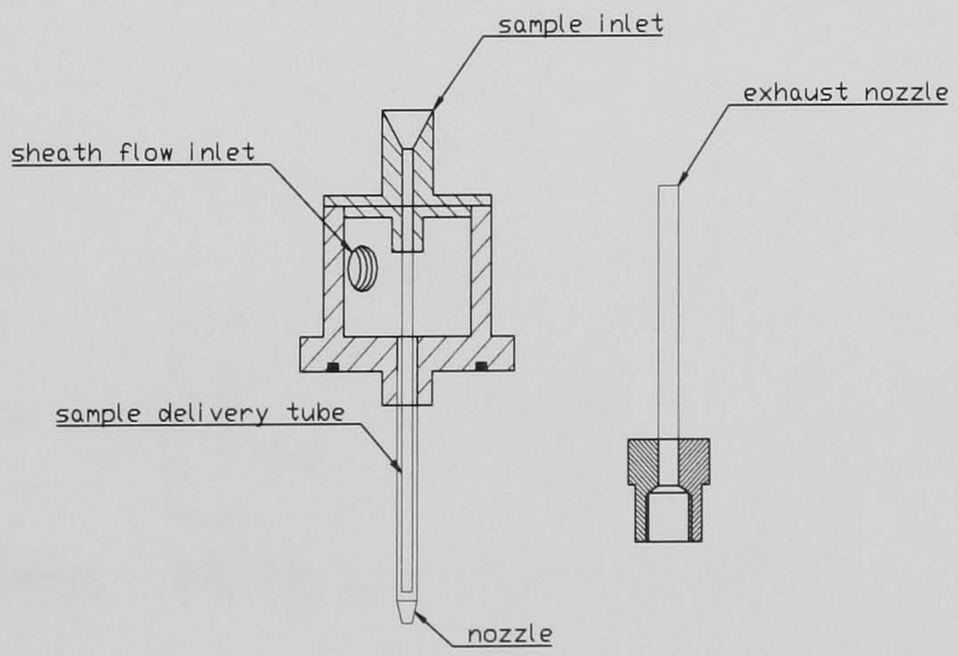



**Figure 5.5 Particle inlet and exhaust sub assemblies. Note that the sample and sheath flow are only combined at the nozzle. See Figure 5.4 for the correct juxtaposition of the sample and exhaust nozzles.**

scavenge the chamber of any particles that may escape from the sample flow. (Fugitive particles such as these tend to ingress into the chamber when the sample flow is first started, due to transient turbulence, or when it is stopped).

Slotted mounting holes allow the position of the inlet sub-assembly on the chamber to be aligned to ensure the sample airflow passes through the focus of the elliptical reflector. The length of the sample tube was initially made long enough for its position along the outer tube to be adjustable by simply sliding more or less through its mounting. The sample aerosol is vented through a simple straight sided exhaust nozzle, the position of which can also be adjusted by movement on its slotted mounting assembly.

Flow control valves are used to control the relative sample, sheath and bleed flow rates. The total flow through the instrument is controlled by a flow-meter in line with a ballast chamber and reciprocating diaphragm pump (Charles Wells and Co., UK).

## 5.4 Data acquisition electronics

The data acquisition electronics consist of 32 identical channels arranged as shown in Figure 5.6 below, (4 identical modules of 8 channels) designed to accept signals from the front end amplifiers associated with the HPD detector or the PMT output preamplifier. The front end amplifiers are configured as transimpedance amplifiers mounted as near as physically possible to the detector, keeping signal noise levels at a minimum. The output of the front end amplifiers is passed to a DC restoration circuit. This is designed to remove any slow speed background signals (time constant ~1 s) that arises from any gradual drift in the laser power level. The DC restored signal is then passed to the integrators. The output of the integrators is multiplexed to the ADC under the control of the MPAM's control logic as described below.

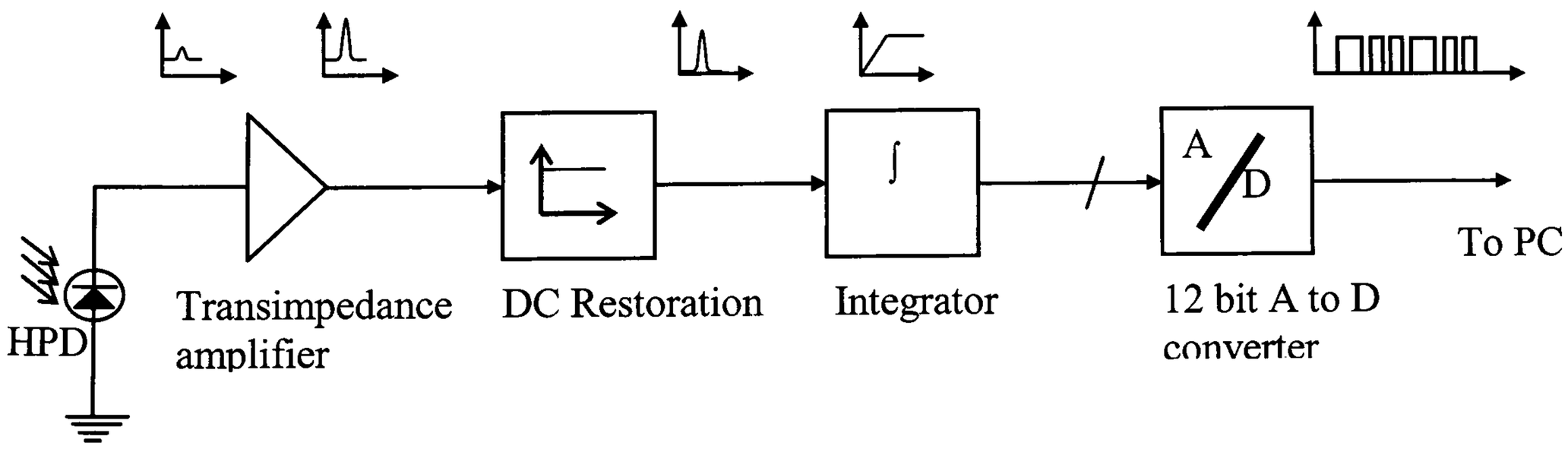



**Figure 5.6 Block diagram representation of data acquisition electronics. An indication of the signal format is shown above each of the main functional blocks.**

The data acquisition boards connect to a single back-plane containing the control logic, implemented in an FPGA (field programmable gate array) device and an RS422 interface to the host PC. Once powered up, the acquisition electronics runs continuously. Any one of the acquisition circuits can also be selected as the trigger channel.

### 5.4.1 Control Logic

The back-plane has a header connector that enables specially written control software to be downloaded directly from the development environment (installed on the host PC) during the initial testing and development stage. The final version of the control software code is written to an onboard EPROM. On reset or power up, the code is loaded from the EPROM into the FPGA. Figure 5.7 is the timing diagram for the control code. The sequence of operation is as follows:

**a.** **Trigger generation.** As a particle enters the laser beam, light scattered to the HPD detector will cause a pulse to appear at the output of the channel amplifier. When the magnitude of this pulse passes the pre-determined trigger threshold, a logic level trigger pulse appears at the output of the trigger channel comparator. This signal is passed to the FPGA **trigger** input.

**b.** This presence of the **trigger** signal causes the FPGA to generate the sample signal to the integrators (int_hld) and start the particle time-of-flight (TOF) measurement count running. (This counts in units of 1 clock cycle, 200 ns, the duration of the transit of the particle through the beam). When the trigger pulse ends, the integrators are taken out of the sample state (int_hld goes high) and the analogue integrated value is held. At the same time the time-of-flight count is stored.

**c.** Each data acquisition board has eight integrators. The outputs of all eight integrators are connected to a single on-board 12-bit analogue-to-digital converter (ADC) chip. An integrator's output is only available when selected by the chip select line. This is an 8-bit bus output labelled **int_sel**. Note that in order to satisfy the timing requirements of the integrators' internal switches all integrators are deselected for a brief period before the next is selected.

**d.** As soon as the first integrator is selected a single pulse is output on the **sync_out** line. This is used to prepare the data logging board in the PC to accept a set of data for a single particle. The ADC chip select signal **ad_cs** then occurs. This causes the ADC to carry out an analogue-to-digital conversion on the next clock (ad_clk). On the next clock pulse a synchronisation bit is output and this is followed on the subsequent 12 clock pulses by the digitised value (most significant bit, MSB, first). The use of a

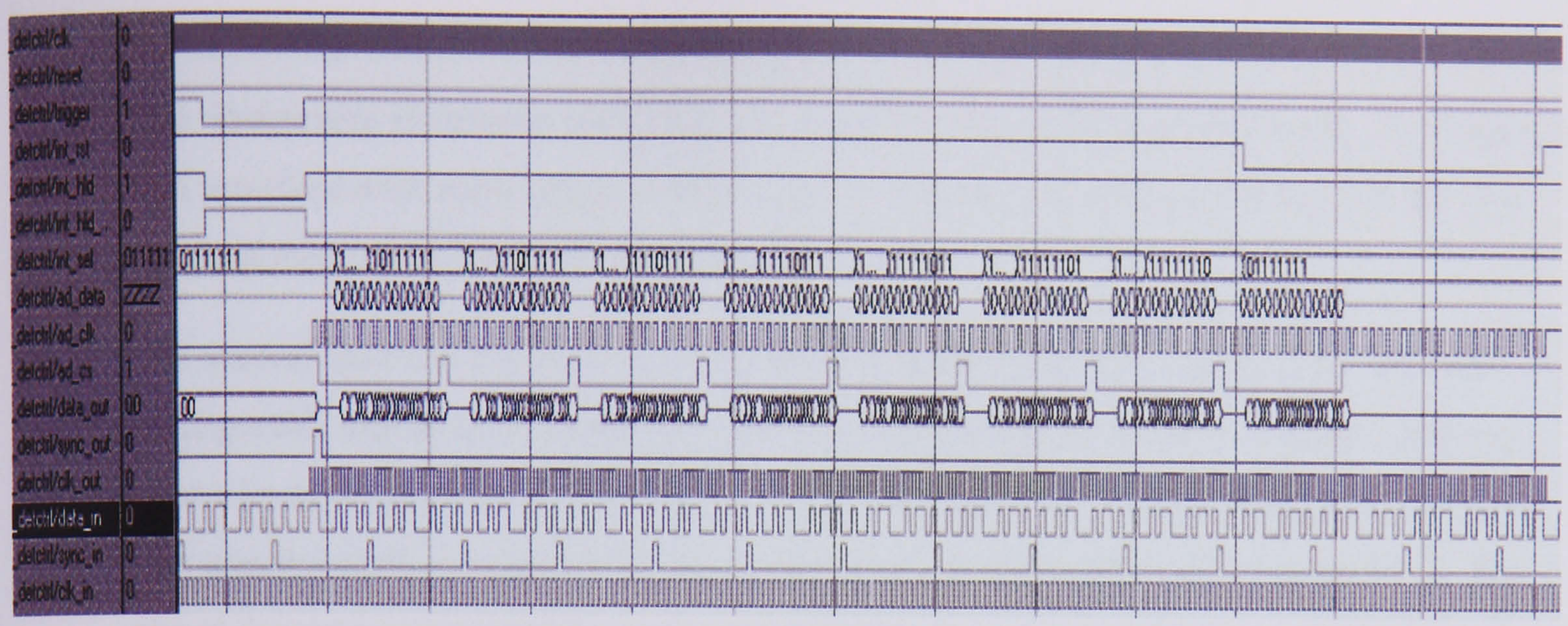



**Figure 5.7 Timing diagram from FPGA code showing the main input, output and control signals. The uppermost trace is the clock input running at ~5Mhz. At the scale shown this appears as a continuous band**

synchronous output ADC means only a single data line is required, significantly simplifying the data communication between the instrument and the PC. When all 12 data bits are output the ADC output is tri-stated.

**e.** The next integrator is now selected and a further data conversion is initiated by the following **ad_cs.** Note that it is the ADC chip select that initiates the conversion process, the clock simply controls the shifting out of the data. The stopping and starting of the **ad_clk** is simply an artefact of the particular implementation used in this application. The **ad_clk** could just have easily been free running.

**f.** This process is continued until the output of the remaining integrators has been output. The integrators are then all reset by the pulse on the **int_rst** control line and the instrument is ready to process the next trigger pulse.

All of the above has assumed that the trigger pulse length, which is determined by the time-of-flight of the particle through the laser beam, has been within an allowed window. Theoretically, all particles carried in the sample airflow will have similar Reynold's Numbers and will therefore be travelling at the same velocity (that of the airflow itself). They should therefore have very similar times-of-flight through the laser beam. However, for a variety of reasons (see Section 5.6.3), the trigger pulse may be too short or excessively long. Either case suggests an 'anomalous particle transit' of some kind, and the validity of the associated scattering and fluorescence data is in doubt. Thus, if the time-of-flight is not valid, the **ad_cs** signals are not generated and consequently no data is output to the PC, and the integrators are reset ready for the next trigger input. The limits of the allowed time-of-flight window value are set by the user and stored as parameters in the control software.

### 5.4.2 Computer Interface Hardware

The instrument interfaces with the host computer using two expansion cards. The first is an ISA interface with appropriate control logic to manage the writing of data from the two onboard memory buffers (FIFOs) to the PC.

The second card is a National Instruments PCI ADC card. This card is fully software configurable and must be initialised before it will operate. It is used to read the 266 nm power level at the output of the MBD. This is detected by a UV enhanced silicone diode that sees the surface reflection of the output window of the MBD (see Figure 4.1.). The signal from this diode is fed to a pre-amplifier in the control electronics of the MBD. Its output is taken by a differential pair analogue connection to the ADC input on the card. This analogue value is then periodically converted to a digital value when a conversion is initiated by the MPAM operating software running on the PC.

### 5.4.3 Computer interface Software

The MPAM control software written for this instrument is an MS Windows-based control program. It supports the standard windows dialogue boxes for the basic house keeping tasks (file opening and closing etc). The software carries out the following functions:

**Initialisation**. When started, the software first reads the default time-of-flight values from the MPAM header file and calculates the value to transmit to the instrument. As stated, the instrument control logic will only perform an A/D conversion and transmission of the data to the host PC if the actual time-of-flight of the particle is within the pre-set window. Initial values for the TOF window must therefore be transmitted by the MPAM software before any data will be output.

The TOF consists of a single 16-bit number: the maximum TOF held in the 9 most significant bits and the minimum TOF in the 7 least significant bits. These two values are concatenated before being transmitted to the control logic. The default minimum and maximum values are 1μs and 15μs respectively. This 16 bit number is only recalculated when the user edits the current TOF values.

A call is then made to the routine to initialise the National Instruments ADC board. When completed, the status of the board is checked, any errors being reported via a popup message box. The software then runs the UV Power Monitor function. This function initiates the ADC on the NI board and is run continually when the instrument is operating. Because the analogue signal from the MBD is noisy the average of a hundred readings is

taken. The (averaged) digitised value is converted to a milliwatt UV power figure according to a previously established calibration function.

If all initialisation is completed successfully the main MPAM program runs. This controls a bar-graph display (see below), the UV power monitoring procedures and data logging.

**Data logging**. This is activated by a drop down menu. The graphical display is stopped and a file to record the data is opened. A file header is written to the file to record the current TOF values, PMT voltage level and the measured UV power. The raw data is then logged to disk, with the particle count written to screen. The UV power level continues to be monitored while data is logged. When logging is stopped, a file footer is written containing the same information as the file header and updated power level reading.

**Data replay.** This allows a previously logged file to be viewed in the same bar graph format as described previously, making a quick visual comparison of current and previously logged data possible.

### 5.4.4 Data display

Whilst the primary function of the host PC is to log particle data for post-experiment analysis, the system has been designed to provide some real-time data display capabilities to allow the user to ensure the instrument is detecting particles satisfactorily and is recording acceptable signal levels. The main element of the display screen is a bar graph that shows on a log or linear scale (user selectable):

- The magnitudes of the scattered light signals to the 24 outer-ring HPD pixels.
- The magnitude of the Trigger signal, Tr (recorded from the HPD inner ring and used to assess particle size).
- The magnitude of the particle fluorescence signal, Fl.
- The magnitude of the particle time-of-flight (TOF) in microseconds.

Figure 5.8 shows a screen-shot of this display screen.

Also shown in the display are:

**The UV power set point.** This is a user settable value via the 'Edit' drop down menu. It is initially set to the default value of 125 mw.

**The current UV power reading.** This is the value returned by the UV power monitor function.

**Lock status.** The UV power level is allowed to vary by +/-4 mw. This is sufficient to accommodate short term variations that occur in the output power of the MBD. Within this range the MBD control electronics will ensure the cavity remains locked on a resonant peak (see Chapter 4.1.1). The range was established by power measurement readings taken at 60s intervals over 30 minutes. Any drift in the power level outside this range requires an adjustment to the pump laser to be made by the operator. If the UV power measured goes outside this range the lock status indicator changes from green to red and a periodic warning beep is sounded.

**Lock lost entirely:** Periodically, lock is lost completely causing the power level to fall to very low values. When a low power level is detected it is assumed that the MBD cannot regain lock and the lock status indicator turns red and a continuous alarm is sounded. Following operator intervention, (or if the MBD control electronics does manage to re-establish lock) the power level returns to its correct level, the alarm is cancelled and the status indicator will return to displaying green, and the audio alarm is cancelled.

**Time of flight limits:** Displays the current minimum and maximum TOF values in µs. These can be modified by the user via the edit drop down menu.

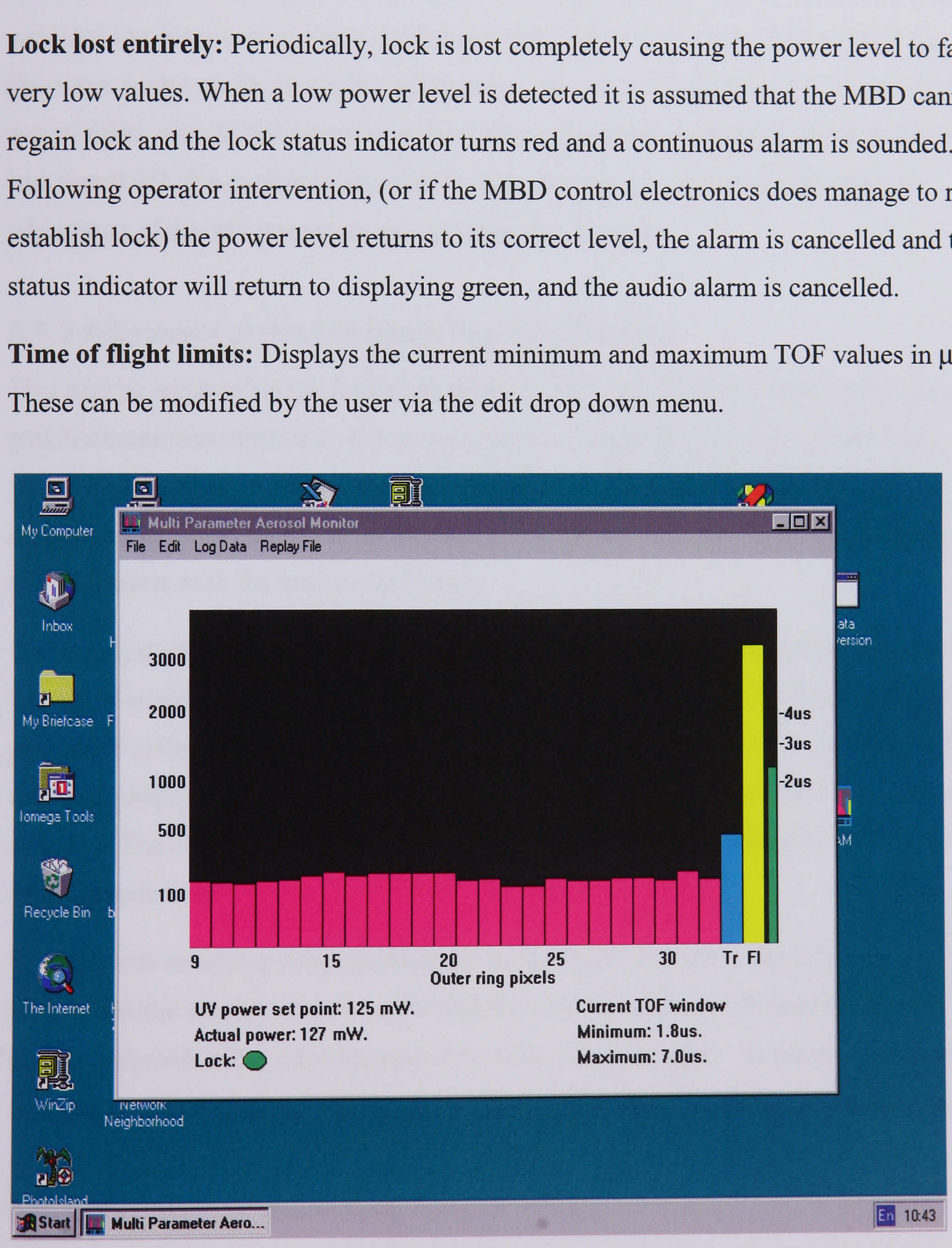



**Figure 5.8 Screen shot of MPAM display showing the bar graph data for the HPD outer-ring outputs, the trigger (Tr), the fluorescence channel (Fl), and the particle TOF. Additional operating parameters such as the UV power set point and status and current TOF window values are also displayed.**

## 5.5 MPAM Set-up and Alignment

Correct opto-mechanical alignment of the MPAM instrument is essential for its efficient operation, and this is achieved through a variety of means:

### 5.5.1 Scattering Chamber Alignment

To cater for slight differences in the alignment of the UV beam output from the MBD, the MPAM chamber is mounted so as to give one angular and two linear degrees of freedom such that the beam can be correctly aligned with respect to the chamber's optical axis. Provision is also made for angular adjustment relative to the plane of the optical table by the use of shims, should this prove to be required. Adjustments to the alignment in this plane of less than 0.05°, the minimum adjustment from the shims, can be accomplished by adjustment of the 45° mirror mounting inside the chamber.

### 5.5.2 Alignment of the UV beam into the chamber

This may be achieved in the following manner. With the chamber moved out of the beam path the beam was projected ~5.5 m onto a piece of paper and the beam centre was recorded. The beam expander was then mounted in front of the MBD output and its position adjusted until the expanded beam image was observed to be fully circular with no flaring and concentric with the unexpanded beam.

The lens systems were initially positioned according to the Zeemax optical design drawings. The tip of an optical fibre mounted on a x-y-z micrometer driven translation stage was positioned at the focus of the ellipsoidal reflector by observation of the symmetry of the reflected image, the lens housing tube having been removed to allow the fibre access to the chamber. The nozzles were then adjusted to bring the tip of the fibre into view in the centre of the sample tube.

With the lens mounting tube repositioned on the chamber, a fibre was introduced into the sample volume via the exhaust nozzle and the scattering pattern allowed to fall onto paper target positioned at the calculated position of the HPD faceplate. Its position was then adjusted so that the image had the correct output diameter. The HPD itself was then fitted.

### 5.5.3 Alignment of the optical components

Because it is not possible to observe the 266 nm beam directly, alignment must be carried out by observing fluorescence induced by the beam itself in a suitably positioned target. The general procedure used to align the UV beam was to observe the fluorescence from the

titanium dioxide whiteners used in standard copier bond, a small piece of which could be positioned in the beam path. This material fluoresces strongly even at the lowest power levels at which the MBD will maintain lock (10 – 15 mw). This method gives a clear indication of the beam's presence, but has the disadvantage that the visible image is diffuse and extends well beyond the $1/e^2$ dimensions of the beam.

The same method was used with a small length of optical fibre positioned in the beam. The fibre is mounted on a thin metal tube. This is attached to a x-y-z micrometer driven translation stage so that the fibre can be moved precisely through the beam. Because the observed fluorescence is occurring inside the fibre some of this fluorescence can itself travel along the fibre before escaping. This makes determining the precise point where the fibre is in the centre of the beam difficult. By carefully observing the fibre as it is translated through the beam several times it is possible to determine an estimate of the beam axis to approximately +/-0.1 mm.

A second method was also used that gives a more precise indication of the true beam cross-section (~ to the $1/e^2$ point) referred to as the paper oxidation method. In this method, a piece of white paper is again placed in the beam path at the point where it is desired to make the observation. The UV beam is then turned on at 75 mw-150 mw until a small plume of smoke is observed to rise from the paper. At this point the beam is quickly switched off, before the reading is invalidated by the oxidation process becoming self-supporting and extending rapidly to the edges of the paper. With the beam now off, the position of the centre of the beam can been seen under normal light as a small circular brown spot on the paper. Measurements of the size of this spot agree with the beam diameter as provided by the manufacturer.

### 5.5.4 Beam profile

The expanded beam had been observed to have a mode structure consisting of the central beam and vertically aligned lobes with nodes spaced at ~1 mm as illustrated in Figure 5.9. The precise mode structure arises from thermal stressing of the MBD crystal, and varies from crystal to crystal.

The structure of the unexpanded beam was measured with a 11 μm pinhole mounted in front of the power meter. This was done by translating the pinhole in X and Y through the beam at 0.02 mm steps. A central minimum orientated along the vertical axis is shown to be present in the central lobe of the beam, as in Figure 5.10.

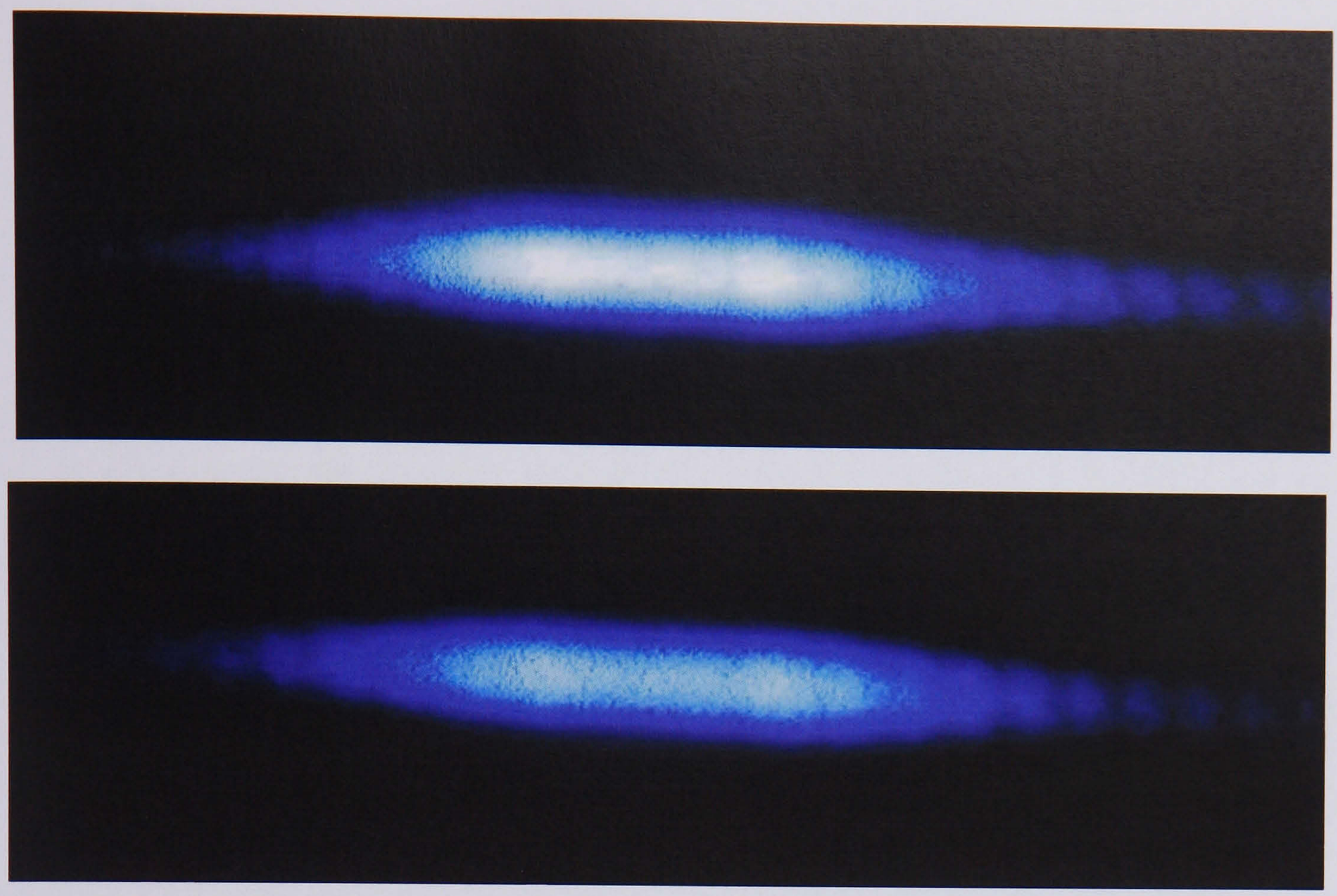


**Figure 5.9. Photos showing beam structure of output beam of MBD. The mode structure in the outer part of the beam can be clearly seen.**

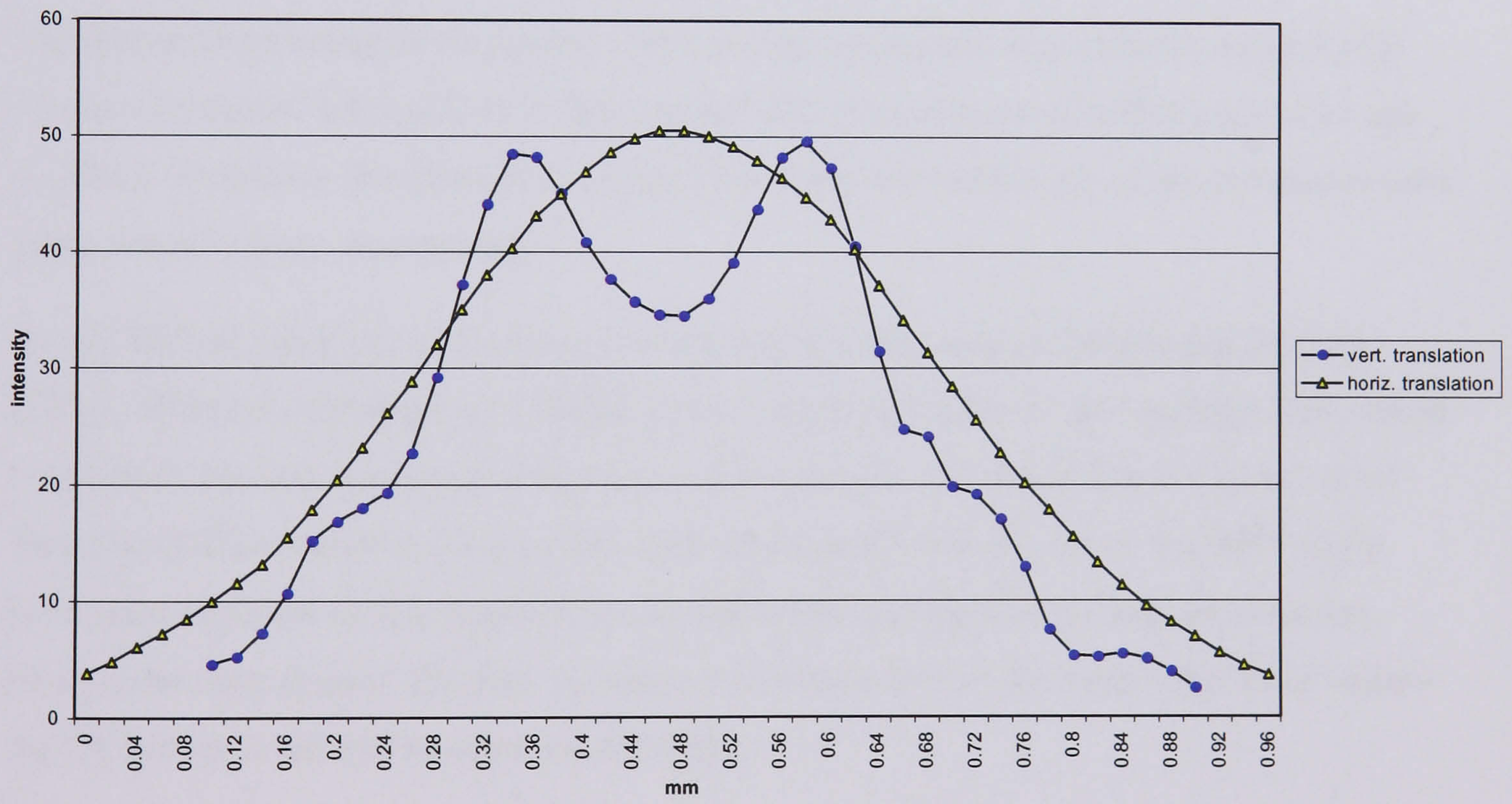



**Figure 5.10 UV beam profile recorded from unexpanded beam. The central horizontally aligned null can be seen in the mode structure of the beam.**

### 5.5.5 Sample flow set-up

The sample flow through the inlet nozzle was set up using a red diode laser beam temporarily aligned in the chamber. Smoke was introduced into the sample flow and the scattering volume was observed with a travelling microscope. Initially the sample flow was an irregular tear drop shape with flaring of the sample into the sheath air. Adjustments to the length of the sample tube and varying the flow rate had little influence on the escape of sample flow into the sheath air, although this did result in variation in the shape adopted by the sample flow. Rotating the sample flow pipe proved to have the greatest influence on the shape taken up and the degree of flaring. Removal of the nozzle and inspection showed that the outlet of the sample tube was not perfectly concentric with the sample pipe. This was adjusted by gentle bending of the sample tube. The improved alignment of the sample and sheath tube produced a near circular sample flow with no visible shedding of sample into the sheath. The scattering volume was shown to be stable over the range of flow rates from 0.8 to 2.0l/min.

## 5.6 MPAM Calibration

The proper functioning of the control software and electronics was initially verified with the use of a pulsed LED placed in front of the HPD to simulate a particle's transit through the beam. Problems identified at this stage were corrected before any of the calibration tests using aerosols were commenced.

As the MPAM made use of both a previously untried laser source (MBD) and detector (HPD), different combinations of MBD power output and detector gain settings were tested to establish the best operating combination. For example, increasing the UV power level from the MBD in order to increase the scattering signal level as seen by the HPD might have been expected to also result in an unsatisfactory increase in the background noise level in the fluorescence channel. A choice would then have to be made whether to reduce the UV power level and increase the HPD gain.

The calibration tests were designed to answer this type of issue and establish the optimum operating settings of the MPAM. The main operating parameters that needed to be established were:

- the trigger level
- the UV power
- HPD gain
- PMT gain
- the minimum and maximum TOF values.

### 5.6.1 Spatial scattering performance

The initial trials of the instrument were carried out using polystyrene latex (PSL) spheres in the size range of 0.72 μm to 3.0 μm in diameter. These were generated using a aerosol nebulizer[1] feeding into a 17.5 litre plastic ballast chamber. Figure 5.11 shows the typical data screen output resulting from a single 3 μm PSL sphere passing through the MPAM scattering volume. For a perfect sphere, scattering theory predicts that the scattering intensity falling on each of the 24 HPD outer-ring pixels should be identical. However, as seen in the Figure, the experimental scattering data exhibits a variation in the scatter intensity seen across the detector. By simply rotating the HPD on the instrument and observing that the pattern moved with this rotation, it was shown that this response variation was actually a characteristic of the HPD detector itself. When a different detector chip was used the pattern was found to be different, although the variation in sensitivity across all pixels was of a similar magnitude. Some correspondence was entered into with the HPD manufacturers, DEP, over the likely cause of this but to date no satisfactory explanation has been found.

Nevertheless, it was a straight forward task to derive a set of correction factors to normalise the response of the detector, and this has been done as part of the data analysis procedure.

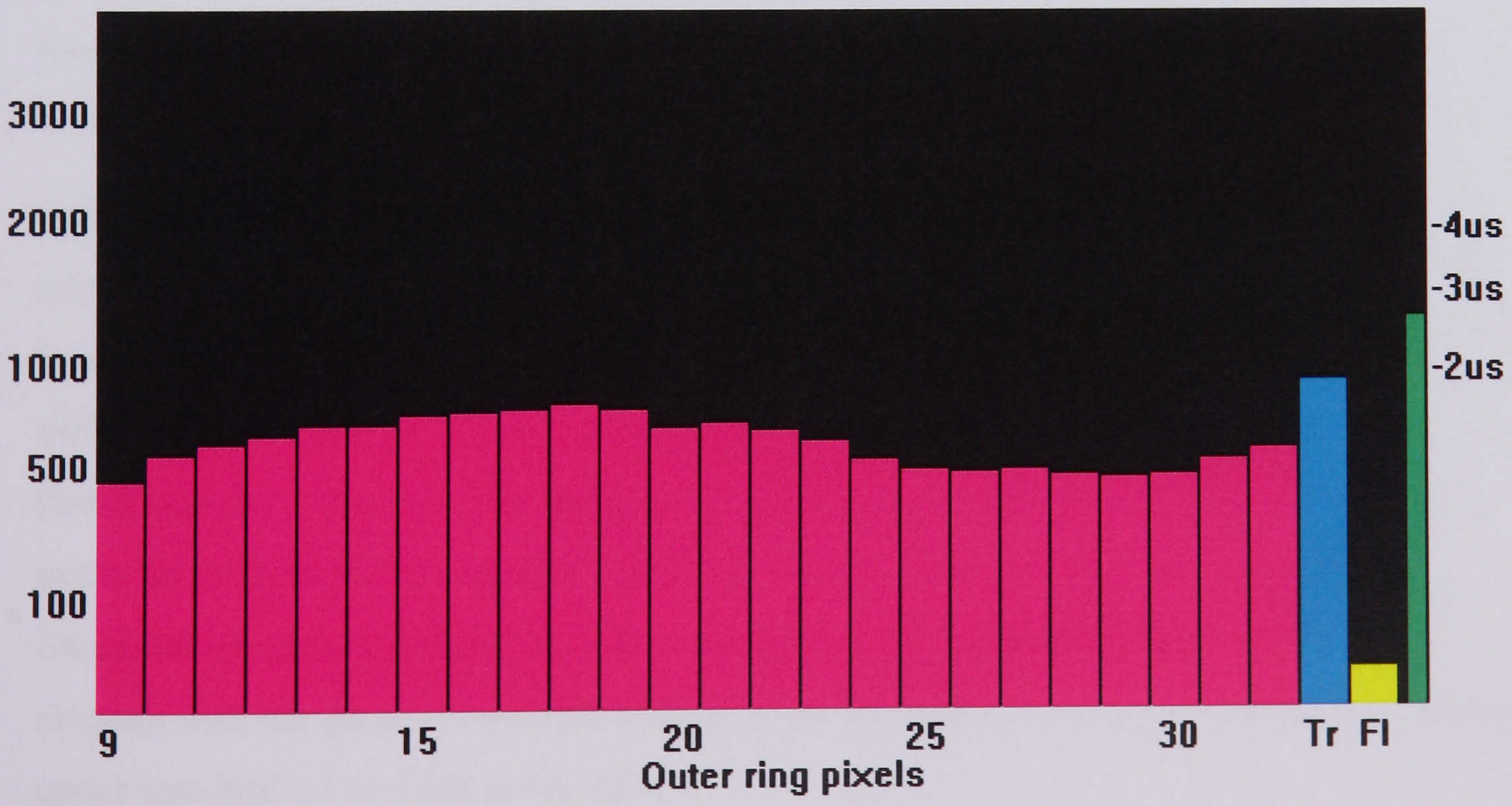



**Figure 5.11. Example of bar-graph output for a single spherical 3μm PSL sphere. Viewed from the front of the detector, pixel 9 is at the 9 o'clock position.**

1 TSI Trijet (TSI Inc. St Paul, MN, USA)

### 5.6.2 Size calibration

A size parameter in arbitrary units has been used in the experimental data analysis. This is defined as the sum of the magnitudes of the scattering recorded by the 24 outer elements of the HPD. To establish the approximate size range over which the instrument could operate and to determine the degree of finesse in size discrimination, data were recorded from a series of test aerosols of known sizes.

A vibrating orifice aerosol generator[2] (VOAG) was used to produce droplets in the size range 1 μm to 7 μm in diameter. The VOAG produces monodisperse droplets by creating mechanical instability in a fine liquid jet emanating from a piezo-electrically vibrated orifice. The liquid is usually a solution of a suitable solute - oleic acid was used in these experiments – dissolved in isopropol alcohol. The droplets created may be comparatively large, up to ~ 100 μm diameter, but reduce considerably in size as the alcohol evaporates to leave pure oleic droplets.

Using a VOAG to produce calibration particles has several advantages over the other most commonly used calibration particles, PSL. In particular, it is easy to generate any particle size by simple adjustment of the ratio of solvent to solute used. Fine adjustment in droplet size can be achieved by varying the vibration frequency of the piezo-electric crystal. No surfactant or fungicide is used in the process and consequently it is not necessary to wash and re-suspend the sample immediately prior to aerosolisation. It is also possible to aerosolise the larger sizes at high enough concentration to keep sampling times short. (With the other nebulizers available for this work, aerosolising larger particles, >~3 μm and above, was difficult).

Solutions were prepared to cover the required droplet size range; the lower size limit that could easily be detected having already been established with the initial tests using PSL. With the grade of IPA used, generating smaller particles than this with the VOAG was difficult since below ~1 μm the solvent residue become an increasingly significant percentage (by volume) of the droplet. Although droplets much larger than 7 μm could easily be produced it was quickly established that the output of the electronic signal integrator saturated for particles above ~6.5 μm. Some data were recorded for 7 μm droplets with the gain on the HPD reduced to prevent saturation, but this was taken as the upper size limit to be used in the tests.

2 TSI VOAG (TSI Inc. St Paul, MN, USA)

The droplet aerosol was generated into the ballast chamber from which a sample could be drawn from two outlet ports. One was connected to a commercial Aerodynamic Particle Sizer (TSI model APS 3320) which was used to verify that the chamber was clean between experimental runs and that the VOAG was adjusted to generate the require size. It was also used to detect the possible presence of unwanted larger sized droplets arising from the coalescence of two or more smaller droplets. The APS port was then shut off and data were logged by the MPAM for offline analysis.

The APS uses a measure of the particle time-of-flight through two closely spaced laser beams and the sample density (if known) to determine an aerodynamic size. Each of the data bins (x-axis) represents an increase in the time-of-flight recorded of 3.5 ns.

The sizing of liquid aerosol droplets by the APS is subject to significant error because the droplets deform to oblate spheroids in the accelerating airflow. Because of this deformation the cross-sectional area presented to the airflow increases and they experience a greater acceleration than would be the case with similar sized rigid spheres such as PSL. Undersizing by up to 25% has been reported in some cases depending on the liquid density and viscosity, see for example Griffiths *et al* [76] and Secker *et al* [77]. The oleic acid droplet sizes used here are given as reported by the APS and consequently the larger droplet sizes are likely to be undersized. It is estimated that for example the 7 μm oleic acid droplets are likely to be undersized by approximately 10%.

Figure 5.12 shows the data as recorded by the APS for a typical sample of oleic acid droplets. The upper plot shows all counts recorded for all 1024 bins. The median size reported is 3.07 μm with a geometric standard deviation of 1.068. The lower plot shows an enlarged view of the central section of the distribution. A small second peak can be seen to the right of the main peak with a reported size of 3.85 μm. This small population of 'doublets' is formed during sample generation or as a result of collisions within the ballast chamber. (A single droplet of 3.06 μm diameter has a volume of 1.497 $x10^{-17} m^3$ and a doublet would therefore have a diameter of 3.85 μm which corresponds with the value reported by the APS).

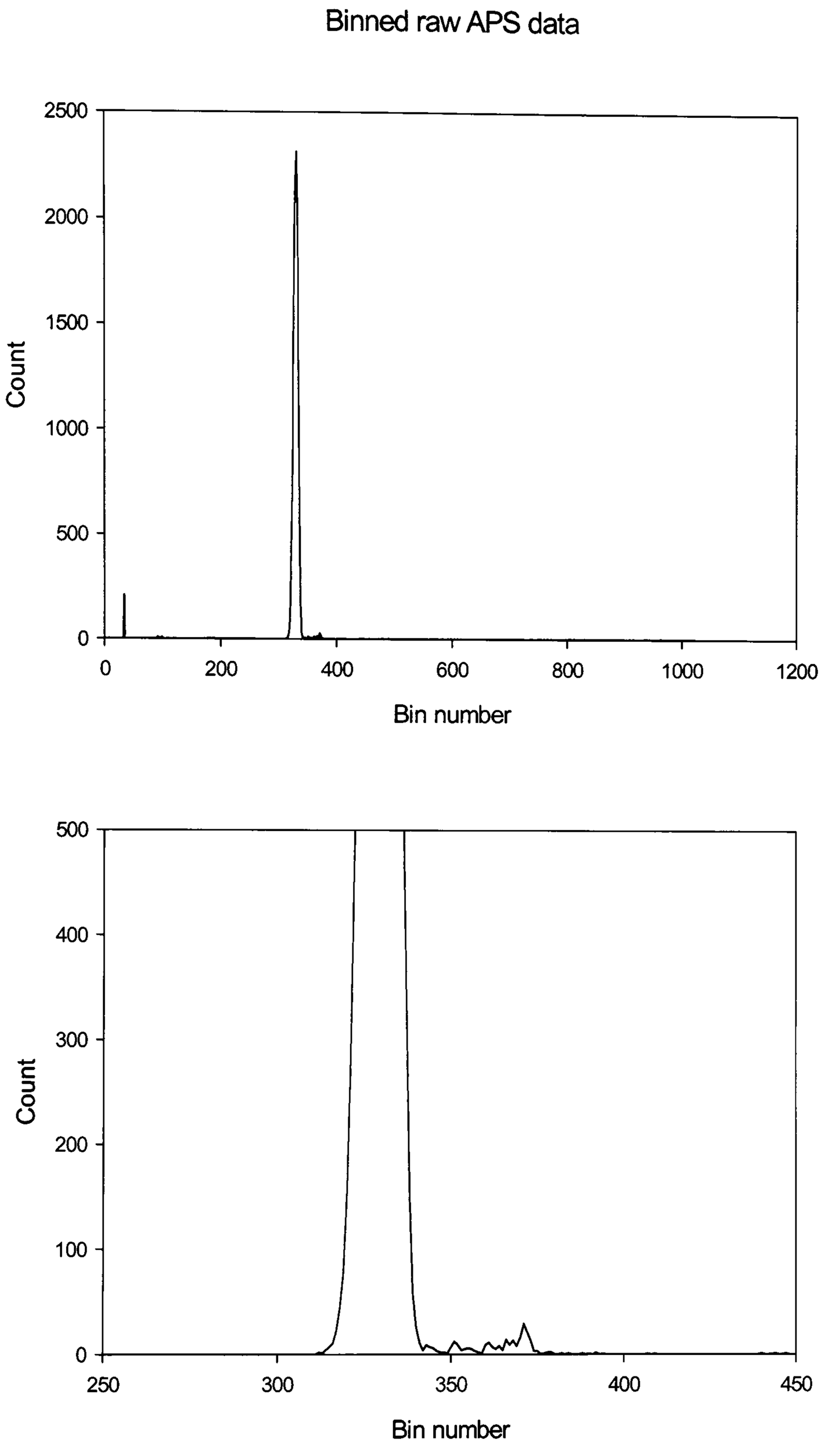



**Figure 5.12 APS data recorded for 3 μm oleic acid droplets generated with the VOAG. The upper plot shows the full range of the data. In the lower plot an enlarged view of the data shows the presence of a small number of doublets to the right of the main peak.**

In Figure 5.13, raw scattering data logged by the MPAM are plotted for the 24 outer-ring HPD pixels. As shown, the bulk of the data lie in a well defined band with a median value of 800 (arbitrary units) and of approximately equal value for all detectors as expected for scattering from a sphere. A distinct band located above the bulk of the data can also be

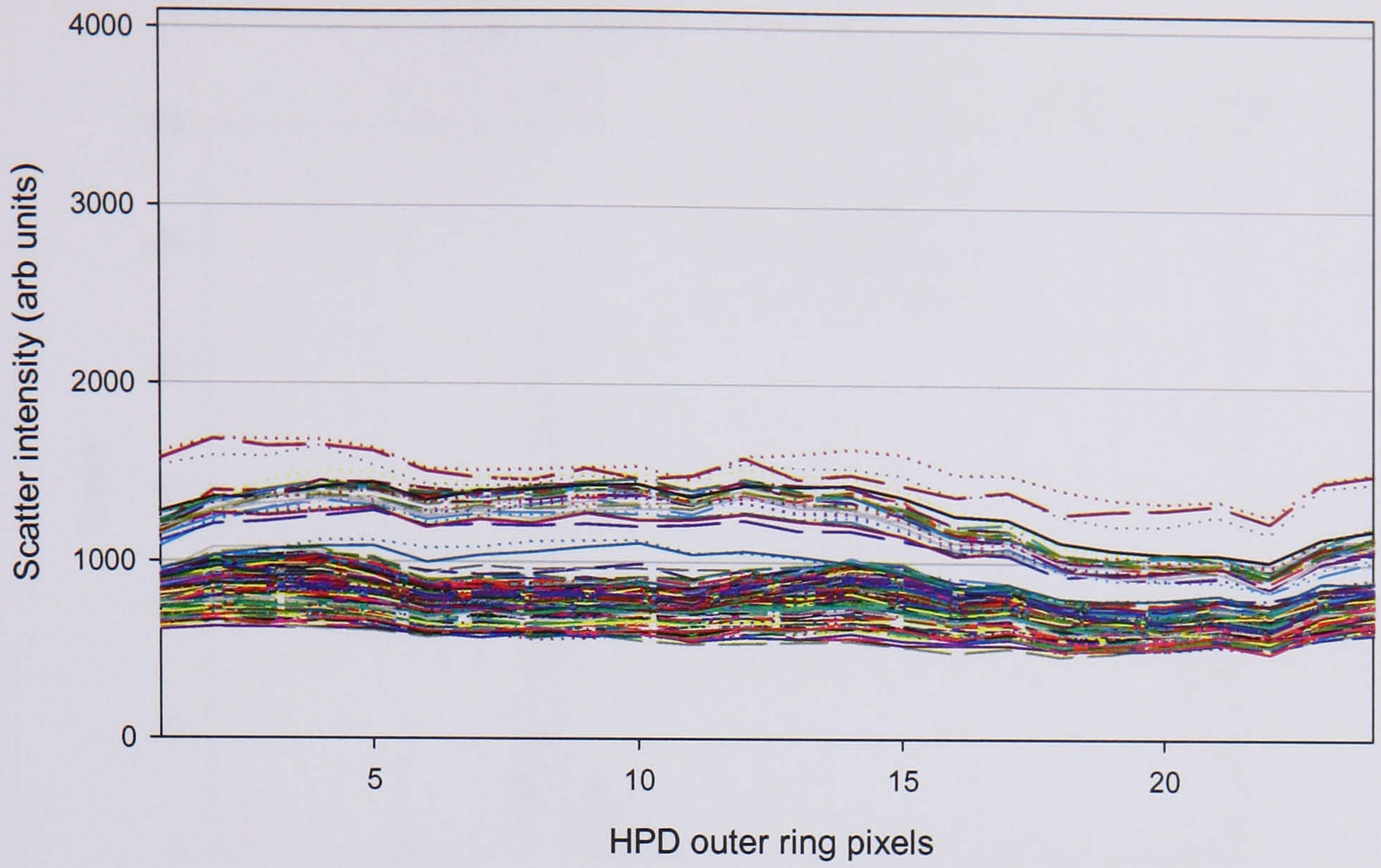



**Figure 5.13. Scatter intensity recorded by the 24 outer-ring pixels of the HPD detector from 3µm oleic acid droplets generated with the VOAG. Each line represents the scattering from a single droplet.**

seen. This arises from scattering from doublets and higher multiples as seen in the APS data. For all the oleic acid calibration data recorded, the number of doublets and triplets is greater than seen by the APS. This arises from the greater residence time of the droplets in the aerosol chamber prior to sampling, increasing the probability of droplets coalescing through collisions.

As stated above, the particle size parameter is defined as the sum of the scattering onto the 24 outer-ring pixels of the detector. For spherical particles, since the scattered intensity falling on all detector elements is, in theory, the same, the magnitude of this size parameter provides a good indicator of the relative diameter of the particle.

To aid comparison of the sample as recorded by the APS and the MPAM, data have been superimposed on the original APS plot of Figure 5.12 and re-plotted in Figure 5.14. The maximum frequency has been normalised to 100 for both sets of data to make direct comparison possible.

Sample data sets of 4,000 particles were taken from all the sizes of oleic acid droplets used in the tests and were subsequently sorted by size. These data have been plotted in Figure 5.15. Each plot shows a number of very small particles (below 1 µm in diameter) followed by a rapid rise to a fairly flat part of the curve representing the bulk of the data at each size. On the right of the curve there is again a small number of particles of a significantly larger size. These are made up of doublets, triplets and higher multiples of the droplet being generated. For the larger sizes there is a flat section where the integrators have saturated at their maximum value.

