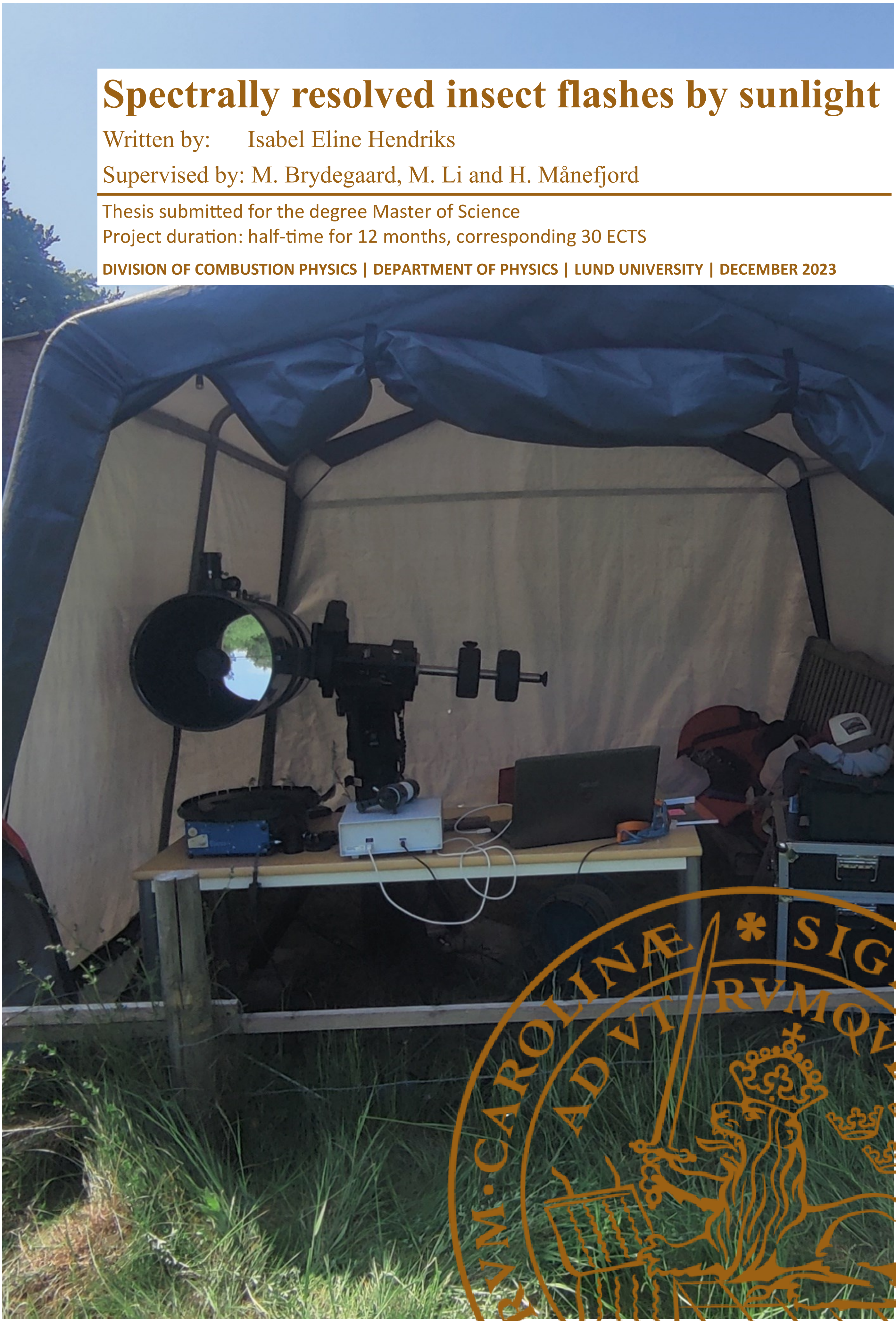
Spectrally resolved insect flashes by sunlight

Written by: Isabel Eline Hendriks
Supervised by: M. Brydegaard, M. Li and H. Månefjord

Thesis submitted for the degree Master of Science
Project duration: half-time for 12 months, corresponding 30 ECTS

DIVISION OF COMBUSTION PHYSICS | DEPARTMENT OF PHYSICS | LUND UNIVERSITY | DECEMBER 2023



Spectrally resolved insect flashes by sunlight

by Isabel Eline Hendriks

LUND
UNIVERSITY

(2023)
Thesis submitted for the degree Master of Science, General Physics
Project duration: Twelve months at half-speed
Thesis advisors: Dr. Mikkel Brydegaard, Meng Li and Hampus Månefjord
Division of Combustion Physics

The data presented in this thesis will be the source of a future publication.

Popular description

Safeguarding insect biodiversity by analyzing the sunlight they reflect

Insects are an incredibly important part of our ecosystem. They serve as pollinators for many of the fruits and vegetables that we eat. Due to climate change, the number of insects and the amount of insect species is declining at an alarmingly high rate, which can have catastrophic consequences for all other life on Earth. It is important to closely monitor this decline so that we can combat it effectively. However, it is difficult to find out how and where this decline is taking place. Current methods for insect monitoring rely on manual counting which is labor intensive. Automatic monitoring of insects could give better insight into this issue, which would allow us to combat it more effectively.

This thesis presents an automatic method to monitor insects by capturing and analyzing the spectrum of sunlight that is reflected by free-flying insects. This method is also referred to as Dark-Field Spectroscopy. A large advantage over manual counting of insects is that this does not harm the insects. Additionally, insects use sunlight to see each other. Therefore, these measurements can also provide clues about the mechanism behind how insects see each other. The research question that forms the basis of this thesis is whether the spectral analysis of the reflected sunlight contains extra information for identifying the species.

The sunlight reflected by insects is captured by a large telescope that is placed in a field. A big black box is placed 100 meters away from the telescope. When there is nothing in the field of view of the telescope, it only sees the black box. When an insect flies through the field of view, it will reflect sunlight into the telescope. The light captured by the telescope is then sent into a spectral analyzer which can determine the different wavelengths that were reflected by the insect, which is also referred to as the spectrum. The spectrometer takes 20,000 measurements every second. The fastest wing beat is from mosquitoes, who beat their wings 1,000 times per second. This means that their spectrum is recorded 20 times during one wing beat.

During three days of recording in the field, 24.000 insects were observed flying through the field of view. The collected data is rich in information. The wing beat frequency is determined by light intensity fluctuations. Opaque wings exhibit thin-film interference, causing spectral fringes from which the wing thickness can be derived. The entire dataset is used to estimate the number of species through clustering, which means grouping similar observations to determine the number of species present.

The data clustering was done with and without the spectral information. Clustering with spectral information resulted in more groups compared to clustering without spectral information. The main conclusion of this project is that the spectral information helps to identify more species. Additionally, the data showed that sunlight reflected by insects has up to eight spectral components. This could be an explanation for why, for example, a dragonfly has five effective spectral bands.

Abstract

Insects are vital for the environment and biodiversity. Optics-based remote techniques can quickly and effectively identify and track insects in their natural habitat. This has the potential to speed up and improve the time-consuming process of manually monitoring insect diversity. This thesis presents a method to remotely sense insects by analyzing the sunlight reflected by free-flying insects. The study proposes a novel approach by simultaneously capturing the spectral and temporal signals of reflected sunlight, containing new information about the spectral properties of insects in the visible wavelength range. For this project, a spectrometer is designed using raytracing, CAD-design, and 3D-printing. The instrument is capable of analyzing light in the visible to near-infrared range (420-840 nm) with 32 spectral bands. The instrument was calibrated and successfully deployed in the field to capture sunlight flashes from free-flying insects, with a sampling rate of 20 kHz. The collected data shows that the instrument successfully captured signals from insects. This data is used to determine if spectrally resolved modulation provides additional information for species differentiation. To achieve this, a hierarchical clustering algorithm is applied to both the spectrally resolved and non-spectrally resolved signals. The number of clusters found when using both spectral and temporal data was larger than the number of clusters found when using only temporal data. The conclusion is drawn that there is additional information for species identification in the spectral data.

Abbreviations

CAD	Computer-aided Design
DFS	Dark Field Spectroscopy
ELM	Electron Leakage Matrix
FoV	Field of View
NIR	Near Infrared
IQR	Interquartile Range
LiDAR	Light Detection And Ranging
MA-PMT	Multi Anode Photon Multiplier Tube
OCS	Optical Cross Section
PCB	Printed Circuit Board
RMS	Root Mean Square
STFT	Short-Time Fourier Transform
SVD	Singular Value Decomposition
VIS	Visible
WBF	Wing Beat Frequency
WISs	Wing Interference Signals

Acknowledgements

I would like to express my gratitude to the following people for their assistance this project.

To Mikkel Brydegaard, for setting up an incredibly interesting project that allowed me to develop many new skills, and for all the input and guidance during this project.

To Meng Li for all of the support in this project, both practical and emotional. For borrowing her car through the fieldwork period, identifying the captured insect species, helping out in the field, and assisting with all parts of the data analysis, especially the data clustering and single value decomposition. On top of that, for making sure there was always candy in her office and always having an open office door to provide moral support throughout the entire project.

To Hampus Månefjord, for sharing his knowledge in both Zemax and Fusion, assisting during the field trials and heroically removing a bee nest from the black box at Stensoffa.

To Lauro Müller, for sharing his knowledge in Zemax and Fusion and assisting during the field trials.

To Alice Asp and Despina Demitriadou, for helping out in the field.

To Love Kildetoft, Henrik Palme, Julia Lövgren, Klara Gerdhem and Alex Petersson for sharing an office with me.

Table of Contents

1 Introduction

1.1 Motivation

In recent years, there has been a growing concern in the scientific community about the decline of insect diversity and abundance [1, 2]. Insects play a crucial role in various ecosystems and are incredibly important to maintain the health of our planet [3]. They serve as pollinators which is a crucial step in food production. The decline of insect abundance can lead to poor agricultural production and a weaker ecosystem in general. Conventional methods for insect monitoring are not sufficient. The existing taxonomic workforce is expensive and labor-intensive. Automated mapping of insect diversity could give clues about the reasons behind this decay, allowing us to mitigate the decline more effectively.

Our biophotonics group has been investigating the potential of optical-based techniques in the field of remote sensing of insects. The advantages of optical remote sensing over manual counting are reduced labor, a larger measurement range, and improved temporal resolution. Additionally, optical remote sensing is non-lethal, meaning that no harm is done to an already fragile group of animals. This makes optical remote sensing an important tool for monitoring the biosphere [4, 5].

A well-established method for remote insect detection is Light Detection And Ranging (LiDAR) [6, 7], which is based on emitting light into the atmosphere and analyzing the backscattered light. It has proven to be a useful method for monitoring gases [8], aerosols [9, 10], vegetation [11, 12], waterbodies [13] and aerofauna [4]. Insect signals can also be measured by utilizing the sun as a light source, a method also known as passive LiDAR or Dark Field spectroscopy [14, 15].

In earlier entomological LiDAR studies, our group has preliminary classified insect signals based on the modulation power spectra derived from LiDAR waveforms, which contain information about polarization and wing beat frequency. Although incorporating polarization features into insect classification has shown promise in enhancing specificity [16], thereby aiding in determining insect diversity, the search for more stable and reliable characteristics of insects during flight, which could improve the accuracy of species classification, remains ongoing. Our team has discovered that the wing interference signals, referred to as fringes in the following text, exhibits significant potential in advancing the identification of insect species and sex [17]. Therefore, in this study, we propose that capturing both the insect dynamic waveforms and wing interference signals simultaneously and integrating these datasets, may lead to even better specificity in insect signal classification in a photonic-based detection system.

Previous LiDAR experiments have shown to be capable of remotely measuring size, wing beat frequency, and wing thickness [18]. Comparing the measured insect features to known insect data can provide insight into which insect was observed.

1.2 State of the art

A recent study on hoverfly specimens showed that spectral fringes, also called wing interference signals (WISs), from clear wings are species and sex-specific among the examined specimens [17, 19]. Such property demonstrates great potential for improving accuracy in insect classification. Figure 1 below shows a scatter plot of the wing modulation and the wing thickness that was found. The data points in the scatter plot have been color-coded to reflect minimized genetic distances between neighboring species. This color-coding helps to visualize species that are genetically closer, allowing for easier interpretation of similarities and differences in their wing properties. As demonstrated by Figure 1, species from the same genera or tribe, identified by similar color coding, tend to exhibit similar fringe properties. Notably, while there is some overlap among neighboring species in the 2D scatter plot, the majority of the examined species show distinct, non-overlapping characteristics.

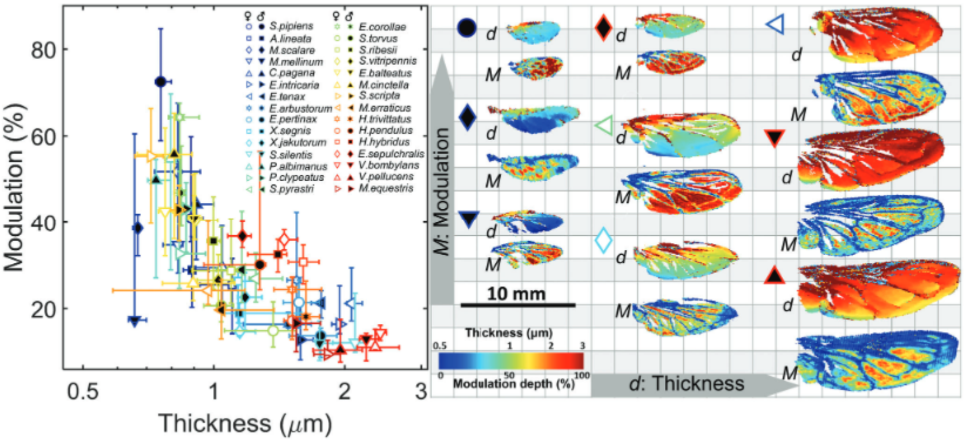


Figure 1: Membrane thickness and wing modulation scatter plot. Figure taken from [17].

In line with this approach, a hyperspectral LiDAR with 64 spectral bands was built that operates at a repetition rate of 500 Hz and is capable of resolving in spatial, spectral, and temporal domains [20]. The spectrometer used in this experiment was designed using ray tracing and fabricated with 3D printing and was shown to have a high signal-to-noise ratio. The device was employed at the Stensoffa field station to measure live insect signals. The experiment was done with a super-continuum white light source, and it was executed at night time to avoid sunlight interfering with the experiment. The instrument was capable of detecting spectral fringes from a damsel fly wing at 87 meters distance from the detector. The retrieved signal corresponded to a thickness of 1412 nm with a 4 nm precision. This finding was consistent with a measurement done in a controlled laboratory environment, confirming the photonic-based system's ability to capture spectral fringes remotely [21].

Furthermore, information about free-flying insects can be obtained by analyzing the sunlight that is backscattered by them. This is also referred to as passive measurements or Dark Field Spectroscopy (DFS). The use of DFS was shown to be a promising concept for inexpensive insect observation [15]. In this study, a spectrometer with a spectral range of 345 to 1050 nm was used with a 50 Hz repetition rate. This experiment showed that an inexpensive setup was able to successfully identify two different damselfly species and distinguish between their sexes. The relative abundances were consistent with manual counts. The experiment also managed to observe movement patterns.

Ranging information on flying insects can be achieved without the need for active illumination [14]. Here, a quadrant detector with a sampling rate of 20 kHz was used. This study revealed that signal patterns across various detector segments could be effectively analyzed to determine the distance to airborne insects. For instance in this study, at close range the signals show significant overlap, but as the distance increases to medium range this overlap reduces, and the signals become clearer. At far range the signals are sharp with very little overlap. Additionally, at close range the recorded signals in adjacent detector elements are identical, demonstrating the system's high sensitivity to objects that are nearby. However, these signals start to differ as the distance increases. This study demonstrates that photonic-based systems can also provide accurate ranging information on flying insects without employing the Scheimpflug LiDAR approach.

1.3 Current Limitations

One limitation of entomological LiDAR is that a light source in the visible wavelength range is disturbing to the surrounding environment. Additionally, it can induce a bias in the results since the insects can see the light and be attracted to it. Therefore, current LiDAR experiments usually use a light source that emits light in the invisible wavelength range. Using the sunlight as a light source allows us to measure scattering in the visible range without disturbing the environment. This will result in an unbiased signal in the visible wavelength range, which is something that has not been measured before. Additionally, signals in the visible wavelength range contain information about clues in visual ecology among insects [22].

However, measuring insect flashes poses a technical challenge. Insect wing beat frequencies can be as high as 1 kHz. The angular size of the sun is 0.5° [23], so a flash duration down to 4 μs can be expected. Resolving these flashes requires an instrument with a sampling range in the kHz region that is susceptible to low light intensities. Such instruments are typically not commercially available and therefore are required to be custom-made.

1.4 Project Aim

The primary objective of this master thesis is the remote measurement of WISs using DFS. The project consists of two parts. First, the design and construction of a spectrometer, and second, the analysis of the collected data. Inspired by the accuracy and performance of the device designed and employed by L. Müller [20], this project will use the same approach for instrument design, namely ray tracing and 3D printing. The detector used in this setup is a Multi-Anode Photo-Multiplier Tube array which is sensitive to low light intensities and can record at high repetition rates. It will be calibrated in the laboratory and then tested to record insect signals over a meadow in Stensoffa field station in Skåne, southern Sweden. The obtained data will be analyzed and the fringe formula will be used to identify any fringes captured in the data.

The main research question in this thesis is: Can spectrally resolved modulation provide additional information for differentiating species?

2 Theory

This section discusses the theoretical concepts that are relevant to this thesis. Firstly the physical scattering mechanisms that will occur in this experiment are discussed. After that, it is explained how this translates into insect signals.

2.1 Light scattering

The backscattered light from insects contains information about the insect such as size, wing beat frequency, and velocity. The more information is extracted from the insect signal, the higher the likelihood of correct classification. In this project the spectral and temporal signals will be captured simultaneously.

Light scattering from a surface can occur with and without energy transfer to the scattering material. Scattering with energy transfer is referred to as inelastic scattering while scattering without energy transfer is referred to as elastic scattering. In this thesis, only the latter is studied. An in-depth discussion of light scattering can be found in [24–27].

2.1.1 Coherent and incoherent scattering

Coherent scattering is scattering during which the phase of the light is preserved. When the phase of the incoming light is lost during the scattering, this is called incoherent scattering. Specular reflection and refraction are examples of coherent scattering.

2.1.2 Specular and Diffuse signals

The coherent reflection of light at the boundary of two materials is called specular reflection. Here, the incident light is divided into a reflected component and a refracted component. The angle of the reflected component is equal to the angle of the incoming light. The angle of the refracted light is determined by Snell's law, and the distribution of power between the reflected and refracted components is governed by the Fresnel equation. Both the angle and the power depend on the difference in refractive index between the two materials. In a clear medium, the refracted light will exit the medium without being scattered. Specular reflections are seen in glossy wings, such as mosquito wings or dragonfly wings. In the time domain, these signals appear as short but high peaks. This is because the light is reflected in one direction, so the signal is only visible when the wing is at precisely the right angle to reflect the light into the detector. This is also referred to as a flash. In the frequency domain, this signal contains a lot of harmonics [27].

Diffuse scattering occurs when the incident light enters the material and propagates through multiple scattering before it exits the material. When the light propagates through the material each photon is scattered at random angles. This causes each photon to take a different, random, path before it exits the material. Therefore, any coherence that the light may have had upon entering the material is lost during this process. The reflection is therefore incoherent. The reflected component does, however, contain information about the absorption spectrum of the sample, since the light was partially absorbed when propagating through the medium [26]. Diffuse scattering occurs in matte wings, such as moths. The angle of the light is not fixed, so these insect signals are similar to a sinusoidal wave in the time domain and result in few peaks in the frequency spectrum. The schematic in Figure 2 below demonstrates the difference between diffuse scattering and specular scattering.

