Erasmus+

COSI Computational Colour and Spectral Imaging

Master in Computational Colour and Spectral Imaging (COSI)

UNIVERSITY OF EASTERN FINLAND

UNIVERSIDAD DE GRANADA

UNIVERSITÉ DE LYON

UNIVERSITÉ JEAN MONNET SAINT-ÉTIENNE

NTNU Norwegian University of Science and Technology

Psychophysical and optical evaluation of transmission haze

Master Thesis Report

Presented by

Maria Santandreu Oliver

and defended at the

Norwegian University of Science and Technology

September 2023

Academic Supervisor(s): Prof. Rafael Huertas Roa
Host Supervisor: Prof. Frédéric Leloup
Jury Committee:

1. Dr. Giorgio Trumpy, Norwegian University of Science and Technology, Norway
2. Dr. Pauli Fält, University of Eastern Finland, Finland

Submission of the thesis: 10$^{\text{th}}$ August 2023
Day of the oral defense: 4$^{\text{th}}$ September 2023

Abstract

This research assesses optical measurement methods used to characterise the transmission haze in correlation with its visual perception. Seven silica-filled amorphous polymer test samples from a developed sample set were selected for this purpose. The optical properties of these polymers were examined using commercial haze meters, along with alternative measuring techniques such as BTDF and luminance camera measurements. The BYK commercial instrument used an integrating sphere to measure the scattered light through the samples, whereas the Rhopoint commercial instrument used an image-based technique to analyse contrast loss through the samples. For the alternative measurements, a Near Field Goniophotometer was used to conduct BTDF measurements at different angles and evaluate possible haze changes for different exit port aperture dimensions. Also, luminance images were acquired and a post-processing image method was used. This enabled assessing the loss of contrast caused by the samples and replicating the approach of the Rhopoint instrument. Following, the different optical methods were analysed and compared. Multiple psychophysical tests presented under different assessment conditions were performed to collect the subjective data and establish visual scales using the Maximum Likelihood Difference Scaling approach. In fact, test samples were backlit through a contrast pattern with three distinct predefined air gap sizes (0mm, 4mm, and 8mm) between the test sample and the contrast pattern. Finally, the results show a clear correlation between the visual and optical haze scales. The samples presented with a 4 mm air gap appear to have the best correlation and to be an optimal configuration to assess haze.

Acronyms

ASTM: American Society for Testing and Materials
ISO: International Organization for Standardization
CIE: Commision Internationale de l'Eclairage
BTDF: Bidirectional Transmittance Distribution Function
NFG: Near-Field Goniophotometer
MLDS: Maximum likelihood difference scaling
GLM: Generalized Linear Model
ROI: Region Of Interest
NTU: Nephelometric Turbidity Units
AU: Arbitrary Units
JND: Just Noticeable Difference

Contents

1 Introduction

This chapter provides relevant terminology for understanding haze and its associated qualities such as transparency and clarity. An overview of the visual perception of these properties, as well as the physical approach is provided. Eventually, the structure and goals of the thesis are detailed.

1.1 Definition of haze

Haze, together with clarity, is one of the attributes of transparency. The transparency of an object is characterised by the transmitted scattered light, and the difference between haze and clarity remains in the scattering angle. International Lighting Vocabulary from CIE (2020) has been consulted as a reference to define this terminology.

- Transparency is defined as the physical property of the materials that makes objects visible through it. It refers to the amount of light that passes through the examined material.

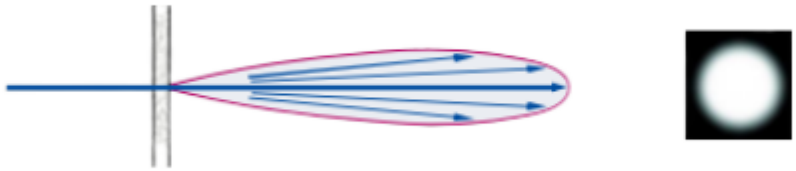
- Clarity of a material is a property of a semi-transparent or transparent material that enables one to perceive through the material, high-contrast images and objects which are positioned at some distance behind the material.

- Haze is defined as an attribute normally related to scattered light, that is perceived as a cloudy appearance caused by a reduction of contrast of objects viewed through a material. It is also considered the percentage of transmitted scattered light that is deflected from the incident beam by more than a certain angle.

1.2 Contextualization of the problem

Light scattering phenomena could explain transparency, considering that the light travelling through transparent materials may be scattered. According to Binsbergen and Van Duijn (1967), the Rayleigh criterion of resolving power facilitates the discrimination between two light sources of identical brightness. The small-angle scattering phenomenon might vary the intensity and resolution of an image of a point source, thus it might impact the clarity of an image. However, the relative contrast is reduced due to the wide-angle scattering when the light comes from a point and passes through a hazy sample. It might also increase the difference in intensity and the resolution angle.

(a) *Wide angle scattering, haze*



(b) *Narrow-angle scattering, clarity*

Figure 1.1: *Difference between perceptual clarity and haze effects. Source: BYK-Gardner (2010)*

Figure 1.1 taken from a report of one of the commercial haze meter brands that measure these transparency attributes shows this perceptual variation caused by the different scattering angles.

Wide-angle scattering results in a loss of contrast between the white and black regions shown in the upper part of Figure 1.1a, making the difference between

these two regions more diffuse. As seen in Figure 1.1b, narrow-angle scattering diminishes the detail resolution. There is no loss of contrast in this particular case. Instead, the edge between the black and white portions is less pronounced.

Years ago, Marasco and Task (1999) introduced a primary approach to the effect that scattered light can have on a target and the impact on visual perception. In Figure 1.2, the effect of haze is represented on a black and white region, background (dark area) and target (bright area), resp.

In the end, the produced effect is a reduction of the contrast that simply can be calculated and compared between samples through any contrast metric. The diagram shows how the apparent contrast between the target and background regions is caused by the apparent luminance from the scattered light, which is the primary consequence of it.

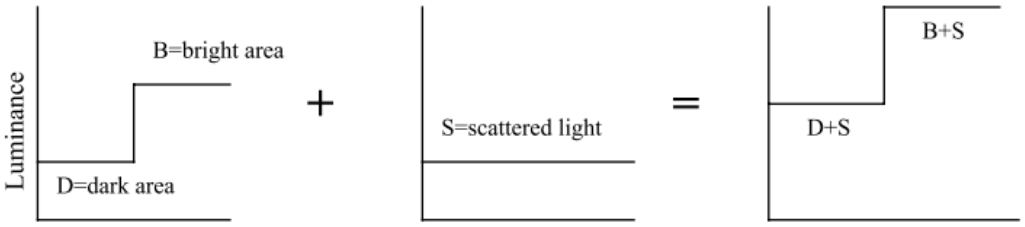


Figure 1.2: *Effect of the scattered light. Source: Marasco and Task (1999)*

Contrary to popular belief, transparency is not always the opposite of haziness. This is because haze depends only on wide-angle scattering, while transparency depends on wide and small-angle scattering. Because of that, some hazy specimens could be more transparent and have better resolution than less hazy counterparts (Webber, 1957; Binsbergen and Van Duijn, 1967)

As stated in Webber (1957) due to the properties of materials, some might scatter broadly the transmitted light but the influence that has on the resolution of the objects seen through them is small. Moreover, transparency correlates well with visual perception when light diffusers, ground glasses, and highly transparent materials are evaluated. When polymers and polyethylene materials are used, this correlation could be altered. Even though the size and properties of the particles of these materials may scatter the light significantly, only a little impact is produced on the resolution of the objects seen through them. One example of this is shown in Figure 1.3 where the sample positioned on the left test chart has a higher haze value, but apparently, also seems to have a better resolution.

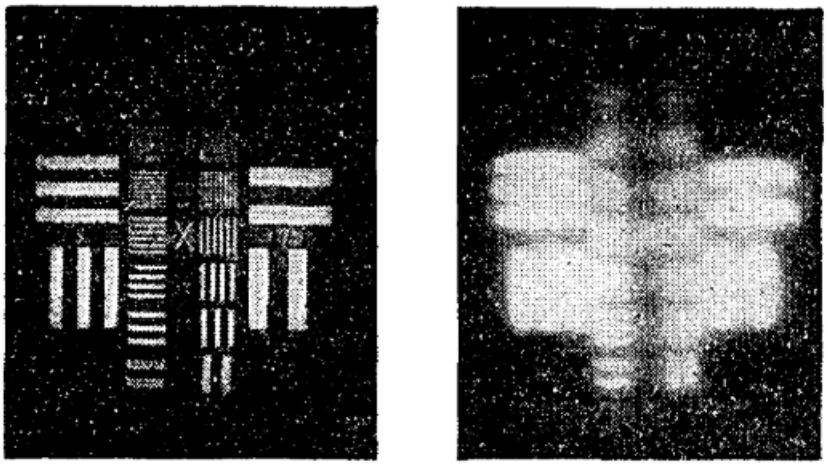


Figure 1.3: *Test charts photographed through different hazy films. Left chart sample properties: Transparency = 12%, Haze = 52%. Right chard sample properties: Transparency = 6.2%, Haze = 15%. Source: Webber (1957)*

The perception of transparency is defined by CIE (2020) as the degree of visibility of an object observed through a medium. However, going into detail, transparency is different from clarity which is the ability to resolve this observed object and consequently is related to the detail resolution. An example of it is tinted glass, which has low transparency and high clarity.

Haze is identified as an apparent reduction in contrast of an object that is viewed through a medium. Depending on the application, a lower or higher degree of haze is desired. For instance, in the case of packaging, a lower haze value is desired to let the observer perceive clearly and evaluate the product (Morris, 2017). On food or beverage packaging, for instance, perceptual appearance is a crucial quality criterion. Contrarily, for optical applications such as light-management layers for displays or photovoltaic technologies, high values of haze are preferred (Wu et al., 2019; Lim et al., 2020).

1.3 Research goals and contributions

Transmission haze has been characterised by standard methods to be measured with optical instruments. Every method considers a specific definition of haze, requirements and setups that should be satisfied to quantify this property. Nevertheless, there is a lack of research addressing the correlation between optical measurements and visual perception of haze, especially on rigid surfaces such as glass or polymers. However, some studies analysed or simulated this correlation on

nonrigid transparent layers or beverages. Notwithstanding, the evaluation of haze is based on the scattering of light for different states, with the procedures varying depending on whether a solid, liquid, or gas is investigated.

Because there is any defined method for analysing haze on rigid surfaces, especially the visual perception of haze, this study attempts to propose a procedure and comprehend this optical phenomenon. Several measurements and experiments are carried out to evaluate the optimum method to correlate the visual perception of haze and optical measurements of haze performed using different instruments and techniques.

Optical haze measurements are obtained on a particular solid set of hazy samples developed for that purpose in a partner laboratory. First, measurements done with commercial instruments to characterise the set of samples were performed by industrial partners of KU Leuven University who own these devices, and the resulting data was provided for this academic research. Moreover, when Prof. Leloup began this research months before the start of this master's thesis, he performed additional characterization measurements. Using a Near Field Goniophotometer that provides a complete study of the pattern of the scattered light through a sample in all possible angles, this characterization procedure was conducted. In addition to these optical measurements, luminance images are acquired with a luminance camera to assess the contrast reduction that the samples induce on a contrast pattern. After, all the optical data is analysed and compared in this thesis.

Psychophysical assessments are performed under different viewing configurations. Three separate sample holders are made and placed on top of a homogeneous diffuser to backlit them. The air gaps set between the sample and the contrast pattern that the observers analysed were 0mm, 4mm and 8mm. The observers should select the biggest difference in contrast reduction between two pairs of samples. Afterwards, these responses are examined and the visual scales are estimated using the Maximum Likelihood Difference Scaling procedure.

Finally, correlations between visual haze perception and optical haze measurement are established. Moreover, based on the obtained results, the most suitable approach to examine haze is proposed.

In this thesis, AI tools, like chat GPT, have not been used. Optical measurements usually yield directly the haze values or simple calculations are needed in order to obtain them from the data of the measurements. These basic equations or rations described in this thesis (cf. infra) are coded using R programming language. Regarding psychophysical investigations, since the MLDS package from the R programming language is used for processing the data and obtaining the visual scales, just minimal manually coded arrangements are made for the analysis of the data and plots.

The outline of the thesis is organised based on the following structure. In

Chapter 2, a literature review explains the definitions of haze proposed in standard methods and the theoretical background on visual haze assessments. In Chapter 3, the methodology applied is detailed, specifying the optical methods used for the measurement of optical haze, as well as the description of the setup and the procedure used for the visual experiments. Chapter 4 describes and analyses the results obtained from the optical measurements of the hazy samples, the visual scales got from the psychophysical tests, and finally, the correlation between both. In the end, the conclusions are exposed in Chapter 5.