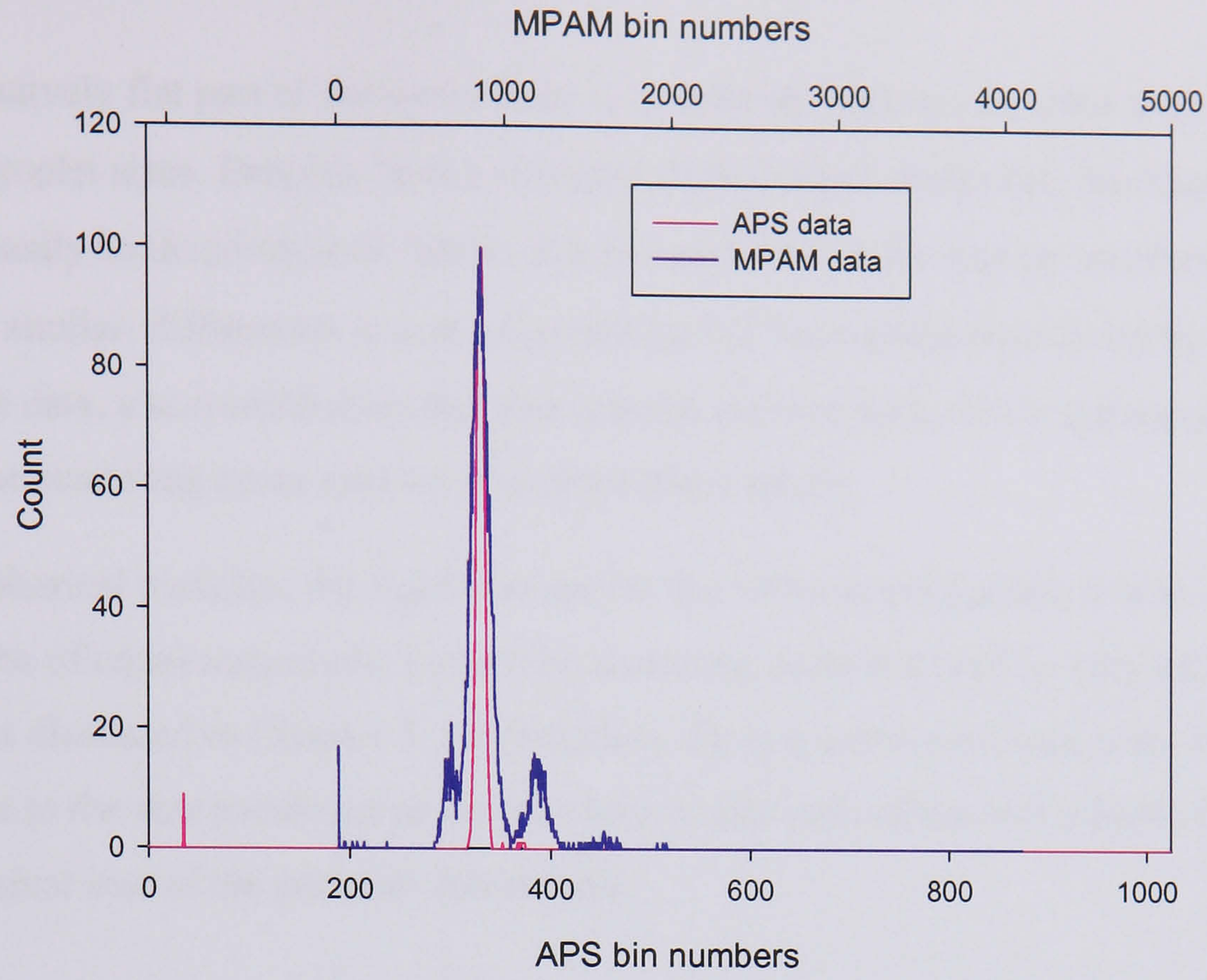



**Figure 5.14. 3 µm oleic acid data as recorded on the MPAM and the APS from the same sample plotted as a frequency histogram. For the purposes of the comparison both sets of data have been normalised to a maximum count**

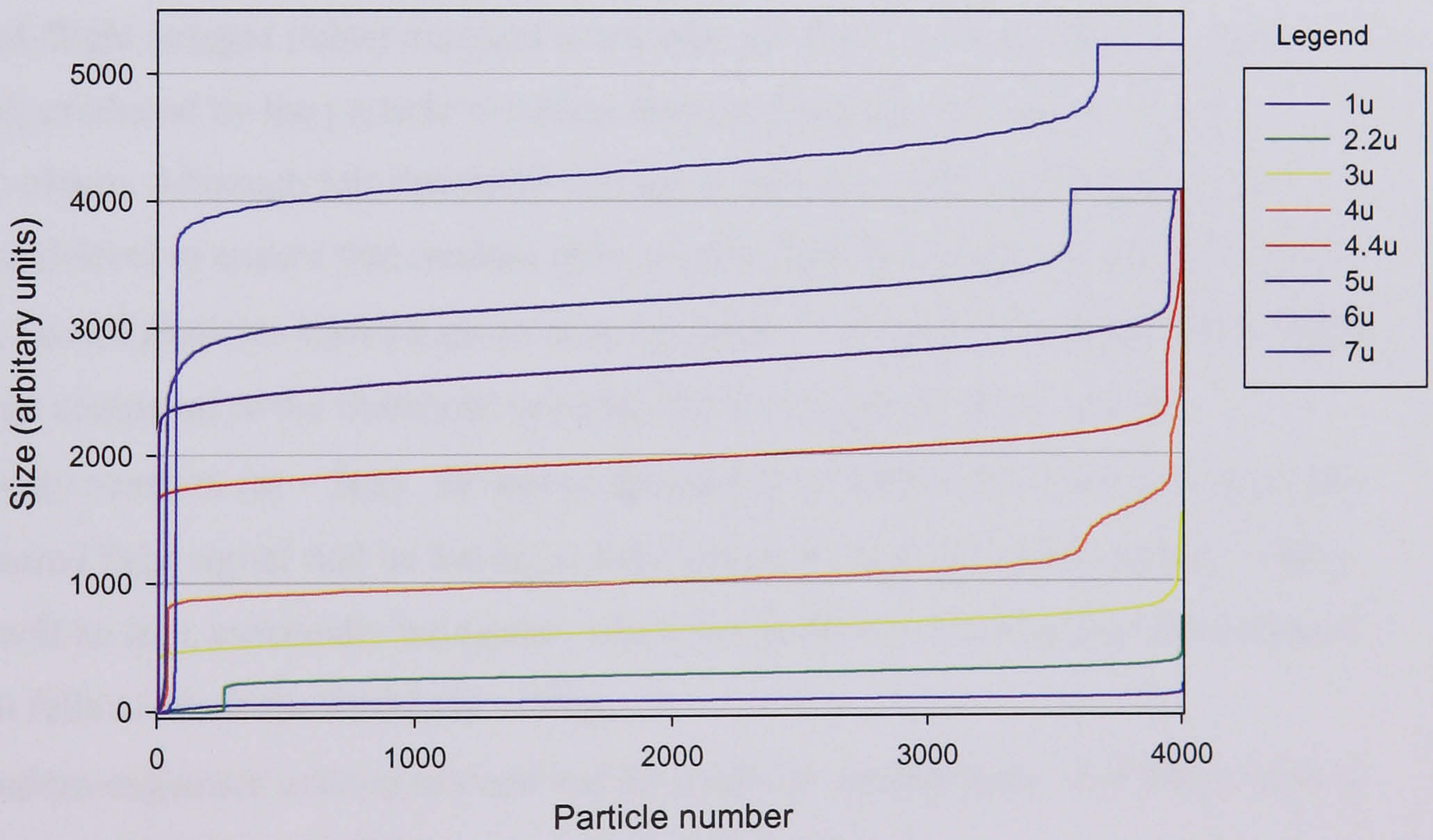



**Figure 5.15 Each line plots a data set of 4000 particles recorded from oleic acid droplets for a range of different sizes. The legend shows aerdoynamic size as reported by the APS. The data has been sorted by size prior to being plotted.**

For the relatively flat part of the curve there is no overlap between the data from the different droplet sizes. Droplets with a diameter difference of somewhat less than 1 μm can therefore easily be discriminated. Given that the gradients of the straight sections of the curves are similar, differences in size of as little as 0.25 μm could potentially be measured. From these data, a size calibration function may be derived with size being expressed in terms of the scattering cross-section of an equivalent sphere.

For non-spherical particles, the light scattered to the 24 outer-ring pixels would, of course, no longer be of equal magnitude. In fact the scattering pattern would be very much more complex as discussed in Chapter 3. Nevertheless, for any particular class of particle, the differences in the size parameter as defined here would still reflect real relative differences in the physical size of the particles themselves.

### 5.6.3 Particle time-of-flight window

As stated in Section 5.4.1, the measured particle 'time-of-flight' signal, equal to the duration of the trigger pulse that results from the detected scattered light, may vary, even though the particles are essentially all travelling at the same velocity through the laser beam. This is for a variety of reasons:

The time-of-flight (trigger pulse) duration is actually the length of time that the scattered light signal, produced by the particle's motion through the beam, is above a pre-set *threshold* voltage. Although this threshold voltage is very low, it has to be sufficiently above ground-level to ensure that random noise on the signal line does not result in false triggering. For all particles above a given size, typically ~ 0.5 μm, the scattered light signal will be large compared to the threshold voltage, and the trigger duration will be approximately constant (at ~ 3μs). However, for particles smaller than this, the magnitude of the scattered light signal will be lower, and the length of time the scatter signal is above threshold will be less, eventually being zero when the particle is so small that the scattered light signal fails to reach the threshold voltage.

Since the micro-organism cells of interest would rarely be smaller than ~ 0.5 μm in size, a **minimum** acceptable time-of-flight is set in the control software.

In a similar way, if the measured time-of-flight is too long, it could be the result of two (or more) particles closely following each other such that the second enters the beam before the first has left it. In such circumstances, the measured spatial scattering pattern and fluorescence signals will be the result of both particles and cannot therefore be interpreted. Additionally, when the pump drawing aerosol through the chamber is switched off or

interrupted for any reason, the laminar flow delivery system breaks down and some particles will diffuse into the scattering chamber. Most of these will sediment out, but some may remain airborne and once the pump is re-started may circulate in the chamber for some time before being captured by the exhaust flow. During this time, they may pass through the laser beam close to the scattering volume but with non-vertical trajectories, leading to long times-of flight. Since these particles have not passed exactly through the scattering volume at the focus of the collection optics, their scattering patterns will be distorted and should be discarded.

Finally, occurrences of the MBD momentarily losing and regaining lock may also cause relatively long optical pulses that result in false elongated times-of-flight in the electronics. In each of the above cases, setting a **maximum** acceptable trigger pulse duration prevents the acquisition of invalid particle scattering and fluorescence data.

The use of a time-of-flight minimum to maximum window prevents such occurrences of invalid data. Saturated values may still occur in the data files as a result of scattering from large particles transiting the scattering volume within the normal window. Clearly such events cannot be eliminated by a windowing function.

The default minimum and maximum values for time-of-flight are 1.8 μs and 7 μs respectively. The minimum value has been set to eliminate triggering on signals which are of sufficient amplitude to cross the trigger threshold but where the intensity scattered to the pixels of the outer ring integrates to values that remain within the background noise limits of the instrument. Although increasing the trigger threshold level would eliminate this problem it could also eliminate some valid particle data.

For those particles whose time-of-flight falls within the preset window their individual times-of-flight are logged along with the scattering and fluorescence data to be used in off-line data analysis procedures.

(The time-of-flight channel was calibrated by use of a light emitting diode (LED) positioned in front of the HPD to trigger the system with a series of pulses of known duration. The digital value output by the time-of-flight channel integrator was written to the data file. In the graphical display mode an equation derived from this calibration is used to display the TOF directly in microseconds).

### 5.6.4 PMT gain calibration

Although it is possible to use the nominal gain values supplied by the manufacturer for the photomultiplier (PMT) used for fluorescence detection, individual tube performance can

vary quite significantly from nominal values. It was therefore decided to calibrate the specific tube used in the instrument.

With the PMT and holder removed from the instrument, an LED was used to illuminate the photocathode with a constant amplitude pulse. The PMT's cathode-to-anode voltage was then varied from 400v to 900v in 50v steps. The output voltage of the PMT preamplifier stage was recorded with the oscilloscope. These values, along with an equation fitted to the data are shown in Figure 5.16. The equation of the line of regression fitted to the data is:

$$G = (2.8 \times 10^{-20})\mathbf{V}^{6.85}$$

The data has also been used to provide a look-up table to normalise the fluorescence channel data to a PMT gain factor equivalent to an applied voltage of 700V. This correction term is applied automatically as part of the data analysis procedures and is described more fully in Chapter 7.

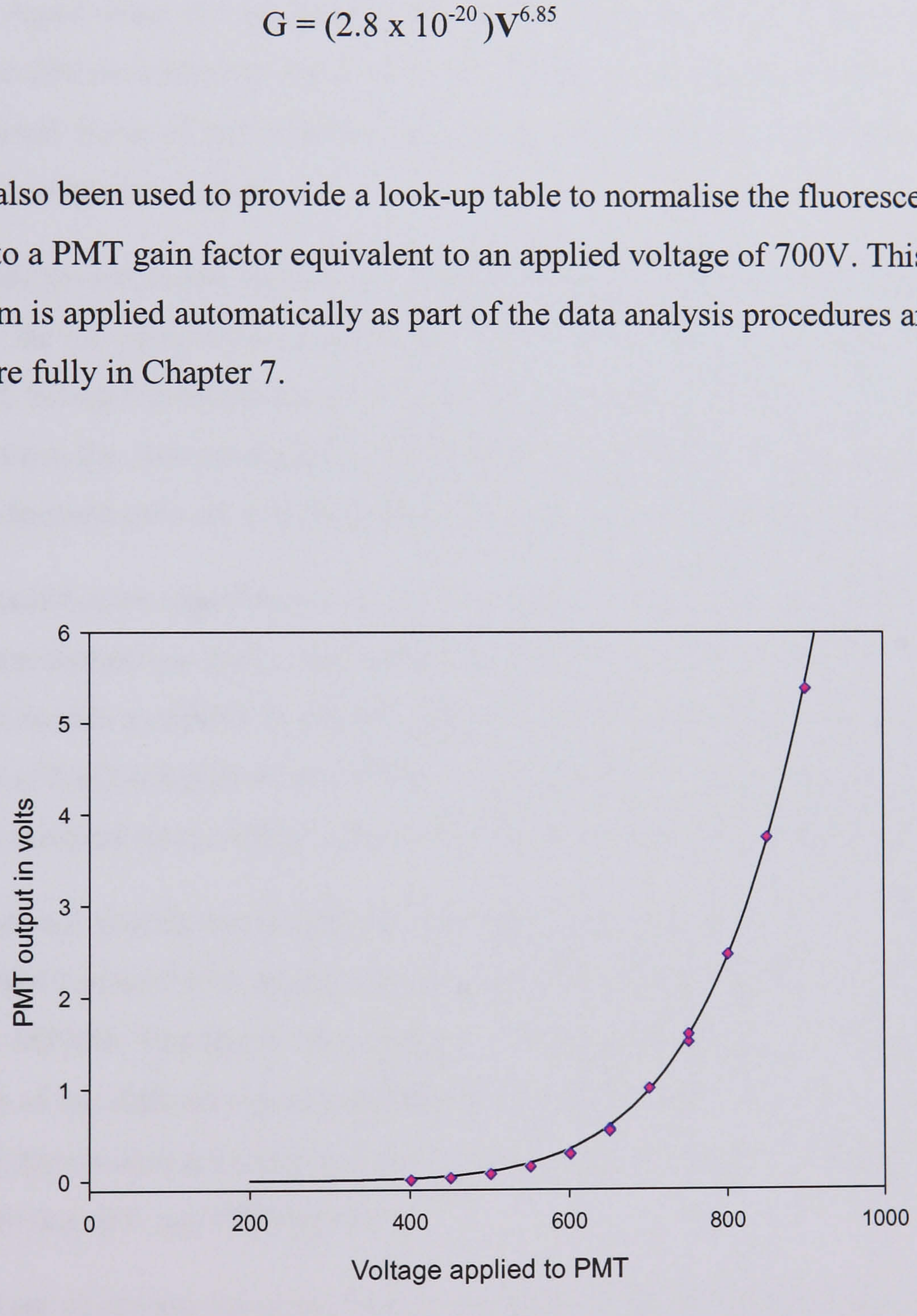



**Figure 5.16 Calibration data recorded for the PMT tube used in the MPAM. The cathode voltage was increased in 50v steps from 400 to 900v. The output was recorded with the oscilloscope.**

### 5.6.5 Testing of fluorescence channel

Because the UV blocking filter used in the instrument is far from ideal, as discussed in Section 4.1.5, tests were carried out to determine the background fluorescence levels seen by the instrument. This fluorescence background can arise in various ways.

Firstly, stray reflections from the optical surfaces result in some radiation, by a process of multiple reflections, finding a path onto the filter. Similarly, the mechanical components also give rise to both stray reflections of the 266 nm radiation and themselves exhibit some fluorescence. Again some of this radiation will find a path to the filter. While it is effective in blocking the 266 nm radiation, the filter itself exhibits some auto-fluorescence in its substrate material. Some of this radiation reaches the PMT giving rise to a 'false-fluorescence' background signal.

Whilst this background signal will be of a constant level and as such can be largely eliminated by the d.c. restoration circuit within the signal acquisition electronics, other a.c. components of background fluorescence cannot be eliminated by such measures. One such source arises from the 266 nm radiation elastically scattered by the aerosol particle that is collected and focused onto the blocking filter, thus giving rise to filter auto-fluorescence.

In order to quantify how significant a part of the signal this scattered component was likely to be, tests were carried out with a non-fluorescent material. The test aerosol chosen was NaCl which is readily available in a highly purified form. By using the appropriate concentration of NaCl solution in the VOAG it is possible to generate a range of crystal sizes. Highly purified water, which is also non-fluorescent, was used as the solute.

As with the aerosol generation method described earlier for the generation of oleic acid droplets, the NaCl aerosol was sampled with the APS for sizing purposes before data were logged on the MPAM. The scatter data sizing was referenced to the size as reported on the APS. Because of the difficulty in controlling the conditions under which the NaCl crystals form, the size distribution achieved is not as narrow as that of the liquid droplets. This is evident in both the APS and MPAM data.

An additional set of fluorescence data was recorded from a polydisperse sample of the high purity water generated by a Venturi spray. Both data sets are data are shown in Figure 5.17.

As shown in the figures, a fluorescence signal may be present up to the digitised value of around 400 irrespective of particle size. Perhaps surprisingly, there seems to be only a small correlation between magnitude of this signal and the particle's size. This was probably because the filter itself was situated in the chamber upstream of the PMT collection optical assembly (see Fig. 5.4). This assembly incorporated a small aperture

which allowed transmission of the focused ray bundle from the elliptical reflector. However, autofluorescence of the blocking filter would be essentially isotropic, so only a very small proportion of this light would pass through the aperture and onto the PMT.

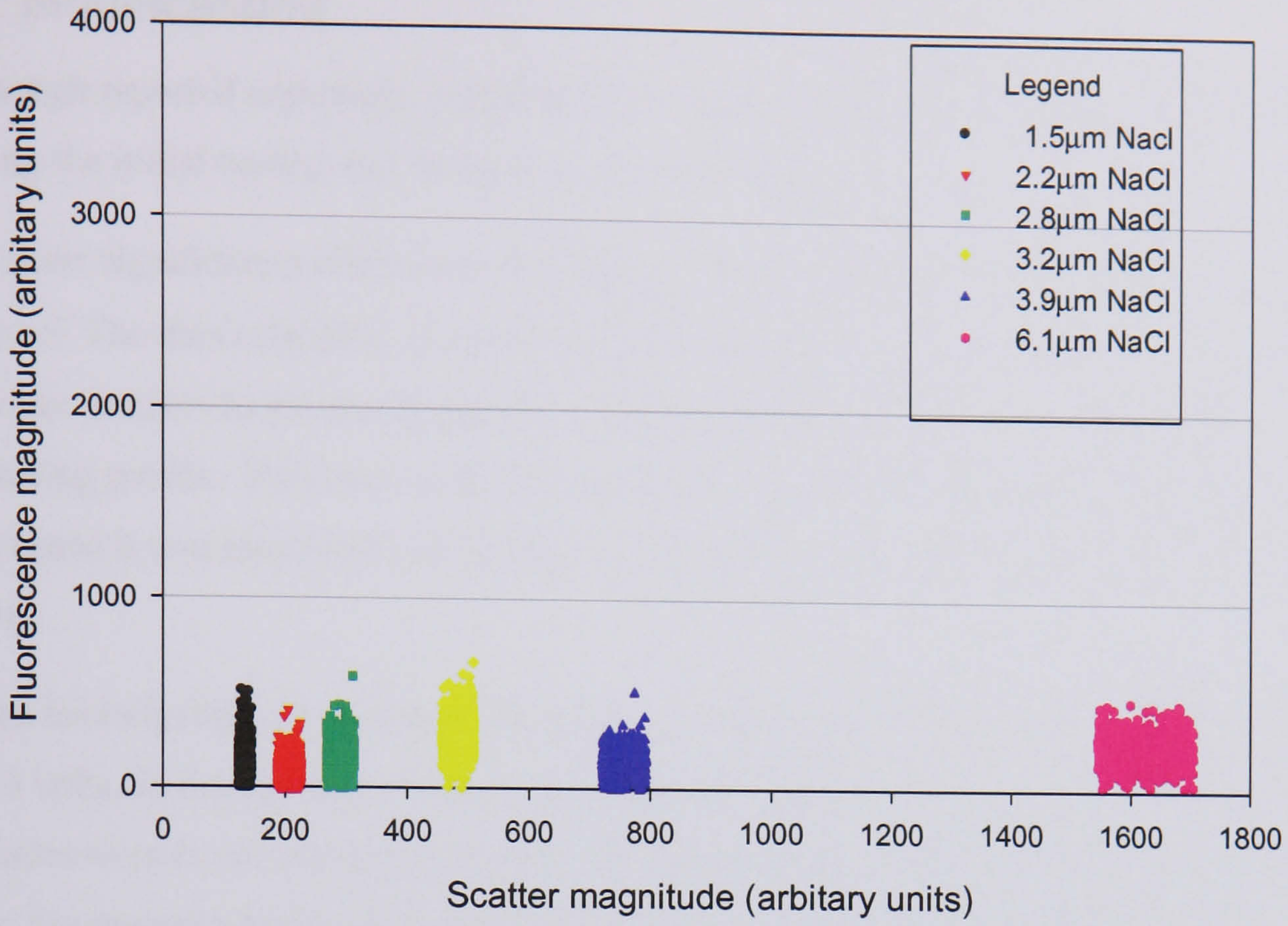
As a result of these tests, any fluorescence signal falling below a digitised value of 400 was regarded as arising from instrumentation noise and not as arising from intrinsic fluorescence from the sample aerosol. Such a correction can easily be made at the data processing stage.

### 5.6.6 Data processing speed

The data buffering on the ISA card meant that the maximum sampling rate with empty buffers was limited by the interval between the resetting of the integrators as shown in the timing diagram (Section 5.4.1). If a particle time-of-flight of 3μs is assumed this would result in a maximum throughput of ~13,000 particles/s. Measurements were carried out with the HPD detector driven by an LED run from a pulse generator. A pulse width of 3-4 μs was used with the duty cycle varied as data was logged to disk. The pulse generator duty cycle was varied until the point at which any further reduction in the pulse period caused the onboard memory-full flag to be set. With the pulse period increased slightly from this point so that the memory-full flag was only set at long intervals (~60s), this rate was then taken as the sampling rate at which data could be logged on a sustained basis. Using a 166MHz PC with no other Windows application running, a data transfer rate corresponding to **5,600 particles/s** was recorded. This would be equivalent to a particle concentration of ~$3.36x10^5$ particles/l. Given the manner in which the Windows operating system manages system resources, if other applications were run simultaneously it would not be possible to predict what sampling rates would be maintained.

autofluorescence of the blocking filter would be essentially isotropic, so only a very small proportion of this light would pass through the aperture and onto the PMT.

As a result of these tests, any fluorescence signal falling below a digitised value of 400 was regarded as arising from instrumentation noise and not as arising from intrinsic fluorescence from the sample aerosol. Such a correction can easily be made at the data processing stage.



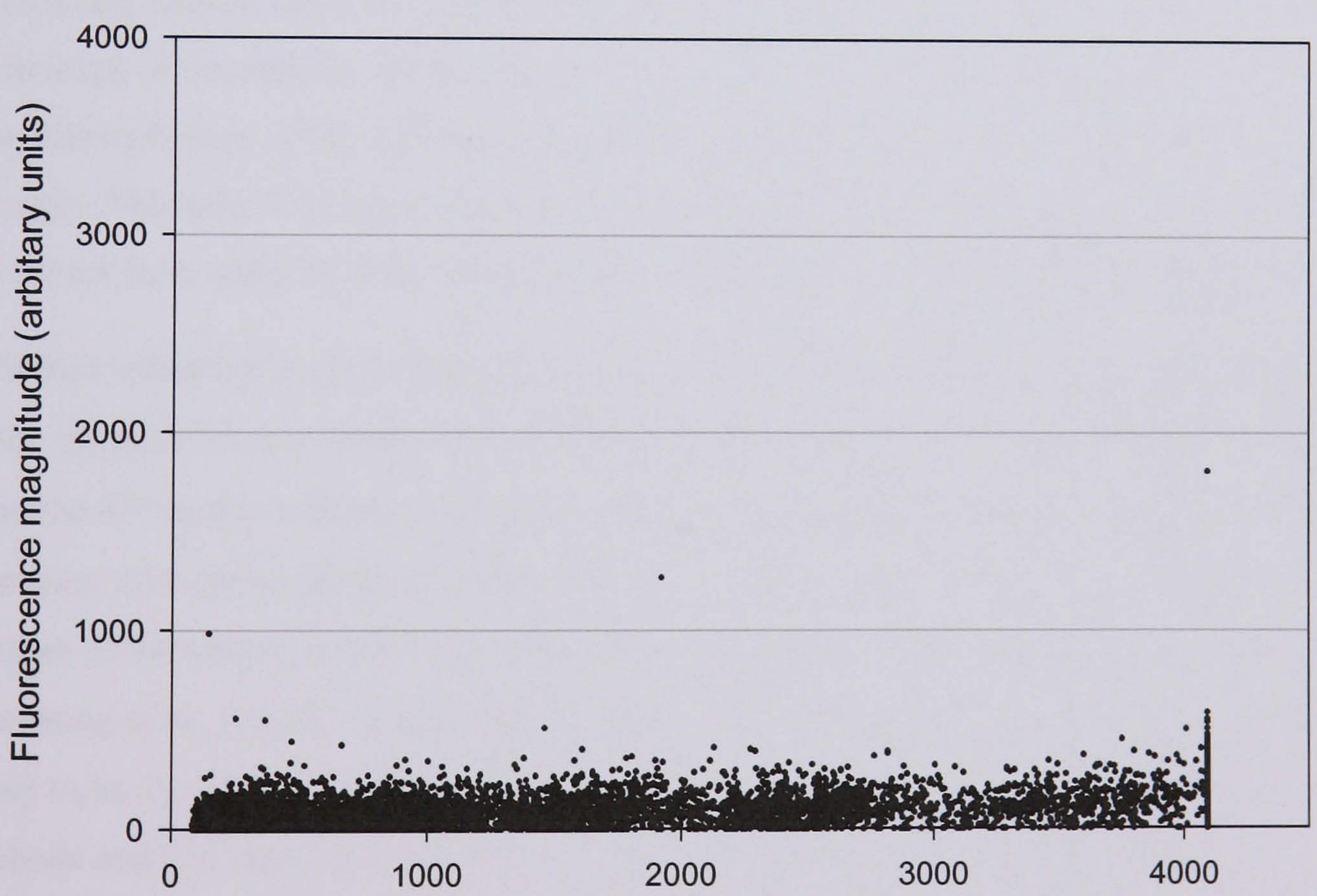




**Figure 5.17 The upper plot shows fluorescence signal plotted against size for NaCl crystals in the size range 1.5 µm to 6.1 µm.**
**The lower graph shows fluorescence plotted against size for a polydisperse distribution of purified water droplets**

## 5.7 Modifications

Although reported separately, modifications were carried out as problems became apparent during the initial testing and calibration of the MPAM.

The most significant problem was the high level of signal noise on the fluorescence channel. The data originally recorded from the instrument using test aerosols of 3 μm PSL or water droplets had a strong signal on the fluorescence channel as well as on the scattering profile. However, when a sample that exhibited much lower level fluorescence was tested it was found that the fluorescence signal was buried within the background noise.

When investigated, it was found that the fluorescence channel PMT had a large DC offset of ~3 volts. In theory, the voltage on this channel should have been zero volts when no particles were in the scattering volume. This indicated that there was a large amount of stray fluorescence being generated in the chamber. Careful visual observation of the chamber through the aperture in the rear of the ellipse or into the ellipse (with the lens tube or PMT mounting plate removed) showed that the 45° mirror used to steer the beam was fluorescing. Inspection of the mirror revealed that the reflective coating did not extend fully to the edge of the surface. In fact there was a 0.5mm uncoated rim around the perimeter. This allowed some of the UV beam to enter the mirror substrate which then fluoresced strongly. This source of stray fluorescence was reduced by painting the uncoated parts of the mirror face, along with the other unsilvered surfaces, with a black absorbing paint.

A further reduction to the background noise was achieved by modifications to the position of the elliptical iris mounted on the 45° mirror mounting head. This was originally 5 mm from the 45° mirror surface and was moved to approximately 20 mm from the mirror by an extension tube surrounding the beam. This had several benefits in terms of fluorescence channel performance. Firstly, any stray fluorescence or 266 nm source radiation that was emanating from the 45° mirror was less likely to escape into the chamber, and was more likely to be absorbed by the tube lining. Secondly, as can also be shown by ray tracing methods and has been observed in the practical construction of various instruments, moving the beam-shroud iris towards the sample volume reduced the number of paths that stray radiation could take via the ellipsoidal reflector to the detector. Of course, this improvement could not be had without cost and there was a corresponding degradation in the noise performance in the forward direction and hence the amount of noise on the HPD signal. Nevertheless, given the high intensity of the scattering relative to the fluorescence signal it was felt that the improved performance in the fluorescence detector would be worth while.

### 5.7.1 General noise reduction

The mechanical surfaces of the chamber were originally painted with a matt black paint. Measurements of the reflection properties of these surfaces showed that there was less reflection from the black felt used in the beam-stop than from the painted surfaces. In an attempt to reduce further the background optical noise in the chamber accessible surfaces were therefore covered with black felt.

Minor modifications were also made to the cylindrical lens holder to allow an adjustable aperture to be mounted between the output of the beam expander and the cylindrical lens, reducing the amount radiation from the periphery of the beam that entered the chamber.

### 5.7.2 System electronics

Under conditions when high concentrations of a monodisperse size sample were present in the aerosol test chamber, the digitised values reported periodically approached double the values typical of the bulk of the sample. This was investigated and was found to arise from the presence of a pulse arriving at the input pin of the integrators while the integrator was in the hold state. The internal switch on the integrators have a low withhold voltage. Any input over this value is therefore accumulated on the integrator's capacitor even when the integrator is in the hold state.

Because the transit of particles through the sample volume are entirely asynchronous events this situation could arise at any time. Consequently whenever a second particle passed through the sample volume while the integrators were in the hold state and prior to being reset the signal from both particles would be integrated as a single particle event. Clearly this type of event is more likely to occur when the sample is generated at a high concentration. When the cause of this problem was identified a small modification was made to the circuit boards to ensure the integrator input was held at ground potential while the integrator was in the hold state. Consequently false values would not be recorded from a second particle passing through the instrument while the integrators were in the hold state.

## 5.8 Summary

Based on the specifications outlined in Chapter 4, the prototype MPAM aerosol monitor was designed, constructed, and calibrated. Following the construction and initial trials, various modifications were carried out to correct or improve the performance of the instrument.

A range of calibration measurements using standardized test aerosols was completed to accurately characterise the performance of the instrument. The optimum HPD and PMT gain settings were established along with minimum and maximum time-of-flight values.

A set of correction factors for the gain normalisation of the HPD detector and gain normalisation lookup table for the fluorescence channel PMT were produced. Testing with non-fluorescent NaCl crystals showed that a single size-independent correction term could be applied to the fluorescence data to compensate for the background signal on the fluorescence channel.

The MPAM instrument was therefore ready for detailed experimental challenges using a variety of biological and non-biological aerosols

# 6 Experimental method and data logging

Apart from the standard spherical particles such as PSL or liquid droplets generated with instruments such as the VOAG, there is not a readily available range of standard test particles that are commonly used in the calibration and testing of aerosol instruments.

Moreover, the scattering profile recorded for any particular sample may be influenced by the circumstances under which the sample was generated. Even for the standardised fluorescent stained PSL, differences have been observed between the level of fluorescence recorded from PSLs with the same specification supplied from different manufacturers. For the biological materials, factors acting long before the sample is aerosolised influence not only the likely size and shape but also the level and spectral detail of any fluorescence recorded. In particular, the characteristics of the culture medium used has been shown to have a significant bearing on the fluorescence [21] [78]. In some cases the detail reported in the spectra of a particular species has later been shown to be entirely the result of the particular culture medium used and not, in fact, intrinsic to the sample organism under investigation.

To overcome some of these limitations in both standard test material availability and the methods of sample preparation, a brief description of each material used is presented below along with a description of the conditions under which it was used.

## 6.1 Sample preparation & aerosol generation

### 6.1.1 Non biological sample materials

Aerosol challenges were carried out at laboratories of both the University of Hertfordshire and Defence Science and Technology Laboratory, Porton Down. Each aerosol was generated independently by the most appropriate method for each particle type and, with the exception of water droplets which were sprayed directly over the MPAM inlet nozzle, were generated into a large ballast chamber.

The ballast chamber used at DSTL comprised a 300 litre clear acrylic box that has its own internal power supply, air circulation fan, light and clean air supply. The chamber had two sample ports which could be controlled independently, one of which was connected to the MPAM and the other to various monitoring instruments such as a UV-Aerodynamic Particle Sizer (TSI Inc., St. Paul MN, USA) or a DSTL B1010 aerosol shape monitor (Biral, Bristol, UK), an instrument for assessing particle size and asymmetry based on the 'triple-detector' scattering chamber design shown in Fig. 3.3. It was therefore possible to

continuously monitor the contents of the chamber while generating samples and ensure the chamber was fully cleaned between experimental runs.

For all experiments, the gain of the MPAM HPD detector was fixed at 2000 (HT = -12KV), with a sample flow-rate of 1.0 l/min. For those (non-biological) particles with a polydisperse size range, a stirred ballast chamber was used (DSTL only), to counteract the preferential loss of larger particles by sedimentation.

### 6.1.2 Spherical particles

#### 6.1.2.1 Polystyrene latex microspheres

Both standard PSLs and fluorescently-dyed PSL microspheres were used. The PSLs were available in a wide range of sizes and distributions, although the maximum size that could be reliably generated using a nebuliser was ~ 3-4 μm. To ensure there was no contribution to fluorescence from the PSL suspending medium (which frequently contained small traces of surfactant and/or fungicide), samples of PSL were washed and re-suspended in distilled de-ionized water. No measurable difference in the florescence recorded on the MPAM could be detected between these samples and their unwashed equivalents.

#### 6.1.2.2 Oleic acid droplets

These were generated using the Vibrating Orifice Aerosol Generator (VOAG) in the size range of 1 μm to 7 μm by mixing the appropriate concentration of oleic acid in IPA solvent. With this method it was simple to generate repeatable samples over a wide range of sizes. The VOAG transport air was used to carry the sample into the ballast chamber.

#### 6.1.2.3 Silica

Samples of dry 3 μm silica ($SiO_2$) sphere powder were aerosolised using compressed air. Although essentially spherical, silica spheres have a slightly faceted surface. They also have a broader size distribution than the oleic acid droplets or PSL spheres.

#### 6.1.2.4 Water droplets

This was the only material not sampled via the ballast chamber. The droplets were sprayed directly over the inlet nozzle of the MPAM with a hand-held nebuliser. The resulting size distribution was highly variable and included droplets up to 10 μm and above. This aerosol was mainly used for checking the operation of the MPAM during the instrument's alignment or when carrying out adjustments.

### 6.1.3 Non spherical particles

#### 6.1.3.1 Cubic and ellipsoidal particles

Low aspect ratio haematite particles ($Fe_2O_3$) were generated from suspensions in water or by dry air jet. These particles were available in aspect ratios of near cubic, 2:1, and 3:1. Although their production process did not give rise to monodisperse sizes, as can be seen in the scanning electron micrographs (SEMs) of Figure 6.1, the bulk of the particles within the sample were of a similar size. For the ellipsoidal particles, the long axis dimension was typically between 1.5 and 3 μm.

At the time of generation, a B1010 aerosol monitor was used to confirm the sample had been generated correctly. These haematite samples had all been sampled previously at DSTL and their typical $Af^a$ and size distribution on the B1010 was therefore known.

In addition, magnetite crystals were used as test particles. These exhibited similar morphology to haematite but with the median size being slightly larger and the size distribution somewhat greater. They were generated directly into the ballast chamber by a dry air jet.

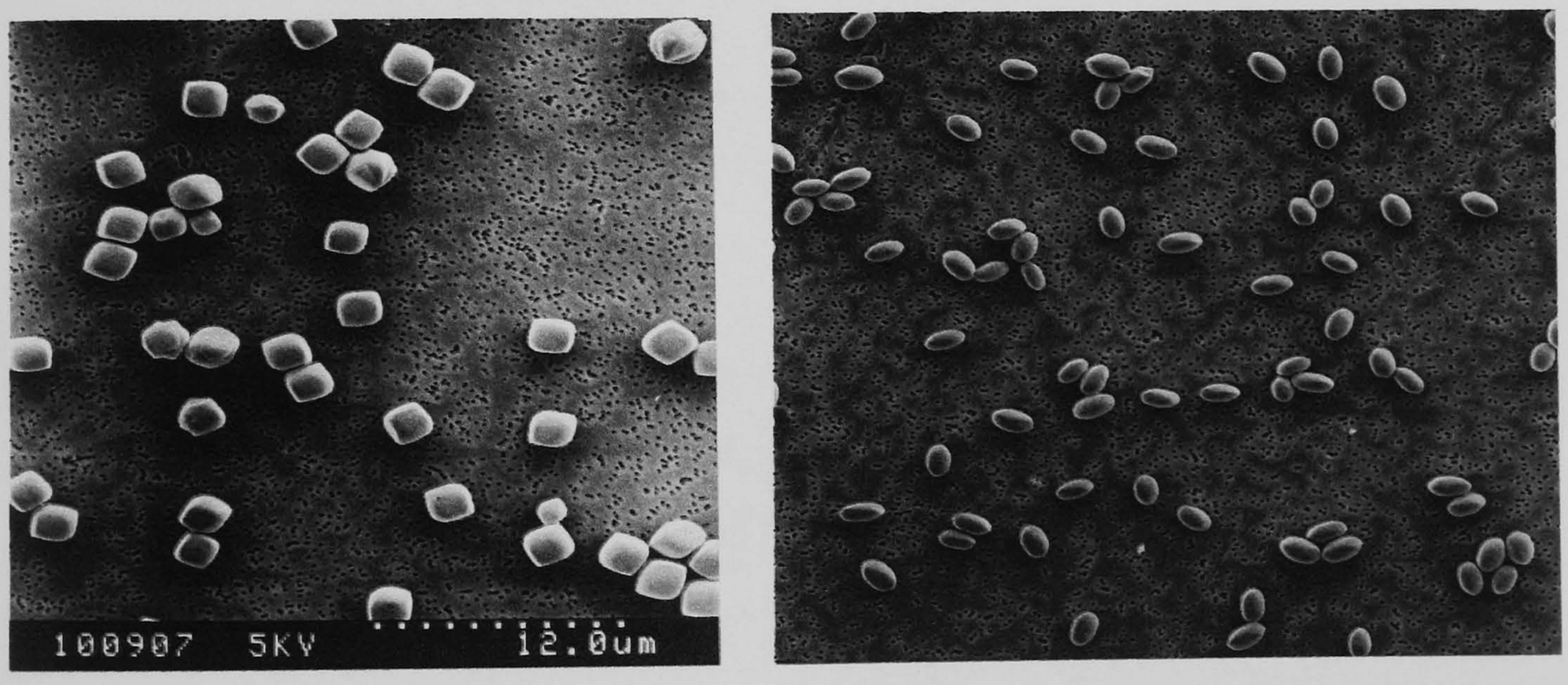



**Figure 6.1 SEM of near-cubic (left) and ellipsoidal haematite crystals. The long axis of the 3:1 aspect ratio crystals shown is around 3 μm.**

a The *Af* is a measure of the asymmetry of a particle's shape. A full definition of the term *Af* is given in the following chapter.

### 6.1.3.2 Cubic NaCl crystals

Cubic particles of various sizes were produced using the VOAG. Because of the essentially uncontrollable micro-environment in the ballast chamber in which the crystals form (as discussed above), the aerosol was monitored with an Aerodynamic Particle Sizer or UV-APS to provide corroboration of the calculated size. Sample crystals are shown in the SEMs of Figure 6.2. Note that the APS was unable to discriminate between single crystals and clusters of crystals forming a single particle of the same aerodynamic size, such as those illustrated to the right in Figure 6.2. These differences could be confirmed only by taking SEMs at the time of aerosol generation.

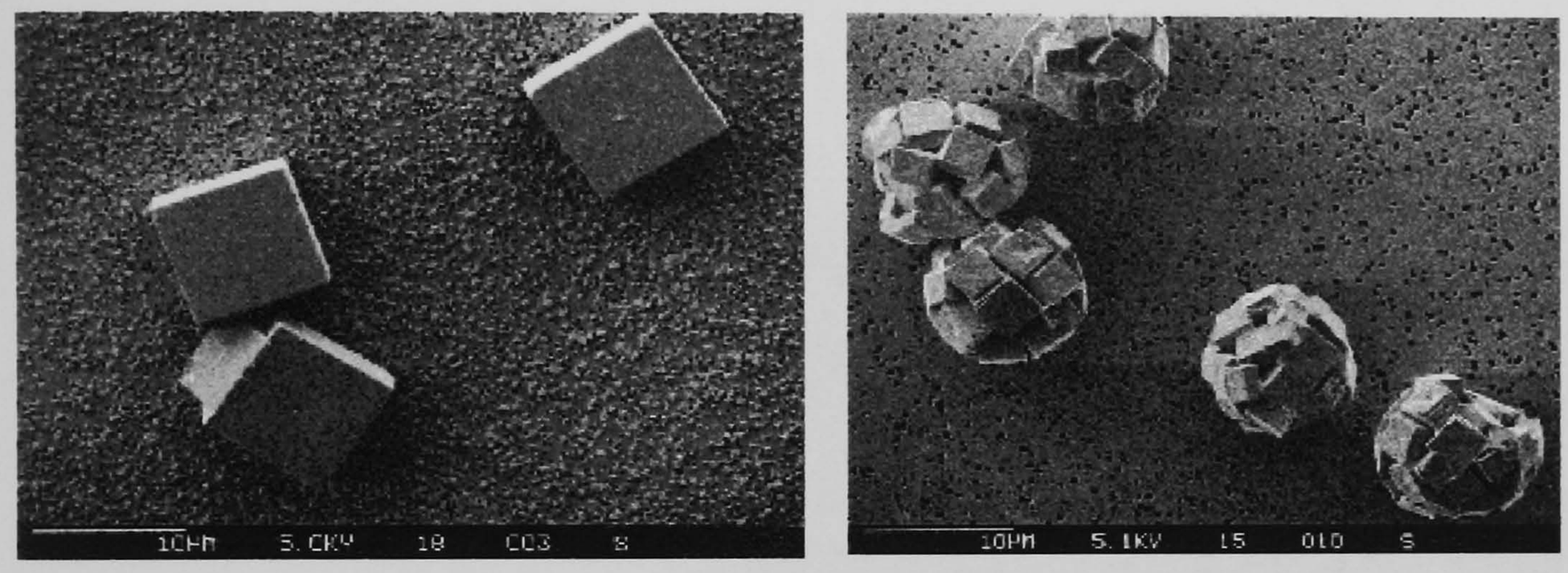
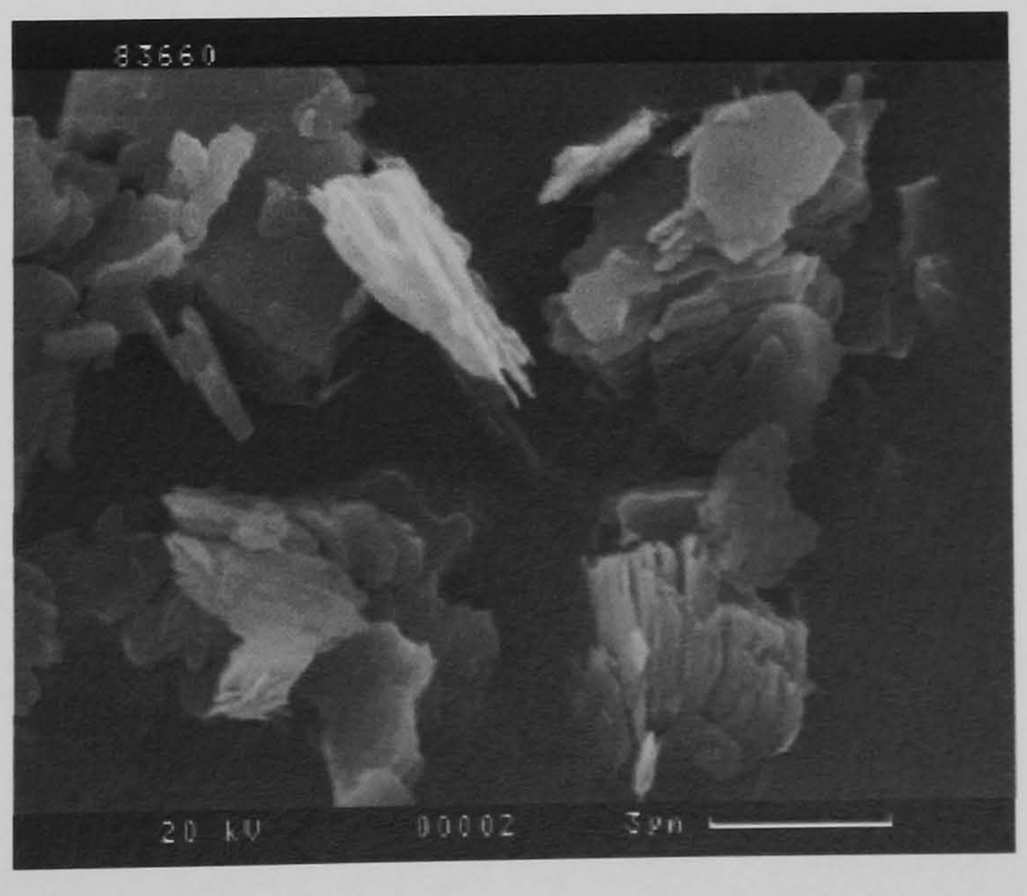



**Figure 6.2 Cubic NaCl crystals (left) and the more typical smaller clusters of crystals (right) that have coalesced to form a single particle.**

### 6.1.3.3 Flakes

Figure 6.3 shows a typical sample of kaolin in which irregular-cubic particles form from a series of stacked plates. The are often broken down into thinner plates as a result of mechanical abrasion and ultimately flake-like particles having face sizes in range from 0.5 μm to 5.0 μm result.

**Figure 6.3 SEM of kaolin showing its stacked plate structure**



### 6.1.3.4 Fibres

Gypsum ($CaSO_4$-2($H_2O$) Hydrated Calcium Sulfate) is a high aspect ratio fibre formed of bundles of long crystals with a maximum dimension of ~10µm. The precise size and aspect ratios seen in any particular sample can vary widely. This variation is largely a result of the mechanical forces applied to the bulk material by the air jet at the time of aerosolisation.

Caffeine fibres ($C_8H_{10}N_4HO_2O_2$) were also generated, this time by sublimation of caffeine crystals from a hot plate operated within the ballast chamber. The generation process was monitored with the B1010. Under the correct circumstances high aspect ratio fibres were formed by condensation. Unlike the gypsum fibres, caffeine was known to be highly fluorescent.

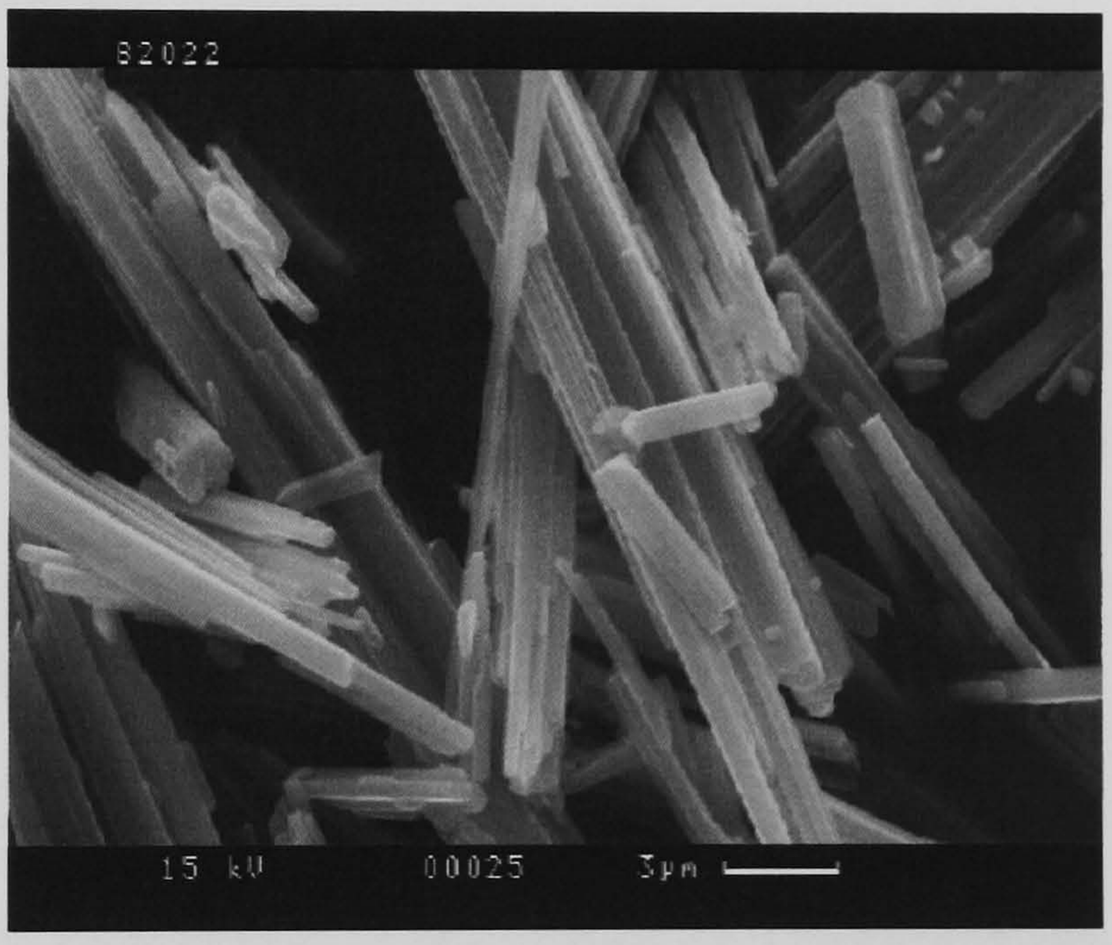



**Figure 6.4 SEM of gypsum showing its fibrous structure.**

## 6.1.4 Biological sample handling

Additional sample handling and aerosolisation procedures were adopted for the biological samples as described below. Quite apart from sample production issues relating to the materials themselves, two other factors come into play when handling biological samples. The first was the need to maintain a safe working environment for the laboratory personnel. As this work was carried out at DSTL a protocol was developed by DSTL personnel in conjunction with their safety officers (see below). The second requirement was to prevent the general contamination of the laboratory environment with live bacterial materials.

The risk of accidental contamination was reduced by having multiple levels of isolation of the sample so that, should the primary ballast chamber isolation fail or be accidentally

broken, secondary containment would ensure no sample escaped into the general laboratory environment.

A dedicated bio-aerosol sample chamber was constructed at DSTL to facilitate MPAM testing. This bio-aerosol chamber was mounted within the standard ballast chamber as described above (6.1.1) that was itself located within a large capacity fume cupboard. This arrangement provided three levels of isolation of the sample from the general laboratory environment. The aerosol sample connection tubes were fixed permanently to the instruments to prevent accidental discharge of the aerosol. All connections were made through valves and HEPA filters that enabled the sample chamber to be completely isolated and purged with the laboratory clean air supply between samples. Sample ports were arranged to allow aerosol sampling by both the UV-APS and the MPAM.

### 6.1.4.1 Sample production and generation

A general protocol for the production of microbiological samples was developed at the University with the advice of the Department of Biology and the University Health and Safety officer. The suggested procedure outlines the steps necessary to produce repeatable experimental samples, in this case for *B.atrophaeus* . The procedure can be applied (with some modification) more generally.

SEMs were to be taken of the aerosol samples at the sampling time along with collection onto appropriate media for culturing to establish viability levels of sample organisms.

University protocol for the production of *B.atrophaeus* spores:

1. Spore suspensions prepared by the Department of Microbiology are suspended in a measured volume of high purity sterile deionized water. (5mg/mL)

2. An aliquot of the sample to be removed aseptically and centrifuged. Remove the supernatant and resuspend the pellet in a volume of sterile deionized water equal to that removed.

3. This procedure must be repeated three times.

4. Break up any clumps by ultrasonic vibration of the sample in ice 15-30s. Clean the tip with 30% ethanol in water before and after use.

5. Dilute the sample by a known amount. (guess initially)

6. Aerosolise with the clean sterile nebulizer at a measured flow-rate.

7. Sterilise/ dispose of safely all sample handling equipment & materials at end of experimental run.

This protocol formed the basis of a similar approach used by DSTL. The precise details were modified by DSTL where necessary to reflect their equipment (e.g. type of centrifuge) and standard laboratory procedures .

### 6.1.5 Biological samples

The samples were all cultured by the microbiology laboratories at DSTL and were supplied both washed (centrifuged and re-suspended) and unwashed (suspended in growth medium). This meant samples from the same organism could be compared to investigate the contribution, if any, of the growth medium to the measured particle scattering profiles or any fluorescence. Where available, information on the culturing and preparation of the biological samples has been provided by DSTL and is given below for each material.

The biological samples used were:

**Bacillus atrophaeus - gram positive (*B.atrophaeus* , Bg).**

This is prolate spheroid approximately 0.8 μm in width by 1.5 μm long. Under adverse conditions it will form long lasting endospores, initially of the same dimensions.

**Preparation:**

Vegetative cells are grown overnight in L broth in an orbital shaking incubator at 37° C.

Spores are prepared from a stock suspension (supplied to DSTL by an outside contractor) and stored refrigerated. Prior to use the suspension is diluted and pasteurised at 70° C for 20 minutes to ensure no vegetative cells are present.

**Escherichia coli - gram negative (*E.coli*).**

A rod shaped bacteria approximately 0.5 μm in width by 1μm long.

**Preparation:**

Cells are grown overenight in L broth in an orbital shaking incubator at 37° C.

Washed cells are centrifuged twice at 7000 x g and resuspended in pure de-ionized water.

**Erwinia herbicola - gram negative (*Erwinia.h*).**

Spherical, less than 0.9 μm in diameter

**Preparation:**

Cells are grown for 24 hours in L broth in an orbital shaking incubator at 28° C.

Washed cells are centrifuged twice at 7000 x g and resuspended in pure de-ionized water.

(The above information regarding cultures and preparation was supplied by DSTL)

**Fungal spores lycoperdon (Puff Ball).**

These are roughly spherical, approximately 2 μm in diameter, with a highly dimpled surface and often with a stem associated with the spore's attachment point, see Figure 6.5.

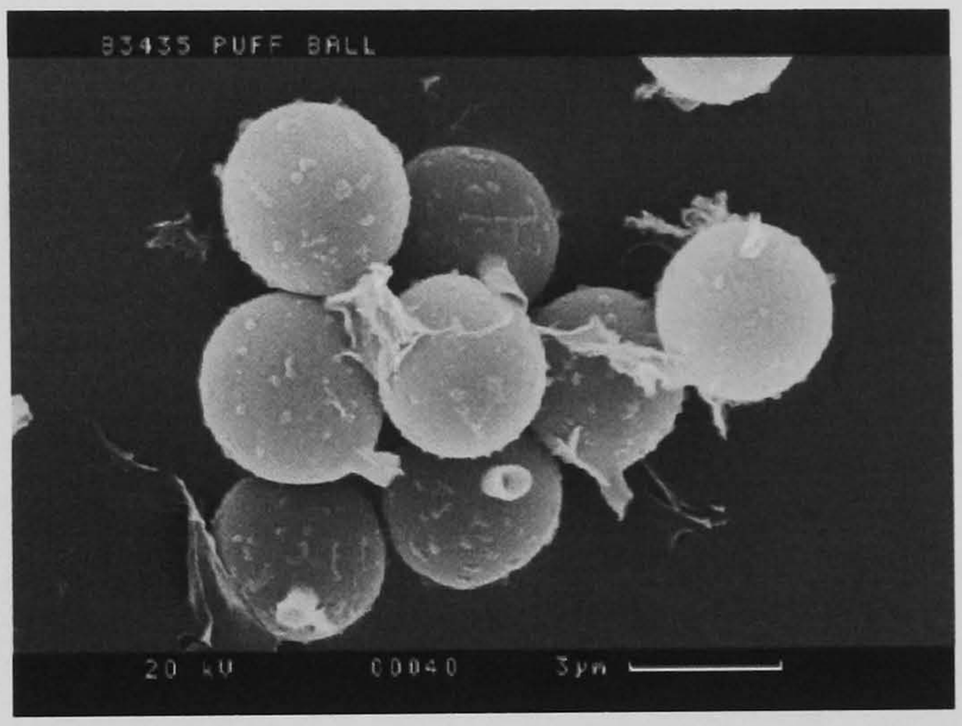



**Figure 6.5 Lycoperdon (puff ball) fungal spores.**

**Paper Mulberry (Broussonetia) pollen.**

Common small tree in southern US with pollen of varying size up to ~10 μm. Because of the method of collection these samples often contain other stray plant material as seen in Figure 6.6.