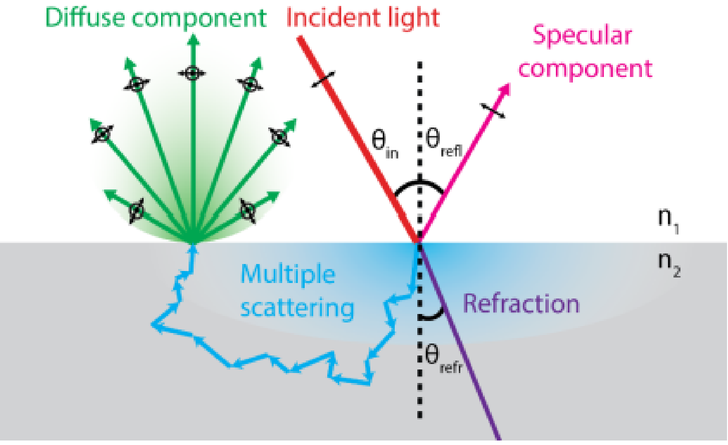


Figure 2: Schematic illustration visualizing the difference between specular scattering and diffuse scattering. In diffuse scattering, the light enters the material and propagates through the material through multiple scattering until it reaches a place to exit the material. In specular scattering, the light is partially reflected and partially refracted. Figure taken from [27].

2.2 Thin-film interference

When a ray of light encounters a thin film, a phenomenon called thin-film interference occurs. A schematic visualization of this process is shown in Figure 3. Thin film interference is caused by light waves reflecting off both the front and rear surfaces. The light reflected from the lower surface has traveled a longer path than the light reflected from the upper surface, inducing a phase shift in the light. The light waves will then interfere with each other. This interference is either constructive or destructive, depending on the phase shift induced by the film, which depends on the thickness of the film and the wavelength of the light. Equation 1 below shows the requirement for constructive interference. Here, $\lambda_{R_{max}}$ is the wavelength for constructive interference, n is the refractive index of the film, d is the thickness of the film and m is an integer number.

$$\lambda_{R_{max}} = \frac{2nd}{(m - \frac{1}{2})} \tag{1}$$

As a result, some wavelengths are enhanced while other wavelengths are suppressed. This leads to an interference pattern in the spectrum of light reflected from a thin film. The separation of the peaks in the fringe is directly related to the thickness of the film. This relationship is shown in equation 2 below.

$$\Delta\lambda = \frac{\lambda^2}{2nL\cos\theta_{in}} \tag{2}$$

Here, $\Delta\lambda$ is the separation between two adjacent peaks in the interference spectrum, λ the central wavelength of the nearest transmission peak, n is the refractive index of the film, L the thickness of the film, and θ_{in} the incident angle of the light. This relationship holds when the refractive index of the film is higher than its surrounding material, which is the case for an insect flying in the air.

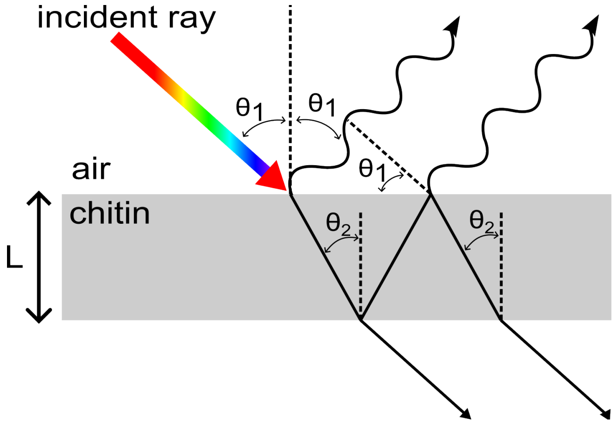


Figure 3: Schematic of thin-film interference

2.2.1 Fringe model

The fringe pattern observed in the spectrum of some insect signals can be used to calculate the wing thickness. The peaks and valleys in the interference pattern are referred to as a spectral fringe. The spacing between these fringes is determined by the wavelength (λ), the membrane thickness (L), and the angle of incidence (θ_{in}). Additionally, the overall reflectance of a thin-film reflector is influenced by the Fresnel equations, that considers factors such as polarization state, incidence angle (θ), and refractive index (n). The analytical model describing the total reflectance ($R_{thinfilm}$) of a thin-film membrane is detailed in equation 3, as provided by reference [28]. Since chitin is a common material for wing membranes, the refractive index of chitin is used in this model.

$$R_{thinfilm} = \frac{4R_{\text{Fresnel}}\sin^2\left(2\pi d\sqrt{n_{chitin} - \sin^2\theta}/\lambda\right)}{\left(1 - R_{Fresnel}\right)^2 + 4R_{Fresnel}\sin^2\left(2\pi d\sqrt{n_{chitin} - \sin^2\theta}/\lambda\right)} \tag{3}$$

The wavelength dependent refractive index for chitin (n_{chitin}) is expressed as $n_{\text{chitin}} = A + B/\lambda$ where $A = 1.517$ and $B = 8800\ \text{nm}^2$ [29]. The refractive index of air is assumed to be 1. Furthermore, d is the thickness of the wing, θ is the incident angle to the wing membrane and λ is the wavelength of the light. The reflection coefficient is governed by the Fresnel equations. The Fresnel equation for both states of polarization is given in equation 4 below. Sunlight is unpolarized, so the mean of the s-polarized and p-polarized Fresnel coefficient is used.

$$R_s = \left(\frac{\cos(\theta) - \sqrt{n_{\text{chitin}}^2 - \sin^2(\theta)}}{\cos(\theta) + \sqrt{n_{\text{chitin}}^2 - \sin^2(\theta)}}\right)^2 \qquad R_p = \left(\frac{\sqrt{1 - \left(\frac{\sin(\theta)}{n_{\text{chitin}}}\right)^2} - n_{\text{chitin}}\cos(\theta)}{\sqrt{1 - \left(\frac{\sin(\theta)}{n_{\text{chitin}}}\right)^2} + n_{\text{chitin}}\cos(\theta)}\right)^2$$

$$R_{\text{Fresnel}} = \frac{R_s + R_p}{2} \tag{4}$$

2.3 Signals from insects

2.3.1 Temporal and Frequency signal

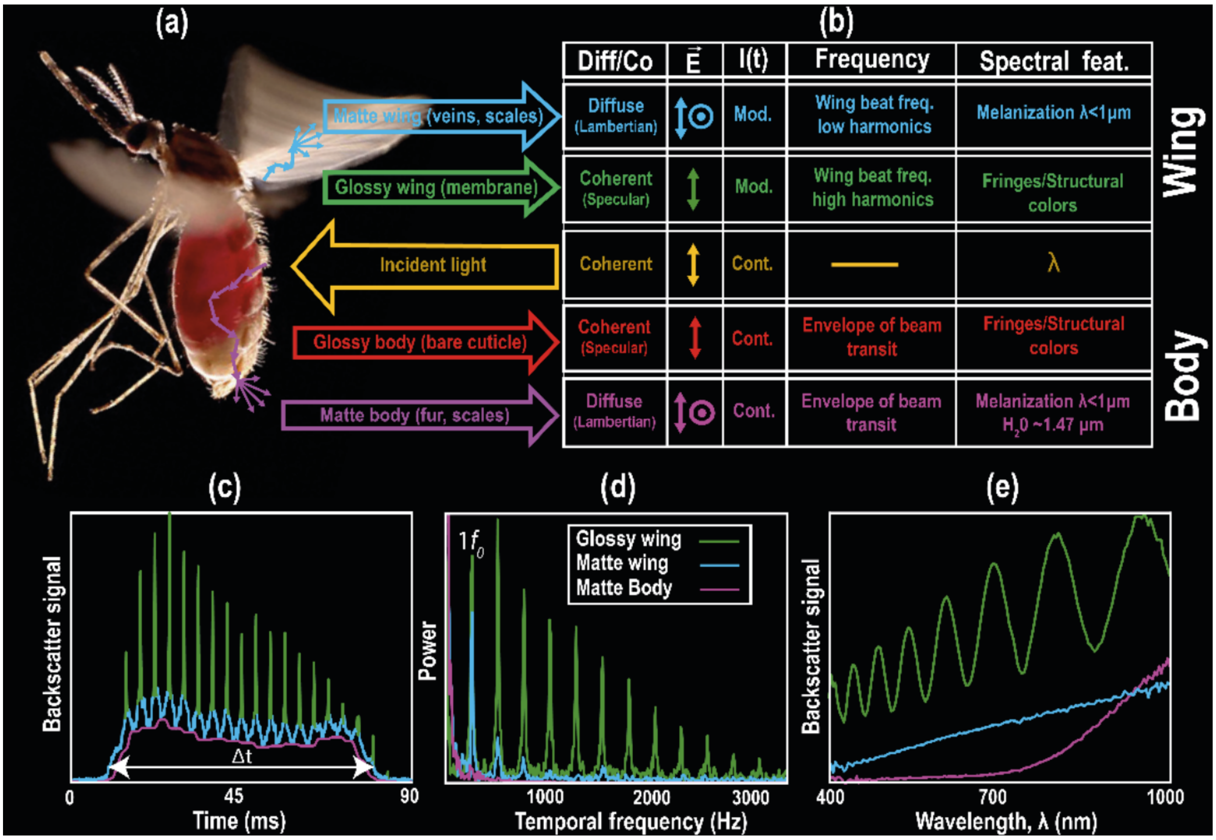
The temporal signal is the measured optical cross section (OCS) over time, which is proportional to the measured intensity counts. This is also referred to as the waveforms. This signal can be converted to the frequency domain through a Fourier Transform.

The temporal and frequency signals from an insect can be divided into a contribution from the body and the wing. The intensity signal is proportional to the OCS. As the insect beats its wings, the total recorded OCS oscillates. The frequency of this oscillation is equal to the wing beat frequency of the insect. The non-oscillating part of the signal is attributed to the body while the oscillatory component is attributed to the wing signal. As discussed in section 2.1.2 above, the matte part and glossy part of the insect can be distinguished by the number of harmonics in the signal. The skewness of a signal is the measure of asymmetry of the peaks. With regards to insect signals, a high skewness corresponds to a flashy signal implying a glossy wing. A low skewness corresponds to a symmetric, more sinusoidal signal, which in turn implies a matte wing.

Figure 4 below shows how the signal from an insect is separated into its components. The table 4(b) summarizes this discussion. Figure 4(c) shows the signal in the time domain where the body is the non-oscillating part, the matte wing is the sinusoidal-like oscillation and the glossy wing corresponds to the short and high peaks. Figure 4(d) shows the same signal in the frequency domain. Here, the body does not have any harmonics. The matte part of the wing has one strong frequency component and the glossy part of the wing has many frequency components.

2.3.2 Spectral signals

The spectral signal of insects is governed by their pigmentation and physical surface structure. Often the signal is a combination of both. Most insect bodies contain melanin, which has an absorption spectrum that shows a linear increase in wavelength starting at 700 nm [30], as displayed in Figure 4(e). The signal of a glossy wing contains information about the wing thickness due to thin-film interference. The spectral fringes due to interference effects are referred to as structural colors. Figure 4(e) below shows how the coherent backscatterd component contains a fringe. The matte wing and body do not contain any structural colors, but their spectrum increases with wavelength which is caused by melanin.



Diff/Co	$\vec{E}$	I(t)	Frequency	Spectral feat.	
Diffuse (Lambertian)	↕ ⊙	Mod.	Wing beat freq. low harmonics	Melanization λ<1µm	Wing
Coherent (Specular)	↕	Mod.	Wing beat freq. high harmonics	Fringes/Structural colors	
Coherent	↕	Cont.	—	λ	
Coherent (Specular)	↕	Cont.	Envelope of beam transit	Fringes/Structural colors	Body
Diffuse (Lambertian)	↕ ⊙	Cont.	Envelope of beam transit	Melanization λ<1µm H_20 ~1.47 µm	

Figure 4: Summary of how the signal from an insect is separated into body and wing signals and glossy and matte signals. The data in Figure (c) and (d) is from a flying fruit fly (Drosophila melanogaster) and the spectra in (e) are from a type of snipe fly. Figure taken from [27].

3 Relevant Components

This section describes all the relevant components used in this experiment. They are listed in the order in which the sunlight encounters the components.

3.1 Sunlight

The spectral radiance of the sun can be calculated using Planck's law which is shown in equation 5 below.

$$B_T(\lambda) = \frac{2hc^2}{\lambda^5}\frac{1}{e^{\frac{hc}{\lambda kT}}-1} \tag{5}$$

Here, $B_T(\lambda)$ is the spectral radiance at a given wavelength λ for a black body at temperature T, h is the Planck constant ($h \approx 6.626 \times 10^{-34}$ J·s), c is the speed of light in a vacuum ($c \approx 3 \times 10^8$ m/s) and k is the Boltzmann constant ($k \approx 1.380649 \times 10^{-23}$ J/K). The temperature of the sun is 5600 degrees Celsius [23].

3.2 Newton Telescope

The backscattered light is captured using a Newton Telescope with an F-number of F/4 ($f = 1200$ mm, $\phi = 300$ mm). The telescope is mounted on a tripod with a motor that is used for the fine adjustment of the position and angle. The telescope is focused on a black termination box at a 100-meter distance from the telescope. The collected signal is focused onto the slit through which it enters the spectrometer, which has an opening of 500μm. The slit is placed at the focal point of the telescope.

The angular field of view (θ) for a telescope with aperture diameter (D) and focal length (f) is given by $\theta = \tan^{-1}\left(\frac{D}{2f}\right)$. For this telescope, substituting these values yields $\theta = \tan^{-1}\left(\frac{300}{2\times1200}\right) \approx \tan^{-1}(0.125) \approx 7.13°$, resulting in a 7° light cone.

3.3 Longpass filter

A colored glass longpass filter is placed after the slit to block shorter wavelengths from entering the spectrometer. Shorter wavelengths will result in higher diffraction orders on the spectrometer signal and should therefore be blocked out. The designed spectrometer is tailored to measure in the wavelength range of 420 nm to 840 nm, so a filter with a cut-on wavelength at 420 nm is used.

3.4 Spectrometer

This section describes the components that are in this spectrometer. An elaborate explanation of how these components were chosen is found in chapter 4. The spectrometer consists of three components, namely a collimating lens that collimates the light cone that comes from the telescope, a reflective diffraction grating, and a focusing lens that focuses the diffracted light onto the detector.

3.4.1 Achromatic doublet lens

A singlet lens has a refractive index that is dependent on the wavelength of the light. Therefore, different wavelengths will have different focal points when transmitted through a singlet. An achromatic doublet circumvents this problem. Here, two elements that are made of different types of glasses with different refractive indices are combined into a doublet lens that has the same focal length for two different wavelengths [31]. Using an achromatic doublet greatly reduces achromatic aberrations when light of different wavelengths is transmitted through a lens.

3.4.2 Diffraction grating

The light captured by the telescope is separated into its wavelength components using a blazed diffraction grating. The angle of the reflected light is according to equation 6 below.

$$m\lambda = \Lambda(\sin\theta_{\text{in}} + \sin\theta_{\text{m}}) \tag{6}$$

Here, m is the diffraction order, λ is the wavelength of the light, Λ is the separation between the grooves, θ_{in} is the angle of the incoming light and θ_{m} is the angle of the m^{th} diffraction order. A blazed diffraction grating is built to maximize the power into one diffraction order. The grating efficiency is optimal when the incident angle θ_{in} is equal to the blazing angle θ_{blaze}.

The diffraction angle for a given wavelength can be calculated by rewriting equation 6. The incident angle θ_{in} must be equal to the blazing angle θ_{blaze} for optimal grating efficiency in the specified order. The angular spread $\Delta\theta$ of the wavelength range can then be calculated as in equation 7 below.

$$\Delta\theta = \sin^{-1}\left(\frac{m\lambda_{\text{max}}}{\Lambda} - \sin\left(\theta_{\text{in}}\right)\right) - \sin^{-1}\left(\frac{m\lambda_{\text{min}}}{\Lambda} - \sin\left(\theta_{\text{in}}\right)\right) \tag{7}$$

Here, λ_{min} is the minimum wavelength (blue) and λ_{max} is the maximum wavelength (red). The periodic spacing of the grooves is denoted by Λ, the diffraction order is denoted by m and the incident angle is denoted by θ_{in}. The required focal length to focus the light on a detector width a width w can then be calculated as in equation 8 below. In this spectrometer, a blazed grating is used which power is optimized in the first diffraction order, so in the equation $m = 1$. Figure 5 shows a schematic of light being diffracted into its wavelength components.

$$f_{\text{foc}} = \frac{w}{2 \times \tan\left(\frac{\Delta\theta}{2}\right)} \tag{8}$$

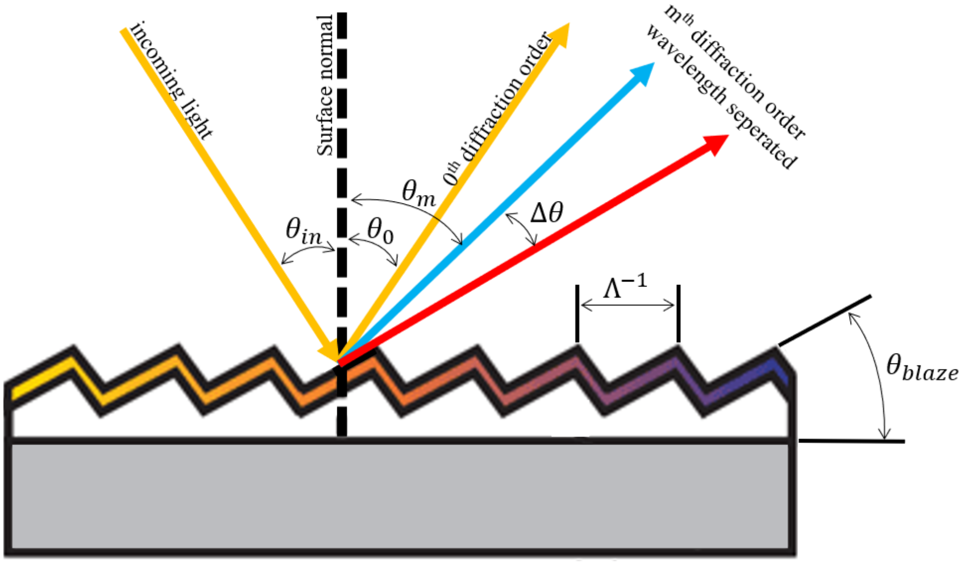


Figure 5: Schematic of light reflected from the blazed reflective diffraction grating. Adjusted version of Figure taken from [32].

3.5 Detector

The separated spectral components are focused onto a Multi Anode Photo Multiplier Tube Array (MA-PMT) (H7260-20, Hamamatsu, Japan), depicted in Figure 6. An MA-PMT operates by converting incoming photons into electrical signals through a series of channels arranged linearly. Each channel amplifies the incoming photons, generating an amplified output signal. The channels are sensitive to low-intensity light, which makes this device suitable for analyzing the low intensity of the backscattered light from insects over a large distance. The MA-PMT used for this experiment can sample at a repetition rate of 200 kHz and is sensitive to light between 300 and 920 nm [33].

The MA-MPT has 32 light-sensitive pixels. The dimension of each pixel is 0.8 mm by 7 mm. The total width of the detection surface is 31.8 mm. The pixels are also referred to as channels later in this thesis.

Figure 6: Picture of MA-PMT

3.6 Signal Processing

Figure 7 below displays a schematic of the different components used for signal processing and how they are connected. The AM-PMT was connected to a printed circuit board (PCB) sensor interface (SIB032D, Vertilon, UK) which was subsequently connected to a PHOTONIQ multichannel data acquisition system (IQSP480, Vertilon, UK) using a micro-coaxial cable assembly. The PHOTONIQ was connected to a data acquisition PC using a USB cable.

In previous experiments with this instrument, the built-in high voltage supply from the PHOTONIQ was used, but this voltage proved to be unstable causing disturbances in the measured signal. Therefore, an external high-voltage power supply was used to apply a voltage to the MA-PMT.