2 Literature Review

In this chapter, the different standards and techniques used to measure haze optically are described, along with an analysis of previous research on the visual assessment of haze.

2.1 Standards

Standardised methods have been defined and introduced to quantify the haze of transparent plastics, such as ASTM Test Method D1003-21 (American Society for Testing and Materials, 2021) and ISO14782:2021 (International Organization for Standardization, 2021). Commercial haze meters have appeared on the market and quantify haze according to ASTM D1003 standard. Recently, a new imaging-based approach has been proposed to evaluate haze and material transmission metrics. This alternative measurement technique is based on the analysis and post-processing of images.

2.1.1 ASTM Test Method D1003

Two methods to measure light transmittance and haze are defined in this standard, implemented by the use of a haze meter and a spectrophotometer, resp. Haze meter measurements normally give more stable values. This work is focused on haze meter functioning and measurements, and therefore, the methods described also centre the attention on this technique and instrument.

ASTM D1003-21 (American Society for Testing and Materials, 2021) quantifies haze as the ratio of light scattered by more than 2.5^{o} from the incident beam direction to the total transmitted light through the specimen, for a unidirectional illumination. A *unidirectional illumination:diffuse viewing* geometry, and a *diffuse illumination:unidirectional viewing* geometry are defined depending on the measurement method. The first one is used for the haze meters, while the second is used when the haze is measured with a spectrophotometer. To determine the

amount of diffuse response and totally transmitted light, an integrating sphere setup is proposed, including an exit port on which a light trap can be mounted.

A schematic of this *unidirectional illumination:diffuse viewing* is provided in Figure 2.1. It reflects the system used for the conventional haze meters that follow this standard, one of the well-known is the *BYK-Gardner GmH* haze meter (Wimmer and Schwarz, 2014).

For measuring the regular luminous transmittance, a clear specimen needs to be placed at some distance from the entrance port of the integrating sphere. Nevertheless, hazy specimens have to be positioned at the entrance port of the sphere, for the total hemispherical luminous transmittance to be measured.

It is important that the materials are free of defects, for instance, scratches or dust, and they should be large enough to cover the entrance port of the sphere.

Furthermore, haze meters need to fulfil some geometric and spectral requirements according to the ASTM D1003 standard. A schematic of the instrument is shown in Figure 2.1. These requirements and elements include an integrating sphere to assemble the transmitted flux, a unidirectional beam to illuminate the specimen which should not be vignetted at either port, and a light source and photodetector giving luminosity response of the 1931 CIE Standard Colorimetric Observer with a CIE Standard Illuminant C, or Illuminant A, provided by a filter.

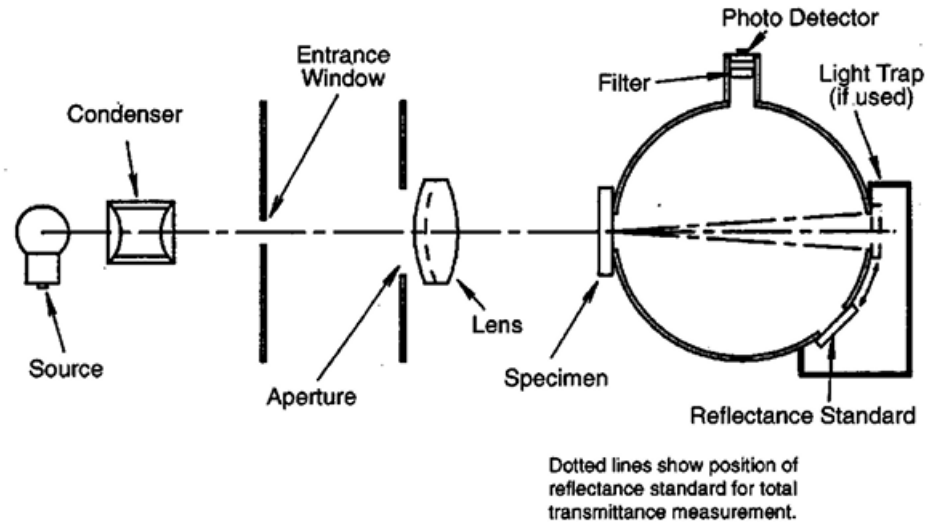


Figure 2.1: *Schematic of a haze meter with Unidirectional Illumination:diffuse viewing geometry. Source: American Society for Testing and Materials (2021)*

An integrating sphere is a spherical hollow covered with a white diffuse reflective coating which provides a uniform scattering effect. Light incident on any point

of the inner surface is distributed equally to all other points through numerous reflections. This optical element also features an entry hole for the light to enter, and an exit port to connect the photodetector and perform the measurements.

The integrating sphere of haze meters also includes a light trap that absorbs all the light in a particular direction and enables measuring the scattered light by more than a certain angle.

The interior surfaces of this sphere, baffles, and white reflectance standard should have the same reflectance, matte, and they should be highly reflecting in the visible spectrum. If the system is designed based on a light trap, this shall totally absorb the beam when there is no specimen placed.

It is important to maintain the photometric stability, temperature and humidity of the elements along the measurements. Moreover, a series of calibrated haze standards for the verification of the accuracy of the instrumental response should be performed.

To calculate haze, the light scattered by the sample is compared to the total transmittance. Only the scattered light will be gathered in the sphere, the non-scattered light will be absorbed by the light trap, which is necessary for this measurement. This light will arrive at the sensor and the scattered light due to the sample investigated will be detected. Finally, the ratio of the scattered light (T_{sct}) and the total transmitted light (T_T) is equivalent to the haze value as defined in Equation 2.1.

$$Haze = \frac{T_{sct}}{T_T} \times 100\% \tag{2.1}$$

On the other side, clarity is supposed to be defined as the light scattered by less than 2.5^{o}. Although BYK measurements follow this supposition, it is not well defined in the ASTM standard.

This standard is valid to quantify haze values that are below 30%. If the percentage of haze is higher, it is considered diffusing light and should be tested in accordance with a different standard (American Society for Testing and Materials, 2019).

As a counterpart to ASTM D1003-21, ISO has introduced ISO 14782:2021 (International Organization for Standardization, 2021) to measure haze in transparent and substantially colourless plastics. In this case, the measurements are applicable to values of haze lower than 40% instead of 30%, but the principle and the instrument of the measurement is like the one described in ASTM D1003-21.

2.1.2 Imaging-based method

One of the inconveniences of the BYK instrument as cited in its patent (Wimmer and Schwarz, 2014) is that it cannot measure and define different optical parameters needed for the optical quality evaluation of products.

Recently, a novel imaged-based technique that allows a more complete optical characterization was proposed by Busato et al. (2021), including optical parameters such as the illumination diffusion haze and sharpness, distance dependency, local defects and impact of the specimen luminescence. The measurement method is based on ISO 14782:2021 (International Organization for Standardization, 2021) for the determination of haze, as well as on ISO 12233:2023 (International Organisation for Standards, 2023) which characterises photographic techniques for measuring the resolution and the spatial frequency response of images.

The proposed image-based method uses a different technique compared to the previously described standard method. In this case, the scattering angle is not considered. Instead, haze is driven by the analysis of the reduced image quality. In Figure 2.2, a schematic arrangement of the instrument is illustrated.

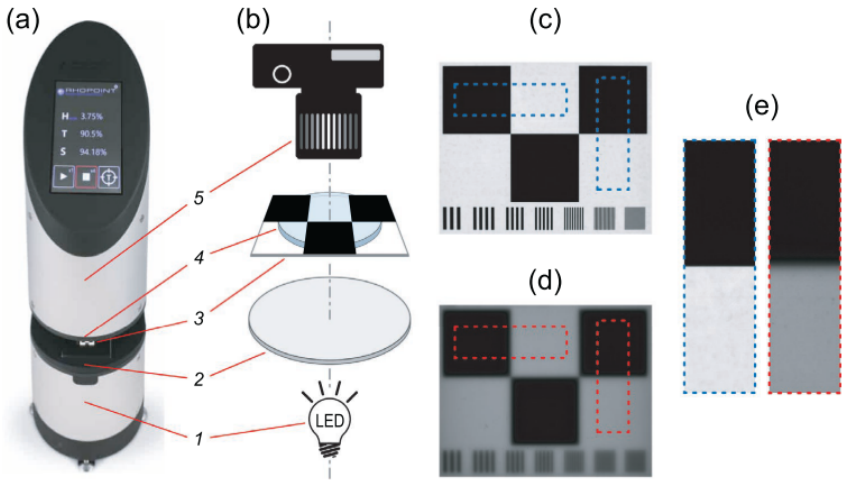


Figure 2.2: *Schematic arrangement of the instrument proposed by Busato et al. 1. Lamp, 2. Diffuser, 3. Knife edge mask, 4. Specimen, 5. Camera. a) Rhopoint instrument, b) Schematic illustration of the components, c) Reference knife-edge mask image without sample, d) Hazy sample in contact with the knife-edge mask image, e) Expanded ROIs. Source: Busato et al. (2021)*

Part a) of Figure 2.2 represents the Rhopoint instrument itself and part b) is a detailed diagram of the components inside the instrument. These components are 1) an LED lamp to backlit the samples, 2) a diffuser to achieve homogeneous illumination, 3) a knife-edge mask to evaluate the properties related to haze and clarity, 4) the evaluated specimen, 5) a camera calibrated and fixed with specific parameters to focus on the knife-edge mask and capture the image of it through the specimen. Parts c) and d) are images taken with the camera of this scenario. Part c) is the reference image without any sample placed on top of it, and part d) is the same capture but with a haze sample positioned on top of the contrast pattern mask. Part e) of the Figure are the same mask areas as parts c) and d) expanding the Regions Of Interest (ROIs) that they want to evaluate.

Illumination Diffusion (ID) is defined as a redistribution of the spatial pattern formed by the graticule, because of diffuse light scattering produced by the evaluated specimen. The features of ID include a loss of contrast and sharpness caused by modifications in the spatial distribution of the transmitted light that affect the quality of the image. These modifications are brought about by scattering and refraction. ID is evaluated with a backlit knife edge observed through a transparent or non-transparent specimen. The knife edge provides a spatial modulation of illumination flux by a high-contrast step-like pattern. A comparative evaluation is done by capturing unmodified graticule images and seeing them through a sample to mimic the visual perception.

Two qualities, described as ID-haze and ID-sharpness are determined with the instrument. ID-haze quantifies the reduction of contrast between two predefined regions, for instance, the transparent and opaque areas close to the knife edge mask, when a test sample is placed on it, (e) in Figure 2.2). It is determined from the ratio of contrast between the test sample (C_{sample}), and the contrast without any sample ($C_{reference}$). The percentage of ID-haze ($H_{ID}(\%)$) may be calculated from this ratio, according to Equation 2.2

$$H_{ID}(\%) = \left(1 - \frac{C_{sample}}{C_{reference}}\right) \times 100 \tag{2.2}$$

C_{sample} and $C_{reference}$ are calculated as the Michelson contrast with the test sample in place and without any sample, resp, as expressed in Equation 2.3. L_{Max} and L_{min} denote the average maximum luminance of the transparent region side of the knife edge mask and the average minimum luminance of the opaque side region of the knife edge mask, resp.

$$C = \frac{L_{Max} - L_{min}}{L_{Max} + L_{min}} \tag{2.3}$$

Next to ID-haze, ID-sharpness ($S_{ID}(\%)$) assesses how the Modulated Transfer Function (MTF) is reduced. MTF depicts the amplitude response of the imaging system to sinusoidal modulations in illumination flux over a range of spatial frequencies. It is normalised to 1 at a spatial frequency of 0, and for ID-sharpness, the spatial frequency range of 5-10 lp/mm (line pairs per millimetre) is assessed. Similar to ID-haze, this percentage $S_{ID}(\%)$, is calculated by dividing the sample MTF value by the reference MTF value.

This new photographic imaging method gives information on the overall contact haze and is therefore more suitable for packaging applications. Moreover, additional information on the microstructure and optical properties is provided with this method.

2.2 Visual haze assessments

Several studies related optical measurements of clarity with the visual perception of the attribute (Webber, 1957; Morris, 2017). Although this correlation being performed, usually, the clarity attribute is considered the same as transparency, meaning that it gives consideration to the whole transmittance of the different samples. Considering CIE definition CIE (2020), clarity is evaluated perceptually as the capacity to see through a specimen and obtain high-contrast images. Contrary, taking into account the definition of Busato et al. (2021), MTF is taken to evaluate the equivalent metric to clarity. As can be seen, this perceptual definition and the distinction between clarity and transparency in visual perception assessments are not well defined, either in ASTM reference standards for light transmission evaluation, For that reason, this work will focus all the attention on haze assessments.