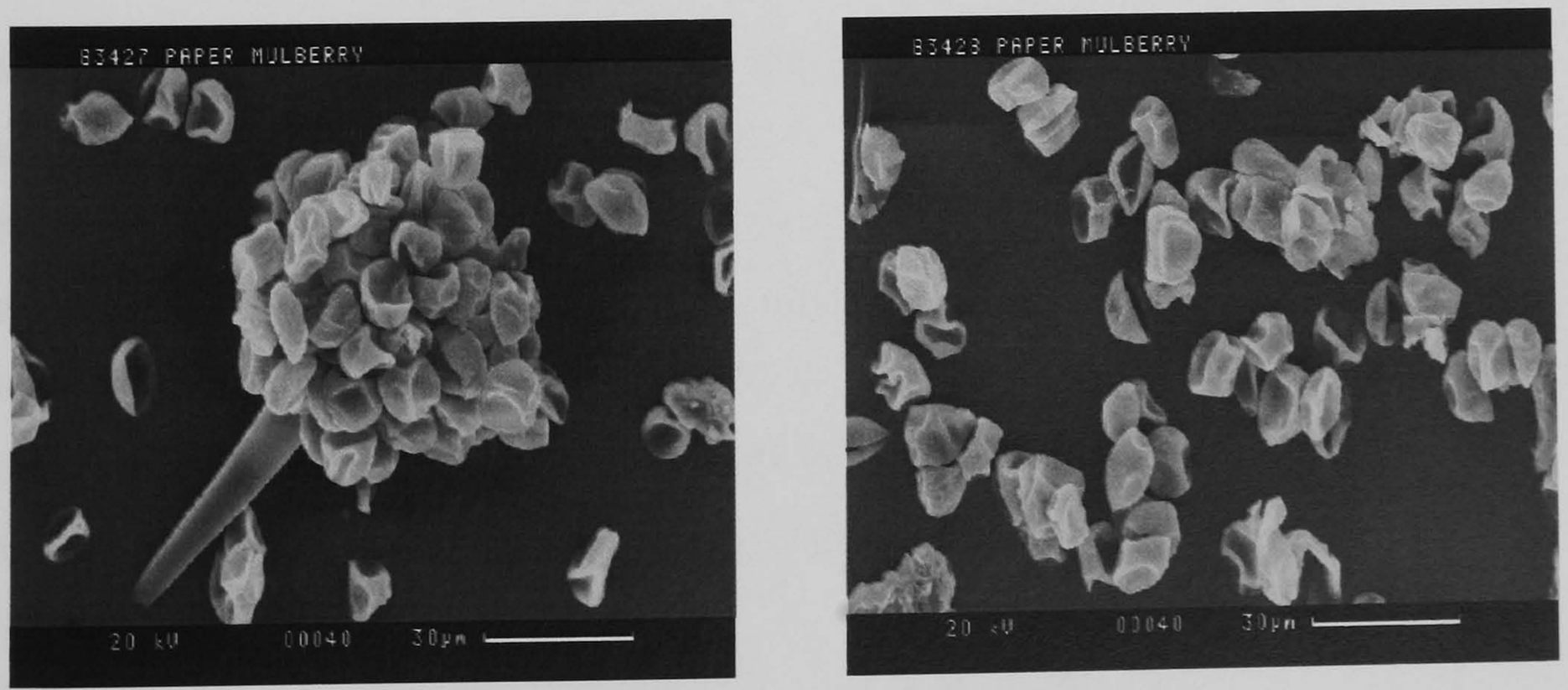



**Figure 6.6 Paper mulberry pollen.**

### 6.1.6 Sample generation procedures

As stated above, each sample was aerosolised independently by the most appropriate method for that material, and this is summarised in the table below. In the case of the bacterial samples the procedure described below was used to ensure there was no cross contamination between the samples.

#### 6.1.6.1 Bacterial sample generation procedure

A new washed and sterilised Bard min-nebulizer (C R Bard, New Jersey, U.S.A.) bottle was used for each sample. The supplied samples were diluted 50 µl in 10 ml of water. The nebulizer was operated at 10 psi. Around 12 s after the generator was started, aerosol was detected by the MPAM instrument. Sample generation was then continued for a further 30 s. This gave a high enough concentration for sample runs to be recorded over the following 10 minutes without any further aerosol being generated. On completion of data logging, a sample of aerosol was drawn onto an inline HEPA filter for 2 minutes by a vacuum pump. This was then isolated and the filter removed for examination by electron microscopy. The inner sample chamber was then cleaned by use of the purge air supply. The aerosol nebulizer was removed and replaced by a clean nebulizer bottle containing water and aerosol was generated until the MPAM detected just water droplets. Finally , the purge air was run again until the chamber was clean, at which point the next sample could be generated.

### 6.1.7 Experimental measurements

The aerosol discussed above were generated on various dates during visits to the DSTL laboratories where the MPAM instrument was permanently located following the constructional stage at the University.

Prior to logging aerosol data, confirmation that the aerosol had been generated properly was made by independent sampling using the B1010 or UV-APS. The proper functioning of the MPAM was monitored via the instrument's graphical display. If necessary, adjustments were made in the generation process to optimise aerosol production. For example, fine tuning of the operating frequency of the VOAG had to be carried out during operation until it reached a stable operating point. This tuning process could not be carried out in advance since the VOAG's output had to be measured in real-time.

| Material | Size (μm) approx | Mono/poly disperse | Aerosol Generation |
|---|---|---|---|
| Oleic acid | 1.0 | Mono | VOAG |
| Oleic acid | 3.0 | Mono | VOAG |
| PSL | 1.0 | Mono | Nebulizer |
| PSL (fluorescent doped) | 1.7 | Mono | Nebulizer |
| Water droplets | 1-10 | Poly | Nebulizer (hand held) |
| Silica | 3-4 | Poly | Dry compressed air |
| Haematite | 1-3 | Poly | Dry compressed air |
| Kaolin | 0.5-5 | Poly | Dry compressed air |
| Magnetite | 1.3 | Poly | Dry compressed air |
| NaCl | 2 | Mono | VOAG (via solution) |
| NaCl | 3 | Mono | VOAG (via solution) |
| Gypsum | 0.5-10 | Poly | Dry compressed air |
| Caffeine | 1-5 | Poly | Thermal evaporation/condensation. |
| BG spores | 1-2 | Poly | Nebulizer |
| BG clusters | 2-5 | Poly | Nebulizer |
| BG vegetative | 1-2 | Poly | Nebulizer |
| E coli (washed) | 1 | Poly | Nebulizer |
| E coli (unwashed) | 1 | Poly | Nebulizer |
| Erwinia (washed) | 1 | Poly | Nebulizer |
| Erwinia (unwashed) | 1 | Poly | Nebulizer |
| Paper Mulberry pollen | 2-4 | Poly | Dry compressed air |
| Puff Ball pollen | | | Dry compressed air |

**Table 6.1 Summary of aerosol test materials and their methods of aerosolisation.**

MPAM aerosol scattering and fluorescence data were logged to disk on the MPAM host PC. Each file contained data from 1,000 to 5,000 particles, taking around one minute to log. For each sample, an initial file was logged with the PMT gain voltage set to 700v; further files were then logged with the PMT voltage increased from 500 to 750v in 50v steps. A final file was logged with the PMT voltage again set to 700v. It was intended that this second file might be useful in identifying any ageing effects in the sample due to drying, differential sedimentation, or any other cause.

Each data file had a header block recording the UV power level as measured by the MPAM at the time of logging, the PMT gain voltage setting, and the minimum and maximum time-of-flight window. Each subsequent row of the file then contained data from the 24 elements of the HPD followed by the trigger channel magnitude, the fluorescence channel magnitude and the actual time-of-flight recorded for that particle, making 27 data values in total per particle. At the end of the file was placed an additional block containing the same information as the header, but with the UV power level recorded from an updated reading

taken at the end of the data logging. Normally, the two UV power levels would agree to within a few milliwatts.

During the experimental phase of the work it was considered useful to have a rapid method of carrying out an initial analysis of the data since this could serve to confirm that the MPAM was continuing to operate correctly. It also meant that any anomalies in the results could be highlighted while the experimental work was under way and further sets of results taken to verify the data from that aerosol. It was decided to employ a previously developed Size vs. $Af$ (these terms are fully described in Section 7.1) data processing algorithm for this purpose. These functions were simple to implement and could be carried out rapidly by the use of spreadsheets. There was also a degree of familiarity with the degree of particle discrimination that could be achieved using this method, at least as far as scattering profiles were concerned, from prior experience with both previous University instruments and data gathered by DSTL scientists.

Detailed analyses of the data were then carried out at a later date. A description of the analysis methods used and the results produced are presented in the following chapters.

# 7 Preliminary data analysis using size, asymmetry and fluorescence parameters

As well as enabling a rapid initial assessment of the MPAM's performance at the time of data logging, these parameters were also used to carry out a preliminary but nonetheless valuable post-experiment analysis of the logged data. A description of the method employed and presentation of the results are given below.

## 7.1 Size vs. Asymmetry Factor

A measure of a particles size (Sz) may be defined as the mean scattering intensity received by the MPAM HPD detector. This may be correlated with a spherical equivalent particle diameter via Mie theory, as described in Chapter 3, and, with the use of spherical calibration particles, an absolute spherical size could be determined for a particular instrument. However, in many situations, and especially where the primary concern is to *discriminate* between different particle types, an absolute measure of size is not required. Therefore, for the purposes of a comparative analysis, the simple measure of the mean scatter intensity expressed in arbitrary units has been used as the size parameter.

The Asymmetry Factor (*Af*) is a measure of the variation in scattered light intensity around the 360° azimuth surrounding the laser beam axis and is therefore governed primarily by particle shape and orientation with respect to the illuminating beam. The parameter has been employed in a wide variety of aerosol measurement instruments developed by DSTL, the University of Hertfordshire, and others. The formula to calculate *Af* is:

$$Af = \frac{k\sqrt{\sum_{i=1}^{n}\left(\overline{E} - E_i\right)^2}}{\overline{E}}$$

where n is the number of pixels in the azimuthal ring, E is the output value of each pixel, E-bar the mean of all E values and k is a constant to render the maximum value of *Af* to be 100, where k is evaluated as:

$$k = \frac{100}{\sqrt{n(n-1)}}$$

Inspection of the formula shows that for scattering from a perfect sphere, such that the output value E is the same for all pixel elements, the *Af* equates to zero. That is, there is no azimuthal asymmetry in the scattering profile. In practice, *Afs* falling between zero and five are usually associated with perfect spheres since measurement errors arising from various noise or quantization sources rarely allow values of zero to be observed. *Afs* up to 10 may result from scattering from slightly roughened or slightly faceted spheres such as silica. At the other extreme, scattering from vertically aligned high-aspect ratio fibres would result ideally in scattering to two diametrically opposed pixels, all other pixels receiving no scattered light (see, for example, the bottom-left scattering profile image in Figure 3.6). This equates to a maximum *Af* value of 70.[a]

The calculation of the *Af* was automated by using a preformatted spread sheet that first applied gain correction factors before processing the data. New data could then simply be pasted into the spreadsheet. Although not providing real time data processing, this method allowed Size versus *Af* plots similar to those given below to be displayed immediately after the data was logged to file.

The *Af* formula takes no account of the output on the fluorescence channel. However by normalising the fluorescence data for measured particle Size, a second plot of Normalised Fluorescence $Fl^{sz}$ versus *Af* could be made. Such a procedure could be justified to the extent that the particle's fluorescence was proportional to particle volume, while scatter intensity was approximately proportional to its cross-sectional area. (This makes a number of assumptions such as that the number of fluorophore molecules relates to the particle's volume[b], which may only hold true for some bioaerosols. It ignores the influence of re-absorption and resonance effects within the particles. Nevertheless, it has proven to be sufficiently useful to make discrimination between otherwise indistinguishable data possible, as is shown below).

Thus, the two plots could easily and rapidly be produced to indicate how the MPAM instrument was performing. The size vs. *Af* plot enabled an immediate assessment of the degree of discrimination that could be achieved on the basis of shape and size alone, whilst the $Fl^{sz}$ versus *Af* plot provided a rapid indication of any further discrimination that could potentially be achieved by incorporation of fluorescence data.

a In theory, scattering to a single pixel would result in an *Af* of 100. In practice, it would be almost impossible for a particle transiting the scattering volume to produce such a scattering pattern.

b A reasonable assumption for homogenous aerosols such as oleic acid droplets or PSL spheres.

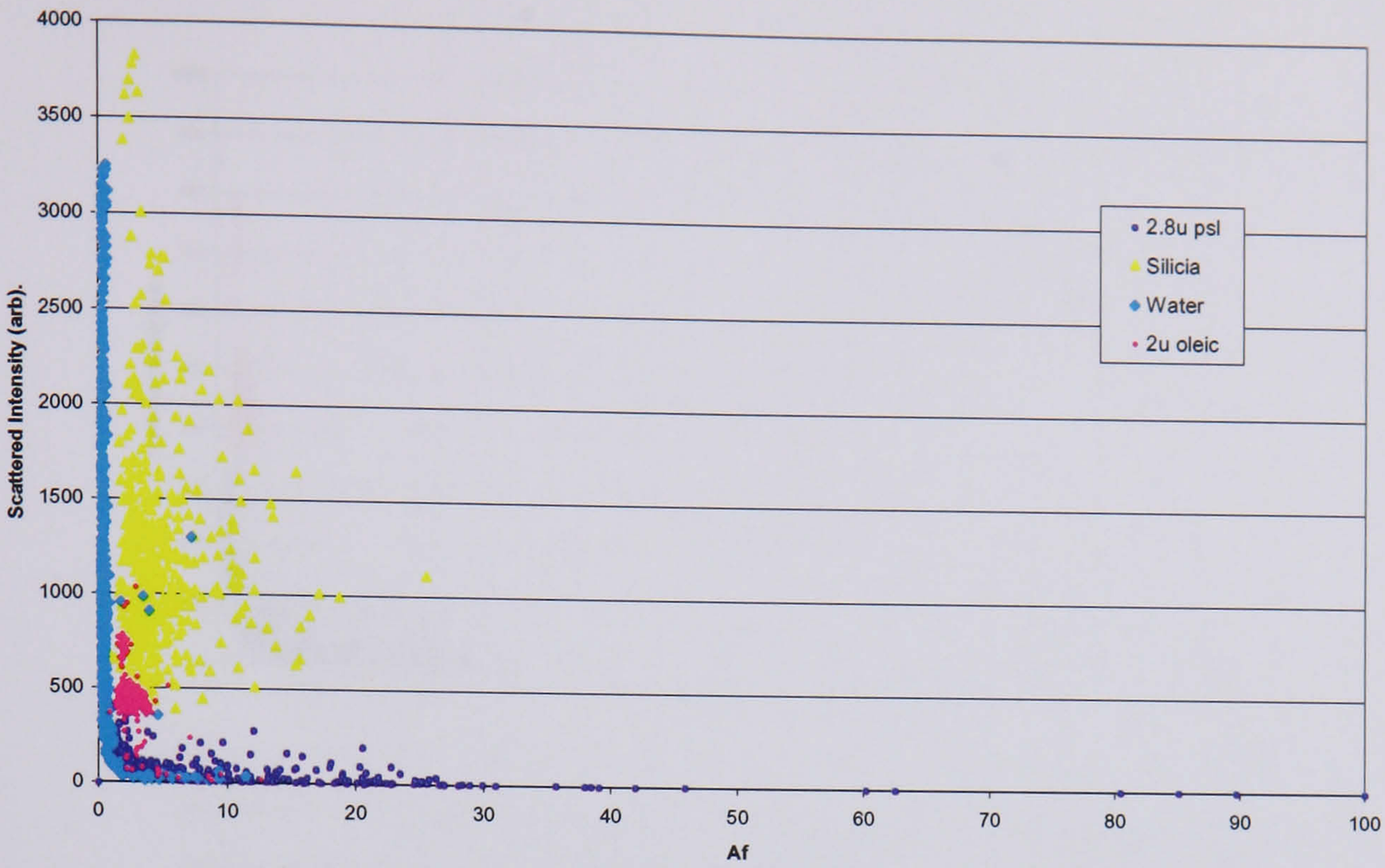



**Figure 7.1 Size versus *Af* plot from 2.8 μ PSL, 2 μ oleic acid droplets, a polydisperse water sample and the slightly faceted silica spheres.**

## 7.2 Af and fluorescence plots

### 7.2.1 Spherical particles

The plot of Figure 7.1 shows size against *Af* for PSL spheres, 3 μm silica 'spheres', oleic acid droplets and a polydisperse sample of water droplets. At the lowest scattering intensities there is a 'tail' of data towards higher *Af* values. This is mostly clearly seen in both the water and PSL data. Several factors contribute to this: some genuine non-spherical particles present in the laboratory air as can be verified by background monitoring; secondly, scattering off dried solid residues from droplets (this is particularly noticeable in the PSL as discussed below); finally, degraded S/N levels from the smallest particles will result in higher calculated *Af*s, even for spherical particles. This is an unavoidable limitation for the smallest particles detectable.

Although the range of *Af*s recorded from the silica is reasonably well constrained, these particles exhibit a larger than expected size range. It appears that the dry silica powder was not broken down into its fundamental particles upon aerosolisation, and some aggregation occurred. Consequently 'sizes' up to the top of the detectable range (~10 μm) were recorded.

The polydisperse water droplets, as expected, result in a data distribution having *Af* values close to zero for all particle sizes except those well below 1 μm in size.

In Figure 7.2, the same data from the oleic acid and PSL has been re-plotted at a larger scale for clarity. The oleic acid droplet data are precisely defined in terms of both *Af* and size

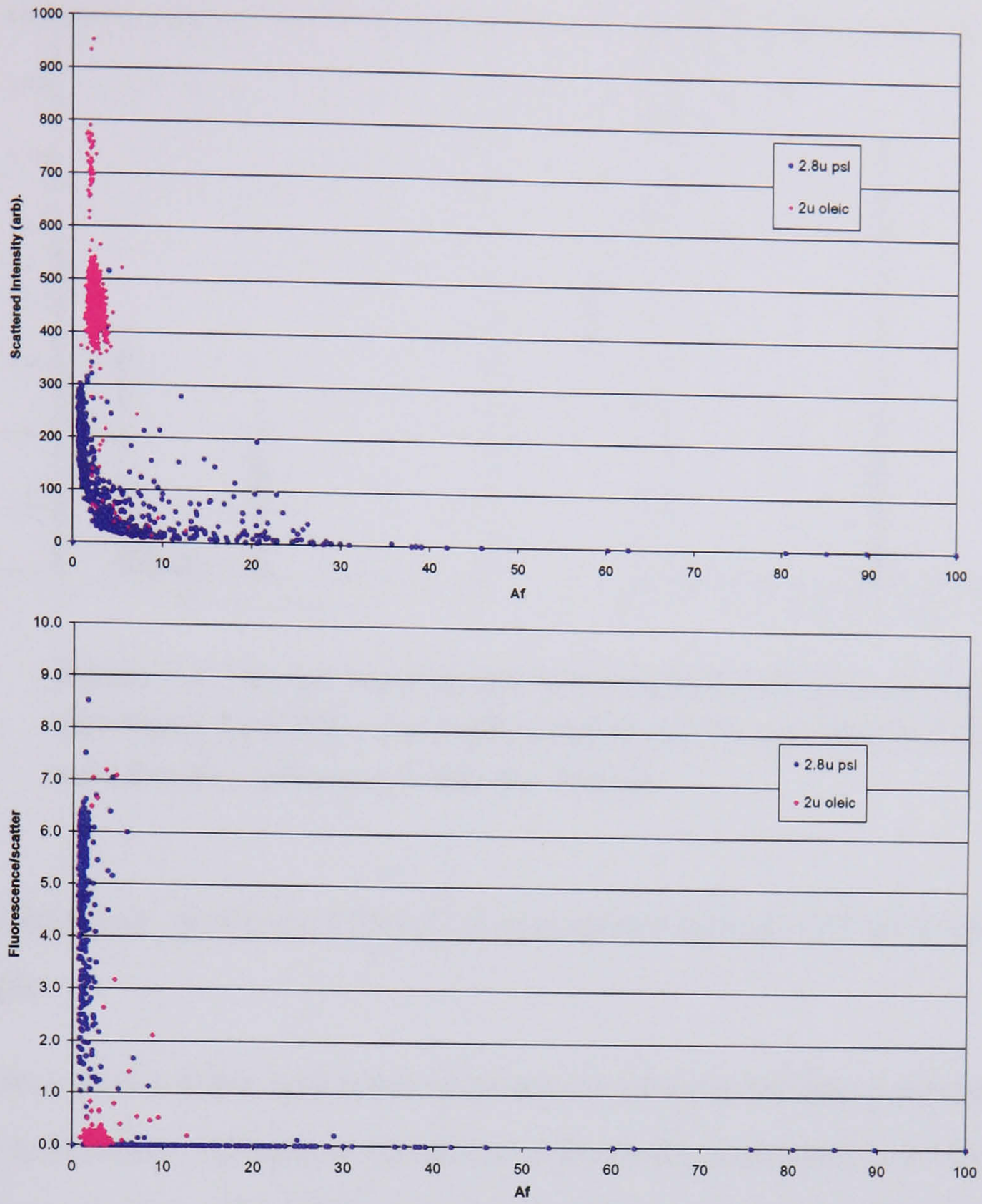



**Figure 7.2. The upper plot shows size versus *Af* for two aerosols, PSL and oleic acid droplets. The PSL data exhibits a long tail of very small particles with increasing *Af* values. The lower plot shows size normalized fluorescence ($Fl^{sz}$) values plotted against *Af* for the same data set.**

distribution. A small distribution of larger oleic acid droplets is clearly distinct from the bulk of the oleic acid data. These are doublets formed from the coalescing single droplets as discussed in the calibration section of Chapter 4. In contrast, the PSL data appears as a single continuous distribution with no clear break between spherical ($Af < 5$ ) and the smaller non-spherical particles. This reflects the difference in the way in which these two particles were produced and aerosolised. The PSL spheres are generally supplied suspended in a liquid containing surfactants and fungicides. If the PSL is aerosolised using the nebuliser without being pre-washed, large numbers of small surfactant or fungicide residual particles are also produced from dried liquid droplets of the suspending medium.

A typical distribution for an unwashed 2.2 μm PSL sample is shown in Figure 7.3 as recorded on an Aerodynamic Particle Sizer. As can be seen from these data, a large number of very small residue particles are generated along with the PSL. Some of these unwanted particles will be large enough to trigger the MPAM and consequently will be recorded in

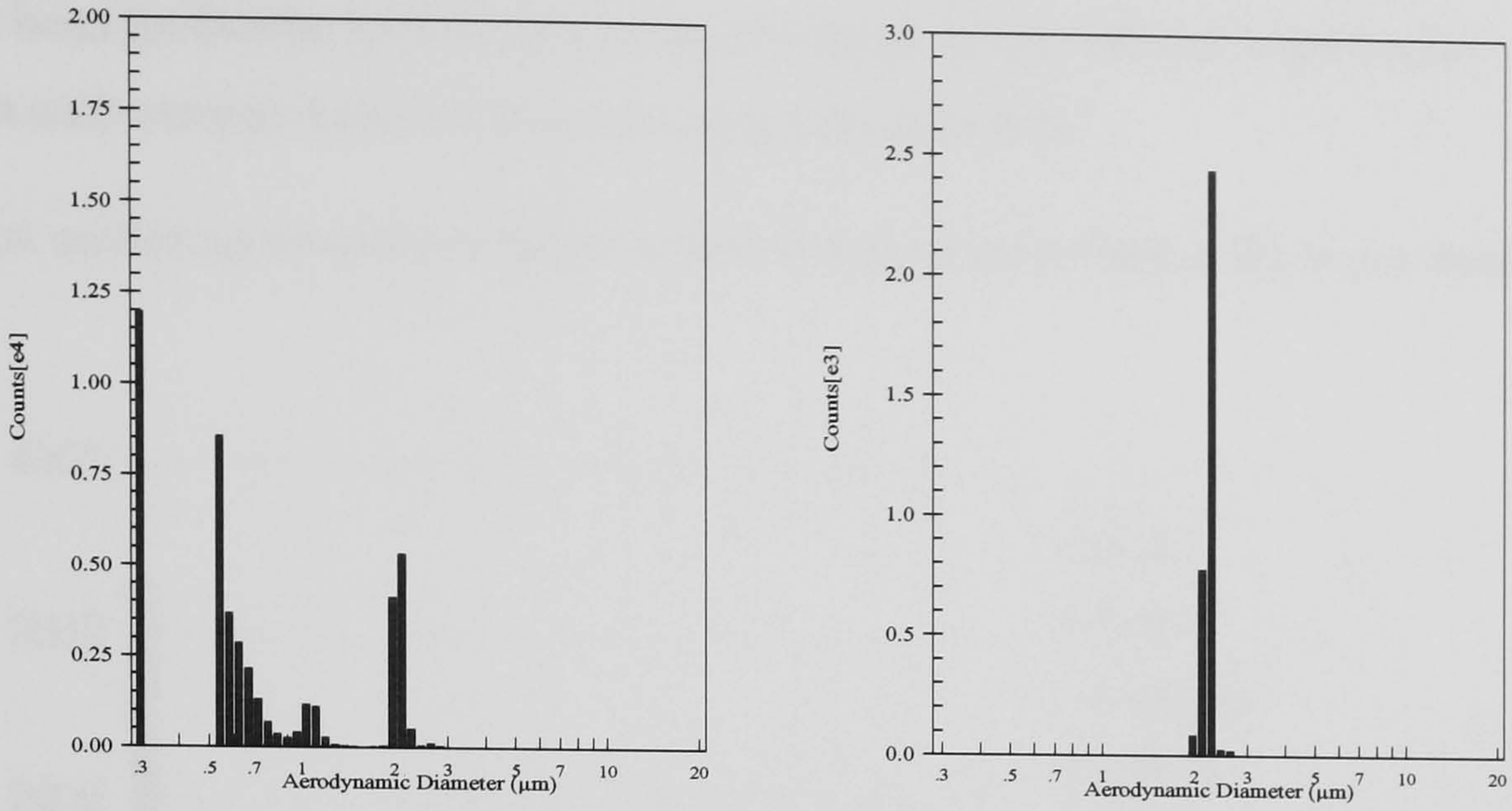



**Figure 7.3 The left hand graph shows typical APS data recorded from unwashed 2µm PSL. The right hand graph is recorded from 2.2µm oleic acid droplets generated with the VOAG**

the data. For comparison purposes, Figure 7.3 also shows typical APS data for a 2.2 µm oleic acid sample.

Although the oleic acid solvent will result in some nanometric residue particles, the number seen above the instrument's detection threshold is generally extremely low. Considering the lower plot of Figure 7.2 showing $Fl^{Sz}$ against *Af*, the oleic acid droplets form a distinct cluster similar to that seen in the Sz against *Af* plot above. In this case, the fluorescence data does not make further discrimination of the sample characteristics possible. In contrast, the PSL data in the lower graph now appears divided into two populations. The bulk of the data has high fluorescence values and *Af*s below 5; the balance of the PSL data has low fluorescence values with *Af*s in the range 0.5 to almost 100. In this example, the $Fl^{Sz}$ versus *Af* plot allows what at first appeared to be a continuous data set to be separated into two distinct sample populations: PSL and residue particles. This use of the fluorescence parameter makes discriminating between the fluorescent PSL and the non-fluorescent surfactant residues possible. Discriminating on the basis of size or *Af* alone does not make it possible to exclude data generated from the residue particles.

### 7.2.2 Non-spherical particles

Experimental data were recorded from a variety of non-spherical particle shapes and are plotted in Figure 7.4. The materials used are as described in Chapter 6 and were generated independently. Data from a polydisperse sample of water droplets are also included for comparison purposes. The haematite grains have an aspect ratio of 3:1. The gypsum and caffeine fibres vary in aspect ratio from ~3:1 to 8:1, and also vary over a large size range. It

would have been preferable to use a monodisperse range of the different aspect ratio particles but such sample materials were not generally available.

At the lowest scattering intensities there is a tail of data towards higher *Af*s as has been

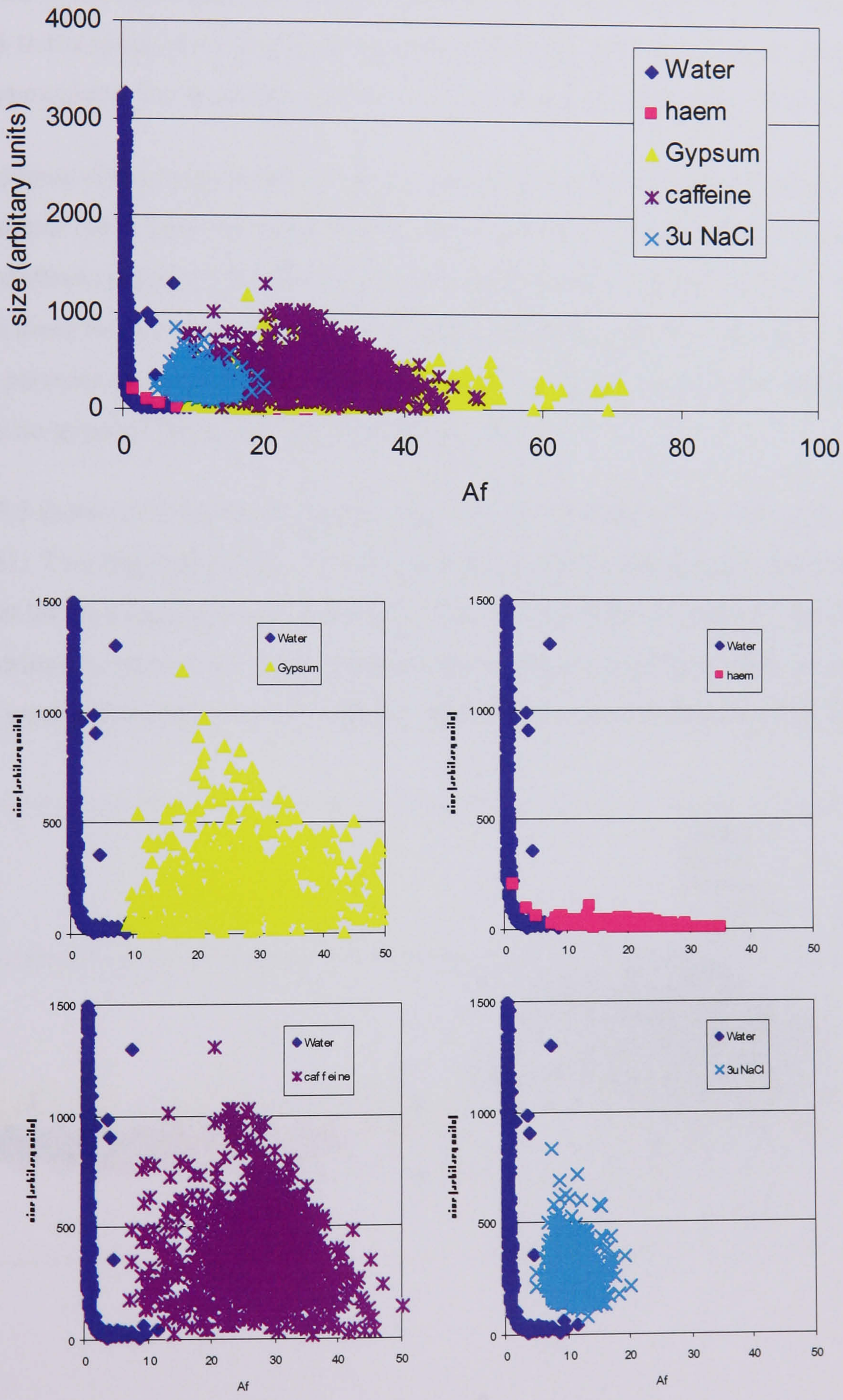



**Figure 7.4 Size against *Af* plots for cubic NaCl crystals, 3:1 ellipsoidal haematite, gypsum and caffeine fibres. The lower thumbnail plots show the sample space occupied by the different materials. The water droplets have been included for comparison purposes.**

discussed above. While this is readily apparent in the data from spherical particles, it will also be present in the data from the smaller of the non-spherical particles, although not so readily apparent. It is therefore to be expected that for the smaller particles, the *Af*s recorded will be slightly higher than might be recorded otherwise.

In contrast to the water droplets, the non-spherical haematite, gypsum and caffeine particles have *Af*s in the range of ~5 to ~70. As previously described, the *Af* formula used will have a value approaching 70 for scattering from a perfectly aligned high aspect ratio fibres.

A high degree of discrimination can easily be established between the spherical and all other particle types. Both the cubic NaCl crystals and the ellipsoidal haematite generally occupy distinct regions of the size versus *Af* sample space making it possible to discriminate between these two particle types. There is some overlap between the haematite and the smaller particles of gypsum. This reflects the common morphology of the smaller low-aspect ratio gypsum fibres and that of the haematite.

Figure 7.5 shows a *Sz* versus *Af* plot for 3 μm and 1μm standard PSL and 1.7 μm fluorescent dyed PSL. Two high aspect ratio fibrous materials, caffeine and gypsum, are also shown. Note that this is a log-log plot to aid clarity. The size distribution of the 1.7 μm fluorescent PSL overlaps, to some extent, the data from the smaller 1μm standard PSL. This suggests that the scattered intensity for the different types of PSL may be influenced by the dying

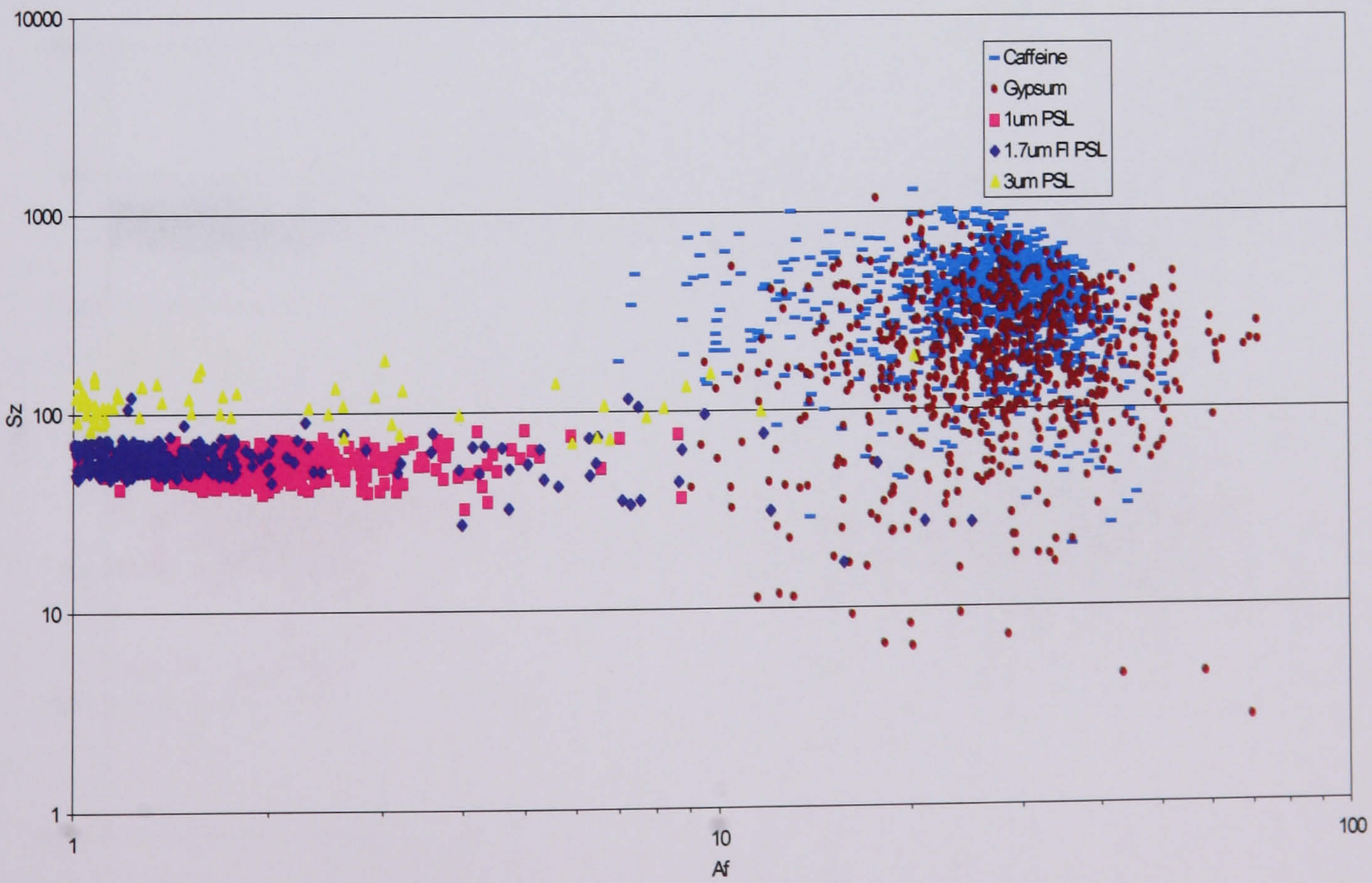



**Figure 7.5 Size versus Af plot of three different spherical particles and two highly fibrous materials, gypsum and caffeine. Note the use of a log log scale**

process, though this possibility has not been investigated further here. In the case of the gypsum and caffeine there is a significant overlap in the data making particle classification uncertain. This is to be expected given their common morphology.

A plot of $Fl^{sz}$ against *Af* for the same set of data is shown in Figure 7.6. Again this has been plotted on a log-log scale. A large dynamic range of 100,000:1 occurs between the lowest *Flsz* values recorded for the gypsum and the highest values from the dyed fluorescent PSL. The use of the fluorescence data makes it easily possible to discriminate between the two previously overlapping sets of data from the 1 μm and 1.7 μm PSL. The use of the fluorescence parameter also makes it possible to classify the two fibre types into distinct classes because of the caffeine fibre's higher level of fluorescence.

### 7.2.3 Biological materials

Size versus *Af* and $Fl^{sz}$ versus *Af* plots for three species of bacteria are presented in Figures 7.7 and 7.8 respectively. These materials were prepared and aerosolised as described previously in Chapter 6.

The data from the bacterial samples have *Afs* in the range of ~10 to ~35, similar to those recorded for low aspect ratio particles such as haematite. This accords well with the known ellipsoidal shape of these bacteria. The sizes recorded cover the range of 30 to 150 (arbitrary units) which would approximate to the scattering from a 2 μm oleic acid droplet;

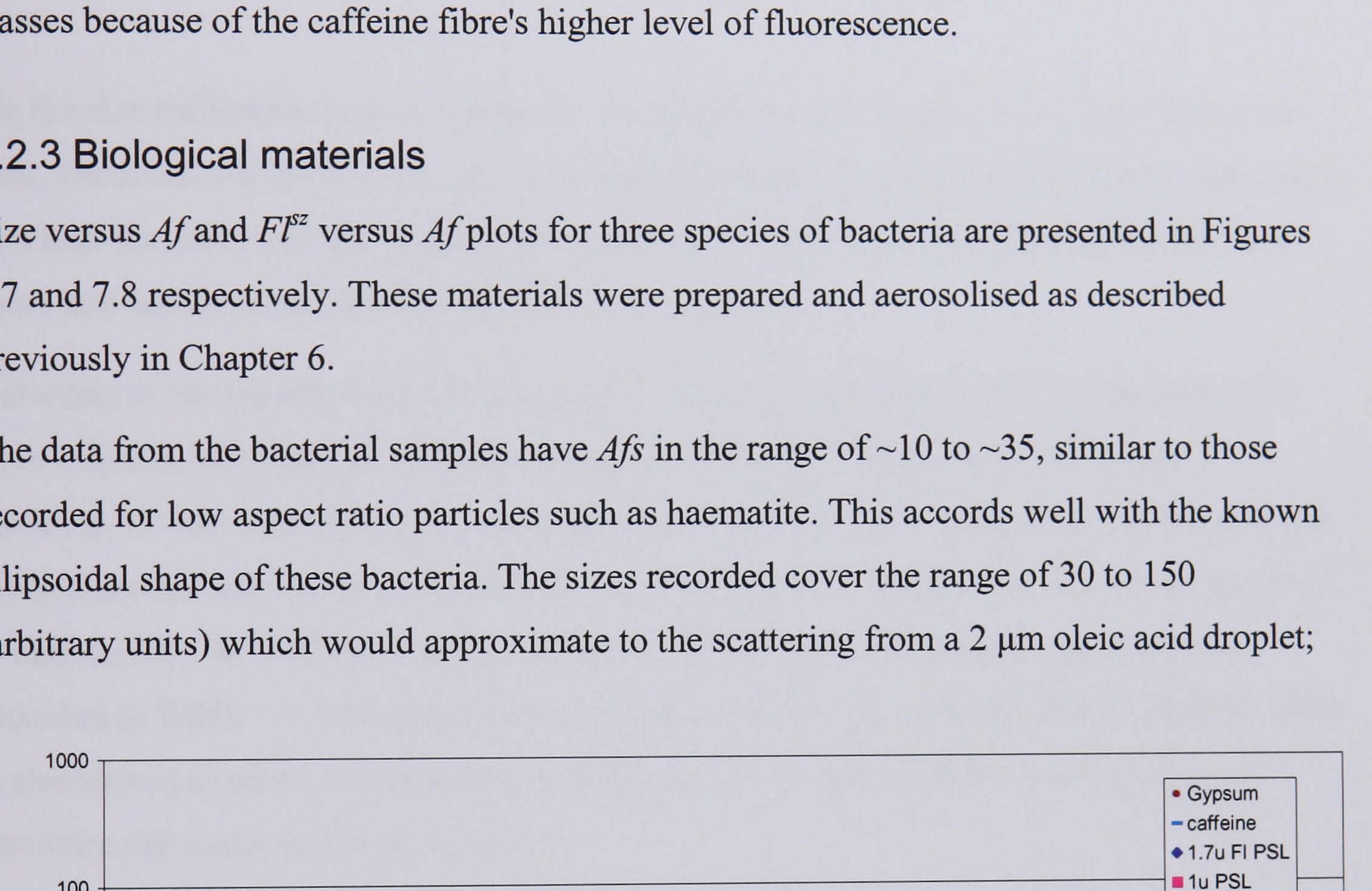



**Figure 7.6 $Fl^{sz}$ versus *Af* plotted for the same range of particles shown in figure 6.5. The different particle types now occupy separate areas of the sample space.**

| | Median | | Lower quartile | | Upper quartile | |
|---|---|---|---|---|---|---|
| | $Fl^{sz}$ | magnitude | $Fl^{sz}$ | magnitude | $Fl^{sz}$ | magnitude |
| Bg | 10.8 | 565 | 6.5 | 343 | 16.3 | 996 |
| E.coli | 24.5 | 1440 | 17.7 | 1069 | 34.4 | 1809 |
| Erwinia | 6.3 | 311 | 44.47 | 222 | 9.02 | 390 |

**Table 7.1 The distribution of fluorescence recorded from three different bacterial species given in terms of their percentile values. Both size normalised values ($Fl^{sz}$) and the absolute magnitude of the fluorescence signal (arbitrary units) are given.**

see the size calibration data of Chapter 4. Although there is significant overlap in the size data, erwinia.h has lower levels of scattering, reflecting its smaller size. However, the extent to which the data from the three bacterial species overlaps makes discrimination between them impossible on the basis of the Sz versus *Af* plot.

Referring to the fluorescence plot of Figure 7.8, the overall range of $Fl^{sz}$ values recorded have a similar distribution for the three species of bacteria. Inspection of the graph, however, reveals that while there is a degree of overlap, the data from each species is in fact centred around a different co-ordinate point within the two-dimensional sample space. This is indicated by the different median values for $Fl^{sz}$ for each of the bacterial species as recorded in Table 7.1. The magnitude of the fluorescence signal, expressed in arbitrary units is also shown to allow a comparison with relative fluorescence levels as reported in the literature, for example Seaver *et al.*[30]

While the Sz versus *Af* data recorded by the MPAM shows there is little difference in the reported size for these bacteria the differences in the fluorescence signal implies that the number of accessible fluorophore molecules varies between the species tested. This suggests that, with the level of sensitivity that can be achieved with the MPAM instrument, some discrimination between the various bacterial species could be achieved, especially if more sophisticated analysis methods were applied to the data.

The relative magnitude of the fluorescence signal detected by the MPAM from the various bacterial samples is similar in magnitude to that recorded from 3 μm PSL microspheres. This has been plotted along with the bacteria for comparison in Figure 7.8.

Two further biological materials, paper mulberry pollen and puff ball spores, have been tested. The results are plotted for Sz versus *Af* and $Fl^{sz}$ versus *Af* in Figures 7.9 and 7.10 respectively. Bg spore data has also been plotted on these figures, this being representative

of the MPAM response to the bacterial samples. As would be expected from the known size and shape of the pollen samples, these materials have *Af* values in the range of ~3 to ~25. This shape parameter encompasses the range from the spherical to those recorded from slightly roughened spheres, such as silica, to cubic NaCl, with a part of the data extending out to the region occupied by the highly fibrous particles such as gypsum.

Examination of the relevant scanning electron micrographs shows that both these samples contain \significant numbers of background particles of varying shapes and sizes. In the case of the paper mulberry these are likely to be background dust particles and other particles of plant material collected accidentally at the time of sample collection. The puff ball sample is also likely to contain particles of the mushroom body, along with short chain like strings of spores resulting from incomplete separation into individual spores as they were ejected.

For the Bg spores, their smaller size and higher *Af* values makes it possible to generally discriminate between the bulk of the bacterial and pollen spores, although it would not be possible to partition the data using Sz versus *Af* data alone. The Bg spore data, being typical of the bacterial data (see Figures 7.7 and 7.8), suggests that the Sz versus *Af* plot can go some way in discriminating between different classes of biological material.

Inspection of Figure 7.10 shows that the partial separation between the bacterial and pollen samples achieved in sample space of the shape discrimination data is increased by the use of fluorescence data. The sample area occupied by the bacterial spores now has virtually no overlap with that of the paper mulberry and puff ball. The bacterial spores exhibit a comparatively high level of fluorescence compared to that from the pollens. The particularly low level of fluorescence recorded from the puff ball is surprising given the high protein content that is generally ascribed to fungi spores [79]. The presence of amino acid residues of tryptophan is commonly thought to be similar to that found in bacterial proteins. It is therefore surprising that significantly different measures of fluorescence are observed by the MPAM between these two classes of spore.

The data from 3 μm NaCl has also been included as a contour line that includes approximately 90% of the NaCl sample. The fluorescence level recorded for both the paper mulberry and puff ball spores have a virtually identical range of values as that recorded from the non-fluorescent NaCl. In particular, in the case of puff ball data in both the $Ff^{sz}$ versus *Af* graph and the Sz versus *Af* graph, the puffball data occupies the same sample space as the NaCl, making discriminating between them impractical.

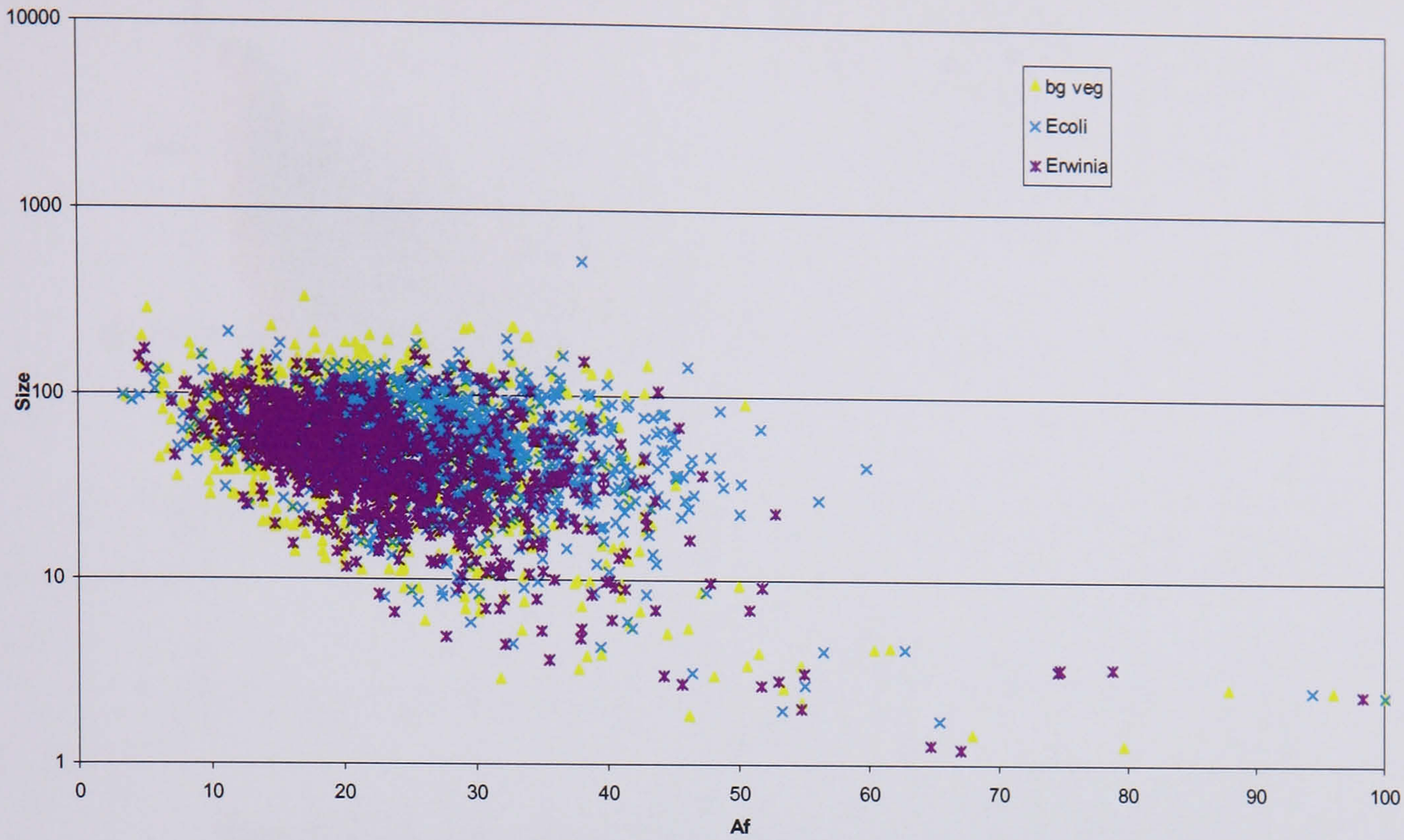



**Figure 7.7 Size versus *Af* plot for the bacteria *erwinia.h, e.coli* and Bg.**

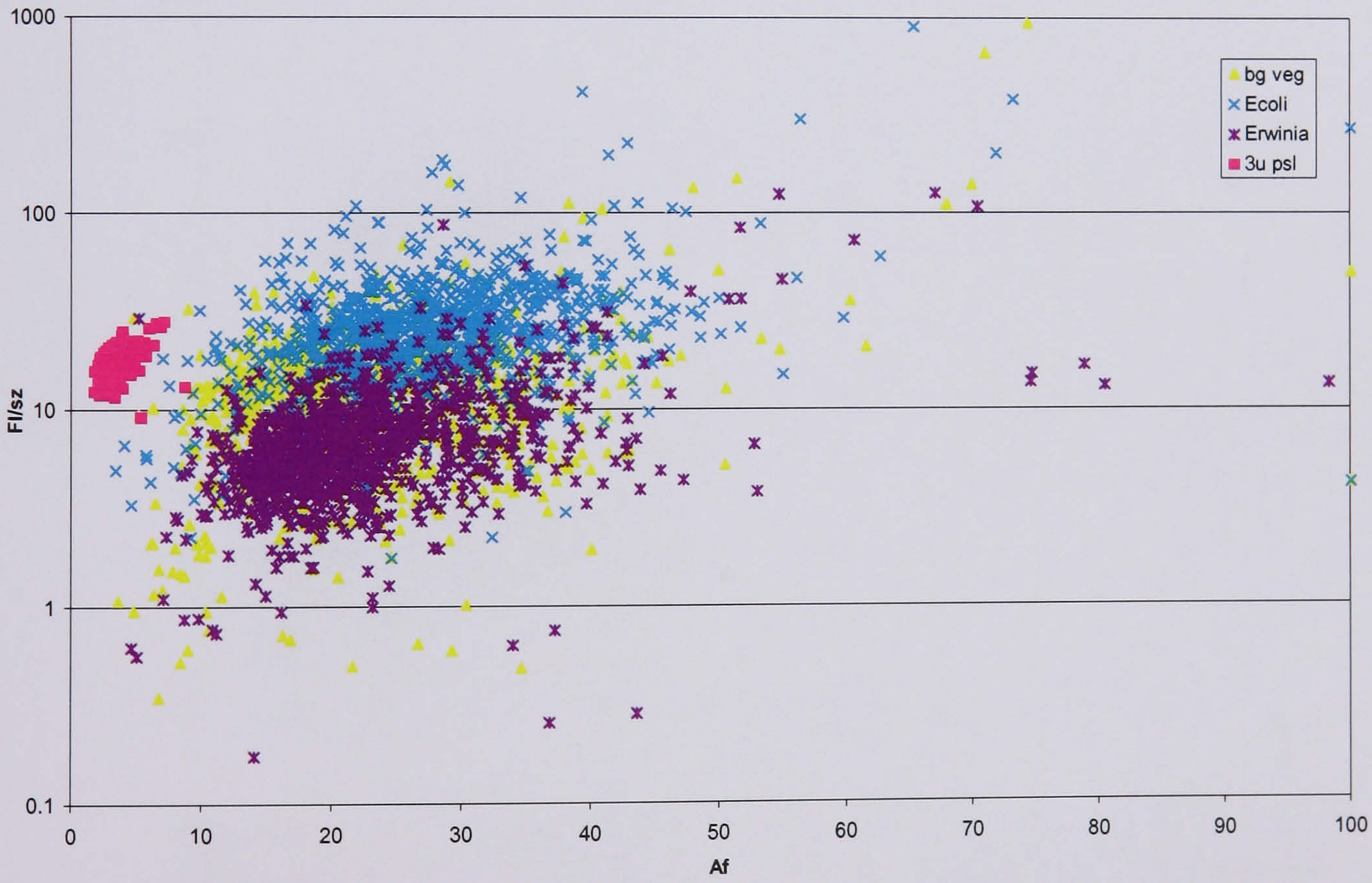



**Figure 7.8 Data from the same bacteria as Figure 7.7on a $Fl^{sz}$ versus Af plot. 3 µm PSl data have also been included for comparison purposes**

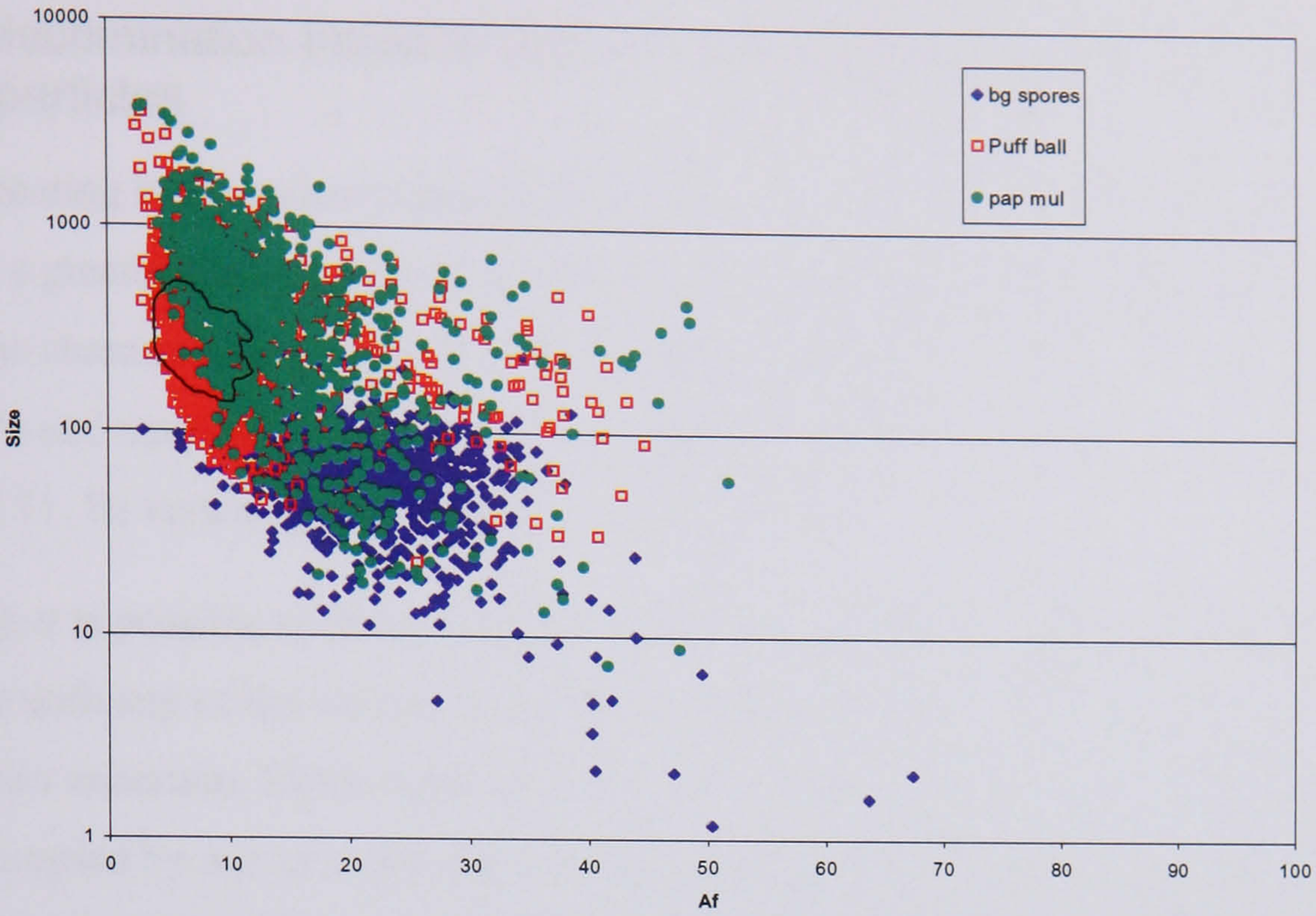



**Figure 7.9 Sz versus *Af* plot of paper mulberry pollen and puff ball spores along with Bg spores. Data from 3 µm NaCl has also been included as a black contour line that includes approximately 90% of the NaCl sample. The individual data points have not been plotted as an aid to clarity**

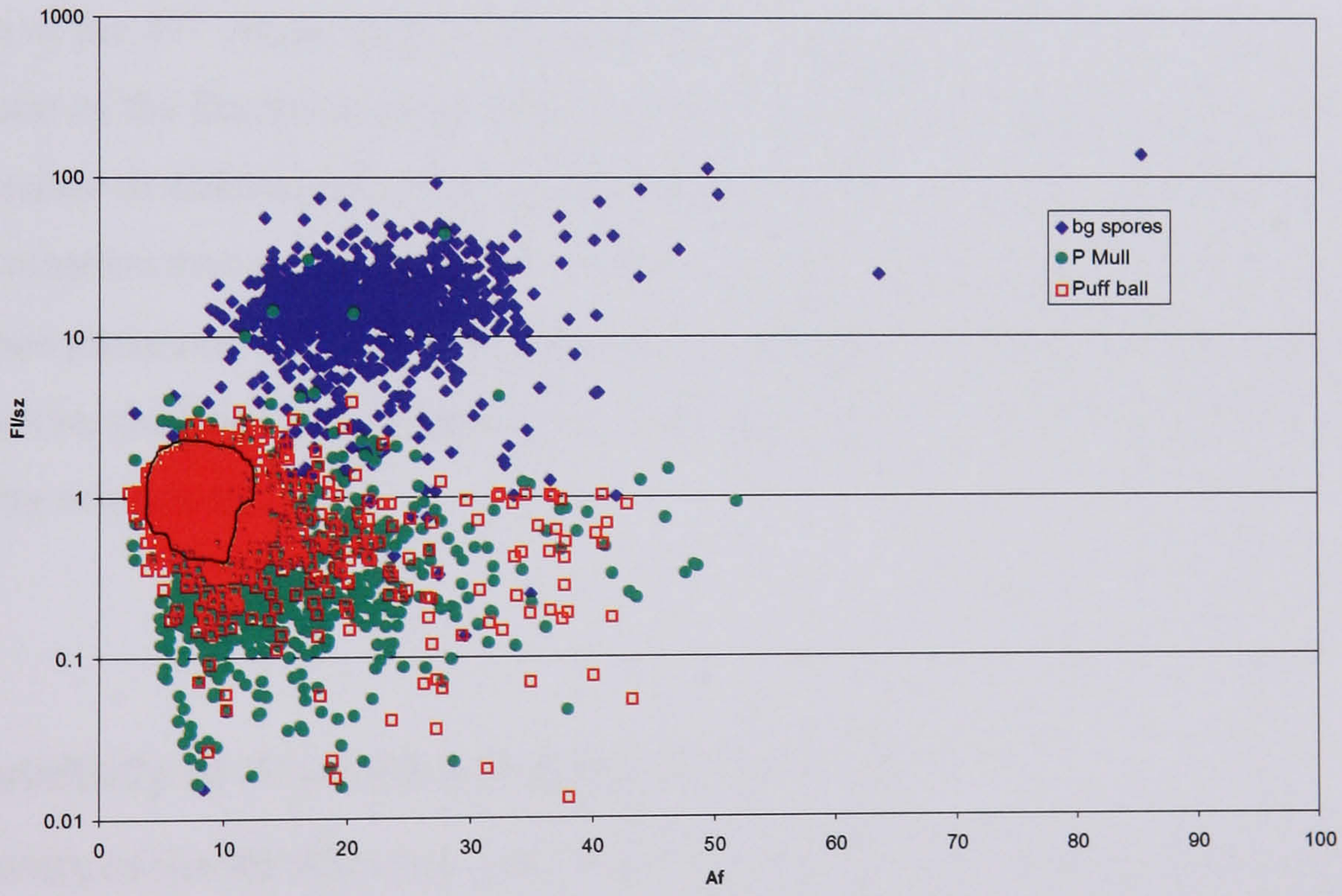



**Figure 7.10 $Fl^{sz}$ versus *Af* plotted for the same samples. The relatively high fluorescence level from the Bg means it can be easily distinguished from the pollen samples. The black contour line again outlines the data from 3 µm NaCl.**

### 7.2.4 Discrimination between bacterial and low aspect ratio non-bacterial particles

Discriminating between low aspect ratio non-bacterial particles and the bacterial particles presents a greater challenge because of the almost complete overlap in their respective size and shape characteristics. Data from two such low aspect ratio ellipsoidal particles, haematite and magnetite, are plotted along with that of a representative bacterial sample in Figure 7.11, Sz versus *Af* and Figure 7.12 $Fl^{sz}$ versus *Af*.

