Department of Physics
Umeå University

Weather effects on short-range LiDAR and their classification

David Blagojevic (ogbl0002@student.umu.se)

September 2, 2022

Abstract

Today we are seeing exciting developments in the field of autonomous vehicles, on both software and hardware. Veoneer is a company making a contribution where research and manufacturing is being done on hardware and active safety. One of the most important aspects in this field is road safety, where understanding the behaviour of sensors used in vehicles is essential. From the point of view of safety, understanding how weather affects the sensors is necessary for a successful deployment. This study is a continuation of previous studies done at Veoneer, and regards how various adverse condition affect the performance of a short-range LiDAR and gives a thorough description of the involved physical processes.
Data collected over a couple of months was analysed and compared to theoretical models in order to establish their validity. In addition, LiDAR measurement were done in a chamber where conditions could be varied in a controlled manner.
Furthermore, analysis methods were used to transform the data into a form potentially more useful for use in machine learning algorithms to estimate the ability to classify conditions based on LiDAR signals.
The used models showed mixed results, with some showing more agreement than others. Models regarding foggy conditions generally showed greater agreement with data than in other conditions, although some variation around the predictions did occur. In regards to the performance of the classification algorithms, there were also mixed results, where the sensitivity in fog was at most 96 % and the precision at most 64 %.
This thesis also enables and suggests further research into the utility of short-range LiDAR both in the field of autonomous vehicle safety as well as in use of other fields such as meteorology.

Supervisors: Emil Hällstig
Examiner: Johan Zakrisson

Preface

This project concludes my studies in engineering physics at Umeå university at the physics department. I am very grateful for Veoneer providing me with the opportunity and making it possible to finalise my degree.
I extend my sincerest gratitude to my supervisor Emil Hällstig for providing me continual guidance and directions during the project, as well as to André Berggren for delivering help where I lacked the necessary knowledge and experience. I also extend my appreciation to my university supervisor Tobias Dahlberg, and my examiner Johan Zakrisson for the given feedback on the project.
Lastly I would like to thank my family for all the given support.

Contents

1 Introduction

Veoneer is a company driving development and research in autonomous vehicles and active automotive safety, as well as creating products in the same field. For this purpose, improvement in sensor technology is of out most importance. The performance of autonomous vehicles(AV) is heavily dependent on sensors, since AVs use technology such as computer vision, lidar ranging, thermal cameras and others, all of which rely on accurate and predictable behaviour of the sensors to present an accurate picture of the vehicles environment.
One of the ranging methods is using LiDARs, which operate on the same principle as RaDARs, but with a much shorter wavelength on the range of visible to infrared light. Light with short wavelength is more easily disturbed by its environment, which can cause light scattering and absorption. This leads to weakening of lidar signals, and potentially worsening the performance of AVs. The particles in various weather conditions such as fog, rain, snow and others can potentially cause loss of signals in lidar ranging.
This study aims to further improve research of LiDAR performance in such adverse weather conditions. Previous studies have been done by Marcus Hedlund[1], and Johan Viklund[2] exploring the behaviour of LiDAR signals. A measurement system was built by Marcus Hedlund consisting of a long ranging LiDAR in 2020, measuring effects on a target 90 m away, along with a weather station for determining conditions such as temperature, luminance and visibility. This measurement system has been gathering data ever since. The measurement data was utilised in Johan Viklunds work. In addition, another short range LiDAR unit was set in 2021 with a target 18 m away, but the data was largely unused.
This study aims to do further research based on the previous works by Marcus Hedlund and Johan Viklund, mainly utilising the measurements from the short-range LiDAR and research how it is affected by adverse conditions.

1.1 Goals and Objectives

This study is further developement from studies done by Marcus Hedlund[1] and Johan Viklund[2]. In his master's thesis, Marcus focused on solving the Radiative transfer equation and develop a numerical method for solving it for different conditions.
Johan Viklunds study went on to study the effect of weather conditions on long-range LiDAR. Towards the end of the study, a short-range LiDAR with a target at 18 meters was installed by Hedlund for future study. In this study, one goal is to determine how the short-range LiDAR signals are affected by weather conditions, and in particular the effect of fog, but also of rain and mixed weather where both may occur, although not many condition fit that criteria. Hedlund also analysed the zero-pulse and how it is correlated with weather parameters such as temperature and visibility. In this project the question of weather the visibility may be estimated based on theoretical ground is being investigated, where the LiDAR equation is of importance. Also, the study includes the general description of light transfer for which will motivate the use of Radiative transfer theory. Yet another question being addressed by this study is if classification of weather

conditions of fog, rain and clear weather may be differentiated by machine learning algorithms. Further development from the previous study is that medium description will be stated in more detail, especially for fog. There the particle/droplet distribution is of great importance, and parameters derived from it, such as effective size and liquid water content, are studied in their effects on short-range LiDAR. For that purpose a collaboration with RISE is done by taking LiDAR measurements in a controlled manner with a climate chamber.

2 Theory

In this section, the necessary theoretical framework will be established. Starting from the descriptions of how radiation is propagated through space and how it varies with time. For this, radiative transfer is central, which involves the radiative transfer equations. For certain applications such as LiDAR technology, the radiative transfer equation may be used to approximate a useful form of the equation in order to describe light propagation. The approximations results in the LiDAR equation, which is the main description of light propagation in this study. The study of the medium in which the radiation is propagating involves Mie theory, which is also addressed in this study. Lastly some data analysis and machine learning concepts are introduced.

2.1 Radiative transfer

There are a few ways light can be represented. In this study, we will mainly take the approach of light being described by a electromagnetic wave[3], although some discussions will regard light being described as a particle called photon[4]. Light propagating in a medium will interact with particles in the medium. This is particularly true if the wavelength of the light is on the same order or smaller than the typical size of the particles in the medium. Common scattering is Rayleigh scattering. A more general type of scattering called Mie scattering[5].

2.1.1 Intensity

In radiation, intensity is a central quantity and describes how energetic the radiation is. In order to derive the necessary quantities for this study, we first consider the quantum physical perspective of radiation being made up of particles called photons.
Consider the function $f_\nu(\vec{\Omega})$ describing the number of photons per unit volume travelling in the $\vec{\Omega}$ direction having the energy in the interval of $h\nu$ to $h(\nu + d\nu)$, which are passing through a unit area and unit time[6]. Here h is plancks constant and ν is photon frequency Now the area element dA with normal $\hat{n}$ makes an angle θ to the direction of propagation. The amount of area projected normal to the direction of travel is $dA\cos(\theta)$ as is shown in figure 1.
In the time dt the photons will have traveled a distance of $c\,dt$, where c is the speed of light, so that the volume swept out is $(dA\cos\theta)\,c\,dt$. The number of photons is this volume is $cf(\vec{\Omega})\,dA\cos\theta\,dt\,d\nu\,d\Omega$. With each photon carrying energy $h\nu$, the change in the total energy dW_ν in the time dt can be determined.

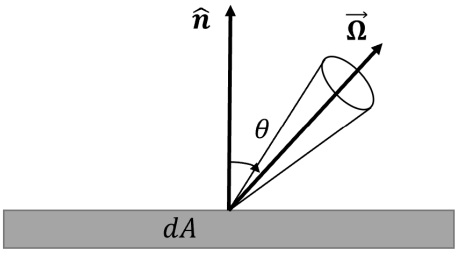

Figure 1 – Figure illustrates the flow of radiation and the geometrical properties used in calculations.

The total energy of the photons is given by

$$dW_\nu = h\nu c f_\nu(\Omega)\cos(\theta)\, dA\, dt\, d\nu\, d\Omega. \tag{1}$$

From equation (1), we define *radiance* or *spectral intensity* as

$$I_\nu(\Omega) = \frac{dW_\nu}{(dA\cos\theta)\, dt\, d\nu\, d\Omega}. \tag{2}$$

In general, the specific intensity shown in equation (2) is a function of direction, position and time so $I_\nu = I_\nu(x, y, z, t, \vec{\Omega})$.

2.1.2 Radiation absorption

Considering a beam of monochromatic radiation propagating as shown in figure 2, inbound on a optically active material with thickness ds. The radiation passing through the material will be partially absorbed.
For absorption we define the absorption coefficient $K_\nu(r)$, which represents the proportion of radiation absorbed by the material per unit length along the direction of incident radiation. The total amount radiation absorbed by the material per unit time in the solid angle $d\Omega$ is $K_\nu(r)\, I_\nu\, d\Omega\, ds$[6]
Matter emits radiation of its own that in general may propagate in all directions, in particular the direction of the beam. This has the effect of increasing the intensity of the beam. The addition of intensity due to matter is a function of temperature T, and is given by Planck's law as

$$B_\nu(T) = \frac{2h\nu^3}{c^2}(e^{\frac{h\nu}{k_B T}} - 1)^{-1}, \tag{3}$$

where k_B is Boltzmann's constant[7].

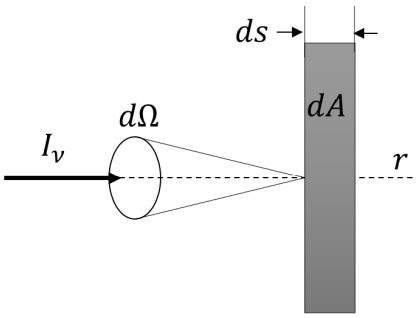

Figure 2 – Figure illustrates how radiation normally incident on a medium of infinitesimal thickness affects the radiation.

2.1.3 Radiation scattering

In additionally, some fraction of photons carrying the energy will be scattered out of the beam, contributing to further decrease in intensity. For scattering, we define the scattering coefficient $\sigma_\nu(r)$ as representing the proportion of the intensity scattered *in all directions* per unit length. The amount of scattered radiation is then $\sigma_\nu(r)\, I_\nu d\Omega$ [8]. The expression does not give any insight into how the scattering directions are distributed. For that, we define the phase function $P_\nu(\vec{\Omega}, \vec{\Omega}')$ so that that $P_\nu(\vec{\Omega}, \vec{\Omega}')d\Omega/4\pi$ is the probability that incident radiation in the direction $\vec{\Omega}'$ with solid angle $d\Omega'$, is scattered in a new direction $\vec{\Omega}$ with solid angle $d\Omega$ as seen in figure 3. Hence we have that

$$\frac{1}{4\pi}\int_\Omega P_\nu(\vec{\Omega}, \vec{\Omega}')d\Omega = 1, \tag{4}$$

which says that the probability of scattering in any direction is 1 as shown in equation (4)[6]. It is assumed that the phase function only depends on the scattering angle θ_0 so that $P_\nu(\vec{\Omega}, \vec{\Omega}') = P_\nu(\cos\theta_0)$.[6].

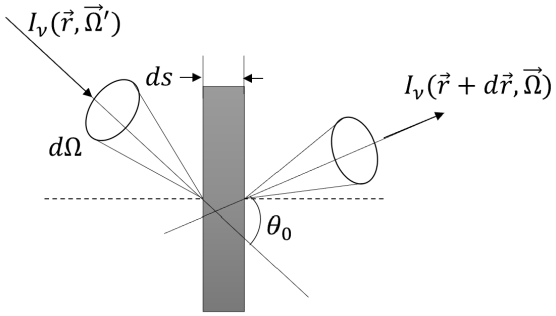

Figure 3 – Figure shows the scattering process of radiation and defines the scattering angle θ_0.

In the case when the scattering particles are much smaller than the wavelength of the radiation, Rayleigh scattering occurs for which the phase function is given by

$$P_\nu(\cos\theta_0) = \frac{3}{4}(1+\cos^2\theta_0), \tag{5}$$

where θ_0 is the scattering angle[9]. The Rayleigh phase function is seen in figure 4, where the length of the arrows is proportional to the probability of scattering in that direction. It can be seen that the probability of scattering backwards is the same as scattering forward, while the probability or scattering perpendicular to the direction of propagation is smaller. In general, most of the scattering happens in the forward direction[10].

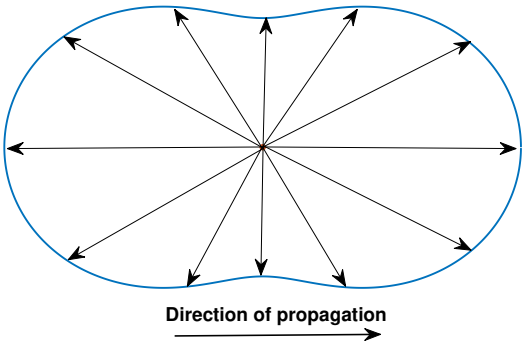

Figure 4 – Figure shows the phase function (5) which applies to Reyleigh scattering in which the forward and backward scattering are equal.

2.1.4 Radiative transfer equation

We will now combine the effects that the material can have on the radiation, which in the end will produce an equation describing how the intensity of the radiation changes as it propagates in the medium.
Starting with a monochromatic beam incident on an emitting, scattering and absorbing medium characterised by its absorption coefficient $K_\nu(s)$ and its scattering coefficient $\sigma_\nu(s)$ as seen in figure 5.

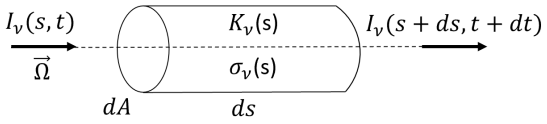

Figure 5 – Figure depicts the intensity change when travelled through a volume element in an absorbing and scattering medium.

The change in radiance is

$$DI_\nu = I_\nu(s+ds, t+dt) - I_\nu(s,t), \tag{6}$$

where the quantity $W_\nu\, ds\, d\Omega\, d\nu\, dt$ denotes the net gain of radiant energy by the volume element in figure 5[6]. From the difference in equation (6) and definition of radiance in equation (2) leads to writing the rate of change in radiance as

$$\frac{DI_\nu(s,t)}{Ds} = W_\nu. \tag{7}$$

The differential operator in equation (7) may be expressed in terms of partial derivative of time and space so that we write

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \frac{\partial I_\nu}{\partial s} = W_\nu, \tag{8}$$

where c is the speed of light[11]. To the right-hand side in (8), we account for four scenarios listed below.

1. $W_{\nu 1}$ energy emitted by the volume element
2. $W_{\nu 2}$ energy absorbed the volume element
3. $W_{\nu 4}$ energy scattered out of the volume element
4. $W_{\nu 1}$ energy scattered into volume element.