Haze is defined by the CIE (2020); Busato et al. (2021), and other studies as a loss of contrast. Thus, its perceptual definition is more consistent. Nowadays, not many studies evaluate how haze is perceived through glass or polymers and correlated to optical measurements in a detailed manner. Regarding the scattering of light in different states of haze, some investigations have studied this relationship in liquids, while other investigations simulated different conditions. They are detailed in the following sections.

2.2.1 Solids

In a technical report of Task and Genco (1985), the haze of plastic windscreens is studied. This is one of the first works studying the relationship between haze measurement and the visual performance achieved using different kinds of windscreen materials.

There, haze perception is described as a loss of contrast caused by scattered light that enters the viewing direction of the observer. It results in a veiling luminance that according to CIE (2020), appears because a luminance superimposes on the retinal image and produces this effect.

Because the ASTM haze standard is only valid for small samples and setups where the scattered light can be gathered in the integrating sphere, while large aircraft windscreens haze had to be determined, they designed a new measurement method. In Figure 2.3 the setup that allows measuring veiling luminance (L) induced by the scattering in the windscreen is presented.

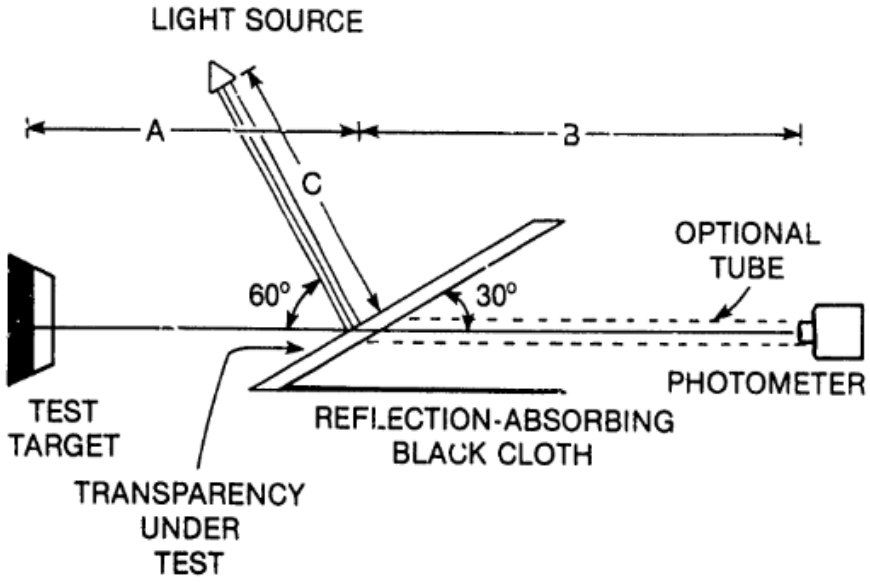


Figure 2.3: *Schematic measurement setup proposed by Task and Genco (1985) to measure haze*

It uses a photometer to capture the black and white areas of the test target through a specimen (transparency under test) while being illuminated by a semi-collimated light source in the laboratory or directly by the sun in the real environment. A black absorbing surface (reflection-absorbing black cloth) is used to absorb the flux transmitted in the regular direction.

A surface that reflectively scatters all incident light is referred to as a Lambertian reflector. When this kind of reflector, such as a Barium Sulphate plate, is placed on the transparency under test, the luminance can be measured with the photometer. Due to the relationship between the definitions of foot-lamberts (lumi-

nance) and foot-candles (illumination), the luminance of this diffusing reflector is mathematically comparable to the illuminance falling on the surface. Consequently, this configuration along with an additional Lambertian reflector might also be used to measure the Illumination on the Surface (E).

Moreover, the transmission coefficient of the windscreen can be measured by positioning a white target instead of the test target. The structure of the system allows measuring for all illumination and viewing angles.

Once the previous parameters are measured, the Haze index (H_i) defined as in Equation 2.4 can be calculated in the laboratory measurements. In the case of real measurements in the sun, the luminance of the black target without the windscreen needs to be measured and multiplied by the transmission coefficient of the sample. After, this amount needs to be subtracted from the veiling luminance for accurate measurements.

$$H_i = \frac{L}{E} \tag{2.4}$$

Considering the luminance of the white and black areas with and without a windscreen, a method similar to the Michelson contrast ratio between the reference and sample measurements is proposed to evaluate the loss of contrast. More transmissive samples have a lower contrast loss, taking the haze index as a reference to compare.

Besides, an analysis of the angles of the source incidence with respect to the sample was done. If the light source emits toward the windscreen, the incident illumination flux is usually higher, and the contrast loss is more significant. Otherwise, when incident illumination flux is lower due to a higher viewing angle, the reduction of contrast is less relevant. The relationship between the angular subtense of the target and the contrast for a dark circular disk on a light background was proved, claiming that as the angle increases, the modulation contrast is lower.

The haze index was related to the comments of the crew that participated in the experiment. In high haze index samples, the opinions were negative, even with unacceptable operability, while the lower haze index specimens obtained the best subjective evaluations. Other remarkable results were that a higher transmittance does not always result in a better contrast perception, coinciding with Webber (1957) and Binsbergen and Van Duijn (1967).

Recently, Marasco and Task (2001) evaluated the visual performance through scattering visors following a procedure of contrast analysis akin to that used in Task and Genco (1985). The observers were shown a black Landolt C on a white background. The experiment began with a bright background, and observers were asked to adjust the target background luminance until they could barely perceive the gap of the letter C and not perceive a closed circle. With hazy samples placed

in front of this test, visual perception was evaluated.

They discovered that for samples with a lower haze value measured using the ASTM D1003 standard, the observers needed a higher background luminance to discern the black character. The background luminance increases approximately linearly when veiling luminance increases. They observed a coefficient of determination of $R^2 = 0.940$ between the predicted and experimental data of the background luminance adjustments of the visual performance. Compared to the haze values of the samples, veiling luminance has a stronger relationship with visual performance. This is because the veiling luminance scattered from a specimen considers the illumination and observation geometry, elements not considered in ASTM D1003 measurements.

2.2.2 Liquids

CIE (2020) defines turbidity as a reduction of transparency in a specimen caused by the presence of particulate matter. Turbidity was also defined years ago as a physical concept by Thorne and Nannestad (1959). It was considered the extinction coefficient due to light scattering, the total scattered light in all directions from the incident beam whilst it goes through a suspension. In their research work, the turbidity of water was addressed, and they proposed that measurements of the scattering of light in the liquid should also consider the size and shape of the particles of this. This principle is based on obscuration, which measures the reduction in light transmission caused by the suspended particulate matter in a liquid.

Carrasco and Siebert (1999) characterised the turbidity describing the haze in liquids using specific instruments and related it to human perception. Pulfrich nephelometer employed to measure turbidity uses Nephelometric Turbidity Units (NTU), the clarity of 23 commercial juice samples of three different colours (clear, yellow and red) was measured using it. Additionally, three tests were carried out, including threshold determination, magnitude estimation, and sensory descriptive analysis. Several regression models were developed to relate the results of the visual assessments with the liquid properties to predict turbidimeter haze values. R-squared values ranging from 0.870 to 0.986 were obtained, and these relationships considered the colour of the solution and its particle size. The turbidimeter measurements were influenced by the particle size and concentration, but the influence of the colour was minimal, obtaining the best $R^2 = 0.986$ between instrument turbidity measurements and predicted values considering particle size and concentration.

2.2.3 Gases

As a gas, haze is semantically more complex than a liquid or a solid. Mist, fog, smoke, volcanic ash, widespread dust, sand, and haze are all categorised as obscuration by International Meteorological regulations (World Meteorological Organization, 2019). When the air is visible and objects in the background are reduced, the dispersion of light is the physical cause of haze.

In Dövencioğlu et al. (2018), the authors simulated some 2D and 3D generative shapes that the observers needed to judge. The fiducial images were extracted from a pre-created data set, and the edges were modified using Matlab. The idea was to distort the borders of the figures to create different eidolons used as stimuli. The fuzzy cloud surrounding the image could be considered as the eidolon (Koenderink et al., 2017).

Observers judged and adjusted the distortions of the eidolons under various simulations of water (liquid), haze (gas), and structured glass (solid). When haze was analyzed, the variability between different subjects was higher. Under the configurations of water and glass, the quality of the edges remained the same, whereas, for haze, the edges were affected by the loss of contrast.

3 Methods

This chapter encompasses the characterization of the samples used for the experiments and the analysis of this data in an optical way using different instruments and techniques, as well as the procedure used for the analysis of the answers obtained in the psychophysical tests. To conclude this chapter, a final section detailing the criteria used to select the alternative optical measurements, the methods of analysis and the elements of the built setup are specified.

3.1 Samples

A dedicated set of 7 circular silica-filled amorphous polymer test samples of 26.6mm diameter and 1.1mm thickness was developed for the study in a partner company laboratory, following the same procedure as Busato et al. (2021). Polymers and different concentrations of additive powder (1.5 µm of SiO_2) ranging between 0.002% and 0.3% were dry-mixed. Next, they were put into a laboratory co-rotating mini-twin-screw extruder at 240^{o}C for approximately 5 minutes at 40rpm under a nitrogen blanket. The molten mixes were extruded into a micro-injection moulder maintained at the appropriate temperature. After each extrusion step, the compounded mix was progressively diluted by adding the necessary amount of neat polymer to make the polymer additive concentrations. This is an optimal manufacturing technique to produce plaques with insignificant wrap and curvature. Ultimately, samples were injected into a plaque mould at 20^{o}C.

Figure 3.1 shows the less haziest sample used in the experiment with a concentration of SiO_2 additive powder of 0.002% on the left side and the haziest sample with a concentration of 0.3% on the right side. They are placed on a sample holder that contains a contrast pattern below and it allows seeing how the sample concentration influences the contrast perceived.

Figure 3.1: *Test samples. SiO_2 concentration of 0.002% (left) and SiO_2 concentration of 0.3% (right)*

3.2 Characterization of the samples

The samples have been characterised through different optical measurements, using commercial instruments that measure haze, measuring the Bidirectional Transmittance Distribution Function (BTDF) with a near-field goniophotometer, and analysing image properties with a luminance camera. BTDF measurement results were compared to measurement results obtained from commercial instruments since both characterise the transmitted scattered light considering a certain scattering angle. Conversely, the analysis with the luminance camera is similar to the method used by Busato et al. (2021) system. Further on, the procedure explained in this work to obtain haze values from the Michelson contrast is applied to compare the results of both methods.

3.2.1 Commercial instruments

Two commercial instruments were used to determine the haze values of the different samples, BYK Haze Gard Plus (BYK-Gardner GmbH, 2015) and Rhopoint ID Transmission Appearance Meter (Rhopoint Instruments Ltd, 2018). The collection of this data was previously done by industry partners and provided to the University of KU Leuven for the study.

BYK instrument measures according to ASTM D1003-21 standard, while Rhopoint instrument measures taking into account the method proposed in Busato

et al. (2021). Both techniques were described in the previous Section 2.1.

3.2.2 BTDF measurements

In addition to commercial optical measurements, Prof. Leloup preciously measured the BTDF of each sample using a Near-Field Goniophotometer (NFG). A description of the procedure used in this technique is provided in this section, and an analysis of the data obtained will be provided in the next Chapter.

BTDF was described for the first time in Bartell et al. (1981) study as an extension of the work already done by Nicodemus et al. (1977). BTDF explains how the scattered light is transmitted through a thin surface and provides an alternative method to quantify optical scatter with a precise angular distribution, according to the theory put forward in Marasco and Task (1999).

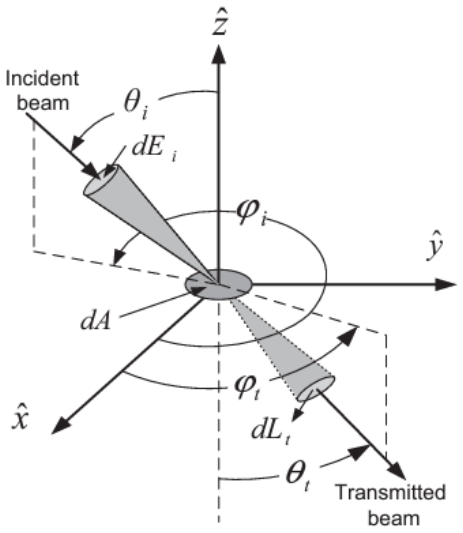


Figure 3.2: *BTDF light transmission geometry. Source: Wang et al. (2015)*

BTDF determines the relationship between the incident and transmitted light intensities for the incoming and outgoing angles, denoted as θ_i and θ_t, resp., in Figure 3.2. This ratio can be characterised by the transmitted radiance, L_t, and the incident irradiance, E_i as specified in Equation 3.1.

$$BTDF(\theta_i, \phi_i, \theta_t, \phi_t) = \frac{dL_t(\theta_i, \phi_i, \theta_t, \phi_t)}{dE_i(\theta_i, \phi_i)} \tag{3.1}$$

Contrarily, Bidirectional Reflectance Distribution Function (BRDF) characterises the reflected light from a material in a way similar to BTDF, but instead of considering the transmitted beam, the incident beam and the reflected are considered. Bidirectional Scattering Distribution Function (BSDF) includes BRDF and BTDF together, thereby describing the interaction of scattered light with surfaces and materials.