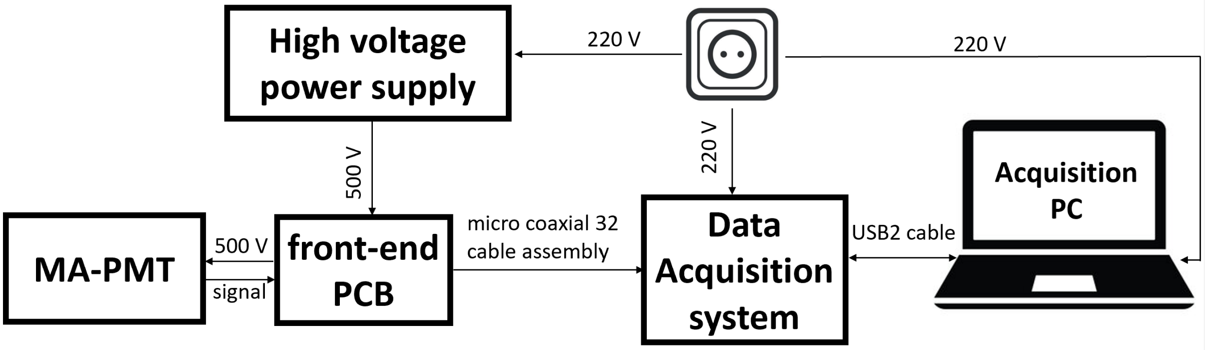


Figure 7: Schematic of components used for signal processing

4 Spectrometer Design

This section describes the design process of the spectrometer, including designs that were not utilized.

4.1 Specification Requirement

The requirements that the spectrometer must adhere to are listed below.

- **F/4 telescope** The largest telescope available to use for this experiment has an F-number of F/4, resulting in a light cone with an angle of approximately 7°. This calculation is shown in section 3.2.
- **Detector Size** The detection surface is 31.8 mm wide and 7 mm high [33]. The spectrum of the captured light should be spread out over the entire detector width.

- **3D printer dimensions** To ensure a stable structure for holding the optical components and to avoid assembly, the spectrometer should be printed as a single piece. Therefore, the size of the spectrometer is limited by the 3D printer dimensions. The 3D printer that is used in this experiment has a printing surface with dimensions 255 mm by 155 mm [34].
- **Timeframe project** The experiment was planned to be done at the beginning of June and the design process started in February. Therefore, components should be ordered from suppliers that have a delivery time within a few weeks. Suppliers that are known to deliver fast are Thorlabs and Edmund Optics. Other suppliers often have a longer lead time.

The performance of the spectrometer is optimized on the factors below.

- **Spectral response** The spectral response is the sensitivity of the detector to different wavelengths. In this project the sun is used as the light source, and its emission spectrum is obtained using Planck's distribution. The efficiency of the MA-PMT is also wavelength-dependent. The emission spectrum of the sun and the efficiency of the MA-PMT are shown in Figure 8. The efficiency of the grating depends on the blazing angle and wavelength. Ideally, the spectrometer should have the same sensitivity for all wavelengths, meaning that the spectral response should be as flat as possible. Therefore, the grating needs to be chosen such that the spectral response is as fast as possible. To improve even further on the spectral response, an equalizer filter can be placed in front of the detector. A cheap alternative to an equalizer filter is to use a colored glass plate. Another option is to use an apodization filter that physically blocks some of the light. This can easily be custom made using a 3D printer.
- **Aberrations** Aberrations are caused by imperfections in the lens which results in the image being distorted. The aberrations that are dominant in this design are spherical and achromatic. The spherical aberrations are caused by lenses being produced spherically, while theoretically the shape should be parabolic. As a result, rays that are transmitted through the edges of the lens are focused in a different spot than the rays that are transmitted through the center. This problem can be solved by using an aspherical lens. Achromatic aberrations are caused by the wavelength dependence of the lens material. This results in a change in focal length for different wavelengths [31].

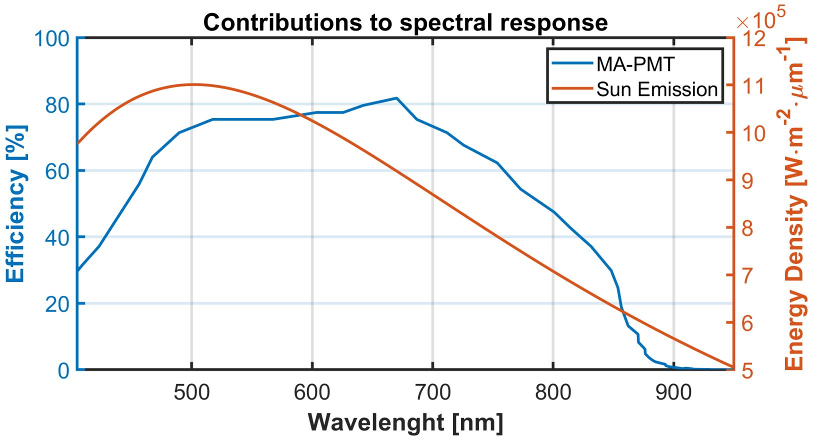


Figure 8: Spectrum of sunlight and wavelength dependence of MA-PMT

- **Resolution** The spot size on the detector must be smaller than the size of an anode, and ideally should be as small as possible to obtain the highest possible resolution.

- **Stray light** When incorporating lenses into a multi-element system, a frequently encountered limitation in overall performance is stray light. Stray light refers to light beyond the clear aperture of an optical system, scattering off the edges of optical or mechanical components, ultimately reaching the sensor as noise rather than a signal.
- **Other diffraction orders** The grating is optimized to concentrate the largest fraction of energy into the first diffraction order. At the center wavelength, approximately 80% of the light intensity is in the first diffraction order, meaning that approximately 20% of the energy is directed into other diffraction orders [35]. For higher diffraction orders it holds that the higher the diffraction order, the lower the intensity in this order. Therefore, only the light in the 0^{th} and 2^{nd} diffraction order is taken into account when designing the spectrometer. Higher diffraction orders are neglected since the intensity of this light is very low [31].

4.2 Choice of components

The spectrometer was designed using an optical raytracing software (Zemax, Ansys, USA). The program was first used to test three different spectrometer layouts to determine which layout was most suitable for this experiment. Once this was decided. The software was used to optimize the design of the chosen configuration and to calculate the position of the lenses, grating, and detector.

4.2.1 Choice of grating

Three different types of gratings were considered for designing this spectrometer, namely a reflection diffraction grating, transmission diffraction grating, and concave reflection diffraction grating. Below is a discussion that compares each grating and discusses their advantages and disadvantages.

reflection grating

Figure 9(a) shows the Zemax raytracing design for a spectrometer with a reflection diffraction grating. The lenses have a focal length of 100 mm and the grating angle is 13.0 °. The diffraction angle of the center wavelength is 37.0 °. The distance from the slit to the collimating lens and the distance from the focusing lens to the detector is equal to the focal length. The distance between the grating and the focusing lens must be sufficiently large such that the lens does not catch the 0^{th} refraction order. Figure 9(b) shows the spot size diagram for this configuration.

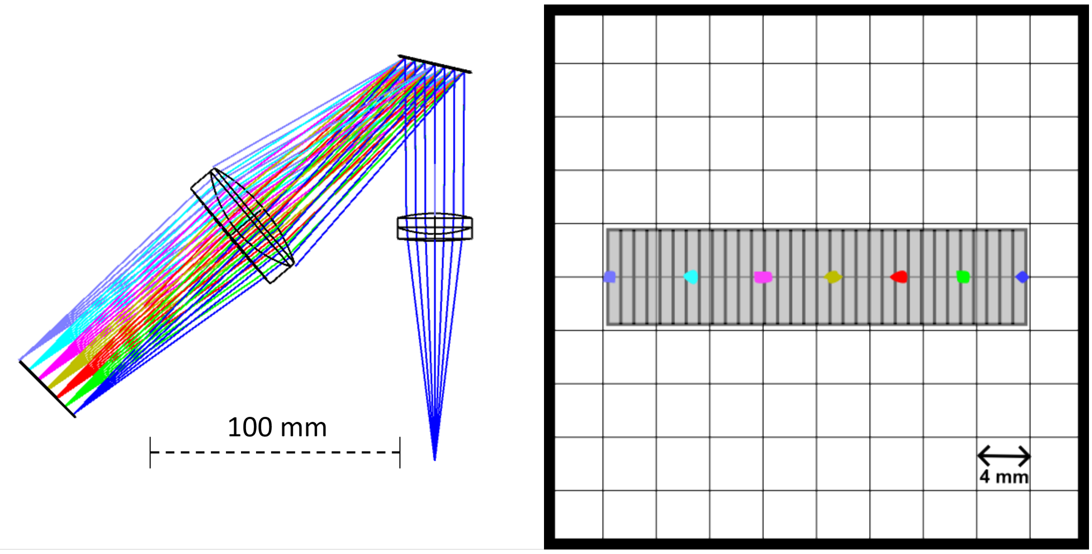
transmission grating
Figure 9(c) below shows the design done in Zemax for a spectrometer with a transmission grating. Figure 9(d) shows the spot size on the detector.

Concave reflection grating
Concave reflection gratings are designed such that the incoming light is dispersed and focused on the detector in one optical component. This means that no collimating lenses and focusing lenses are required, which eliminates spherical and achromatic aberrations [36].

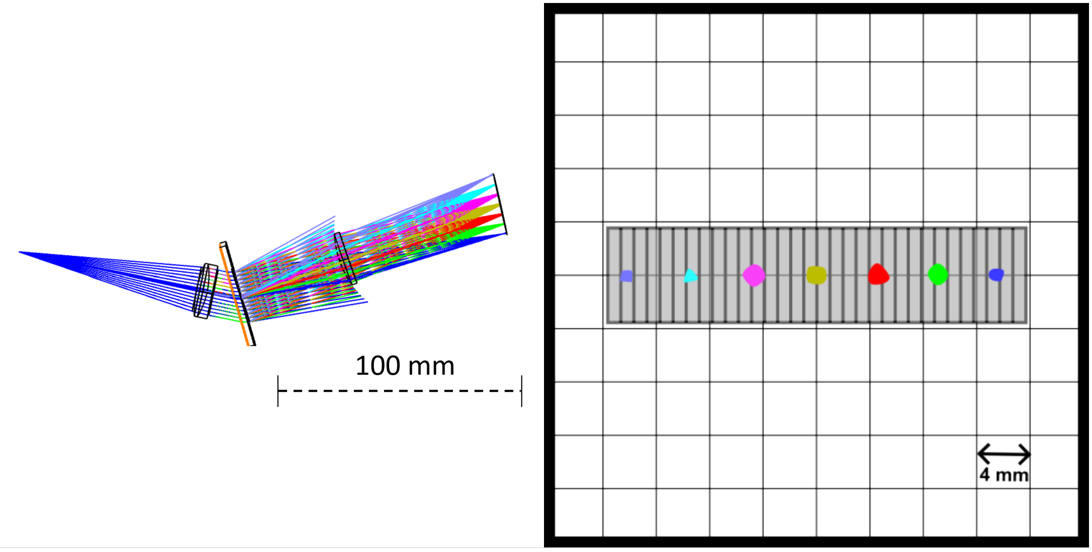
The concave reflection grating has a defined distance to the slit, distance to the detection plane, and incoming and outgoing angles. This means that it is not possible to change the dispersion angle to match the detector size. Therefore, the choice is limited by the commercially available characteristics. Due to the set time frame of this project, it is not feasible to order from any irregular suppliers as these often have a longer lead time. Therefore, the options are limited by the offer from Thorlabs and Edmund Optics. The best option was a concave diffraction grating with 600 grooves/mm and a blaze wavelength of 230 nm (11-541, Edmund Optics, USA). The raytracing results are shown below. It is not possible to calculate a spectral response curve for this grating since the grating sensitivity was not disclosed by the fabricator. Figure 9(e) and (f) show the results of raytracing in Zemax using the concave reflection grating by Edmund Optics.

Comparisson
In Figure 9(e), all components are placed at the specified position described in the datasheet of the grating. Figure 5(f) shows that the spot size is too large for the detector. Additionally, using a grating with a blaze wavelength that is far off will result in both a low grating efficiency and the spots being placed far off-center. Edmund Optics only offers concave gratings with a blaze wavelength in the UV region. Therefore, this configuration will not be used for this experiment [31].

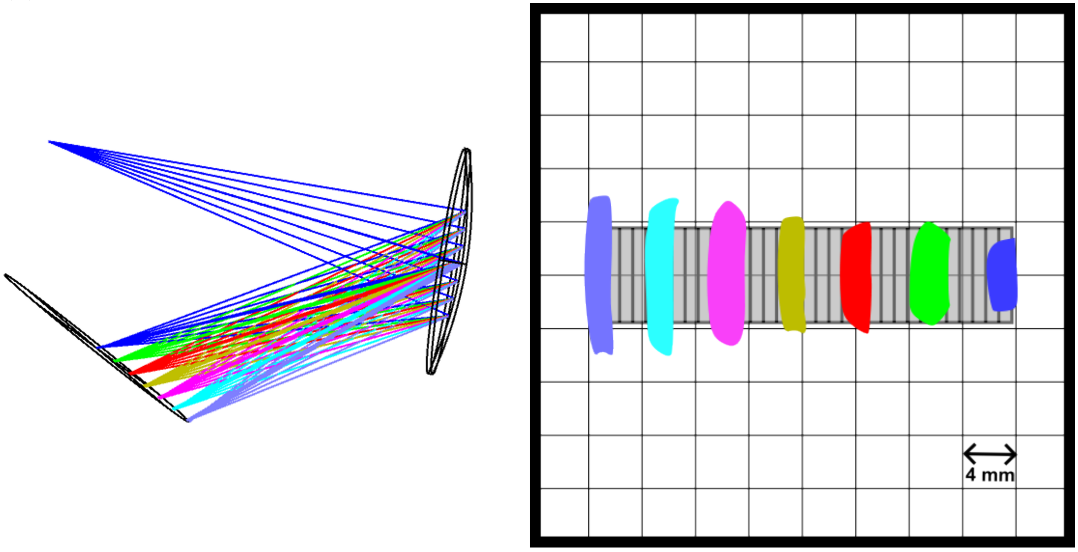
Two relevant differences between a spectrometer with a reflection grating and a transmission grating are the different spectral responses and the size of the final instrument when using a different configuration.
A comparison between the total spectral sensitivity when using different types of reflection and transmission gratings in the instrument is shown in Figure 10 below. The curves in this Figure are a multiplication of the sun spectrum, detector efficiency, and grating efficiency. The first two are depicted in Figure 8. This Figure shows that the reflection grating with 600 grooves per millimeter and a blaze angle equal to 13.0 ° has the flattest spectral response. Therefore, the reflection configuration is the more favorable choice.
The instrument will be more compact when using a reflection grating. This makes it more likely that it will fit in the printing surface of the 3D printer. The transmission grating will require a longer instrument which likely would not fit on the printing surface resulting in a less robust instrument.



(a) Ray diagram reflection grating (b) Spot diagram reflection grating



(c) Ray diagram transmission grating (d) Spot diagram transmission grating



(e) Ray diagram concave grating (f) Spot diagram concave grating

Figure 9: Ray tracing diagram and obtained spot sizes of three different spectrometer configurations. The wavelengths in the simulations are seven values linearly interpolated between 420 nm and 840 nm. The detector is depicted in the figure.

Considering both arguments, the configuration with a reflection diffraction grating was chosen for the design of the instrument.

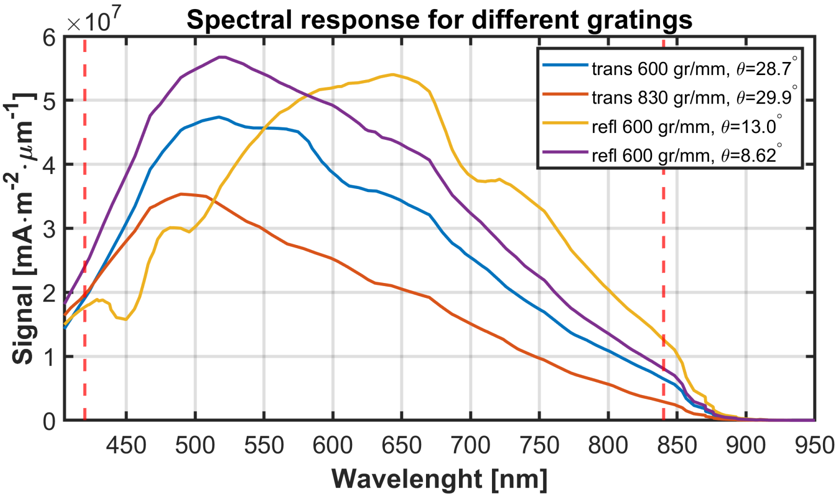


Figure 10: Spectral response of instrument for four different diffraction gratings. Each curve is a multiplication of the two curves in Figure 8 and the grating efficiency for the specified grating. The red lines mark the wavelength range that was chosen for the design of the detector.

4.2.2 Choice of lenses

Figure 10 shows that the reflection grating with $\theta_{\text{Blaze}} = 13.0°$ and has 600 grooves per millimeter has the flattest spectral response, so this grating will be used in the spectrometer (55-258, Edmund Optics, USA). The wavelength range that the spectrometer will analyze was also determined using Figure 10. The wavelength range is chosen so that the signal strength at the minimum wavelength and maximum wavelength are equal. This will result in the most flat spectral response. For this curve, this corresponds to a spectral range from 420 nm to 840 nm. These boundaries are depicted with red dashed lines in Figure 10. The width of the detector surface is 31.8 mm. Using equation 8 the required focal length of the collimating lens is calculated to be 98.65 mm. The spectrometer will not magnify the image, so the focal length of the focusing lens will also be 98.65 mm.

The required focal length is conveniently close to 100 mm, which is a commonly used value for off-the-shelf lenses. For both lenses, an achromatic doublet lens was used to minimize the achromatic aberrations. This resulted in the choice of a VIS-NIR coated (400 nm to 1000 nm) doublet. The edges of the lenses are blackened, as a way to mitigate stray light.

The collimating lens diameter, denoted as D, must exceed the diameter of the light cone at the point of incidence. It can be computed using the focal length of the lens (f_{col}) and the light cone angle (θ_{cone}) as follows: $D = 2 f_{\text{col}} \tan(\theta_{\text{cone}})$. Given $f_{\text{col}} = 100$ mm and $\theta_{\text{cone}} = 7°$ for the F/4 Newtonian telescope, this yields a diameter of 24.6 mm. Therefore, a lens diameter of 30 mm is selected for the collimating lens (49-374-INK, Edmund Optics, USA). The focusing lens should capture as much of the diffracted light in the 1$^{\text{st}}$ diffraction order as possible. Hence, a lens with a diameter of 50 mm is chosen for the focusing lens (49-390-INK, Edmund Optics, USA).

4.3 Raytracing

The system was designed using raytracing in Zemax. This was used for the calculation of the lens positions, grating position, and detector position and angle. Zemax includes a merit function tool that enables the optimization of the system to attain the smallest spot size for each wavelength. The merit function can adjust variables such as lens separations, detector position, and tilt to achieve this goal.

The merit function can include constraints. Two constraints were imposed on the merit function, namely minimum distances between the focusing lens and the grating, and between the collimating lens and the grating. The minimum distances were determined to prevent the lenses from the capturing light from the 0^{th} diffraction order. The grating angle was not included as a variable in the merit function. The grating angle was set to align with the blaze angle of 13.0 °, maximizing resolving power. Additionally, a small vertical tilt was applied to the detector to prevent reflection interference back into the spectrometer.

Utilizing the merit function, optimization was carried out to minimize the spot size, a parameter characterized by the root mean square (RMS) value. The variables in the optimization were the distances between lenses, the grating, and the detector, as well as the chief ray angle of the reflected light and the detector angle. The results of this optimization process are illustrated in Figure 9(a).

To minimize chromatic aberrations, the detector was positioned at an angle. Chromatic aberrations are caused by the fact that lenses have a wavelength-dependent focal length. Tilting the detector counteracts this variation, thereby reducing aberration. The RMS as a function of wavelength is displayed in Figure 11. The spot size needs to be smaller than the anode size to achieve a high resolution. However, achieving a spot size beyond this threshold does not lead to improved performance. The anode width of the MA-PMT is 800 μm, meaning that the spot size is well below the anode size.

The largest contribution to aberrations in this system is spherical aberrations. These aberrations stem from the lens's spherical shape deviating from the ideal parabolic form, causing image distortion. The contribution from aspherical aberrations is larger than from the achromatic aberrations since they are compensated for using the achromatic doublet. Using aspherical lenses is a known solution for this problem. However, an aspherical lens with the right focal length and diameter was not available as an off-the-shelf product. Consequently, the system may still exhibit some remaining spherical aberrations. However, since the RMS spot size is well below the anode size, the aspherical aberrations do not impact the performance of the device.