The first term $W_{\nu 1}$ is given by the Planck's radiation in equation (3) multiplied by the absorption coefficient. This is only valid under the assumption that the material is in local thermal equilibrium[7]. The second term is given by the proportion of the intensity absorbed in the volume as $W_{\nu 2} = -K_\nu(s)I_\nu(s,t)$. The amount scattered out of the volume is similarly written as $W_{\nu 3} = -\sigma_\nu(s)I_\nu(s,t)$.
The final term represent the amount of radiation scattered from all directions into the direction of propagation $\vec{\Omega}$. This is given by

$$W_{\nu 4} = \frac{\sigma_\nu(s)}{4\pi}\int_{\Omega'} P_\nu(\cos\theta_0)I_\nu(s,\vec{\Omega}',t)d\Omega' \tag{9}$$

where $P_\nu(\cos\theta_0)$ in equation (9) is the phase function describing the scattering distribution. Now, the attenuating terms and the adding terms are put together to produce

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \frac{\partial I_\nu}{\partial s} = -(K_\nu(s) + \sigma_\nu(s))I_\nu(s,t) + j, \tag{10}$$

where j in equation (10) is called the source term corresponding to the amount of intensity added to the beam[6]. The source term is explicitly as shown

$$j = K_\nu(s)B(T) + \frac{\sigma_\nu(s)}{4\pi}\int_{\Omega'} P_\nu(\cos\theta_0)I_\nu(s,\vec{\Omega}',t)d\Omega' \tag{11}$$

In equation (11), all the terms that add to the energy are collected into a single expression. The total attenuation can be summarised as the sum of the absorption and scattering coefficients, $\beta(s) = K_\nu(s) + \sigma_\nu(s)$, where $\beta_\nu(s)$ is appropriately named extinction coefficient. Defining $J = j/\beta_\nu(s)$, the change in specific intensity is

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \frac{\partial I_\nu}{\partial s} = -\beta_\nu(s)(I_\nu - J). \tag{12}$$

Equation (12) is the general radiative transfer equation. In many applications, the time dependence term is much smaller compared to the other terms and can be ignored. If the source term J is small, a simple solution may be obtained from equation (12)

$$I_\nu(s) = I_\nu(0)e^{-\beta_\nu s}, \tag{13}$$

which is known as Beer-Lambert law. Here, the extinction coefficient is assumed to not vary in space. Some particular solutions are seen in figure 6.

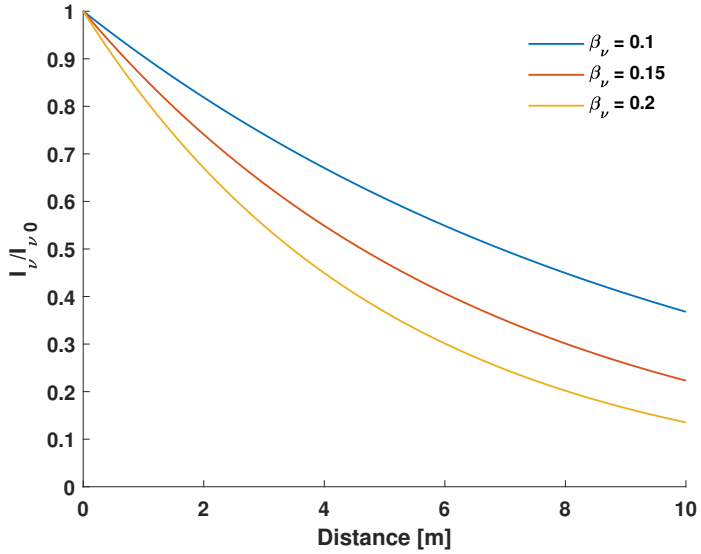

Figure 6 – Figure shows the attenuation of intensity as a function of distance from predicted by Beer-Lambert law (13) for various extinction coefficients.

2.1.5 LiDAR equation

Here we discuss the equation used to describe the LiDAR signal propagation in a medium with an extinction coefficient $\beta_\nu(s)$. We proceed from the point described by the Radiative equation (12).
As was discussed in the previous chapter, the temporal term is usually smaller than the other terms and may in many applications be ignored. One must however be cautious when applying this assumption when dealing with pulsed light, since then the temporal term is larger for very short pulses. However in LiDAR application this assumption is sufficient. Even if the rate of change of intensity with time is comparable to the spatial derivatives, it is divided by the speed of light making the overall contribution very small[12].
Another assumption made is that $I_\nu >> J$. That is, the intensity is much larger than the source term. This is true if the scattering from the medium into the direction of signal propagation is small, along with if the medium is at a relatively low temperature. This is under the assumption the only single scattering occurs. The LiDAR equation relates how the power of a light pulse is attenuated in a propagating medium. It is given by

$$P(s) = P_0 \frac{c\tau}{2} \frac{A\, O(s)\, P_\nu(-1)}{s^2} e^{-2\beta_\nu s} \tag{14}$$

where P_0 is the power output of the LiDAR, τ is the pulse duration, $O(r)$ is the overlap function of the sensor and laser beam and A is the area of the sensor receiving the beam. In this thesis, the detector parameter and distance are kept constant, making it possible to normalise the LiDAR equation so to isolate the effect of the extinction coefficient better.

In figure 7, the amplitudes re shown normalised at 10 m and extinction coefficient of 0 m^{-1} for a few distances and extinction coefficients.

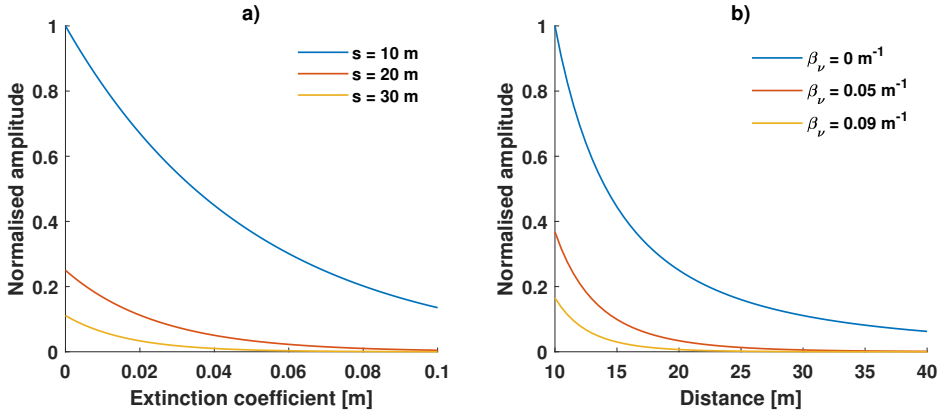

Figure 7 – Figures show the normalised amplitudes modelled by the LiDAR equation (14) as a function of extinction coefficient and distance to target. In a), the normalised amplitude is shown for distances of 10, 20 and 30 m, and in b), normalised amplitudes are shown for extinction coefficients of 0, 0.05 and 0.09 m^{-1}

2.2 Medium description

As can be seen in the LiDAR equation (14), the effect of the atmospheric medium through which the light propagates is summarised by the extinction coefficient β_ν. In general, the LiDAR equation described the attenuation in a medium for which the properties may vary with position, but a common assumption is that the medium is homogeneous, hence the extinction coefficient is a constant. A study of the extinction coefficient is essential for understanding effect on LiDAR signals. Therefore in the preceding chapters general descriptions along with approximations and empirical models are discussed in more detail.

2.2.1 Mie Theory

There are two ways of descibing the medium in which light is propagating. One way is from a macroscopic perspective. There the important quantity is the extinction coefficient. It is known that the atmosphere is made out of a large number of different gases, particles and aerosols. Various adverse conditions can be described from the microscopic perspective. In the microscopic view, the medium is described by the sizes of the particles, how they are distributed and from which matter the particles are made of. Figure 8 illustrates the different perspectives.

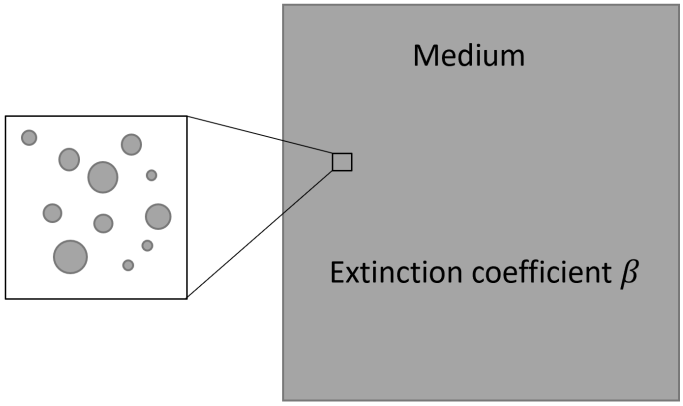

Figure 8 – Figure illustrates the propagating medium. From a macroscopic view it is enough to know the extinction coefficient. In the microscopic realm, the medium is made out of particles with varying size where the size distribution may be used to derive useful quantities such as extinction coefficient.

The study of how these particles affect radiation is described by *Mie theory*, and is a way to connect the microscopic and macroscopic perspectives.
There are many different distributions that are commonly assigned to atmospheric particles, with one of them being the log-normal distribution [13]. The random variable R is normally distributed if $\ln(R) \sim \mathcal{N}(\mu, \sigma^2)$, and has a distribution $n(r)$. Figure 9 shows a typical shape of log-normal variable which could be the radius of a particle.

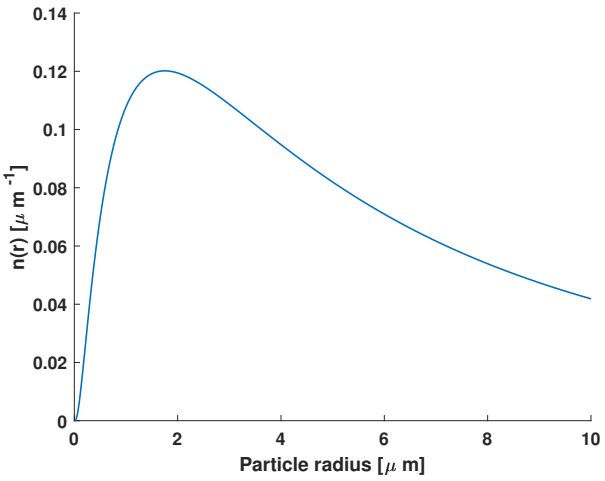

Figure 9 – Figure shows probability distribution of log-normally distributed particle radii with parameters $\mu = 2$ and $\sigma = 1.2$.

From the above figure, we can expect that much of the particles are described by smaller size, there are a smaller number of them with larger size. There are other distributions that are commonly assigned to particles, such as the modified gamma distribution[14] or Rosin-Rammler distribution[15]. Both of the distributions have the quality that a smaller sizes are more numerous.
Using Mie teory, the extinction coefficient of a medium is expressed as

$$\beta = \int_0^\infty Q_{ext}(r,\gamma)n(r)\pi r^2 \mathrm{d}r, \tag{15}$$

where $Q_{ext}(r,\gamma)$ is the extinction efficiency, and $n(r)$ is the particle size distribution. Although the integration limits are from 0 to ∞, in practice we can integrate from some minimum size r_{min} to the largest size r_{max}.
In order to determine the extinction coefficient in equation (15), the extinction efficiency is to be determined firstly. It is given by

$$Q_{ext} = \frac{2}{x^2}\sum_{n=1}^{\infty}(2n+1)\mathrm{Re}(a_n + b_n). \tag{16}$$

The coefficients a_n and b_n in equation (16) are calculated using the Basel and Ricci functions, and $x = 2\pi r/\lambda$, with r being the particle radius and λ the radiation wavelength[16]. However in for typical droplet sizes the extinction efficiency is approximately constant and more suitable measures may be used to determine the extinction coefficient[14][17]. These are discussed in the following section.

2.2.2 Particle distribtion

In this section, some useful properties in describing particles are shown, starting with the particle distribution itself. The particles size distribution in air, and particularly if fog conditions are represented by the distribution $n(r)$ which describes the number of particles in a unit volume of radius r per class width dr. The distribution $n(r)$ then is the number of particles per unit volume and length(usually in $1/(\mathrm{m}^3\ \mu\mathrm{m})$).
There are many suggested distributions that are adequate models for the particle sizes, with one of them being the log-normal distribution as shown in the previous section in figure 9. While log-normal distribution is a good description for general aerosols, for conditions such as fog or clouds, the modified gamma distribution is a good choice[14]. Its mathematical representation is given by

$$n(r) = ar^\alpha \exp\left[-\frac{\alpha}{\gamma}\left(\frac{r}{r_c}\right)\right], \tag{17}$$

where α and γ are parameters varying the slope of the distribution, r_c is the mode of the distribution and a is a proportionality factor. An example of a modified gamma distribution is shown in figure 10.

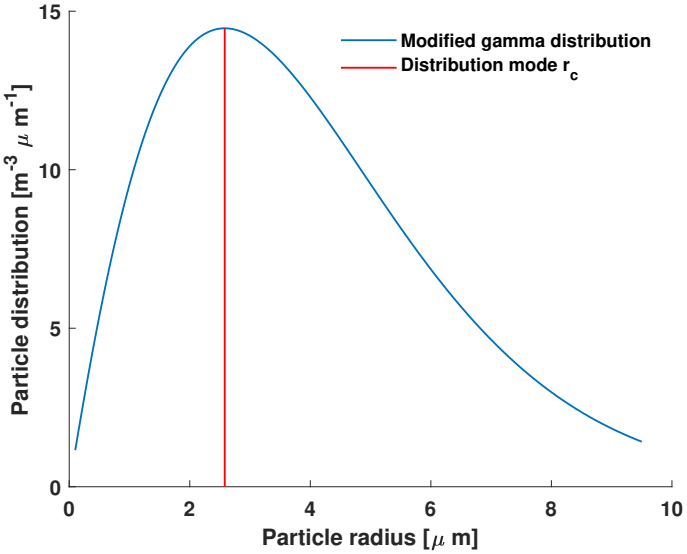

Figure 10 – Figure shows the modified gamma distribution (17) with parameters $r_c = 2.58$ μm, $\gamma = 1.37$, $a = 11.63$ and $\alpha = 1$.