Although it is possible to discriminate between the haematite and magnetite particles, their presence with any of the bacteria makes discrimination relying solely on Sz and *Af* parameters uncertain. Either material overlaps, to a greater or lesser extent, with the sample space occupied by one or more of the bacterial samples. In the case of the haematite, a partial discrimination between it and the bulk of the bacteria may be made. However, the presence of both haematite and magnetite effectively masks the majority of the sample space occupied by the bacteria. Consequently their presence renders reliable discrimination on the basis of these two parameters impossible.

Inspection of the $Fl^{sz}$ versus *Af* plot for the same particles, Figure 7.12 illustrates the effectiveness of the fluorescence parameter in separating the sample types. Using this plot it is now possible to discriminate between the bacteria and the magnetite, greatly improving the discrimination that can be achieved. It should be noted, however, that in this case the fluorescence parameter has made discrimination between the haematite and some of the bacteria worse, the haematite sample space now falling almost identically within that occupied by *erwinia.h*.

### 7.2.5 Sensitivity of *Af* to the sampling environment

The sensitivity of the MPAM-measured shape parameter *Af* to the environment into which the sample is generated is demonstrated in Figure 7.13. In this case, *erwinia.h* was generated over an extended period of time. Initially the relative humidity (RH) in the ballast chamber containing the aerosol was low[c]. As the sample generation continued the RH increased. In the left most plot of Figure 7.13 the initially dry sample has *Af* values typical

c Although no RH or temperature probes were attached to the sample chamber, previous work by the author with a chamber of similar volume which included RH and temperature probes, had a RH of ~15% at the beginning of a sample run. The temperature stayed within ±2° C of the laboratory ambient temperature of 20° C. During typical sample generation times these values did not change appreciably.

of *erwinia.h*; data logged as the chamber becomes more saturated, however, shows the *erwinia.h* appearing to become both more spherical and grow in size, consistent with water adsorption onto the surface and, to a lesser extent, absorption into the cell.

The right hand plot showing *Fl* versus *Af* shows the fluorescence magnitude lying within the same band for both the wet and dry samples. Since the water does not contribute to the fluorescence signal the fluorescence recorded arises solely from the *erwinia.h*. In neither graph is there any indication of the formation of simple water droplets in the wet sample. Such droplets would appear as a low *Af* (<5) and possibly of varying size in the Sz versus *Af plot.* They would also appear as a cluster of low fluorescence spheres on the *Fl* versus *Af* graph.

That the fluorescence has not increased for the wet sample also suggests that few of the individual *erwinia.h* particles have formed clusters through collisions within the sample chamber. While such clusters might be more spherical they would also exhibit an increase in both size, as occurs, and a corresponding increase in fluorescence that does not occur in this case.

In this situation, as the RH in the chamber has increased, the water vapour has started to condense out, with the bacteria acting as nucleation centres. The results confirm that the water vapour is condensing upon the *erwinia.h* particles and not simply forming simple water droplets.

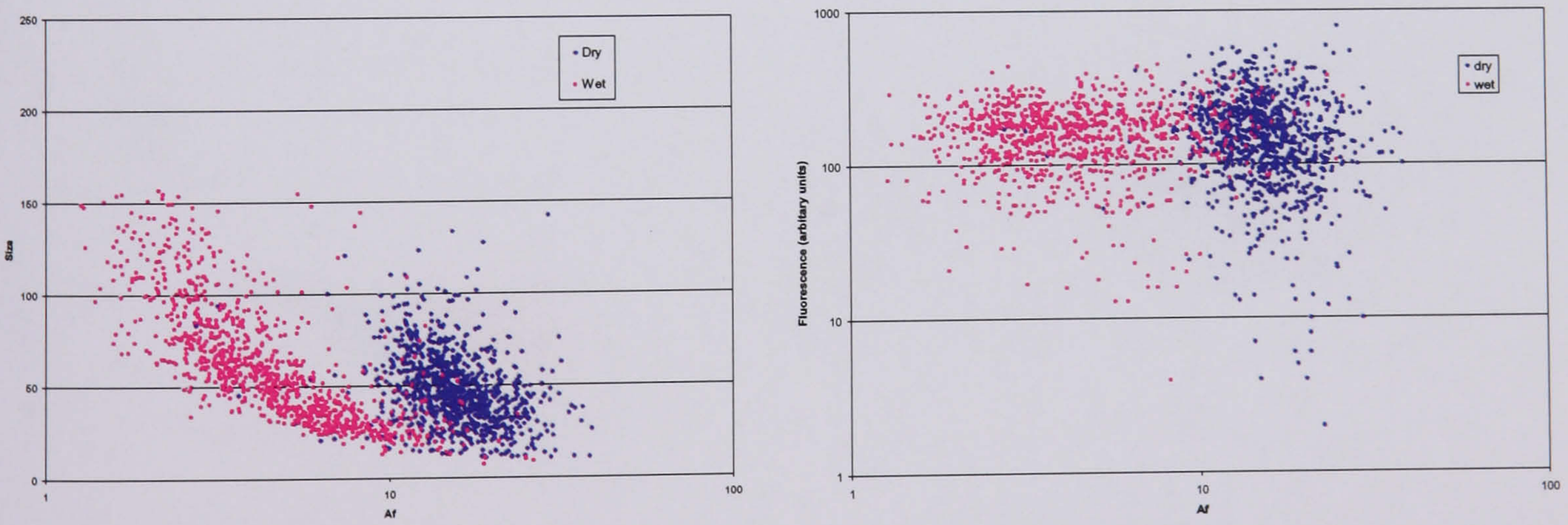



**Figure 7.13. Wet and dry *erwinia.h* plotted showing Sz versus *Af*, left most figure and *Fl* versus *Af*. Note *Af* is shown on a log scale.**

### 7.2.6 Influence of culture media on *Af* and the fluorescence signature

As discussed in Chapter 2, there have been some indications that the culture media used can influence the fluorescence spectra when unwashed samples are used. To investigate this, samples of *erwinia.h* and *e.coli* were placed in the nebuliser suspended in the diluted culture broth ( L broth). Typical results for these unwashed samples are shown in Figure 7.14, along with samples prepared using the normal procedure as described in section 6.1.4.

In the Sz against *Af* graph, the unwashed sample is both more variable in size and has a more extensive range of *Af*s. This increased variance in the sample is probably due to cell debris, broth residue particles and broth coating the exterior of the spores. This occurs when the sample is aerosilized, any liquid rapidly evaporating leaving a media culture residue coating the particle, influencing both its size and shape. The right hand graph shows fluorescence magnitude against *Af*. Note that in this case the fluorescence value has not been normalized for size. Although there is some overlap in the undispersed fluorescence recorded from the washed and unwashed samples, in general the unwashed sample tends to exhibit a higher level of fluorescence. The median value fluorescence recorded from the unwashed sample was 477 compared with 310 from the washed sample. This difference in fluorescence magnitude is similar to that reported by others as has previously been discussed in Section 2.2.

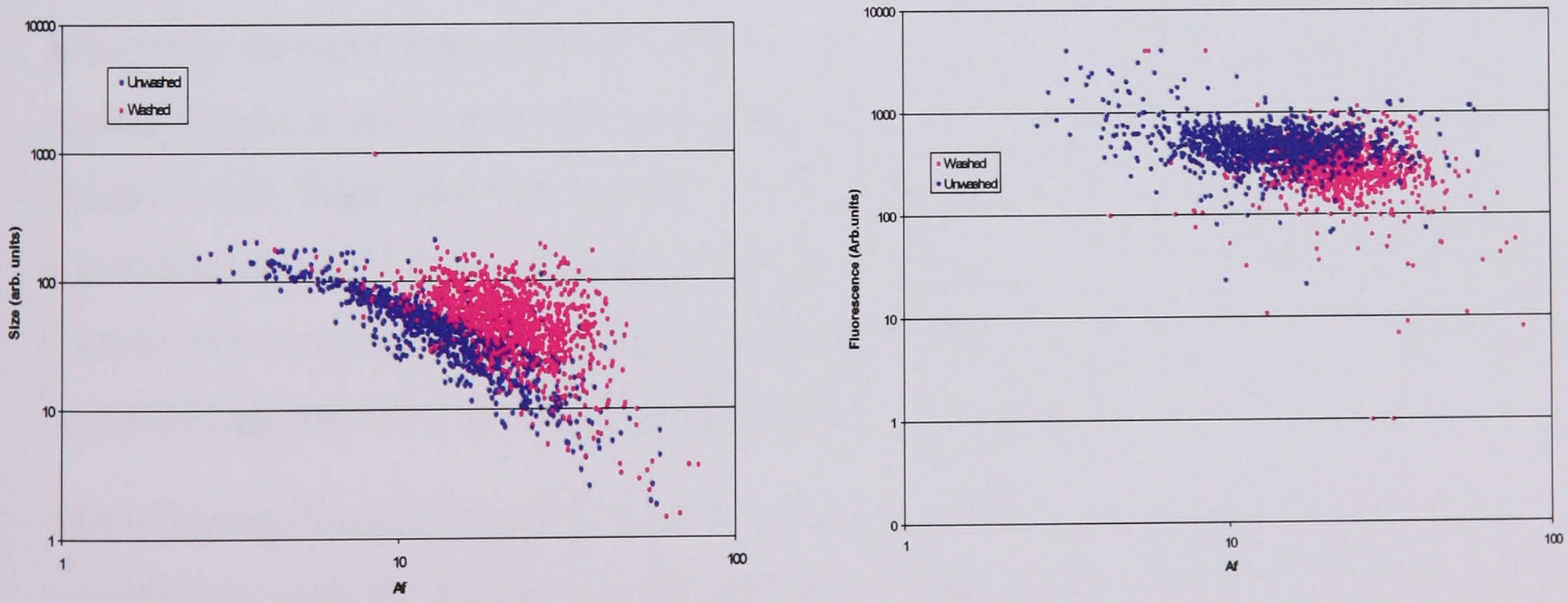



**Figure 7.14 Size versus *Af* for unwashed and washed sample of *erwinia.h* along with a plot of fluorescence against *Af* for the same sample. Note that in this case the fluorescence measure has not been normailized for size.**

## 7.3 Discussion

Despite the simplicity of the data processing required, the combination of the *Sz* versus *Af* and $Ff^{sz}$ versus *Af* plots makes discrimination between a wide range of particle types possible. The multi-parameter approach facilitates discrimination between particles of common shape that exhibit different levels of fluorescence. Likewise, there are cases where the levels of fluorescence recorded do not aid discrimination but where the shape parameter allows discrimination between these classes of particle.

However there are also cases where neither plot can achieve sufficient separation to enable discrimination between particle classes. For example, in the plots of haematite, magnetite and the various bacteria, there is an overlap between the sample space occupied in both the Sz versus *Af* plot and the $Ff^{sz}$ versus *Af*, making discrimination between these particles impossible. While the simplicity of these methods of analysis means they can be easily applied, they are ultimately limited by being unable to bring all aspects of the data to bear simultaneously in carrying out the analysis. The degree of useful classification that can be had with such simple methods suggests that more advanced methods capable of mobilizing all the data in a single analytical process superior levels of classification would be possible from the MPAM data.

Moreover, it is recognised that the Asymmetry Parameter *Af,* whilst easy to compute and convenient to use, does not take advantage of all of the relevant information contained within the azimuthal scattered light pattern. *Af* involves a root-mean-square calculation and is thus simply a measure of the variation in scattered light intensity about the azimuth; potentially far more information is available. For example, a hypothetical particle that scattered light in such a way that alternate pixels in the 24-pixel outer-ring of the HPD detector were 'high' and 'low' intensity, would give the same value of *Af* as a particle which produced 12 'high' pixels in a contiguous block and 12 'low' pixels in the remainder. Clearly, the two particles would have significantly different morphologies to produce these scattered light patterns, yet the *Af* parameter fails to recognise this.

The following Chapter therefore examines the use of several more sophisticated probabilistic analyses for particle classification. These utilise all of the discrete elements within the spatial scattering and fluorescence data to fully realise the potential of the MPAM approach.

# 8 Data analysis using probabilistic methods

## 8.1 Introduction

Ideally, the data analysis method used must be able to interpret the data generated by the MPAM instrument in such a way that the particle in question can be classified into one of a number of predetermined classes. Additionally, the method used must be able to run sufficiently fast as to allow particle classification in real-time at potential data rates up to several thousand particles per second.

A wide variety of data analysis methods could be applied, from traditional statistical methods to the more sophisticated approaches employed by neural networks or genetic algorithms. One further approach would be the application of expert systems methods that may combine some aspects of statistical data processing along with a rule-based decision making tree. Such rules being derived from the analysis by (human) experts of similar data sets.

To investigate all of these methods would be a major undertaking in its own right. Consequently, the methods investigated here will be limited to taking a classical statistical approach to the problem of data analysis. Four methods have been employed: Normal density distribution, Linear discriminant, K-nearest neighbours, and Fuzzy-K nearest neighbours. All attempt to classify the unknown particle by matching it to a set of predetermined particle classes, each derived from 'exemplar' data. The particle is ultimately ascribed to the class for which the probability of correct classification is highest. In effect, the exemplar data is used to 'train' the analysis process and thus render it capable of recognising and classifying unknown particles into one of the classes for which it had been trained.

Software has been written for each classification algorithms to allow their comparative assessment in terms of both particle classification efficiency and execution speed. A brief description of the four algorithms is provided below; a more detailed mathematical development can be found in the relevant literature [80, 81, 82].

## 8.2 Classification algorithms

### 8.2.1 Normal Density Distribution (Norm Dist)

This program classifies the input data into one of several classes using discriminant functions based on the normal distribution. For each class, the mean vector, inverse covariance matrix, and the log of the determinant of the covariance matrix must be

computed. The multidimensional sample space is divided into decision regions which can be hyperplanes, hyperspheres, hyperellipsoids, hyperparaboloids or hyperhyperboloids depending on the calculated parameters of all the classes. In addition, the regions can be discontinuous. While this technique is potentially powerful it has the disadvantage of the requiring the multidimensional data elements to be normally distributed. The Norm Dist parameters for each class can be calculated with a Matlab™ procedure using exemplar data from each class (see later for discussion of generation of exemplar data). The program works by taking the data for each particle and using the previously calculated parameter data to compute the value of the discriminant function for each classification class. The class with the largest discriminate function is then assigned as the most probable class. The program output is a text file where each row contains the classification result from a single particle. The first column of the file is the actual class number assigned to the data, which is input to the program. The second column contains the class number as computed by the program. For this method and the three described in the following sub-sections, data are output in a common format to make comparison straightforward.

### 8.2.2 Linear Discriminant (Lin Disc)

This method classifies particle data using Linear Discriminant functions. The effect of this type of classification is to partition the multidimensional feature space into class regions separated by hyperplanes. Unlike the normal distribution discriminant functions, the decision boundaries are never curved and the regions are always continuous, consequently two separate regions cannot belong to the same class. As a result, the method has some limitations where the data requires that the decision boundary is curved, or the distribution may contain sub-classes necessitating non-contiguous sample regions. However, unlike the normal distribution discriminant functions, the method has the advantage that there is no requirement for the data to be normally distributed.

There are several ways of determining the values of the discriminant coefficients, not all of which are suitable for multi-class non-separated data of the type produced by the MPAM. The method that is used here is a minimum-squared-error method. This method attempts to position the hyperplanes to minimise some specified error criterion. A Matlab™ function to compute the coefficients from the exemplar data has been used. The program computes the discriminant function value for each data input using the discriminate coefficients previously calculated and the class with the largest value is determined as the most likely class for the particle. The program output is as for the Norm Dist program.

### 8.2.3 K-nearest neighbours (KNN)

This program uses the k-nearest neighbours algorithm which has been extensively used as a powerful non-parametric method of classifying data. This technique computes the multidimensional Euclidean distance from the input data to the k-nearest values in the exemplar data for all classes. The modal class, i.e. the most frequently occurring class, in the k exemplars is then selected as the most probable estimate of the actual particle class. The effect of the process is to locate a hypersphere in multidimensional space centred on each data point. Its radius is then increased until k exemplars are encompassed. The number of exemplars per class divided by k is then taken to be an estimate of the probability of the class at the data point. The class occurring most often will therefore have the highest probability. Deciding k, the number of neighbours to include, is determined by the need to balance several factors. In theory the greater the number of neighbours included the better the accuracy of the classification. This gives rise to two problems however. Firstly, checking large numbers of high dimension samples is very time consuming. Secondly, the number of exemplar samples must be large compared to k, such that the actual number of sample data available, distinct from the test data, may constrain the maximum value that can be assigned to k. This arises because of the requirement that (ideally) the exemplar data occupies the entire sample space for that class. This ensures that the probability estimates for all the classes at any point in sample space are close to the true but unknown values. In practice then, the need to process the data at speed and the limited size of the exemplar data file imposes a practical limit on the value chosen for k.

The KNN program operates by reading the data to be classified one particle at a time and computing the distance from this point to all the data in the sample file. The sample file contains all the exemplar data from each particle class concatenated into a single data file. The nearest k are then processed to determine the modal class. The processing therefore requires one scalar product computation per example and k comparisons for each input. The speed of the computation thus depends on the dimension of the data, the number of exemplars required to give a good classification and k, the number of neighbours. Unlike the calculation of the discriminant functions for the Lin Disc and Norm Dist, this processing cannot take place off-line in advance of the data processing run. In its present form, the KNN could be too slow to operate in a real-time application.

### 8.2.4 Fuzzy-K nearest neighbours (Fuzzy-K)

The Size versus *Af* and $Fl^{sz}$ versus *Af* plots previously examined suggest that for some of the data there may be a genuine overlap in the sample space occupied by different materials. This reflects real similarities in either the morphology of the particles (in the case of the

shape data) or their chemical composition (in the case of the fluorescence parameter). There may then in reality be an ill-defined i.e. "fuzzy" overlap in the sample space occupied by two or more particle types. By allowing data to be allocated a weighted membership of more than one class, the Fuzzy-k approach attempts to "second guess" the classification when there is ambiguity in the distribution of the data.

This program is a fuzzy logic adaptation of the k nearest neighbours classification procedure. In this approach each of the samples in the exemplar class is allocated a degree of membership to each of the other classes. So that, in the simplified case of a two-class exemplar sample set consisting of say NaCl and haematite, the subsequent classification of a NaCl particle could result in 0.94 membership of the NaCl class and 0.06 membership of the haematite class as a consequence of the overlap in their respective distributions.

In operation, for the test data, the k-nearest neighbours are found as described for KNN above. The degree of membership of each test particle is then found by taking the sum of the inverse distances weighted by a previously calculated example membership term. The test particle is then allocated to the class which has the maximum membership.

The membership term is computed from the training data using a Matlab function which, for each sample, finds the nearest n samples and then allocates memberships based on the number of samples occurring in each class divided by n. In this case n has been set equal to k.

The technique requires the same amount of computation as the KNN method with a small additional overhead arising from the membership allocation process.

## 8.3 Selection of exemplar data

As discussed above, each of the various classification algorithms requires an exemplar dataset representing all of the possible particle classes. Within each dataset, it is important that the exemplar for a given particle type accurately represents that aerosol. The left-hand column of Figure 8.1 illustrates five examples of MPAM raw data recorded from 3 μm oleic acid droplets as they would be displayed on the MPAM instrument. It can be seen that the first four samples have a high degree of similarity, while the final sample is very different. Since the scattering profile expected from spherical particles is known it is a simple process to select valid examples for inclusion in the training file. Similarly with the data from the gypsum fibres (middle column in Figure 8.1) the characteristic scattering profiles make the selection of exemplar data a straightforward and relatively unambiguous procedure. For the

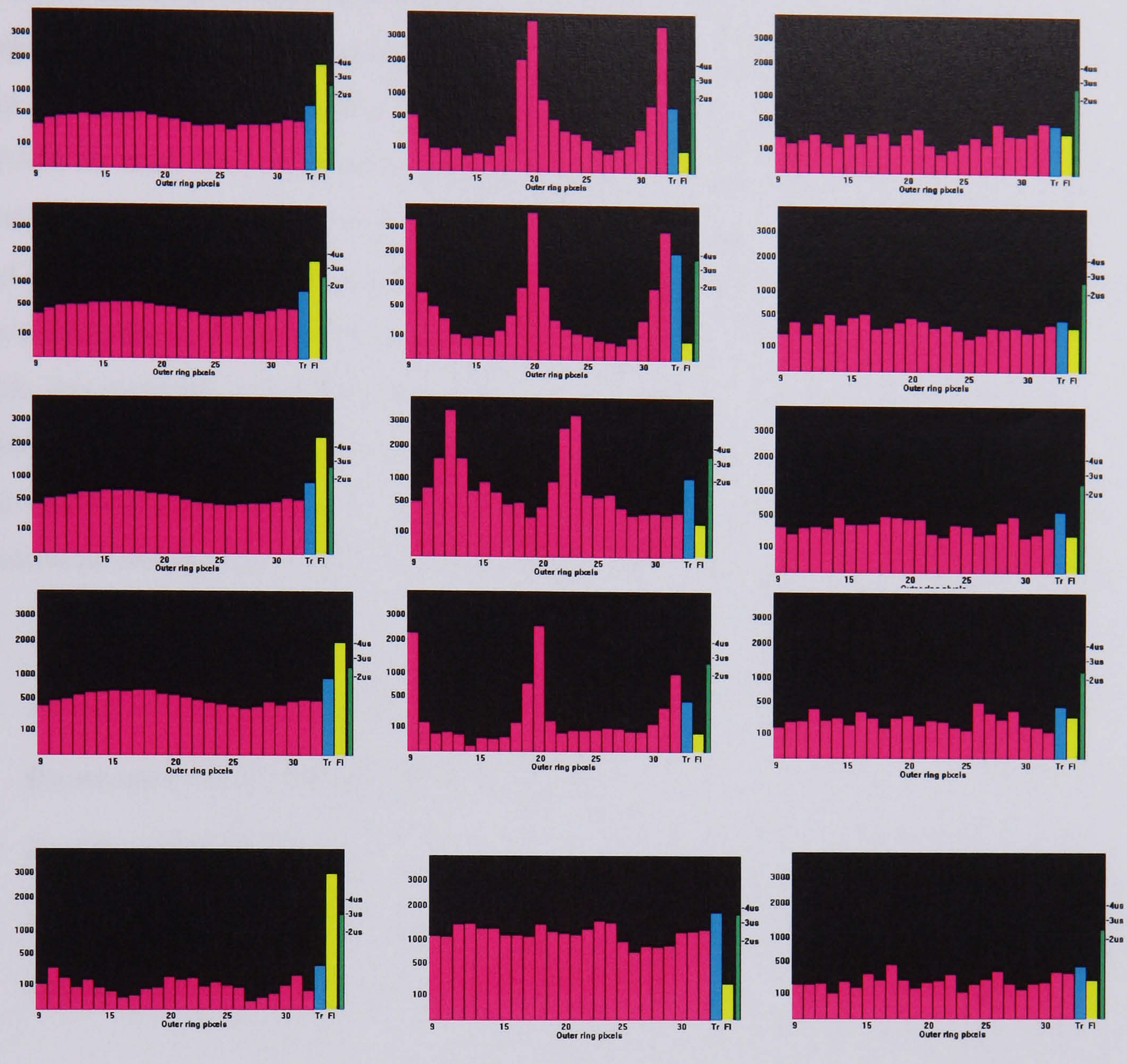



**Figure 8.1 MPAM output as seen on the MPAM data display screen. The left most column is from a 3 μm oleic acid data file, the middle column is taken from a gypsum file and the right hand column is from an NaCl data file. In the case of oleic acid and gypsum, the first four images have been chosen to represent typical examples. The fifth example represents the type of data that would be excluded from the exemplar training files, it being probably caused by a contaminant/background particle in the aerosol.**

majority of the materials the MPAM data is more complex as illustrated in the right hand column of Figure 8.1 for NaCl.

Account must also be taken of the fluorescence values recorded in the selection process. It seems reasonable to expect the level to be similar from particle to particle for a specific material, subject to particle size. With extended periods of observation of the data it becomes possible to make reasonable judgements as to the typical nature of the data from any individual particle of a known source, and therefore allocate it to the exemplar data file or not.

Returning to the plots of Chapter 7, it is clear that for many of the samples there is considerable overlap in scattering profiles and or fluorescence levels. (If this was not the

case it would not be necessary to attempt more advanced methods of data analysis). Consequently, the selection of typical sample data for these aerosols becomes more difficult. Ideally, the exemplar data should occupy the entire sample space as recorded by the MPAM for each class of particle. To try to achieve this, a selection process was established whereby all data were included in the exemplar file unless they were ruled out by a set of pre-defined criteria for that class of particle. The exclusion criteria were based on the known attributes of the individual aerosol being sampled. Using haematite as an example, this material is roughly ellipsoidal in habit, of a small size and not fluorescent. Therefore, if the scattering from an individual particle within the aerosol showed a highly spherical scattering profile it could be excluded as being unlikely to be haematite but rather some contaminant particle. Taking this approach, the exclusion criteria for haematite might be:

- Average scattering magnitude over 1000 i.e. large particle
- fluorescence level > 350 i.e. exhibits fluorescence
- 2μs< time-of-flight <4μs
- exhibits spherical scattering pattern
- exhibits two distinct peaks i.e. fibre scattering.

Any one of these conditions would be sufficient to exclude the sample from the exemplar file. For the controlled conditions under which the aerosol samples were generated it would be expected that most data would, in fact, be valid, though the presence of contaminant / background particles in the aerosol could not be ignored.

### 8.3.1 Selection software

Before the data analysis can be applied to the data, the format of the file has to be made suitable for handling by the data analysis programs as well as generating the training and test data files. This process is carried out by software written to make the selection of the exemplar, or 'training', data.

#### 8.3.1.1 MPAM data viewer

The raw data as logged by the MPAM is saved in a binary format that cannot be easily read by other applications. A Microsoft Windows program has been written to read the logged data file and display the data in a bar chart format on a particle-by-particle basis. The software allows the operator to allocate each sample to the exemplar (training) files

according to the pre-determined selection criteria. The exemplar file size has been set at 150 particles for each particle type. When this number have been selected, seven text-formatted output files (see below) are automatically saved. Note that nothing is written to the original binary data file which is opened for reading only.

#### 8.3.1.2 MPAM data viewer output files

A brief description of the data format of each of the output files is provided below under the heading of the default file name created by the viewer software.

1. data.trn

The file contains data from the selected images only. It has a file header block listing the UV power level as measured by the MPAM at the time of logging, the PMT gain voltage setting and the minimum and maximum TOF set on the MPAM. Each subsequent row of the file then contains data from the 24 pixels of the HPD outer-ring followed by the trigger channel, the fluorescence channel and the time-of-flight recorded for that particle, making 27 data values in total per particle.

2. data.tst

This file contains data from all those particles that have not been included in the training data set or deleted. In all other respects, its format is the same as the data.trn file.

3. trn_data.sca

Contains only selected images data. The header block has been stripped from the file. Each row consists of a unique class number that has been assigned to each aerosol material, as listed in Table 8.1. The class number is followed by 24 pixel data values from the HPD detector normalised for size on a particle-by-particle basis. No trigger channel, fluorescence data or TOF data are included.

4. trn_data.fl

Each row in the file corresponds with a row in the trn_data.sca file. It contains three values for each particle. The first is the size term found by summing the data from the 24 elements of the detector. The second is the fluorescence value. The software reads from the header block (in data.trn) the PMT gain voltage that was used at the time of logging. A gain correction term from a look up table is then applied to give the effective value that would have been recorded if the data had been logged at the MPAM's nominal PMT voltage of

700v. The third and last term in the row is the particle-by-particle time-of-flight recorded at the time of logging.

5. tst_data.sca

This file has the same format as the file trn_data .sca file but consist exclusively of test data.

6. tst_data.fl

This file has the same format as the file trn_data .fl file but consist exclusively of test data.

7. selected.txt

This is a compete list of all selected and deleted rows. It makes it easy to re-examine the original binary data file in the MPAM viewer to view both selected and deleted data samples.

## 8.4 Data combinations tested

The data format of the MPAM is unique in that it contains data relating to three parameters: particle shape, particle size and the particle's intrinsic fluorescence. Consequently no systematic study assessing the relative importance of each of the three parameters (shape, size & fluorescence) on the classification process has been possible prior to this work. Five different data combinations were thus selected to enable a detailed examination of the relative performance of these parameters for particle classification. They are summarised below.

1. Combination 1: classification based on **Shape** alone.

   Combination 1 was based on exemplar and test files (ie: data for which particle classification was sought) which included only light scatter data: for each particle the data format comprised 24 azimuthal scattering values normalised for particle size. Since no particle size or fluorescence data were included, this approach gives an insight into the degree to which particle classification can be achieved on the basis of shape alone .

2. Combination 2: classification based on **Shape + Size**

   This is as combination 1, but where the 24 azimuthal light scattering values were supplemented by a single particle size value derived from the mean of all scattering values.

3. Combination 3: classification based on **Shape + Fluorescence**

   This is as combination 1, but includes the particle fluorescence data as described for the trn_data.fl file format. No particle size data were included.

4. Combination 4: classification based on **Shape + Size + Fluorescence**

   All available data incorporated into the classification process.

5. Combination 5: classification based on **Size + Fluorescence**

   This final format was examined so that comparison could be made with proprietary aerosol analysis systems such as UV-APS (TSI Inc.) that use both a size and fluorescence parameter in their particle classification process. By replicating this data combination here and comparing the results to those of Combination 4 above, it would be possible to assess the importance of incorporating particle shape into the classification process

## 8.4.1 Sample grouping and classes

The experimental data previously logged as described in Chapter 5 has been assigned to both a **group** and a **class**. The individual particle classes have been grouped into five broad groups: Spherical, Non-spherical, Fibres, Bacteria, and Pollen. The group assignment does not solely reflect a physical attribute of the sample, for example Bg could legitimately be allocated to the ellipsoidal group on the basis of its shape. Rather it reflects the basic groups into which it was hoped to be able to classify the aerosols. Each particle type has also been allocated a distinct class number that is appended to the data as described for the trn_data.sca file format. The group and class information is summarised in Table 8.1.

| Group | Class # | Material | Size (μm) | Aerosol Generation Method |
|---|---|---|---|---|
| Spherical | 1 | Oleic | 1.0 | VOAG |
| | 1 | Oleic | 3.0 | VOAG |
| | 2 | PSL | 1.0 | Nebulizer |
| | 2 | PSL (fluorescent | 1.7 | Nebulizer |
| | 3 | Water droplets | 1-10 | Nebulizer (hand held) |
| | 4 | Silica | 3-4 | Dry compressed air |
| Non-spherical | 5 | Haematite | 1-3 | Dry compressed air |
| | 6 | kaolin | 0.5-5 | Dry compressed air |
| | 7 | Magnetite | 1.3 | Dry compressed air |
| | 8 | NaCl | 2 | VOAG (via solution) |
| | 8 | NaCl | 3 | VOAG (via solution) |
| Fibres | 9 | Gypsum | 0.5-10 | Dry compressed air |
| | 10 | Caffeine | 1-5 | Thermal sublimation/condensation. |
| Bacteria | 11 | BG spores | 1-2 | Nebulizer |
| | 12 | BG clusters | 2-5 | Nebulizer |
| | 13 | BG vegetative | 1-2 | Nebulizer |
| | 14 | E coli (washed) | 1 | Nebulizer |
| | 15 | E coli (unwashed) | 1 | Nebulizer |
| | 16 | *Erwinia* (washed) | 1 | Nebulizer |
| | 17 | *Erwinia* (unwashed) | 1 | Nebulizer |
| Pollen | 18 | Paper Mulberry | 2-4 | Dry compressed air |
| | 19 | Puff Ball pollen | | Dry compressed air |

**Table 8.1. Group and class classification assigned to the sample particles.**

## 8.4.2 Classification procedure

For each of the input data combinations described a file from each aerosol class was processed by the four data processing algorithms. The processing of the appropriate exemplar training file, and calculation of any discriminant functions required etc., had already been carried out. In the case of the KNN algorithm some limited initial trials were carried out with k the number of nearest neighbours set at 5, 10 and 15 respectively. Inspection of these results suggested that k = 5 resulted in a poor classification while k = 15 did not produce significantly better results than achieved with K= 10, but with the disadvantage of requiring substantial extra processing time. It was decided to use k =10 for the bulk of the data for this and the Fuzzy-k method.

The output from all of the four methods has been further processed to yield a classification 'accuracy' output in percentage terms. Ideally, test data from a given particle type would be

classified 100% into its own **class**, while a successful if not ideal classification would allocate the test particle to the correct **group**.

For example, if the test aerosol was of BG spores, a PERFECT classification accuracy for the MPAM data would be reflected in a 100% figure in the 'BG spores' class (see Table 8.1 again). A more realistic result might be, say, 70% in the 'BG spores' class and 20% in the 'BG clusters' class, with the remaining 10% spread across the remaining five classes (ie: all data still correctly classified with respect to the Group – 'Bacteria'). A poor classification accuracy result might be 40% in the 'BG spores' class, 34% in the 'BG clusters' class, and 16% in the 'Haematite' class, with the balance again spread across the other groups - ie: the classification process has not even managed to identify the aerosol particles as all belonging to the same Group, let alone Class of particle type.

In most real circumstances, any test aerosol data would normally be allocated into a variety of output classes or groups depending on several factors such as the purity of the aerosol, the natural variability of the aerosol, as well as the efficiency of the algorithm used.

## 8.5 Numerical results classification methodologies

The use of nineteen different test aerosols, five combinations of size, shape, and fluorescence parameters, and four different classification methods for each parameter combination gave rise to a very large body of classification results. Moreover, there were frequently specific observations to be made regarding the performance of a particular classification algorithm with a particular class of particle, and corresponding postulates offered in an attempt to explain the observations. These detailed numerical data tables and their accompanying observations were considered too large to be incorporated within the thesis at this point, and instead they are included in full in Appendix 1. However, it was still necessary to find a method by which the extensive body of results contained within Appendix 1 could be distilled down further so as to reveal simply and clearly the overall performance of each classification algorithm. This has been achieved by the use of a graphical approach, as is described in the remainder of the Chapter.

## 8.6 Graphical presentation and interpretation of results

Different graphical methods were investigated in an attempt to give an overview of the classification performance landscape that had been created for each analysis approach. Ultimately, an approach based on contour plots was found to be most effective, and this is detailed below.

Simplification of the data presentation has been achieved by restricting the output classes plotted to the five broader groups: Spherical, Non-spherical, Fibres, Bacteria and Pollen. A further simplification has been made through including only a single class from each aerosol. So, for example, 3 μm oleic acid droplet data is omitted as it presents no additional classification challenge to that presented by 1 μm oleic acid droplets. Similarly, the 1.7 μm PSL aerosol data has been omitted as the fluorescence values it gave were orders of magnitude higher than those recorded for natural un-doped materials, and to incorporate the doped material data would have disproportionately biased the classification process. Finally, the 3 μm NaCl aerosol has been omitted from the classification process as examination of the raw data indicated the aerosol was of solution droplets rather than dried crystals. All of these omitted datasets are included in Appendix 1 for examination if desired.

Figure 8.5 shows how a the contour plot model would appear with a 100% correct classification of all aerosol particles. This is the 'target' response we are looking for. The plot has been produced using SigmaPlot™ graphing software from Excel™ spreadsheet data.

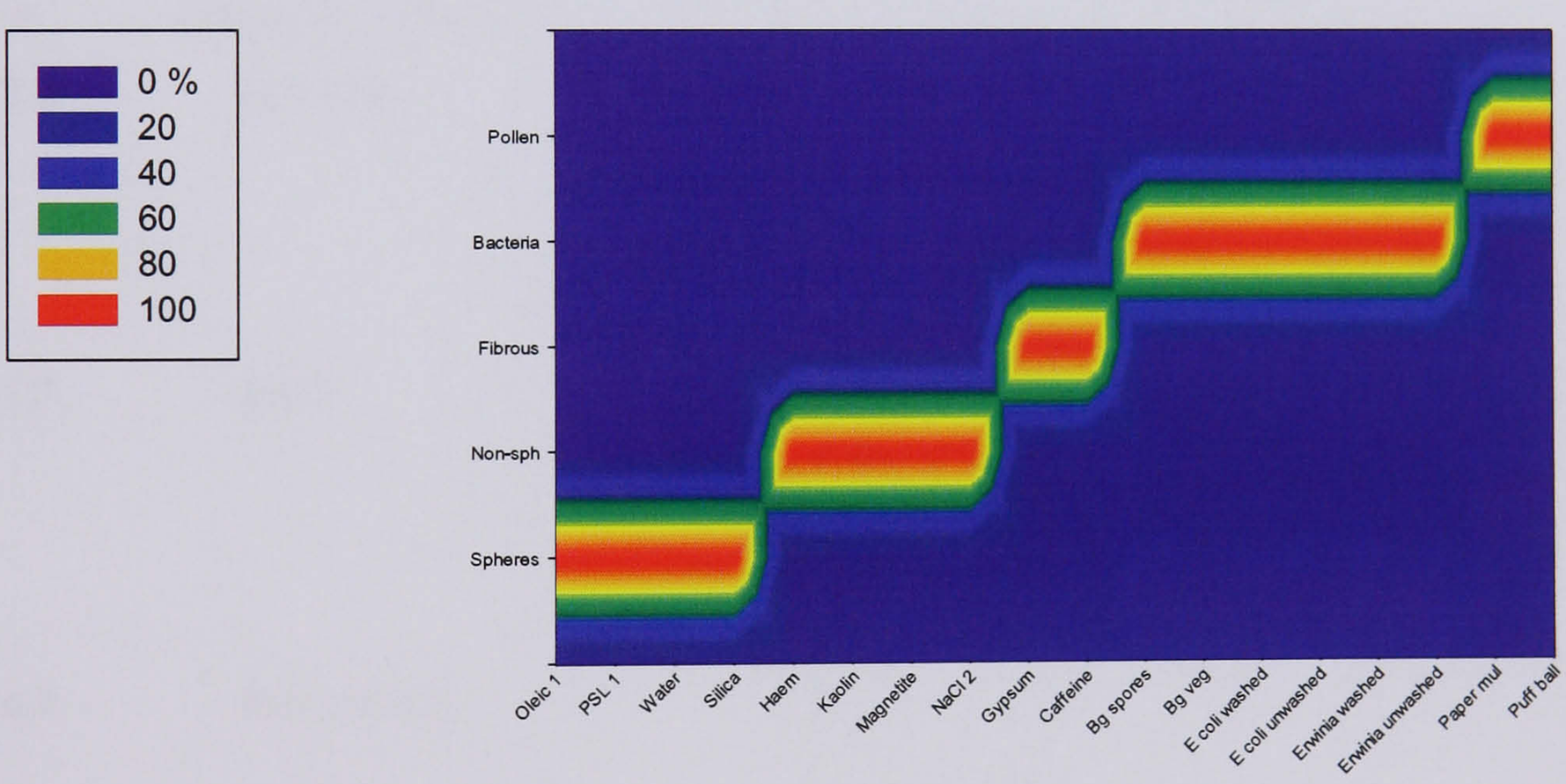



**Figure 8.5 Illustration of a 'perfect' result in which all aerosol types are correctly classified into their respective Groups.**

An artefact of the contour graphing process is the appearance of the full spectrum of colours in Figure 8.5. With 100% correct classification in each class, one might expect Figure 8.5 to show a series of short horizontal *red* lines (ie: 100% accuracy achieved) rather than the horizontal spectral bands as shown. The colours result from the smoothing algorithm used by the graphing software which doesn't allow abrupt step-wise changes in the data. Whilst

this might seem to be misleading, it should be remembered that the purpose of these contour plots is to give an arbitrary visual representation of the overall performance of a specific classification methodology. Since every graph will be subject to the same 'smoothing' function, the visual comparison of two or more graphs will remain a valid approach to assessing the relative performance of the classification methodologies.

A brief descriptive assessment of the performance of each of the four classification algorithms when processed using each of the five data parameter sets is provided below, each assessment being followed on a separate page by the relevant graphical Figure, 8.6 to 8.10, respectively. For convenience, Figure 8.11 provides a summary overview of ALL graphical classification outputs.

Finally, Figure 8.12 show a direct comparison of the best performing Algorithm-Parameter combination to the 'Perfect' classification performance shown in Figure 8.5 above.

The graphical Figures are thus arranged as :

| Figure # | Classif'n Algorithm | Parameters used in Classification |
|---|---|---|
| Figure 8.6 | Lin. Disc. | Shape |
| | | Shape and Size |
| | | Shape and Fluorescence |
| | | Shape, Size, and Fluorescence |
| Figure 8.7 | KNN | Shape |
| | | Shape and Size |
| | | Shape and Fluorescence |
| | | Shape, Size, and Fluorescence |
| Figure 8.8 | Fuzzy-KNN | Shape |
| | | Shape and Size |
| | | Shape and Fluorescence |
| | | Shape, Size, and Fluorescence |
| Figure 8.9 | Norm Dist'n | Shape |
| | | Shape and Size |
| | | Shape and Fluorescence |
| | | Shape, Size, and Fluorescence |
| Figure 8.10 | Norm. Dist'n. | Size and Fluorescence |
| Figure 8.11 Summary Table | All classification algorithms | All combinations of data |
| Figure 8.12 | Comparison of 'Best' with 'Perfect' classification performance. | |

## 8.7 Results of classification processes

### 8.7.1 Input data format: Shape Only

See Figure 8.6 overleaf.

Algorithm: **Linear Discriminant**

This algorithm has performed poorly for virtually all particle types except larger spheres and the ellipsoidal particles (magnetite, haematite and kaolin). Given the more complex morphology of the ellipsoidal materials there is a general decline in the percentages classified correctly when compared to the spherical results. For this group Lin Disc performed best overall, although all the algorithms achieved similar levels of classification. With only Shape as the discriminating parameter, the pollens have been mis-classified as Spheres and Non-spheres (see inset in Fig. 8.6) as might be expected from their varying spheroidal shapes.

Algorithm: **KNN**

For the spheres better than 95% of the sample is correctly classified. with exception of the 1μm PSL. This reflects the poor signal-to-noise present in the 1 μm PSL scattering data, and the fact that normalising the very low scattering values accentuate the variability across the 24 azimuthal values. The Bacteria are classified with a higher accuracy than when using Lin Disc. Pollen are mis-classified, largely into the Non-sphere group.

Algorithm: **Fuzzy-K**

Similar to KNN above.

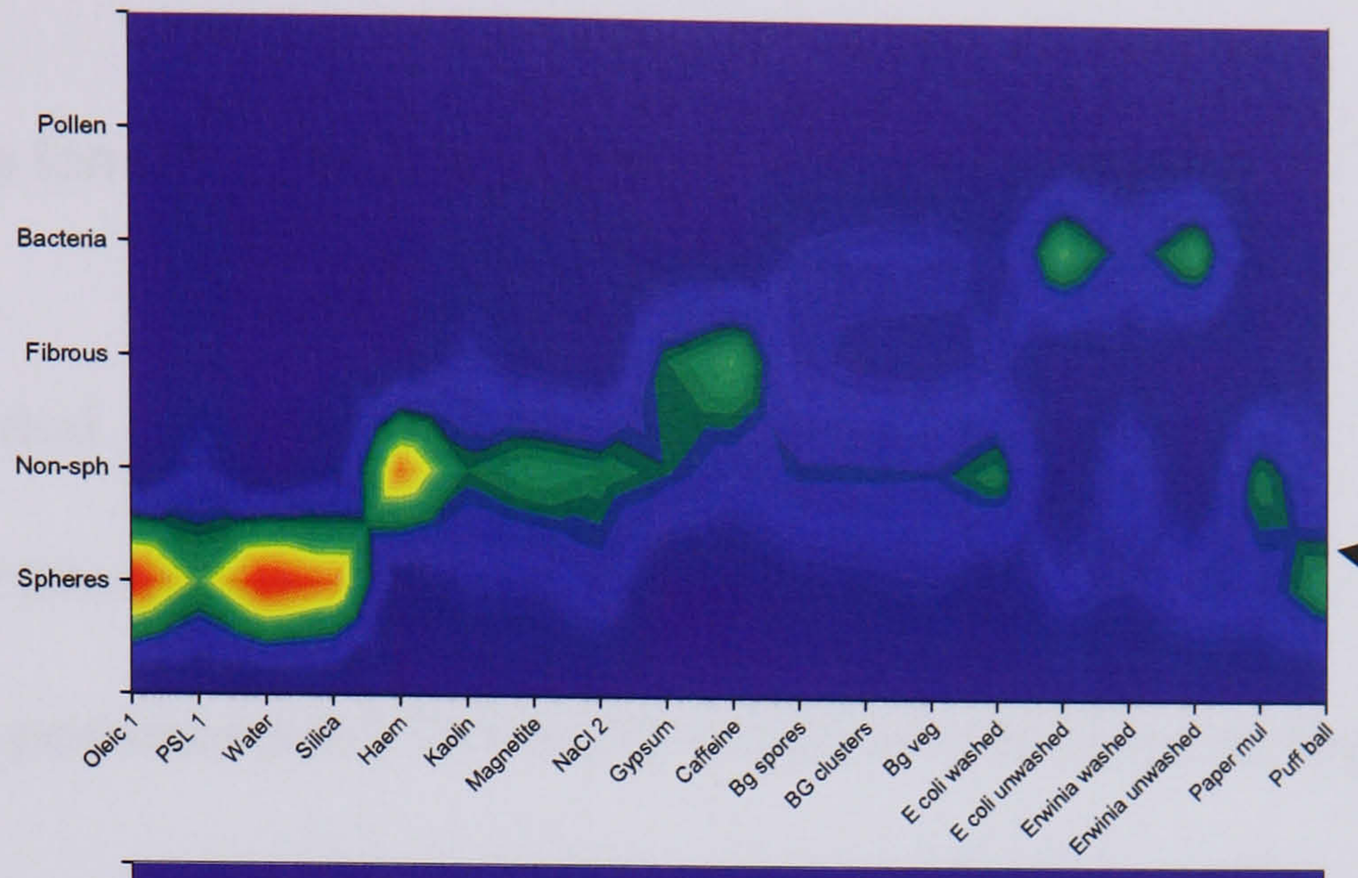
Algorithm: **Normal Distribution**

Performs nearly as well as KNN in the classification into the Sphere group. Although the 1 μm oleic has been relatively successfully classified as Spheres, the noisier1μm PSL has again has been poorly classified across Sphere, Non-sphere, and Pollen groups. Water and silica are correctly classified as spheres as would be expected. Haematite and magnetite are reasonably well classified as Non-spheres, but Kaolin is spread across several classes, a result of the wide morphological variation (from irregular cubic, to flake-like and fibre-like) present in the kaolin particles. The NaCl (2 μm) aerosol has been poorly classified with a significant proportion being incorrectly classed as Bacteria. Clearly this would be a problem

for an instrument relying on particle shape analysis alone. Most of the bacterial based aerosols have been reasonably well classified as Bacteria, although the classification efficiency falls to around 60% for *E. coli* (washed). Paper mulberry and Puff ball pollens have been very poorly classified. Mis-classification into all 4 groupings is present, although a preponderance falls into the Non-sphere group in the case of the Paper mulberry.