Figure 13 displays how the 0^{th} , 1^{st} and 2^{nd} diffraction order propagate through the instrument. Light from the 0^{th} diffraction order could be avoided by adjusting the distance of the focusing lens. Mitigating the 2^{nd} diffraction order poses a challenge. Placing the lens significantly further away to avoid the light would result in an impractically large instrument. Moreover, a larger focusing lens would be required which would introduce more achromatic and aspherical aberrations. Fortunately, the power in the second diffraction order is small, minimizing its impact on the system's performance.

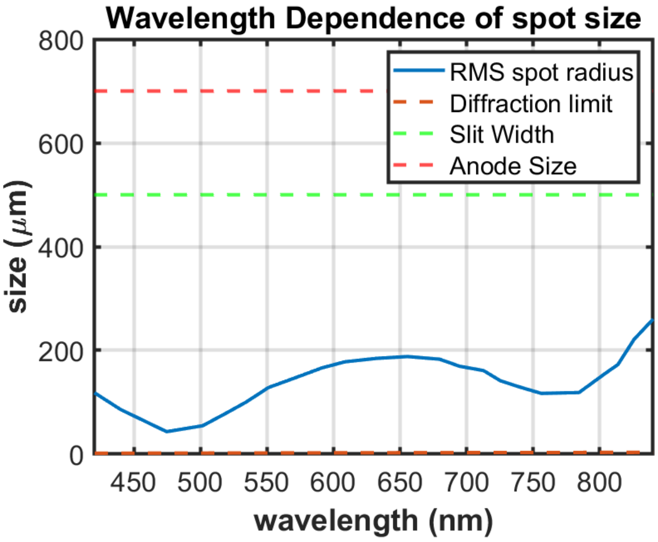


Figure 11: The RMS as a function of the wavelength. The anode size is 800 μm, meaning that the RMS spot size is well below the anode size.

4.4 CAD design

The raytraced optical system described in section 4.3 was imported into a Computer-aided Design (CAD) software (Fusion 360, Autodesk, USA). An opto-mechanical system was designed around the optical components. The structure is constructed with a 3D-printed sandwich design, complemented by elements from the Thorlabs cage system [21, 37]. Namely four rods (ERx, $\phi = 6$ mm, Thorlabs, USA), a cage plate with removable filter holder (CFH2R/M, Thorlabs, USA), and a slit with an adjustable opening (VA100C/M, Thorlabs, USA). The distance between the slit and the collimating lens can be adjusted by sliding the slit over the rods. This can be used to optimize the focus of the instrument. Figure 12 shows a depiction of the spectrometer.

Disturbances in the spectrometer include stray light and light from the 0^{th} and 2^{nd} diffraction order. The spectrometer is printed in black plastic to maximize the amount of stray light that is absorbed. However, a small fraction of the light will be reflected. An effective approach often employed to mitigate stray light is to include baffles that capture and absorb the stray light. The baffles to combat stray light are placed between the slit and the collimating lens. The baffle diameter is slightly larger than the diameter of the light cone at that position so that none of the intensity of the signal is absorbed. The light from the 0^{th} and 2^{nd} is absorbed by baffles with an angle of 30°. This angle will make the light reflect several times until the intensity is reduced to nearly zero. Figure 13 displays how the diffraction orders propagate through the spectrometer and how the baffles are placed to absorb this light.

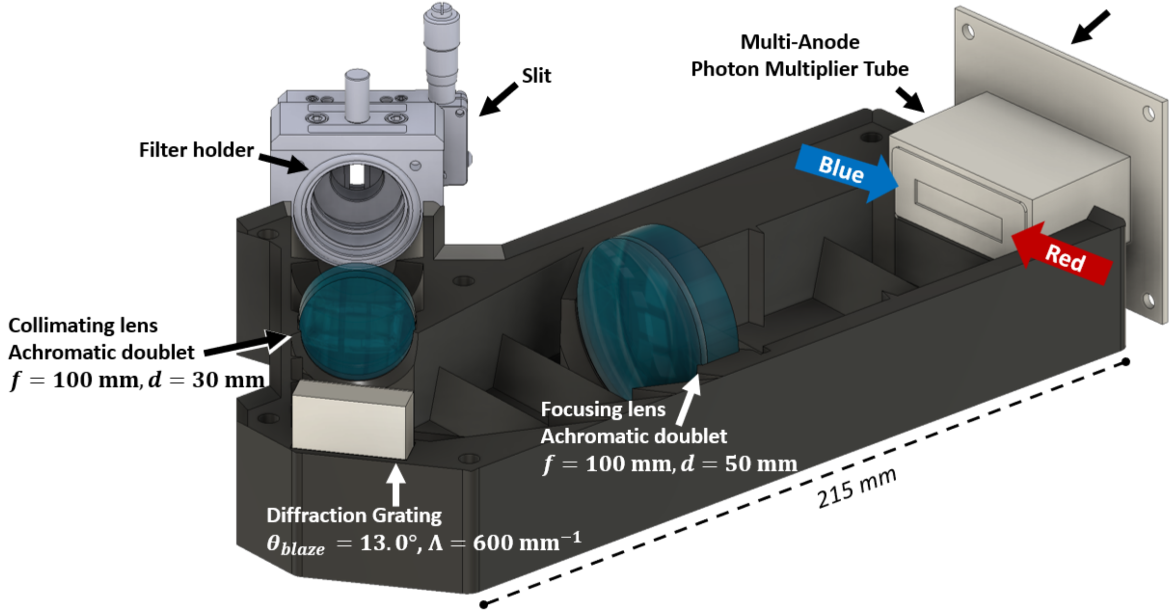


Figure 12: Depiction of spectrometer in CAD software with all components annotated

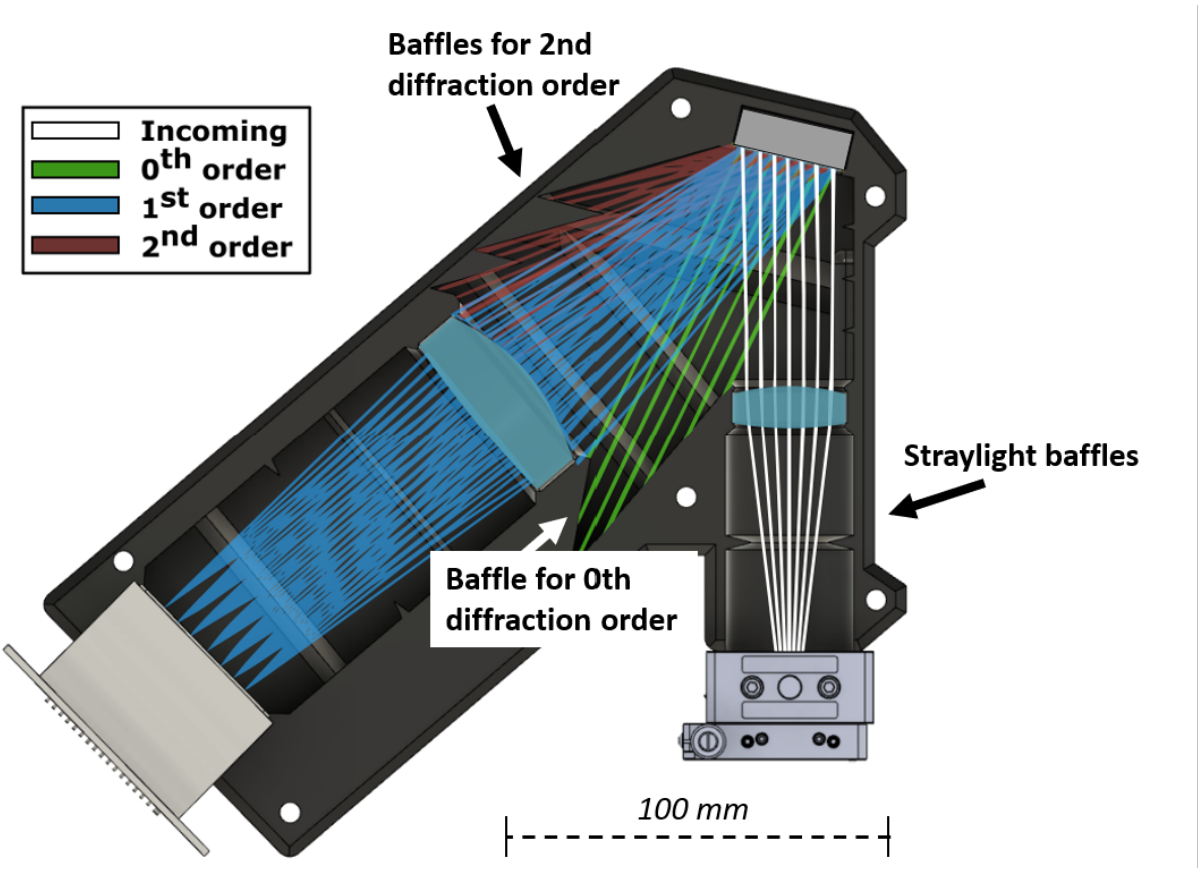


Figure 13: Depiction of spectrometer from CAD software with different diffraction orders and baffles annotated.

4.5 3D printing

The completed sandwich structure was fabricated utilizing a commercially available 3D printer (3D45, Dremel, USA) using black Polylactic acid (PLA-DF=02, Dremel, USA) with a 20% fill factor and a layer height of 200 μm. Black plastic was used since this absorbs stray light. The printed structure is illustrated in Figure 14 below.

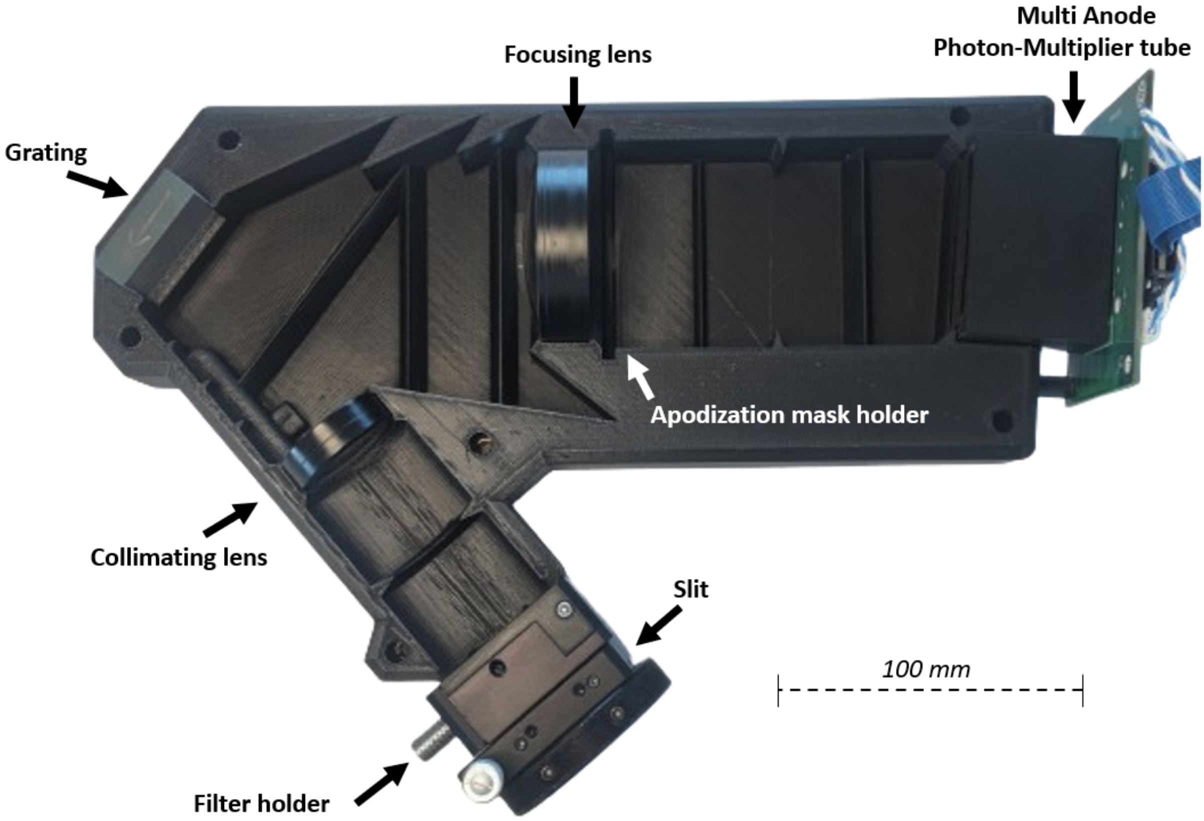


Figure 14: Photo of the 3D printed structure with all components inserted.

5 Instrument Calibration and Characterization

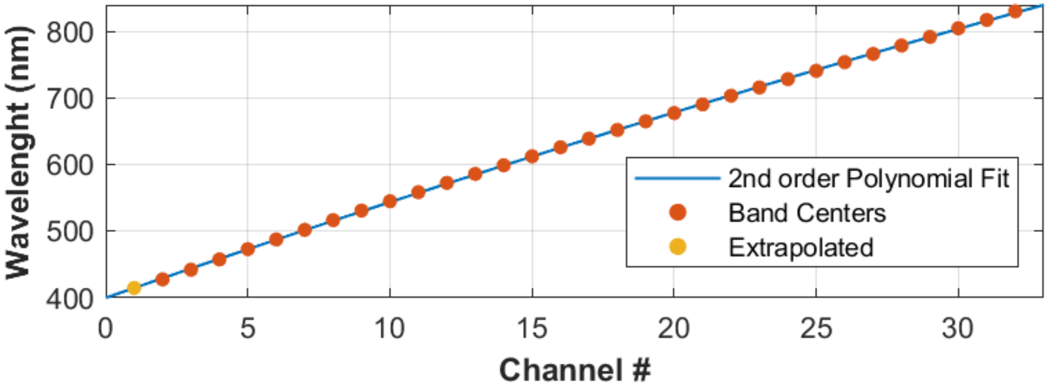
5.1 Spectral Calibration

This section describes the calibration of the instrument. The chapters are listed in the chronological order of when they were executed.

5.1.1 Monochromator

The spectrometer was calibrated using a monochromator. The spectrometer was set up to read out the light emitted from the monochromator. The monochromator was used to scan the channels of the PMT and find the edge of each channel. The edge of channel is the point where the readout of two adjacent channels is equal. The center of each band is then determined by linear interpolation to find the center value between the two edges. During the calibration, the longpass filter was not placed in the spectrometer. Therefore, it was not possible to determine the edge of the first and second channel since the harmonic was visible on the last channel, making it difficult to determine where exactly the edge was. The value of the first channel was therefore determined by extrapolation from the center wavelengths of the other channels. The channel number and associated wavelength are shown in the upper graph in Figure 15 below, together with the constructed fit. The interpolated value for the first channel is shown in yellow. The lower Figure shows the residual of the fit, which shows that the error margin of the calibration is within ±2 nm, which is an acceptable range for this experiment.

(a) Fit plot



(b) Residual plot

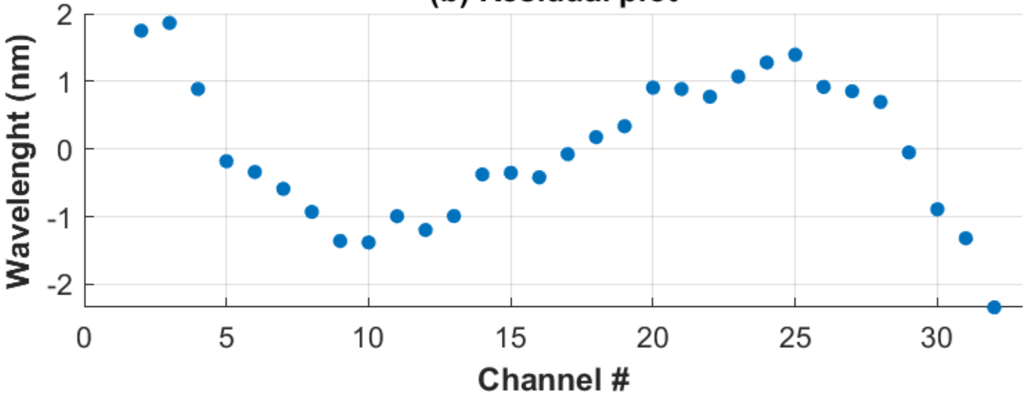


Figure 15: Calibration of the wavelength of each channel on the PMT and the residual.

5.1.2 Emission spectra

To confirm the quality of the calibration, the emission spectrum of several different spectral lamps was measured using the spectrometer. An integrating sphere was used to diffuse the light. Figure 16 below shows the measurement of the cadmium spectrum. In this figure, the known peaks of the emission spectrum of cadmium are clearly visible at the right wavelengths. This measurement and the measurement of several other emission spectra confirmed that the spectral calibration in section 5.1.1 was successful.

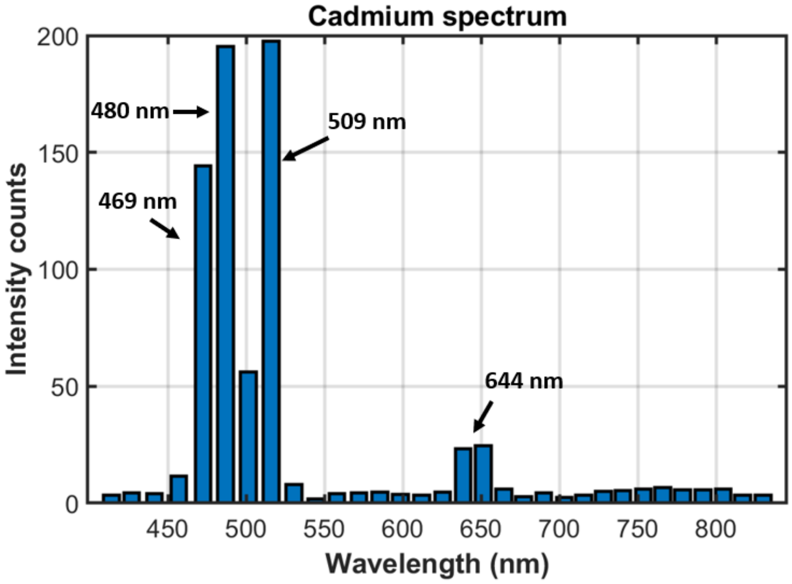


Figure 16: The emission spectrum of a cadmium lamp measured using the spectrometer and the calibration curve obtained with the measured with the monochromator. The cadmium peaks are retrieved from NIST [38]. The ticks on the horizontal axis correspond to the wavelength at the edge of each channel. Therefore, there is no tick on the first boundary.

5.1.3 Electron leakage

A photomultiplier tube can experience electron leakage due to imperfections or defects in the insulating materials surrounding the anodes, leading to unintended electron currents. This leakage reduces the precision of the device by introducing additional cross-talk. It is possible to compensate for the electron leakage in the processing of the data by multiplying it with an Electron Leakage Matrix (ELM) like so: $I_{\text{corr}} \cdot \text{ELM} = I_{\text{raw}}$, where I_{corr} is the corrected data, and I_{raw} is the raw data from the measurement. The data sheet of the instrument provided an ELM that can be used to compensate for this leakage. Additionally, the electron leakage was measured using the monochromator and this data was used to set up another, experimentally obtained, ELM.

Figure 17 below shows both electron leakage matrices applied to the measured emission spectrum of cadmium. The experimentally obtained ELM overcompensates for the leakage since the curve contains negative values and the edges of the peaks are lower than the background level. Therefore the ELM from literature is used when processing the data.