As can be seen, this distribution has similar shape as the log-normal distribution, but the number of larger particles drops faster than in the log-normal distribution.
Although these distributions are appropriate in many cases, they are very limited in the since that in many cases, the sizes are bimodaly distributed(and even trimodally). In that cases the fog is considered as mixtures of various types. In that case there are conventionally two types of fog listed as

1. Maritime fog

2. Continental fog

Maritime fog is characterised as having large droplet bigger that 5-10 μm. It can be considered as advection fog. Continental fog is characterised by small droplets of a few micrometers($<$ 5-10 μm) and is a good representation of radiation fog[14].
Even though the distribution of fog may not necessarily be described by a unimodal distribution, henceforth it will be assumed in this study. We now discuss some properties of the distributions. First being the total number of particles within a unit volume given by

$$N = \int_0^\infty n(r)dr. \tag{18}$$

where N is the number density. The probability distribution is given by $n(r)/N$. Therefore the average size is given by

$$r_{mean} = \frac{\int_0^\infty rn(r)dr}{N}, \tag{19}$$

where the radius is integrated over all sizes. In practice, one can integrate from some meaninful smallest and largest size.
Yet another useful property is the total amount of mass per unit volume. For a spherical particle with density ρ and radius r, the mass is given by $4\pi r^3\rho/3$. For a medium such as fog, the density is constant. The quantity of *liquid water content* describes the amount of mass per unit volume usually in grams per cubic meter. It is calculated as

$$\text{LWC} = \int_0^\infty \frac{4\pi r^3 \rho}{3} n(r) dr. \tag{20}$$

Similarly, one can determine the total surface area S in a unit volume as

$$S = \int_0^\infty 4\pi r^2 n(r) dr. \tag{21}$$

A very important property is the effective size(diameter) which gives the typical size of the droplets in fog. It is the diameter that droplets would have if they were all the same and producing the same surface area and liquid water content[14]. The effective diameter is given by

$$D_{eff} = \frac{\int n(D) D^3 dD}{\int n(D) D^2 dD}. \tag{22}$$

In terms of Surface area and liquid water content it is written as

$$D_{eff} = \frac{6\,\text{LWC}}{\rho S}. \tag{23}$$

These quantities apply generally to any kind of medium such as for example ice crystals, though the equations might be altered slightly since the shape of ice crystals may not be spherical[14].
Now, the most general expression for the extinction coefficient is given by equation (15). If the droplets are much bigger than the radiation wavelength, the extinction efficiency is approximately constant and equal to 2[2][17]. For wavelength in the visible range, the extinction coefficient may be expressed in terms of liquid water content and effective size as

$$\beta = \frac{3\,\text{LWC}}{\rho D_{eff}} \tag{24}$$

2.2.3 Empirical models

As it is often not feasible to describe the medium by its microscopic properties such as particle size distribution, many other useful measures and empirical models are used to determine the extinction coefficient.
MOR(Meteorological Optical Range), also called *visibility*, is a measure of light attenuation in a medium. Visibility is defined as the distance where the contrast between a

entirely black and an entirely bright surface is 5 %, although initially it was set at 2 % [14]. An empirical relationship between visibility and coefficient was found by Kim and Kruse, based on the wavelength of the used light[18]. The relationship is described by

$$\beta = \frac{3.91}{V}\left(\frac{\lambda}{\lambda_0}\right)^{-q}, \tag{25}$$

where V is the visibility, λ_0 is 550 nm, and the parameter q is dependent on visibility and is given by

$$q = \begin{cases} 1.6,\ V > 50\ \text{km} \\ 1.3,\ 6 < V < 50\ \text{km} \\ 0.16V + 0.34,\ 1 < V < 6\ \text{km} \\ V - 0.5,\ 0.5 < V < 1\ \text{km} \\ 0,\ V < 0.5\,\text{km} \end{cases}$$

Another two models for fog attenuation proposed by Naboulsi[19], which discriminate between two main types of fog. These are

1. Advection fog

2. Radiation fog

These were discussed in the previous chapter when describing the characteristics of particle distributions. In advection fog, the droplets are characterised by typically smaller droplets of a few micrometers($<$ 5-10 μm), while radiation for is typically made of larger droplets bigger that 5-10 μm[14]. For advection fog the extinction coefficient is linear in wavelength and the empirical model is given as

$$\beta_{adv} = \frac{0.11478\lambda_{\mu m} + 3.8367}{V}, \tag{26}$$

where $\lambda_{\mu m}$ is the used wavelength in micrometers. For the other type of fog, namely radiation fog, the extinction coefficient is given by

$$\beta_{rad} = \frac{0.18126\lambda_{\mu m}^2 + 0.13709\lambda_{\mu m} + 3.7502}{V} \tag{27}$$

An example of the models are shown in figure 11.

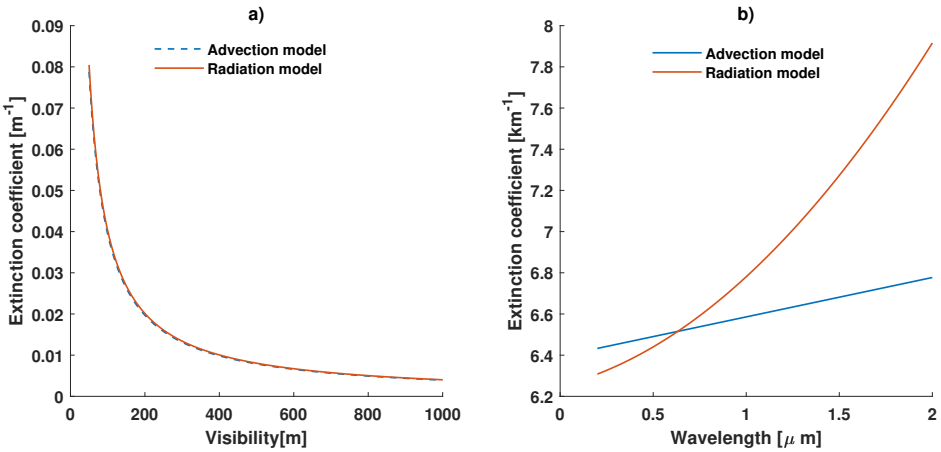

Figure 11 – Figure shows the extinction coefficient according to Naboulsi models for advection and radiation fog. In a) the advection and radiation model are shown with varying visibility at a wavelength of 905 nm, while b) shows varying wavelength for a visibility of 600 m.

The behaviour shown in figure 11 indicates that the distinction between the two types are negligible when using wavelength of 905 nm. However the difference between the two models grows larger with larger wavelength. This is mainly due the the difference in typical size of the droplets.
As with visibility there are models for predicting extinction coefficient based on precipitation. Most notably the Carbonneu attenuation for rain[2][20][21]. These are described by

$$\beta = (aR^b)\left(\frac{\ln(10)}{10}\right), \tag{28}$$

where R is the precipitation in mm/h and β is the extinction coefficient in km^{-1}. The constants a and b empirically determined mainly for two types of rain described by these models. One is for low intensity rain with smaller droplets[2]. In that case the constants are determined to be $a = 1.076$ and $b = 0.67$. For rains characterised with high intensity and large droplets, the constants are $a = 0.365$ and $b = 0.63$[20]. The resulting models are shown in figure 12. These types of showers are more common in tropical parts of the world, and are hence named tropical rains.

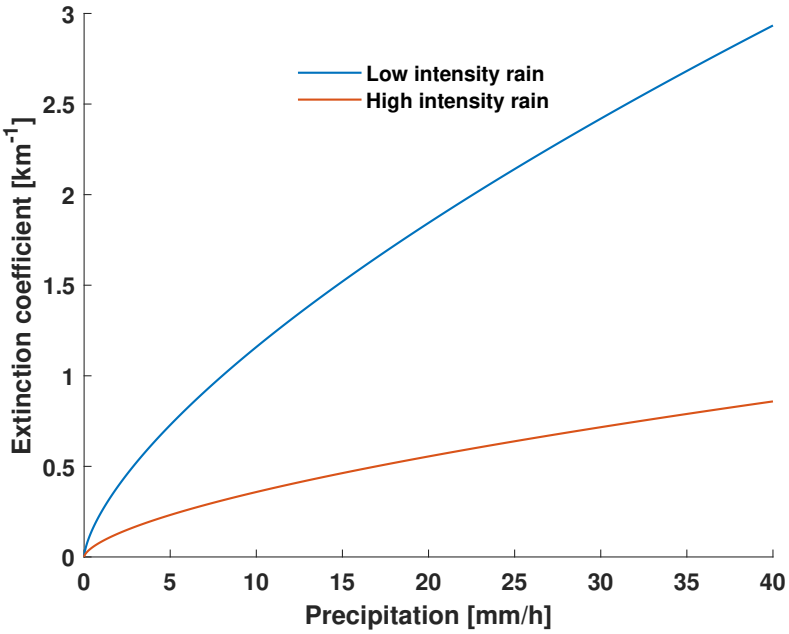

Figure 12 – Figure shows the expected extinction coefficient based on the Carbonneu model (28) for attenuation due to precipitation.

As is shown in the above figure, the attenuation is expected to larger for lower intensity rain. It is postulated that the main reason is that in low intensity rain, the rain is closer to the conditions of fog for the same precipitation, which would scatter and absorb the radiation more steadily.

2.2.4 Visibility estimation

Here, a method of estimating visibility is proposed based on how the amplitude of the signal is attenuated. That is, if visibility can be established solely based on a LiDAR signal. Starting from LiDAR equation (14), it is assumed that the only parameter varying is the extinction coefficient, while other parameters such as distance, detector functions and reflectivity of the target are not changing as shown in figure 13. Consider that the calibration β_{cal} is known, and that a measurement is done producing a result of P_{cal}. Changing the extinction to β and measuring P, the ratio of P_{cal} and P is formed and is given by

$$\frac{P_{cal}}{P} = e^{-2s(\beta_{cal}-\beta)}. \tag{29}$$

Solving for β in equation (29) yields

$$\beta = \beta_{cal} + \frac{1}{2s}\ln\left(\frac{P_{cal}}{P}\right). \tag{30}$$

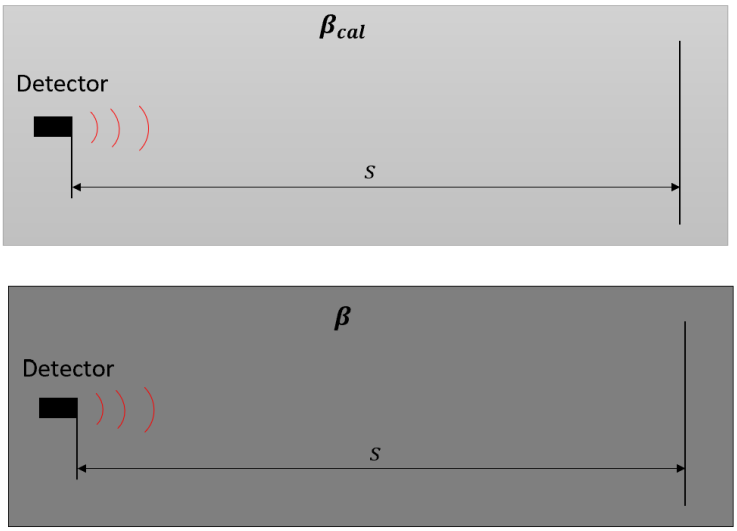

Figure 13 – Figure illustrates the method for estimating visibility. Once a calibration is done it is possible to estimate the extinction in a different conditions using equation (30).

One might recall power is total intensity spread over an area. If the detectors have the same measuring surface, the intensity may be substituted for power.
For visibility, we might use any suitable models such as Al Naboulsi fog models for fog described by equations (26) and (27). Also Kim-Kruse visibility model can be used in (25), however Naboulsi models are computationaly simpler since for the the visibility can be explicitly calculated. For the Kim-Kruse model the visibility needs to be solved for numerically.
The performance of this method is expected to vary at different distances, with better results at larger distances. In figure 14, the normalised amplitudes calculated from the Kim-Kruse model are shown.
From the point of view of the curve for 15 meters, it is harder to distinguish visibilities over 4 km since they all have about the same amplitude. When random variation is introduced, some amplitudes with higher visibility might have smaller amplitudes.
As can be seen in figure 14, increasing the distance at which the measurement is performed spreads out the amplitude making it easier to distinguish visibilities as well as being less sensitive to random variations.

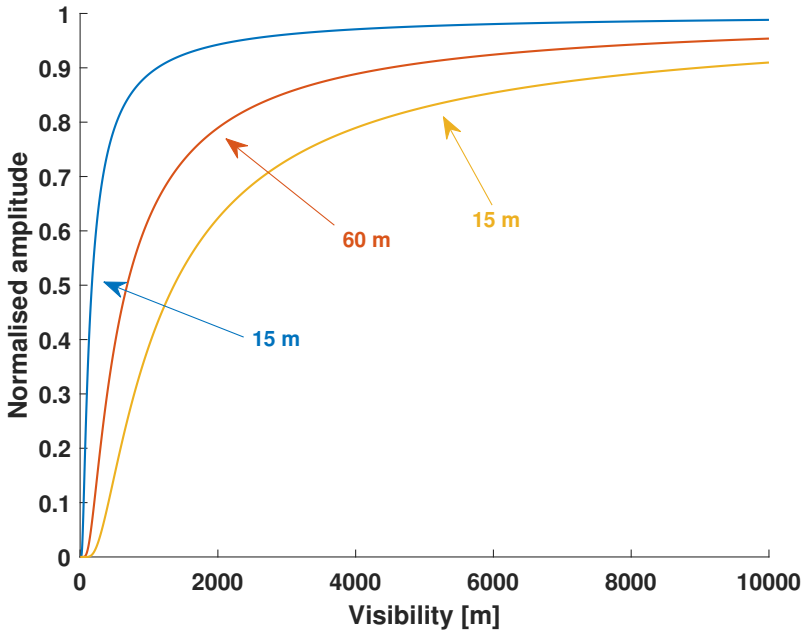

Figure 14 – Figure shows the *normalised amplitude* as a function of visibility as predicted by the LiDAR equation (14) and Kim-Kruse model (25), for different distances to target.

Yet another important remark about this method is that, when considering taking the ratio P_{cal}/P is that the detector might be very nonlinear when measuring power(or intensity). The actual signal is a voltage, so if equation (30) is to be used by taking the ratio of amplitudes, one must either make an approximation that in the interval of amplitudes that are studied, the relationship between measuring voltage and intensity is approximately linear, or if a calibration curve is known in advance.

2.3 PCA

In this study, some special data analysis techniques and machine learning methods will be applied. The rest of this chapter goes through some fundamentals of those methods, starting with principal component analysis. *Principal Component Analysis*, commonly referred to as PCA, is a useful technique which is used in many areas of science, engineering and other fields where gathering and extracting patterns from data is necessary. Often times, measurements and data is not presented in its most useful form where possible relationships are hard to detect. PCA is a method of transforming data into a more useful form, from which some pattern may be easily expressed[22].
Here we introduce the theoretical framework for principal component analysis.

2.3.1 Data matrix

In statistics and data analysis, data is usually represented using vectors and matrices, where each component of the data vectors represent a variable or distinct feature of the object being measured. These are expressed as column vectors $\mathbf{x} = [x_1, x_2, x_3, ..., x_p]^\top$. Now, in common use of PCA, each component in the above vector corresponds to a feature that is meaningful and usually represents something in the real world. These could be for example height, weight and age of a human being. However if the PCA is to be used on a time signal, in particular a LiDAR signal, the components then represent the detector signal at a time when the component was measured.
Taking many measurements, say N of them, we can represent with a single data matrix $X_{N\times p} = [\mathbf{x}_1^\top, \mathbf{x}_2^\top, \mathbf{x}_3^\top, ..., \mathbf{x}_p^\top]^\top$. The data matrix X is an N by p matrix where the rows correspond to single measurements, and columns represent the variables, or in the case of a signal, the measurement at each time step.