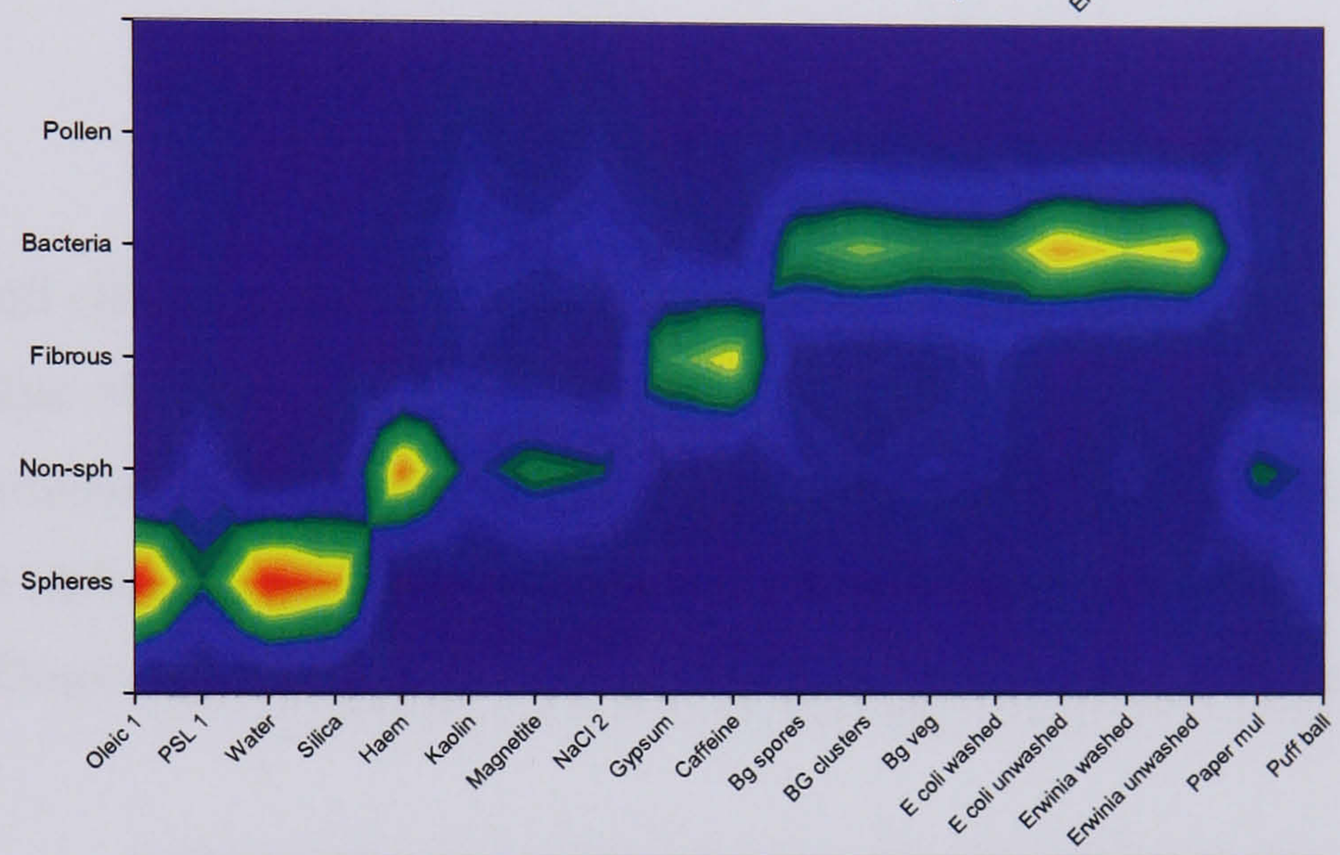
## Shape

Lin. Disc.

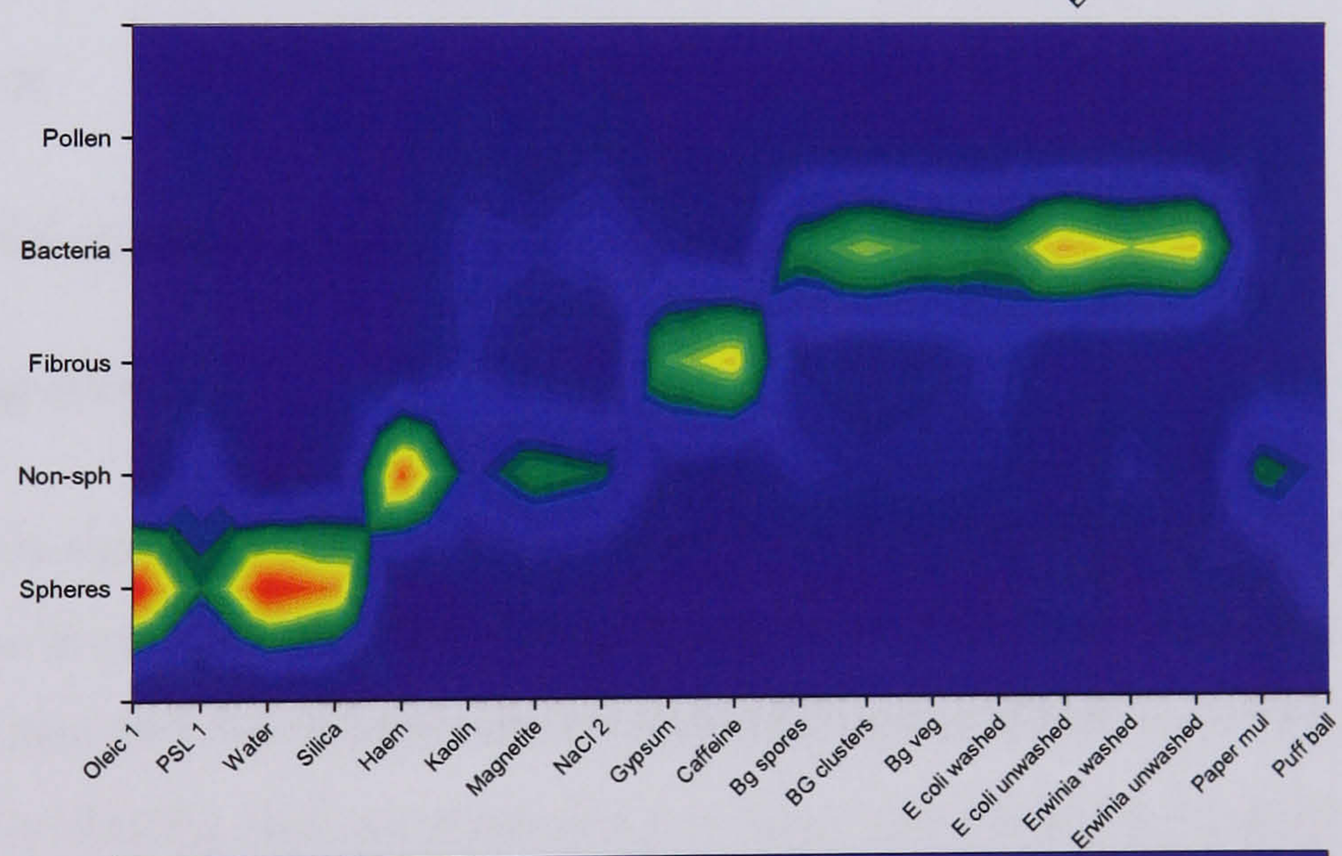



Example: Puff Ball pollen mis-classified into the 'Spheres' and 'Non-spheres' classes instead

KNN




Fuzzy K




Normal Dist'n.

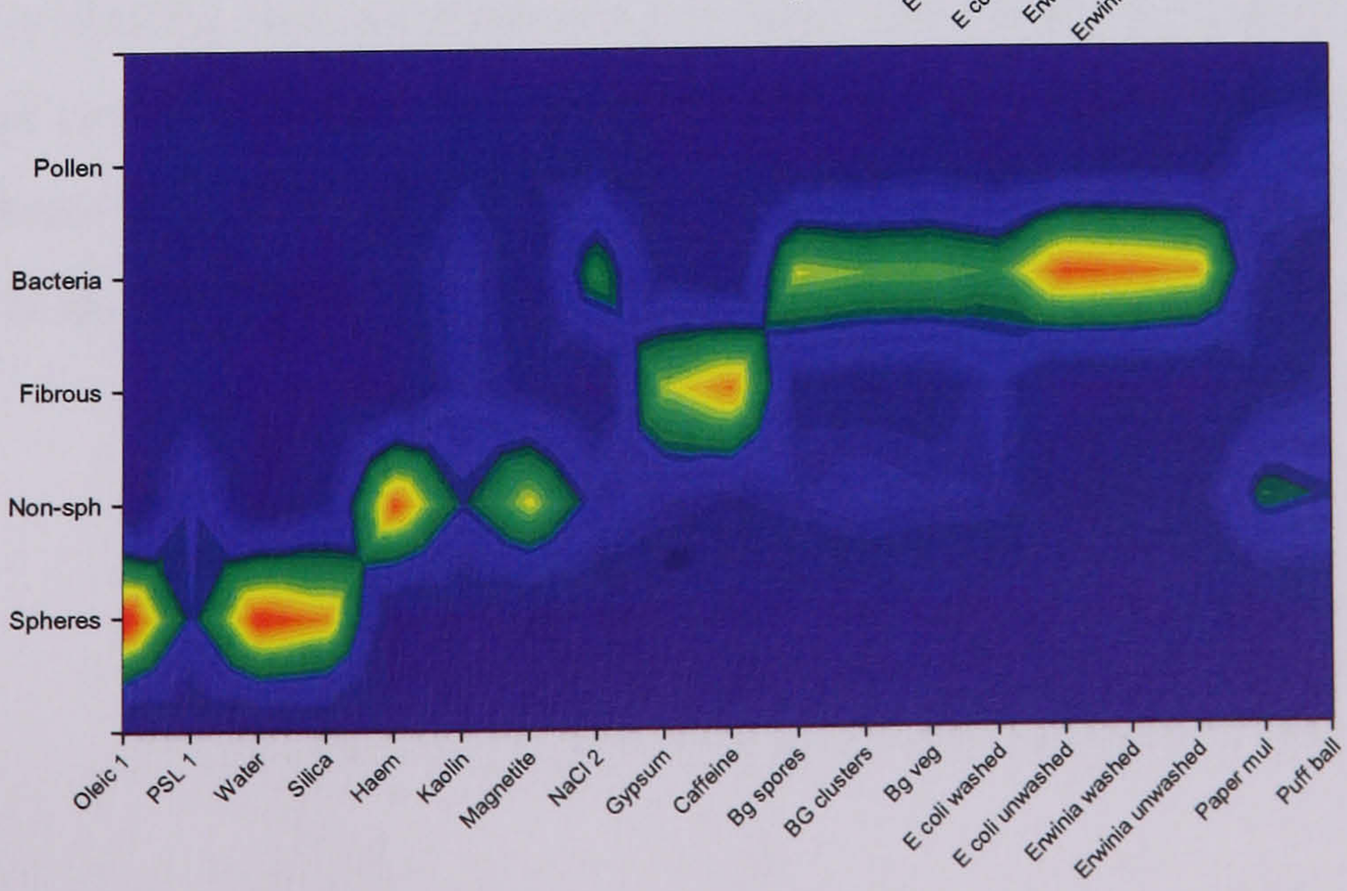



**Fig. 8.6 Results of particle classification for each algorithm using only SHAPE data.**

### 8.7.2 Input data format: Shape + Size

See Figure 8.7 overleaf.

Algorithm: **Linear Discriminant**

This algorithm has performed poorly for virtually all particle types except the larger spheres.

Algorithm: **KNN**

The Spheres are well classified with a small improvement in the classification of the 1 μm PSL. In general, those materials that naturally have a broad distribution in size have shown the greatest improvement. This can be most clearly seen in the polydisperse water and silica results. Overall the inclusion of the Size parameter has resulted in a small improvement across all groups. Despite this the Pollen are still largely mis-classified across virtually all available classes.

Algorithm: **Fuzzy-K**

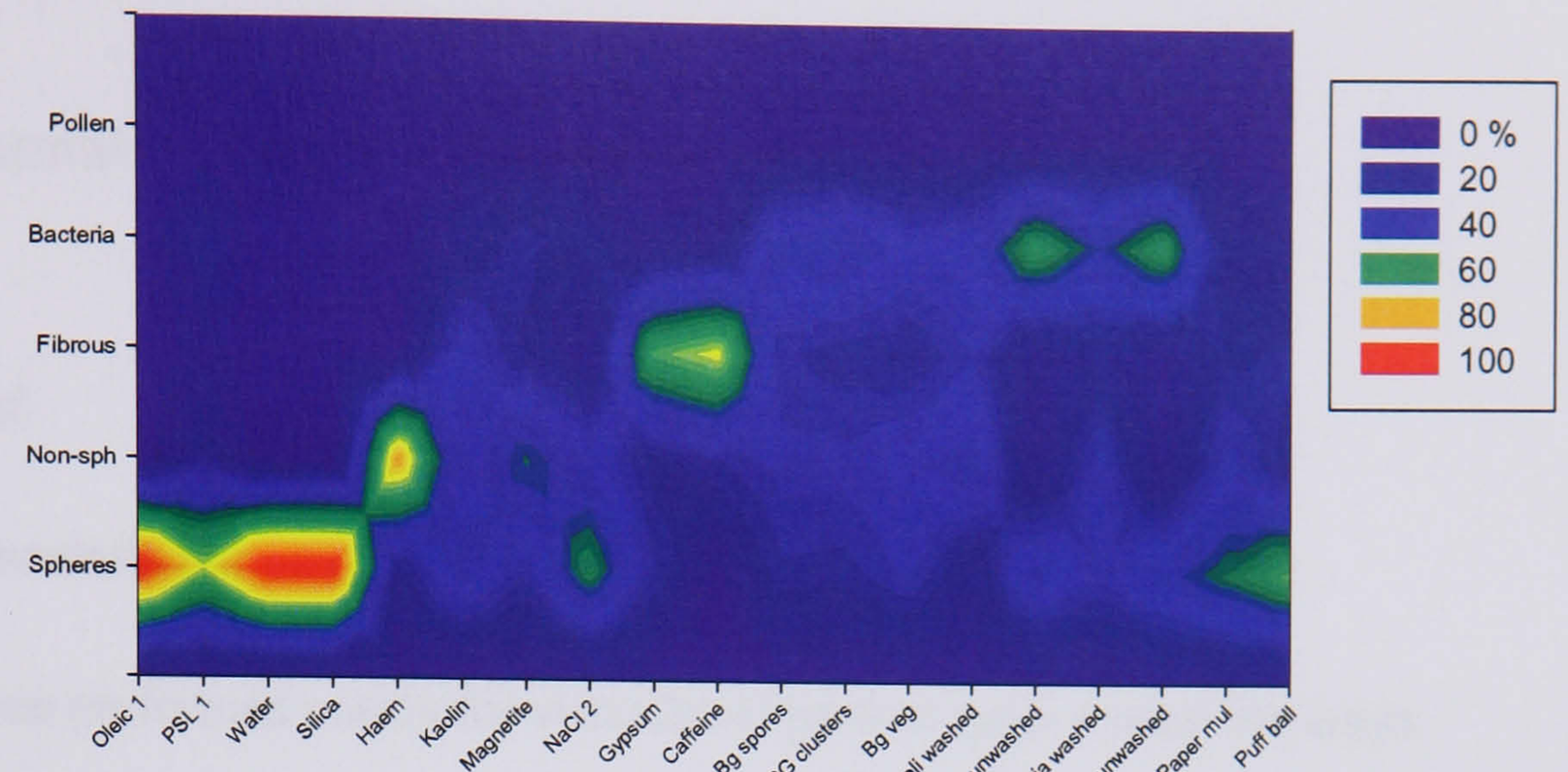
Very similar to KNN above.

Algorithm: **Normal Distribution**

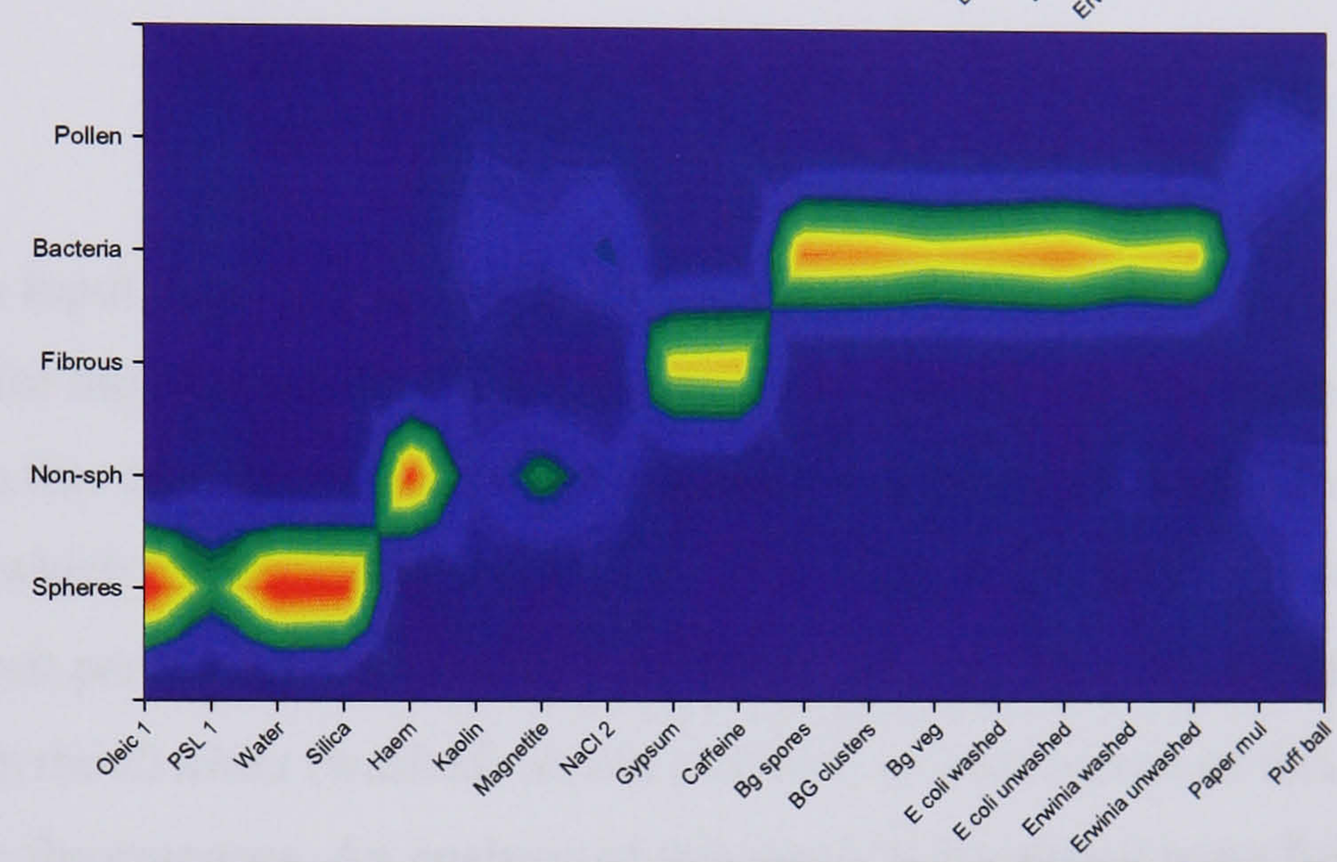
For most groups this algorithm does not perform as well as KNN although the differences are small (as shown in the data of Appendix 1). The graphical output shows little improvement over that derived for the case of Shape data alone (Section 8.7.1). One significant change is that the NaCl particles are no longer classed as Bacteria but in this case are correctly classed as Non-spheres. For Pollen, the inclusion of the size parameter has resulted in a significant improvement in classification. This is the only group where there is a clear divergence in the performance of the different algorithms, as can be seen in Figure 8.7.

## Shape + Size

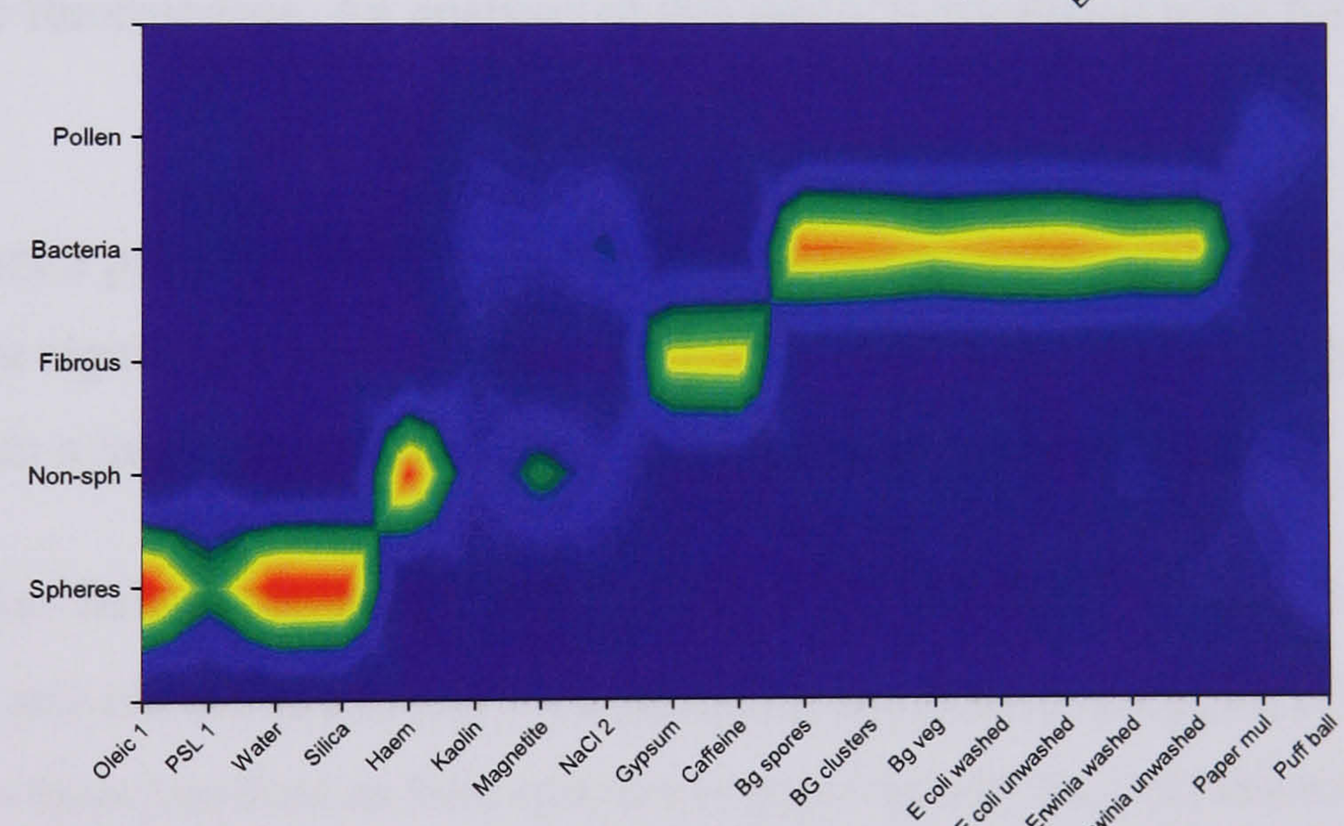
Lin. Disc.




KNN




Fuzzy K




Normal Dist'n.

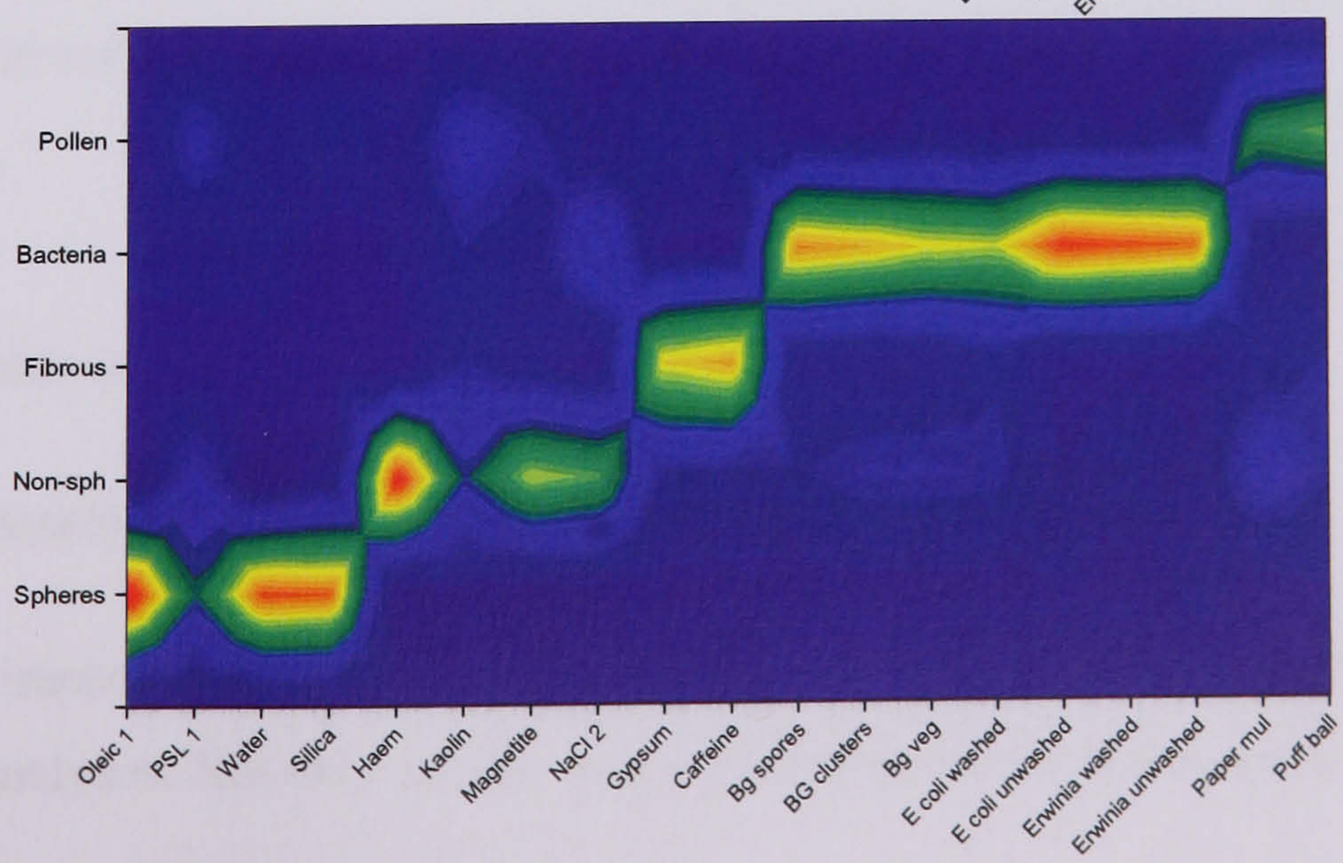



**Fig. 8.7 Results of particle classification for each algorithm using SHAPE and SIZE data.**

### 8.7.3 Input data format: Shape + fluorescence

See Figure 8.8 overleaf.

Algorithm: **Linear Discriminant**

This algorithm has again performed poorly for virtually all particle types except the larger spheres.

Algorithm: **KNN**

In contrast to previous Input data Formats using the KNN algorithm, the performance here has greatly improved for the Non-sphere group. However, the Bacteria group has become less well classified.. In this case the mis-classification exhibits a characteristic pattern. The largest single class to which the bacteria are misclassified is NaCl (Non-spheres group), with a further significant percentage assigned to 1 μm Oleic spheres. This misclassification is most noticeable with the *Erwinia* (washed) results probably a consequence of their relatively low intrinsic fluorescence. An analysis of this result is developed more fully in Appendix 1.

For an instrument where a primary aim is to discriminate between biological and non biological particles, the significant level of misclassification shown by KNN in this case would be likely to mean a large proportion of biological material would be missed.

The KNN algorithm also fails to correctly classify either of the varieties of Pollen, with the Puff Ball again badly mis-classified into the Non-sphere group. Interestingly, the 1 μm Oleic particles are also mis-classified as Non-spheres suggesting that the fluorescence data are having a detrimental influence on classification of this particle type.

Algorithm: **Fuzzy-K**

Very similar to KNN above.

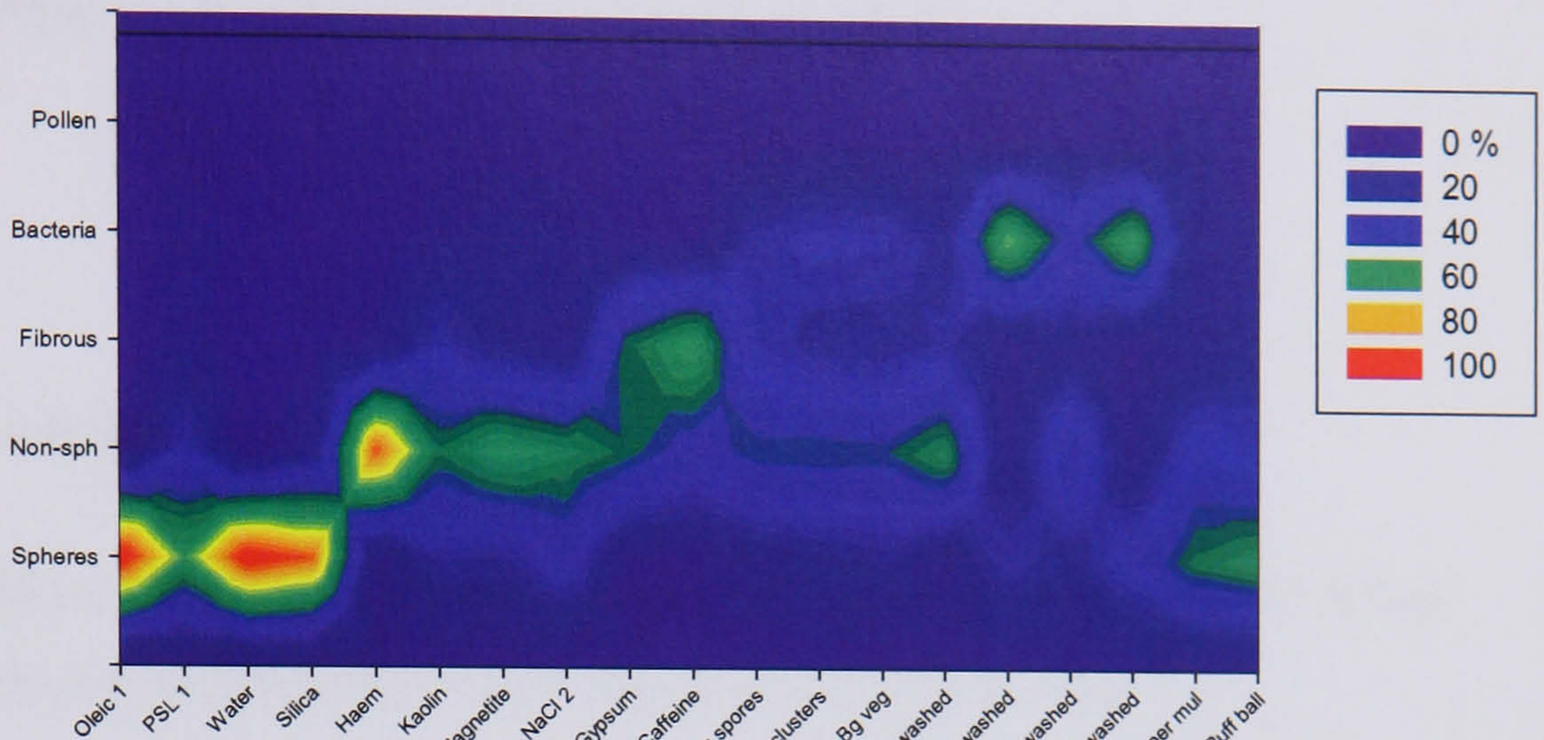
Algorithm: **Normal Distribution**

This shows a marked improvement in classification efficiency over previous input data formats / Norm Disc analyses. Not only are the various bacterial particles classified with

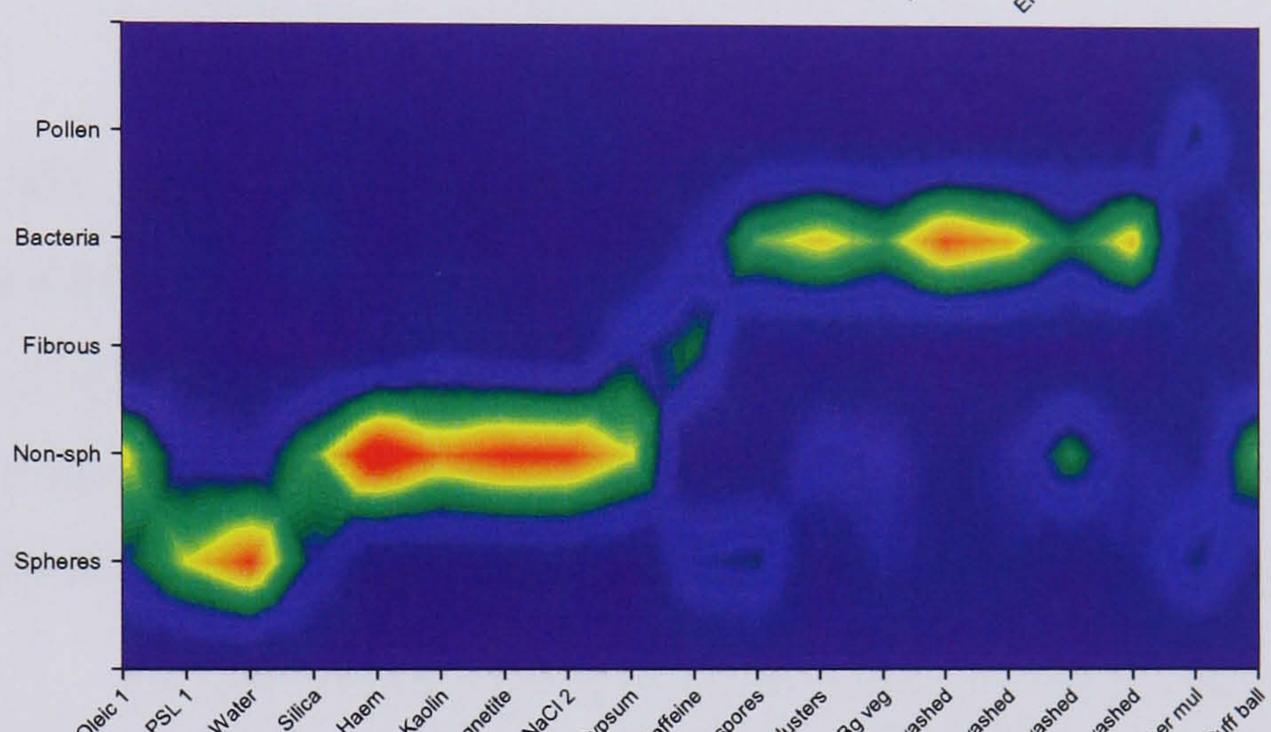
high accuracy into the Bacteria group, but, for the first time, a pollen (Paper Mulberry) has been correctly classified into the Pollen class with reasonable accuracy (~60%). The other pollen, Puff Ball, exhibits low intrinsic fluorescence and as a result has been mis-classified into other classes, principally the Non-sphere class. Most other particle types are classified with reasonable accuracy, with the exception again of 1 μm PSL.

**Shape + Fluorescence**

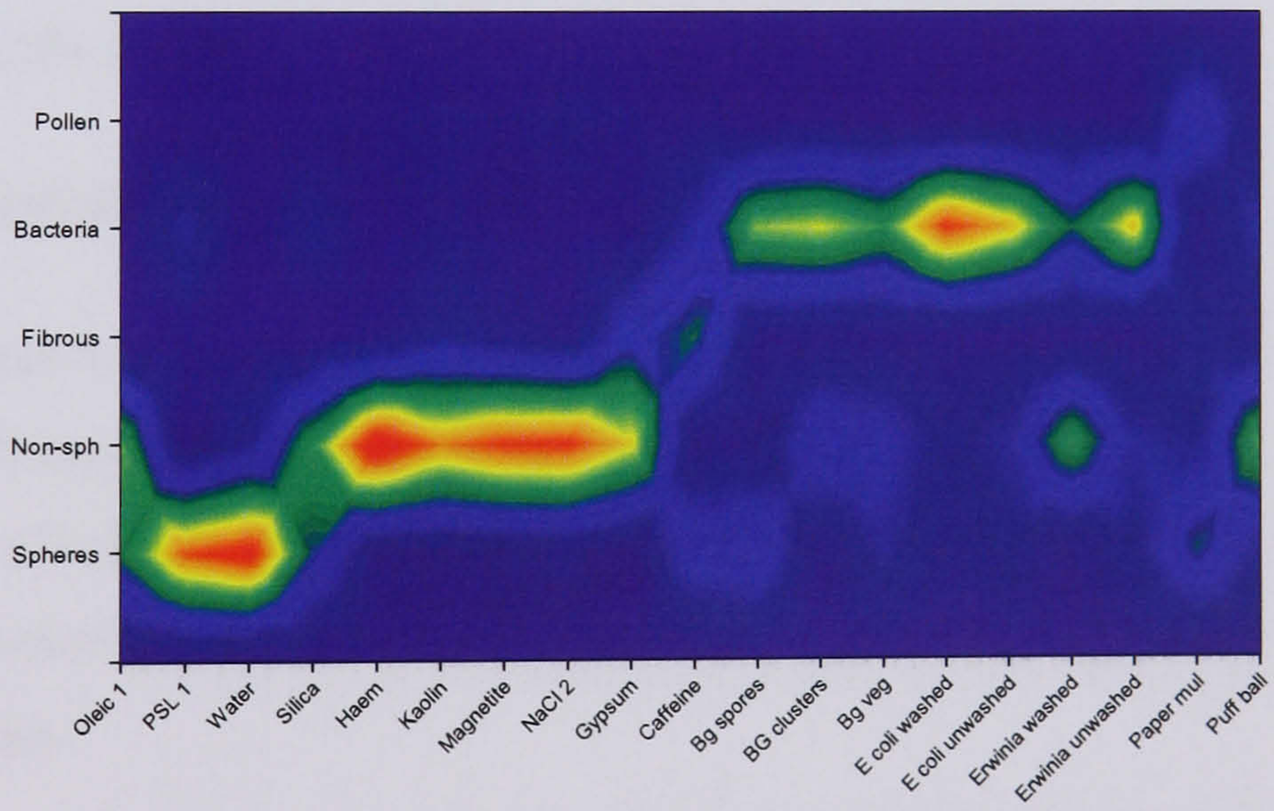
Lin. Disc.




KNN




Fuzzy K




Normal Dist'n.

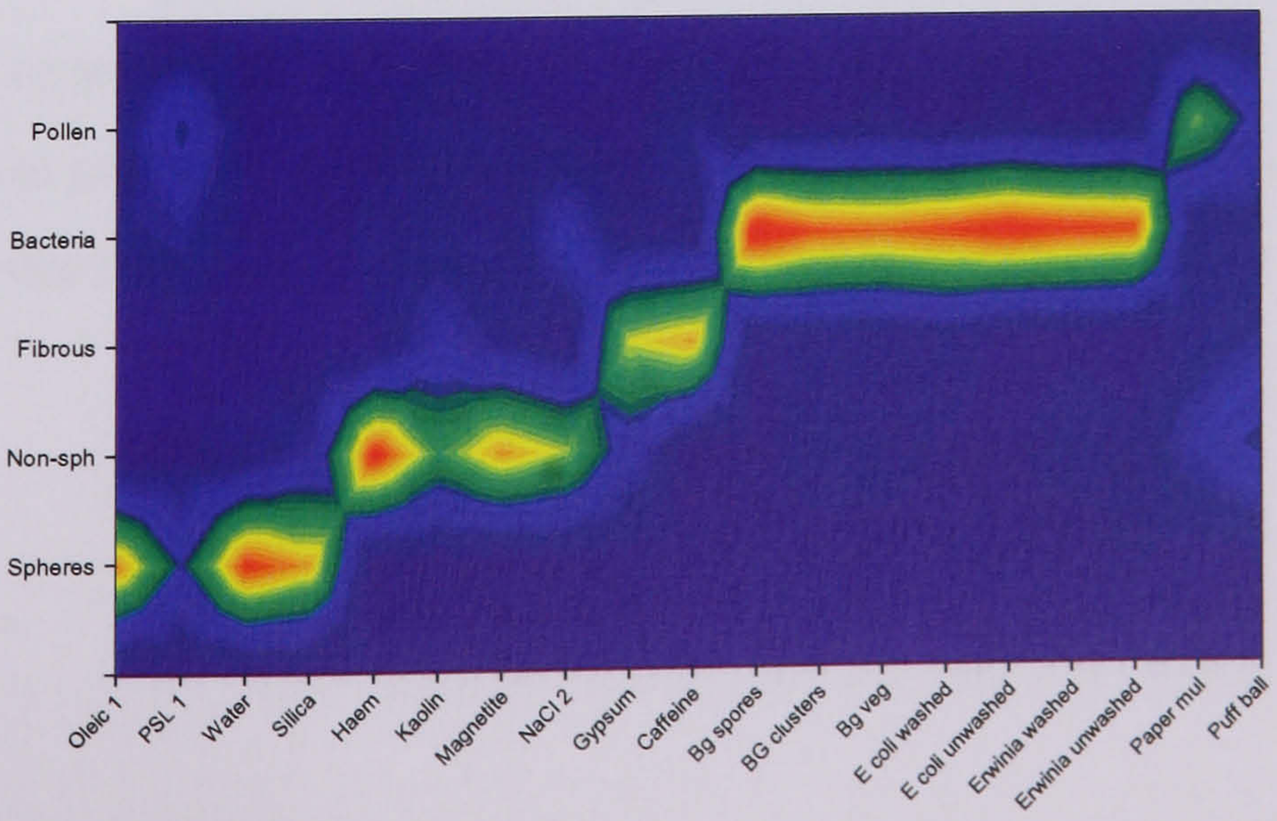



**Fig. 8.8 Results of classification for each algorithm using SHAPE and FLUORESCENCE data.**

### 8.7.4 Input data format: Shape + size + fluorescence

See Figure 8.9 overleaf.

Algorithm: **Linear Discriminant**

Even for this rich data input, this algorithm has again performed poorly and, despite its fast execution speed, is clearly not satisfactory for use in processing MPAM data.

Algorithm: **KNN**

Over all KNN performs well. In the case of the Non-spheres group it performance is better than any of the other algorithm / data format combinations. The previously described erroneous classification of 1μm Oleic (Section 8.7.3) is remedied using this input data format.,. However, as for the other data formats considered so far, neither of the pollens are well classified, the additional parameter providing no real improvement in this instance.

Algorithm: **Fuzzy-K**

Again, very similar to KNN above.

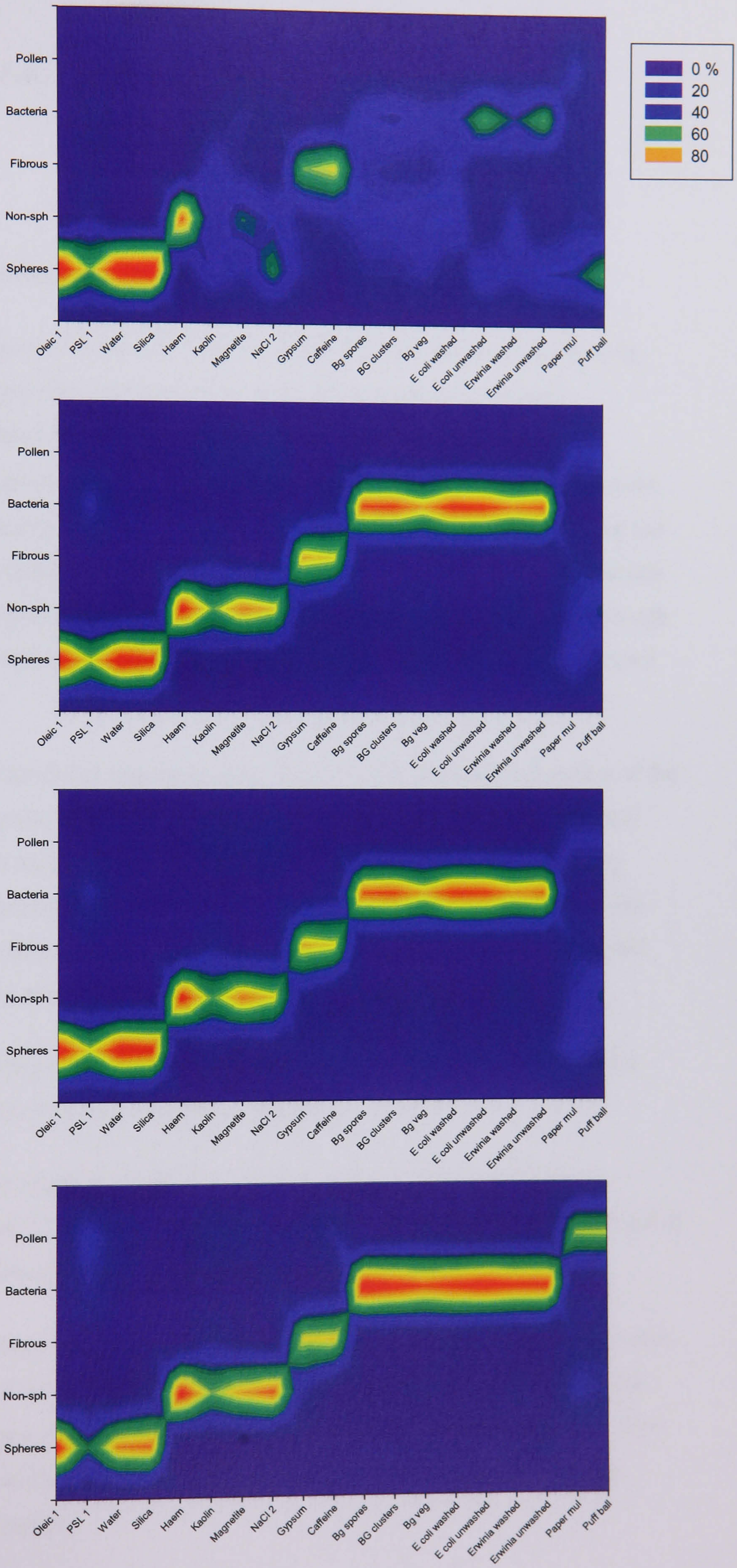
Algorithm: **Normal Distribution**

For all the particle varieties examined, this algorithm performs well. All the bacterial are especially well classified into the Bacteria group. For the first time, both Puff Ball and Paper Mulberry are classified with reasonable accuracy, with ~67% being allocated to the Pollen group. Also importantly less than 2% of these materials is mistakenly classified into any of the bacterial classes.

Again, with the known exception of 1 μm PSL, the overall performance of the classification procedure is excellent, reinforcing the fact that all three types of data – shape, size, and fluorescence – are important contributors to the classification process.

**Shape + Size + Fluorescence**

Lin. Disc.




KNN

Fuzzy K

Normal Dist'n.

**Fig. 8.9 Results of classification for each algorithm using ALL data parameters.**

### 8.7.5 Input data format: size + fluorescence

See Figure 8.10 below.

Algorithm: **Normal Distribution**

The example given in Figure 8.10 shows the best classification performance when using Size and Fluorescence input data, and again this was achieved using the Normal Distribution method. Overall, the classification accuracy is reasonable, although a significant proportion (approx. 34%) of the 1 μm PSL was mis-classified into the Bacteria group. This was not unexpected given the similarities in size and fluorescence between the two particle types, and it highlights the fundamental weakness of classification approaches which do not incorporate particle shape as a discriminatory factor. (This was noted through the presence of fuel droplets causing false positives in recent field trials of US bio-aerosol detection systems).

The results for the non-fluorescent water are also instructive. A significant proportion of the water droplets are misclassified into one of the non-spherical groups. The Non-spherical group receives the largest single allocation with 39%: six percent more than is actually classed as spherical. A further quarter of the water sample is classed as fibres. Again this reflects the similarity in both the scattering size and fluorescent of the water droplets and the non-spherical particles examined (ie: haematite, kaolin, magnetite, NaCl).

The lack of a shape parameter can clearly give rise to some significant confusions in the classification of some common types of aerosol particles.

Also, although caffeine is reasonably well classified, the gypsum, which is far less fluorescent, is poorly classified across several classes. This again highlights the drawback of not having shape information in the classification process.

Finally, two types of bacteria, BG (veg) and *Erwinia* (washed) are poorly classified. Again this is probably a result of the lower intrinsic fluorescence of these particle types and the resulting confusion between them and particles such as haematite and magnetite. It is also worthy of note that the low-fluorescence paper mulberry spores are mis-classified in a similar was for the same reasons.

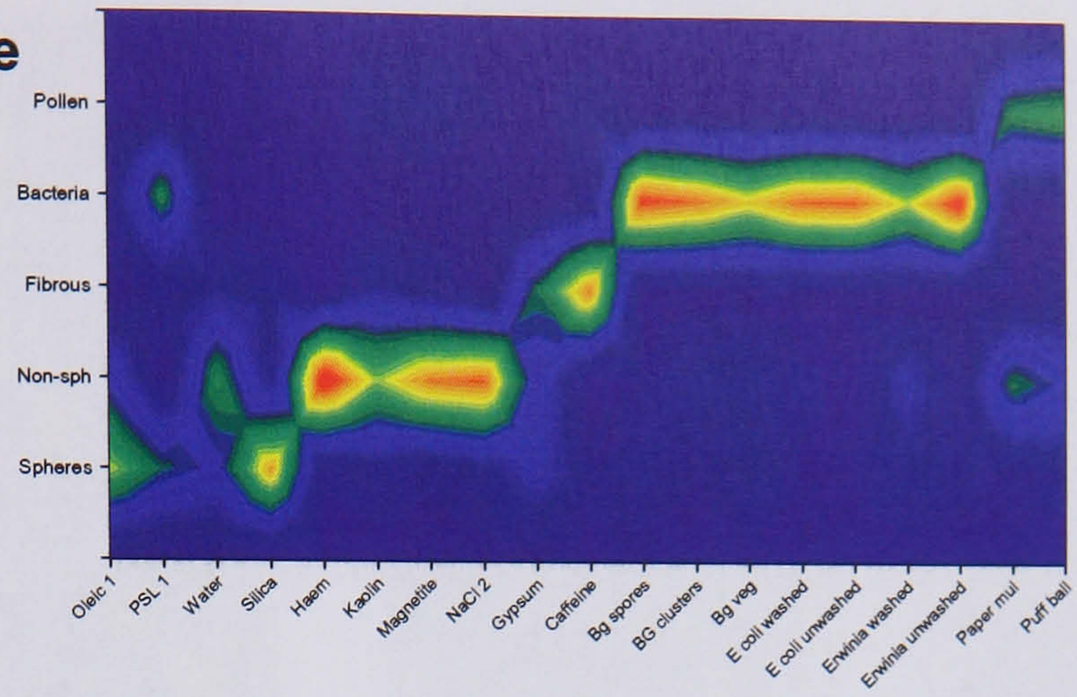



**Fig. 8.10 Results of classification for Norm. Dist. Using Size and Fluorescence parameters.**

The Lin Disc, KNN, and Fuzzy-k classification performance graphs have not been reproduced here as all were significantly inferior to the Norm Dist algorithm shown above.

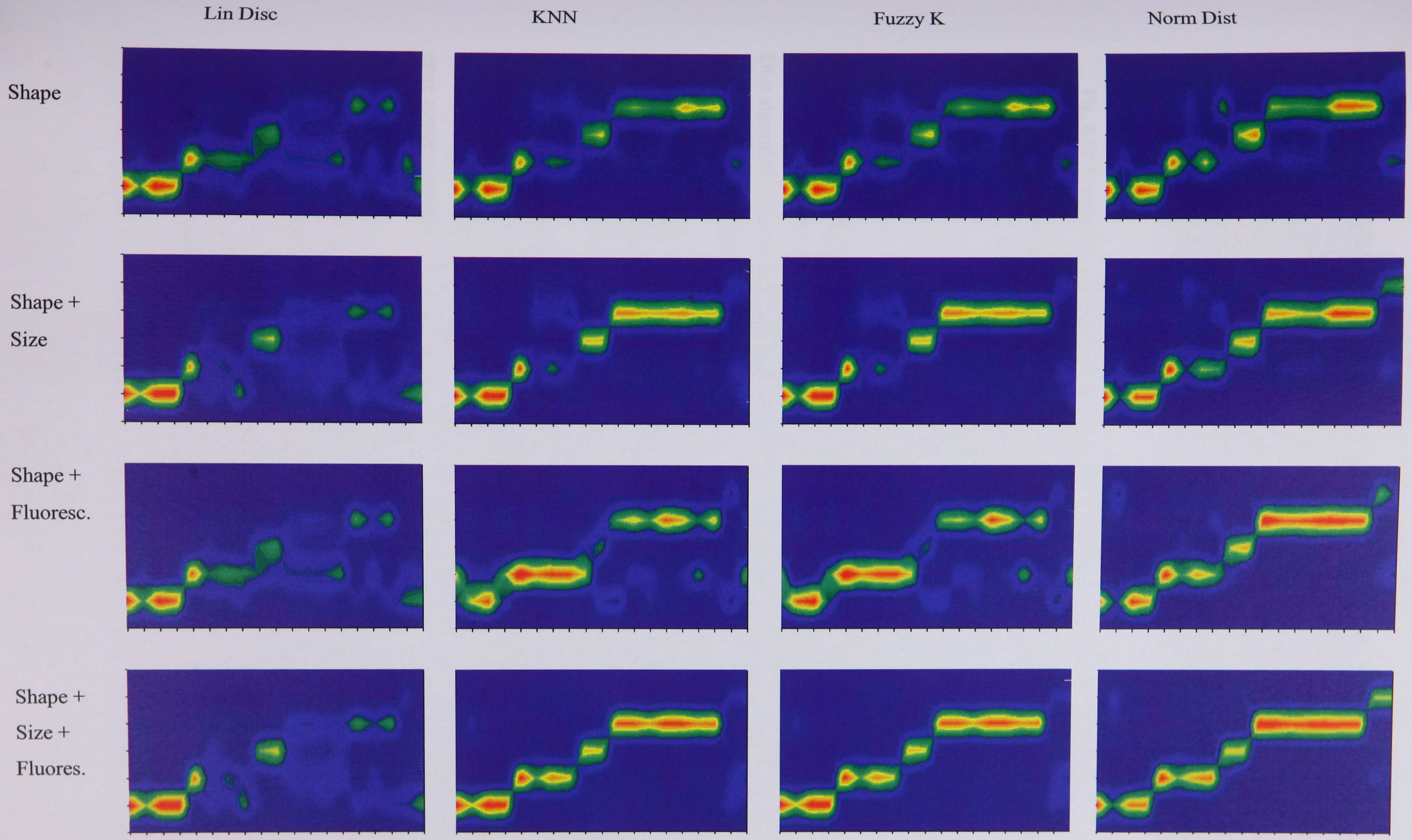



**Fig. 8.11 Summary of classification results for ALL algorithms and ALL data combinations used.**

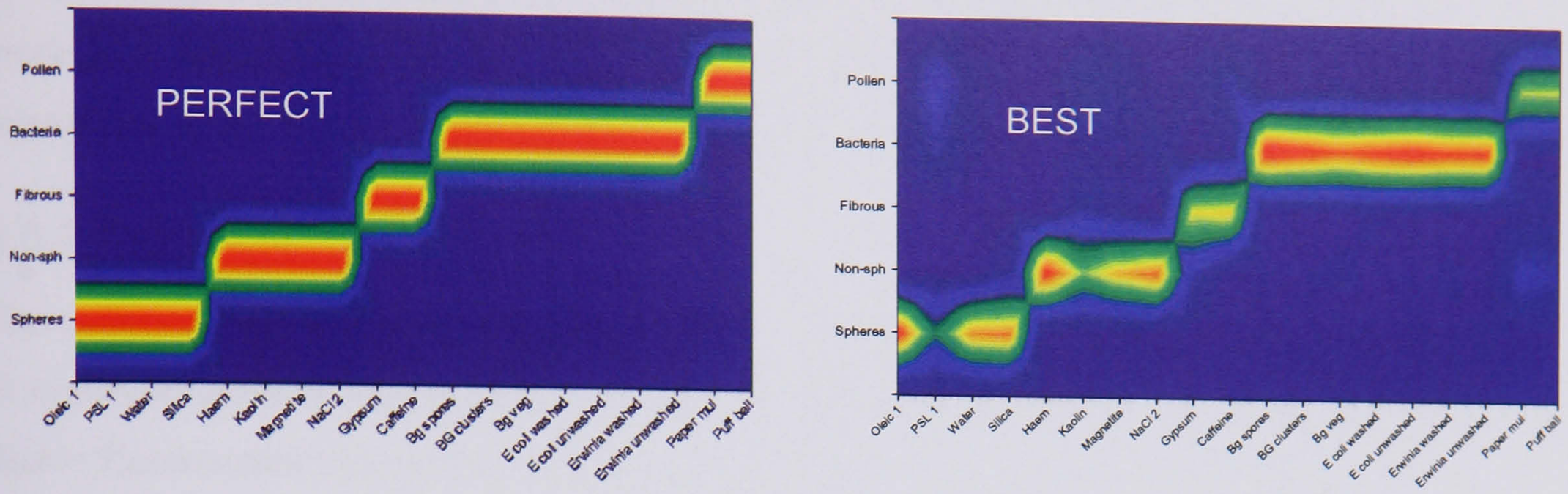



**Fig. 8.12 Comparison of the 'Perfect' classification result with the best performing real classification achieved through the application of the 'Norm. Dist.' algorithm to MPAM data relating to particle Size, Shape, and Fluorescence. Most particle classes are correctly classified, with the Bacteria class being especially well represented.**

## 8.8 Discussion

The detailed consideration of the data shows that no single algorithm or data combination performs universally well. The efficiency of the classification when undertaken by the Lin Disc algorithm shows little improvement between that of the single parameter input of shape and any of the other input combinations.