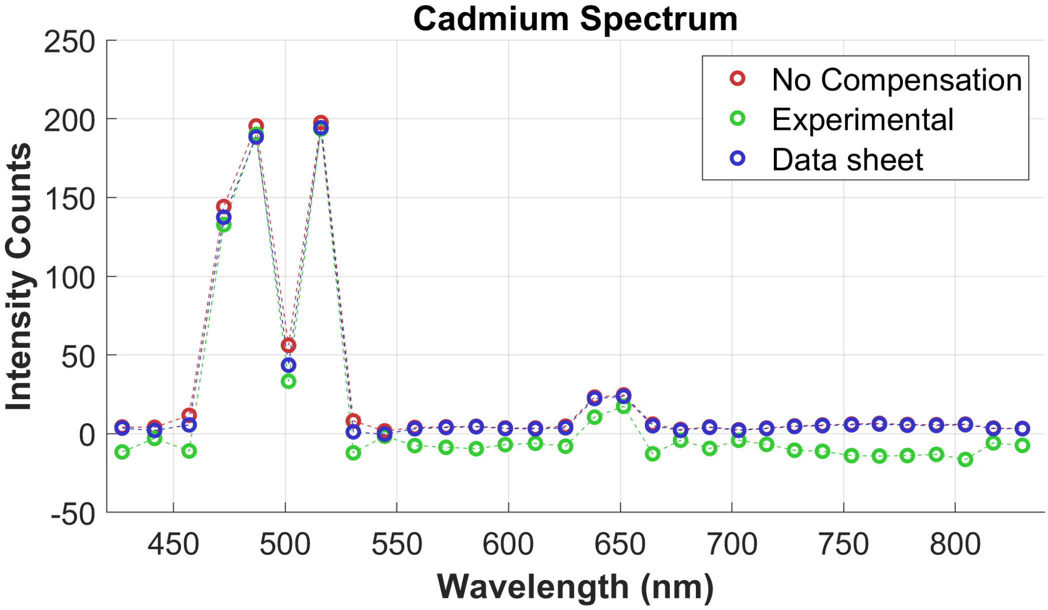
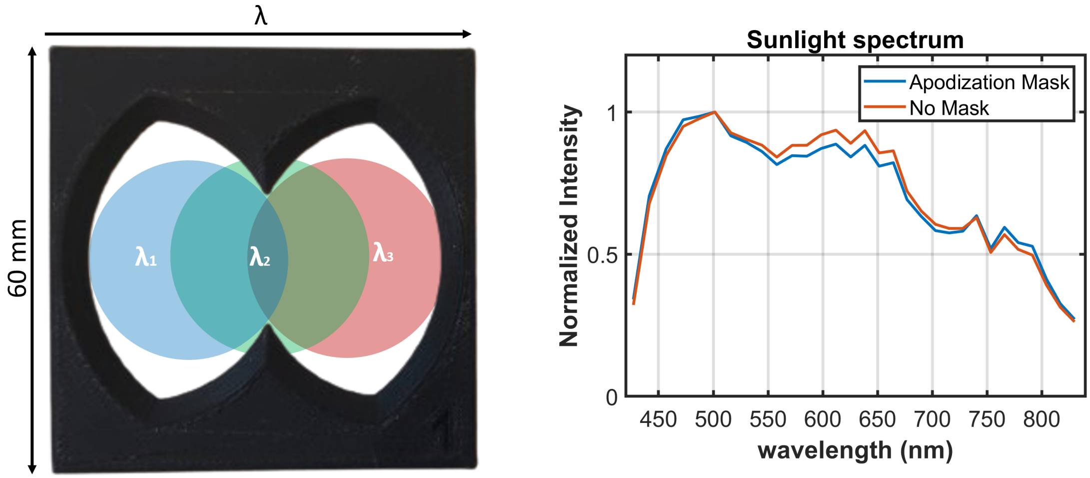


Figure 17: Comparison between two different methods of compensating for the electron leakage in the PMT.

5.2 Apodization mask

An apodization mask was placed behind the focusing lens as a way to flatten the spectral response further. The position is marked in Figure 14. The apodization mask was produced through 3D printing. The theoretical spectral response is displayed in 10 in yellow. Behind the focusing lens, the wavelength components are separated. Therefore, placing a mask that blocks the light at the center of the beam will reduce the intensity of the wavelengths at the center. This mask is shown in Figure 18(a). The mask was tested by measuring the spectrum of sunlight with and without the apodization mask. The result of this is displayed in Figure 18(b). Here, the apodization mask does reduce the intensity of the center wavelength, but it appears that there is a peak at a shorter wavelength. The curve without an apodization mask seems to me more flat, so therefore it was decided to conduct the experiment without the apodization mask.



(a) Depiction of apodization mask.

(a) Measured spectrum sun

Figure 18: The apodization mask is displayed in Figure (a), where the propagation of three different wavelengths is displayed. Figure (b) shows the spectrum of sunlight measured with and without the apodization mask.

5.3 Temporal and Angular calibration

A glass slide was attached to a drill and rotated in the field of view (FoV). Figure 19(a) shows how the experiment was executed, and Figure 19(b) shows a picture of the glass slide in its 3D printed holder. The glass slide will create a signal similar to the flash of a specular insect wing. This measurement was used to calculate the angular size of the sun to confirm signal quality. Figure 20 shows the measured signal. The small peaks in this graph are reflections from the 3D-printed holder of the Plexiglas slide. The foot that is visible at the bottom of the peaks is likely due to reflections from light scattered by the vegetation. The distance between two peaks is equivalent to half a rotation of the slide, so 180 degrees. Using this conversion, the full-width half max (FWHM) of each peak is calculated. The average value of the FWHM of the five largest peaks is 0.667°. The angular size of the sun is 0.5 degrees [23], meaning that the result is within an acceptable range of the angular size of the sun.

(a) Executing the experiment

(b) Depiction of glass slide in holder

Figure 19: Experimental setup for calculating the angular size of the sun with a drill

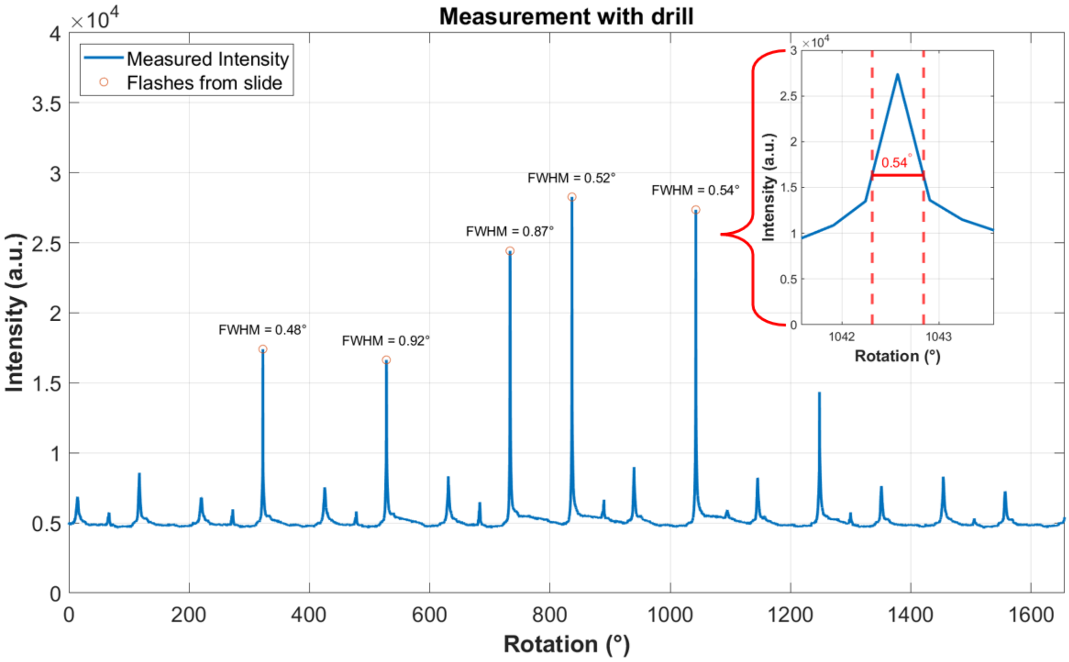


Figure 20: Recorded intensity when rotating the glass slide in the FoV.

5.4 Sunlight Calibration

This experiment uses sunlight as the light source. To compensate for the changes in sunlight throughout the day, the intensity and spectrum of the sunlight were measured. This was done by moving a white stick throughout the FOV every two hours, resulting in five data points each day. Figure 21 shows an example of such a measurement. Figure 21(a) shows the broadband temporal signal and Figure 21(b) shows the spectral signal. This data was used to create a model of the change in sunlight intensity throughout the day. This measurement does not compensate for temporal changes in the sunlight intensity or spectrum, such as a cloud covering the sun. This remains a source of error in this experiment.

Figure 22(a) displays the measured intensity and the model that was made based on this data. Figure 22(b) shows the sunlight spectrum. The measured intensity is the sum of all spectral bands. This value is corrected for the background. The fit is based on a cosine wave that is centered at solar noon (13:05) [39]. The sunlight spectrum is the mean of the solar spectrum measured at solar noon each day since this is the time with the least disturbance from vegetation. The measurement taken at 9:00 was disregarded since it was shaded. The incoming light is then modeled by extracting the intensity at the measurement time from the fit and then scaling each channel with the spectrum of sunlight.

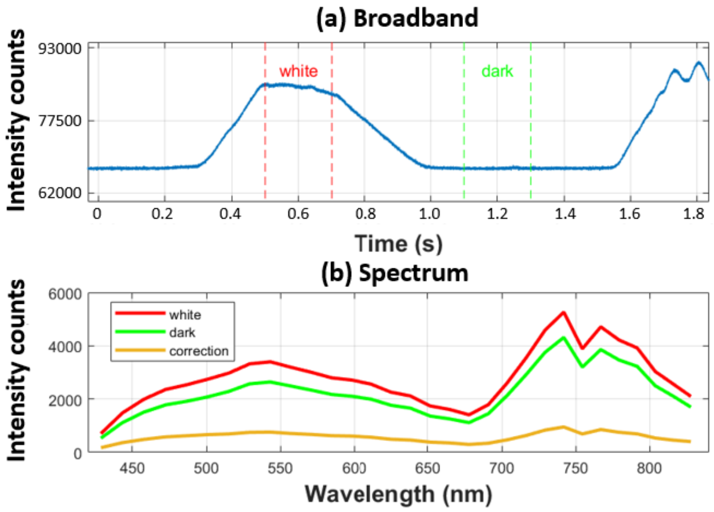


Figure 21: Example of temporal and spectral measurement from white stick

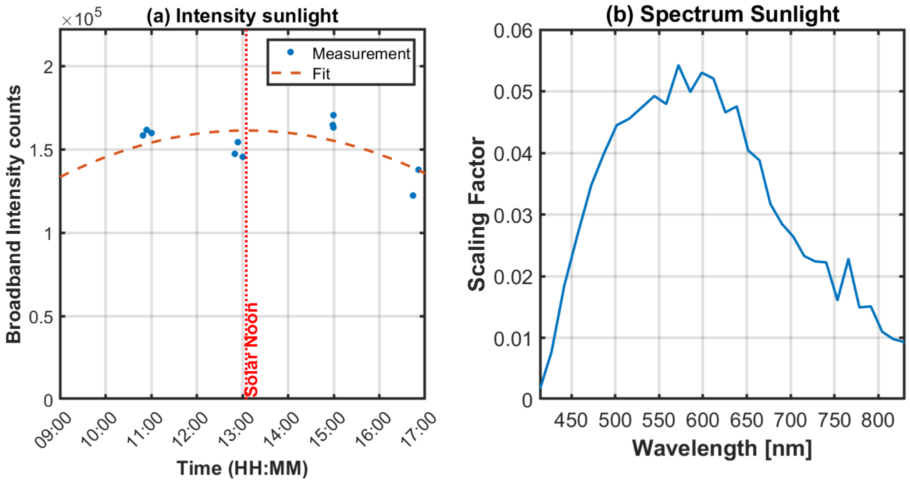


Figure 22: Model for compensating for the change in sunlight throughout the experiment. The left Figure shows the measured intensity values and a model

6 Field Deployment

6.1 Stensoffa site

Stensoffa field station is located in Skåne, southern Sweden(55°41’44”N 13°26’50”E). The experiment was performed on June 6th, 7th and 8th. Each day had sunny weather, with little to no cloud coverage and no precipitation. The station experienced sunrise at 04:26 and sunset at 21:45 on those days, with solar noon occurring at 13:05. The maximum and minimum temperatures recorded were approximately 24°C and 7°C, respectively. The average wind speed was about 2.2 m/s [39, 40]. The altitude and heading of the sun throughout the day are displayed in Figure 23 below.

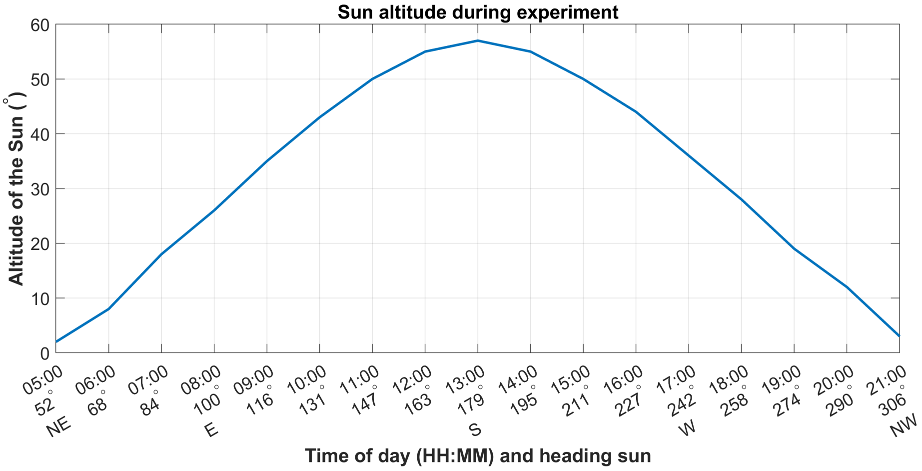


Figure 23: Altidude and spread of the sun during the experiment. The data in this Figure is taken from [39]

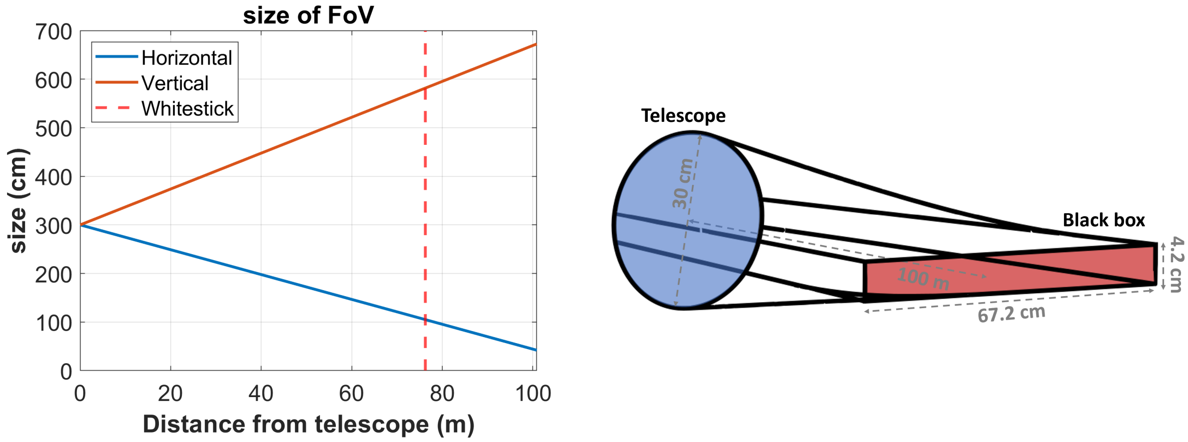
6.2 Experimental setup

The DFS system was oriented northwards with the telescope focused on the entrance of the black box termination, which is 100.86 meters away from the telescope. The experimental setup is depicted from above and from the side in Figure 25 below. The setup was placed in a tent to protect it from any rain or wind. The telescope was mounted on a tripod. During the alignment process, a view finder was attached to the eyepiece of the telescope to make sure that the telescope was looking straight into the black termination box. Following the alignment, the viewer was replaced with the custom-designed spectrometer. The components used for signal processing and how they are connected are displayed in Figure 7. The voltage applied to the MA-PMT was fixed at 500 V during the entire experiment. This value was determined at solar noon when sunlight intensity is maximum. The voltage was chosen such that the strongest flashes are just below the point where it saturates the MA-PMT. The driver was connected to a computer, allowing for settings and control to be configured directly through the computer interface.

To gain insight into the insect species that were present during the measurement period, a malaise trap was placed in close vicinity to the FoV, marked in Figure 26. A malaise trap is an entomological tool used to collect flying insects. It consists of a tent-like structure with a sloping design that guides insects into a collecting container. The malaise trap does not use any bait to trap the insect, giving the least biased catches compared to other trapping methods. However, larger insects often seem to be too intelligent to be trapped by a malaise trap, reducing the number of *Lepidoptera* and *Odonata* in the trap.

6.3 Field of View

The telescope is designed to project the image of a spectrometer's slit onto a termination black box. This slit captures the light from the telescope's FoV. At the telescope the FoV is circular, matching its 30 cm diameter. Over the distance to the termination, this FoV transforms into a rectangular shape. The result observed at the termination is a magnified rectangular version of the slit. The intersection of these circular and rectangular fields of view generates a shape similar to a 'toothpaste tube'. This shape is depicted in Figure 24(b) below. The magnification of the spectrometer is one, so the channels of the MA-PMT are only magnified by the telescope. The magnification of the telescope is $M = \frac{d}{f} = \frac{100.86 \text{ m}}{1.2 \text{ m}} \approx 84$, where d is the distance between the telescope and the black box. Given that the width of the slit is 500 μm, and the height of the pixel is 7.0 mm, the dimension of the FoV at the black box becomes 4.2 cm by 67 cm. The dimensions of the FoV over the distance from the telescope is displayed in Figure 24(a) below.



(a) Size of FOV as a function of distance (b) Depiction of shape FOV

Figure 24: Figure (a) display the size of the FoV in the experiment, and Figure (b) displays the shape. The red line in Figure (a) is the distance at which the white stick calibration measurement was done.

6.4 Daily routine

Measurements were taken while the sun was illuminating the black box, which was between 9:00 and 17:00. The day was split up into four two hour segments. Between each measurement, a white stick was moved through the FoV close to the termination, at a distance of 76 m from the telescope. This resulted in 5 data points each day. The position is marked in Figure 26(a) and indicated with a red line in Figure 24(a). This data will be used to reconstruct how the sunlight intensity changed throughout the day. The malaise trap was set up on the morning of the 7th of June and was emptied twice on the last two days of the experiment. The specimens were collected in a freezer and later counted and identified at family level.

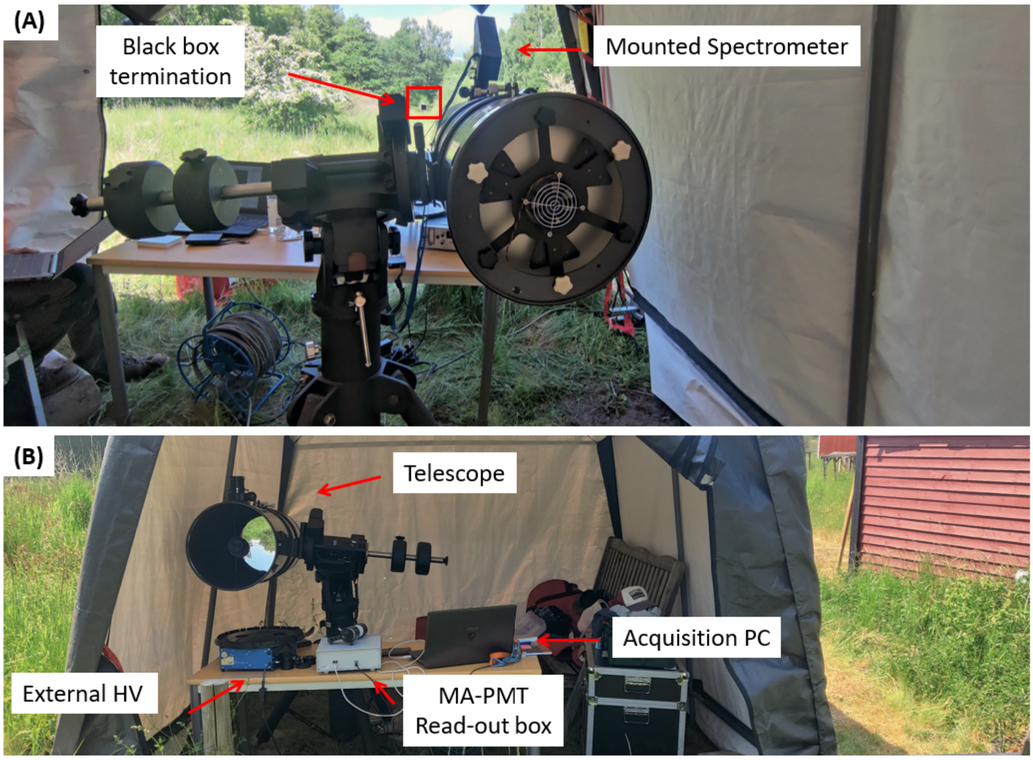


Figure 25: Experimental setup in the tent at the Stensoffa site. The Figure marks all the components mentioned in the text above.