2.3.2 Data transformation

As stated earlier, the goal of PCA is to transform the data into a more useful form. The type of transformation that PCA performs is called *linear transformation*[22].
Consider a data vector $\mathbf{x}$. We want to create a new variable y that is a linear combination of the components in $\mathbf{x}$, which can be written as

$$y = w_1x_1 + w_2x_2 + w_3x_3 + ... + w_px_p \tag{31}$$

The linear transformation in equation (31) can be seen as projecting the vector $\mathbf{x}$ on to an axis specified in the direction of $\mathbf{w} = [w_1, w_2, w_3, ..., w_p]^\top$. The we can write the new variable as $y = \mathbf{w}^\top\mathbf{x}$. The variable y is often times referred to as *pca score*, or simply as score, while the vector $\mathbf{w}$ is referred to as *pca vector*. In general we have p orthogonal directions to project on to, giving p pca vectors. Collecting all the PCA vectors into the matrix $W_{p\times p}$, we get the full transformed data vector as

$$\mathbf{y} = W\mathbf{x}. \tag{32}$$

The method of PCA then concerns with the calculation of the transformation matrix W in equation (32). The PCA vectors are chosen such that they are all mutually orthogonal, and such that the projection has the largest possible *variance*[22]. The final result is that the pca vectors can be obtained by *single value decomposition* of the covariance matrix of the matrix $X_{N\times p}$. The PCA vectors have the property that $\mathbf{w}_i \cdot \mathbf{w}_j = \delta_{ij}$ where δ_{ij} is the Kronecker delta function[22].

2.4 Discriminant Analysis

Discriminant analysis is a statistical approach used for classification of different groups. The decision for the classification is based on probability[22][23]. Consider variable G that describes k classes, and $\mathbf{x}$, which is a feature vector. The variable G is the variable describing each class and $1 < G < k$. We classify a feature vector belonging to the class

with the largest posterior probability $\Pr(G|\mathbf{x})$. Let also the prior probabilities of class i be $\pi_i = \Pr(G = i)$, with the condition that the sum of all π_i; $i = 1, 2, 3, ..., k$ is equal to 1. Also, the class-conditional density of $\mathbf{x}$ in class i is given by $f_i(\mathbf{x})$.
The posterior probability is then given by

$$\Pr(G = i|\mathbf{x} = \mathbf{x}_0) = \frac{f_i(\mathbf{x}_0)}{\sum_{j=1}^{k} f_j(\mathbf{x}_0)\pi_j} \tag{33}$$

Equation (33) represent the probability that $G = i$ given that the feature vector $\mathbf{x}_0$ was measured. Therefore, that vector is classified to the class for which the probability is highest. An assumption is made that the conditional probabilities are normally distributed with each class having independent mean vectors and covariance matrices[22][23].
If the covariance matricies of each class are all equal, the decision boundary will be linear. Therefore this method of classification is referred to as *Linear Discriminant Analysis*, or LDA. If no assumption is made about the covariance matricies, the method is called *Quadratic Discriminant Analysis*, or QDA[22].

3 Method

3.1 Experimantal setup

In this study, LiDAR measurements were made along with a weather station at Veoneer's facility in Skellefteå. The measurement system was firstly set up by Marcus Hedlund[1] for his initial Master's thesis in 2020 where a long range LiDAR was used. Additional modifications were made by Johan Viklund for his Master's thesis in 2021 where a short range LiDAR was added. In this study, the data gathered from the short range LiDAR is used. The measurement devices used are

1. Weather station
   (a) Vaisala PWD22 Weather and visibility sensor
   (b) Vaisala PWL111 Luminance sensor
2. Short range LiDAR Torunn

The weather station is used to record the environmental conditions of surroundings. In it is included measurements of visibility, precipitation and temperature as well as classifying the conditions according to codes given by WMO table 4680, METAR table 4678 and NWS table. The station measures the the luminance of the surroundings as well as visibility from for up to 20 km. Other important measures gained by the weather station is temperature(-40 °C to 70 °C), precipitation(0 mm/h to 99 mm/h)
The short range LiDAR is based on sending a light pulse with a wavelength of 905 nm, and a pulse width of 2.5 ns.

Figure 15 – Figure shows measurement system setup in Skellefteå by Marcus Hedlund which includes the LiDAR unit and the weather station[1][2].

3.1.1 RISE facility

To be able to perform analysis about particle distributions and how they affect LiDAR performance in a controlled manner, a climate chamber was used in collaboration with RISE in Borås. There artificial fog was generated varying parameters such as liquid water content and particle distribution. The distribution parameters such as liquid water content and particle size were measured in conjunction with a short range LiDAR, and measurements were made once a minute for various periods while conditions were varied. Distribution parameters were measured using a spectrometer and shining a lamp with a spectrum of wavelengths. The disturbance of the lamp on the LiDAR was determined to not affect the signal. For the LiDAR, a wooden plane target was set at approximately 6 meters away. The entire chamber had the approximate dimensions of $20 \times 4 \times 3$ meters with a volume of 240 m^3. The general setup is shown in figure 16.

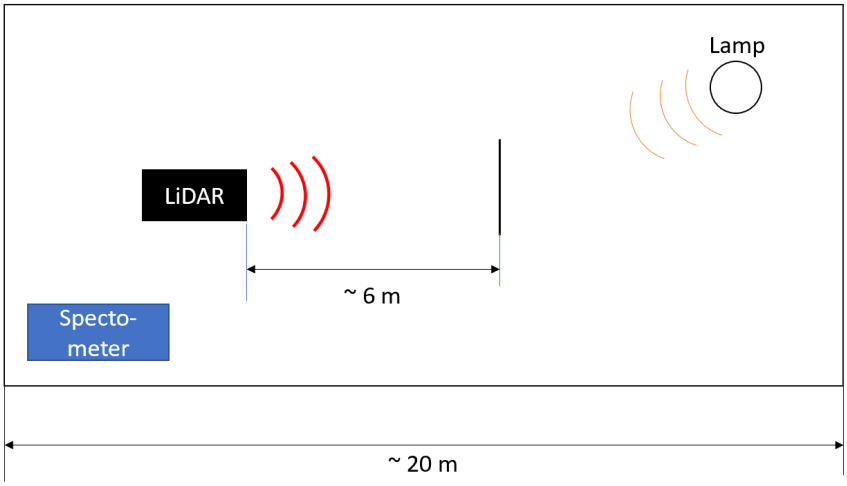

Figure 16 – Figure shows the general setup during experiments at RISE where both the LiDAR and the spectrometer were used to gather data about droplet distribution.

During the measurement procedure, the vaporised water formed fog and was inserted into the chamber for a minute until it filled the entire volume, simulating the conditions of fog. During this time, measurements on the LiDAR were taken every minute where 15 frames were taken so that an average signal could be created. The measurements were taken until the conditions returned to nominal values. The physical setup is seen in figure 17.

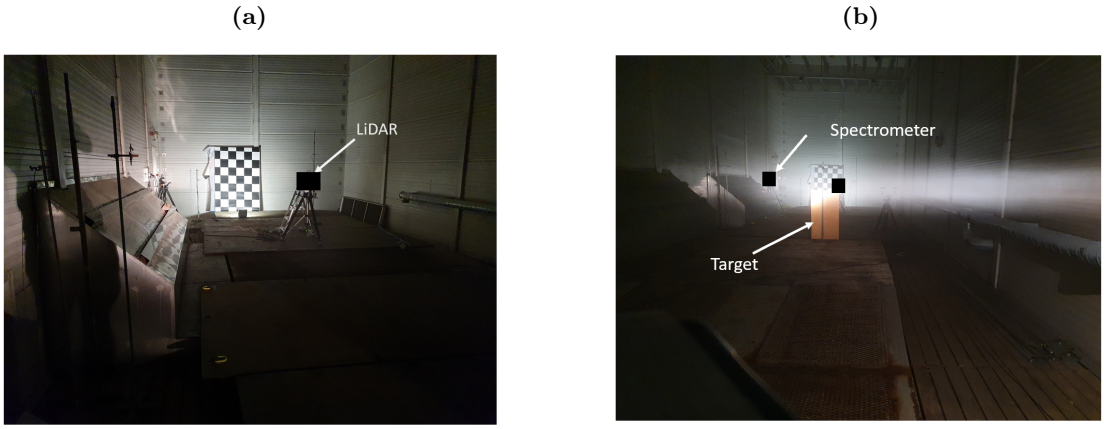

Figure 17 – Figure shows the setup used during the experiment where the LiDAR and spectrometer. In a) the left view of the experiment setup shows the LiDAR unit and b) shows the location of the target and spectrometer.

As was mentioned previously, the distribution parameters were measured in conjuncion with the LiDAR measurements, and they were done so independently. Before starting with measurements, the time was firstly synchronised between the instruments so that could later be compared at the same time. The distribution parameters were measured once every couple of minutes, while the LiDAR took 15 frames once or twice a minute on average.

3.2 Data analysis

Here we discuss the analysis principals applied to the gathered data from measurement. Firstly, to determine the amplitude, the location of the peak must be found. Consider a general signal amplitude as shown in figure 18.

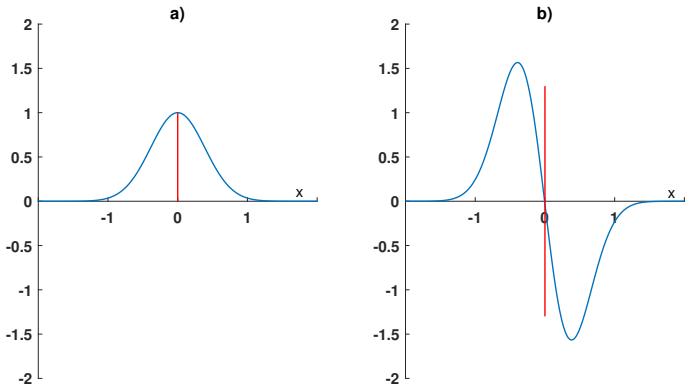

Figure 18 – Figure shows a) a generic function and b) its derivative which is used for peak detection by locating the position at which the derivative crosses zero.

To find the amplitude of the peak, we can look at where the derivative is zero and evaluate the signal at that point. This is method that may also be used to determine the distance to object since in the case of LiDAR, and in general other ToF techniques using pulsed lasers, the peak location will represent time for the light to go from the laser to the object and back to the detector. Knowing the time delay enables one to determine the distance to objects. Although looking at where the derivative is zero is a reasonable method in determining the distance to objects, Marcus Hedlund suggest that for such purpose, the time at which the second derivative crosses zero is a better method[1]. The method of finding where the derivative is zero may easily be implemented manually, however MATLABs native function *findpeaks* is a useful implementation of this method. Especially since a signal generally will have many local maxima and minima produced by noise. In that case the MATLAB native function in particulary useful since parameters such as pulse width and prominence may be put for which the unwanted local minima and maxima may be ignored.
The measurements from the short range LiDAR were stored in a database in conjunction with the data gathered by the weather station. The data were then compared for the same times in order to get predicted amplitudes from theoretical models.
For application of principal component analysis and predictive models, the data was organised into a data matrix where each row represented a signal. The MATLAB native function *pca* was used to extract the principal component scores. In order to train the discriminant models the functions *lda* and *qda* were used which are a part of the machine learning package provided by MathWorks. In order to train the neural network, the deep learning package was used and the methodology of how it is applied is provided at Mathworks documentation site[24]. For a review of the workings of neural networks, see appendix A.

4 Results

This chapter includes the results obtained relevant to the study, starting with the discussion about the properties of the measurement system and the data gathered from it and how it is used for further analysis. The analysis is compared to theoretical models discussed in previous chapters. Firstly from the data gathered from the year-round measurement made in Skellefteå, and later the measurement done in the RISE fascility in Borås.

4.1 Measurement system

As was discussed earlier, the measurement system includes a short range LiDAR sending a light pulse of width of 2.5 ns at a target 18 meter away. The detector samples with a sampling frequency of 3 GHz. The signal returns to nominal level after about 6 meters, making it appropriate to place a target att least 6 meters away. A typical signal is shown in figure 19. As was discussed in Johan Viklunds and Marcus Hedlund's theses[1][2], in order to remove random noise from the data, 15 frames were taken and averaged over to get an average signal. This meant a reduction of random noise present in the singular measurements. This has two primary effect. Firstly it lowers the Signal-to-Noise ratio making it easier to assign causes to attributes of the signal to weather conditions. Secondly, it makes for easier peak detection.

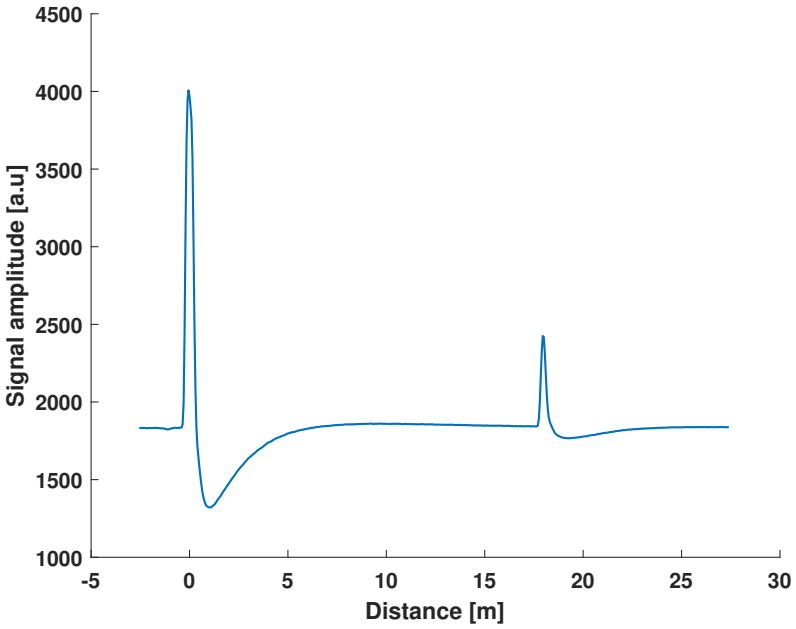

Figure 19 – Figure shows a typical signal generated from the LiDAR on a target 18 m away. The initial peak is the zero-pulse generated by the internal reflections in the LiDAR, while the second peak is the return pulse from the target.

For peak detection, it is possible to differentiate the signal and solve for zero crossing, certainly occurs at the peaks. For this study, the MATLAB function *findpeaks* in the *Signal Processing toolbox* is used for peak detection, which with a prominence parameters finds all peaks with a minimum peak prominence of 50, which is certainly the case in most signals as shown in figure 19.

4.1.1 Temperature dependence

As in many measurement systems, the the measured signals are dependent on both the systems temperature and the ambient temperature. In order for consistent comparison with conditions of varying temperature, the signals should not be affected by temperature, or a temperature compensating factor should be determined. For the LiDAR unit investigated by Johan Viklund[2], the signal amplitude was heavily dependent on ambient temperature.
To determine a potential temperature effect, measurements made in clear weather defined as visibility over 20 km are used, along with minimal luminance of 4 cd/m^2 from the surroundings.
Furthermore, the temperature of the components in the measurement system may affect the signal amplitude. Similarly, measurements in visibility over 20 km and surrounding luminance of 4 cd/m^2 were used to determine a potential temperature effect.
In this regard, the signal amplitude does drop significantly for both detector and laser temperature as shown in figure 20.