It is possible, however, to conclude that for Norm Dist, KNN and Fuzzy-k, the accuracy of classification improves with the number of parameters as can most easily be seen in the summary of Figure 8.11.

Both KNN and Fuzzy-k succeed in classifying all particle types reasonably well with the exception of pollens. Where their performance does diverge, Fuzzy-k performs worse. Given that it involves additional processing, there is no advantage to using this approach over the more straight forward KNN.

The Norm Dist algorithm generally performed as well as the KNN/Fuzzy-k algorithms for all classes. In particular, for those aerosol particles exhibiting more complex morphologies or size distribution (polydisperse water) Norm Dist out performs the other algorithms. When used with the multi- parameter data it is the only algorithm that succeeds in classifying the pollens with reasonable accuracy. In addition, the Norm Dist algorithm has the advantage of far faster execution speed than either of the KNN variants.

Over the range of aerosol materials investigated, the Norm Dist achieves the most efficient levels of classification with the lowest computational overheads. For any detection system intended to operate in near real time this is a important advantage.

### 8.8.1 Data input combinations

The inclusion of a fluorescence parameter (along with shape or size data) alone is insufficient to avoid confusion between some of the aerosol samples as discussed above. In particular the Size + Fluorescence data combination gives a misleading classification of non-biological material into the Bacteria group, highlighting the drawback of not having particle shape information to complement size and fluorescence data in the classification process.

The Size + Fluorescence data input combination also gives rise to a large under-reporting of the biological sample (Section 8.7.5), such that an initial build up in the bacterial concentration within the ambient aerosol could be missed.

## 8.9 Summary

The potential for using MPAM data to achieve discrimination between various aerosol particle types within an ambient atmosphere has been assessed using four well-established classification algorithms and various combinations of the Size, Shape, and Fluorescence data parameters available from MPAM. Each algorithm required 'training' via exemplar data from a wide variety of known particle types.

The results of the assessment showed the multi-parameter approach of Shape, Size and Fluorescence combined with the Norm Dist analysis algorithm performs at least as well as any other data input combination and in some important cases achieves better results. It is the only approach to succeed in classifying the pollens with a reasonable level of accuracy. It provides the highest levels of accurate classification of the bacteria both at the level of the group and in terms of individual classes. The problem of misclassification of bacteria to other classes is effectively removed with never more than 5% assigned to any other single class (typically ~1 - 2%), well below any reliable detection threshold level likely to be used.

Perhaps the most important feature is the ability to discriminate non-biological materials that exhibit both size and fluorescence attributes similar to those of the bacterial samples tested. The risk of reporting of a non-existent bacterial presence arising from the presence of interferents is therefore reduced.

Although a limited range of materials has been tested, the results demonstrate that the Normal Distribution algorithm, when provided with particle data relating to all three parameters (Size, Shape, and Fluorescence), provided the best performance in terms of particle classification / discrimination capability, and this augers well for the potential use of MPAM in a real-life field monitoring application.

(More details of these classification experiments, including classification efficiency data for each combination of algorithm, parameter set, and particle type, are given in Appendix 1)

# 9. Conclusions and further work

The primary objective of the project was to develop a monitoring instrument capable of achieving enhanced levels of discrimination between different types of ambient aerosol particles down to the sub-micrometre size-range, with the specific aim of differentiating biological from non-biological particles.

This objective has been largely achieved by the construction of a unique instrument, the MPAM (Multi-parameter Aerosol Monitor). The MPAM is capable of simultaneously recording the spatial light scattering profile and the UV-excited intrinsic fluorescence from individual aerosol particles at high throughput rates. The initial specification called for the instrument to operate over a range of sizes down to the sub-micron size range. In practice the lowest size range at which the MPAM can reliably operate is ~1 μm (see section 9.2.3).

The development and testing of the MPAM instrument, along with the investigation of different data input combinations and processing algorithms has allowed some general conclusions relating to performance and particle discrimination capability to be drawn. These conclusions also provide a useful indication as to those areas where further work can most usefully be undertaken.

## 9.1. Conclusions

The following conclusions may be drawn concerning the overall performance of the MPAM instrument.

The instrument has been shown to be capable of recording useful data from individual particles in the approximate size range ~1 μm to 12 μm, corresponding to that into which most airborne biological particles of interest are found.

The instrument is capable of recording data from ~6,000 particles per second. With a sample volumetric flow rate of ~1.0 l/min., this data capture rate corresponds to all particles in an ambient aerosol for concentrations up to ~ $3.6 \times 10^5$ particles/l.

MPAM is the first instrument capable of recording simultaneous data relating to particle size, shape and intrinsic fluorescence.

Trial data analyses using four popular classification algorithms: Linear Discriminant, K-Nearest Neighbours, Fuzzy-K Nearest Neighbours, and Normal distribution, have shown that Normal Distribution provides the greatest degree of particle discrimination for the range of particle types tested.

The performance of the Normal Distribution algorithm improves with increasing richness of input data, with the best results being achieved with the unique set of Size, Shape and Fluorescence parameters produced by MPAM. This combination of algorithm and input data has been shown to provide excellent overall particle discrimination for the range of aerosols investigated. In particular:

- It was the only combination able to classify the bacterial samples unambiguously.
- It performed best in discriminating between the different classes of biological Particle (bacteria, pollens) tested.
- It was the only combination able to classify the pollen with any degree of success.
- It was the only combination that did not classify the non-biological materials tested into the biological or pollen groups, thus avoiding 'false-positive' outcomes.

True real-time hardware/software implementation of the Normal Distribution classification algorithm was not implemented in practice, although processor timing tests on pre-recorded data showed that this algorithm could indeed achieve real-time classification at the particle throughput rates achievable in MPAM when implemented on a typical gigahertz processor available in contemporary personal computers.

## 9.2. Discussion

Taking these conclusions together shows that the MPAM was superior to the other approaches reviewed in Chapters 2 and 3 for discriminating the biological component of an ambient aerosol in real time. Although it was not part of the original aim to provide an improved level of discrimination *between* the different biological materials, the multi-parameter results also show that it is in fact possible to go some way towards providing a degree of classification at the level of the individual species.

However, since by their nature the parameters being recorded by MPAM are, to a large extent, generic to bacteria at the level of the genus and are not influenced strongly by features expressed at the level of the species, it seems unlikely that any combination of these parameters will be able to provide species level detection. Discrimination at the species level will require detection methods able to operate at the level of individual

proteins or the nucleic code. This non-specificity of the parameters, nevertheless, may be advantageous. In the context where a main requirement is the detection of biological warfare agents, the ability to detect a biological organism at the most general level may be important, since these aspects are difficult to disguise. Consequently, one strength of the multi-parameter approach may lie in recognizing previously unknown agents as being a biological threat. This may be particularly important when the detection instruments are confronted with a challenge from a biologically engineered material intended to avoid detection by the more specific methods such as PCR (Polymerase Chain Reaction).

The capacity to discriminate biological material at a general level, and hence potentially to detect previously unknown biological agents, is further indicated by the MPAM response to the unwashed biological samples. When unwashed bacterial samples have been tested, the samples are either classified to their respective class or to the other unwashed classes, rather than to the washed class of the same bacteria, or indeed to any other of the bacteria. That this phenomenon is present in this way illustrates the influence of the culture medium on the observed characteristics of the bacterial aerosol. The sample is, however, still recognised as biological in origin.

In a situation where the primary requirement is the detection of a specific known biological organism, such as anthrax for example, additional means of detection will be required. As discussed above, such a level of discrimination is unlikely to be possible with instruments such as the MPAM. A hierarchical solution with the MPAM operating in real-time as a high level classifier alerting a slower, lower-level, species specific classifier to the presence of biological material could be developed. In this scenario it would be possible to exploit the relatively high degree of specificity demonstrated by the MPAM to reject a large quantity of the common bioaerosol background, making the job of the lower level classifier easier.

Pan *et al* [83] have demonstrated one solution to this problem by developing a system that diverts particles according solely to the magnitude of their intrinsic fluorescence into a second chamber where an automated PCR analysis is carried out. Such an approach could also be employed, arguably more efficiently in terms of particle selection, with an instrument such as the MPAM.

One further method that has been reported to increase the specificity of fluorescence based bio-detection techniques is to record a detailed particle intrinsic fluorescence spectrum. In the majority of these instruments (as described in Chapter 2), the dominance of tryptophan

fluorescence, including when other fluorophores are present, has meant that even with high resolution spectra, the fluorescence signature between different organisms has remained essentially indistinguishable. Consequently, despite the additional complexity of recording such spectra on a particle-by-particle basis, such approaches have not been able to demonstrate superior levels of discrimination. This indicates that future work intended to improve discrimination based on a more detailed spectral analysis of the bulk fluorescence recorded by the MPAM will have little advantage over the undispersed fluorescence output.

As discussed in Chapter 2, different instruments have excited intrinsic fluorescence at other than the 266nm wavelength used in the MPAM. In particular, the UVAPS instrument operating at 355nm can detect intrinsic fluorescence from flavins and NAD compounds as described by Agranovski *et al* [35] and others. A recent instrument developed by Sivaprakasam *et al* [43] where multiple wavelength excitation and detection was employed, shows that the use of multiple wavelength excitation may give enhanced levels of discrimination between different species. The MPAM has already shown that it is possible to provide some inter-species discrimination. The use of a multiple wavelength excitation-emission matrix (EEM) might improve the classification efficiency at this level even further.

One area of future work therefore would be the modification of the MPAM instrument to include excitation at an additional wavelength.

### 9.2.1. Improved false positive rate

When used with any of the two-parameter combinations (e.g.: Size + Shape, Shape + Fluorescence), the Normal Distribution algorithm has been shown to offer good levels of classification accuracy. In the case of the spherical group, because of their well defined and consistent morphology, extremely good levels of classification can be achieved. However, even in this apparently simple situation, the dual parameter combination has an important weakness. In particular, the Size + Fluorescence combination, as used in several of the biological detection instruments developed by other workers, can give rise to a confusion between spherical particles that exhibit fluorescence and the bacteria. The lack of a shape parameter in this case precludes the possibility of any reliable discrimination being made. In the situation where an instrument based on these parameters is used to provide warning of a biological offensive against an airfield for example, fluorescence from unburnt aviation fuel droplets (as commonly released from jet engines) could give rise to false-

positive alarms. The inclusion of the shape parameter, along with those for size and fluorescence eliminates this confusion from the classification. Consequently, the likelihood of this type of spherical droplet being incorrectly flagged as a bacterial aerosol is appreciably reduced.

Given the natural variability within the majority of test aerosols used, and instrumental factors such as signal noise etc, a small (<3%) percentage of each aerosol is typically classified into all groups and can be treated as noise in the classifier's output. In a number of instances, the dual input combinations resulted in a grouping of the mis-classified aerosol into a specific group. For example, when only size and fluorescence data are used as inputs, a significant percentage of the spherical PSL is misclassified to the bacteria group. This tendency to mis-classify into a specific group is reduced in the multi-parameter results.

Consequently the number of events leading to false positives is significantly reduced with the multi-parameter input. This is of great importance in any instrument designed to trigger an alarm when an unusually high concentration of a specific aerosol occurs in the ambient atmosphere. Significant numbers of false positive alarms can rapidly lead to a loss of confidence in the performance of the instrument, ultimately resulting in genuine alarm events being ignored.

### 9.2.2. Classification of pollens

Of all the groups, the pollen proved the most difficult to classify accurately. This reflects their heterogeneous nature, especially in the case of the paper mulberry pollen. Additionally, the low level of fluorescence observed from the puff ball spores suggests that the fluorescence parameter did not play a strong role in the classification of this particle type. It appears that excitation at 266 nm may not be optimum wavelength to elicit fluorescence from pollens.

Although none of the two parameter combinations did particularly well with the pollen, the Size + Shape combination had the highest success rate, illustrating the significance of the shape parameter in achieving an accurate classification.

The Size + Fluorescence combination, as well as poorly classifying pollen, misclassified a significant part of the pollen as NaCl. In a outdoor air sampling situation, given the ubiquitous nature of pollen and the distinct possibility of NaCl being present in the free air (especially in coastal regions) the inability to discriminate between pollen, salt like

materials and other biological materials is a weakness of detection methods based solely on these two parameters.

The lack of a fluorescence signature from the puff ball pollen, when excited at 266 nm, combined with the poor discrimination that the size term seems to confer on the classification of pollen implies that improved measures of fluorescence will not help in the identification of this class of biological material. The results from the Shape + Size suggests that a shape parameter plays an important role in the accurate classification of this class of aerosol.

That only the multi-parameter shape, size and fluorescence combination was able to classify the pollen with a good degree of success illustrates the importance of being able to apply the maximum level of information to the classification for aerosols of a more complex or variable morphology.

### 9.2.3. Instrument performance

In the initial specification for the instrument it was hoped that particles down into the sub-micron range could be classified. In practice the lowest size range at which the MPAM can comfortably operate is ~1 μm. Below this size, the noise level degrades the instrument's performance. The main reason for this is the high level of optical noise present in the chamber. This noise is largely the result of the lack of effective anti-reflective coatings being available for the lenses at 266 nm wavelength, giving rise to a higher level of background illumination in the chamber than might otherwise be expected. The lack of a highly efficient non-fluorescing absorbing material for coating the mechanical surfaces also contributes to the high level of noise. With the continuing development of UV laser and diode sources, new materials better suited to working at these wavelengths are likely to follow. When they become available, the application of such coatings would reduce the lower size limit at which the instrument could successfully operate, thus improving the instrument's discrimination of smaller particles.

Although the choice of a single beam design to simultaneously generate both spatial scatter and fluorescence data has advantages in terms of allowing fairly straight forward optical and mechanical design, the use of a single source does introduce some disadvantages. Whilst a pulsed UV laser would have been cheaper and more compact than the CW laser used, it would have necessitated the use of another continuous-wave source (such as a diode laser), to interrogate the sample space, detect the presence of the aerosol particle, and

subsequently trigger the pulsed laser. This approach has been used in many systems developed by other workers and is known to be problematic in terms of consistency of timing of the laser pulse (i.e.: ensuring the target particle is in exactly the correct position when the pulse occurs). At the time of the MPAM instrument construction, continuous-wave UV laser technology was still largely at the prototype stage. The complexity of the mono-block doubler (MBD) and its sensitivity to vibration, along with the difficulty of selecting and aligning active areas of the crystal, meant there were problems in keeping the laser operating in a reliable manner.

The second generation of MBD technology which is now becoming available, with improved mechanical isolation and automated management of the crystal, has largely overcome these limitations.

Apart from issues of the reliability of the MBD technology one further disadvantage of the single beam design is that it is not possible to optimise the beam power for both the excitation of fluorescence and the generation of scattering profiles. Experience with the MPAM has shown that a fluorescence signal of sufficient magnitude to be detected could be produced at considerably lower UV power levels than were actually used. The fluorescence channel PMT detector was in practice only required to operate at low gain. The UV power level was determined by the need to have strong enough scattering signals arriving at the hybrid-photodiode (HPD) detector - in part a consequence of the limited gain of the HPD compared with that of a PMT.

An alternative design allowing the use of multiple sources, for example a green laser for the generation of scattering profiles, with the UV laser for the excitation of fluorescence, would make it possible to optimise each of these sources independently for both the fluorescence and scattering profile signals. By reducing the overall power level in the chamber such a solution would also help in the reduction of the background optical noise in the chamber, hopefully allowing the recording of improved scattering profiles from sub-micron aerosols. Offsetting this advantage would be the reduction in sensitivity of the recorded scattering patterns to particle shape, especially for the smaller particles, as a result of the lower particle size parameter (See Section 3.1).

## 9.3. Future program of work

As with all efforts to develop a new instrument, experience has suggested a number of improvements that would enhance instrument performance. Some of these are of a

relatively minor nature, such as improvements to the electronics to expand the dynamic range, and as such could be carried out easily. Other aspects of the instrument's performance would require more extensive modifications or a programme of experimental work to be undertaken. The main areas of any such future work are outlined below.

### 9.3.1. Laboratory trials

Both a consideration of the literature in this field and the experimental work described here will quickly reveal that only a very small number of different biological organisms have been used in the experimental work. This is largely a result of the difficulties of handling of bioaerosols in a safe manner within the laboratory environment, necessitating a careful selection of the organisms to be aerosolised.

With the very limited range of bioaerosol materials sampled, there is a possibility that the data analysis methods become finely tuned to some unsuspected features of these particles. Therefore, before additional work is carried out on improving the analysis methods, further experimental work should be undertaken to increase the range of primary data recorded from a more extensive range of materials. The experimental data can be improved in three areas:

It is commonly known that a wide range of factors, including the aerosolisation method employed, the culture medium, the relative humidity, and the presence of common environmental pollutants influences the culturability of bioaerosols. These factors may also influence the characteristic fluorescence or scattering profile observed. The data recorded to date indicates that, at the very least, the presence of culture media can be detected. Further work along these lines should be carried out to determine if the influence of these factors on the bioaerosols are detectable by the MPAM.

Secondly, the possibility of some commonly occurring particles confusing the instrument's potential to detect the presence of bioaerosols has already been alluded to. No experimental work has so far been carried out to determine the MPAM's response to the presence of such materials. A program could be undertaken to test the MPAM's performance with a range of potential interferents. These might include aviation fuel droplets, diesel exhaust, and signal smokes.

Finally, all the experimental samples used have been generated independently. It is possible that mixtures of simultaneously generated aerosols may influence the scattering or fluorescence signature recorded, particularly if electrostatic forces or collisions between particles of different materials result in the formation of aggregate particles. Materials that are commonly found naturally and influence the fluorescence and scattering signatures may form aggregates with bacteria. Such materials include NaCl crystals (wind blown sea salt) and haematite that may be carried hundreds of kilometres by high level air streams, having been picked up from iron rich soils. Some attributes of such aggregates may mask features of the aerosols seen when generated independently. Experimental work to investigate this process and its influence on the output of the MPAM would be useful in developing a fuller characterization of the instrument's capabilities.

### 9.3.2. Field trials

The issues surrounding the release of large quantities of bacterial material into the outdoor environment makes the carrying out of systematic field trials difficult. However, in some facilities such as that at the US Army Dugway Proving Ground, Utah, it is possible for the large scale release of known bioaerosols to be carried out under carefully monitored conditions. Such trials were not available to the author during the completion of the MPAM project, partly because of the unsuitability of the MPAM's prototype UV laser for prolonged transport over comparatively rough ground. Should this situation change, it would be useful to test the MPAM under such circumstances, allowing its performance to be more directly compared to that of other instruments and detection methods intended to operate in the field.

### 9.3.3. Outdoor trials

Ultimately, the capacity of instruments such as the MPAM to operate effectively within the real world environment is of interest. In the outdoor environment, except under unusual circumstances, bioaerosols are unlikely to present themselves for examination as carefully prepared fully washed samples. They are most likely to arrive as single organisms, as clusters, or associated with other material. Furthermore, the bioaerosol loading of the atmosphere varies widely from place to place and time to time. Longer term outdoor trials would be of interest in assessing the utility of this type of instrumentation compared to other techniques.

### 9.3.4. Multiple wavelength operation

Although requiring more extensive modification to the instrument, the introduction into the scattering volume of the third harmonic of the Verdi pump laser would enable the MPAM to be operated in a multiple wavelength EEM mode, with excitation wavelengths at both 266 nm and 355 nm. A further round of experimental work could then be undertaken with the specific objective of investigating the extent to which such enhanced optical detection methods might be able to achieve improved species level discrimination.

### 9.3.5. Investigation into the shape parameter

Even for an instrument employing multi-wavelength EEM, the MPAM results suggest the shape parameter will continue to provide improved levels of detection and discrimination. This may particularly be the case with pollens which often have a very complex and distinctive shape. For instruments intended to operate outside the laboratory environment, pollens are likely to form a large part of the biological component of the ambient aerosol. The ability to discriminate reliably between pollens and bacteria would be advantageous.

It has been assumed that the more detailed the recorded scattering profile, the better the level of particle discrimination achievable, and indeed this has been borne out in practice through the results from the various instruments described in Chapter 3. The disadvantage of increased detector resolution is, of course, the higher demands made of the data computation employed and, potentially, the lower particle analysis rate. However, no detailed systematic study has been carried out into the optimum scattering pattern detection resolution and of issues such as the relative importance of polar as opposed to azimuthal intensity variations in the pattern. (The MPAM detector is, of course, only able to cover a very limited range of polar angles). With on-going increases in available computation power, the demands placed by the inclusion of high resolution spatial scattering data in any particle classification algorithm, may now make viable the use of detector resolutions that even two years ago would have been impractical, and a detailed study of the performance of various detector geometries of increasing resolution would be valuable.

Data produced from such an exercise, when combined with the fluorescence data, could be reprocessed by the same data processing algorithms described in this thesis, allowing a fairly direct comparison to be made.

## 9.4. Summary

The MPAM instrument has demonstrated the capability for real-time discrimination between bioaerosols and other non-biological particles. When used with the best of the data processing algorithms, the MPAM data has demonstrated an enhanced level of discrimination between the various bacteria and pollens investigated. Of equal importance, the MPAM multi-parameter approach shows the least confusion between the non-biological aerosols and the bacteria and pollens, thus improving both the likelihood of detection and reducing the probability of false positive bio-detection. This represents an important further step towards the ultimate goal of high speed classification of individual species.

While for the reasons outlined above it remains unlikely that species level discrimination will be possible by optical methods alone, the MPAM has demonstrated that it is possible to go some way along that road. The results presented point the way to both a more extensive experimental programme and instrumental modifications that could further lift the MPAM's performance another step nearer to that goal.

John E. Barton

London, 2005

# Appendix

# A1 Numerical results

The classification procedures as described in chapter 8 when applied to the nineteen different test aerosols, five combinations of size, shape, and fluorescence parameters, and four different classification methods for each parameter combination gave rise to a very large body of classification results. Such detailed numerical data tables and their accompanying observations were considered too large to be incorporated within the main body of the thesis, and instead they are presented here in full.

In Chapters 7 and 8 graphical methods have been used to provide an overview of the classification landscape in an attempt to present the results in a simple clear format, making it possible to gain an overall understanding of the relative performance of each of the five data combinations. Such presentation methods whilst revealing the overall performance have inevitably required some simplifications to be made in the presentation of the data. A more detailed consideration of the results, however, is instructive as to the sensitivity of the different algorithms to each of the data parameters. Moreover, there were frequently specific observations to be made regarding the performance of a particular classification algorithm with a particular class of particle, and corresponding postulates have been offered in an attempt to explain the observations.

## A1.1 Format of Results

The results for each of the input data combinations are presented in a common tabular form. The first column identifies the material tested and the class to which it is allocated. This is followed by the analysis method applied. The third column gives the percentage allocated correctly to the group associated with that class. So that for example in Table A1.1 for kaolin (Class 6) the Norm Dist algorithm has allocated 5.75% of this sample incorrectly to magnetite (Class 7). However, magnetite is in the same group as kaolin and consequently this is included in the group classification. As an aid to clarity the area of the table included in each group has been highlighted.

The results for the five different data format combinations are presented in the tables as listed below. The tables themselves can be found at the end of the Appendix.

| Combination | Data format | Table |
|---|---|---|
| Combination 1 | Shape | A1.1 |
| Combination 2 | Shape + Size | A1.2 |
| Combination 3 | Shape + Fluorescence | A1.3 |
| Combination 4 | Shape + Size + Fluorescence | A1.4 |
| Combination 5 | Size + Fluorescence | A1.5 |

# A1.2 Discussion of results

## A1.2.1 Combination 1: classification based on Shape

### A1.2.1.1 Spherical group

As would be expected all four algorithms performed well for the spherical group classification, with generally better than 90% correct classification being achieved. An indication as to the sensitivity of the different methods to the influence of noise in the data is given by the fall off in performance seen for the 1μm PSL (class 2). The Norm Dist correctly classified only ~35% of this class to the spherical group.

Looking in more detail at the classification into classes it can be seen that the high quality data from oleic acid droplets results in a virtually 100% correct classification by class, irrespective of the algorithm, reflecting the relative simplicity of the task. Turning to the results for the 1 μm PSL the percentage correctly classified ranges from 28% for Norm Dist to 37% by KNN. The class that has the next largest tranche of the data is NaCl (class 8) with more than 20% of the 1 μm PSL allocated to it.

The noisier data at the smaller size makes the scattering data appear similar to that from the more cubic NaCl. This effect has already been observed in the *Af* plots for PSL shown in Chapter 7.

The polydisperse water and the silica are particularly successfully classified by the Lin Disc with 99% correct by group for the water. For this data combination i.e. where the sample size distribution cannot influence the outcome, the different algorithms show quite different levels of classification across the four classes in the spherical group. Lin Disc, for example, places 99% of the water correctly in the spherical group but only 0.25% are actually classified as water.

### A1.2.1.2 Non-spherical group

Given the more complex morphology of these materials there is a general decline in the percentages classified correctly when compared to the spherical results. All the algorithms achieve similar levels of classification with the Lin Disc performing best overall. All four methods classify the haematite to within a few percentage points of 85% to the group. For the NaCl on the other hand the best that is achieved is 51% by Lin Disc, with Norm Dist managing only 33% correct. For the individual classes in this group no single algorithm manages to achieve high levels of classification for all of the materials.

Generally the balance of the classification is spread over the entire range of the remaining classes[a], including in some cases the spheres. This reflects the real overlap in the shape of many of these materials, and gives an indication of the limitations of classifying on the basis of shape alone.

### A1.2.1.3 Fibre group

This group contains only two particle types: gypsum and caffeine. In both cases the Lin Disc performs poorly and the results from this algorithm are not considered further. With the other techniques a minimum of 60% of the gypsum is classified correctly, while for caffeine over 72% of

a This contrasts with the spherical group where often misclassifications are grouped into one or two classes. See the data for 1μm PSL as discussed earlier.

the test data is classified correctly with no significant differences between the classification methods. None of the algorithms miss classify the data as spherical.[b] The balance of the output is fairly evenly spread across all other classes. Considering the class by class results generally the largest single percentage is assigned to the correct class, with the caffeine being more accurately classified at this level than the gypsum, this no doubt reflecting the greater spread in the gypsum particle shape and size as seen in the *Af* data of Chapter 7.2.

#### A1.2.1.4 Bacteria group

This group is made up of seven classes. For this group the Lin Disc algorithm has the worst performance for all data combinations and consequently will not be considered further in this section.

All other methods perform reasonably well although with quite a wide variation between the different classes. The %s classified correctly ranges from 54% for E.coli washed (class 14) to 82% correct for E.coli unwashed, (class 15). Within the group the highest percentage correctly classified also falls into the correct class, although there is a broad spread across all classes in the group.

#### A1.2.1.5 Pollen group

This group has been poorly classified by all methods. Again the Lin Disc performs particularly poorly with none of the pollen correctly classified. A preponderance of the pollen is misclassified to the Non-spherical group, with a large percentage classed as NaCl.

### A1.2.2 Combination 2: classification based on shape & size

#### A1.2.2.1 Spherical group

Again all methods achieve high levels of classification by group. Considering the more detailed classification by class results shows a noticeable improvement on the percentage correctly allocated to their respective class. For example, in the shape only results, 41% of the polydisperse water (class 3) is correctly classified; with the inclusion of the size term this figure increases to 55% for KNN. A similar pattern can be observed for the silica data. Overall for this group the inclusion of the additional size parameter gives a small improvement in the already high levels achieved on the basis of shape alone, with a larger improvement in the class data.

#### A1.2.2.2 Non-spherical group

For the haematite inclusion of the size parameter has resulted in an improvement in the classification irrespective of the algorithm. For example the percentage correctly classified goes up from 74% to 89% with KNN.

For the remaining materials in this group the situation is less clear cut. In 6 cases the classification is improved, although generally only by a small margin (4-7 %), while in another 5 cases the classification is worse, again by a small percentage. No change occurs in the balance of the cases.

#### A1.2.2.3 Fibres group

b For the sake of brevity fractional percentage values have been ignored throughout in order that some understanding of overall performance can be gained. So that for example the 0.25% of the gypsum data that has been classified by KNN as class 1 is disregarded.

The addition of the size parameter resulted in a small improvement in the efficiency of the classification. Fuzzy-K gave the best overall performance with 74% and 77% of the gypsum and caffeine respectively being correctly classified. Inclusion of the size parameter has also improved the classification by class for both materials. No single algorithm performed appreciably better than any other for this group, although overall the Lin Disc performed poorly .

#### A1.2.2.4 Bacteria group

The inclusion of the size parameter resulted in an improvement for the classification of all classes in the bacteria group. For example 87% of Bg spores are correctly classified to the bacteria group by the Fuzzy-K algorithm. Although no single algorithm performs best for all classes in the group, Fuzzy-K showed the best performance in 5 of the 7 classes. .

#### A1.2.2.5 Pollen group

The inclusion of the size parameter resulted in a significant improvement in the success rate for this classification. This is the only group where there is a clear cut divergence in the performance of the different algorithms can be detected. Norm Dist has the greatest success rate with ~58% accurately classified. Lin Disc does spectacularly badly with only 0.25% of the Puff ball being classified as pollen.

### A1.2.3 Combination 3: classification based on Shape & Fluorescence

With the exception of the larger spheres the Lin Disc performed poorly for all aerosol types and consequently will not be considered in the rest of the discussion.

#### A1.2.3.1 Spherical group

The results overall for this group are generally worse than with the shape or shape plus size data combinations. Previously all methods achieved near 100% correct classification for the 1 μm oleic acid. Inclusion of the fluorescence parameter resulted in only 34% of this sample being classified correctly. A more detailed examination of the results, however, is instructive as to the sensitivity of the different algorithms to each of the parameters.

Looking at the Silica (class 4) results both KNN and Fuzzy-K previously classified 95% of this particle type as spherical. Inclusion of the fluorescence parameter has resulted in this falling to ~35%. The majority of the data, 51% by KNN and 50% Fuzzy-K, has in fact been misclassified as NaCl (class 8). A similar although less pronounced shift has taken place in the results from Norm Dist with 17% of the data now being classified as NaCl as against 0.25% previously.

The results for both water and 1μm oleic acid show a similar pattern, with a reduction in the percentage classified correctly as spherical and a sizeable percentage misclassified as NaCl. In each of these cases this shift is most pronounced for both KNN and Fuzzy-K algorithms.

For 1μm PSL the Shape + Size combination gave a moderately good level of classification. For example KNN gets 55% correct. All the algorithms classify around 10% of this particle data as NaCl because of the effect of noise as discussed previously. When the Shape + Fluorescence parameter combination is used ~75% of the sample is now correctly classified by KNN & Fuzzy-K, with the percentage classified as NaCl falling to less than 2%. At the same time the percentage misclassified to the bacteria group has increased from 5% (combination 1 KNN) to 24%. This shift reflects the roughly equivalent level of intrinsic fluorescence of the bacteria and PSL. Again these changes are most pronounced with the KNN and Fuzzy-K algorithms.

| particle type | fluorescence channel value | | No particles | | |
|---|---|---|---|---|---|
| | | | logged | Fl. >500 | Fl.< 500 |
| | median | max | | | |
| NaCl | 309 | 1652 | 5590 | 320 | 5270 |
| water | 53 | 4088 | 26026 | 397 | 25706 |
| 1μm PSL | 1716 | 4091 | 4214 | 4066 | 148 |

**Table A1.6 Fluorescence recorded from the non-fluorescent 3.2 μm NaCl and polydisperse water compared to that recorded from 1μm PSL.**

A more detailed inspection of the data helps to explain these shifts in the distribution. In the case of water the values recorded on the fluorescence channel are similar in magnitude to those recorded for NaCl, as described in Chapter 7 and as summarised in Table A1.6. These values reflect the system noise on the fluorescence channel. Although these values are low for this channel they are large relative to the size normalised data values of the scattering profiles. In the case of the 1 μm PSL the magnitude of fluorescence is well above the noise and in this case the effect is to shift the classification towards particles exhibiting the higher levels of fluorescence.

The fact that this shift is most pronounced for KNN and Fuzzy-K suggests that these methods are particularly sensitive to the scaling of the data. Since it is not possible to normalise the fluorescence data this may be a limitation of these types of algorithms with this type of data.

### A1.2.3.2 Non-spherical group

#### Norm Dist

This shows an improvement over the previous input combinations. However the percentage correctly classified by class has tended to decrease. For example 74% of the magnetite is correctly classified to the Non-spherical group compared with 64% for the previous data combinations. However only 49% of the sample is correctly classified to the magnetite class, compared with 55% with the size & shape combination.

#### KNN & Fuzzy-K

Again these two algorithms have a similar level of performance, with both showing a marked improvement in the classification overall with the exception of haematite where their performance is worse. The results show a shift in the classification towards the NaCl. Presumably this reflects the low level of fluorescence characteristic of these particles.

For kaolin which has generally been poorly classified this combination shows a large improvement with 70% correctly classified (28% previous combinations). Why the inclusion of the fluorescence parameter has given this improvement is not clear. The kaolin itself has only low levels of fluorescence and no other materials in the group exhibit significant levels of fluorescence.

## A1.2.3.3 Bacteria group

### Norm Dist

The results for this group show a marked improvement in classification efficiency over previous formats. For example 93% of the Bg Spore data is correctly classified. In addition the largest classification for each class falls in its own class. So, for example, for *Erwina.h* washed ( class 16) 86% falls in the bacteria group, with 48% being classified as class 16.

### KNN & Fuzzy-K

In contrast to the results for Norm Dist the efficiency is poorer than that of the shape + size combination. For each class the percentage correctly classified is around 20% less than previous combinations. Figure A1.1 shows the characteristic pattern of misclassification that occurs. Note that often the largest single class to which the bacteria are allocated is NaCl. A further significant percentage is assigned to class 1 spheres. For comparison purposes a typical classification distribution is shown for the Shape + Size data combination format.

This outcome is unexpected and the reason for its occurrence is not apparent. It had been expected that intrinsic fluorescence exhibited by the bacteria would result in an improvement in the discrimination achieved between materials of otherwise similar morphology, such as the Non-spherical particles and the bacteria.
The fluorescence data from the bacteria shows high levels of intrinsic fluorescence being recorded. Table A1.7 shows that even for *Erwinia.h*, which has the lowest levels of fluorescence, the values recorded are well above the fluorescence channel noise threshold and indeed values logged from the 2 μm NaCl crystals.

Further more it might be expected that for those bacteria classified as spherical the fluorescence parameter would cause the classification to fall into the 1 μm PSL class, since this exhibits fluorescence levels similar to those of the bacteria. In practice virtually none of the bacteria are assigned to this class, while ~10% are classified as the weakly fluorescent 1 μm oleic acid (class 1).

These results indicate that these types of classifiers are not effective when handling this type of data.

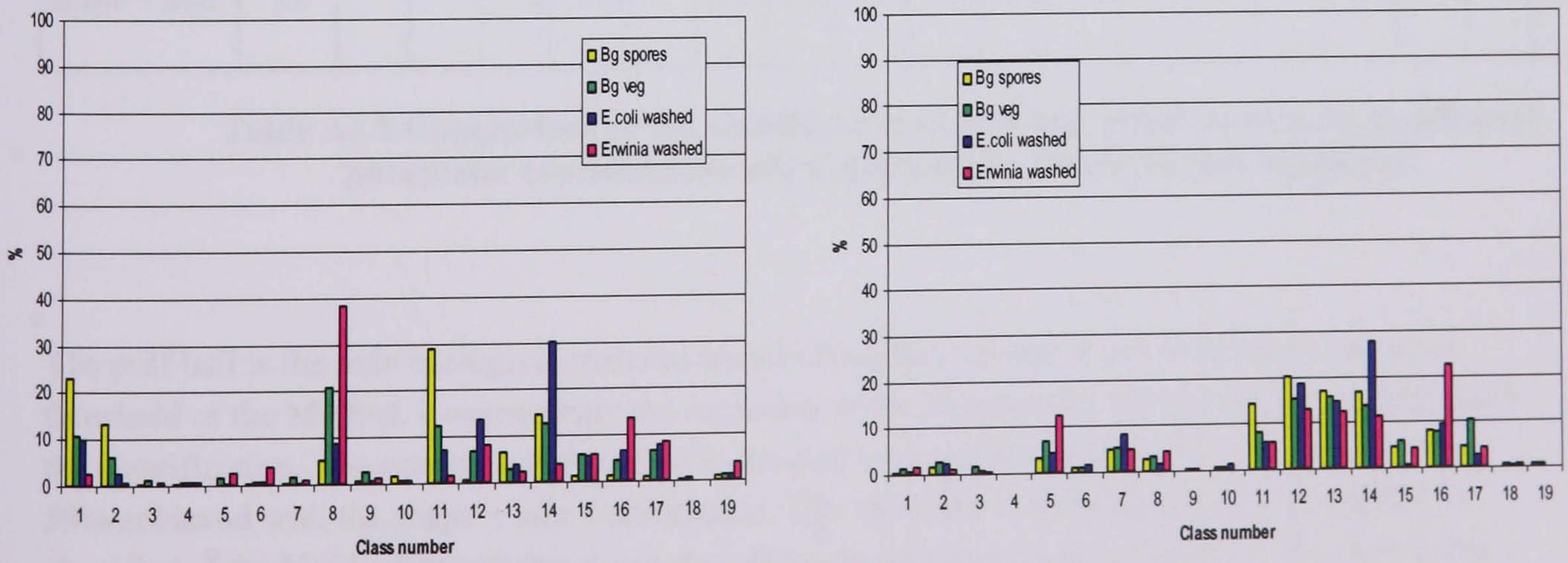



**Figure A1.1 The left hand plot shows the classification achieved by the KNN algorithm when using the Shape + Fluorescence parameter combination. The miss classification to NaCl (class8) is readily apparent. The right hand plot is the results from the Size + Shape data combination.**

| material | median fluorescence recorded | 3rd quartile |
|---|---|---|
| Bg spores | 938 | 1172 |
| E.coli | 615 | 765 |
| Erwinia | 496 | 597 |
| NaCl | 96 | 134 |

Table A1.7 Relative levels of fluorescence recorded from bacterial samples and NaCl.

### A1.2.3.4 pollen group

Norm Dist

The classification of Paper Mulberry is noticeably improved over previous data input / algorithm combinations with 60% correctly classified as a pollen. A further 30% of this sample is classified to the Non-spherical group. The classification of the sample into these two distinct groups represents a good indication of the real nature of the sample as collected, as discussed in chapter 6. The level of classification for puff ball, the other member of the pollen group, is worse than achieved previously.

| data input combination | group % | class % | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
| Shape only | 34 | | 3.5 | 0.5 | 1.25 | 1.5 | 8.25 | 4.25 | 23.5 | 7 | 10.8 | | 2 | 0.25 | 0.5 | 1 | 0.25 | 1.75 | 7.5 | 26.3 |
| Shape + fl | 30 | | 3.5 | | 1 | 1.3 | 11 | 1.8 | 26 | 8.3 | 7.5 | 1.5 | 2.8 | 0.3 | 0.3 | 1.5 | 0.5 | 2.8 | 2.5 | 28 |
| Shape + size | 59 | | | 1 | 2.5 | | 10.8 | 3.25 | 4.25 | 7 | 8.5 | 1 | 1.5 | 0.5 | | 0.5 | 0.25 | 0.5 | 17.8 | 40.8 |

**Table A1.8 Comparison of the classification of puff ball achieved with three different parameter combinations when classified by the Norm Dist algorithm.**

The puff ball is the only biological material tested which has fluorescence levels below the noise threshold of the MPAM. Consequently the inclusion of the fluorescence parameter has not improved the classification. The percentage classified to the pollen group has fallen to 30% compared with 59% achieved with the shape + size combination. The class into which the next largest fraction is classified is the NaCl, although this is not true when the size parameter is included. The figures are reproduced above in Table A1.8 for convenience from the relevant tables.

KNN & Fuzzy-K

The paper mulberry is poorly classified by these methods. Fuzzy-K only manages to classify 34% of the paper mulberry as pollen, an identical result to that achieved by the shape + size input format. The fluorescence parameter has resulted in a further 27% being classified as 1μm PSL, reflecting the effect of the high fluorescence values of some of this sample. The results for KNN are again very similar to those for Fuzzy-K.

For the puff ball these algorithms perform particularly badly, managing to classify only ~3% as pollen. As with the other data combinations the class that has the highest rate of classification is again NaCl.

## A1.2.4 Combination 4: classification based on Shape & size & fluorescence

With the possible exception of the spherical group the Lin Disc algorithm performs poorly with this data input combination. For the bacteria it has a success rate of around one third of the level achieved by the other methods. Its performance for the pollen group is even worse. Consequently the results for the Lin Disc algorithm are not given further consideration.

### A1.2.4.1 Spherical group

Norm Dist

The inclusion of all three parameters gives rise to a small improvement in the already high percentage of the data being correctly classified for this group.

KNN & Fuzzy-K

For the polydisperse water 100% is correctly allocated to this group, a marginal improvement on 98% correctly classified by both the shape and shape + size combinations, although the classification by class is significantly improved for both water and the silica class.

### A1.2.4.2 Non-spherical group

Norm Dist

The multi-parameter input results in improved levels of classification for all classes although overall does not perform quite as well as the KNN algorithms. The most significant improvement occurs with the NaCl class with 84% of the data correctly classified, compared with 65% for the next best combination (Shape +Fluorescence). The class by class classification also showed a 20% improvement.

KNN & Fuzzy-K

These algorithms achieve the best performance for this group, although the margin is typically only ~4% better than the Norm Dist results. Considering the class by class classification the results are more mixed. For haematite 92% is correctly classified compared with 82% for Norm Dist. For kaolin and NaCl however the classification is significantly poorer than that achieved by Norm Dist (24% correct with Fuzzy-K, 41% with Norm Dist).

### A1.2.4.3 Fibres group

All the algorithms perform well for this group, although the multi-parameter approach has not improved on the success rate shown by the simpler classification of shape + size. For both methods ~70% of the data is correctly classified. In the case of the caffeine fibre the pattern of misclassification with the Norm Dist is of interest. Norm Dist correctly classifies 73% of the caffeine to the fibre group. Of the remaining 27% of the sample 20% is classified as pollen.

If this is compared to the caffeine results from the shape + size combination it will be seen that the success rate of the classification is very similar. However in this case no more than 6% of the remaining 25% of the data falls into any single group. That this tendency to misclassify the caffeine into a single group is a consequence of the inclusion of the fluorescence parameter can be seen by considering the results of the Shape + Fluorescence combination where, although less significant, a similar pattern is starting to emerge.

The results for KNN are similar although the grouping that occurs within the misclassified part of the sample is less pronounced. In this case only 1% is classified as pollen. However 13% is classed as spherical (mainly to the fluorescent PSL class).

For caffeine the multi-parameter approach does not give rise to any real improvement in the correct classification of the sample. The inclusion of each additional parameter reinforces the tendency of part of the sample to be misclassified into a single group (as pollen by Norm Dist or spherical by KNN). For any instrument where a primary aim is to discriminate between biological and non biological particles, the significant level of misclassification shown by Norm Dist would be likely to trigger an alert of a high biological presence which in fact would be false.

### A1.2.4.4 Bacteria group

#### Norm Dist

In this case the inclusion of the fluorescence parameter gives an improvement in the classification over all other data combinations. A minimum of ~90% of the data is classified into the bacteria group. For the seven bacterial classes tested the largest percentage of each sample is correctly classified into its own class[c]. The highest classification performance is for *Erwinia* washed with 51% correctly classified to this class. Unlike for the fibre group, where inclusion of the fluorescence parameter has resulted in a large misclassification of caffeine into the pollen group, the reverse does not hold true. In most cases less than 1% of the bacteria are classed as caffeine. Perhaps most importantly in no case is there a significant misclassification of a bacterial material into one of the non biological classes. Figure A1.2 shows the results for *E.coli* which are typical of the group overall.

The comparative results from the washed and unwashed samples is indicative of the influence of the culturing media on the classification. For *Erwinia* unwashed 43% of the sample is correctly classified as *Erwinia* unwashed with the next most significant class being *E.coli* unwashed with 31%. Surprisingly only 4.5% is classified as *Erwinia* washed. This same pattern is repeated in the case of *E.coli* unwashed with only 2.5% classified as *E.coli* washed and 41% being classified as *Erwinia* unwashed.

The situation for the washed samples is similar. In the case of *Erwinia* washed 51% is correctly classified with only 4% classed as *Erwinia* unwashed. For *E.coli* washed only 1.25% is classed as *E.coli* unwashed, with significant percentages being classified as Bg. These results confirm the significance of preparation methods on the fluorescence profile of the bacteria as discussed in Chapter two. The improved ability of the multi parameter approach to discriminate these features is important in the ability of the instrument to detect bacteria generated from a common anthropogenic source.

c Except *E.coli* unwashed where slightly more is classed as *Erwinia*.


**Figure A1.2 The results of the Norm Dist classification of *E.coli* for the multi-parameter Shape + Size + Fluorescence data combination. Note the very low levels of misclassification to non bacterial classes.**

KNN & Fuzzy-K

The results for these two algorithms are similar to those described above, although on average the group classification is ~ 5% poorer than those achieved by Norm Dist. As the same general observations apply the discussion will not be repeated.

### A1.2.4.5 Pollen group

Norm Dist

These results show a marked improvement in the efficiency of the classification over those achieved with any previous input formats. Both the puff ball and paper mulberry are classified with reasonable accuracy of ~67% to the pollen group. The classification into classes has undergone an even greater improvement compared to the next best data input combination of Shape + Size. For both pollens 59% are classified to their own class, compared with 40% by the next best data combination. Less than 2% of the sample is mistakenly classified into any of the bacterial classes making it unlikely that pollens would be incorrectly reported as of bacterial origins. For both paper mulberry and puff ball the next largest class is kaolin (~11%) with 30% being assigned to the Non-spherical group overall. The fact that a large fraction of both these widely different pollens are placed in the same class further indicates that the original sample materials are likely to be made up of a mixture of at least two distinct components,. as discussed in Section 8.4.1.

KNN & Fuzzy-K

These methods perform poorly with this data input with only 28% of the sample data being correctly classified to the pollen group, noticeably worse than for the Shape + Fluorescence combination which achieved 37% for paper mulberry. As with Norm Dist ~24% are classified as Non-spherical and a further large percentage are classified as spheres (~29% in the case of paper mulberry). This compares with none being misclassified as spheres by Norm Dist.

## A1.2.5 Combination 5: classification based on Size & fluorescence

Once again, with the exception of the results for the spherical group, Lin Disc has the worst performance for every single class. No further consideration will therefore be given to the performance of this algorithm in this section.

### A1.2.5.1 Spherical group

#### Norm Dist

This algorithm performs poorly when compared to the other input data combinations. The 1μm PSL is particularly poorly classified (only 49% allocated to the spherical group). Over 15% of the sample is classed as caffeine fibres resulting from the similarity of the fluorescence signal from the PSL and caffeine. Of greater significance is the 34% of the sample assigned to the bacteria group. In particular 15% of the sample is classed as Bg veg, a result of both the similarity in the magnitude of the scattering and fluorescence recorded from Bg.

The results for the non fluorescent water are also instructive. A significant proportion of the water droplets are misclassified into one of the non-spherical groups. In total one quarter of the sample is classed as fibres of which 10% of is gypsum. The Non-spherical group has the largest single allocation with 39%: six percent more than is actually classed as spherical. Again this reflects the similarity in both the scattering size and fluorescent of the water droplets and the Non-spherical particles. The non inclusion of a shape parameter can clearly give rise to some significant confusions in the classification of some common types of aerosol particles.

#### KNN & Fuzzy-K

Both these methods have similar results to those of Norm Dist. The same confusion of PSL with the bacteria group and the classification of water as both fibres and ellipsoids is present.

### A1.2.5.2 Non-spherical group

#### Norm Dist

This group of materials is reasonably well classified with the percentage correctly classified typically a few percentage points below that achieved for the multi-parameter combination. A significant proportion of the NaCl is misclassified into the spherical group. Similarly 12% of the kaolin is classified as pollen.

#### KNN & Fuzzy-K

These results are marginally poorer than those for Norm Dist. The misclassification follows the same general pattern. For example 16% of both the magnetite and NaCl are classified as spherical, again demonstrating the disadvantage of not having shape information included in the classification.

### A1.2.5.3 Fibre group

#### Norm Dist

The efficiency of classification achieved is markedly different for the two materials in this group. The non-fluorescent gypsum is poorly classified with a quarter classed as spherical. The percentage of caffeine correctly classified (78%) is better than any other data combination. Of equal significance is the reduction in the percentage misclassified as pollen, compared to the multi-parameter combination. This is illustrated in Figure A1.3 where the results for this and the multi-parameter data combinations are plotted. The probability of the presence of biological material being falsely reported from the presence of caffeine like fibres is reduced by this data combination.

KNN & Fuzzy-K

These algorithms achieve poor levels of classification for these aerosol types especially compared with multi-parameter combination.

## A1.2.5.4 Bacteria group

Norm Dist

The efficiency of classification achieved is poorer than achieved by the multi-parameter approach. For example 86% of Bg class 11 is placed within the bacteria group compared to 97% by the multi-parameter combination. In the case of *Erwinia* only 24% of the test aerosol was recognised as belonging to this class, with a further 20% classified as haematite. Finally *Erwinia* (washed) is poorly classified. Again this is probably a result of the lower intrinsic fluorescence of this particle type and the resulting confusion between *Erwinia* and particles of a similar size, such as haematite and magnetite.

KNN & Fuzzy-K

For all the data combinations reported so far the performance of KNN and Fuzzy-K have followed each other to within two or three percentage points. Consequently the results for these two algorithms have been treated together. For this group a clear difference in the efficiency of class between the two algorithms appears. The results for KNN are all close to those for Norm Dist. The Fuzzy-K results however are worse than for any of the other input combinations with the classification falling across all groups.

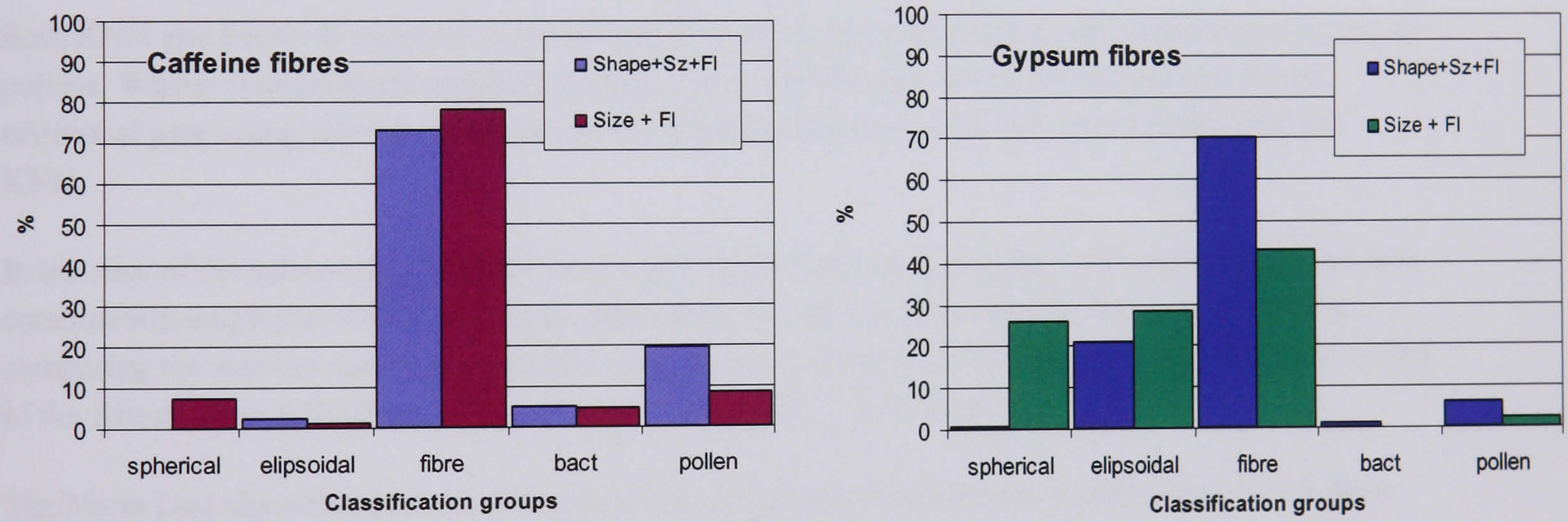



**Figure A1.3 Results of the classification of fibres by the Norm Dist algorithm when applied to the Shape + Size + Fluorescence and the Size + Fluorescence data combinations. Note the 20% of the caffeine misclassified as pollen by Shape + Size + Fluorescence data combination and the misclassification of gypsum to the Non-spherical and spherical groups by the Size + Fluorescence combination**

### A1.2.5.5 pollen group

#### Norm Dist

The level of classification achieved is inferior to that achieved by the multi-parameter approach being roughly equivalent to the levels of discrimination achieved by the other dual parameter input combinations. The paper mulberry spores are particularly poorly classified with half the sample classed as either fibres or Non-spherical.

#### KNN & Fuzzy-K

As in the bacteria group these algorithms show a distinct difference in the efficiency of discrimination achieved by the two approaches, the success rate of Fuzzy-K being again less than half that for KNN. For both cases the classification is too poor to be of practical use.

## A1.3 Summary

The detailed consideration of the data shows that no single algorithm or data combination performs the absolute best in every single case. The efficiency of the classification when undertaken by the Lin Disc algorithm shows little improvement between that of the single parameter input of shape and any of the other input combinations. In some cases Lin Disc had a worse performance with the Size + Fluorescence parameter combination than it managed with the shape parameter alone.

For Norm Dist, KNN and Fuzzy-K however, the accuracy of classification improves with the number of parameters. This general trend can be seen in the contour plots of Chapter 8.

Both KNN and Fuzzy-K succeed in classifying all particle types reasonably well with the exception of pollens. Where their performance does diverge Fuzzy-K performs worse. Given that it involves additional processing there is no advantage to using this approach over the more straight forward KNN.