Bidirectional distribution functions try to reproduce the actual behaviour of the surfaces, and they have been used for visual effects modelling that mimic the actual light behaviour and the appearance of materials. Because of that, they have been recently implemented in computer vision applications (Nomura et al., 2011; Steinberg et al., 2022).

Even though these models are challenging to define, they are helpful for the standardisation and characterisation of optical properties of the material, and nowadays, this property can be measured using commercially available instruments like the NFG.

The goniophotometer uses a measuring technique based on an image-resolving CCD that could provide the Luminous Intensity Distributions (LID), luminous flux and ray data. A goniometer moving around the examined element allows to determine the lighting parameters at all possible angles on a spherical surface. This goniometer has a CCD camera sensor attached, and its field angle limits the radius of the sphere examined around the element. The rotating parts are incorporated into a free-standing rack as illustrated in Figure 3.3. The system used in the laboratory is formed by a Goniophotometer RiGO801-600 (TechnoTeam Bildverarbeitung GmbH, 2016) that includes a photometer detector acting as a far-field detector. It was adapted with a laser source to illuminate the samples. In this case, the examined element were the samples, and the values obtained were for the scattered transmitted light through them.

In order to assess possible haze changes for various exit port aperture dimensions, measurements were obtained for normal incidence and hemispherical viewing. Different exit port aperture sizes of 1.5^{o}, 3.0^{o} and 4.5^{o} were taken into account to calculate haze values and to compare them with haze values obtained with the commercial instrument (BYK).

From BTDF data, haze percentage is calculated as described in Equation 3.2, the ratio between the diffuse transmittance as defined from transmitted scattered light by more than the examined exit port aperture from the incident beam direction (T_{dif}), and the total transmission of the sample (T_T).

$$H = \frac{T_{dif}}{T_T} \times 100 \tag{3.2}$$

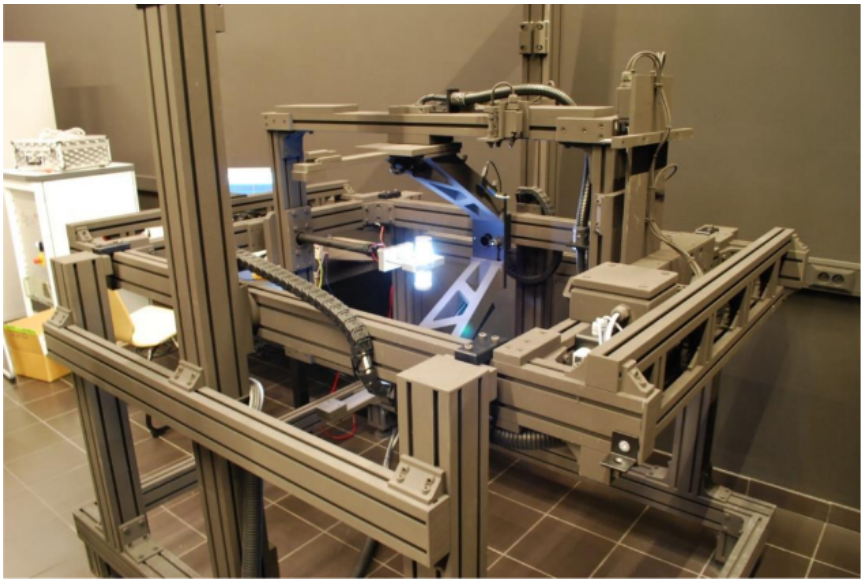


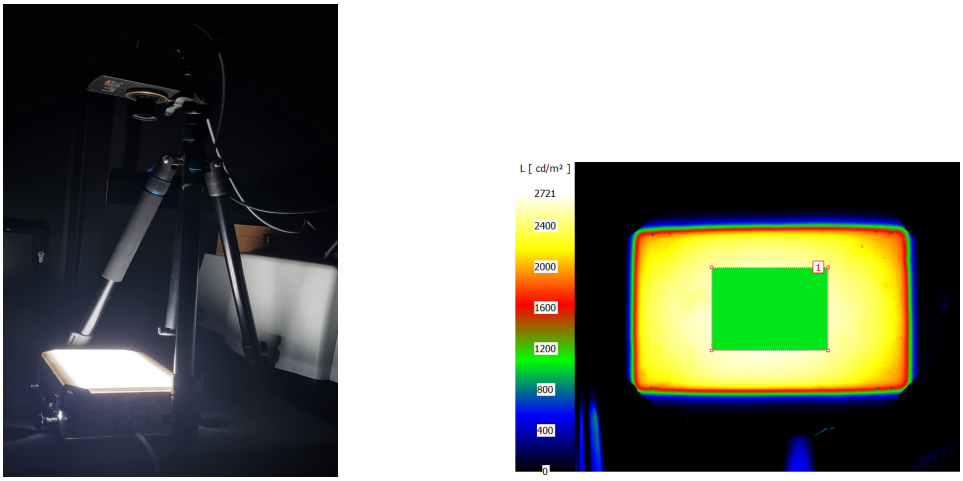
Figure 3.3: *NFG used to measure BTDF values of the set of samples*

3.2.3 Luminance camera measurements

A luminance camera (LMK) has been used to determine the thermal stability of the setup built for the visual evaluation of haze (cf. infra) and for evaluating how the contrast pattern is seen through the different samples used in the experiments. From the acquired images, the haze percentage was calculated as proposed by Busato et al. (2021) using the Michelson contrast metric.

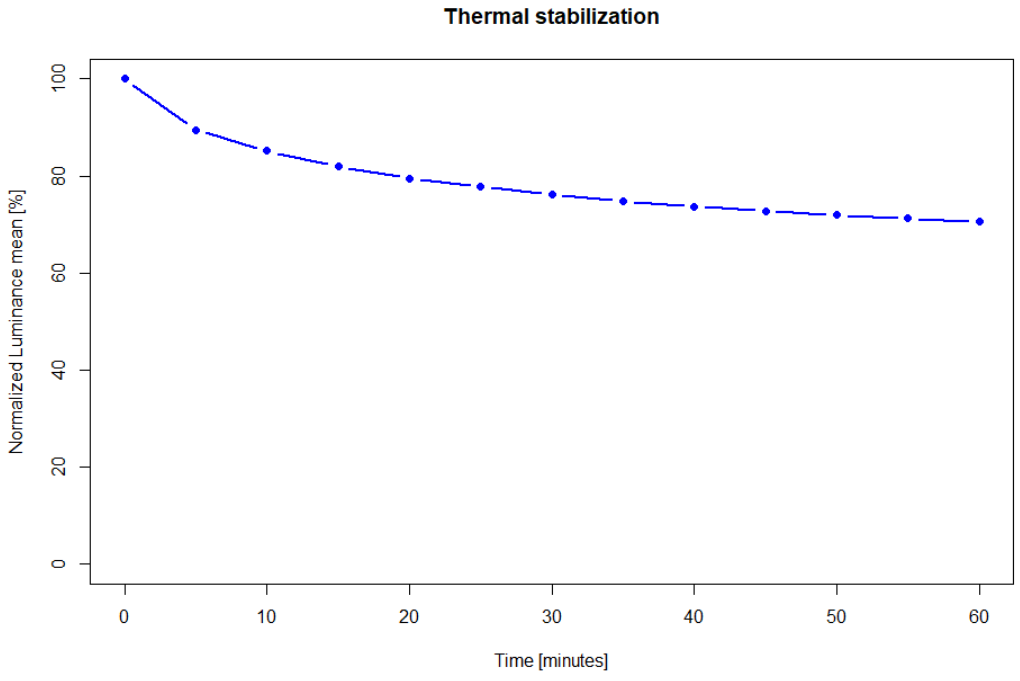
Luminance images were measured using an LMK 5-5 Color luminance camera (TechnoTeam Bildverarbeitung GmbH, 2017). It includes a colour filter wheel that may be used for colorimetric measurements and evaluating the brightness according to how humans perceive it (adapted with a $V(\lambda)$ spectral filter glass). The camera consists of a 2448×2050 pixel (5-megapixel) CCD sensor with a 14-bit resolution. The system enables the measurement of image luminance. Moreover, the software provides statistical data from the entire image or particular ROIs selected for prospective analysis.

The measurement arrangement to evaluate the stability of the built test setup for the visual assessment of haze is shown in Figure 3.4a. Several pictures taken for an hour at 5 minutes intervals determined when the luminance given by the source was entirely constant and stable. The ROI for the stability evaluation selected in the software includes the middle area of the diffuser covering the area of the sample holder placement, see Figure 3.4b.

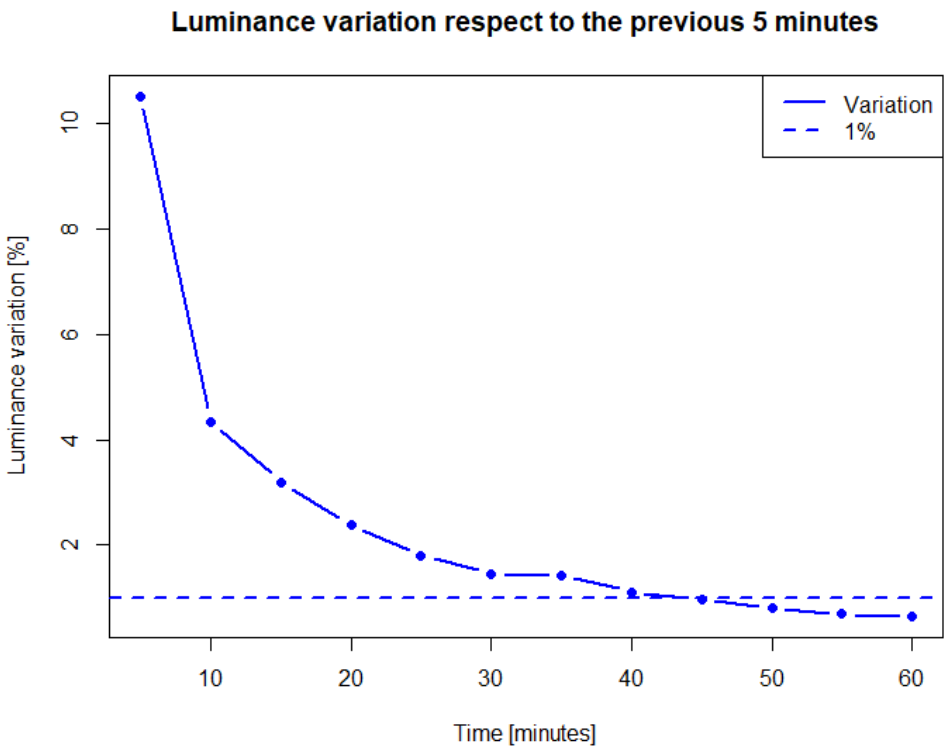


(a) *LMK setup measurement* (b) *ROI selected*

Figure 3.4: *LMK setup measurement and ROI selected*



(a) *Normalised luminance changes during one hour*

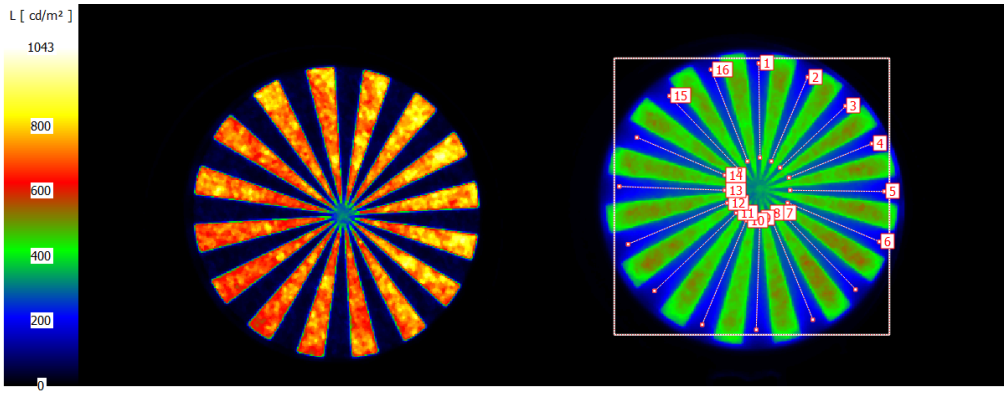


(b) *Variation in percentage of the difference with respect to the previous 5 minutes*