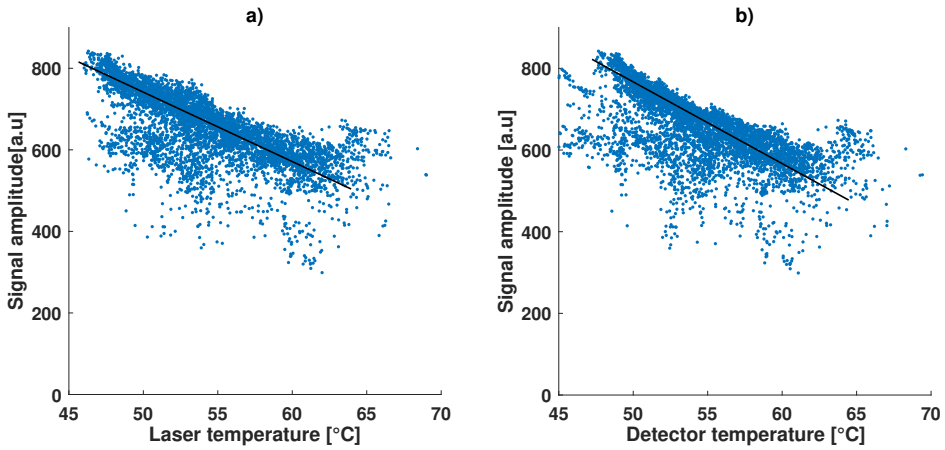

Figure 20 – Figure shows amplitude dependence on temperature in low luminance, high visibility environment for both the a) laser and b) detector temperature. Lines represent linear regressions done on the data.

The laser and detector temperature are highly correlated(sample correlation of 0.97 and a slope not significantly different from 1). If temperature compensation is needed, only one of the temperatures are needed, and here the laser temperature is used.
As is seen in figure 20, the signal amplitude as a function temperature in the operating

region is adequately modelled by a linear model. There are however many outliers that seemingly lie significantly from the regression lines. When creating the compensation factor, these outliers were excluded.

4.1.2 Angle dependency

Yet another factor that may impact the signal amplitude is the incident angle sent from the LiDAR. The figure 21 shows a point cloud create by the LiDAR at the targets location 18 m away, which has a bright and a dark portion.

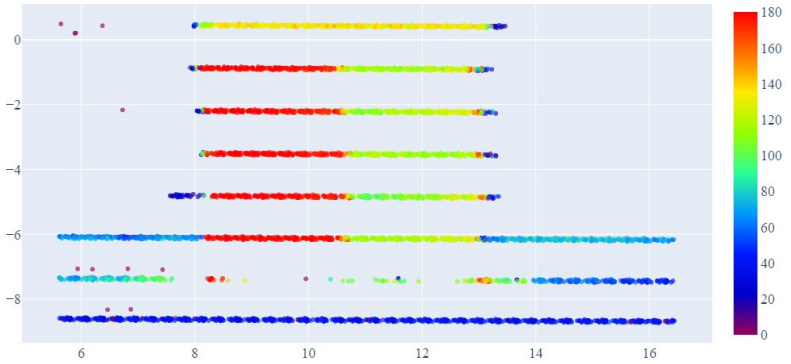

Figure 21 – Figure shows point cloud and signal strength generated around the target located at a distance of 18 m.

At the targets edges the amplitude of the signal is drastically decreased, which may correspond to the outliers in figure 20. The angular dependency in analysed and shown in figure 22. The changing angle in both azimuthal and polar direction does not affect the amplitude in a predictable manner, where the amplitude varies uniformly across all angles. This means that the outliers in figure 20 are not explained by angular dependency

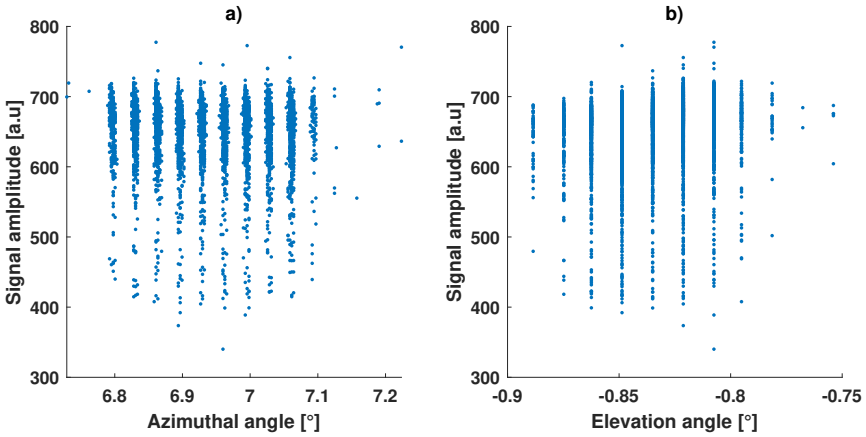

Figure 22 – Figures show how the signal amplitude varies as a function of a) azimuthal and b) elevation angles.

4.2 Amplitude variation with visibility

Here the general results of measurements done in Skellefteå on the short range LiDAR are presented. Firstly, we examine the pattern exhibited by the various weather types detected by the weather station and how the signal amplitude is determined. The amplitude is normalised so that the normalisation factor is the average amplitude in clear weather with visibility of 20 km or more. The amplitudes has also been compensated by a factor dependent on the laser(or detector) temperature.
The results of the analysis is shown in figure 23.

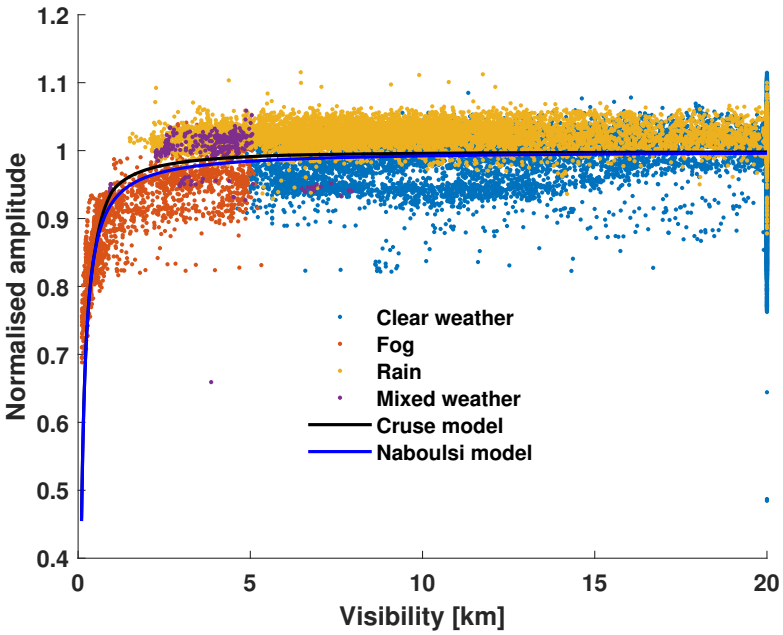

Figure 23 – Figure shows the data gathered by the measurement system for the short range LiDAR and with weather station in various weather conditions along with expected attenuation predicted from the LiDAR equation (14), the Kim-Kruse model (25) and Naboulsi models(11).

As can be seen in figure 23, the amplitude of the signals generally follow the expected behaviour predicted by the LiDAR equation (14) combined with the empirical models based on visibility described by equations (25), (26) and (27).
One noteworthy remark is that in rainy weather, the trend for the amplitude is slightly larger that other types of weather types, which is to says that the amplitudes in rainy conditions are not centred around the predicted models. Although the exact cause of this discrepancy remains unknown, a hypothesis may be that rainy conditions change the reflectivity of the target, hence increasing it. A proposition also made by Marcus Hedlund in his thesis[1]
Another remark is that the maximum visibility that could be measured is 20 km, hence

producing a seemingly vertical line. This is solely due to the fact that the amplitude varies even for higher visibilities for which can not be measured.
Overall, the measurement do follow the predictions from the visibility models with a moderately amount of variation, and a sligtley smaller variation for really low visibility's of less than 200 m.

4.3 Amplitude variation with precipitation

As with the visibility models, the precipitation model described by equation (12) may be used to predict the behaviour of LiDAR signals in rainy conditions, for both tropical high intensity rain and low intensity rain. The result of the analysis is shown in figure 24.

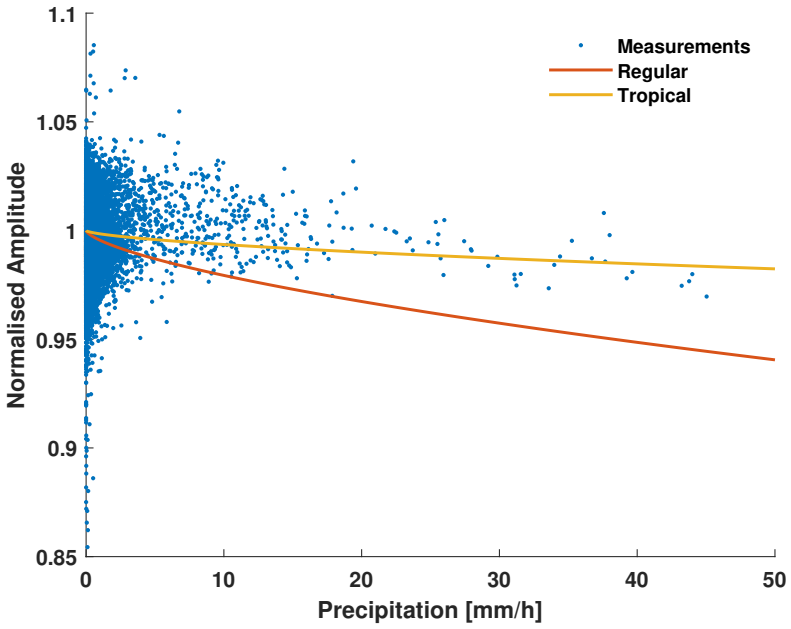

Figure 24 – Figure shows the data gathered in rain along with the two Carbonnoue models (28) for low intensity and high intensity(tropical) rain.

From the results shown in the figure above, it leads to draw the conclusion that the models do not have any predictive power when it comes to precipitation of rain. A possible better predictor might instead be the size of the droplets, together with precipitation.
From the measurements, it can be seen that the amplitude varies drastically even at very small precipitation. A noteworthy remark here, with consideration of application of LiDAR for distance detection. The amplitude does not vary much more than 5 % from the nominal peak(normalised at 1), even at very high precipitation rates of over 40 mm/h, for a target 18 meters away. A much better predictor for the attenuation in rain might be visibility as is shown in figure 23.

4.4 Visibility estimation

A possible method of estimating visibility with the help of models was proposed in previous sections, and the results of the method is analysed here. The validity of the empirical models of Naboulsi and Kim-Kruse were shown to be adequate in predicting signal amplitude as is shown in figure 23. There was also was great variation around the models, especially for larger visibilities bigger than about 400 meters. Already from that we can expect that estimating visibility from the models will have large variation around the model. For that reason, only visibility in fog is considered. There (30) is applied with a calibration amplitude taken to be the same normalisation factor that produced figure 23, and a calibration extinction coefficient of $1.96 \cdot 10^{-4}$ m^{-1}, which is calculated from equation (25).
To convert to estimated visibility, the Naboulsi model in equation (26) was used since solving for visibility does not require numerical methods. Figure 25 shows the result of the analysis.

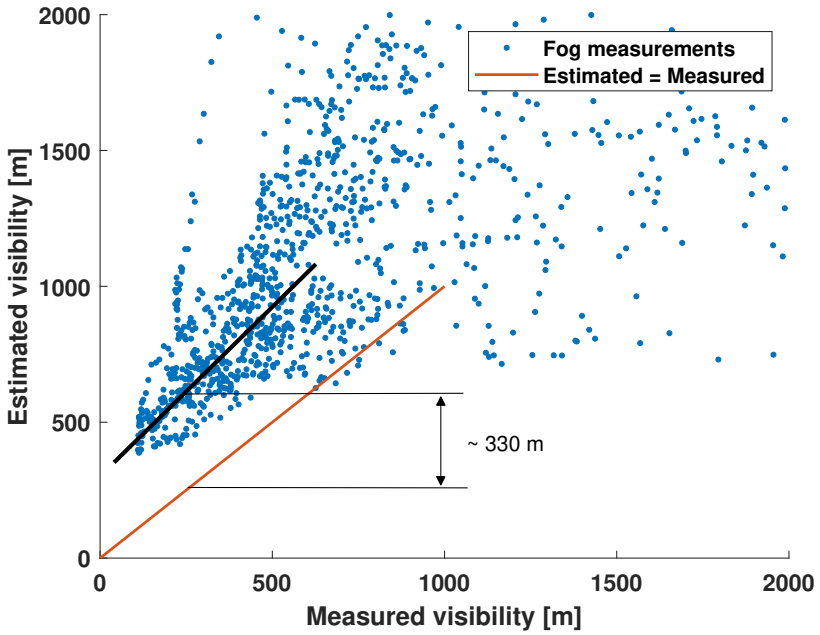

Figure 25 – Figure shows the comparison between the estimated visibility based on equation (30) and using the Naboulsi model (27). Linear regression is done on the data with measured visibility below 500 m.

A regression was made on measurements under 500 meters measured visibility which produced a slope of 1.04 and a bias of 330 m. The reason only points below 500 meter visibility were considered is since the increasing trend disappears and the variation in estimated variability grows unbound.
As the figure 13 suggest, equation (30) predicts that the slope of the created line is 1.

However there seems to be a systematic bias of around 330 meters.
Another remark is than even though that the regression line has a bias of around 330 meters, the measurements vary a bit around the regression line($\sim$ 150 m), meaning even if the bias was removed, the uncertainty in estimated visibility would then be about 150 m, and growing for increasing visibilities, making this a somewhat uncertain method of producing visibility estimates.
Therefore the application of this method in the automotive industry is nonexistent. Especially since there, the distance of the target is varying, the reflectance of the target is not known. Another big issue is to determine how the calibration should be done. In this study the calibration was done on measurements of above 20 km visibility and the Kim-Kruse model was used to get the calibrated extinction coefficient. There is however no evidence that it is the best choice of calibration or even a remotely a good one. Another reason is that the method is not expected to work well for short distances as discussed in previous sections.
However if the reflectance of the target is kept the same, and the distance of the target is larger with a suitable calibration, the method is expected to produce a more adequate results.
Another approach to possibly estimate visibility was made using machine learning method of LSTM neural networks[25] where it was trained on the measured signals and the training outputs were measured visibilities. The output of the neural network were then the estimated visibility, and the mean squared error was minimised.
That study produced a dis-satisfactory result that the neural networks output was always the average visibility(7 km) of all measured samples regardless of what the input signal was. This suggested that using raw signals was not an adequate method to determine weather conditions based on LiDAR signals, and that some transformation method like principal component analysis might transform the signal into more suitable data to predict weather conditions.