In the case of the spherical group KNN and Fuzzy-K perform exceptionally well regardless of the data combination employed. Of the fifteen possible cases for oleic acid, 1 μm PSL and water reported, comparing the success rate of the classification for Norm Dist and KNN/Fuzzy-K shows that in twelve of the fifteen cases KNN/Fuzzy-K achieves higher levels of classification.

The Norm Dist algorithm performed in general as well as the KNN/Fuzzy-K algorithms for all other classes. In particular, for those aerosol particles exhibiting more complex morphologies or size distribution (polydisperse water) Norm Dist out performs the other algorithms. When used with the multi- parameter data it is the only algorithm that succeeds in classifying the pollens with reasonable accuracy. In addition, the Norm Dist algorithm has the advantage of far faster execution speed than either of the KNN variants.

Over the range of aerosol materials investigated, the Norm Dist achieves the most efficient levels of classification with the lowest computational overheads.

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| oleic 1u | Norm Dist | 92.25 | 30.25 | 59.75 | 2.25 | | | | | 0.25 | | | | | | | 4.25 | | 2.5 | | 0.75 |
| Class 1 | Lin Discrim | 100 | 81.5 | 18.5 | | | | | | | | | | | | | | | | | |
| | KNN | 100 | 99.5 | 0.5 | | | | | | | | | | | | | | | | | |
| | Fuzzy K | 100 | 99.75 | 0.25 | | | | | | | | | | | | | | | | | |
| oleic 3u | Norm Dist | 100 | 100 | | | | | | | | | | | | | | | | | | |
| Class 1 | Lin Discrim | 100 | 100 | | | | | | | | | | | | | | | | | | |
| | KNN | 100 | 99.5 | | 0.5 | | | | | | | | | | | | | | | | |
| | Fuzzy K | 100 | 99.75 | | 0.25 | | | | | | | | | | | | | | | | |
| PSL 1u | Norm Dist | 35.5 | 1.75 | 28.25 | 4.75 | 0.75 | | 9.75 | 4 | 20.75 | 10.5 | 0.25 | 0.5 | 0.5 | 1 | 0.25 | 1 | 0.25 | 1.75 | 3.75 | 10.25 |
| Class 2 | Lin Discrim | 62.75 | 32 | 30.75 | | | 1 | 0.25 | 1 | 22 | 1.25 | 4.25 | | 1 | 0.25 | 0.25 | 3.75 | 0.5 | 1.75 | | |
| | KNN | 48 | 10.25 | 36.75 | 1 | | 0.75 | 4.5 | 3 | 23.75 | 1.5 | 3.75 | 0.5 | 0.25 | 0.75 | 1.25 | 1.25 | | 0.75 | 3.75 | 6.25 |
| | Fuzzy K | 48 | 11.25 | 35.75 | 0.5 | 0.5 | 1 | 3 | 2.5 | 27 | 2 | 4.25 | 0.5 | 0.25 | 1 | 1.25 | 0.75 | | 1 | 3.25 | 4.25 |
| PSL 1.7u | Norm Dist | 90.25 | 2 | 69.25 | 16.75 | 2.25 | 0.25 | | | 2.25 | | | 0.25 | | 0.25 | | 1.25 | 0.25 | 2.75 | 0.5 | 2 |
| Class 2 | Lin Discrim | 99.75 | 69.75 | 30 | | | 0.25 | | | | | | | | | | | | | | |
| | KNN | 97.75 | 33.5 | 57.5 | 5.25 | 1.5 | | | | 0.75 | | | | 0.5 | | | | | 0.25 | 0.5 | 0.25 |
| | Fuzzy K | 97.75 | 45.5 | 50.25 | 1.25 | 0.75 | 0.25 | | | 1.5 | | | | | | | | | 0.25 | 0.25 | |
| PSL 3u | Norm Dist | 98.75 | 12.5 | 85.75 | | 0.5 | | 0.25 | | | | | | | | | 0.25 | | 0.5 | | 0.25 |
| Class 2 | Lin Discrim | 100 | 75.25 | 24.75 | | | | | | | | | | | | | | | | | |
| | KNN | 99.75 | 6.75 | 92.75 | 0.25 | | | | | 0.25 | | | | | | | | | | | |
| | Fuzzy K | 99.5 | 7.5 | 92 | | | | | | 0.25 | | | | | | | | | | | 0.25 |
| Water poly | Norm Dist | 86 | 46.5 | 12.75 | 24.5 | 2.25 | | | | 4.25 | | | | | | | 1.5 | | 4.25 | | 4 |
| Class 3 | Lin Discrim | 99.25 | 90.25 | 8.75 | 0.25 | | | | | | | | | | | | 0.5 | | 0.25 | | |
| | KNN | 96.5 | 43.25 | 10.75 | 40.75 | 1.75 | | | | 1.5 | | | | | | | 1.5 | | 0.25 | | 0.25 |
| | Fuzzy K | 97.5 | 53.25 | 11.25 | 31.75 | 1.25 | | | | 1 | | | | | | | 0.75 | | 0.5 | | 0.25 |
| Silica | Norm Dist | 75.5 | 2.5 | 2.5 | 9.5 | 61 | 0.75 | 2.25 | | 12.5 | 0.5 | | 0.25 | 0.25 | 0.5 | | | 1 | 0.25 | 3.25 | 3 |
| Class 4 | Lin Discrim | 87.75 | 86.25 | 1.5 | | | 1 | | | 10.5 | | 0.5 | | | | | | | 0.25 | | |
| | KNN | 82.75 | 30.75 | 9.5 | 9.25 | 33.25 | | | 0.25 | 13.25 | | | | | 0.25 | | 0.5 | 0.5 | | 1.5 | 1 |
| | Fuzzy K | 83.25 | 42 | 8.5 | 7.25 | 25.5 | 0.25 | | 0.25 | 13.75 | | | | | 0.25 | | 0.25 | 0.25 | | 1 | 0.75 |
| Haem | Norm Dist | 81.75 | | | 0.25 | | 69 | 4.75 | 4.75 | 3.25 | 2.5 | 0.25 | 2.75 | 0.5 | 3 | 1 | | 2.25 | 0.25 | 4 | 1.5 |
| Class 5 | Lin Discrim | 86 | 1.25 | 8.75 | | | 62.25 | | | 23.75 | 1.5 | 2.5 | | | | | | | | | |
| | KNN | 84.75 | | | 0.5 | | 74.75 | 1.5 | 4.25 | 4.25 | 2.75 | 0.25 | 3.25 | | 1.75 | 2.25 | | 1.75 | 0.25 | 0.75 | 1.75 |
| | Fuzzy K | 87.75 | | | 0.5 | | 76.75 | 1 | 4.5 | 5.5 | 4 | 0.25 | 1.5 | | 1.25 | 1 | | 0.75 | 0.25 | 0.75 | 2 |

**Table A1.1 Combination one: shape data**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Kaolin | Norm Dist | 38.5 | | | | 0.5 | 1.5 | 17 | 5.75 | 14.25 | 17.25 | 5.75 | 6.5 | 5.5 | 3.25 | 5.25 | 0.25 | 3.25 | 0.5 | 7.75 | 5.75 |
| Class 6 | Lin Discrim | 47.75 | 14.25 | 7.5 | | | 4.75 | | 1.5 | 41.5 | 4 | 19 | 0.25 | 2.75 | | 1.5 | 2 | 0.5 | 0.5 | | |
| | KNN | 38.25 | 1 | 0.5 | | 0.75 | 3.25 | 9 | 3.25 | 22.75 | 8 | 8.75 | 4.5 | 5.5 | 5.25 | 5 | 0.75 | 3 | 2 | 6.5 | 10.25 |
| | Fuzzy K | 37.25 | 2 | 1.25 | 0.25 | 1 | 3.25 | 8 | 3 | 23 | 11 | 10.75 | 4.5 | 4 | 4.75 | 5.25 | 0.75 | 3 | 2.25 | 4.75 | 7.25 |
| Magnetite | Norm Dist | 67 | | | | | 6.25 | 4.75 | 49.75 | 6.25 | 10.25 | 3 | 1.25 | 1.5 | 1.75 | 1 | 0.25 | 0.75 | 1.5 | 3.75 | 8 |
| Class 7 | Lin Discrim | 56.75 | 12 | 11.25 | | | 18.25 | | 2.5 | 36 | 3 | 7.75 | | 1.5 | 0.25 | 0.25 | 4 | 1.5 | 1.75 | | |
| | KNN | 52 | 4 | 2.75 | 0.5 | 0.5 | 14.5 | 1.5 | 25.5 | 10.5 | 5.25 | 3.5 | 3.75 | 7.25 | 2.25 | 7 | 0.5 | 0.5 | 2 | 2.25 | 6 |
| | Fuzzy K | 52.75 | 6 | 2.5 | 0.5 | 0.5 | 16.25 | 1.75 | 25 | 9.75 | 5 | 3.75 | 2 | 8 | 2 | 6.75 | 0.5 | 0.25 | 2 | 1.5 | 6 |
| NaCl 2u | Norm Dist | 32.25 | | | | 1.25 | 1 | 12.75 | 2 | 16.5 | 1.25 | 0.25 | 7.25 | 6 | 11.5 | 5.75 | 0.5 | 19.5 | 0.75 | 11.75 | 2 |
| Class 8 | Lin Discrim | 51.5 | 19.5 | 15.75 | | | 6.75 | 0.5 | 2.75 | 41.5 | 1 | 4 | | 1 | | 1 | 4 | 1 | 1.25 | | |
| | KNN | 45.75 | 2.5 | 2.5 | | 1.25 | 5 | 6.75 | 1.5 | 32.5 | 1 | 1 | 3.25 | 3.5 | 6.5 | 3.5 | 2.25 | 6.25 | 3.5 | 11 | 6.25 |
| | Fuzzy K | 47 | 4.25 | 2.25 | | 0.5 | 5.25 | 4 | 2.75 | 35 | 0.75 | 1.25 | 3.5 | 2.75 | 6 | 5.25 | 2.25 | 5 | 5 | 7.5 | 6.75 |
| NaCl 3u | Norm Dist | 62 | | 1.75 | | 0.5 | 0.75 | 6.75 | 3.75 | 50.75 | 0.25 | 0.5 | | 1.25 | 0.25 | | 0.5 | | 3.25 | 9.5 | 20.25 |
| Class 8 | Lin Discrim | 25.5 | 55 | 16.5 | | | 2.25 | | | 23.25 | | 1 | | 0.25 | | | 1 | | 0.75 | | |
| | KNN | 55.75 | 13.5 | 12.75 | 1.25 | 8 | 0.25 | 1.75 | 0.75 | 53 | | | | | | 0.25 | 0.5 | | | 3 | 5 |
| | Fuzzy K | 50.5 | 19.5 | 14.75 | 0.5 | 6.5 | 0.5 | 1.5 | 1 | 47.5 | | 0.25 | 0.75 | | | 0.25 | 0.75 | | | 2.75 | 3.5 |
| Gypsum | Norm Dist | 63.75 | | | | 0.5 | 2.75 | 5.75 | 11 | 1.25 | 21.75 | 42 | 1.25 | 1.25 | 2.5 | 5.25 | | 1.25 | | 2.5 | 1 |
| Class 9 | Lin Discrim | 47.5 | 2 | 1.75 | | | 4.75 | | 0.25 | 39.75 | 14.25 | 33.25 | | 2.75 | | | 1 | | 0.25 | | |
| | KNN | 59.75 | 0.25 | 0.25 | | 0.5 | 3.75 | 3 | 7 | 2.5 | 34 | 25.75 | 3.5 | 2.75 | 4.25 | 6.75 | 0.25 | 2.25 | 0.25 | 0.5 | 2.5 |
| | Fuzzy K | 61.5 | 0.5 | 0.5 | | 0.25 | 4.25 | 2.5 | 6.5 | 2.75 | 33.75 | 27.75 | 1.5 | 4 | 4.5 | 6.5 | 0.5 | 2 | | 0.25 | 2 |
| Caffeine | Norm Dist | 77.5 | | | | | 2.75 | 3.75 | 7 | 1.75 | 10.25 | 67.25 | 0.25 | 0.5 | 0.5 | 1.75 | | 0.25 | 0.25 | 2.75 | 1 |
| Class 10 | Lin Discrim | 60.75 | 1 | 1.75 | | | 2.25 | | 0.5 | 30.5 | 15 | 45.75 | | 2.5 | | | 0.25 | 0.5 | | | |
| | KNN | 72.75 | | | | | 1.5 | 1.75 | 5.25 | 1.75 | 12.75 | 60 | 2.5 | 1.5 | 3 | 4.5 | 0.25 | 0.75 | 0.25 | 0.75 | 3.5 |
| | Fuzzy K | 74.5 | | | | | 1.5 | 1.25 | 6 | 2 | 12.25 | 62.25 | 0.75 | 1.5 | 2 | 6 | 0.25 | 0.25 | 0.25 | 0.75 | 3 |
| Bg spores | Norm Dist | 64.5 | | | | | 3.5 | 6.25 | 7.75 | 2.5 | 11.25 | 1.25 | 9.25 | 19.25 | 10.25 | 10.5 | 1.25 | 12 | 2 | 1.5 | 1.5 |
| Class 11 | Lin Discrim | 22.25 | 2.75 | 10.25 | | | 7.5 | | 2.25 | 32.25 | 6.75 | 16 | | 5.25 | | 1 | 10 | 1 | 5 | | |
| | KNN | 58 | 0.25 | 0.25 | 0.25 | 0.25 | 5.5 | 4 | 5.25 | 6 | 8.5 | 4.25 | 7.25 | 17.75 | 11.5 | 9.5 | 2.75 | 5.25 | 4 | 3.25 | 4.25 |
| | Fuzzy K | 54.25 | 0.25 | 0.25 | | | 6.75 | 4.25 | 5.75 | 6.25 | 9.25 | 5.75 | 6.25 | 17.75 | 9.25 | 9 | 3 | 5 | 4 | 2.5 | 4.75 |
| Bg sp | Norm Dist | 59.75 | | 0.25 | 0.25 | | 3.75 | 5.5 | 12 | 3.5 | 9.5 | 1.5 | 11 | 19.5 | 5 | 9 | 1.25 | 8.5 | 5.5 | 2.25 | 1.75 |
| clusters | Lin Discrim | 25 | 4.25 | 11.25 | | | 6 | 0.25 | 2 | 34 | 4.25 | 13 | | 6.5 | | 1 | 10.5 | 0.5 | 6.5 | | |
| Class 12 | KNN | 67 | 1 | 0.25 | 0.25 | 0.75 | 4 | 3.75 | 4.25 | 5.25 | 6.5 | 2 | 11.5 | 25.25 | 7 | 9.25 | 4.5 | 6.5 | 3 | 0.75 | 4.25 |
| | Fuzzy K | 68 | 1.25 | 0.25 | | 0.75 | 4.5 | 2.75 | 4.25 | 6 | 6.75 | 2.5 | 9 | 28.25 | 6.75 | 9.75 | 4 | 7 | 3.25 | 0.25 | 2.75 |

**Table A1.1 Combination one: shape data (continued)**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bg veg | Norm Dist | 60.25 | | 0.5 | | 0.25 | 2.5 | 4.75 | 7.5 | 5.75 | 8 | 2.25 | 6.75 | 9 | 11.5 | 8.75 | 4.75 | 12.5 | 7 | 6.25 | 2 |
| Class 13 | Lin Discrim | 24.5 | 5 | 13.75 | | | 11.25 | 0.25 | 2.5 | 26.5 | 4.25 | 12 | | 3.25 | | 2.75 | 9 | 1 | 8.5 | | |
| | KNN | 59 | 1 | 1.75 | | 1 | 5.75 | 3.5 | 4 | 7.5 | 10.25 | 2 | 6 | 8.75 | 13.5 | 11.75 | 4.25 | 5.75 | 9 | 2 | 2.25 |
| | Fuzzy K | 58.5 | 1.75 | 1.5 | | 0.75 | 6.25 | 2.75 | 3.5 | 7.75 | 10 | 3.25 | 5.25 | 9.25 | 14.25 | 11.5 | 5.25 | 5 | 8 | 2.25 | 1.75 |
| Ecoli | Norm Dist | 54.75 | | 0.25 | 0.25 | | 3.25 | 5.5 | 8.5 | 2.75 | 18 | 4 | 7.25 | 9.5 | 9 | 17 | 0.5 | 9.25 | 2.25 | 1.75 | 1 |
| washed | Lin Discrim | 16.5 | 1.5 | 8.5 | | | 11.5 | | 1.5 | 39 | 4.5 | 17 | | 4.75 | 0.5 | 1.25 | 5.5 | 2 | 2.5 | | |
| Class 14 | KNN | 57.5 | 0.5 | 0.75 | | 0.5 | 3.75 | 3 | 6 | 4 | 11.75 | 9 | 6 | 10 | 13 | 17.5 | 0.5 | 8.5 | 2 | 1.5 | 1.75 |
| | Fuzzy K | 54.5 | 0.75 | 0.75 | | 0.5 | 6 | 2.25 | 5.25 | 3.5 | 13.25 | 10.25 | 3 | 11 | 12.25 | 18.5 | 0.25 | 7.75 | 1.75 | 1.25 | 1.75 |
| E coli | Norm Dist | 82 | | 2 | 1.5 | 0.25 | 0.25 | 2 | 0.5 | 3.25 | 2 | 0.25 | 2.75 | 3.5 | 0.75 | 1.75 | 38 | 1.75 | 33.5 | 2 | 4 |
| unwashed | Lin Discrim | 63 | 16.75 | 9.75 | | | 0.75 | | 0.75 | 4.25 | 1.25 | 3.5 | | 0.25 | | | 29.25 | 0.25 | 33.25 | | |
| Class 15 | KNN | 80.25 | 8 | 2.75 | 1.25 | 1 | 1 | 0.5 | 1 | 1.5 | 1.25 | 0.75 | 2 | 2.75 | 2 | 1.5 | 34 | 1.75 | 36.25 | | 0.75 |
| | Fuzzy K | 78.5 | 9 | 2.75 | 1.25 | 1 | 1 | 0.75 | 1.25 | 1.25 | 1.75 | 1 | 1.75 | 2.25 | 1 | 2.25 | 32.5 | 1.75 | 37 | 0.25 | 0.25 |
| Erwinia | Norm Dist | 78.5 | | 0.25 | | | 2.25 | 5 | 0.75 | 3.75 | 2.25 | 0.5 | 6 | 11 | 10 | 7.5 | 3 | 37.5 | 3.5 | 5.75 | 1 |
| washed | Lin Discrim | 34.5 | 3.75 | 16.5 | | | 12.5 | | 2.25 | 20.25 | 1.75 | 8.5 | | 1.75 | 0.25 | 4.75 | 15.5 | 6 | 6.25 | | |
| Class 16 | KNN | 68.5 | 0.25 | 0.75 | 0.25 | 0.5 | 5 | 2.25 | 3.75 | 8.75 | 4.25 | 1 | 6.25 | 13 | 14.25 | 5.75 | 3.75 | 21.5 | 4 | 3.25 | 1.5 |
| | Fuzzy K | 67.75 | 0.5 | 0.75 | 0.25 | 0.5 | 5.5 | 2.25 | 3 | 11 | 3.5 | 2 | 4.5 | 13.75 | 15 | 6.75 | 4.25 | 19.75 | 3.75 | 2.5 | 0.5 |
| Erwinia | Norm Dist | 73.75 | | 2.5 | 1.75 | 1.75 | 0.75 | 2.25 | 1.5 | 7.25 | 0.75 | | 1.5 | 2.25 | 1 | 0.75 | 30.75 | 3.75 | 33.75 | 3 | 4.75 |
| unwashed | Lin Discrim | 58.5 | 19.25 | 15.25 | 0.25 | | 0.75 | | 1.5 | 2.5 | 0.5 | 1.5 | | 0.25 | | 0.5 | 30.75 | 0.5 | 26.5 | | |
| Class 17 | KNN | 74.5 | 7.5 | 3.5 | 2.25 | 2 | 0.75 | 1 | 1.75 | 4.75 | 1 | 0.25 | 0.75 | 3.75 | 2.25 | 1 | 32 | 2.5 | 32.25 | | 0.75 |
| | Fuzzy K | 74.5 | 8.25 | 3.25 | 2 | 1.75 | 0.75 | 0.75 | 1.75 | 5 | 1.25 | 0.25 | 1.25 | 3.5 | 1.75 | 1 | 31 | 2.5 | 33.5 | | 0.5 |
| P mullbery | Norm Dist | 27 | | 0.25 | | 1.25 | 3.5 | 11.5 | 8.25 | 22 | 6.25 | 3.25 | 2.75 | 3.75 | 3.5 | 2.5 | 1 | 2 | 1.25 | 14 | 13 |
| Class 18 | Lin Discrim | 0 | 22 | 12.25 | | | 7.75 | | 0.75 | 43.25 | 1.5 | 6.25 | | 1.5 | | | 3 | | 1.75 | | |
| | KNN | 15 | 3.75 | 4 | 0.25 | 3.5 | 4.5 | 6 | 3.75 | 34 | 2.5 | 3.5 | 5 | 4 | 4 | 1.75 | 1.5 | 2.25 | 0.75 | 8.25 | 6.75 |
| | Fuzzy K | 11.25 | 5 | 4.5 | | 3.5 | 5.5 | 4.5 | 3.25 | 35.75 | 3.25 | 4 | 4 | 4.25 | 4.5 | 2 | 2 | 1.75 | 1 | 5.5 | 5.75 |
| Puff ball | Norm Dist | 33.75 | | 3.5 | 0.5 | 1.25 | 1.5 | 8.25 | 4.25 | 23.5 | 7 | 10.75 | | 2 | 0.25 | 0.5 | 1 | 0.25 | 1.75 | 7.5 | 26.25 |
| Class 19 | Lin Discrim | 0 | 42.75 | 15.25 | 0.25 | | 2.75 | | | 23.5 | 2 | 9.5 | | 1.25 | | | 1.75 | | 1 | | |
| | KNN | 15 | 14.5 | 11 | 1.5 | 6.25 | 0.75 | 3.75 | 1.75 | 29.75 | 3.75 | 4 | 0.75 | 2.75 | 0.25 | 1.75 | 1 | 0.5 | 1 | 2.75 | 12.25 |
| | Fuzzy K | 11.75 | 19.5 | 10.25 | 0.75 | 5 | 0.75 | 4.25 | 2 | 30.25 | 3.5 | 4.75 | 0.75 | 2.5 | 0.25 | 1.75 | 1 | 0.25 | 0.75 | 1.5 | 10.25 |
| | Class: | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |

**Table A1.1 Combination one: shape data (continued)**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| oleic 1u | Norm Dist | 93 | 91.25 | 1.75 | | | | | | | | | | | | | 5 | | 1.25 | | 0.75 |
| Class 1 | Lin Discrim | 99.5 | 17.75 | 81.75 | | | | | | | | | | | | | | | 0.5 | | |
| | KNN | 100 | 99 | 1 | | | | | | | | | | | | | | | | | |
| | Fuzzy K | 100 | 98.75 | 1.25 | | | | | | | | | | | | | | | | | |
| PSL 1u | Norm Dist | 41.75 | 0.75 | 37 | 3.5 | 0.5 | | 10.75 | 7 | 9.25 | 1.25 | 0.25 | 1.5 | 0.75 | 1 | 0.25 | 1.75 | 0.25 | 0.75 | 4.25 | 19.25 |
| Class 2 | Lin Discrim | 68.5 | 2 | 66.5 | | | 4 | 0.25 | 1 | 9.25 | 2 | 4.25 | | 1 | 0.5 | 0.25 | 5 | 0.75 | 3.25 | | |
| | KNN | 56.75 | 4.5 | 50.75 | 1.5 | | 0.5 | 0.75 | 5.5 | 10.5 | 0.25 | 1.75 | 3.25 | 1.25 | 1.75 | 2.75 | 2 | 0.5 | 1 | 1.25 | 10.25 |
| | Fuzzy K | 57 | 5.5 | 50.25 | 1.25 | | 1 | 0.25 | 5.5 | 10.5 | 0.25 | 2 | 3 | 1.25 | 2 | 3 | 2 | 0.75 | 1 | 0.5 | 10 |
| Water poly | Norm Dist | 82.5 | 2.25 | 9.25 | 70.25 | 0.75 | | 0.25 | | | | | | | | | 3.25 | | 7 | | 7 |
| Class 3 | Lin Discrim | 98.75 | 7.25 | 63.75 | 0.25 | 27.5 | | | | | | | | | | | 0.75 | | 0.5 | | |
| | KNN | 97.5 | 9.25 | 20.75 | 55.25 | 12.25 | | | | | | | | | | | 1.5 | | 1 | | |
| | Fuzzy K | 97.75 | 11.75 | 25.5 | 48.75 | 11.75 | | | | | | | | | | | 1 | | 1.25 | | |
| Silica | Norm Dist | 80.75 | | | 13.75 | 67 | | 6.25 | | 0.25 | 0.75 | | | | | | | | | 11.5 | 0.5 |
| Class 4 | Lin Discrim | 98.75 | 3.25 | 8.75 | | 86.75 | 0.25 | | | 0.25 | 0.75 | | | | | | | | | | |
| | KNN | 94.75 | 4.5 | | 21.75 | 68.5 | | 1.25 | | 0.5 | 0.5 | | | | | | | | | 2.75 | 0.25 |
| | Fuzzy K | 95.75 | 5.5 | | 20.5 | 69.75 | | 0.25 | | 0.5 | 0.5 | | | | | | | | | 2.75 | 0.25 |
| Haem | Norm Dist | 88.5 | | | 0.25 | | 79.5 | 4.25 | 2.75 | 2 | 0.5 | 0.25 | 0.75 | 1 | 2.25 | | | 0.75 | 0.25 | 2.75 | 2.75 |
| Class 5 | Lin Discrim | 81.5 | 0.75 | 13 | 0.5 | | 71.5 | | 0.25 | 9.75 | 1.25 | 1.75 | | 0.25 | | 0.5 | | 0.25 | 0.25 | | |
| | KNN | 91.5 | 0.25 | 1.5 | 0.75 | | 89.75 | | 1.75 | | | | 0.5 | 2 | 1.25 | 1 | 0.25 | 1 | | | |
| | Fuzzy K | 91.5 | 0.5 | 1.25 | 0.75 | | 90 | | 1.5 | | | | 0.75 | 1.75 | 1.25 | 1 | 0.25 | 0.75 | | | 0.25 |
| Kaolin | Norm Dist | 38.5 | | | | 0.5 | 0.25 | 25.5 | 5.75 | 7 | 9.5 | 6.25 | 3.75 | 4.25 | 2.75 | 4 | | 1.5 | | 18 | 11 |
| Class 6 | Lin Discrim | 24.5 | 6.75 | 12.5 | 0.5 | 16.75 | 7.25 | 0.75 | 2.75 | 13.75 | 8.25 | 19.75 | 0.5 | 3.75 | 0.25 | 2.75 | 1.75 | 0.5 | 1.25 | 0.25 | |
| | KNN | 28.25 | 0.75 | 0.75 | 0.5 | 7.75 | 1.5 | 8.25 | 5.25 | 13.25 | 6 | 8.5 | 5.75 | 4 | 7 | 6.75 | 0.25 | 3.25 | 1.5 | 14 | 5 |
| | Fuzzy K | 26 | 0.75 | 0.75 | 0.5 | 8.75 | 1.75 | 5.75 | 5.75 | 12.75 | 5 | 11.25 | 6 | 4.75 | 7 | 6.5 | 0.5 | 2.5 | 1.5 | 13.25 | 5 |
| Magnetite | Norm Dist | 64 | | | | | 1 | 4.75 | 55 | 3.25 | 5.25 | 2.25 | 1.75 | 2 | 1.75 | 2.5 | | 1.5 | 0.25 | 7 | 11.75 |
| Class 7 | Lin Discrim | 45 | 12.25 | 7.5 | 0.25 | 2.25 | 25.75 | | 11.25 | 8 | 6.25 | 8.75 | 0.25 | 3.25 | 0.25 | 1.5 | 6.5 | 2 | 3.75 | | 0.25 |
| | KNN | 52.5 | 0.75 | 0.75 | 0.75 | 1 | 5.75 | 0.5 | 36 | 10.25 | 0.75 | 0.5 | 5.25 | 9.25 | 3.75 | 6.5 | 0.25 | 1.5 | 0.75 | 6 | 9.75 |
| | Fuzzy K | 52.75 | 1.25 | 1.25 | 0.75 | 1.25 | 7.25 | 0.5 | 33.75 | 11.25 | 0.5 | 0.25 | 4.5 | 11.5 | 4.25 | 6.25 | 0.5 | 1 | 0.75 | 4.5 | 8.75 |

**Table A1.2 Combination two: shape + size data**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NaCl 2u | Norm Dist | 58 | | 0.25 | | 0.25 | | 3.5 | 2.25 | 52.25 | 0.25 | 0.5 | 10.75 | 3.5 | 8 | 5 | 0.25 | 8 | | 3.75 | 1.5 |
| Class 8 | Lin Discrim | 30.5 | 21.5 | 31.25 | | | 11.75 | | 3.25 | 15.5 | | 3.75 | | 2.75 | 0.5 | 1.25 | 4.25 | 2.5 | 1.75 | | |
| | KNN | 32 | 0.75 | 8.5 | 1 | | 1.75 | 1.75 | 3.75 | 24.75 | 0.25 | 0.25 | 9.75 | 6.25 | 8.25 | 5.5 | 2.5 | 9.25 | 2.25 | 0.25 | 13.25 |
| | Fuzzy K | 31 | 1.25 | 10 | 1 | | 2.25 | 0.75 | 3.75 | 24.25 | | 0.25 | 8.75 | 6 | 9.75 | 5.25 | 3.5 | 8 | 2.25 | 0.25 | 12.75 |
| Gypsum | Norm Dist | 69 | | | | 0.75 | | 9.75 | 4 | 1.5 | 36.5 | 32.5 | 0.25 | 1 | 1.25 | 0.5 | | 0.25 | | 9 | 2.75 |
| Class 9 | Lin Discrim | 62.25 | 0.25 | 2.5 | | 17 | 3.75 | 0.25 | 2.25 | 6.75 | 32.25 | 30 | | 3 | | 1.25 | 0.5 | | 0.25 | | |
| | KNN | 74.5 | 0.25 | 0.25 | 0.5 | 4.5 | | 4.25 | 2.75 | 1.25 | 47 | 27.5 | 0.5 | 1 | 2.5 | 2 | | 0.5 | | 4 | 1.25 |
| | Fuzzy K | 74.5 | 0.25 | 0.25 | 0.25 | 4.75 | | 3 | 2.5 | 1 | 43.5 | 31 | 0.25 | 1.25 | 3.5 | 2.25 | | 0.5 | | 5 | 0.75 |
| Caffeine | Norm Dist | 75.25 | | | | | 0.5 | 6.5 | 6 | 0.75 | 8.25 | 67 | 0.25 | | 0.25 | 1.75 | | 0.25 | | 5.25 | 3.25 |
| Class 10 | Lin Discrim | 71 | 0.25 | 2 | | 0.5 | 5.75 | 0.25 | 1.75 | 11.25 | 24.5 | 46.5 | | 3.25 | 0.25 | 2 | 1 | 0.5 | 0.25 | | |
| | KNN | 77 | | | | 0.25 | | 1 | 3.25 | 2 | 15 | 62 | 0.5 | 0.25 | 2.5 | 4.5 | | 0.25 | | 4.75 | 3.75 |
| | Fuzzy K | 77.25 | | 0.25 | | 0.25 | | 0.5 | 4 | 2.75 | 12.5 | 64.75 | 0.25 | 0.25 | 2 | 5.25 | | 0.5 | | 4.5 | 2.25 |
| Bg spores | Norm Dist | 76.75 | | | | | 0.75 | 2.25 | 7.5 | 4.75 | 4.25 | 1.25 | 13.75 | 23.25 | 10.25 | 14.25 | 1.25 | 11.25 | 2.75 | 1 | 1.5 |
| Class 11 | Lin Discrim | 34.5 | 4.75 | 14 | | | 9.5 | | 4 | 12.25 | 5.5 | 15.5 | | 10.75 | | 6.25 | 10.5 | 2 | 5 | | |
| | KNN | 85.25 | 0.25 | 1.5 | | | 3.25 | 1 | 5 | 2.75 | | 0.5 | 14.25 | 20.25 | 17 | 16.5 | 4.5 | 8.25 | 4.5 | | 0.5 |
| | Fuzzy K | 87.25 | 0.5 | 1 | | | 4.5 | 0.25 | 3.75 | 1.5 | | 0.75 | 15 | 23.75 | 15.5 | 15.5 | 5 | 8 | 4.5 | | 0.5 |
| Bg sp clusters | Norm Dist | 72.5 | | | 0.5 | | 2 | 1.75 | 11.5 | 5.5 | 3.5 | 0.75 | 9 | 30.25 | 3.75 | 11.25 | 1.75 | 11 | 5.5 | 0.75 | 1.25 |
| Class 12 | Lin Discrim | 40 | 5.5 | 12.5 | | | 7.5 | | 8.25 | 9.75 | 4.25 | 12.25 | 0.25 | 12.75 | | 6.5 | 11.5 | 1 | 8 | | |
| | KNN | 82.75 | 0.5 | 2.5 | 0.5 | | 4 | 0.5 | 5.5 | 2 | 0.5 | | 10.75 | 35.75 | 7.25 | 12 | 4.5 | 8.25 | 4.25 | | 1.25 |
| | Fuzzy K | 83.75 | 1.5 | 2.25 | 0.5 | | 5.25 | | 5.25 | 1.5 | | | 8.75 | 39.5 | 5.75 | 12.5 | 5.25 | 8.25 | 3.75 | | |
| Bg veg | Norm Dist | 66.5 | | 0.25 | 0.25 | | 2 | 3.5 | 8.5 | 7.25 | 5 | 1.25 | 8.75 | 13.75 | 11.75 | 9 | 5.75 | 10.5 | 7 | 3.25 | 2.25 |
| Class 13 | Lin Discrim | 29 | 7.75 | 17.25 | | | 15 | 0.25 | 6.75 | 9 | 4.5 | 10.5 | 0.5 | 4.75 | 0.25 | 5.75 | 8.75 | 1.5 | 7.5 | | |
| | KNN | 76.5 | 1.25 | 2.5 | 1.5 | | 6.75 | 1 | 5.25 | 3.25 | 0.25 | 0.5 | 8 | 15.25 | 15.5 | 13.75 | 5.75 | 7.75 | 10.5 | 0.75 | 0.5 |
| | Fuzzy K | 76.25 | 2.5 | 2.5 | 1.25 | | 8.25 | 0.5 | 5 | 2.5 | 0.25 | 0.25 | 7.5 | 17.25 | 17.75 | 12 | 5.75 | 6.5 | 9.5 | 0.5 | 0.25 |
| Ecoli washed | Norm Dist | 65 | 0 | 0.5 | 0.25 | 0 | 1 | 3 | 11.5 | 3.25 | 11.5 | 2.25 | 4.5 | 16.25 | 9.5 | 20.75 | 1 | 10.5 | 2.5 | 0.75 | 1 |
| Class 14 | Lin Discrim | 31.25 | 1 | 8.25 | 0 | 0 | 15.75 | 0 | 6.25 | 13.75 | 6.75 | 17 | 0.5 | 7.5 | 0.75 | 9.75 | 5.25 | 3.75 | 3.75 | 0 | 0 |
| | KNN | 79.5 | 0.5 | 2.25 | 0.5 | 0 | 4.25 | 1.5 | 8 | 1.75 | 0.25 | 1.25 | 5.75 | 18.5 | 14.75 | 27.75 | 0.75 | 9.5 | 2.5 | 0.25 | 0 |
| | Fuzzy K | 81 | 0.75 | 2.25 | 0.25 | 0 | 5 | 0.25 | 7.5 | 1.25 | 0.25 | 1.25 | 6.25 | 18.75 | 15.25 | 29 | 0.75 | 8.25 | 2.75 | 0.25 | 0 |

**Table A1.2 Combination two: shape + size data (continued)**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| E coli unwashed | Norm Dist | 85.5 | 1.5 | 0.25 | 1.5 | 0.5 | 0.25 | 1.5 | 0.75 | 1.75 | 0.75 | 0.5 | 2.25 | 3.5 | 1.5 | 1.5 | 37 | 1.25 | 38.5 | 1.25 | 4 |
| Class 15 | Lin Discrim | 58 | 22 | 13 | | | 1 | | 1 | 1.5 | 1 | 2.5 | | 0.5 | 0.25 | 0.25 | 29 | 0.5 | 27.5 | | |
| | KNN | 84 | 7.25 | 3 | 2.25 | | 0.75 | | 1 | 0.5 | 0.25 | 0.5 | 1.25 | 4.25 | 0.75 | 2.25 | 33.5 | 2.25 | 39.75 | | 0.5 |
| | Fuzzy K | 82.5 | 9.25 | 3 | 1.75 | | 1 | | 0.75 | 0.5 | 0.25 | 0.5 | 1 | 4.5 | 0.75 | 2.5 | 33.25 | 1.5 | 39 | | 0.5 |
| Erwinia washed | Norm Dist | 83.25 | 0 | 0.25 | 0 | 0 | 3.25 | 2 | 1 | 4.75 | 1.25 | 0.5 | 12.5 | 12.5 | 8.25 | 6.5 | 2.5 | 36.75 | 4.25 | 2.5 | 1.25 |
| Class 16 | Lin Discrim | 40 | 7.5 | 16.75 | 0 | 0.25 | 14.75 | 0 | 1.5 | 10.5 | 1.5 | 7.25 | 0 | 2.75 | 0.25 | 6.75 | 15.5 | 7.5 | 7.25 | 0 | 0 |
| | KNN | 73.75 | 1.75 | 1 | 0.25 | | 12.5 | | 5 | 4.25 | | | 6 | 13 | 12.5 | 11.25 | 4.25 | 22.5 | 4.25 | 0.75 | |
| | Fuzzy K | 75.25 | 2.25 | 1 | | | 13.75 | | 4 | 3 | | | 3.5 | 13.5 | 16.25 | 12 | 5.25 | 20.75 | 4 | | |
| Erwinia unwashed | Norm Dist | 80 | 0.5 | 1.5 | 2.25 | 1.25 | 1 | 1.25 | 1.75 | 1.25 | 0.75 | | 1.75 | 4.5 | | 0.25 | 30.5 | 4.25 | 38.75 | 2.5 | 6 |
| Class 17 | Lin Discrim | 58 | 19 | 16.75 | | | 1.5 | | 1.75 | 1 | 0.75 | 1.25 | | 0.75 | | 0.5 | 30 | 0.75 | 26 | | |
| | KNN | 77.5 | 7.25 | 5 | 5.75 | | 1.75 | 0.25 | 2 | 0.5 | | | 1 | 5.25 | 1 | 0.75 | 32.25 | 1.5 | 35.75 | | |
| | Fuzzy K | 77.25 | 8.75 | 4.5 | 5.25 | | 1.75 | | 2.25 | 0.25 | | | 0.75 | 5.25 | 1 | 0.5 | 30 | 2 | 37.75 | | |
| P mullbery | Norm Dist | 52 | | | | 2.25 | | 17.75 | 6.75 | 5.5 | 5.5 | 3.25 | 1.75 | 1.5 | 1.75 | 1.25 | 0.25 | 0.5 | | 35.25 | 16.75 |
| Class 18 | Lin Discrim | 0.75 | 7.25 | 15.25 | 0.25 | 25.75 | 12.5 | 0.25 | 3.25 | 10.5 | 7.25 | 7.25 | | 1.75 | 0.25 | 1.5 | 4.5 | 0.25 | 1.5 | | 0.75 |
| | KNN | 35.75 | 2 | 1.25 | 0.25 | 12.5 | 0.75 | 5.25 | 4.5 | 15.75 | 4.75 | 3.75 | 3.25 | 1.75 | 5 | 2 | 0.5 | 0.75 | 0.25 | 27 | 8.75 |
| | Fuzzy K | 34 | 3.25 | 1.5 | 0.25 | 12.5 | 1 | 3.5 | 4.25 | 17.25 | 4.5 | 4.75 | 3 | 1.25 | 4.75 | 3 | 0.75 | 0.5 | | 25.75 | 8.25 |
| Puff ball | Norm Dist | 58.5 | | | 1 | 2.5 | | 10.75 | 3.25 | 4.25 | 7 | 8.5 | 1 | 1.5 | 0.5 | | 0.5 | 0.25 | 0.5 | 17.75 | 40.75 |
| Class 19 | Lin Discrim | 0.25 | 15.75 | 29.25 | 0.25 | 16.5 | 5.25 | | 0.25 | 8 | 7.5 | 9.5 | | 2.25 | | 0.5 | 3 | 0.25 | 1.5 | | 0.25 |
| | KNN | 23.75 | 8.25 | 13.25 | 4 | 9.75 | | 4.5 | 2.5 | 17.25 | 4 | 7 | 1.5 | 0.75 | 0.75 | 1.25 | 0.5 | 0.25 | 0.75 | 8.5 | 15.25 |
| | Fuzzy K | 20.75 | 11 | 13 | 4 | 10.25 | | 3 | 1.75 | 17.5 | 5.25 | 7.75 | 1.75 | 0.75 | 0.75 | 0.5 | 0.75 | 0.25 | 1 | 7.25 | 13.5 |
| | Class: | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |

**Table A1.2 Combination two: shape + size data (continued)**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| oleic 1u | Norm Dist | 78 | 42.75 | 19.25 | 7.5 | 8.5 | | | 0.25 | 0.75 | | | | | | | 7.75 | | 5.25 | | 8 |
| Class 1 | Lin Discrim | 99.75 | 92.75 | 7 | | | | | | | | | | | | | | | 0.25 | | |
| | KNN | 34.25 | 20.75 | | 8 | 5.5 | | 0.25 | | 62 | | | | | | | 1.25 | | 1.25 | | 1 |
| | Fuzzy K | 43 | 30.75 | | 7.25 | 5 | | | 0.25 | 54.75 | | | | | | | 0.25 | | 1 | | 0.75 |
| oleic 3u | Norm Dist | 100 | 100 | | | | | | | | | | | | | | | | | | |
| Class 1 | Lin Discrim | 100 | 100 | | | | | | | | | | | | | | | | | | |
| | KNN | 97 | 93.25 | 3.75 | | | | | | | | 0.75 | 2 | | | 0.25 | | | | | |
| | Fuzzy K | 98.25 | 93 | 5.25 | | | | | | | | 0.75 | 0.5 | | 0.5 | | | | | | |
| PSL 1u | Norm Dist | 36.5 | 3 | 33.5 | | | | 0.25 | | 2.25 | 0.5 | 2.5 | 8.75 | 1 | 4 | 1.75 | | | 2.75 | 37 | 2.75 |
| Class 2 | Lin Discrim | 59.75 | 39 | 20.75 | | | 1.5 | 0.25 | 1 | 23.5 | 1.25 | 5 | | 1 | | 0.25 | 4 | 0.5 | 2 | | |
| | KNN | 72.5 | 54.25 | 18 | 0.25 | | | | 0.25 | 1.25 | | 2.5 | 10.25 | 1.25 | 3.75 | 6.25 | 0.75 | 0.75 | 0.5 | | |
| | Fuzzy K | 76.25 | 57.75 | 18.25 | 0.25 | | | | 0.25 | 1.5 | | 3 | 8.25 | 1.25 | 3 | 4.75 | 1 | 0.25 | 0.5 | | |
| PSL 1.7u | Norm Dist | 93.25 | | 93.25 | | | | | | | | | | | | | 0.25 | | | 6.25 | 0.25 |
| Class 2 | Lin Discrim | 100 | 4.25 | 95.75 | | | | | | | | | | | | | | | | | |
| | KNN | 88 | 0.25 | 87.5 | 0.25 | | | | 0.25 | | | 1.25 | | | | 0.25 | | | | 10.25 | |
| | Fuzzy K | 84.5 | | 84.25 | 0.25 | | | | 0.25 | | | 1 | 0.25 | | | 0.25 | | | | 13.75 | |
| PSL 3u | Norm Dist | 99.5 | | 99.5 | | | | | | 0.25 | | | | | | | | | 0.25 | | |
| Class 2 | Lin Discrim | 99.75 | 66.75 | 33 | | | | | | | | | | | | | | | 0.25 | | |
| | KNN | 81.25 | 0.5 | 80.25 | 0.5 | | | | | | | 18.75 | | | | | | | | | |
| | Fuzzy K | 77.25 | 0.75 | 76 | 0.5 | | | | | | | 19.5 | | | | | | | | 3.25 | |
| Water poly | Norm Dist | 86.25 | 41 | 2 | 41 | 2.25 | | 0.5 | | 8 | | | | | | | | | 1.25 | | 4 |
| Class 3 | Lin Discrim | 98.25 | 94.75 | 3.5 | | | | | | | | | | | | | 1 | | 0.75 | | |
| | KNN | 75.5 | 6.75 | | 63.25 | 5.5 | | 0.5 | 1.25 | 22.25 | | | | | | | 0.25 | | | 0.25 | |
| | Fuzzy K | 84 | 7 | | 68.5 | 8.5 | | 0.75 | 1.25 | 13.75 | 0.25 | | | | | | | | | | |
| Silica | Norm Dist | 74 | 1.5 | 4.5 | 4.5 | 63.5 | 0.75 | 2.75 | 0.25 | 17 | 0.5 | | | | | | | 1.5 | | | 3.25 |
| Class 4 | Lin Discrim | 88.25 | 87.75 | 0.5 | | | 1 | | | 10 | | 0.5 | | | | | | | 0.25 | | |
| | KNN | 31.5 | 15.25 | 0.25 | 9.25 | 6.75 | 0.75 | 0.75 | 0.25 | 51.5 | | 0.25 | 0.5 | 1.75 | 0.25 | 1 | 3.5 | 0.5 | 5.75 | | 1.75 |
| | Fuzzy K | 36.75 | 16.5 | 0.75 | 11 | 8.5 | 0.25 | 0.25 | 0.5 | 50 | 0.25 | | 0.5 | 1.25 | 0.25 | 0.25 | 3 | 0.75 | 4.5 | 0.25 | 1.25 |

**Table A1.3 Combination three: shape + fluorescence data**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Haem | Norm Dist | 89 | | | 0.25 | | 73 | 7 | 4.25 | 4.75 | 3.75 | | | 1 | 1.25 | | | 2 | 0.25 | 0.25 | 2.25 |
| Class 5 | Lin Discrim | 89.25 | 1.5 | 4.75 | | | 61.75 | | | 27.5 | 1.5 | 3 | | | | | | | | | |
| | KNN | 86.5 | 3.25 | | 0.5 | 0.75 | 27.25 | 2.25 | 3.25 | 53.75 | 5 | | | | 0.25 | 0.25 | 0.25 | 0.5 | | 0.25 | 2.5 |
| | Fuzzy K | 88.25 | 4 | | 0.75 | 0.75 | 22 | 2.25 | 3.25 | 60.75 | 3.25 | | | | | | | 0.25 | | 0.25 | 2.5 |
| Kaolin | Norm Dist | 54.5 | | | | 0.25 | 2 | 26.5 | 8 | 18 | 23.25 | 1.5 | 1.5 | 2.25 | 3 | 0.5 | 0.25 | 3.25 | 0.5 | 0.5 | 8.75 |
| Class 6 | Lin Discrim | 49 | 15.75 | 5.25 | | | 4 | | 1.25 | 43.75 | 3.75 | 18.75 | 0.25 | 2.75 | | 0.75 | 2.25 | 0.5 | 0.75 | | 0.25 |
| | KNN | 70.5 | 3 | | 1 | 1 | 4 | 7.25 | 11.75 | 47.5 | 15.25 | | 0.25 | 0.75 | 0.25 | 1 | 1.75 | 0.25 | 0.75 | 2.25 | 2 |
| | Fuzzy K | 72 | 2.5 | | 0.75 | 1.5 | 3 | 3.25 | 12.5 | 53.25 | 14.25 | | 0.25 | 0.75 | 0.25 | 0.75 | 1.5 | 0.25 | 0.75 | 2.25 | 2.25 |
| Magnetite | Norm Dist | 74.75 | | | | | 7.25 | 5.75 | 49.5 | 12.25 | 13 | 1.25 | 0.25 | 1.25 | 1 | | 0.25 | 0.75 | | | 7.5 |
| Class 7 | Lin Discrim | 59.75 | 12.75 | 6.5 | | | 17.5 | | 2.75 | 39.5 | 3.25 | 8.25 | | 1.5 | 0.25 | 0.25 | 3.75 | 1.5 | 2.25 | | |
| | KNN | 78.25 | 4.25 | 0.25 | 2.25 | 0.25 | 3.75 | 3.75 | 34 | 36.75 | 10.25 | | | | | | 0.75 | 0.75 | | 1 | 2 |
| | Fuzzy K | 79 | 4.25 | 0.25 | 2.5 | | 3.25 | 3.5 | 35 | 37.25 | 8.75 | | 0.25 | | | | 0.75 | 0.75 | | 1.25 | 2.25 |
| NaCl 2u | Norm Dist | 62.75 | | | | 2 | 4.5 | 24 | 4.75 | 29.5 | 3.25 | | 2.75 | 3 | 5.5 | | 0.25 | 17 | 0.75 | | 2.75 |
| Class 8 | Lin Discrim | 53.75 | 21.5 | 10.75 | | | 8 | 0.5 | 1.75 | 43.5 | 1 | 4.75 | | 0.75 | | 0.5 | 4 | 1 | 2 | | |
| | KNN | 76.75 | 2.5 | | 3.5 | 1.75 | 6 | 6 | 11.25 | 53.5 | 5.75 | | | 0.5 | | | | 0.75 | | 5 | 3.5 |
| | Fuzzy K | 80 | 1.5 | | 4.5 | 1.75 | 3.75 | 4 | 12.5 | 59.75 | 4 | | | 0.25 | | | | 0.25 | | 5 | 2.75 |
| NaCl 3u | Norm Dist | 65.75 | | | | 0.25 | 0.25 | 8 | 0.25 | 57.25 | 0.75 | 0.25 | 0.75 | 1.5 | | | 0.5 | | 4 | 0.25 | 26 |
| Class 8 | Lin Discrim | 27 | 59.25 | 10.75 | | | 2 | | | 25 | | 0.75 | | 0.25 | | | 1 | | 1 | | |
| | KNN | 57.5 | 5.5 | | | 4.5 | 0.25 | 0.25 | | 57 | | 0.25 | 0.75 | 3 | | 1.5 | 8.5 | 4 | 9.25 | | 5.25 |
| | Fuzzy K | 57.5 | 6.25 | 0.25 | | 3.75 | | 0.25 | | 57.25 | | 0.25 | | 3 | | 1.5 | 9.5 | 3.75 | 9.5 | | 4.75 |
| Gypsum | Norm Dist | 65.75 | | | | 0.5 | 3.25 | 9.75 | 11.25 | 3.25 | 39.5 | 26.25 | | 0.5 | 2 | 1.25 | | 1.75 | | | 0.75 |
| Class 9 | Lin Discrim | 46.75 | 2.25 | 1.25 | | | 4.25 | | | 41.75 | 14 | 32.75 | | 2.75 | | | 0.75 | | 0.25 | | |
| | KNN | 31 | 1.25 | | 0.5 | 0.25 | 5.75 | 11.5 | 16.75 | 26.75 | 31 | | 0.25 | 0.5 | | | 0.75 | 0.25 | 0.25 | 0.75 | 3.5 |
| | Fuzzy K | 29 | 0.75 | | 0.75 | 0.25 | 5 | 8.75 | 17.25 | 32.75 | 29 | | 0.25 | 0.25 | | | 0.5 | 0.5 | 0.25 | 0.75 | 3 |
| Caffeine | Norm Dist | 74 | | | | | 1 | 0.75 | 1.25 | 0.25 | 0.75 | 73.25 | 1.5 | 0.5 | 1.75 | 3 | | 0.25 | | 15.5 | 0.25 |
| Class 10 | Lin Discrim | 59.75 | 1 | 1.5 | | | 2.5 | | 0.5 | 31.75 | 14.75 | 45 | | 2.25 | | | 0.25 | 0.5 | | | |
| | KNN | 41.5 | 13.5 | 19.75 | | | 0.25 | | 0.5 | 3 | 0.25 | 41.25 | 6.5 | 1.5 | 1 | 9 | 2.25 | 0.25 | 1 | | |
| | Fuzzy K | 40.75 | 13.25 | 17 | | | 0.25 | | 0.5 | 3 | 0.25 | 40.5 | 6.25 | 1.5 | 2 | 8.25 | 2 | 0.5 | 0.75 | 4 | |