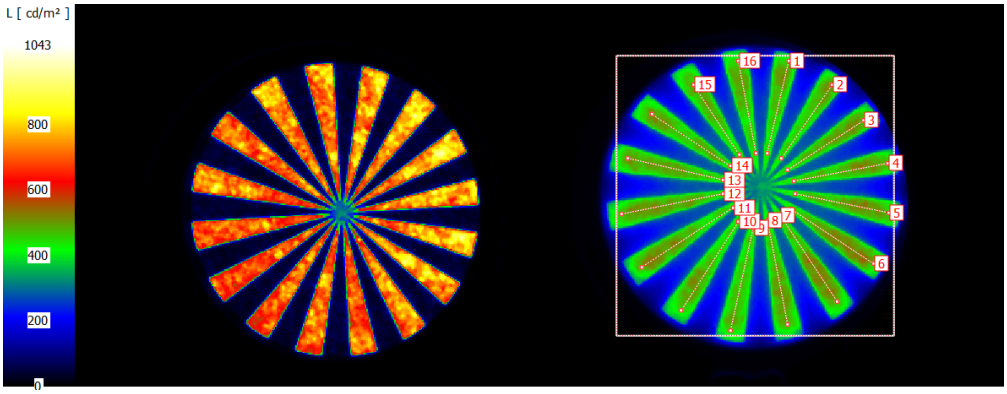
Figure 3.5: *Stabilization results from luminance variance evaluated with the LMK*

The outcomes of this normalised stabilisation luminance are represented in Figure 3.5a. The variation between contiguous measurements after 40 minutes is less than 1% as illustrated in Figure 3.5b. As a result, before beginning any tests, a stabilization period of 40 minutes was provided for the test setup.

Furthermore, a Siemens star was used as a contrast pattern. All their black and white stripes were chosen in LMK software as ROIs for haze analysis. The ROIs chosen on the Siemens star, black and white areas, are shown in Figure 3.6a and 3.6b, resp. In the presented example, the scale next to the pictures shows that the stimulus on the left has a higher luminance range than the stimulus on the right. Indeed, the left stimulus consists of a Siemens star presented without any test sample in place, while the right stimulus consists of the Siemens star presented behind a hazy test sample.



(a) *Black ROIs selected*



(b) *White ROIs selected*

Figure 3.6: *ROIs selected for the calculation of Michelson contrast*

The following steps were taken to replicate the procedure described in Busato et al. (2021). The average across all areas (black vs white) without placing any sample was computed to get the reference contrast. The contrast after the introduction of each sample was computed in the same manner. Michelson contrast was calculated using this data as defined in Equation 2.3, while from Equation 2.2 the haze percentage could then be determined.

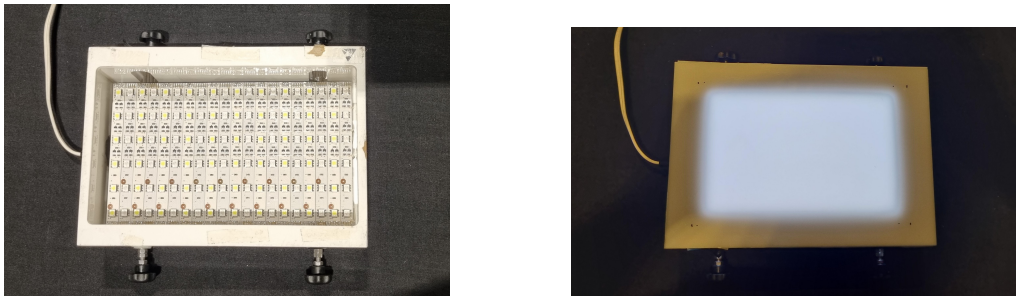
Three different configurations, including no air gap, an air gap of 4mm and an air gap of 8mm between the sample and the contrast pattern, were used in the visual experiments. More information will be detailed further in the description of the experimental setup, but all of them are considered for the characterization of the samples using the luminance camera.

3.3 Psychophysical experiments

Psychophysical experiments were performed using the set of selected samples. The Maximum Likelihood Difference Scaling (MLDS) procedure was used to derive visual scales from the responses of the observers in each test condition, as explained further. Finally, the obtained visual scales were compared with the instrumental scales obtained previously through optical measurements.

3.3.1 Setup

For the psychophysical experiments, a light booth (Figure 3.7a) with a diffuser on top was employed to backlight the sample holder and specimens, as shown in Figure 3.7b. The diffuser was chosen to offer a uniform luminance over the entire surface. On the top of the diffuser, a sample holder was placed, including four compartments to asses 2 pairs of samples simultaneously.



(a) *Light booth* **(b)** *Light booth with the diffuser*

Figure 3.7: *Light booth used without and with the diffuser on top*

The light booth and sample holder were positioned on a lectern, so the observers could place themselves approximately 50 cm away from it and look perpendicularly at the samples. Sinusoidal Siemens star target (Loebich et al., 2007), as presented in Figure 3.8, was used as a contrast pattern that the observers could assess through the presented 2 pairs of samples.

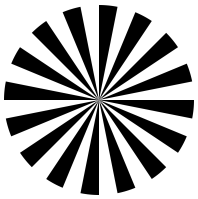
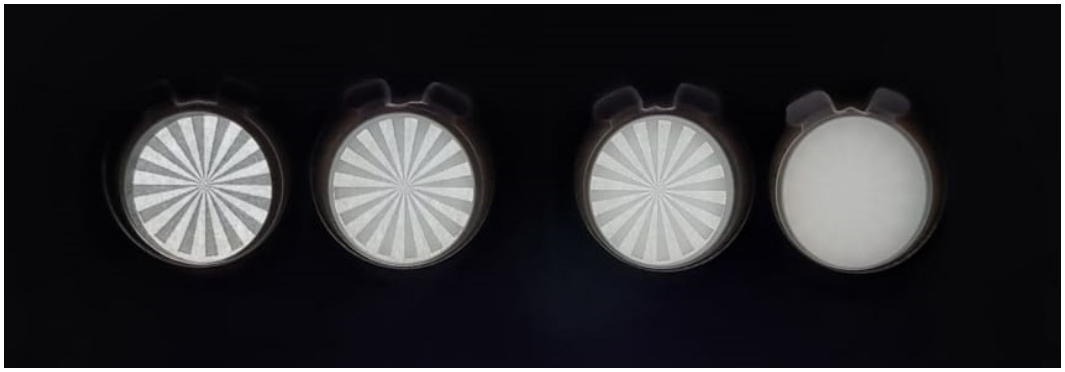
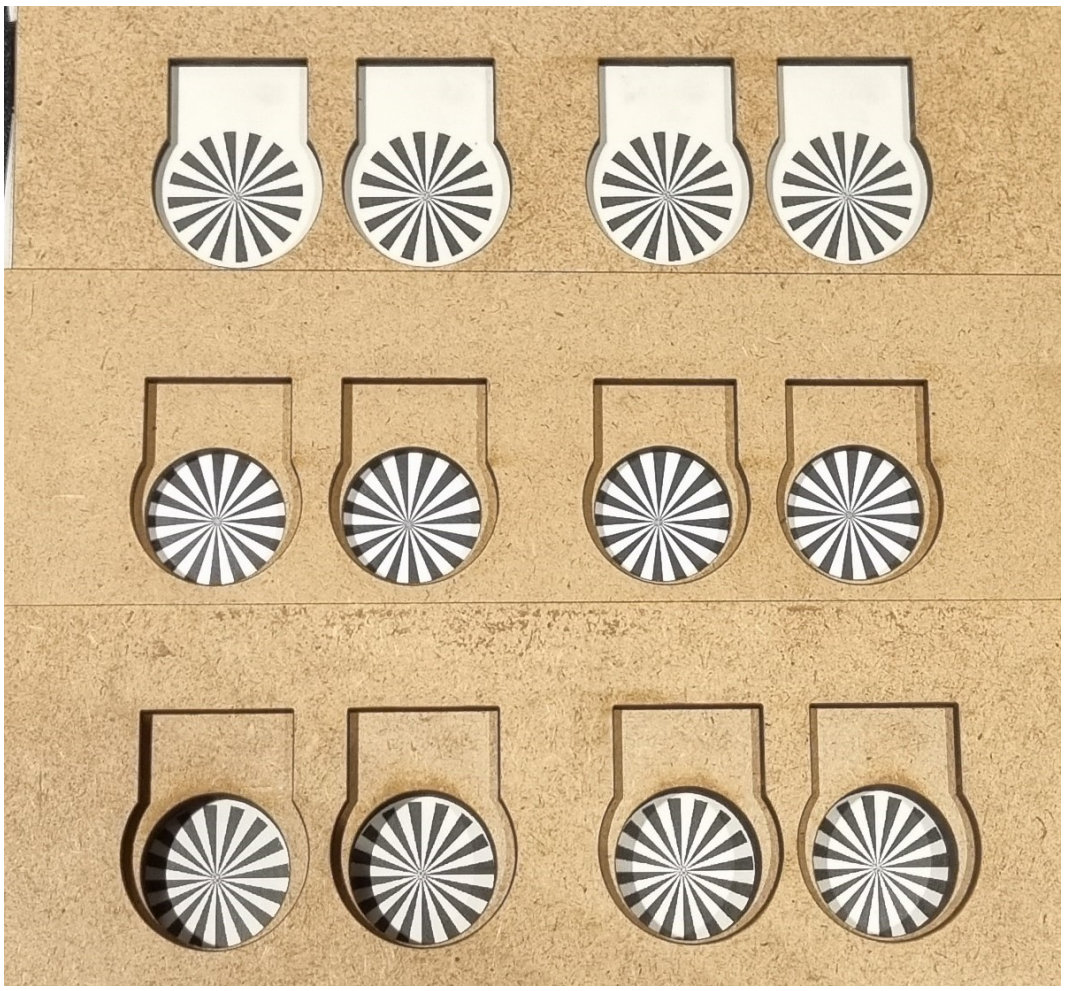


Figure 3.8: *Siemens star target used in the experiments*



(a) *Example of 2 pairs of samples presented in the experiments*



(b) *Sample holders. From Top to bottom: 0mm, 4mm, and 8mm air gap*

Figure 3.9: *Setup*

Two pairs of samples were presented for each evaluation, as seen in Figure 3.9a. The observer was asked to indicate in which sample pair (left or right) the contrast difference between the samples was the largest.

The quadruples used for the pairs were chosen in a specific manner to design the experiment correctly. The first step was choosing just the non-overlapping quadruples using the MLDS algorithm, which will be described later. This way, 35 combinations were obtained from the seven samples used.

The order for the pair presentations was selected following the method proposed by Ross (1934, 1939) that removes time and space errors, prevents frequent repetitions which may affect the judgment of the observers, and presents pairs involving the same stimulus as widely apart as feasible. Because there were 7 selected samples utilised in this study ($n = 7$), pairs with the same sample were separated by a minimum of $(n - 3)/2$ pairs and a maximum of $(n - 1)/2$ pairs in any order for the chosen elements.

35 paired comparisons were presented to each observer participating in the test, while ten random combinations were presented a second time to check the intra-observer consistency. Each participant thus evaluated 45 combinations of 2 pairs of samples, which took around 30 minutes on average.

Three different sample holder configurations were created. These arrangements produced air gaps of 0mm, 4mm and 8mm between the sample and the contrast pattern, as shown in Figure 3.9b. For each of these three configurations, 10 volunteer observers, aged between 22 and 59 years, participated in the experiment. In each experiment, the distribution of the 10 observers was 6 men and 4 women. All tests were conducted in a dark room.

3.3.2 Maximum Likelihood Difference Scaling

The MLDS scaling procedure was proposed by (Maloney and Yang, 2003). MLDS is a two-alternative forced choice (2AFC) method where two supra-threshold pairs of stimuli need to be evaluated by the observer. This model is a stochastic procedure about how the observer decides which pair contains the larger perceptual difference. Afterwards, using the maximum likelihood criterion, the parameters of the approach needed to obtain the perceived visual scale are estimated.

First of all, the quadruples used for the pairs (formed by four samples, i.e. a, b, c, d) must be non-overlapping, $a < b < c < d$, to avoid extra artefacts in the experimental design. It is important to avoid repetitions in the list of indices.