4.5 PCA

In this chapter, we discuss the results obtained by applying PCA principals on the data collected. The signals were 600 samples long, which creates 600 PCA eigenvectors defining the new directions to project the data.
As was discussed in the theory section, the PCA eigenvectors are ordered according to decreasing variance. This way, the first few PCA components called scores might be used as feature extractions.
As was discussed in the previous chapter, in the autonomous vehicle industry, the target distance is varying and the reflectance is not known. Therefore, two studies were done where PCA was applied. One where the whole signal including the return pulse of the target was kept. This is so the general predictive power is established.
After that the signal was cut before the target response so that a more specific performance could be established more suitable in applications. The signals with return pulse included are expected to give more satisfactory results since great information about the conditions are contained in it.

4.5.1 PCA on complete signal

Here the result for applying PCA on entire signals including the return pulse are presented.
In PCA, there is always a choice at which components one might find useful. Based on theory, the first few are expected to hold most information, but there is no guarantee that the various classes are appropriately separated in those components. Therefore, a manual visual inspection in needed to determine suitable components. One does not however need for suitable components which are of higher order since they are almost for certain not suitable as features since they correspond to random noise. For the first few components one might look for within the first 15 or less for suitable separation.
In figure 26, the 3rd, 4th and 10th components are plotted for the classes of mixed weather, clear weather, rain and fog.

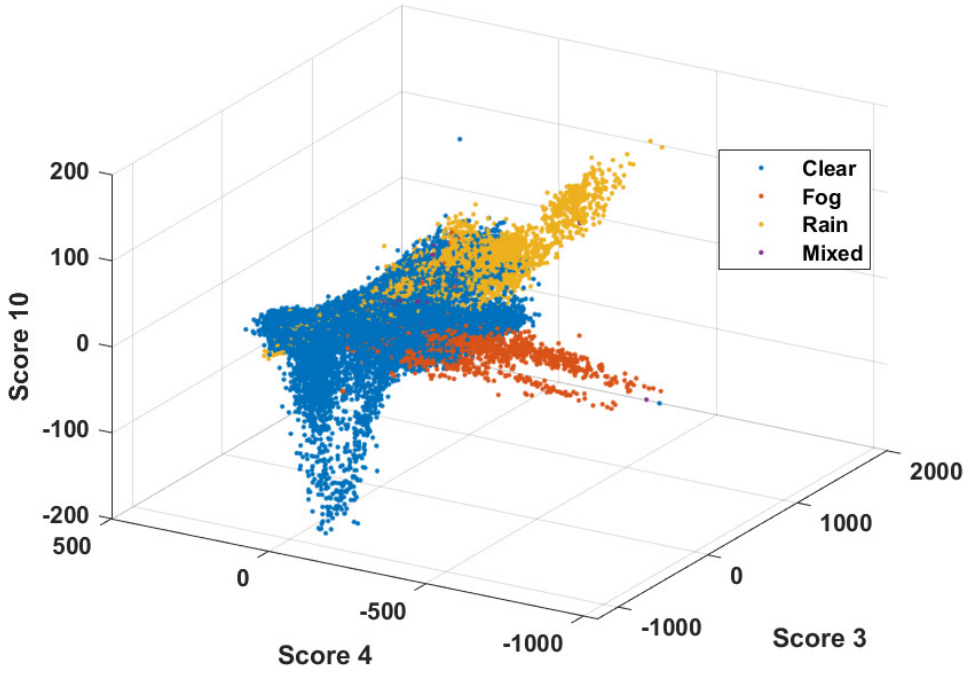

Figure 26 – Figure shows the results for applying PCA on entire signals and plotting the 3rd, 4th and 10th principal component scores.

We can see from figure 26 that some separation of the different types occurs, even though there is some overlap. Even though there is a certain degree of overlap, we can see that there are spaces where the different types are certain, suggesting that some prediction about the types may be made by using some models. Additionally, we can see that some components are suited in separating one class from the other. There are also cases when a component is not suitable in providing any information about the type. An example

of this is that the 4th principal component is useful in separating fog from other types, as is indicated in figure 27.

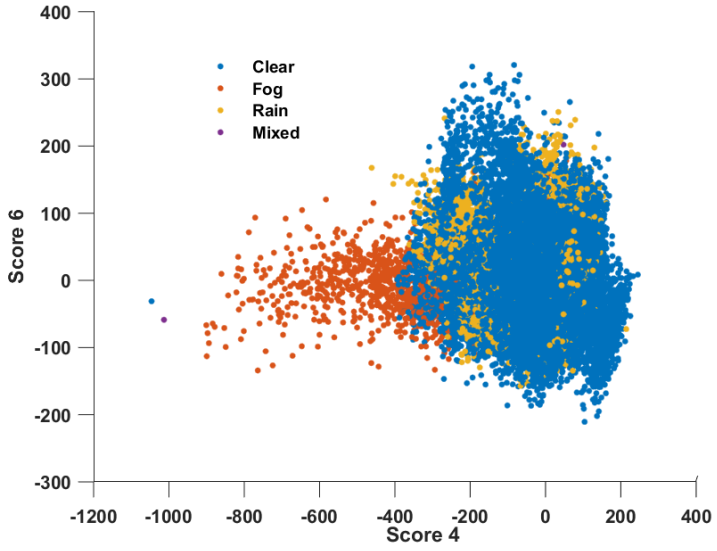

Figure 27 – Figure shows the scores which separate mostly rain from the rest of the classes moderatly well.

We can see in figure 27, that in the 4th component, fog and clear weather(and mixed weather) are lumped together. If we were only given the 4th component, there would be not way to differentiate between fog and clear weather. However is does separate rain from the other types. For example a measurement might have a value of -600 for the 4th pca component. This would almost certainly indicate that the measurement was taken in rainy conditions.
On the other hand we can study the predictive power of the 6th pca component. All the points can be projected onto the 6th component and conclude that all the classes would be spread over the same interval of about -200 to 200, making it impossible to make predictions about new measurements. Although we can say that a measurement yielding the 6th component of say 150 is most likely not during rainy conditions.

4.5.2 PCA on signal with removed return pulse

As it was mentioned previously, when it comes to usefulness in applications to autonomous vehicles, or prediction weather conditions in dynamic situations where the distance to a target is changing with time, the reflectance of objects are not known along with ignorace about other possible dynamical factors, applying PCA on that type of signals that contains the return pulse might not prove to be any useful. Although in areas where conditions are held static and measurement parameters are known, the results from the previous section may be more applicable. It may be more useful in fields such

as meteorology.
For the reason mentioned above, the signal was cut efter a distance of 15 meters, which removed everything after 15 m including the return pulse at 18 meters.
Applying the method of PCA produced the result shown in figure 28, in which only the 6th, 7th and 8th PCA scores are shown.

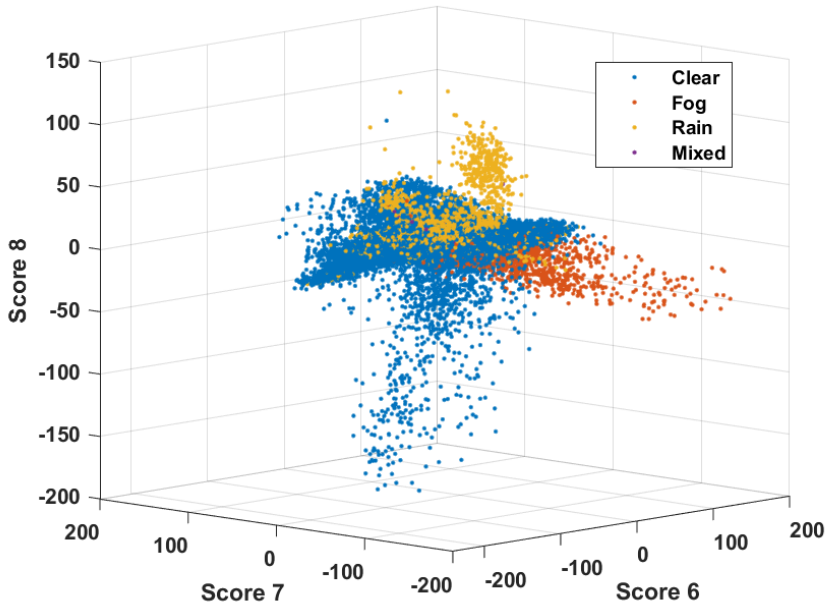

Figure 28 – Figure shows the result of applying PCA on signals with removed return pulse with the 6th, 7th and 8th scores plotted.

We can indeed as in previous section that there seems to be a certain degree of separation between the different types, although just by visual inspection it is hard to see the degree of separation, or if how many of the points in each type are outside the common cluster. Just by inspection there seems to be less points in the class of fog and rain that are outside of the common cluster. This suggests that if trained for prediction, the data without the peaks may produce somewhat worse results.
This may be more easily seen when shown with two pca scores. In figure 29, the 6th and 8th scores are arbitrarily chosen and shown. Here both rain and fog are shown a degree of separation from clear weather, although not as much as rain in figure 27.

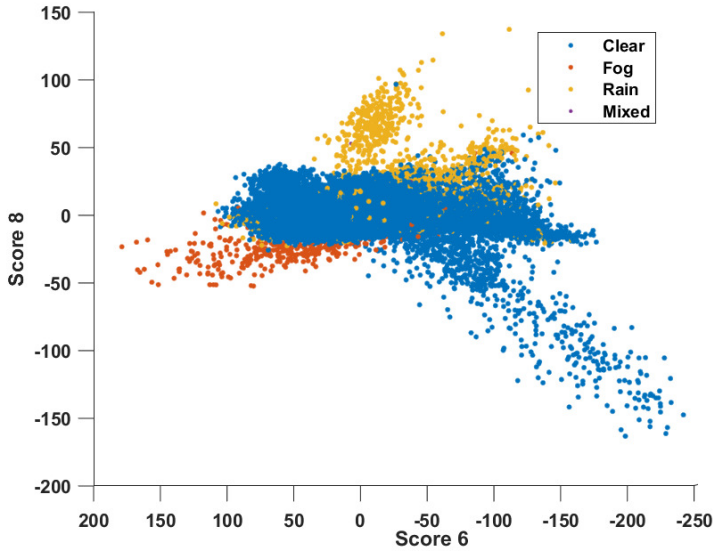

Figure 29 – Figure shows the result of applying PCA on signals with removed indicating some possible separation between the classes for PCA scores 6 and 8.

4.6 Predictive models

In this section, the results are presented when the various machine learning schemes are applied to signals transformed by PCA, with and without the return pulse. As was discussed in the previous chapter, there are different ways of measuring performance of predictive algorithms when it comes to classification. Since the data that is used is randomly selected into a training set consisting of 80 % of the data, and 20 % used for for testing, the results will vary to some degree each time the algorithms are applied. Therefore measures that are used to measure performance are random variables. To get average performance, we train a number of algorithm of the same type and we average the results. The algorithms used in this study are discriminant analysis and "regular" vanilla neural network. An important remark is that the mixed type as seen in figures 26 and 28 has been neglected since not many data points are obtained of that type. A typical result is summarised in the confusion matrix shown in figure 30.

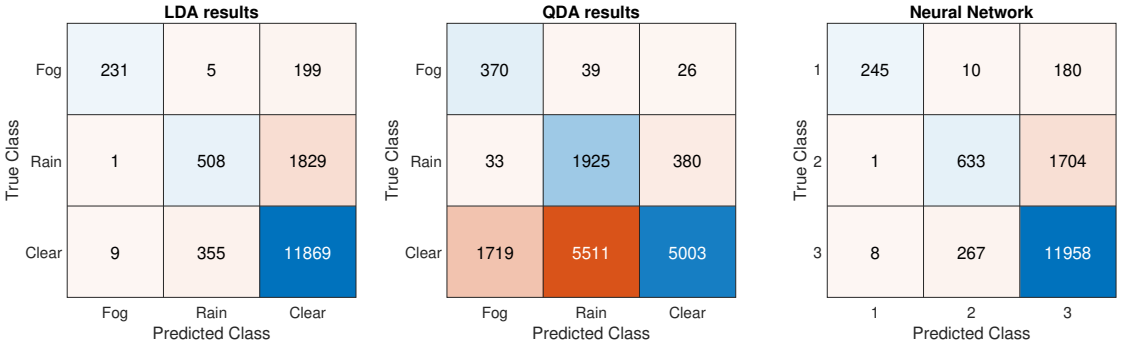

Figure 30 – Figure shows the a typical confusion matrix for trained models with the 20 first PCA scores created from entire measured signals.

From the confusion matrix, the relevant performance measures may be estimated. Accuracy is a common way to measure performance, and is estimated by the fraction of correct classifications. In the case shown above, the methods produce a somewhat high accuracy(above 80 % with exception of QDA with about 50 %). This is however not a good way to measure the predictive power in this case. This is due to the fact that the data set consist of a large majority of the same type, namely clear weather type. The model of guessing clear weather type regardless of the input signal will produce a higher accuracy just due to the fact that most of the data points are of the clear weather type. Arguably another reason for not using accuracy is that we are not interested in predicting clear weather. Instead the predictive power of adverse conditions is of more interest such as rain and fog.
For that purpose, more suitable measure of performance is sensitivity S and precision P. For sensitivity we measure how accurate the models are when predicting a certain class. That may be estimated by taking the fraction of predictions within a class that the model classified correctly. From the confusion matrix it is calculated by the fraction of correct classifications within a column.
Precision is a measure of accuracy of all data points within a class(as opposed to accuracy of predictions as in sensitivity). From the confusion matrix it may be estimated by calculating the accuracy of a type within a row. The general results for fog and rain are shown in tables 1.

Table 1 – Tables shows the estimated average sensitivity and precision for fog and rain when training 100 models on the first 20 principal component scores from signals with return pulse kept intact for a) fog and b) rain.

(a)

	$S(\%)$	$P(\%)$
LDA	96 ± 2	52 ± 5
QDA	20 ± 4	85 ± 4
Neural Network	97 ± 1	59 ± 2

(b)

	$S(\%)$	$P(\%)$
LDA	61 ± 3	22 ± 6
QDA	26 ± 5	82 ± 3
Neural Network	70 ± 2	27 ± 2

From tables 1, we can see that LDA and neural networks generally have a high sensitivity fog(> 90 %) and less for rain(> 60 %), although neural network has slightly lower variability. In contrast, the QDA method has substantially lower sensitivity which also varies slightly more than the other methods.
In regards to precision, the situation seems to be reverse for both fog and rain. The precision for LDA and neural networks are generally lower than for QDA. There seems to be no obvious apparent reason why the reverse metrics are good for QDA.
The results stated were for entire signals with the return pulse kept in the signal. The same analysis was done for signals which were cut just before the return pulse so that a comparison can be made and perhaps a possibility of application established. A typical confusion matrix is shown in figure 31

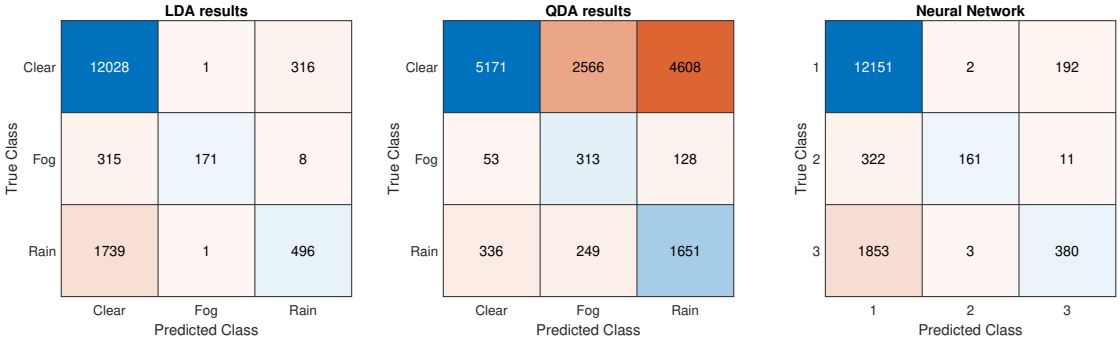

Figure 31 – Figure shows the a typical confusion matrix for trained models with the 20 first PCA scores created from signals with removed return pulse.