**Table A1.3 Combination three: shape + fluorescence data (continued)**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bg spores | Norm Dist | 93.25 | | | | | 0.75 | 0.75 | | 0.25 | 0.5 | 3 | 37.5 | 9.75 | 25.75 | 12.75 | 0.25 | 4.75 | 2.5 | 0.75 | 0.75 |
| Class 11 | Lin Discrim | 22 | 3.25 | 8 | | | 7.5 | | 2 | 33.25 | 7.25 | 16.75 | | 5.5 | | 1 | 9.5 | 0.75 | 5.25 | | |
| | KNN | 53.75 | 23 | 13.25 | 0.25 | 0.25 | | 0.25 | 0.25 | 5.25 | 0.75 | 1.75 | 28.75 | 0.5 | 6.5 | 14.5 | 1.25 | 1.25 | 1 | 0.25 | 1 |
| | Fuzzy K | 57.5 | 20.75 | 12 | 0.25 | | 0.25 | 0.25 | 0.25 | 5.25 | 0.5 | 2.25 | 30.25 | 1.25 | 7.25 | 14.5 | 1.5 | 1.5 | 1.25 | 0.25 | 0.5 |
| Bg sp clusters | Norm Dist | 82.75 | | 0.25 | 0.25 | | 2.25 | 2.75 | 1 | 3.25 | 2 | 2.75 | 4.75 | 42.75 | 6 | 10 | 1.75 | 12.25 | 5.25 | | 2.75 |
| Class 12 | Lin Discrim | 25.5 | 4.75 | 8.75 | | | 5.75 | | 1.75 | 36.25 | 4 | 13.25 | | 6.5 | | 0.75 | 10.25 | 1 | 7 | | |
| | KNN | 63.5 | 2.75 | | 0.5 | | 0.75 | 1 | 1.25 | 25.5 | 0.75 | 0.75 | 2.75 | 19 | 2 | 6.75 | 9.75 | 14.25 | 9 | | 3.25 |
| | Fuzzy K | 60.75 | 2.5 | 0.25 | 0.5 | | 0.5 | 0.5 | 1.25 | 30.5 | 0.5 | 1 | 2.5 | 18.25 | 2.5 | 6.25 | 11 | 10.75 | 9.5 | | 1.75 |
| Bg veg | Norm Dist | 81.25 | | 0.5 | | 0.25 | 0.5 | 2.75 | 0.5 | 4.25 | 1.5 | 2.5 | 15.25 | 15.25 | 17.25 | 8.5 | 3.75 | 12.75 | 8.5 | 4 | 2 |
| Class 13 | Lin Discrim | 24.5 | 6.25 | 10.25 | | | 11.75 | 0.25 | 2.75 | 28.5 | 4 | 11.75 | | 3.25 | | 2.75 | 8.75 | 1 | 8.75 | | |
| | KNN | 51.75 | 10.75 | 6.25 | 1.25 | 0.5 | 1.5 | 0.75 | 1.5 | 20.75 | 2.75 | 0.5 | 12.25 | 6.75 | 3 | 12.75 | 6 | 4.5 | 6.5 | 0.5 | 1.25 |
| | Fuzzy K | 51.25 | 11.25 | 5.5 | 1.75 | 0.25 | 1.25 | 0.25 | 1.75 | 22.5 | 2.25 | 0.75 | 11 | 7.75 | 3.75 | 13 | 5 | 4 | 6.75 | 0.25 | 1 |
| Ecoli washed | Norm Dist | 84.75 | | | 0.25 | | | 1 | 0.5 | 2.25 | 1.5 | 7.5 | 6.75 | 18.75 | 16.25 | 30.5 | 1 | 9.5 | 2 | 1 | 1.25 |
| Class 14 | Lin Discrim | 17.25 | 2.25 | 5.5 | | | 11.75 | | 2 | 40 | 4.25 | 17 | | 4.75 | 0.5 | 1.25 | 5.75 | 2.25 | 2.75 | | |
| | KNN | 74.75 | 9.75 | 2.5 | | 0.5 | 0.25 | 0.75 | 0.25 | 8.75 | 0.5 | 0.75 | 7.25 | 13.75 | 3.75 | 30.25 | 5.5 | 6.5 | 7.75 | | 1.25 |
| | Fuzzy K | 79.5 | 5.25 | 1.25 | | 0.5 | | 0.5 | 0.25 | 10.25 | 0.25 | 1.25 | 10 | 14.25 | 4.75 | 30.5 | 5.75 | 6.5 | 7.75 | | 1 |
| E coli unwashed | Norm Dist | 90 | 0.75 | 1.75 | | | | 0.5 | | 1 | | 0.5 | 4.5 | 5.25 | 2.25 | 2.25 | 39 | 0.75 | 36 | 1.5 | 4 |
| Class 15 | Lin Discrim | 64.25 | 19 | 5 | 0.25 | | 0.75 | | 1 | 5 | 1.25 | 3.5 | | 0.25 | | | 29.5 | 0.25 | 34.25 | | |
| | KNN | 65.5 | 11 | 1.25 | | 0.25 | | 0.25 | | 19.5 | | 0.5 | 2.75 | 12 | 1 | 11.75 | 16.75 | 5.75 | 15.5 | | 1.75 |
| | Fuzzy K | 66.5 | 9.25 | 1.25 | | 0.25 | | 0.25 | | 20 | | 0.75 | 3.25 | 10.5 | 0.75 | 11.5 | 19 | 5.5 | 16 | | 1.75 |
| Erwinia washed | Norm Dist | 86.25 | | 0.25 | | | 1.75 | 3.5 | 0.75 | 4 | 1.25 | 0.25 | 5.75 | 14.75 | 7.25 | 4.25 | 2 | 48.25 | 4 | 0.5 | 1.5 |
| Class 16 | Lin Discrim | 35.5 | 3.75 | 11 | | | 13 | | 2 | 24 | 2 | 8.75 | | 1.75 | 0.25 | 4.75 | 16 | 6 | 6.75 | | |
| | KNN | 44.25 | 2.75 | 0.75 | 0.75 | 0.5 | 2.75 | 3.75 | 1 | 38.25 | 1.25 | | 1.75 | 8.25 | 2.25 | 4 | 6 | 13.5 | 8.5 | | 4 |
| | Fuzzy K | 41 | 2 | 0.5 | 0.75 | 0.5 | 2.25 | 3.25 | 1 | 45.5 | 0.75 | 0.25 | 1 | 8.5 | 2 | 4.25 | 6.5 | 11 | 7.75 | | 2.25 |

**Table A1.3 Combination three: shape + fluorescence data (continued)**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Erwinia unwashed | Norm Dist | 84.25 | 0.25 | 3.75 | | 1 | 0.25 | 0.25 | 0.25 | 4 | 0.5 | 0.25 | 3 | 5.25 | 0.25 | 0.5 | 29.5 | 4.25 | 41.5 | 2.25 | 3 |
| Class 17 | Lin Discrim | 61 | 22.5 | 8.25 | | | 1.25 | | 1.5 | 3.5 | 0.75 | 1.25 | | | | 0.5 | 30.5 | 0.25 | 29.75 | | |
| | KNN | 65 | 12.75 | 1 | | 1 | 0.25 | 0.25 | | 18 | 0.25 | 0.25 | 3.75 | 12 | 1.75 | 12 | 14.5 | 7.25 | 13.75 | | 1.25 |
| | Fuzzy K | 66.25 | 11.5 | 1.25 | | 0.5 | 0.25 | | | 19.25 | | 0.25 | 3.75 | 10.75 | 2 | 11.5 | 15.25 | 6.75 | 16.25 | | 0.75 |
| P mullbery | Norm Dist | 59.5 | | 0.5 | | | 0.5 | 8.75 | 6 | 10.25 | 5.25 | 1.75 | 2.75 | 1 | 2.25 | 0.25 | | 0.75 | 0.5 | 54.25 | 5.25 |
| Class 18 | Lin Discrim | 5.5 | 8.25 | 41.75 | | | 4.5 | | 0.5 | 27 | 1.75 | 5.5 | | 1.25 | | 0.25 | 2.25 | | 1.5 | 5.5 | |
| | KNN | 37 | 6 | 27 | 4 | | | 1.25 | 11.75 | 1 | 4.75 | 4.75 | 0.5 | 0.5 | | 0.5 | 0.5 | 0.25 | 0.25 | 37 | |
| | Fuzzy K | 34 | 6.25 | 26.75 | 5.75 | | | 0.75 | 13.25 | 1.25 | 5.75 | 4 | 0.25 | 0.5 | 0.25 | 0.5 | 0.5 | | 0.25 | 34 | |
| Puff ball | Norm Dist | 30.5 | | 3.5 | | 1 | 1.25 | 11 | 1.75 | 25.75 | 8.25 | 7.5 | 1.5 | 2.75 | 0.25 | 0.25 | 1.5 | 0.5 | 2.75 | 2.5 | 28 |
| Class 19 | Lin Discrim | 0 | 47.25 | 8.25 | 0.5 | | 2.75 | | 0.25 | 24.25 | 2.25 | 9.75 | | 1.25 | | | 2.25 | | 1.25 | | |
| | KNN | 3.5 | 8.75 | 1.25 | 2 | 2.5 | 1.75 | 2.5 | 2.5 | 47.5 | 4.25 | 0.5 | 3.5 | 3.25 | 0.25 | 3.25 | 4.75 | 2.75 | 5.25 | 1.25 | 2.25 |
| | Fuzzy K | 3.25 | 10.75 | 1.5 | 2.5 | 1.75 | 0.25 | 1.25 | 2.5 | 51.25 | 4 | 0.5 | 3.5 | 3.25 | 0.25 | 2.5 | 3.5 | 2.25 | 5.25 | 0.75 | 2.5 |
| | Class: | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |

**Table A1.3 Combination three: shape + fluorescence data (continued)**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| oleic 1u | Norm Dist | 93.25 | 92.5 | 0.5 | 0.25 | | | | | | | | | | | | 3.75 | | 1.5 | | 1.5 |
| Class 1 | Lin Discrim | 99 | 72 | 27 | | | | | | | | | | | | | 0.25 | | 0.75 | | |
| | KNN | 100 | 99.5 | | 0.5 | | | | | | | | | | | | | | | | |
| | Fuzzy K | 100 | 99.5 | | 0.5 | | | | | | | | | | | | | | | | |
| PSL 1u | Norm Dist | 99.04 | 0.24 | 98.08 | 0.72 | | | | | | | | | | | | 0.24 | | | 0.48 | 0.24 |
| Class 2 | Lin Discrim | 64.25 | 3 | 61.25 | | | 4.25 | 0.25 | 1.25 | 12.25 | 1.75 | 4.75 | | 1.25 | 0.5 | 0.25 | 4.5 | 0.75 | 3.75 | | 0.25 |
| | KNN | 69.5 | 0.5 | 69 | | | | | 0.25 | | | 3.5 | 10 | 2.25 | 7 | 3 | 2.25 | | 1.5 | | 0.75 |
| | Fuzzy K | 70.5 | 0.5 | 70 | | | | | 0.25 | | | 3.5 | 10.25 | 2.25 | 5.75 | 3.5 | 1.75 | | 1.5 | | 0.75 |
| Water poly | Norm Dist | 84 | 1.25 | 2.5 | 77.75 | 2.5 | | 0.25 | 0.25 | 0.25 | | | | | | | 0.75 | 0.25 | 5 | | 9.25 |
| Class 3 | Lin Discrim | 98 | 24.25 | 42.75 | 0.75 | 30.25 | | | | | | | | | | | 1.25 | | 0.75 | | |
| | KNN | 100 | 4.5 | | 85 | 10.5 | | | | | | | | | | | | | | | |
| | Fuzzy K | 100 | 5.75 | | 84 | 10.25 | | | | | | | | | | | | | | | |
| Silica | Norm Dist | 82.5 | | | 6 | 76.5 | | 11.75 | | 0.25 | 1.75 | | | | | | | | | 3.75 | |
| Class 4 | Lin Discrim | 98.5 | 4.75 | 5 | | 88.75 | 0.25 | | | 0.5 | 0.75 | | | | | | | | | | |
| | KNN | 94.25 | | | 20 | 74.25 | | 1.75 | | 1 | 1.5 | | | | | | | | | | 1.5 |
| | Fuzzy K | 95.75 | | | 19.5 | 76.25 | | 1 | | 1 | 1 | | | | | | | | | | 1.25 |
| Haem | Norm Dist | 93.25 | | | 0.25 | | 82.75 | 5 | 2.75 | 2.75 | 0.75 | | | 0.75 | 1 | | | 0.75 | 0.25 | 0.25 | 2.75 |
| Class 5 | Lin Discrim | 83.5 | 1.25 | 10.5 | 0.5 | | 72.5 | | 0.25 | 10.75 | 1.25 | 1.75 | | 0.25 | | 0.5 | | 0.25 | 0.25 | | |
| | KNN | 95.5 | 1.75 | | | | 92 | 0.25 | 2.25 | 1 | | | | 0.25 | 0.25 | | | 2 | | | 0.25 |
| | Fuzzy K | 96 | 1.75 | | | | 92.75 | 0.25 | 2 | 1 | | | | | | | | 2 | | | 0.25 |
| Kaolin | Norm Dist | 63.25 | | | | 0.25 | 0.5 | 41.5 | 7.5 | 13.75 | 14.25 | 2 | 0.5 | 0.75 | 1.5 | 0.25 | | 1.25 | | 2.5 | 13.5 |
| Class 6 | Lin Discrim | 25.75 | 7.75 | 9.25 | 0.5 | 17.25 | 7 | 0.75 | 2.5 | 15.5 | 8.75 | 20.5 | 0.25 | 3.5 | 0.25 | 2.5 | 1.75 | 0.75 | 1 | | 0.25 |
| | KNN | 64 | 0.75 | | 0.5 | 9.25 | 1.75 | 30 | 16.75 | 15.5 | 12.5 | | | | 0.25 | | | 0.25 | | 2.75 | 9.75 |
| | Fuzzy K | 62.5 | 0.75 | | 0.5 | 10.25 | 2.5 | 24 | 18.75 | 17.25 | 14 | | | | | | | | | 2.25 | 9.75 |
| Magnetite | Norm Dist | 77 | | | | | 0.75 | 10.5 | 60.75 | 5 | 6 | 1 | 0.25 | 0.5 | 0.75 | 0.5 | | 0.5 | 0.25 | 0.25 | 13 |
| Class 7 | Lin Discrim | 44.5 | 15.25 | 4.75 | 0.5 | 2 | 25.25 | | 10 | 9.25 | 6 | 9.25 | 0.25 | 3 | 0.25 | 1.25 | 6.75 | 2 | 4 | | 0.25 |
| | KNN | 83.75 | 0.25 | 0.25 | 2.75 | 1 | 5.75 | 7.5 | 58.75 | 11.75 | 1.5 | | | | | | | 0.25 | | 1.25 | 9 |
| | Fuzzy K | 85 | 0.25 | 0.25 | 3 | 1 | 6.75 | 6.75 | 59.75 | 11.75 | 1 | | | | | | | | | 0.75 | 8.75 |
| NaCl 2u | Norm Dist | 84.5 | | | | 0.25 | | 7 | 4.5 | 73 | 1 | | 1.25 | 2.25 | 3.25 | 0.5 | | 3.25 | 0.25 | | 3.5 |
| Class 8 | Lin Discrim | 33.5 | 24.5 | 25.5 | | | 12.75 | | 2.75 | 18 | 0.25 | 4 | | 2.5 | 0.5 | 1.25 | 4.25 | 2.25 | 1.5 | | |
| | KNN | 75.25 | 3 | | 4.25 | | 4.25 | 6 | 14 | 51 | | | | 0.5 | | | | 1.5 | 0.5 | 2 | 13 |
| | Fuzzy K | 74.25 | 5 | | 5.25 | | 5 | 3.25 | 17 | 49 | | | | 0.25 | | | | 1.75 | | 1.5 | 12 |

**Table A1.4 Combination four: shape + size + fluorescence data**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gypsum | Norm Dist | 70.75 | | | | 0.75 | | 14.25 | 4.75 | 2 | 54.25 | 16.5 | | | 0.75 | 0.25 | | 0.25 | | 3 | 3.25 |
| Class 9 | Lin Discrim | 61.75 | 0.25 | 1.5 | | 17.5 | 3.75 | 0.25 | 2.25 | 7.5 | 32 | 29.75 | | 3.25 | | 1.25 | 0.5 | | 0.25 | | |
| | KNN | 78.25 | | | 0.5 | 5.5 | | 7 | 5.75 | 2 | 78.25 | | | | | | | 0.25 | | 0.25 | 0.5 |
| | Fuzzy K | 77.75 | | | 0.5 | 6 | | 5 | 7.5 | 2.5 | 77.75 | | | | | | | | | 0.25 | 0.5 |
| Caffeine | Norm Dist | 73 | | | | | 0.25 | 0.25 | 1.75 | | | 73 | 1 | | 1.75 | 2 | | 0.25 | | 18.5 | 1.25 |
| Class 10 | Lin Discrim | 70.75 | 0.25 | 2.25 | | 0.25 | 5.5 | 0.25 | 1.25 | 11.75 | 23.75 | 47 | | 3 | 0.25 | 2 | 1 | 0.25 | 0.5 | 0.75 | |
| | KNN | 73 | 3 | 10.25 | | | | 0.25 | 1 | 1.25 | | 73 | 2.5 | 1 | 2.25 | 3 | | 1.5 | | | 1 |
| | Fuzzy K | 73.25 | 3 | 10.25 | | | | 0.25 | 1 | 1 | 0.25 | 73 | 2.75 | 1.75 | 1.75 | 2.25 | | 1.75 | | | 1 |
| Bg spores | Norm Dist | 96.75 | | | | | 0.25 | 0.25 | 0.25 | 0.75 | | 0.75 | 39.75 | 8.75 | 26.25 | 13.5 | 0.25 | 4.75 | 3.5 | 0.75 | 0.25 |
| Class 11 | Lin Discrim | 33.75 | 6 | 10.5 | | | 10 | | 3.5 | 14.5 | 5.5 | 16.25 | | 10.5 | | 6 | 10.25 | 2 | 5 | | |
| | KNN | 89.5 | 0.5 | 4.5 | 0.75 | | 1.5 | | 0.25 | 1.75 | | 0.5 | 47 | 2.75 | 20.5 | 14.25 | 1.5 | 2.5 | 1 | | 0.75 |
| | Fuzzy K | 89.25 | 0.75 | 4.75 | 0.5 | | 2 | | 0.25 | 1.5 | | 0.5 | 47.75 | 2.5 | 19.75 | 14.25 | 1.5 | 2.5 | 1 | | 0.5 |
| Bg sp clusters | Norm Dist | 93.25 | | 0.25 | 0.25 | | 1 | | 1 | 1.75 | 0.5 | 0.25 | 6.25 | 48.75 | 6 | 12.25 | 1.5 | 13 | 5.5 | 0.5 | 1.25 |
| Class 12 | Lin Discrim | 38.5 | 7 | 9.75 | | | 8.75 | | 8.25 | 11.25 | 4.25 | 12.25 | 0.25 | 12.25 | | 5.25 | 11.75 | 1 | 8 | | |
| | KNN | 91 | 1.25 | 0.25 | 0.75 | | 3 | 0.25 | 1.75 | 1.75 | | | 1.75 | 53.25 | 2.75 | 7.75 | 7.25 | 14.5 | 3.75 | | |
| | Fuzzy K | 90.5 | 2 | 0.5 | 0.75 | | 3 | 0.25 | 1.75 | 1.25 | | | 2 | 56 | 1.25 | 7.5 | 7.75 | 12.75 | 3.25 | | |
| Bg veg | Norm Dist | 86.5 | | 0.5 | | | 1 | 1.5 | 0.5 | 2.25 | 1 | 2 | 16 | 16.25 | 20.25 | 8.5 | 3.5 | 13 | 9 | 2.5 | 2.25 |
| Class 13 | Lin Discrim | 29 | 8.5 | 14.5 | | | 15.25 | 0.25 | 6.5 | 10.25 | 4.75 | 11 | 0.5 | 4.75 | 0.25 | 6 | 8.75 | 1.5 | 7.25 | | |
| | KNN | 79.5 | 4 | 3.75 | 2.75 | | 2.75 | 0.75 | 3.75 | 2.25 | | 0.25 | 16.25 | 18.25 | 10.75 | 15 | 6.25 | 10.25 | 2.75 | | 0.25 |
| | Fuzzy K | 78.5 | 4.75 | 4 | 2.75 | | 3.5 | 0.25 | 4.25 | 1.75 | | 0.25 | 16 | 18.75 | 9.75 | 14.75 | 6.75 | 10 | 2.5 | | |
| Ecoli washed | Norm Dist | 90.5 | | 0.25 | 0.25 | | | 1.25 | 0.5 | 0.75 | 0.5 | 4.75 | 8 | 19.75 | 16.25 | 33.5 | 1.25 | 9.75 | 2 | 0.75 | 0.5 |
| Class 14 | Lin Discrim | 31.5 | 1.75 | 6.75 | | | 15.5 | | 6.5 | 14.25 | 7 | 16.75 | 0.5 | 7.5 | 0.75 | 9.75 | 5.25 | 4 | 3.75 | | |
| | KNN | 93.5 | 2.5 | 0.25 | | | 1.5 | | 0.75 | 1 | | | 9.75 | 21.75 | 9.25 | 41.25 | 2 | 7.75 | 1.75 | | 0.5 |
| | Fuzzy K | 93 | 3.25 | 0.25 | | | 1.75 | | 0.75 | 1 | | | 9.75 | 24 | 5.75 | 43.25 | 1.75 | 7 | 1.5 | | |
| E coli unwashed | Norm Dist | 94.25 | 0.5 | 2.75 | | | | | | | | 0.5 | 4 | 4.25 | 3 | 2.5 | 38.25 | 1 | 41.25 | 1 | 1 |
| Class 15 | Lin Discrim | 58 | 26.25 | 8 | | | 1.25 | | 1.25 | 1.75 | 1 | 2.5 | | 0.5 | 0.25 | 0.25 | 29 | 0.5 | 27.5 | | |
| | KNN | 93 | 3.25 | 3.25 | | | 0.5 | | | | | | 4.5 | 3.75 | 1 | 4 | 40.75 | 1.25 | 37.75 | | |
| | Fuzzy K | 92.25 | 4 | 3.25 | | | 0.5 | | | | | | 4 | 3.25 | 1.25 | 4.75 | 40 | 1.25 | 37.75 | | |

**Table A1.4 Combination four: shape + size + fluorescence data (continued)**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Erwinia washed | Norm Dist | 89.75 | | 0.25 | | | 2.5 | 1.75 | 0.75 | 1 | 0.25 | 0.5 | 5.5 | 16 | 6.5 | 5 | 1.75 | 51 | 4 | 0.75 | 2.5 |
| Class 16 | Lin Discrim | 40 | 7.5 | 16.75 | | 0.25 | 14.75 | | 1.5 | 10.5 | 1.5 | 7.25 | | 2.75 | 0.25 | 6.75 | 15.5 | 7.5 | 7.25 | | |
| | KNN | 83.5 | 1.25 | 0.5 | 1 | | 8.25 | 1.25 | 1.25 | 2.5 | | | 0.25 | 16 | 3.75 | 5.25 | 7.5 | 45.5 | 5.25 | | 0.5 |
| | Fuzzy K | 82.25 | 1.5 | 0.5 | 1 | | 9.75 | 0.75 | 1.5 | 2.5 | | | 0.5 | 20 | 2 | 5 | 8 | 42.75 | 4 | | 0.25 |
| Erwinia unwashed | Norm Dist | 89.75 | 0.25 | 5 | | 0.5 | 0.25 | 0.25 | 0.25 | | | 0.25 | 3.75 | 6.25 | 0.25 | 0.75 | 31.25 | 4.5 | 43 | 1.5 | 2 |
| Class 17 | Lin Discrim | 57.75 | 23.25 | 12.75 | | | 1.5 | | 1.75 | 1.25 | 0.5 | 1.25 | | 0.75 | | 0.5 | 29.5 | 0.75 | 26.25 | | |
| | KNN | 87.75 | 4.5 | 6.75 | 0.25 | | 0.5 | 0.25 | | | | | 4 | 5.75 | 0.5 | 2.5 | 41.75 | 2.25 | 31 | | |
| | Fuzzy K | 86.25 | 5 | 7.5 | 0.25 | | 0.75 | 0.25 | | | | | 3.75 | 6.25 | 0.25 | 2.25 | 42.5 | 2 | 29.25 | | |
| P mullbery | Norm Dist | 66.25 | | | | | | 11.5 | 6.25 | 7 | 4.25 | 2.25 | 1.25 | 0.25 | 1 | | | | | 59 | 7.25 |
| Class 18 | Lin Discrim | 24.75 | 3.25 | 24.25 | 0.25 | 11.5 | 8.25 | | 2 | 7 | 4.5 | 6.25 | | 1.25 | 0.25 | 1.5 | 3.25 | 0.25 | 1.5 | 24 | 0.75 |
| | KNN | 28.25 | 3.5 | 23.75 | 2.5 | 1 | 0.75 | 5.5 | 13.75 | 3.75 | 6.25 | 9 | | | 0.75 | 0.5 | 0.25 | 0.5 | | 27.75 | 0.5 |
| | Fuzzy K | 27.5 | 3.75 | 22.75 | 3 | 1.25 | 0.75 | 4.75 | 14.75 | 4.25 | 6.75 | 8.25 | 0.25 | | 0.75 | 0.25 | 0.25 | 0.5 | 0.25 | 27 | 0.5 |
| Puff ball | Norm Dist | 67.25 | | | | 0.25 | | 11.75 | 2.25 | 4.75 | 6.25 | 5.75 | | 0.75 | | | 0.25 | 0.25 | 0.5 | 8.5 | 58.75 |
| Class 19 | Lin Discrim | 0.25 | 19.75 | 21.75 | 0.75 | 16.75 | 5.75 | | 0.5 | 10 | 7.5 | 9.75 | | 2.25 | | 0.5 | 2.75 | 0.25 | 1.5 | | 0.25 |
| | KNN | 29.75 | 11.5 | | 3.75 | 2.75 | 0.5 | 12 | 3.75 | 28 | 5.25 | 2.25 | | | | | 0.25 | 0.25 | | | 29.75 |
| | Fuzzy K | 27.75 | 12.5 | | 3.75 | 3.5 | 0.5 | 10.5 | 4.25 | 28.75 | 6 | 2 | | | | | 0.25 | 0.25 | | | 27.75 |
| | Class: | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |

**Table A1.4 Combination four: shape + size + fluorescence data (cintinued)**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| oleic 1u | Norm Dist | 63.75 | 63.75 | | | | 15.25 | | 4.5 | 10.75 | | | | | | | | 5.75 | | | |
| Class 1 | Lin Discrim | 100 | | 100 | | | | | | | | | | | | | | | | | |
| | KNN | 56 | 55.25 | | 0.75 | | 9.5 | 0.25 | 5 | 13.25 | | | | | 0.25 | | | 14 | | 0.25 | 1.5 |
| | Fuzzy K | 55 | 53 | | 2 | | 8 | 0.75 | 4.25 | 11.5 | 0.25 | 0.5 | 0.75 | 0.25 | 1.75 | | 0.25 | 13.5 | 0.5 | 0.25 | 2.5 |
| PSL 1u | Norm Dist | 49.64 | 1.20 | 48.44 | | | | | 0.72 | 0.24 | | 15.11 | 10.31 | | 14.87 | 6.24 | 0.48 | 2.16 | 0.24 | | |
| Class 2 | Lin Discrim | 100 | 0.5 | 99 | | 0.5 | | | | | | | | | | | | | | | |
| | KNN | 63 | 0.75 | 62 | 0.25 | | | | 0.25 | | | 3.25 | 11 | 0.75 | 11.25 | 4.5 | 2.25 | | 3.5 | | 0.25 |
| | Fuzzy K | 58.5 | 0.5 | 58 | | | | | 0.25 | | | 5.5 | 13.25 | 0.75 | 11 | 4.5 | 1.75 | 0.25 | 4 | | 0.25 |
| Water poly | Norm Dist | 33.25 | 2 | | 26 | 5.25 | 2.25 | 1.75 | 35.75 | 10.25 | 15 | | | | | | | | | | 1.75 |
| Class 3 | Lin Discrim | 97.5 | 9 | 49.5 | | 39 | | | | 2.5 | | | | | | | | | | | |
| | KNN | 40.5 | 2.75 | | 26.75 | 11 | 1.25 | 6.75 | 16 | 14 | 14.25 | | | | | | | | | 5.5 | 1.75 |
| | Fuzzy K | 40 | 2.5 | 0.75 | 25 | 11.75 | 1.5 | 10.5 | 12.75 | 14 | 14.25 | | 0.25 | 0.25 | | | | 0.25 | | 4.75 | 1.5 |
| Silica | Norm Dist | 75 | | | 15.75 | 59.25 | | 11.25 | 0.25 | | 11 | | | | | | | | | 0.75 | 1.75 |
| Class 4 | Lin Discrim | 99.75 | | 0.25 | | 99.5 | | | | 0.25 | | | | | | | | | | | |
| | KNN | 83.75 | | | 18.5 | 65.25 | | 4.75 | | 0.5 | 9.25 | | | | | | | | | 0.25 | 1.5 |
| | Fuzzy K | 80.5 | | | 17.25 | 63.25 | | 6 | | 0.5 | 11.5 | | | | | | | | | | 1.5 |
| Haem | Norm Dist | 93 | 4.75 | | | | 91.5 | | 1.25 | 0.25 | | | | | | | | | 2 | | 0.25 |
| Class 5 | Lin Discrim | | | 100 | | | | | | | | | | | | | | | | | |
| | KNN | 88 | 5.5 | | | | 83 | 0.25 | 3.75 | 1 | | | | 0.5 | 0.25 | | 0.75 | 2 | 3 | | |
| | Fuzzy K | 87.25 | 5.75 | | 0.25 | | 83 | 1 | 2.5 | 0.75 | | | 0.25 | 1 | 0.25 | | 0.75 | 1.25 | 3.25 | | |
| Kaolin | Norm Dist | 62 | 2.75 | | 1 | 6 | 1.75 | 13.5 | 30.25 | 16.5 | 16.25 | | | | | | | | | 1 | 11 |
| Class 6 | Lin Discrim | 5.25 | 12.75 | 48.5 | | 33.5 | | | | 5.25 | | | | | | | | | | | |
| | KNN | 53 | 2.5 | | 6.75 | 10.5 | 1.75 | 16.5 | 12.75 | 22 | 15.75 | | | | | | | | | 1.75 | 9.75 |
| | Fuzzy K | 51.5 | 2.75 | 0.25 | 8.5 | 11.5 | 1.25 | 18 | 12.25 | 20 | 14 | | | | | | | 0.25 | | 2.5 | 8.75 |
| Magnetite | Norm Dist | 78.75 | 4.5 | | | 0.25 | 13.75 | 7.25 | 41.5 | 16.25 | 7.75 | 0.25 | | | | | | | | | 8.5 |
| Class 7 | Lin Discrim | 4.25 | 11.5 | 74 | | 10.25 | | | | 4.25 | | | | | | | | | | | |
| | KNN | 66.25 | 3.25 | 0.25 | 12.25 | 0.25 | 7.75 | 12.25 | 24.75 | 21.5 | 7.75 | | | | | | | | | 3 | 7 |
| | Fuzzy K | 66 | 3.75 | 0.5 | 10.25 | 0.5 | 7.5 | 16 | 22.75 | 19.75 | 9 | 0.25 | | 0.25 | | | | | 0.25 | 3.5 | 5.75 |
| NaCl 2u | Norm Dist | 82.25 | 12.25 | | | | 4 | | 31.5 | 46.75 | 0.25 | | | | | | | | | | 5.25 |
| Class 8 | Lin Discrim | 2.75 | 1.25 | 96 | | | | | | 2.75 | | | | | | | | | | | |
| | KNN | 64.75 | 8.25 | | 8 | | 3.25 | 5.5 | 10 | 46 | 2 | | | | 0.25 | | | 0.25 | | 2.5 | 14 |
| | Fuzzy K | 62.25 | 7.5 | | 7.5 | | 3.25 | 4.5 | 13.25 | 41.25 | 3.5 | | 0.5 | | 1.25 | | | 0.75 | | 1.75 | 15 |

**Table A1.5 Combination five: size +fluorescence data**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gypsum | Norm Dist | 43.5 | 0.25 | | 9.75 | 16 | 0.25 | 4.5 | 19.75 | 3.75 | 43.5 | | | | | | | | | 0.25 | 2 |
| Class 9 | Lin Discrim | | 13 | 19 | | 62 | | | | 6 | | | | | | | | | | | |
| | KNN | 34 | 0.25 | | 8.5 | 29.5 | 0.25 | 10 | 5.5 | 4.75 | 34 | | | | | | | | | 5.25 | 2 |
| | Fuzzy K | 35.75 | | | 9.25 | 27.75 | 0.25 | 9 | 4.75 | 5 | 35.75 | | | | | | | | | 6 | 2.25 |
| Caffeine | Norm Dist | 78.25 | 1.5 | 6 | | | 0.25 | | 0.5 | 0.5 | | 78.25 | 1.25 | | 1.75 | 0.5 | | 0.5 | 0.5 | 4.5 | 4 |
| Class 10 | Lin Discrim | | 9.25 | 54.25 | | 35.75 | | | | 0.75 | | | | | | | | | | | |
| | KNN | 58.75 | 10 | 17.25 | 0.25 | | | | | 6 | | 58.75 | 1.5 | | 2.25 | 1.25 | 0.25 | 0.75 | | 0.25 | 1.5 |
| | Fuzzy K | 61.25 | 9 | 14.75 | | | | | 0.5 | 6.25 | | 61.25 | 1 | | 2.5 | 1.5 | 0.25 | 0.75 | | 0.75 | 1.5 |
| Bg spores | Norm Dist | 86 | 3 | 6.75 | | | 1.25 | | 1 | 1.5 | | 0.5 | 56.25 | | 3.75 | 16 | 3.5 | 2 | 4.5 | | |
| Class 11 | Lin Discrim | | | 100 | | | | | | | | | | | | | | | | | |
| | KNN | 67 | 2 | 25.75 | | | 0.75 | 0.25 | 1.25 | 2.5 | | 0.25 | 36 | 0.75 | 9.5 | 15 | 2.25 | 1.75 | 1.75 | | 0.25 |
| | Fuzzy K | 41 | | 1.25 | 12.75 | | 0.5 | 16.5 | | | 7.75 | | | 27.25 | | 0.25 | 13.5 | | | 20.25 | |
| Bg sp | Norm Dist | 80.25 | 4 | 0.25 | | | 12.5 | | 1.75 | 0.75 | | 0.5 | 1 | 3.25 | 2.25 | 4.75 | 37.75 | 17 | 14.25 | | |
| clusters | Lin Discrim | | | 100 | | | | | | | | | | | | | | | | | |
| Class 12 | KNN | 82.75 | 3 | 0.75 | 1 | | 10 | | 1 | 1.25 | | 0.25 | 1 | 21 | 3.5 | 5.25 | 22 | 12.75 | 17.25 | | |
| | Fuzzy K | 35.25 | | 1 | 14 | | 0.75 | 24.5 | | 1.25 | 12.75 | | | 23 | | 0.25 | 12 | | | 10.5 | |
| Bg veg | Norm Dist | 69.25 | 8.25 | 6 | | | 10.25 | | 3.75 | 1 | | 1.5 | 14 | 1 | 5.25 | 11.5 | 18.5 | 9 | 10 | | |
| Class 13 | Lin Discrim | | | 100 | | | | | | | | | | | | | | | | | |
| | KNN | 69.5 | 3.25 | 11.75 | 0.25 | | 9.25 | | 2.75 | 2 | | 0.5 | 10.75 | 8 | 5.75 | 12.5 | 10.25 | 9.75 | 12.5 | | 0.75 |
| | Fuzzy K | 36.5 | | 1.75 | 16.75 | | 0.25 | 19 | | 1.25 | 12.25 | | | 21.75 | | 0.5 | 14 | | 0.25 | 12.25 | |
| Ecoli | Norm Dist | 91.25 | 3 | 0.25 | | | 4 | | 0.5 | 0.5 | | 0.5 | 5.5 | 2.25 | 4.75 | 32.5 | 26.25 | 6.25 | 13.75 | | |
| washed | Lin Discrim | | | 100 | | | | | | | | | | | | | | | | | |
| Class 14 | KNN | 91 | 1.5 | 2.25 | | | 3.75 | | | 1.25 | | | 8.25 | 10 | 3.75 | 33.5 | 16.5 | 5.5 | 13.5 | | 0.25 |
| | Fuzzy K | 45.5 | | 0.75 | 20 | | | 16 | | | 9.75 | | | 23.75 | | 0.25 | 21.5 | | | 8 | |
| Ecoli | Norm Dist | 92 | 2 | 2.25 | | | 3 | | | | | 0.75 | 3.25 | 1.25 | 2.75 | 3 | 40.5 | 6.75 | 34.5 | | |
| unwashed | Lin Discrim | | | 100 | | | | | | | | | | | | | | | | | |
| Class 15 | KNN | 93 | 1.75 | 2.75 | | | 2.25 | | | | | 0.25 | 4 | 14 | 3 | 7.5 | 32.25 | 7.5 | 24.75 | | |
| | Fuzzy K | 36 | | | 22.75 | | | 15.25 | | | 19.25 | | | 23.25 | | | 12.75 | | | 6.75 | |

**Table A1.5 Combination five: size +fluorescence data (continued)**

| Mat | Method | Group % | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Erwinia washed | Norm Dist | 64.25 | 12 | | | | 20.75 | | 2.25 | 0.25 | | 0.25 | 0.25 | 1 | 1.5 | 3.25 | 19.5 | 24 | 14.75 | 0.25 | |
| | Lin Discrim | | | 99.75 | | 0.25 | | | | | | | | | | | | | | | |
| Class 16 | KNN | 76 | 6.75 | 0.75 | 1 | | 13.5 | | 1.25 | 0.5 | | | 0.5 | 12 | 3 | 4 | 13 | 30.75 | 12.75 | 0.25 | |
| | Fuzzy K | 35.75 | | 1.5 | 18 | | | 14.5 | | 0.5 | 15 | | | 26.5 | | | 9.25 | | | 14.75 | |
| Erwinia unwashed | Norm Dist | 89 | 2 | 3.75 | | | 4.75 | | | | | 0.25 | 2.5 | 1.25 | 3.5 | 8 | 32 | 8.25 | 33.5 | | 0.25 |
| | Lin Discrim | | | 100 | | | | | | | | | | | | | | | | | |
| Class 17 | KNN | 91 | 2 | 3.5 | | | 2.75 | | 0.25 | 0.25 | | | 5.5 | 11.5 | 2.5 | 11.5 | 26.75 | 6.25 | 27 | | 0.25 |
| | Fuzzy K | 38.25 | | | 21 | | | 14 | | | 18 | | | 19.5 | | | 18.5 | | 0.25 | 8.75 | |
| P mullbery | Norm Dist | 48.5 | | 0.75 | 1.25 | | 1 | | 25.25 | | 11.75 | 10.75 | | | 0.25 | 0.25 | | | 0.25 | 47.75 | 0.75 |
| Class 18 | Lin Discrim | 0.25 | 6.5 | 70 | | 21 | | | | 2.25 | | | | | | | | | | 0.25 | |
| | KNN | 30.25 | 4.25 | 24.5 | 4.5 | 1 | | 1 | 14 | 0.5 | 10.5 | 8.25 | | | 1 | | 0.25 | | | 29.75 | 0.5 |
| | Fuzzy K | 9.25 | | 4.75 | 9.75 | | 0.75 | 9.5 | | 0.25 | 19.75 | | 4.25 | 27.75 | | 0.75 | 11.75 | | 1.5 | 9.25 | |
| Puff ball | Norm Dist | 53.75 | 7 | | | 0.75 | | 6.25 | 10.5 | 15.75 | 4.25 | 1.75 | | | | | | | | 8.25 | 45.5 |
| Class 19 | Lin Discrim | | 15.75 | 44.5 | | 35.5 | | | | 4.25 | | | | | | | | | | | |
| | KNN | 29 | 4.75 | | 1.5 | 3.25 | 0.25 | 13.5 | 4 | 36.5 | 4 | 3 | | | | | | 0.25 | | 0.5 | 28.5 |
| | Fuzzy K | 14.75 | | 7.75 | 12.5 | | 5.5 | 3.75 | | 4.25 | 1.25 | | 11.5 | 7.25 | | 13.5 | 4.5 | | 13.5 | 14.75 | |
| | Class: | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |

**Table A1.5 Combination five: size +fluorescence data (continued)**

# References


1. Marple, V.A., *History of Impactors-the First 110 Years.* Aerosol Science and Technology, 2004. **38**(3): p. 247-292.

2. Harrison, R.M. and R.E.v. Grieken, eds. *Atmospheric Particles*. IUPAC Series on Analytical and Physical Chemistry of Environmental Systems, ed. J. Buffle and H.P.v. Leeuwen. Vol. 5. 1998, John Wiley & Sons.

3. Burge, H.A., ed. *Bioaerosols*. Indoor Air Research Series, ed. M. Eisenberg. 1995, Lewis Publishers: London.

4. *European Working Group for Legionella Infections*. www.ewgli.org/ Last updated 11/02/2005.

5. Fluckiger, B., T. Koller, and C. Monn, *Comparison of airborne spore concentrations and fungal allergen content.* Aerobiologia, 2000. **16**: p. 393-396.

6. Levetin, E., *Fungi*, in *Bioaerosols*, H.A. Burge, Editor. 1995, Lewis Publishers: London.

7. Making the UK safer: detecting and decontamination chemical and biological agents, The Royal Society. 2004

8. Miller, S.L., Y.S. Cheng, and J.M. Macher, *Guest Editorial.* Aerosol Science and Technology, 1999. **30**(2): p. 93-99.

9. Li, C.S., M.L. Hao, W.-H. Lin, C.-W. Chang, and C.-S. Wang, *Evaluation of Microbial Samplers for Bacterial Microorganisms.* Aerosol Science and Technology, 1999. **30**: p. 10-108.

10. Shahamat, M., M. Levin, I. Rahman, C. Grim, J. Heidelberg, G. Stelma, and R. Colwell, *Evaluation of media for recovery of aerosolized bacteria.* Aerobiologia, 1997. **13**(14): p. 219-226.

11. Griffiths, W.D., I.W. Stewart, J.M. Clark, and I.L. Holwill, *Procedures for the characterisation of bioaerosol particles. Part II: Effects of environment on culturability.* Aerobiologia, 2001. **17**: p. 109-119.

12. Roselle, D.C., M. Seaver, and J.D. Eversole, *Changes in Intrinsic fluorescence during the production of viable but non culturable Escherichia coli.* Journal of Industrial Microbiology & Biotechnology, 1998. **20**: p. 265-267.

13. Lighthart, B. and A.J. Mohr, *Atmospheric Microbial Aerosols*. 1994, London: Chapman & Hall.

14. Fox, A., *Report of the "Bioterrorism Workshop" Duke University April 2-4, 2002.* Journal of Microbiological Methods, 2002. **51**: p. 247-254.

15. Pan, Y.-l., S. Holler, R.k. Chang, S.C. Hill, R.G. Pinnick, N. S, and J.R. Bottiger, *Single-shot fluorescence spectra of individual micrometer sized bioaerosols iluminated by 351 or a 266nm ultraviolet laser.* Optics Letters, 1999 January 15. **24**(2): p. 116-118.

16. Hairston, P.P., J. Ho, and F.R. Quant, *Design of an Instrument for Real Time Detection of Bioaerosols Using Simultaneous Measurement of Partice Aerodynamic Size and Intrinsic Fluorescence.* Journal of Aerosol Science, 1997. **28**(3): p. 471-482.

17. Seaver, M., J.D. Eversole, J.J. Hardgrove, W.k. Cary.Jr, and D.C. Roselle, *Size and Fluorescence Measurements for Field Detection of Biological Aerosols.* Aerosol Science and Technology, 1999. **30**(2): p. 174-185.

18. Du., H., R.A. Fuh., J. Li., A. Corkan., and J.S. Lindsey., *Photochemcad: A computer aided design and research tool in photochemistry.* Photochemistry and Photobiology, 1998. **68**: p. 141 -142.

19. Campillo, A.J. and H.B. Lin, *Absorption and Fluorescence Spectroscopy of Aerosols*, in *Optical Effects Associated With Small Particles*, P. WBarber and R.K. Chang, Editors. 1988, Worl Scientific: Singapore.

20. Brown, S.B., ed. *An Introduction to Spectroscopy for Biochemists*. 1980, Academic Press: London. 403.

21. Kunnil, J., S. Sarasanandarajah, E. Chacko, and L. Reinisch, *Fluorescence quantum efficiency of dry Bacillus globiggi spores.* Optics Express, 2005. **13**(2): p. 8969 - 8979.

22. Li, J.K., *Monitoring Cell Concentration and Activity by Multiple Excitation Fluorometry.* Biotechnology Progress, 1991. **7**(1): p. 21-27.

23. Dalterio, R.A., W.H. Nelson, D. Britt, J. Sperry, D. Psaras, J.F. Tanguay, and S.L. Suib, *Steady-State and Decay Characteristics of Protein Fluorescence from Bacteria.* Applied Spectroscopy, 1986. **40**(1): p. 86-90.

24. Dalterio, D., W.H. Nelson, D. Britt, J. Sperry, D. Psaras, J.F. Tanguay, and S.L. Suib, *Steady-State and Decay Characteristics of Protein Tryptophan Fluorescence from Bacteria.* Applied Spectroscopy, 1986. **40**(1): p. 86-90.

25. Dalterio, D., W.H. Nelson, D. Britt, J. Sperry, J.F. Tanguay, and S.L. Suib, *The Steady-State and Decay Characteristics of Primary Fluorescence from Live Bacteria.* Applied Spectroscopy, 1987. **41**(2): p. 234240.

26. Kinkennon, A.E., *Fluorescence lifetime Characterization of Bacteria using Total Lifetime Distribution Analysis with the Maximum Entropy Method.* Journal of Fluorescence, 1997. **7**(3): p. 201-210.

27. Seaver, M., D.C. Roselle, J.F. Pinto, and J.D. Eversole, *Absolute emission spectra from Bacillus subtilis and Escherichia coli vegetative cells in solution.* Applied Optics, 1998 August 1. **37**(22): p. 5344-5347.

28. Cheng, Y.S., E.B. Barr, B.J. Fan, P.J. Hargis, D.J. Rader, T.J. O'Hern, J.R. Torczynski, G.C. Tisone, B.L. Preppernau, S.A. Young, and R.J. Radloff, *Detection of Bioaerosols Using Multiwavelength UV Fluorescence Spectroscopy.* Aerosol Science and Technology, 1999. **30**: p. 186-201.

29. Faris, G.W., R.A. Copeland, K. Mortelmans, and B.V. Bronk, *Spectrally resloved absolute fluorescence cross sections for bacillus spores.* Applied Optics, 1997 February 1. **36**(4): p. 958-967.

30. Seaver, M., J.F. Pinto, and J.D. Eversole, *Fluorescence Spectroscopy of B.subtilis (globigi) and E.coli in solution.* OSA TOPS on Biomedical Optical Spectroscopy and Diagnostics, 1996. **3**: p. 263-268.

31. Niessner, R. and A. Krupp, *Detection and Chemical Characterization of Polycyclic Aromatic Hydrocarbon Aerosols by Means of Laser-Induced Fluorescence.* Part.Part.Syst. Charact., 1991. **8**: p. 23-28.

32. Cignoli, F., S. Benecchi, and D. Tencalla, *Atlas Of Fluorescence Spectra Of Aromatic Hydrocarbons*. Vol. Vol. 1: Istituto Tecnologie dei Materiali e dei Processi Energetici. 212.

33. Reyes, F.L., T.H. Jeys, N.R. Newbury, C.A. Primmerman, G.S. Rowe, and A. Sanchez, *Bio-Aerosol Fluorescence Sensor.* Field Analytical Chemistry and Technology, 1999 27 May. **3**(4-5): p. 240 - 248.

34. Agranovski, V., Z. Ristovski, P. Blackall, and L. Morawska, *Real-time detection of bioaerosols at a piggery.* Journal of Aerosol Science, 2000. **31**(Supplement 1): p. 739-740.

35. Agranovski, V., Z.D. Ristovski, P. Blackall, and L. Morawska, *Field evaluations of the UVAPS performance.* Journal of Aerosol Science, 2004. **35**(Supplement 1): p. 599-612.

36. Pinnick, R.G., S.C. Hill, P. Nachman, J.D. Pendleton, G.L. Fernandez, M.W. Mayo, and J.G. Bruno, *Fluorescence Particle Counter for Detecting Airborne Bacteria and Other Biological Particles.* Aerosol Science and Technology, 1995. **23**: p. 653-664.

37. Pinnick, R.G., S.C. Hill, P. Nachman, G. Videen, G. Chen, and R.K. Chang, *Aerosol Fluorescence Spectrum Analyzer for Rapid Measurement of Single Micrometer Sized Airborne Biological Particles.* Aerosol Science and Technology, 1998. **28**(2): p. 95-104.

38. Chen, G., P. Nachman, R.G. Pinnick, S.C. Hill, and R.k. Chang, *Conditional-firing aerosol-fluorescence spectrum analyzer for individual airborne particles with pulsed 266nm excitation.* Optics Letters, 1996 August 15. **21**(16): p. 1307-1309.

39. Hill, S.C., R.G. Pinnick, S. Niles, Y.-L. Pan, S. Holler, R.K. Chang, J. Bottiger, B.T. Chen, C.-S. Orr, and G. Feather, *Real-Time Measurement of Fluorescence Spectra from Single Airborne Biological Particles.* Field Analytical Chemistry and Technology, 1999. **3**(4-5): p. 221-239.

40. Brosseau, L.M., D. Vesley, N. Rice, K. Goodell, M. Nellis, and P. Hairston, *Differences in Detected Fluorescence Among Several Bacterial Species Measured with a Direct-Reading Particle Sizer and Fluorescence Detector.* Aerosol Science and Technology, 2000. **32**(6): p. 545-558.

41. Ristovski, Z. and V. Agranovski, *Real-time monitoring of viable bioaerosols: capability of the UVAPS to predict the amount of individual microorganisms in aerosol particles.* sbmitted to Journal of Aerosol Science, 2004.

42. Eversole, J.D., J.J. Hardgrove, W.K.C. Jr, D.P. Choulas, and M. Seaver, *Continuous, Rapid Biological Aerosol Detection with the Use of UV Fluorscence; Outdoor Test Results.* Field Analytical Chemistry and Technology, 1999. **3**(4 - 5): p. 249 - 259.

43. Sivaprakasam, V., A.L. Huston, C. Scotto, and J.D. Eversole, *Multiple UV wavelength excitation and fluorescence of bioaerosols.* Optics Express, 2004. **12**(19): p. 4457-4466.

44. Willeke, K. and P.A. Baron, eds. *Aerosol Measurement.* 1994, Van Nostrand Reinhold.

45. Bohren, C.F. and D.R. Huffman, *Absorption and Scattering of Light by Small Particles.* 1983: Wiley Inter-Science.

46. Arakawa, E.T., P.S. Tuminello, B.N. Khare, and M.E. Milham, *Optical Properties of Erwinia herbicola Bacteria at*
*0.190–2.50 um.* BIOPOLYMERS, 2003. **72**(5): p. 391-398.

47. Bazhan, V. Scatlab www.scatlab.com/index.html. Last visited:30/09/2005

48. Hirst, E., P.H. Kaye, and J.R. Guppy, *Light scattering from nonspherical airborne particles: experimental and theoretical comparisons.* Applied Optics, 1994 October 20. **33**(30): p. 7180-7186.

49. Wriedt, T., *A Review of Elastic Light Scattering Theories.* Particle and Particle Systems Characterization, 1998. **15**: p. 67-74.

50. Jones, A.R., *Light scattering for particle characterization.* Progress in Energy and Combustion Science, 1999. **25**(1): p. 1-53.

51. Ye, M., S. Wang, Y. Lu, T. Hu, Z. Zhu, and Y. Xu, *Inversion of particle -size distribution from angular light-scattering data with genetic algorithms.* Applied Optics, 1999. **38**(12): p. 2677-2685.

52. Umhauer, H. and M. Bottlinger, *Effect of particle shape and structure on the results of single-particle light scattering size analysis.* Applied Optics, 1991. **30**(33): p. 4980 - 4986.

53. Sloot, P.M.J., A.G. Hoekstra, H.a.d. Liet, and C.G. Figden. *Arbitrarily shaped particles measured in flow through systems.* in *Proceedings of the 2nd international congress on optical particle sizing*. 1990.

54. Spinrad, R.W. and J. Brown, *Effects of asphericity on single-particle polarized light scattering.* Applied Optics, 1993. **32**(30): p. 6151- 6158.

55. Latimer, P., A. Brunsting, B.E. Pyle, and C. Moore, *Effectrs of asphericity on single particle scattering.* Applied Optics, 1978. **17**(19): p. 3152 - 3158.

56. Mugnai, A. and W.J. Wiscombe, *Scattering from nonspherical Chebyshev particles. Variability in angular scattering patterns.* Applied Optics, 1989. **28**(15): p. 3061 -3073.

57. Bartholdi, M., G.C. Salzman, R.D. Hiebert, and M. Kerker, *Differential light scattering photometer for rapid analysis of single particles in flow.* Applied Optics, 1980. **19**(10): p. 1573 - 1581.

58. Ludlow, I.L., *Particle analysis by light scattering.* UK Ministry of Defence report ER1A/9/4/2161/CDE, 1982.

59. Kaye, P.H., K. Alexander-Buckley, E. Hirst, and S. Saunders, *A real-time monitoring system for airborne particle shape and size analysis.* Journal of Geophysical Research, 1996 August 27. **101**(D14): p. 19,215-19,221.

60. Wyatt, P.J., J.R. Bottiger, K.L. Schehrer, and S.D. Phillips, *Aerosol particle analyzer.* Applied Optics, 1988. **27**(2): p. 217-221.

61. Dick, W.D., P.H. McMurry, and J.R. Bottiger, *Size- and Compositional - Dependent Response of the DAWN-A Multiangle Single - Particle Optical Detector.* Aerosol Science and Technology, 1994. **20**: p. 345-362.

62. Kaye, P.H., N.A. Eyles, J.M. Clark, and I.K. Ludlow. *A laser scattering instrument for airborne particle size and shape classification.* in *2nd International congress on optical particle sizing*. 1990 March. University of Arizona Phonex USA.

63. Hirst, E. and P.H. Kaye, *Experimental and theoretical light scattering profiles from spherical and nonspherical particles.* Journal of Geophysical Research, 1996 August 27. **101**(D14): p. 19,231-19,235.

64. Pan, Y.-L., k. B, Aptowicz, R.K. Chang, M. Hart, and J.D. Eversole, *Characterizing and monitoring respiratory aerosols by light scattering.* Optics Letters, 2003. **28**(8): p. 589 - 592.

65. Real-time detection, Characterisation and Monitoring of Respirable Airborne Particles, University of Hertfordshire
Bristol Industrial and Research Associates Ltd. 1997 January

66. Kaye, P.H., E. Hirst, and Z.W. -Thomas, *Neural-network based spatial light scattering instrument for hazardous airborne fiber detection.* Applied Optics, 1997 August 20. **36**(24): p. 6149-6156.

67. Zerull, R.H., R.H. Giese, and K. Weiss, *Scattering functions of nonspherical dielectric and absorbing particles vs Mie Theory.* Applied Optics, 1977. **16**(4): p. 777 - 778.

68. Sachweh, B.A., W.D. Dick, and P.H. McMurry, *Distinguishing Between Spherical and Nonspherical Particles by Measuring the Variability in Azimuthal Light Scattering.* Aerosol Science and Technology, 1995. **23**: p. 373-391.

69. Wiscombe, W.J. and A. Mugnai, *Scattering from nonspherical Chebyshev particles. Means of angular scattering patterns.* Applied Optics, 1989. **27**(12): p. 2405 - 2421.

70. Barton, J. and P.H. Kaye Multi-Parameter Aerosol Monitor Final Report, University of Hertfordshire. 1999

71. Ulanowski, Z., Z.N. Wang, P.H. Kaye, and I.K. Ludlow, *Application of neural networks to the inverse light scattering problem for spheres.* Applied Optics, 1998. **37**(18): p. 4027-4033.

72. Wang, Z., Z. Ulanowski, and P.H. Kaye, *On solving the inverse scattering problem with RBF neural networks: noise-free case.* Neural Computing Applications, 1999. **8**(2): p. 177-186.

73. Pan, Y.L., S.C. Hill, J.P. Wolf, S. Holler, R.K. Chang, and J.R. Bottiger, *Backward-enhanced fluorescence from clusters of microspheres and particles of tryptophan.* Applied Optics, 2002. **41**(15): p. 2994-2999.

74. Hirst, E., *Airborne Particle Size and Shape Classification from Spatial Light Scattering Profiles*, in *STRC*. 1993, University of Hertfordshire.

75. Murrell, W.G., *Chemical Composition of Spores and Spore Structures*, in *The Bacterial Spore*, G.W. Gould and A. Hurst, Editors. 1969, Academic Press: London. p. 215-273.

76. Griffiths, W.D., P.J. Iles, and N.P. Vaughan, *The behaviour of liquid droplet aerosols in an APS 3300.* Journal of Aerosol Science, 1986. **17**(6): p. 921-930.

77. Secker, D., P. Kaye, and E. Hirst, *Real-time observation of the change in light scattering from droplets with increasing deformity.* Optics Express, 2001. **8**(6): p. 290-295.

78. Rubel, G.O. and K.H. Fung, *Influence of saline media on the fluorescence emission of Bacillus spores.* Applied Optics, 1999. **38**(31): p. 6673-6676.

79. www-micro.msb.le.ac.uk/224/myc2.html, *Reproduction of Fungi*. Last updated 02/02/2005, University of Leicester.

80. Duda, R.O. and P.E. Hart, *Pattern Classification and Scene Analysis*. 1973, New York: Wiley-Interscience. 428.

81. Hand, D.J., *Discrimination and Classification*. 1981, Chichester: John Wiley and Sons. 218.

82. *Pattern Recognition for HCI*
www.cs.princeton.edu/courses/archive/fall04/cos436/Duda/PR_home.htm,
Last updated 1997, Department of Electrical Engineering San Jose State University.

83. Pan, Y.-L., V.e. Boutou, J.R. Bottiger, S.S. Zhang, J.-P. Wolf, and R.K. Chang, *A Puff of Air Sorts Bioaerosols for Pathogen Identification.* Aerosol Science and Technology, 2004. **38**(6): p. 598-602.