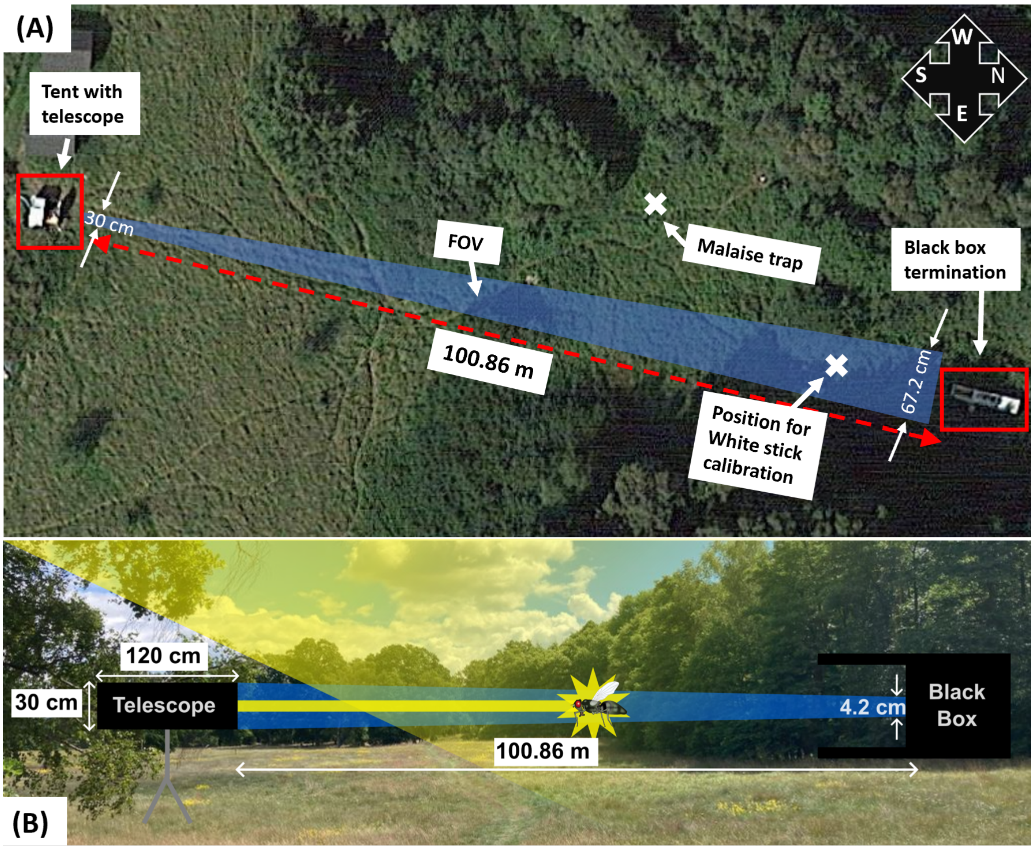


Figure 26: Overview of experimental setup in the field. Figure (A) shows the experiment from above, Figure (B) shows the experimental setup from the side. These figures are not to scale, all the important distances are marked.

7 Data Analysis

The data was analyzed using a programming language (Matlab, MathWorks, USA). The sections below will discuss the different steps that were taken to analyze the data.

In this thesis, the intensity refers to the sum of intensity counts in all spectral bands. The temporal signal is based on the intensity signal and can either be represented in the time domain or frequency domain. These two are directly related by a Fourier transformation, which is calculated using the `Pwelch()` command in Matlab. The spectral signal is the reflectance as a function of wavelength.

7.1 Data Format

The data was assembled at a 20 kHz rate with 32 channels. One capture contains 39 channels in total. Each measurement was approximately taken for 2 hours, and four measurements were taken each day for three days, resulting in 12 files. The approximate size of each file is 10.46 GB, as calculated in equation 9 below. This means that in total approximately 125.5 GB of data was collected during the field measurements.

$$\text{File size (GB)} = \frac{39 \times 20000 \text{ s}^{-1} \times (2 \times 3600 \text{ s}) \times 2 \text{ bytes}}{1024 \times 1024 \times 1024} \approx 10.46\,\text{GB} \tag{9}$$

7.2 Single Event Detection

The first step in the data analysis was to extract single insect observations out of the data set. A signal from an insect can be distinguished from other signals, such as leaves or seeds flying through the FoV, because of the vibration from the wing beating. These vibrations have harmonic overtones which can be seen in the frequency domain. The insect observations can be extracted by applying a mask to the frequency domain with a threshold.

To visualize the frequency domain of the signal, the signal is cut up into smaller segments of 20 seconds. For each segment, the spectrogram is calculated. A spectrogram is a visual representation of the time and frequency. In Matlab, this is calculated by splitting the signal up in overlapping time windows and applying Welch's method over each window to find the power spectrum [41]. The calculated Power spectrum is then visualized using a heat map. A mask is defined using image dilation for data points that exceed the threshold. The threshold is based on the frequency-dependent median and interquartile range (IQR). If the number of mask elements exceeds the threshold, which is set to five times the IQR, the observation is saved as a 2D array with dimensions 32xN, where N is the number of captures. This is stored in a structure which is saved to an external variable. Figure 27 below shows an example of a spectrogram calculated for a 20-second segment of data with an applied mask.

This algorithm extracted approximately 2,000 insect observations from each 2-hour-long measurement, resulting in a total of 24,745 insect observation. The data size is 4.44 GB.

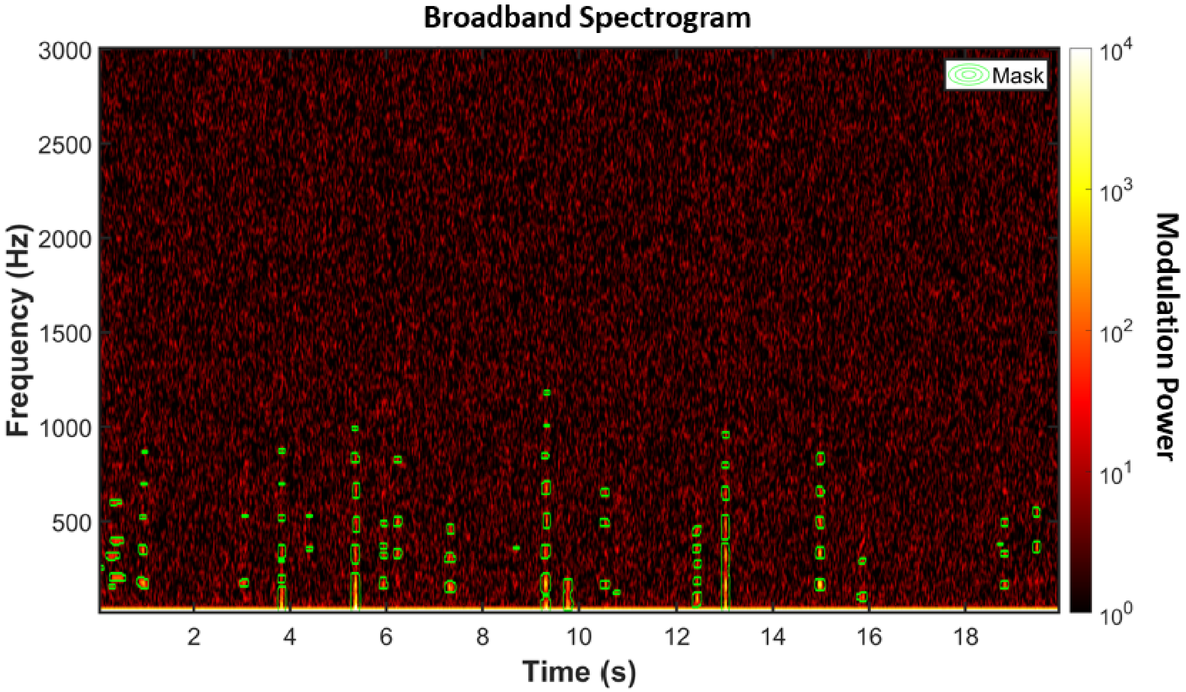


Figure 27: Spectrogram calculated for a 20-second segment of data. The green square indicate the mask that is used to extract insect observations

7.3 Event analysis

After extracting the single events, they are analyzed individually. A model is created for both the dynamic and spectral signal to reduce the noise. Firstly, the center of mass and the spread are calculated using the baseline corrected intensity signal. The signal is then trimmed to the boundaries of the spread to only analyze the signal from the insect. Then, the intensity counts are converted to the optical cross section (OCS). Equation 10 below describes the relation between OCS (σ), area (A) and reflectance (R). The number of intensity counts for the white stick can be extracted from the model displayed in Figure 21. The reflectance of the white stick is assumed to be 80%. The width of the white stick is 10 mm, and the height of the FoV at the position where the stick was transmitted of the beam is extracted from Figure 24(a) to be 124 mm. This results in the area of the stick in the FoV being 1,240 mm^2. The ratio between OCS and intensity counts of the white stick is then used to relate the intensity counts from the insect observation to the OCS of the insect observation.

$$\sigma = A \cdot R \tag{10}$$

7.3.1 Wing beat frequency calculation

The Wing Beat Frequency (WBF) is one of the identifying features that can be used for species identification. The WBF is calculated using the intensity signal. An initial guess is established by calculating the mean distance between the peaks. The peaks are identified using the built-in `findpeaks()` command. This value is optimized by applying a parameterization based on harmonics to a range of WBF values centered around the initial guess. The WBF that results in the highest correlation between the parameterized signal and measured signal is then chosen [42, 43]. The parameterization code is shown in appendix B.

7.3.2 Wing thickness calculation

The wing thickness is calculated by fitting the fringe formula to the measured spectrum of the recorded flash. To extract the thickness of the wing from the fitted fringe the incident angle of the sunlight on the wing is required. This is calculated using the altitude angle θ_{altitude} and heading angle θ_{heading} at the time of observation. This can be found in Figure 23. The incident angle of the sunlight on the wing is then calculated by $\theta_{\text{in}} = \arccos\left(\cos\left(\theta_{\text{altitude}}\right) \cdot \cos\left(\pi - \theta_{\text{heading}}\right)\right)/2$.

To suppress noise, a sinusoidal parameterization with 20 harmonics is applied to each spectral band. The reflectance (R) is calculated for each spectral band. This represents the fraction of light that is backscattered by the measured sample. Equation 11 below shows how the reflectance is calculated for each spectral band.

$$R_\lambda = \frac{I_{\text{flash}_\lambda} - I_{\text{body}_\lambda}}{I_{W_\lambda} - I_{D_\lambda}} \tag{11}$$

Here, R_λ is the reflectance as a fraction, I_{flash_λ} the intensity at the peak of the flash, I_{body_λ} the intensity of the body signal, I_{W_λ} the intensity for the measurement of the white stick and I_{D_λ} the dark background intensity. The reflectance is calculated individually for each spectral band.

The signal is compensated for the intensity and spectral shape of the sunlight in the denominator of equation 11. This information is extracted from the model in Figure 22. The code that calculates the reflection is displayed in C.

The wing thickness is calculated by fitting the fringe equation displayed in equation 3 to the reflectance spectrum. From the fit, the thickness of the wing can be extracted. The fitting is done using the curve fitting tool in Matlab. Global fitting is applied by fitting with a range of starting conditions, to avoid local minima. The code that applied this method is displayed in appendix D.

7.4 Statistics

The number of insect species present in the data set can be approximated by grouping similar data points, which is referred to as clustering. Data clustering is an unsupervised method used to organize information into groups based on similarities, allowing for the identification of patterns or trends within a large data set. The hierarchical clustering algorithm is used for this data set.

Hierarchical clustering is used for classifying insect species through the analysis of intensity or spectral signals. The Euclidean distance between two signals is the straight-line length in multi-dimensional space, representing the overall difference between corresponding components of the signals. This is used as a measure of similarity between two signals. In hierarchical clustering, each signal is initially treated as a distinct cluster. The hierarchical clustering algorithm systematically merges clusters with the smallest Euclidean distance, revealing natural associations between insect species based on their intensity patterns [44].

Hierarchical clustering is a preferred method for measurements of free-flying insects since the user is not required to specify the number of clusters beforehand. Each cluster is equivalent to an insect species, and the number of insect species is not known. Another advantage of hierarchical clustering is its deterministic nature, as it is not based on stochastic processes. This means that each time the algorithm is applied to the same data set, it will produce the same hierarchical structure.

Hierarchical clustering can be applied to the dynamic signal and the spectral signal. To determine whether measuring the spectral data leads to an increased species specificity, the data set is clustered twice. First with only the dynamic signal, and then with both the dynamic and spectral signal. This will tell us if there was extra information about the species in the spectral signal.

The data needs to be delivered to the hierarchical clustering algorithm as a set of data points where each data point is a one-dimensional array with a fixed size. Each data point is one insect observation. Therefore, each observation needs to be represented such that it is the same size. This is achieved by representing the dynamic signal in the frequency domain, which is an array with dimensions $(1 \times F)$, where F is the number of frequency bins. When the clustering algorithm is applied to both the dynamic and spectral signal, the Fourier transform is applied to each spectral band. The data is then stored in a one-dimensional array of size $(1 \times 32 \cdot F)$. The Euclidean distance between each observation is then calculated as shown in equation 12 below [16]. The power spectra are denoted by $P_{n,f}$ where n represents the n^{th} observation and f represents the frequency space in the frequency vector.

$$D = \sqrt{\sum_{f}^{f_{\mathrm{lim}}} (\log P_{a,f} - \log P_{b,f})^2} \tag{12}$$

Determining the number of clusters present in the data set poses a challenge. In this thesis, the number of clusters is determined by defining a turning point in the linkage, which is the Euclidean distance between the clusters. A linear model is fitted to the values and a range within plus or minus the IQR of the distance between each data and the linear model is defined. The turning point in the data is the point where the steady decrease enters the specified range from this fit. Figure 28 illustrates this method.

7.5 Singular Value Decomposition

Singular Value Decomposition (SVD) is applied to the raw data to determine how many spectral components each observation displays during interception with the probe volume. SVD breaks down the data into three parts: U, S, and V matrices. The S-matrix contains singular values. Higher singular values carry more information whereas singular values of noise decay exponentially. For the singular values, the defined turning point indicates the number of spectral channels where additional information becomes negligible, so the turning point indicates the number of spectral components that were present in the measurement. The turning point is determined using the same method for finding the turning point in the linkage, described in paragraph 7.4 above. The code for the singular value decomposition is presented in appendix E.

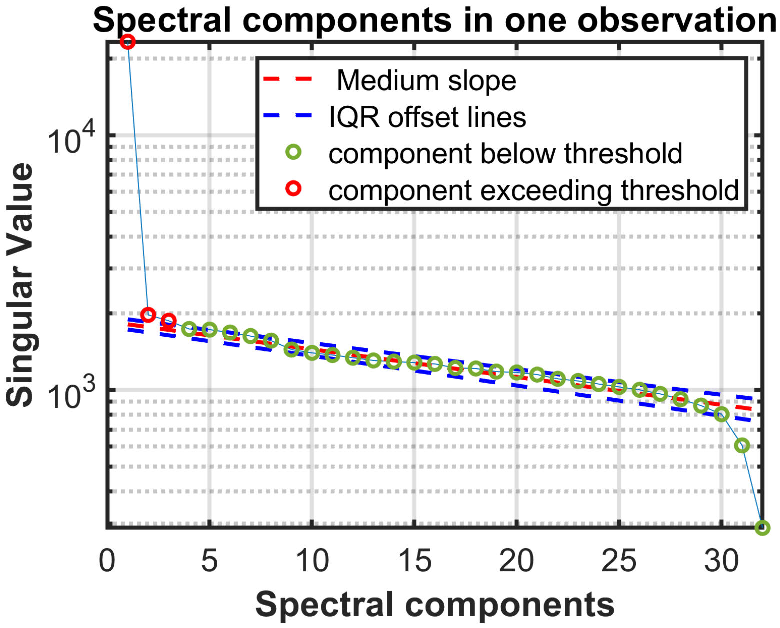


Figure 28: Example of defining the turning point in the singular values. In this example, there are three singular values above the fitted line.

The number of spectral components represents the number of distinct spectra during an observation. For an insect, this is at least two, which can be attributed to the wing and the body. More spectral components can be caused by a clear wing that exhibits a fringe, or an insect that has a different color on the top and bottom of the wing. The number of spectral components represents the number of spectral bands which could be useful to identify the species. This knowledge is useful for designing new multiband instruments for insect identification [5]. It could also relate to why insects have a additional spectral bands [22].

8 Results and Discussion

8.1 Specular insect observations

Figures 29, 30, and 31 below display three observations that were recorded during the experiment. In all figures, (a) displays the recorded broadband dynamic signal. The grey dots show the raw data and the blue line shows the fit that was made with 20 harmonics. Figure (b) displays a single period of the fit made in Figure (a). The fit is not apodized, so it does not include any signal from the body. The position of the flash and the body are calculated using the broadband signal and these times are denoted in the figure. The reflectance at the flash and the body are displayed in Figure (c). The fit for the flash is also shown in Figure (c). Figure (d) shows a surface plot of the frequency domain of each spectral band. This is based on the raw data. Figure (e) shows the broadband frequency domain of the broadband signal. The grey box in this Figure marks the visualization limits that are used for Figure (d).

All three observations are specular observations, since the dynamic signals displayed in (a) and (b) show sharp high peaks resulting in a high skewness value. Consequently, the frequency domain displayed in Figure (e) shows many overtones. All three observations contain a fringe in Figure (c). The correlation between the data and the spectral fit in Figure 29 is 63.19%, 74.69% in Figure 30, and 66.22% in Figure 31. This indicates the presence of a fringe in the data, meaning that the insect had a homogeneous wing thickness.

In Figure 30(c), the reflectance of the body starts to increase linearly at about 700 nm. This corresponds to the spectrum of melanin, suggesting a melanized body. Additionally, the negative values for the Optical Cross Section (OCS) in Figure 30(a) imply that the body is darker than the background. Melanin has a darkening effect, which aligns with the melanized body hypothesis. Figure 29(c) also exhibits a slight increase in the reflectance at wavelengths higher than 700 nm, and the OCS has negative values. This implies that the body is melanized, but less melanized than the body of the insect in observation 30.

Noise from external sources such as aerosols or leaves passing through the FoV do not vibrate and therefore have no overtones in the frequency domain. As a result they are only visible at loew frequencies, below 100 Hz. Any noise at higher frequencies, so above 500 Hz, is caused by defects in the experimental setup. Noise is observed at frequencies higher than 500 Hz at the far red region (> 700 nm), for all observations in figures 29(d), 30(d), and 31(d). The increased noise at the far red spectral region is likely explained by the high reflection of chlorophyll above 700 nm [45]. The baffles in the spectrometer partially absorb stray light, but a considerable amount appears to have reached the detector. This is also seen in the reflectance spectra in Figure (c), where the longer wavelengths are over-represented. This is likely caused by the insect wing reflecting both the sunlight and light that is reflected by the vegetation.

The intensity of light reflected by the vegetation is dependent on two variables: the time of insect recording, since the incident angle of the sunlight changes during the day, and the position at which the insect intersects the FoV, due to the varying vegetation at Stensoffa across the FoV. This experiment does not contain ranging information, so it is not known at which position the insect intersected the FoV. Therefore, it is not possible to accurately compensate for the chlorophyll reflectance with the data obtained in this experiment.