As explained in Maloney and Yang (2003), the experiment has to be designed based on distinct indices $1 \leq a \leq b \leq c \leq d \leq N$, N being the total number of samples. The total possible number of combinations is $\binom{N}{4}$ and could be described

and calculated following Equation 3.3. Taking into account that we used seven samples (N=7), our number of possible combinations is 35.

$$D(N) = \frac{1}{24}[(N-4)^4 + 10(N-4)^3 + 35(N-4)^2 + 50(N-4) + 24] \tag{3.3}$$

Moreover, every stimulus has an assigned number, this allows to have absolute differences between scale values and the judgments of the observer could be predicted accurately. In our experiment, the samples are arranged based on the concentration of additive powder that they contain, which change their haziness. Assuming each quadruple has four specimens (S_a, S_b, S_c, S_d), the quadruples will always be ordered so that the scale values assigned $(\psi_a, \psi_b, \psi_c, \psi_d)$ follow Equation 3.4. Because the scale of the provided values is arranged in ascending order for convenience, $\psi_b > \psi_a$ and $\psi_d > \psi_c$.

$$|\psi_b - \psi_a| > |\psi_d - \psi_c| \tag{3.4}$$

Randomizing the right/left locations of each stimulus in the pairings makes no sense since the observer can readily determine the sequence of the pair stimuli, $a < b$ and $c < d$, but the right/left positions of the pairs are more significant.

Considering the length scale difference value of each pair $l_{ab} = \psi_b - \psi_a$ and the $l_{cd} = \psi_d - \psi_c$, the decision variable used by the observer is represented in Equation 3.5.

$$D(a, b; c, d) = l_{ab} - l_{cd} \tag{3.5}$$

The observer should select the first interval ab if D is positive; else, the interval ab. Certain judgements in the responses could be inconsistent if D is tiny in comparison to the Gaussian standard deviation. In order to prevent that, an additional element called ε is introduced to Equation 3.5. ε denotes a Gaussian random variable having a zero mean and a standard deviation greater than 0 $(\sigma > 0)$, as defined in Equation 3.6.

$$D(a, b; c, d) = l_{ab} - l_{cd} + \varepsilon \tag{3.6}$$

The Gaussian variable and its standard deviation have no bearing on the differences of stimuli compared or their magnitude. Despite this assuming a homogeneous variance observer, the non-homogeneous variance was explored in Maloney and Yang (2003), and it was discovered that changing this parameter had

no impact on the MLDS fitting process. In real experiments, the additive error ε might have different distribution forms, not necessarily Gaussian. In Maloney and Yang (2003), they simulated alternative distributions for the probability density function of the results, such as Uniform on the interval, Laplacian and Cauchy while keeping a constant variance Gaussian error assumption. The results showed that these modifications in the distribution do not create significant changes in the estimations of free parameters compared to the results obtained with the original MLDS method. Although few biases could be identified, the scale values assigned remained the same. Therefore, the distributional robustness of the MLDS approach was validated.

Small biases and narrow confidence intervals for the estimated parameters could also appear whether a small amount of data is used. That was assessed using only a percentage of the possible trials calculated using Equation 3.3. This fact conditions the standard deviation of the results of the estimated scale. MLDS is asymptotically unbiased and efficient. It has a minimum variance, meaning that the use of more data reduces until vanishing, according to Maloney and Yang (2003). The variance of the maximum likelihood estimations converges to the minimum unbiased estimator as the amount of data grows.

The differences estimated by the observers may be altered by stochastic errors. However, the MLDS approach is successful if the performance of the observers is consistent in line with the difference scaling model.

Using MLDS could be set that $\psi_1 = 0$ and $\psi_N = 1$, obtaining $N - 1$ free parameters and the standard deviation of the error term without losing generality ($\psi_2, ..., \psi_{N-1}$ and σ). That is possible because any linear transformation of the scale values assigned parameters $(\psi_1, ..., \psi_N)$ along with the corresponding scaling of the standard deviation σ gives a group of parameters that preserve the predictive performance of the original model.

The probability of the subjective judgments is computed to estimate these free parameters using MLDS. After getting the responses of the observers for a particular quadruple, the probability of this response will be stated, given any option of the free parameters. For instance, Equation 3.7 describes the likelihood that the decision variable for the quadruple *a, b, c, d* is positive if the response is that the larger interval is the first one, *ab*.

$$P[D(a, b; c, d) > 0 | \psi_2, ..., \psi_{N-1}, \sigma] \tag{3.7}$$

Considering $x = l_{cd} - l_{ab}$, and $\phi_\sigma(x)$ the cumulative standard normal distribution function of the Gaussian random variable ε, these probabilities could be calculated following Equation 3.8.

$$\phi_\sigma(x) = P[\varepsilon \leq (x)] = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} \exp -\frac{u^2}{2\sigma^2} du \tag{3.8}$$

If the observer considers the second option *cd* as the larger interval, the decision variable is negative, and the probability will be $1 - \phi_\sigma(x)$.

So, the probability of the observers' answer could be described as in Equation 3.9. The difference in length on each trial (t) will be defined as $\Delta_t = \psi_b - \psi_a - \psi_d + \psi_c$, and $R = 1$ is used in case the first interval is classified as the largest one, and $R = 0$, in case that the second interval is selected.

$$P(r|\psi_2, ..., \psi_{N-1}, \sigma) = \phi_\sigma(\Delta_t)^R \times (1 - \phi_\sigma(\Delta_t))^{(1-R)} \tag{3.9}$$

When it comes to analysing the probability of a pattern of responses, R_t along the different trials, $t = 1, ..., T$, for any selection of free parameters, the probability of the pattern observed is written in Equation 3.10.

$$P[R_1, ..., R_T|\psi_2, ..., \psi_{N-1}, \sigma] = \prod_{t=1}^{T} \phi_\sigma(\Delta_t)_t^R [1 - \phi_\sigma(\Delta_t)]^{(1-R_t)} \tag{3.10}$$

Given the previously known response $(R_1, ..., R_T)$, it is also the likelihood of any specific selection of free parameters $L[\psi_2, ..., \psi_{N-1}, \sigma|R_1, ..., R_T]$. This likelihood could also represent a Bernoulli variable. The maximum likelihood estimations of the free parameters that permit the maximization of Equation 3.10 are obtained using conventional numerical optimization techniques.

The work of Knoblauch and Maloney (2008) and MLDS *R* package were followed to compute the visual scale once all the subjective data from the psychophysical experiments were collected. The package follows the structure of MLDS proposed in Maloney and Yang (2003), and next, the data is arranged in a Generalised Linear Model (GLM) defined in McCullagh and Nelder (1989).

Equation 3.11 defines a GLM, with X being the model matrix, β being a vector of coefficients, and $E[Y]$ is the expected value of the response vector Y which is distributed as a member of the exponential family. Finally, η is the link function that transforms $E[Y]$ to the scale of a linear predictor.

$$\eta(E[Y]) = X\beta \tag{3.11}$$

Due to the nature of the experimental data, the responses of the observers may be interpreted as Bernoulli variables and described using a binomial distribution. To construct the matrix X, the weights of the perceptual scale values (ψ) are set as

1, -1, -1, 1, corresponding to the order of the four stimuli. The dimensions of the X matrix are $n \times p$, n being the number of quadruples tested (45 in this case, 35 combinations and 10 repetitions), and p the number of physical levels evaluated (7 in this case, corresponding to the 7 samples). An example of this X matrix having the following combinations of quadruples

$$\begin{matrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 6 \\ 2 & 4 & 5 & 7 \end{matrix}$$

is this one:

$$X = \begin{pmatrix} 1 & -1 & -1 & 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 & 1 \end{pmatrix}$$

The direct optimization method considering $p-1$ parameters removes the first column of the X matrix ($\beta_1 = 0$) and it is identified with the Equation 3.12.

$$\phi^{-1}(E[Y]) = \beta_2 X_2 + \beta_3 X_3 + ... + \beta_p X_p \tag{3.12}$$

Since several optimization methods were evaluated in Maloney and Yang (2003), with no discernible changes in the estimated parameters, the *probit* link function is used in Koenderink et al. (2017) package. This link function employed in the GLM approach appears to be suitable for the majority of MLDS applications.

In a recent research (Pastor et al., 2022), an extension of the MLDS approach was developed in an effort to improve the accuracy and robustness analysis of the psychophysical data when perceptual scales from different sources of stimuli need to be analysed. Reducing the number of comparisons required in experiments of this type with numerous stimuli is another further application of this improvement work. Nevertheless, the standard MLDS technique is utilised in the present study to derive the perceptual scales from the psychophysical data. The studied stimuli are always the same for the three configurations (0mm, 4mm and 8mm air gap) and the number of stimuli used is not very large.

3.4 Criteria

This section includes a few observations about the criteria used for the setup decisions as well as the assessment, measurements and analysis of the data for this study.

- Rather than a knife edge pattern as in Busato et al. (2021), in this study, sinusoidal Siemens star targets were used as a contrast pattern. The slanted edges contrast pattern, as described in ISO 12233:2023 (International Organisation for Standards, 2023), should be used to characterise electronic still images. Nevertheless, in Busato et al. (2021), orthogonal edges are sufficient because the characterization relies on values retrieved from pictures with specimens in comparison to those from a reference image.

- Observers were viewing the samples perpendicularly. That was because haze perception changes with different viewing angles. Moreover, most of the methods used, including ASTM D1003 standard and Rhopoint instrument, measure haze values at normal incidence. Therefore, in order to assess the impact of haze on visual performance and optimise the usefulness of the measurement techniques, it is crucial to evaluate the subjective perception under the same conditions (Marasco and Task, 1999).

- All the possible 35 pair combinations for this work, coming from using 7 different samples, are used in the experiments. Following the recommendation of Maloney and Yang (2003), when a small number of stimuli is assessed with the MLDS method, it is suggested to repeat all the possible combinations as many as necessary times.

- As noted in Goos and Großmann (2011), levels of a single quality, as in this case haze evaluation, have been compared frequently using paired comparison methods in psychophysical experiments. The responses were restricted to a two-alternative forced choice technique, and a paired comparison approach was adopted in order to make the task easier for the observers, enabling the application of the MLDS method next. In addition, proposing a paired comparison approach enhances the discriminability of the given stimuli by requiring observers to simply state the preference for each pair (Pastor et al., 2022).

- Regarding the scaling method

  - In other similar studies like Carrasco and Siebert (1999), haze thresholds have been ascertained using the Ascending Method of Limits, a forced-choice technique that evaluates when the observer perceives a difference between samples. Thurstonian scaling method (Thurstone, 1927) is another well-known alternative that may be used to scale the difference between samples on a one-dimensional continuum. It is based on scale-related confusions between nearby stimuli, for instance, finding the Just Noticeable Difference (JND) in a certain physical scale. There is no

evidence proving that techniques like JND between stimuli could foretell supra-threshold perceptual differences. Otherwise, MLDS compares large and small intervals.

  - Another way to determine a visual scale is by using a ranking method, but when a linear variable cannot be used to accurately reflect the quality being assessed, the ranking system has a significant flaw (Kendall and Babington Smith, 1940).
  - After studying the different perceptual scaling approaches and weighing the benefits and drawbacks of each method, the MLDS technique was considered the best method to obtain the perceptual visual scale of the experiments.

- BTDF measurements were performed simulating different aperture sizes of the exit port to compare the data of haze percentage with the values obtained using the ASTM D1003 standard. It can be seen that Equation 3.2 follows the same format as Equation 2.1. The main difference is that the last one is fixed for $2.5^{\underline{o}}$, while for Equation 3.2, the angle varies each time depending on the measurements.

- LMK camera images were acquired to check whether the Michelson contrast could be extracted from the data to obtain the haze values and correlations with the visual data were similar to that of the Rhopoint image-based technique.

4 Results and Discussion

Results and discussion are detailed in this Chapter, starting with the results of the characterization of the samples, continuing with the psychophysical results, and finally comparing both of them to study the correlation for each of the three configurations defined (0mm, 4mm and 8mm air gap between the samples and the contrast pattern).

4.1 Characterization of the samples

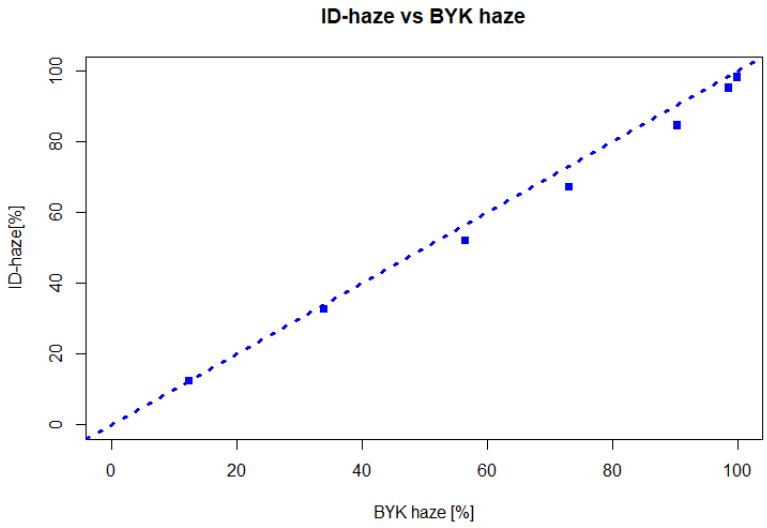
The outcomes of haze values acquired using the different measuring techniques will be shown through different graphs and tables in this section. Additionally, all the methods will be compared with one another.