From the confusion matrix such as in figure 31, the performance metrics are estimated. As previously general accuracy is not a good metric since a substantial majority of the data is of the same type, and sensitivity and precision are a more appropriate measure of performance. The results of the analysis is shown in table 2

Table 2 – Tables shows the estimated average sensitivity and precision for fog and rain when training 100 models on 20 principal component scores from signals with return pulse being removed for a) fog and b) rain.

(a)

	$S(\%)$	$P(\%)$
LDA	94 ± 1	37 ± 3
QDA	13 ± 3	64 ± 4
Neural Network	96 ± 2	41 ± 4

(b)

	$S(\%)$	$P(\%)$
LDA	60 ± 4	21 ± 3
QDA	25 ± 3	69 ± 6
Neural Network	65 ± 5	20 ± 2

The results shown in table 2 show that some metrics were reduced when removing the return pulse, mainly the precision for both fog and rain. The sensitivity of LDA and neural networks were not greatly affected by the removal of the return pulse for fog, while for QDA the sensitivity for fog dropped to 13 %.

The precision in fog was greatly reduced for all models, with QDA still having a largest of 64 %, a drop from 85 %. For rain the results are a little varied. The LDA and QDA model produced a similar sensitivity as before, while for neural network it was reduced by 5 percent points.
In summary, removing the return pulse reduced the performance of the models, in particular the precision was most affected metric.
The results shown in this section is when the algorithms were applied to the first 20 PCA scores. We can see how choosing different number of PCA scores affect the performance by estimating metrics while varying the number of PCA components. In figure 32 the sensitivity of the entire signal. a

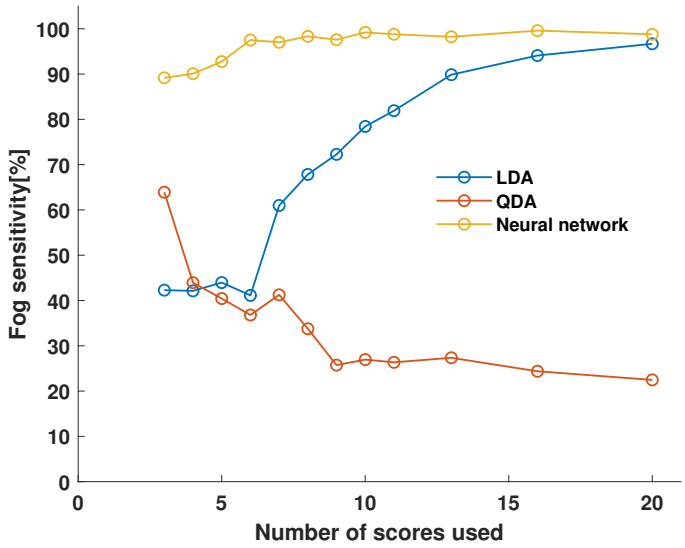

Figure 32 – Figure shows the sensitivity in fog for the models used as the number of PCA scores are increased.

Figure 32 shows how the sensitivity of predicting fog is affected by the number of pca scores included in the training the models. For the neural network, the sensitivity with only 3 scores is already fairly high at about 95 percent and steadily increasing with if more pca scores included. For all models the sensivity stabilises at about 20 PCA scores. Examining precision instead of sensitivity yield similar results where the precision of LDA and neural networks increases and for QDA increases and stabilises at 20 PCA scores.

4.7 LiDAR performance in artificial conditions

Here the results from measurements done in the climate chamber are presented. The results are a culmination of varying the fog conditions in the climate chamber, measuring the droplet size along with the liquid water content in the air. Some typical measurements are seen in figure 33 and some visual description of conditions are shown below the signals.

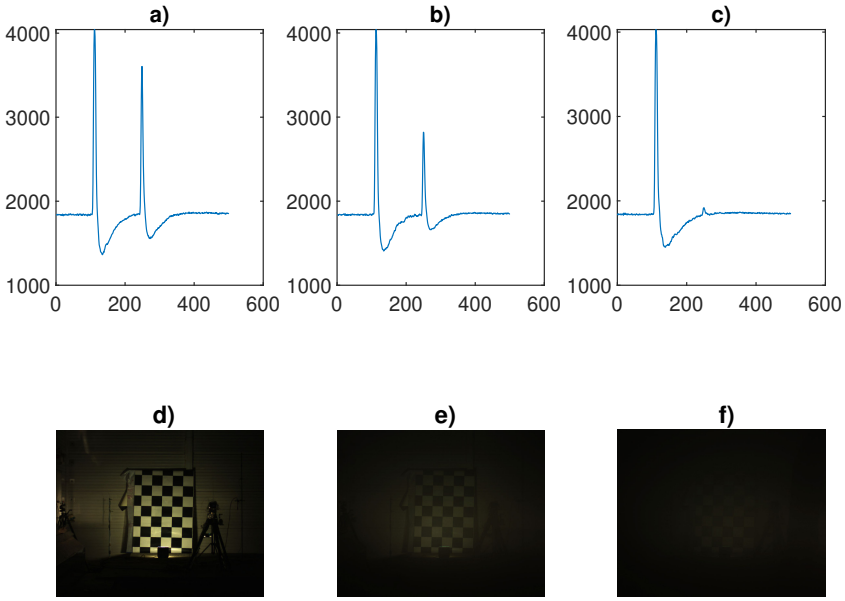

Figure 33 – Figure shows examples of LiDAR signals measured during the experiment shown in a), b) and c). The visual conditions are shown in d), e) and f) corresponding to the LiDAR signals shown in a), b) and c).

From the contrast on the black and white squares, the lowest visibility produced in the climate chamber were below 10 m. At that point, the signal of the short-range LiDAR was dampened enough to almost make the return pulse indistinguishable from possible noise in the signal. Along with that, it may be seen from figure 33 that very low visibility produces more noisy signals. This however may be due to inhomogenous fog since it was created by a spray creating patches of high density fog in some regions. This will make the extinction coefficient vary in space.
In order to estimate the expected attenuation, the particle distribution parameters were measured. The results is shown in figure 36.
The distribution of the droplets may appropriately described by log-normal distribution effective diameter which was varying between 1 to 8 μm. At the same time the amount of water was measured in the for of liquid water content which measures the mass of

water in a cubic meter. The water content varied from 0.02-0.16 g/m^3. The results are shown in figure 34.

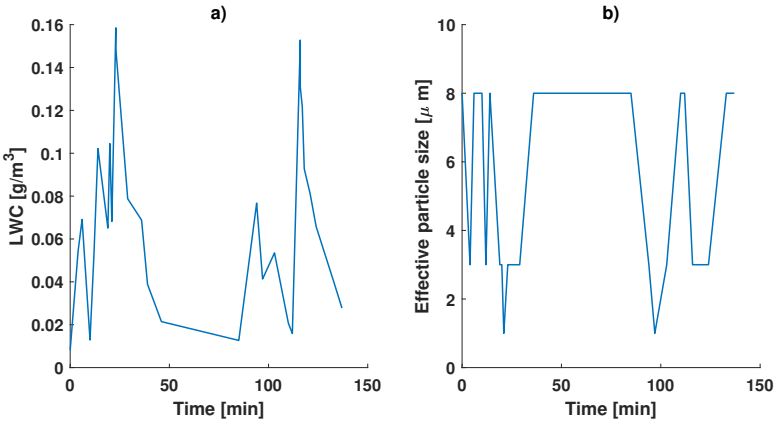

Figure 34 – Figure shows the a) the liquid water content as it varied during the time of the experiment along with b) the effective particle size during the same time interval.

Varying the distribution parameters shown in figure 34 produces an estimated extinction coefficient. This coefficient is calculated using equation (24). From the extinction coefficient, visibility may be estimated using any of the empirical models such as Kim-Kruse model in equation (25) or the Naboulsi models (26) and (27). The visibility resulted from the control experiment is shown in figure 35.

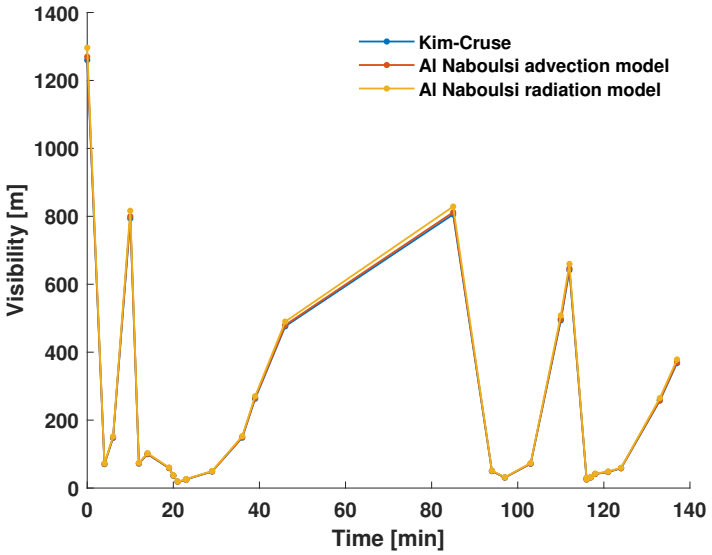

Figure 35 – Figure shows the visibilities reached by varying the distribution parameters in the chamber, where then the visibility was estimated using the Kim-Kruse model (25), and Naboulsi models for advection fog (26) and radiation fog (27).

We can see per definition of fog, which says that fog is generally defined as a weather condition for which visibility is under 1 km, that the generated conditions were indeed fog[17]. However there seems to be no significant distinction between the different models when estimating visibility. This may not be true when using larger wavelength since the two Naboulsi models diverge for increasing wavelength(smaller wavelengths also but it is assumed that shorter wavelength are not an option for LiDAR's since they would be approaching the visible range) as shown in figure 11.
As well as estimating what the visibility is in the created fog conditions, the LiDAR measurements are also compared with the expected amplitude which are calculated using the normalised LiDAR equation (14). The comparison between measurements and theoretically calculated amplitudes are shown in figure 36.

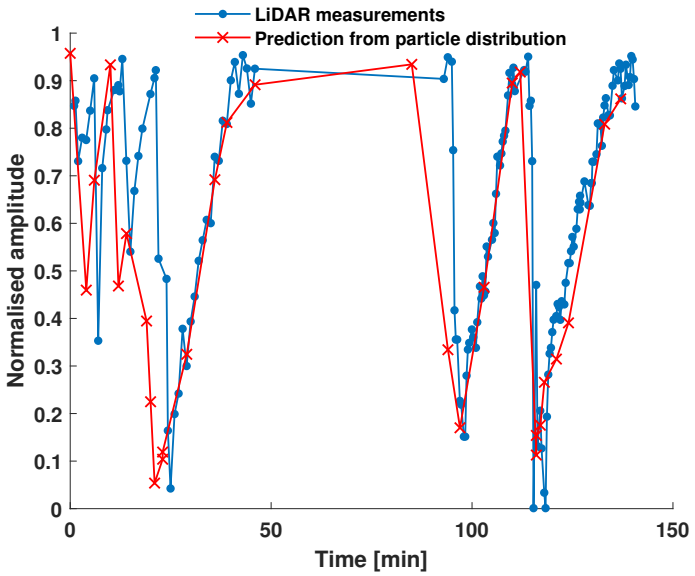

Figure 36 – Figure shows the measured data by the LiDAR during the experiment along with expected attenuation using predictions made from Mie theory approximation in equation (24).

The results above show that the models derived from distributions descriptions which results in equation (24) is a good way determine what kind of attenuation one might expect.
There seem to be some discrepancy between the measured amplitudes and the theoretical expectation, in particular at the beginning of the experiment. The reason for this is hypothesised to be due the the fact that the instrumentation measuring the distribution properties were situated a distance a few meters away from the short-range LiDAR. At the start of the experiment the fog was very localised so that the LiDAR measured different conditions while the rest of the instrumentation measured another, so that the LiDAR has a slight delay in the beginning. This delay was then reduced once the fog became homogenous everywhere in the chamber.

Finally, the reverse might possibly be done so that LiDAR amplitude might be useful for potentially determine distribution properties. This is done by running the the LiDAR equation in reverse to estimate the extinction coefficient β. However the extinction coefficient does not depend only on one parameters. For larger wavelengths, the extinction only depends on liquid water content, and the effective size through equation (24). Since there are two unknown, only the combination of LWC/D_{eff} might be estimated.