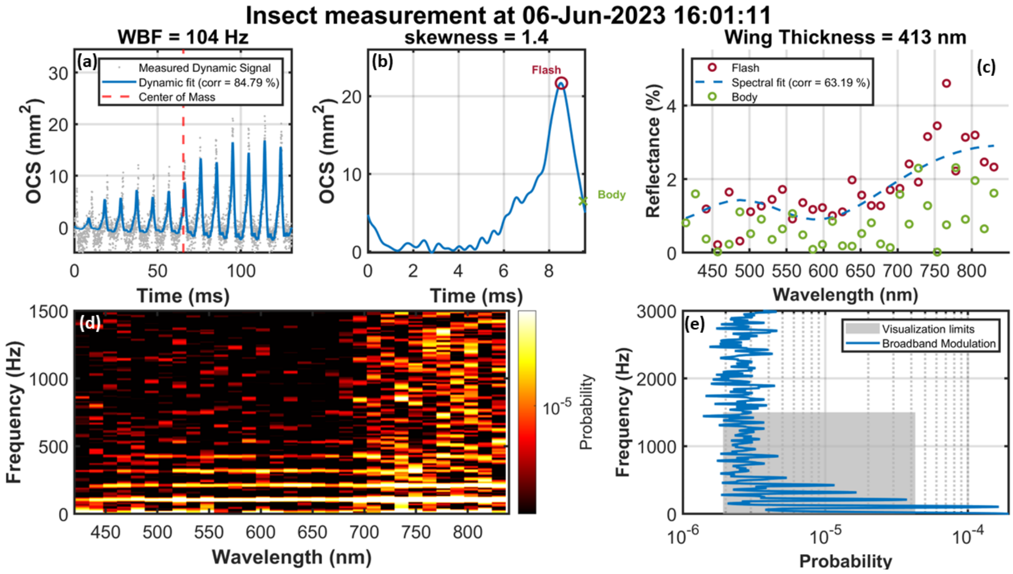


Figure 29: Insect observation 1

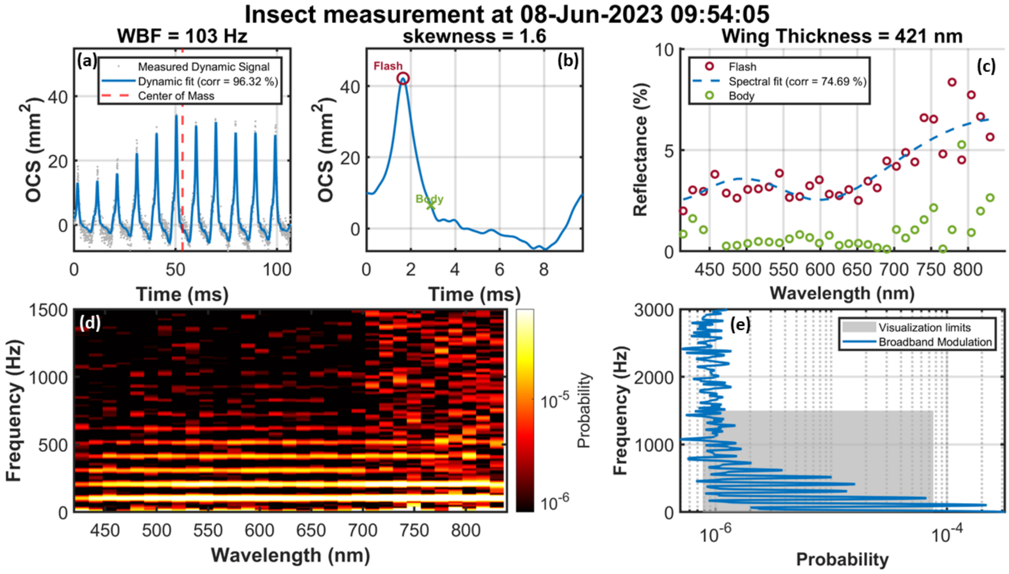


Figure 30: Insect observation 2

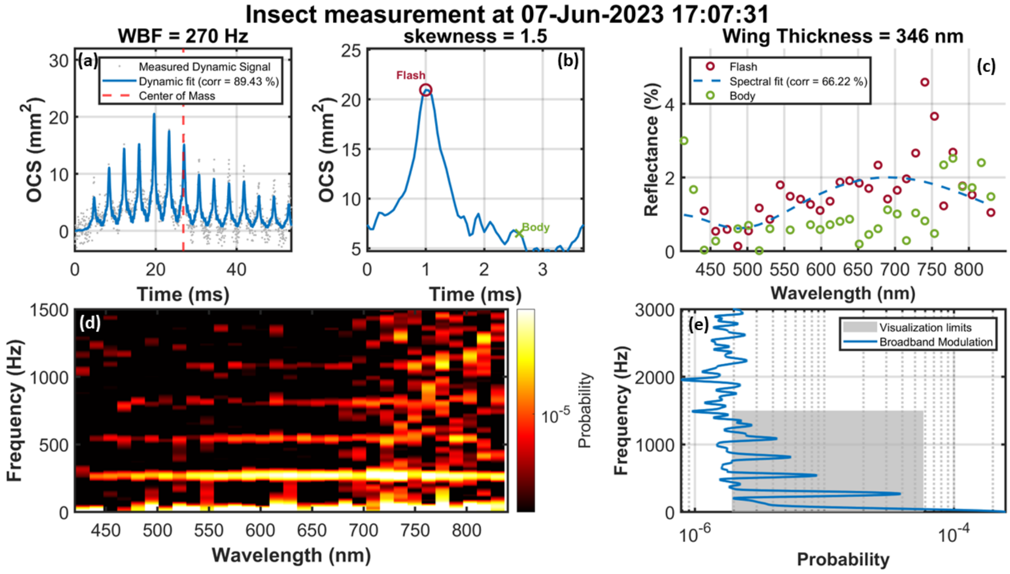


Figure 31: Insect observation 3

8.2 Insect identification

The wing beat frequency and specular modulation from the insect observations can be used to determine the insect order. To use the wing thickness as a parameter for insect identification a hyperspectral analysis of the captured samples would be required, which was not feasible within the time frame of this project.

Figure 33 displays the insects that were captured in the malaise trap during the experiment, meaning that these insect orders could have been observed during the experiment. Larger insects are often misrepresented in the malaise trap, so *Odonta* or *Lepidoptera* should also be considered for observations, despite not being captured.

All three observations have specular signals, indicating that the insect has a clear and homogeneous wing. The orders that contain species which such wings are *Diptera*, *Hymenoptera* and *Odonta* [19]. Figure 32 displays the wing beat frequency for different insect orders. Figure 29, 30 and 31 have a WBF of 104, 103, and 270 Hz, respectively. All three frequencies fall within the range of *Diptera* and *Hymenoptera*, meaning that all three observations were likely of an insect species in one of those two orders. A more accurate conclusion could be drawn by hyperspectral analysis of the captured samples.

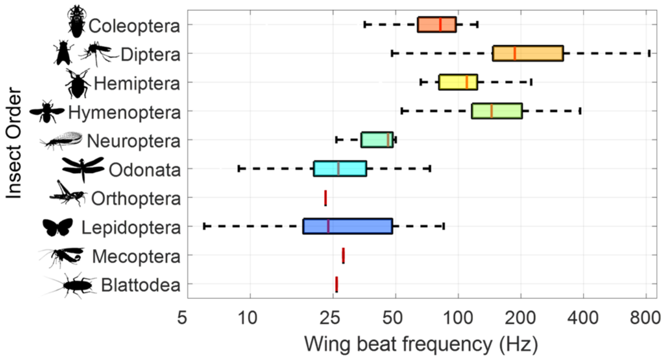


Figure 32: Wing beat frequencies measured for different insect orders. Taken from [46].

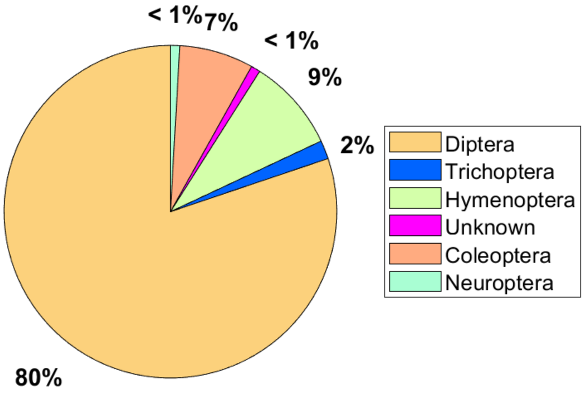


Figure 33: Pie chart displaying the insect order and relative quantity that was found in the collected insect samples captures with the malaise trap.

8.3 Clustering

The linkage for each clustering process is shown in Figure 34 below. The clustering using only the temporal data resulted in 2,098 clusters. Using only the spectral data resulted in 58 clusters. Using both the spectral and temporal data resulted in 3,040 clusters. This indicates that the spectral data enables the differentiation of a greater number of species. Figure 35 and 36 display three temporal and spectral clusters as an example. The grey lines indicate the range of the clustered data.

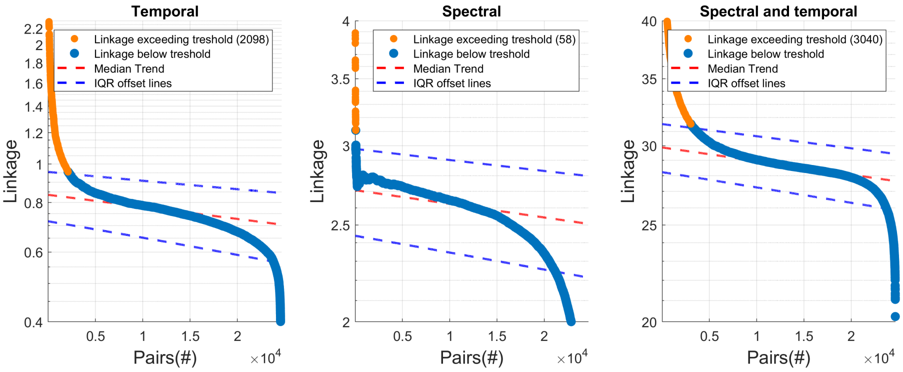


Figure 34: Linkage curves of clustering using only the data from the temporal domain, spectral domain, and combined spectral and temporal domain.

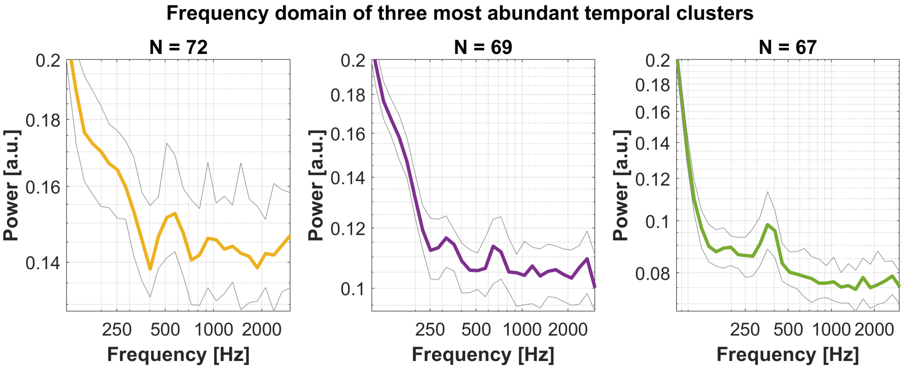


Figure 35: The frequency domain of the three most abundant temporal clusters. The number of clusters in the group is denoted in the Figure title. The grey lines depict the boundaries of the range within this cluster.

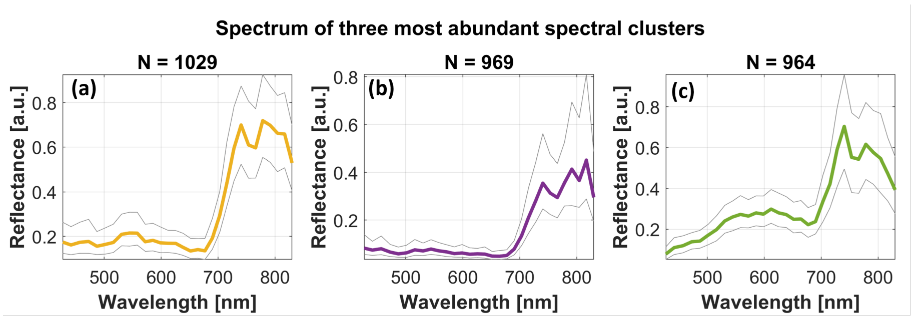


Figure 36: The spectra of the three most abundant clusters in the spectral domain. The number of clusters in the group is denoted in the Figure title. The grey lines depict the boundaries of the range within this cluster.

In the temporal clusters displayed in 35 up to three frequency overtones are visible. The WBF of an insect's wing varies within a 25% range [18]. Consequently, the second overtone displays a 50% fluctuation, the third overtone a 75% fluctuation, and any higher frequencies will have 100% or higher fluctuation in WBF, causing them to diminish. The clusters in Figure 35 follow this theory.

Figure 36(a) contains a fringe, indicating that this cluster belongs to an insect with a clear and homogeneous wing. Figure 36(b) has a flat and low reflectance until 700 nm, where a steep increase is visible. This suggests an insect that is completely melanized. All three clusters have a plateau at 700 nm onwards, stemming from the chlorophyll spectrum from sunlight reflected by the vegetation.
The presence of the chlorophyll spectrum in insect observations can lead to clustering of observations based on the time and location of intersection, as these factors influence the chlorophyll spectrum. Consequently, a single species can lead to multiple clusters. Therefore, eliminating the chlorophyll spectrum will give more dependable clustering results.

8.4 Singular Value Decomposition

Figure 37 below contains a histogram displaying the number of spectral components present in every insect observation. There were 1,800 observations that have zero spectral components, which is not possible. This can be attributed to noise entering the data pipeline due to imperfections in the event detection algorithm. These observations are therefore not displayed in Figure 37. The observations with two spectral components can be attributed to the signal from an insect with a diffuse wing, where one spectral component is from the body and one from the wing. The observations with more spectral components can be attributed to insects with structural colors from clear wings, or insects with different colors on their wing or body.
There are four or less spectral components in 18,112 observations. This corresponds to 79% of all observations that are not regarded as noise. Therefore, an instrument for insect identification in the same circumstances as this experiment needs four spectral bands to classify the insect species for 79% of the observations. In [17], a model for spectral fringes is set up with four degrees of freedom. This data confirms that this is enough to identify most of the insect fringes. Dragonflies have five effective spectral bands [22]. Figure 37 reveals that insects can contain up to nine spectral components. Dragonflies are predators that hunt other insects. This data suggests that the additional spectral bands could aid them to visually identify prey while flying.

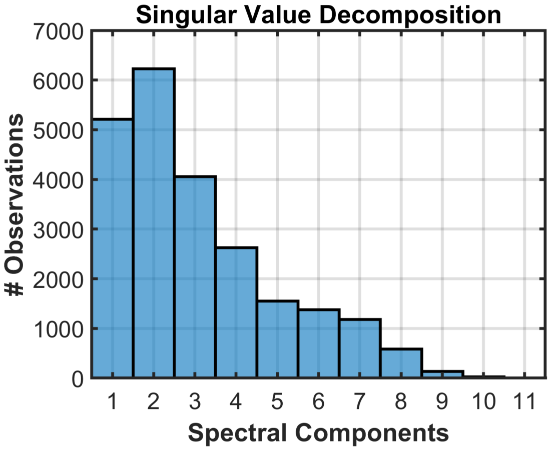


Figure 37: Histogram of the spectral components in each insect observation

9 Conclusion

There are two parts to this project. The first part is the development and deployment of the spectrometer and the second part is the analysis of the data. This section concludes both parts and uses the results to answer the research question. Additionally, some limitations that arose during the experiment are discussed.

The spectrometer design and production were accomplished effectively. Spectra obtained from the spectral lamps affirmed the proper spectral functionality of the instrument, while the measurement of the angular size of the sun confirmed the temporal calibration. The field trials went smoothly and the spectrometer captured dynamic and spectral signals from free-flying insects. During the field trials, a total of approximately 24,000 insect observations were made, resulting in approximately 125.5 GB of data.

During the data analysis, the harmonics within the temporal signals were successfully extracted, and fitting sinusoidal waves yielded high correlations. The temporal signal was used to determine the WBF of the insect observations, and the found values corresponded well with the WBF values for the insect catches. Fringes in specular signals, indicative of insect wing thickness, were fitted using the fringe formula and gave decent correlations to the measured values. To draw precise conclusions about wing thickness, an analysis of captured insect samples is necessary.

More data clusters are found when the spectral data and temporal data are combined. The number of clusters present in the data indicates the number of species that are differentiated in the data. The answer to the research question is that spectrally resolved modulation provides additional information for differentiating species.

An identified limitation in this experiment is the presence of the chlorophyll reflectance spectrum in the data. The unexpectedly strong signal, which is also visible at higher frequencies, suggests that insect wings reflect not only sunlight but also light reflected from the vegetation into the telescope. Removing the signal from the chlorophyll is not possible since this experiment does not contain ranging information. Despite the presence of the chlorophyll spectrum, it was still possible to use the data to answer the research question.

The application of singular value decomposition (SVD) to the spectral modulation data has provided insight into the spectral components that are present in the insect observations. By analyzing the magnitude of the singular values, it becomes evident that only a limited number of spectral components are required to adequately describe the signal. This analysis gives clues about the minimal number of spectral bands necessary to accurately describe the resolved spectra in the observations. Therefore, this study enhances our understanding of visual ecology among insects and provides information on the number of spectral bands required for future multiband instruments.

The sampling frequency of 20 kHz proved sufficient, resolving wing beat frequencies up to 500 Hz. A higher sampling rate would not provide additional information but it would compromise the captured light intensity.

In conclusion, this study successfully captured and analyzed sunlight flashes from free-flying insects, achieving the objectives of measuring wing beat frequency and wing thickness. The collected data provided an answer to the research question and contained clues about visual ecology among insects.

10 Outlook

During this experiment, 125.5 GB of data was collected containing about 24,000 insect observations. Only three of these insect observations were properly analyzed and studied. There is still a lot of information hidden in this data that could be extracted. The data obtained during this experiment is the first data of insect fringes in the visible wavelength range. This means that the data could contain valuable clues about visual signals between insects. Further investigations of these signals could lead to exciting findings in visual ecology. Additionally, clues about insect behavior can be extracted from this data, as done in [15].

Future experiments could include a hyperspectral analysis of insect samples that were collected during the field trial. The fringes recorded in the laboratory can be compared to the fringes recorded in the field to aid identification of the species that were recorded, as done in [21].

The chlorophyll reflection spectrum resulted in disturbance in the spectral domain. The experiment could be redone with a more accurate model to compensate for the chlorophyll reflectance spectrum. As discussed in section 8.1, the chlorophyll reflectance depends on the position where the insect intersected the FoV and the time of the observation. To be able to accurately compensate for the chlorophyll reflectance spectra, the experiment could be redone with ranging to obtain the position where the insect intersects the FoV. It is shown in [14] that it is possible to obtain ranging data from a dark-field spectroscopy experiment. Additionally, more measurements from white stick calibration throughout the day could have given a better insight into how the intensity from the spectrum of chlorophyll reflectance changes throughout the day. Using this data, an accurate model of the sunlight and the light reflected from the vegetation can be made, allowing for accurate compensation of the chlorophyll reflectance spectrum.

The intensity of the chlorophyll reflectance spectrum could also be reduced by avoiding vegetation. This can be done by marking the FOV clearly and removing nearby vegetation to reduce the amount of light reflected from the vegetation that the insect wings reflect into the telescope. However, this is not always an option. A measurement over a grass meadow without any nearby trees or bushes would give a homogeneous chlorophyll reflectance, which would make it possible to compensate for the chlorophyll reflectance spectrum without the need for ranging information. Another way to completely avoid vegetation is to execute the experiment above a lake, like done in [47].

References

[1] C. A. Hallmann *et al.*, "More than 75 percent decline over 27 years in total flying insect biomass in protected areas," *PLoS ONE*, vol. 29, no. 4: Biophotonics, pp. 1–8, 2017. DOI: `https://doi.org/10.1371/journal.pone.0185809`.