4.1.1 Commercial instruments

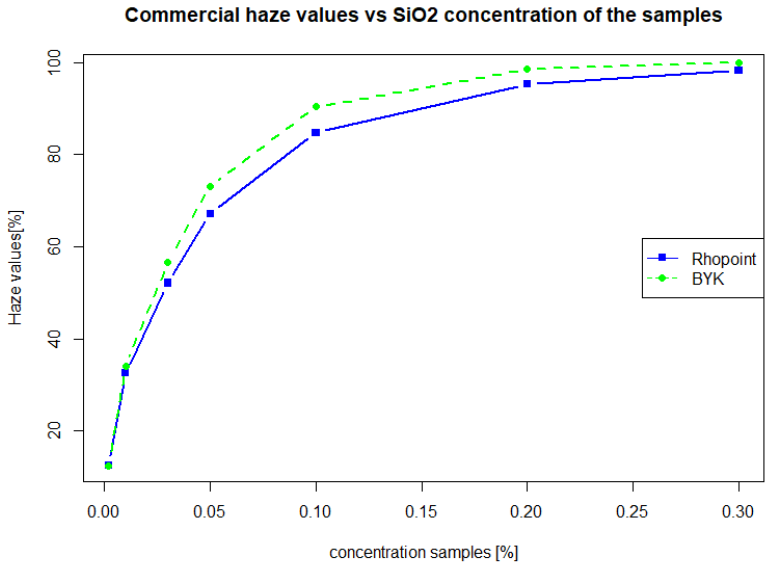
The set of seven samples used for the experiments was characterised using Rhopoint and BYK instruments. The results corresponding to haze values measured by these commercial instruments are stated in Table 4.1. Usually, there is a one-to-one correlation between the ID-haze measured with the Rhopoint instrument at around 8mm air gap distance between the sample and the contrast pattern, and the ASTM D1003 standard for values of haze lower than 30% (Busato et al., 2021). That is the reason why the commercial Rhopoint haze values obtained for the selected set of samples are also measured and reported at this distance.

For silica samples above 30% of haze, Busato et al. (2021) claimed that the Rhopoint instrument underestimates the values in comparison to the BYK instrument that measures according to the ASTM D1003 method. Results of optical measurements of the samples used in this experiment are plotted in Figure 4.1. Indeed, samples with less than 30% of haze show a clear match between both instruments, see Figure 4.1a. However, as the haze percentage rises a discrepancy between the haze values obtained with both instruments is observed, Rhopoint

instrument underestimating the values obtained with the BYK instrument (cf. Figure 4.1b).



(a) *ID-haze [%] vs BYK haze [%]*



(b) *Commercial instrument haze values*

Figure 4.1: *Commercial instruments characterization*

Table 4.1: *Commercial instrument haze values*

Sample concentration of SiO_2 [%]	Rhopoint ID-haze [%]	BYK Haze Gard Plus [%]
0.002	12.5	12.4
0.01	32.6	34.0
0.03	52.1	56.6
0.05	67.2	73.1
0.1	84.7	90.4
0.2	95.3	98.6
0.3	98.3	100

4.1.2 BTDF measurements

In this section, the results of haze calculations obtained from BTDF data for each sample are listed, except for 0.002% and 0.01% concentrations samples. BTDF values for the samples containing this additive amount of contaminant could not be used due to the low scattering value.

As previously stated, the BTDF data was used to determine three haze values for each test sample based on whether the light was scattered by more than 1.5°, 3.0°, or 4.5° from the incident beam direction. Haze values calculated from BTDF measurements (further denoted as BTDF haze values) are depicted for each of the 5 test samples measured in Figure 4.2.

Furthermore, BTDF haze values detailed in Table 4.2 concur with the findings of Task and Genco (1985) that measurements at a higher viewing angle result in lower contrast reduction. BTDF haze values calculated for a 4.5^{o} exit port aperture are lower than those calculated for a 1.5^{o} exit port aperture, which corresponds to less contrast loss.

It can be observed that for all the samples, a higher haze value is obtained at a lower measuring angle of 1.5^{o} compared to a higher angle of 4.5^{o}. This is coherent with the ASTM D1003 haze operations. By simulating an exit port with small aperture size, less light exits the integrating sphere, resulting in more scattered light detected inside of it. Hence, the haze will likewise increase when accounting for the ratio indicated in Equation 2.1.

Besides, this reduction is more noticeable for lower concentration samples (i.e. 0.03%), their scattering decreases more quickly and significantly with increasing measuring angles from 1.5^{o} to 4.5^{o} than for higher concentration samples (i.e. 0.3%). For instance, there is a decrease of 4% for the sample with the lowest SiO_2 concentration of contaminant measured with this technique, while for the sample

with 0.3% SiO_2 concentration of contaminant, this decrease is only 1.3%. Marasco and Task (1999) already suggested that this phenomenon could happen with certain materials, under some conditions and at certain viewing angles.

Table 4.2: *Haze values calculated from BTDF measurements*

Sample concentration of SiO_2 [%]	BTDF 1.5º	BTDF 3.0º	BTDF 4.5º
0.03	59.5	57.4	55.5
0.05	75.4	73.4	71.4
0.1	89.9	88.1	85.8
0.2	98.4	97.6	96.3
0.3	99.6	99.2	98.3

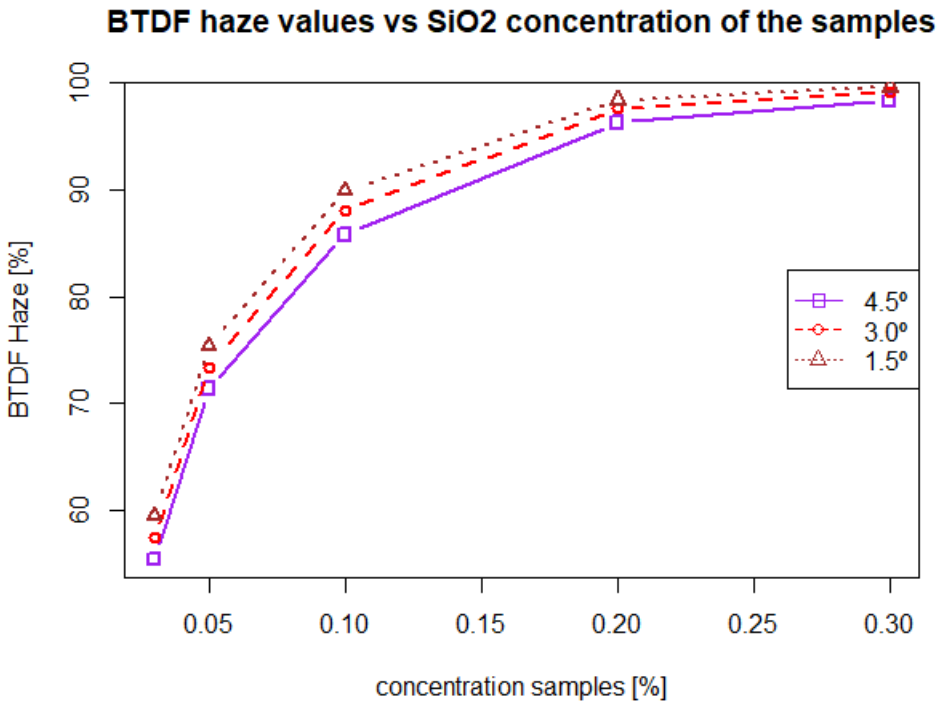


Figure 4.2: *BTDF characterization*

4.1.3 LMK camera

Following the calculations of Busato et al. (2021), measurements with the luminance camera were made, and the LMK haze values were computed using the previously described method for the defined ROIs.

The results are reported in Table 4.3. It can be observed that even though they vary for each air gap configuration, 4mm and 8mm results follow the same pattern. However, LMK haze values obtained for the 0mm configuration are exceptionally low, even for the sample with the highest contaminant concentration (0.3% of SiO_2). A haze percentage of 19.68% is obtained for the most hazy sample which almost corresponds to the value of the lowest contaminant concentration sample measured with an air gap of 4mm or 8mm.

Table 4.3: *Haze values calculated from LMK measurements*

Sample concentration of SiO_2 [%]	Haze [%] at 0mm	Haze [%] at 4mm	Haze [%] at 8mm
0.002	1.56	15.34	21.62
0.01	1.73	26.78	38.11
0.03	2.48	50.86	60.79
0.05	3.35	64.30	73.37
0.1	6.16	82.37	89.28
0.2	11.54	93.82	97.52
0.3	19.68	97.49	99.41

These distinctions are visually depicted and more evident in Figure 4.3. Additionally, it is apparent that the relation between contaminant concentration and haze varies in each case. The rise in haze percentage with contaminant concentration is approximately linear for the 0mm setup, while for the other air gap configurations a logarithmic pattern is observed.

An example of a picture obtained with the LMK during the haze analysis of each sample is presented in Figure 4.4. Since it is an image directly acquired from the setup, it may also somewhat mimic how the observer perceives the evaluated samples when they are placed on the different sample holders. Each image has numbers from 1 to 7 representing the contaminant concentration of the sample from 0.002% to 0.3% SiO_2, resp.

Figure 4.4a shows the case of a 0mm air gap between the target and sample, and it can be observed that the perceived contrast difference between the samples is small, which is consistent with the low haze values obtained for this configuration in Table 4.3. Even so, from sample 3 onward, there is a noticeable loss of contrast when examining the contrast pattern seen through the samples in this configuration; prior to that, these changes are hardly noticeable.