5 Discussion and future work

The study of the LiDAR signal and how it was affected by weather conditions produced some interesting result. First we described the theoretical framework for describing LiDAR pulse propagation, in which the Radiative transfer equation is most general description of the phenomena of radiation transfer. There the central quantity in regards to attenuation due to weather conditions was the extinction coefficient, which represents the total attenuation which is due to scattering and absorption of photons. The radiative transfer model could be simplified into a form most useful for LiDAR application called the LiDAR equation. In the LiDAR equation, the signal amplitude is dependent on multiple measurement system parameters, the distance travelled and the extinction coefficient. Since only the extinction coefficient varied for the measurements, the LiDAR equation was used.
Next, the description of the medium was given where Mie theory is central. With certain conditions regarding particle size distribution, the extinction could easily be determined with parameter such as liquid water content and effective particle size. Some empirical models regarding visibility and precipitation were considered for determining extinction coefficient.
In the examination of the parameters that could affect the signal, it was found that as opposed to the long-range LiDAR, the short-range LiDAR was not affected significantly by the ambient temperature, which might be cause of isolation or self regulation of temperature. However when looking at the internal temperature of the laser and the detector, the signal amplitude was significantly decreased in those operating temperatures. A correction factor linear in temperatures was created and added to the signal amplitude.
In the temperature dependence, there were many outliers from the linear trend. We examined if angular dependency was a potential cause but that did not seem to be the case since analysing angles showed no systematic variation. A possible task for future assignment is to verify that the outliers are not just specific to this study, but also persist in the future studies. If then it is verified that there seems to be outliers, study what is the cause of creating so many outliers. There the angle dependency should be verified. A potential cause could be reflectivity of the target.
Next we analysed the data gathered in clear weather, fog, rain and weather with mixed conditions, where the empirical models of Kim-Kruse and the Naboulsi models were compared to data. Just as in the previous studies of Marcus Hedlund and Johan Viklund, the empirical models showed to be adequate in prediction of normalised LiDAR amplitude. There was quite a bit of variation around the models though even when temperature compensation was applied. A more interesting result is perhaps that in rainy conditions, the amplitude shows difference from the other conditions where the amplitudes are systematically larger than in other conditions. It is postulated that a possible cause could be that rain changes reflectivity of the target, or that it affects the temperature of the measurement system significantly. A future study could be to determine if those hypotheses are false to some degree. Also as mentioned the amplitudes show substantial

random variation around the models(except in rain). An additional task is to determine if those seemingly random variation may be attributed to some other factors that are not extinction coefficient, temperature, or surrounding luminance.
In this study, the data that was used was taken when the ambient light was at minimal luminance. This means that measurements were taken in complete or almost complete darkness. Since autonomous vehicles are suppose to operating in all types of times during the day, a study to determine if the effect of ambient light could be included as a factor or if there is any pattern to effects of light on the signal amplitude.
In addition to verifying that the visibility models are adequate even for short-range LiDARs, the Carbonboue models with precipitation were examined. There it seems to be a greater discrepancy between the data and models. From the analysis, the Carbonnue models do no seem to represent the pattern in the data sufficiently well and precipitation rate does not seem to be a good predictor for determining the attenuation of signal in rain. A perhaps better predictor in rain would be the droplet size. It is postulated that smaller droplets would decrease the signal amplitude more than bigger droplet size for same precipitation, since decreasing droplet size would resemble fog more than larger droplets. A better predictor for rain is visibility when applied to LiDAR signals. Also, with application to automotive industry, the decrease in signal amplitude in rain in very small even in very heavy rain of 50 mm/h. Therefore, regarding *only* the attenuation of the amplitude due to rain does not seem to be a large safety factor. However other effects could occur in rainy conditions such as changing the reflectivity which are significant.
In regards to snowy conditions, for this study the effects of snow were not examined since the short-range LiDAR was dis-functional during snowy conditions at the time. With the LiDAR being returned to working conditions, another potential task for future work is to determine the effects of snow for the short-range LiDAR.
Next the method of determining visibility was examined. The application of the method gave mixed results and showed to work better for lower visibilities and the variation of the estimation increased greatly after 500 m visibility. There also seemed to be a systematic bias of about 330 m. As was discussed in the previous chapter, the application of this method to the automotive industry seems to be minimal, and is perhaps more suited for fields such as meteorology since it requires a calibration and the reflectivity of the target must be kept constant. Another problem is that it is suspected to work better for targets further away and are static, which is certainly not suited for moving vehicles. Nevertheless, the task for future is to analyse the method for larger distances and determine the performance.
The next result were applying principal component analysis in order to transform the LiDAR signals into feature vectors that could be used to separate the different conditions. There we showed a significant separation between the conditions were the return pulse was kept within the analysis. Removing the return pulse still yielded some separation, but perhaps not as great as with the return kept intact.
Predictive models of LDA, QDA and vanilla neural network were trained to determine the predictive power. Training on the data with the return pulse yielded very high sensitivity in LDA and neural networks with lower sensitivity in QDA. For precision the

situation reversed where LDA and neural networks had slightly lower precision while QDA had higher precision. Establishing the performance with the return pulse it could be compared to results without the return pulse. Removing the return pulse from the signals lowered both sensitivity and precision which was expected. The potential list of future task in this area is extensive. The first task would perhaps be to verify the results gotten in this study and perform the same analysis on the long-range LiDAR. Another potential modification is to determine if there are other transformation methods of the data that are more suitable than PCA and if there are more suitable predictive models to use other than discriminant analysis and neural networks. Here, support vector machines is a possible alternative among others. In this study we used PCA on the whole signal for feature extraction. In future study some other possible parameters could be manually extracted as suggested by Johan Viklund along with analysing the derivative of the signal.
Finally, experiments were done in collaboration with RISE to determine the effects of fog by varying particle distribution parameters. There the measurements done on the size distribution proved to be good in describing the effect on attenuating LiDAR signals.
Yet another potential future task as suggested by Johan Viklund is to analyse single waveform and their noise. From that one can analyse at what distance the return pulse will of comparable size to the noise.

References

[1] Marcus Hedlund 2020. *Weather Influence on LiDAR Signals using the Transient Radiative Transfer and LiDAR Equations*. Master's thesis. Dept. of Computer Science, Electrical and Space Engineering, Luleå University of Technology.

[2] Johan Viklund 2021. *Atmospheric Attenuation for LiDAR Systems in Adverse Weather Conditions*. Master's thesis. Dept. of physics, Umeå University.

[3] David J. Griffiths, *Introduction to Electrodynamics*. 4th ed. Pearson:(2013).

[4] David J. Griffiths, *Introduction to Quantum Mechanics*. 2nd ed. Pearson:(2005).

[5] Lockwood D.J. (2016) *Rayleigh and Mie Scattering*. In: Luo M.R. (eds) Encyclopedia of Color Science and Technology. Springer, New York. https://doi.org/10.1007/978-1-4419-8071-7_218(Accessed 2022-03-15)

[6] James J. Buglia 1986. *Introduction to Theory of Athmospheric Radiative Transfer*. NASA, Hampton Virginia

[7] Ralph Baierlein 2005. *Thermal Physics*. 1st ed. Cambridge University Press.

[8] Chandrasekhar S. 1950. *Radiative transfer*, 1st ed. Claredon press, Oxford

[9] Caitlin Griffith, *Light Scattering*, Lecture notes. University of Arizona. https://www.lpl.arizona.edu/classes/Griffith_656A/AT10+13-scattering.pdf(Accessed 2022-04-03)

[10] Diogo A.R Lopes, António Ramires Fernandes (2014). *Athmosperic Scattering - State of the Art*. Universidade do Minho, Portugal. https://www.researchgate.net/publication/309399089_Atmospheric_Scattering_-State_of_the_Art(Accessed 2022-04-09)

[11] Frank L. Pedrotti, Leno M. Pedrotti. Leno S. Pedrotti *Introduction to Optics*. 3rd ed. Cambridge University Press:(2017).

[12] B.T.N. Evans, *On the inversion of the LiDAR equation*. Centre de recherches pour la défense:(1984), Québec, Canada.

[13] Jost Heintzenberg (1994). *Properties of Log-Normal Particle Size Distribution*. Institute for Tropospheric Research, Leipzig Germany. https://doi.org/10.1080/02786829408959695(Accessed 2022-02-25)

[14] Dense, *Characteristics of Adverse Weather Conditions*. European Commision:(2017). https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5b0632e8d&appId=PPGMS(Accessed 2022-03-15)

[15] P. Aarne Vesilind (1980). *The Rosin-Rammler Particle Size Distribution*. Duke University, North Carolina.
https://doi.org/10.1016/0304-3967(80)90007-4(Accessed 2022-04-23)

[16] H. C. van de Hulst, *Light Scattering by Small Particles*. Dover Publications:(2012)

[17] Atmosphere. (2020). *Fog Classification by Their Droplet Size Distributions: Application to the Characterization of Cerema's Platform*
https://doi.org/10.3390/atmos11060596(Accessed 2022-03-15)

[18] Laurence D. McGlauchlin and Richmond B. McQuistan Paul W. Kruse, *Elements of Infrared Technology: Generation, Transmission, and Detection*. John Wiley & Sons:(1962)

[19] Maher C. Al Naboulsi, Frederique De Fornel, Hervé Sizun, Michael Gebhart, Erich Leitgeb, S. Sheikh Muhammad, Benno Flecker, and Christoph Chlestil, *Measured and predicted light attenuation in dense coastal upslope fog at 650, 850, and 950 nm for free-space optics applications*, Optical Engineering 47(3), 036001 (1 March 2008).
https://doi.org/10.1117/1.2899023(Accessed 2022-03-12)

[20] Mazin Ali A. Ali (2013). *Analysis Study of Rain Attenuation on Optical Communications Link*. Al-Mustansiriya University https://www.researchgate.net/publication/265384431_Analysis_Study_of_Rain_Attenuation_on_Optical_Communications_Link

[21] J. S. Marshall, W. Mc K. Palmer (1948). *The distribution of raindrops with size*. The institute for Advanced Study. https://doi.org/10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2(Accessed 2022-04-04)

[22] Richard A. Johnson, Dean W. Wichern. 2007. *Applied Multivatiate Statistical Analysis*. 6th ed. Pearson Education.

[23] Trevor Hastie, Robert Tibshirani, Jerome Friedman. 2007. *The Elements of Statistical learning*. 2nd ed. Springer.

[24] Mathworks. *Neural Networks*
https://se.mathworks.com/help/stats/neural-networks-for-classification.html?s_tid=CRUX_lftnav(Accessed 2022-05-10)

[25] Shai Shalev-Shwartz, Shai Ben-David. 2014. *Understanding Machine Learning: From Theory to Algorithms*. New York:Cambridge University Press.

A Neural Network

Here some discussions about neural networks is discussed, since they were used in order to analyse the predictive power of some machine learning models, of which neural network were one of three used. The other two being linear discriminant analysis and quadratic discriminant. The theory for both of these fall under the same umbrella, namely Gaussian discriminant analysis. The theory for discriminant analysis is fairly simple and could appropriately explained in the main bulk of the paper. The theory of neural networks is a little bit more involved, therefore should either be skipped or put in the appendix. The term neural network encompasses a large number of models and statistical methods, giving an enormous catalogue of literature. Therefore the rudimentary theory of neural networks used in the main study is presented here.
In this discussion, a restriction is put in a sense to only talk about a classic neural network commonly called single hidden layer network.
In essence, a neural network is just a nonlinear model, that is usually depicted by the graph structure shown in figure

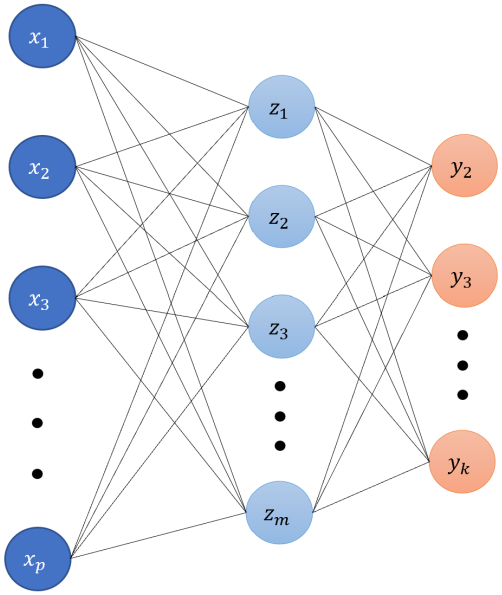

Figure 37 – Figure illustrates the basic architecture of a one-layer neural network.

The entire goal of the neural network is to map a certain feature vector X to the output vector Y, where these vectors are given by

$$\begin{cases} X = [x_1, x_2, x_3, ..., x_p]^\top \\ Y = [y_1, y_2, ..., y_k]^\top \end{cases}$$

The mapping from the feature vector X to the output vector Y is determined by the

connections in the network. The connections are represented by weights. For classification, the classes have to be encoded in a vector so that Y(k=K) = 1 and the rest of the elements are 0. This represents the output vector of the class K.
The derived vector Z which represents the hidden layer, is created by taking linear combinations of the feature vectors according to the weights and applying an activation function to the linear combination. A commonly used activation is the sigmoid function given by

$$\sigma(z) = \frac{1}{1 + e^{-sz}},$$

where s is a parameter determining the slope of the activation function, which may be seen in figure 38.

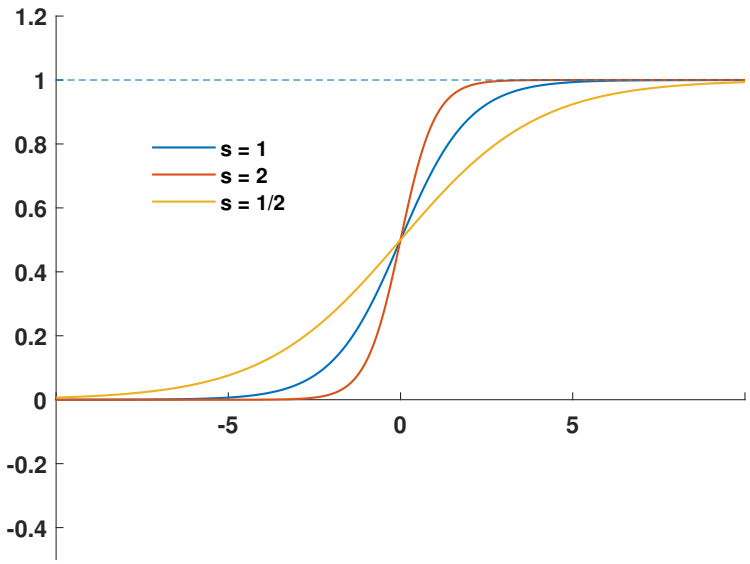

Figure 38 – Figure shows the sigmoid activation function for a few cases of the slope parameter s.

Given an input layer X, the connection weights between input and hidden layer along with the activation function, we can calculate the vlues of the hidden layer as applying the activation function to the linear combination as follows

$$Z_i = \sigma(\alpha_{0i} + \alpha_m^\top X),\ i = 1, 2, 3, ..., m,$$

For the next step, we determine the outputs of the hidden layer as

$$T_j = \beta_{0j} + \beta_j^\top Z,\ j = 1, 2, 3, ..., k,$$

where β_j are the weights from the hidden layer to the output layer.
Finally, to get the output vector Y, we apply the final transformation function $g_j(T)$, so that

$$Y_j = g_j(T),\ j = 1, 2, 3, ..., k$$

This mapping process is essentially summarised as taking the input vector X and getting and output vector as a function $Y_j = f_j(X)$, or more formally

$$f|\mathbb{R}^p \longrightarrow \mathbb{R}^k.$$

. It is summerised by the following algorithm

1. Apply linear transformation to input vector

2. Apply activation function σ to linear transformation to get output of the hidden layer

3. Apply linear transformation to the output of the hidden layer.

4. Apply final transformation function g to get the final output vector.

The algorithmic process above described the general approach to mapping the input vector to the output vector. We could in principle insert more layers by inserting a linear transformation and applying the activation function after the hidden layers.
There are multiply choices one can make when chosing the function $g_j(T)$. A possible choice is that the final transformation does not do anything, in which case

$$g_j(T) = T_j.$$

This is a reasonable choice, however if one wishes to interpret the output as being the probability so that $g_j(T)$ is the probability that the input layer corresponds to the class j, then a more appropriate choice is the softmax function

$$g_j(T) = \frac{e^{T_j}}{\sum_{l=1}^{k} e^{T_l}}$$

In order to give the model predictive power, one has to choose the optimal weights in the network. This method is referred to as backpropagation which is a optimisation algorithm involving gradient descent or stochastic gradient descent. Those algorithms will not be explained in this appendix and a detailed description is given in *Understanding Machine Learning: From theory to Algorithms*[25].