[2] S. Seibold *et al.*, "Arthropod decline in grasslands and forests is associated with landscape-level drivers," *Nature*, vol. 574, no. 7780, pp. 671–674, 2019, Epub 2019 Oct 30. DOI: `10.1038/s41586-019-1684-3`.

[3] J. C. Miller, "Insect natural history, multi-species interactions and biodiversity in ecosystems," *Biodiversity & Conservation*, vol. 2, no. 3, pp. 233–241, Jun. 1993, ISSN: 1572-9710. DOI: `10.1007/BF00056670`. [Online]. Available: `https://doi.org/10.1007/BF00056670`.

[4] M. Brydegaard and S. Svanberg, "Photonic monitoring of atmospheric and aquatic fauna," *Laser & Photonics Reviews*, vol. 12, no. 12, p. 1 800 135, 2018.

[5] K. Rydhmer, E. Bick, L. Still, *et al.*, "Automating insect monitoring using unsupervised near-infrared sensors," *Sci Rep*, vol. 12, p. 2603, 2022. DOI: `10.1038/s41598-022-06439-6`. [Online]. Available: `https://doi.org/10.1038/s41598-022-06439-6`.

[6] V. A. Kovalev and W. E. Eichinger, *Elastic lidar: theory, practice, and analysis methods*. John Wiley & Sons, 2004.

[7] M. Brydegaard, B. Kouakou, S. Jansson, J. Rydell, and J. Zoueu, "High dynamic range in entomological scheimpflug lidars," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 27, no. 4, pp. 1–11, 2021. DOI: `10.1109/JSTQE.2021.3062088`.

[8] J. Larsson *et al.*, "Atmospheric co2 sensing using scheimpflug-lidar based on a 1.57-μm fiber source," *Opt. Express*, vol. 27, no. 12, pp. 17 348–17 358, Jun. 2019. DOI: `10.1364/OE.27.017348`. [Online]. Available: `https://opg.optica.org/oe/abstract.cfm?URI=oe-27-12-17348`.

[9] G. Zhao *et al.*, "Particle profiling and classification by a dual-band continuous-wave lidar system," *Appl. Opt.*, vol. 57, no. 35, pp. 10 164–10 171, Dec. 2018. DOI: `10.1364/AO.57.010164`. [Online]. Available: `https://opg.optica.org/ao/abstract.cfm?URI=ao-57-35-10164`.

[10] M. Brydegaard *et al.*, "Short-wave infrared atmospheric scheimpflug lidar," in *EPJ Web of Conferences*, EDP Sciences, vol. 176, 2018.

[11] X. Wang, Z. Duan, M. Brydegaard, S. Svanberg, and G. Zhao, "Drone-based area scanning of vegetation fluorescence height profiles using a miniaturized hyperspectral lidar system," *Applied Physics B*, vol. 124, no. 11, p. 207, Oct. 2018, ISSN: 1432-0649. DOI: `10.1007/s00340-018-7078-7`.

[12] R. Boateng *et al.*, "Remote vegetation diagnostics in ghana with a hyperspectral fluorescence lidar," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 29, no. 4: Biophotonics, pp. 1–7, 2023. DOI: `10.1109/JSTQE.2023.3234022`.

[13] G. Zhao *et al.*, "Inelastic hyperspectral lidar for aquatic ecosystems monitoring and landscape plant scanning test," in *EPJ Web Conferences*, 2018, pp. 1–4.

[14] S. Jansson and M. Brydegaard, “Passive khz lidar for the quantification of insect activity and dispersal,” *Anim Biotelemetry*, vol. 6, no. 6, 2018. DOI: `https://doi.org/10.1186/s40317-018-0151-5`. [Online]. Available: `https://doi.org/10.1186/s40317-018-0151-5`.

[15] A. Runemark, M. Wellenreuther, H. H. E. Jayaweera, S. Svanberg, and M. Brydegaard, “Rare events in remote dark-field spectroscopy: An ecological case study of insects,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 18, no. 5, pp. 1573–1582, 2012. DOI: `10.1109/JSTQE.2012.2184528`.

[16] Xu, Zhicheng, *Insect Diversity Estimation in Entomological Lidar*, Student Paper, 2022.

[17] M. Li, A. Runemark, J. Hernandez, J. Rota, R. Bygebjerg, and M. Brydegaard, “Discrimination of hover fly species and sexes by wing interference signals,” *Advanced Science*, vol. n/a, no. n/a, p. 2 304 657, DOI: `https://doi.org/10.1002/advs.202304657`. [Online]. Available: `https://onlinelibrary.wiley.com/doi/abs/10.1002/advs.202304657`.

[18] M. Li *et al.*, “Bark beetles as lidar targets and prospects of photonic surveillance,” *Journal of Biophotonics*, vol. 14, no. 4, e202000420, 2021. DOI: `https://doi.org/10.1002/jbio.202000420`. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/jbio.202000420`. [Online]. Available: `https://onlinelibrary.wiley.com/doi/abs/10.1002/jbio.202000420`.

[19] M. Li, A. Runemark, N. Guilcher, J. Hernandez, J. Rota, and M. Brydegaard, “Feasibility of insect identification based on spectral fringes produced by clear wings,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 29, no. 4: Biophotonics, pp. 1–8, 2023, Early online date: 2022 Oct 31.

[20] L. Muller, “Elastic hyperspectral lidar for detecting coherent backscatter from insects,” 30 ECTS, Master’s Thesis, Lund University, Jun. 2022.

[21] L. Müller *et al.*, “Remote nanoscopy with infrared elastic hyperspectral lidar,” *Advanced Science*, vol. n/a, no. n/a, p. 2 207 110, DOI: `https://doi.org/10.1002/advs.202207110`. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/advs.202207110`. [Online]. Available: `https://onlinelibrary.wiley.com/doi/abs/10.1002/advs.202207110`.

[22] T. Cronin, S. Johnsen, N. Marshall, and E. Warrant, *Visual Ecology*. Princeton University Press, 2014, ISBN: 9781400853021. [Online]. Available: `https://books.google.se/books?id=v21_AwAAQBAJ`.

[23] NASA. “The sun.” NASA, Jan 22, 2013. Image Article. (2013), [Online]. Available: `https://solarsystem.nasa.gov/sun/overview/`.

[24] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*. John Wiley & Sons, 2019.

[25] S. Svanberg, *Atomic and Molecular Spectroscopy: Basic Aspects and Practical Applications*. Springer Science & Business Media, 2012, vol. 6.

[26] S. Jansson, “Entomological lidar,” Ph.D. dissertation, Lund University, 2020.

[27] E. Malmqvist, “From fauna to flames,” Ph.D. dissertation, Lund University, 2019.

[28] H. Yin *et al.*, “Iridescence in the neck feathers of domestic pigeons,” *Phys. Rev. E*, vol. 74, p. 051 916, 5 Nov. 2006. DOI: `10.1103/PhysRevE.74.051916`. [Online]. Available: `https://link.aps.org/doi/10.1103/PhysRevE.74.051916`.

[29] D. G. Stavenga, “Thin film and multilayer optics cause structural colors of many insects and birds,” *Materials Today: Proceedings*, vol. 1, pp. 109–121, 2014, Living Light: Uniting biology and photonics – A memorial meeting in honour of Prof Jean-Pol Vigneron, ISSN: 2214-7853. DOI: `https://doi.org/10.1016/j.matpr.2014.09.007`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S221478531400008X`.

[30] A. K. Gebru, M. Brydegaard, E. Rohwer, and P. Neethling, “Probing insect backscatter cross-section and melanization using kHz optical remote detection system,” in *Remote Sensing and Modeling of Ecosystems for Sustainability XIII*, W. Gao and N.-B. Chang, Eds., International Society for Optics and Photonics, vol. 9975, SPIE, 2016, p. 997 504. DOI: `10.1117/12.2236010`. [Online]. Available: `https://doi.org/10.1117/12.2236010`.

[31] J. F. James, *Spectrograph design fundamentals.* Cambridge University Press, 2007, ISBN: 9780521864633.

[32] E. Optics. “Reflective ruled diffraction gratings.” Accessed: 16/11/2023. (2023), [Online]. Available: `https://www.edmundoptics.com/f/reflective-ruled-diffraction-gratings/12220/`.

[33] *Linear array multianode pmt assemblies and modules*, H7260, HAMAMATSU, May 2015.

[34] *Dremel® digilab 3d45 3d printer*, `https://3pitech.com/products/dremel-digilab-3d45-3d-printer`, Accessed on 14/11/2023, 2023.

[35] Edmund Optics Inc. “600 grooves, 30mm square, 750nm ruled diffraction grating, technical information.” Accessed on December 13, 2023. (2023), [Online]. Available: `https://www.edmundoptics.com/p/600-grooves-30mm-square-750nm-ruled-diffraction-grating/12086/`.

[36] J. F. Wu, Y. Y. Chen, and T. S. Wang, “Flat field concave holographic grating with broad spectral region and moderately high resolution,” *Appl. Opt.*, vol. 51, no. 4, pp. 509–514, Feb. 2012. DOI: `10.1364/AO.51.000509`. [Online]. Available: `https://opg.optica.org/ao/abstract.cfm?URI=ao-51-4-509`.

[37] H. Månefjord *et al.*, “3d-printed fluorescence hyperspectral lidar for monitoring tagged insects,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 5: Lidars and Photonic Radars, pp. 1–9, 2022. DOI: `10.1109/JSTQE.2022.3162417`.

[38] *Nist atomic spectra database (asd) - cadmium*, `https://physics.nist.gov/PhysRefData/Handbook/Tables/cadmiumtable2_a.htm`, Accessed: [10/11/2023].

[39] “Södra sandby, skåne county, sweden — sunrise, sunset, and daylength, june 2023.” Accessed on: 11/10/2023. (), [Online]. Available: `https://www.timeanddate.com/sun/@2675692?month=6&year=2023`.

[40] “Weather data for svedala, skåne, sweden.” Accessed on: 11/10/2023. (), [Online]. Available: `https://www.wunderground.com/history/weekly/se/svedala/ESMS/date/2023-6-6`.

[41] MathWorks. “Matlab documentation: Spectrogram function.” Accessed: 13/12/2023. (2023), [Online]. Available: `https://se.mathworks.com/help/signal/ref/spectrogram.html`.

[42] E. Malmqvist, S. Jansson, S. Török, and M. Brydegaard, “Effective parameterization of laser radar observations of atmospheric fauna,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 22, no. 3, pp. 327–334, 2016. DOI: `10.1109/JSTQE.2015.2506616`.

[43] S. Jansson *et al.*, “Real-time dispersal of malaria vectors in rural africa monitored with lidar,” *PLOS ONE*, vol. 16, no. 3, pp. 1–21, Mar. 2021. DOI: `10.1371/journal.pone.0247803`. [Online]. Available: `https://doi.org/10.1371/journal.pone.0247803`.

[44] F. Nielsen, “Hierarchical clustering,” in *Introduction to HPC with MPI for Data Science*, ser. Undergraduate Topics in Computer Science, Published on 04 February 2016, Springer, 2016, pp. 195–211, ISBN: 978-3-319-21902-8. DOI: `10.1007/978-3-319-21903-5_8`. [Online]. Available: `https://doi.org/10.1007/978-3-319-21903-5_8`.

[45] I. Moya *et al.*, “A new instrument for passive remote sensing: 1. measurements of sunlight-induced chlorophyll fluorescence,” *Remote Sensing of Environment*, vol. 91, no. 2, pp. 186–197, 2004, ISSN: 0034-4257. DOI: `https://doi.org/10.1016/j.rse.2004.02.012`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0034425704000756`.

[46] M. Li, Unpublished PhD thesis, Ph.D. dissertation, Lund University, 2024.

[47] S. Jansson *et al.*, “Spatial monitoring of flying insects over a swedish lake using a continuous-wave lidar system,” *Royal Society Open Science*, vol. 10, May 2023. DOI: `10.1098/rsos.221557`.

A Single Event Detection

The code below shows how insect observations are extracted from the large set of data. The data is loaded as an array called `D` of dimensions (32 x 400000), for the 32 spectral bands during 20 seconds of measurement.

```matlab
% Define relevant parameters
fs=20000; % Sampling rate (Hz)
L=20*fs; % Width of window that is analyzed (20 seconds)
fMax=3000; % Maximum Frequency for spectogram (Hz)
width=50E-3; %  window width (s)
ol=95; % percentage of overlap for spectrogram
WL= width*fs; % sample points in window

% Baseline Correction
% Remove the baseline drift from the signal.
Dbl = D - imerode(D, ones(WL, 1)); % Baseline-corrected signal
Ibl = sum(Dbl, 2); % Baseline intensity

% Define parameters for the spectrogram computation.
nol = floor(2 * WL * ol / 100); % Number of samples overlapping in the
↪  window
W = normpdf(-WL:WL, 0, WL / (4 * sqrt(2 * log(2)))); % Gaussian window
F = linspace(1/width, fMax, 2 * fMax * width); % Frequency bins

% Compute the spectrogram of the baseline-corrected signal.
[S, F, T, P] = spectrogram(Ibl, W, nol, F, fs);

% Define threshold for mask
trsh=10; %threshold in respect to noise IQR
maskTrsh=5; % number of mask elements
Trsh = repmat(median(P')' + trsh * iqr(P')', [1 size(P, 2)]); % Threshold

% Generate a binary mask for the power spectrum.
mask = imdilate(imerode(P, shrink), grow) > Trsh; % Mask thresholding
Q = bwconncomp(sum(mask) > maskTrsh); % Connected components

% Extract and store relevant observations.
entry = struct('t', [], 'D', []);
for n = 1:Q.NumObjects
    Tind = Q.PixelIdxList{n};
    tind = find(t < T(Tind(1)), 1, 'last'):find(t > T(Tind(end)), 1,
    ↪  'first');
    entry(n).t = ft + t(tind(1)) / (60 * 60 * 24); % Time in days
    entry(n).D = Dbl(tind, :); % Baseline-corrected signal
end
obs = [obs entry]; % Append to the observations list
save([fn(1:(end-3)) 'mat'],'obs','-v7.3') % Save variable
```

B Parameterization of Dynamic signal

The code below shows how the parameterization in the temporal domain is applied to the raw data.

```matlab
function [Ihat,Ihat_apod]=parameterize(I, t_vec, wbf, f, n)
    % Parameterize signal in sin wave with 5 harmonics.
    % Input parameters:
    % I:     Intensity signal
    % wbf:   wing beat frequency (Hz)
    % f:     sampling frequency (Hz)
    %
    % Output parameters
    % Ihat: Parameterized intensity signal
    % Ihat_apod: Parameterized intensity signal multiplied by envelope

    envl=(imerode(I,ones([round(f/wbf) 1]))+imdilate(I,ones([round(f/wbf) 1])))/2;
    Window=normpdf(0:(2*round(f/wbf)),f/wbf,f/(6*wbf));
    envl=conv(envl,Window,'same');
    envl=envl/max(envl);

    regr=t_vec'.^0; % column of ones

    for h=1:n
        regr=[regr [sin(2*pi*wbf*h*t_vec') cos(2*pi*wbf*h*t_vec')]];
    end

    regr_apod=regr.*repmat(envl,[1 size(regr,2)]); % apodize
    th=regr_apod\I;
    Ihat_apod=regr_apod*th;

    Ihat=regr*th;

```

C Reflectance calculations

This code calculates the reflectance. Here, `spec_1305` is a 32x1 array containing the mean of the spectrum measured at 13:05 each day, `I_fit` is the fit displayed in figure 22, `t_mod` the modulated time of the measurement in serial format, `Dhat` is an 32xN array containing the parameterized spectrum, where `N` is the length of the observation multiplied by the sampling frequency.

```matlab
i_period = interp1(t_vec,1:length(t_vec),1/(wbf/1e3),'nearest');  % Index of one wing beat period
Dpar = Dhat(1:i_period,:); % Spectra of one wing beat period
Ipar = sum(Dpar,2); % Intensity of one wing beat
[flash, i_flash] = max(Ipar); % Location of flash
body = median(Ipar);
```

```
i_body = interp1(Ipar,1:i_period,median(Ipar),'nearest'); % Location of
 ↪  body

% Calculate white stick spectrum at given time
whitestick = zeros(1,32);
spec_1305_sum=sum(spec_1305);
for l=1:32
    scale = spec_1305(l)/spec_1305_sum;
    whitestick(l) = feval(I_fit, t_mod)*scale;
end

% Calculate Reflectance
Flash_spec =  Dpar(i_flash,:); % Flash spectrm
Body_spec = Dpar(i_body,:); % Body spectrum
R = (Flash_spec-Body_spec) ./ whitestick; % (white - dark_cal); %
 ↪  Calculate Reflectance
Body_plot = Body_spec./ whitestick; % Spectrum of body for plotting
```

D Fitting the fringe formula

The code below shows how the fringe formula is fitted to the spectral data. This is done using the curve fitting toolbox that is built in to Matlab. The reflectance values are stored in a list called `R` of dimensions 32 × 1. The corresponding spectral values are stored in a list called `bandCntrs2` of dimensions 32 × 1.

```
% Fit fringe formula
wingFit = fittype(
 ↪  'fringeFunc(d,w,th)*amp*w.^2.71./(w.^2.71+w0.^2.71)+bias*w0.^2.71./(w.^2.71+w0.^2
 ↪  'coefficients',{'d' 'amp'  'w0' 'bias'}, 'problem', th);

fitOpt=fitoptions(wingFit);
fitOpt.Robust='on';
fitOpt.Lower=[0.3  0   0   0  ];
fitOpt.Upper=[ 0.9  100 100 100 ];

initial_guesses =   [0.3, 3, 3, 0.8;
                    0.4, 3, 3, 0.8;
                    0.5, 3, 3, 0.8;
                    0.6, 3, 3, 0.8;
                    0.7, 3, 3, 0.8
                    0.8, 3, 3, 0.8
                    0.9, 3, 3, 0.8];

best_gof = -inf;  % Initialize the best goodness of fit
best_fitRes = [];

for i = 1:size(initial_guesses, 1)
    start_d = initial_guesses(i, 1);
```

```
    start_amp = initial_guesses(i, 2);
    start_w0 = initial_guesses(i, 3);
    start_bias = initial_guesses(i, 4);
    fitOpt.StartPoint = [start_d, start_amp, start_w0, start_bias];

    % Fit the data with the current initial guess
    [fitRes, gof] = fit(bandCntrs2/1e3, R', wingFit, itOpt);
    % Check if this fit has a better goodness of fit
    if gof.rsquare > best_gof
        best_gof = gof.rsquare;
        best_fitRes = fitRes;

        disp(['New best R^2: ' num2str(best_gof) ' with starting thickness ' num2str(start_d*1e3) ' nm'])
    end
end

% Fringe function
function RsFilm=fringeFunc(d,w,th)
    n=1.517+8800./((w*1E3).^2); %Chitin


    Rs=((cosd(th)-sqrt(n.^2-sind(th).^2))./(cosd(th)+sqrt(n.^2-sind(th).^2))).^2; % Fresnell eq. for S-pol
    Rp = ((sqrt(1 - (sind(th)./n).^2) - n.*cosd(th)) ./ (sqrt(1 - (sind(th)./n).^2) + n.*cosd(th))).^2; % Fresnell eq. for P-pol
    Rf = (Rs + Rp)/2;
    term=(4*Rf.*sind(2*180*d.*sqrt(n.^2-sind(th).^2)./w).^2);
    RsFilm=term./((1-Rs).^2+term);
end
```

E Singular Value Decomposition

This code shows how the singular value decomposition was achieved. In this code `D` is a matrix containing of size (32xN).

```
[U, S, V] = svd(D, 'econ');
ss=(1:length(S))'; % Spectral components
medSlope=median(diff(log(diag(S)))); % % Median slope
medS=S(length(S)/2,length(S)/2); % Median singular value
Shat=median(diag(S))*exp((ss)*medSlope)/exp(length(S)/2*medSlope); % Exponential slop
Siqr=median(diag(S))*iqr(diag(S)./Shat); % IQR offset
Ns = diag(S)>Shat+Siqr ; % Number of values above line
i = 1;
NoS = 0;
while Ns(i)==1
    NoS = NoS +1;
    i=i+1;
end
```