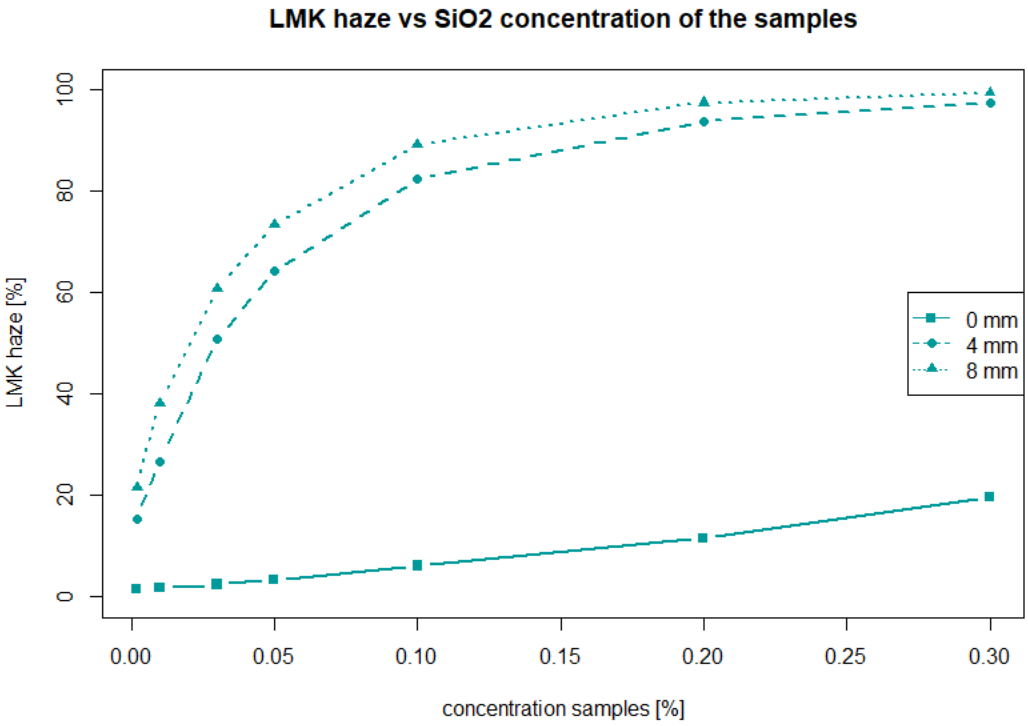
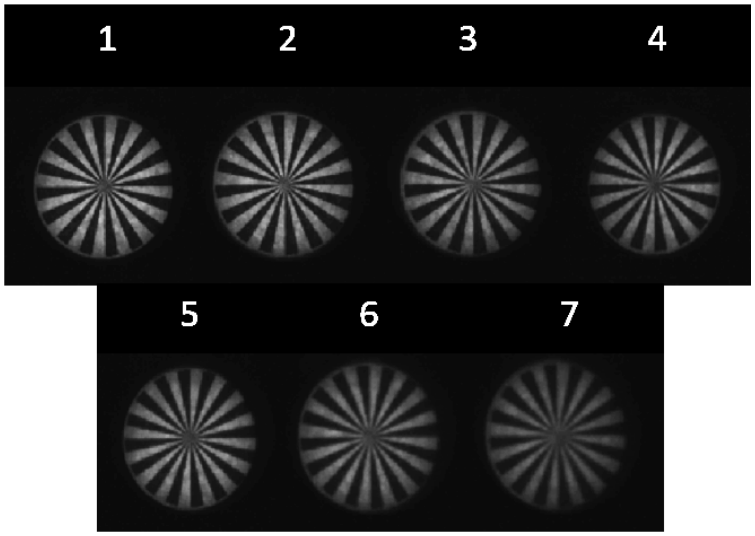
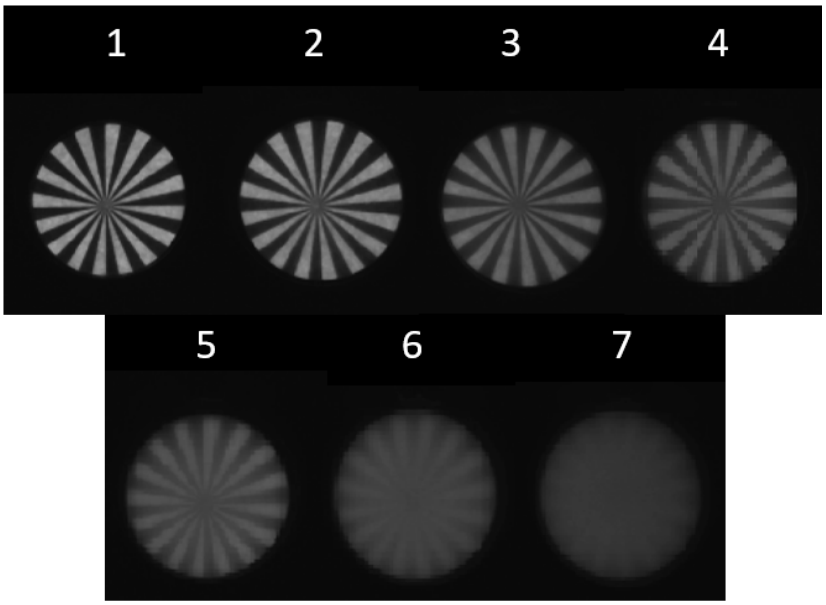


Figure 4.3: *LMK characterization*

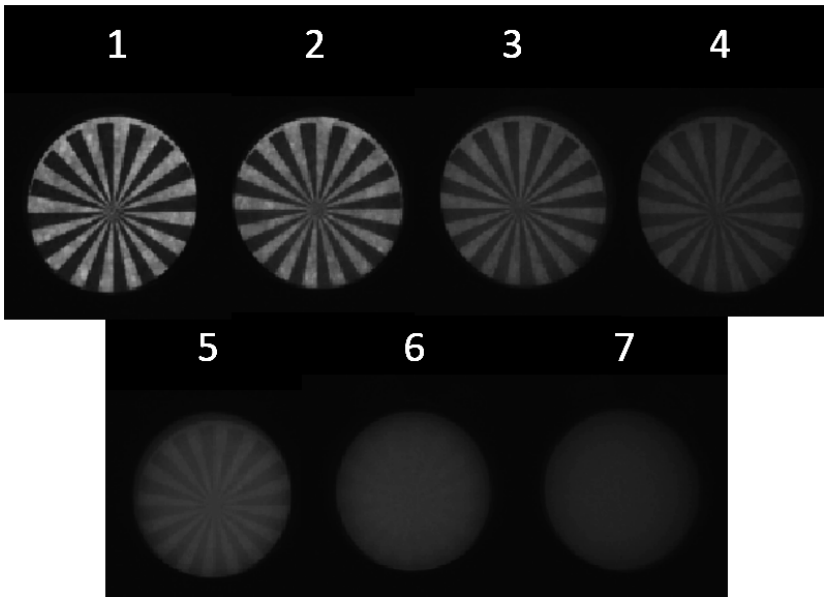


(a) *0mm air gap*

Figure 4.4: *LMK captured images*



(b) *4mm air gap*



(c) *8mm air gap*

Figure 4.4: *LMK captured images*

Perceptual changes produced by the different samples are more clear in Figures 4.4b (4mm air gap configuration) and 4.4c (8mm air gap configuration), in which the reduction of contrast is more discernible as the contamination level increases. In none of the examples there is an obvious distinction between samples 1 and 2, but from sample 2 on, noticeable contrast changes appear. Although, these observed changes are not always discernible, especially between the most hazy samples. Indeed, the difference between samples 6 and 7 in the 8mm configuration (see Figure 4.4c) appears to be almost nonexistent.

4.1.4 Comparison between measurement techniques

A comparison between haze values obtained with the commercial instruments and haze values calculated from the BTDF measurements is presented in Figure 4.5.

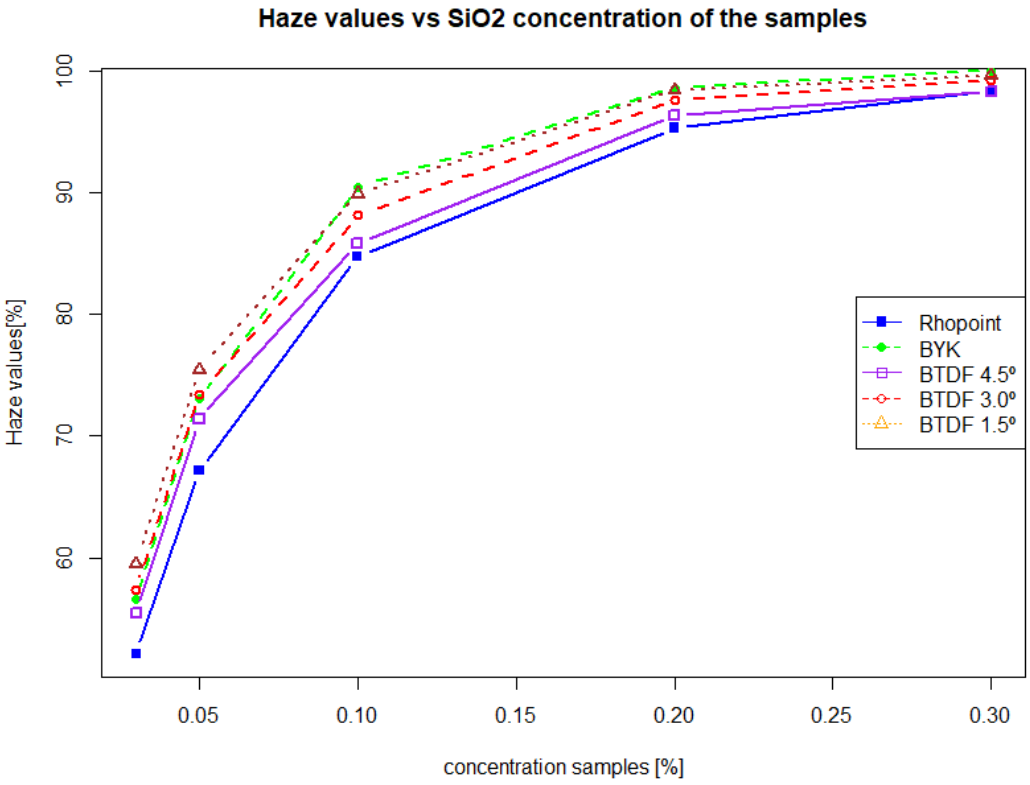


Figure 4.5: *Commercial and BTDF characterization*

Although the measuring techniques employed in each instrument are totally different, the variation in haze across commercial devices can be connected to the

simulated aperture size of the exit port. In fact, the results obtained with the BYK instrument closely match the haze calculated from the BTDF data using a simulated exit port aperture of 1.5°. On the contrary, the Rhopoint instrument results better agree with haze calculated from the BTDF data with a simulated exit port aperture of 4.5°. This explains the previous observation that, from a certain degree of scattering on, haze values obtained with the BYK instrument are greater than the corresponding haze values obtained with the Rhopoint device. Indeed, a larger exit port aperture size reduces the ratio of the diffuse to total transmitted flux.

Figure 4.6 depicts all instrumental haze values obtained for the different samples. As expected from the results in Table 4.3, the 0mm LMK measurements do not correspond with any other measurement result obtained from the commercial and BTDF data. However, the other LMK haze measurements come close to matching the commercial values of Rhopoint and BYK.

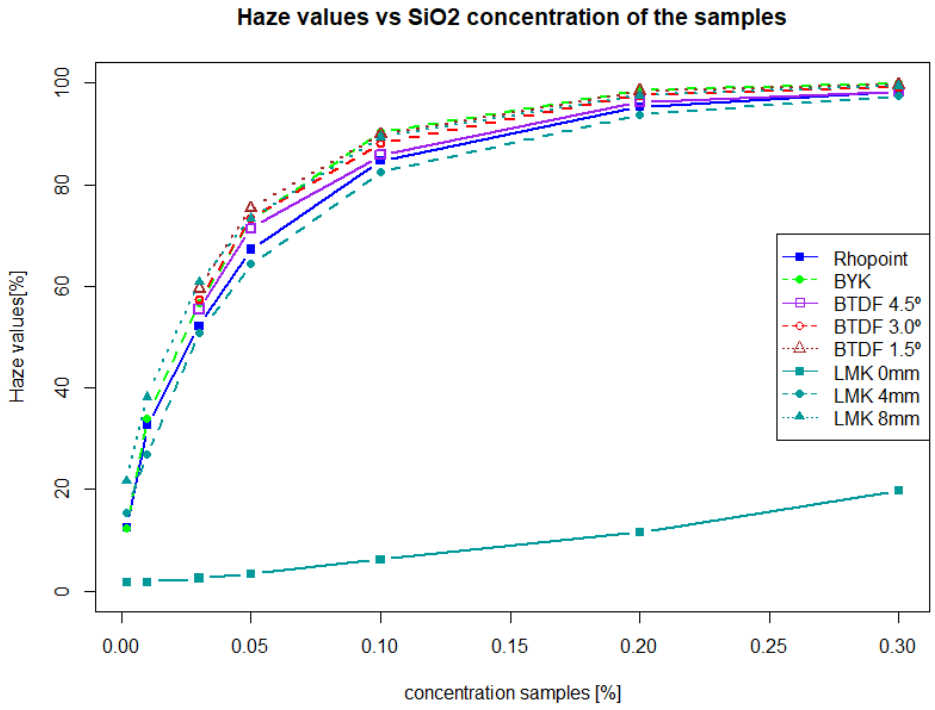


Figure 4.6: *All characterization methods data*

LMK haze at 8mm seems most similar to BYK and BTDF haze at 1.5º, and LMK haze at 4mm seems to better correspond with Rhopoint data and, BTDF haze calculated for a 4.5º simulated exit port aperture.

The first coincidence is consistent with the assertion made in Busato et al. (2021) that haze measurements taken with photographic technique using an air gap of 8mm between the sample and the contrast pattern often correspond with ASTM D1003 haze measurements which is the standard used by BYK instrument. Furthermore, it appears that when the air gap widens, the haze value increases and matches observations made with a narrower exit port aperture or viewing angle. Smaller air gaps are correlated with a bigger aperture size of the simulated exit port, and lower haze values.

4.2 Psychophysical results

This section includes an analysis of the perceptual scales derived from the MLDS procedure as well as observer variability and the impressions and judgments of the observers.

4.2.1 Pilot study

First, a pilot study was conducted to evaluate how many observers were required to participate in the experiment. This was done by continuing the initial work of Prof. Leloup using the configuration of the 4mm air gap previously assessed for 10 observers. Another set of psychophysical experiments with 10 more observers was carried out to have a total of 20 observers for this configuration. The results of the visual scales obtained from these two experiments are plotted in Figure 4.7.

Following a comparison of the MLDS perceptual scale data, estimates and standard deviations for both pools of observers are provided in Table 4.4.

Table 4.4: *Visual scale estimated through MLDS method and standard deviation for 4mm air gap, 10 vs. 20 observers*

		10 observers		20 observers	
Sample number	SiO_2 concentration	Estimated scale	Standard deviation	Estimated scale	Standard deviation
1	0.002	0.00	0.00	0.00	0.00
2	0.01	1.49	0.15	1.20	0.13
3	0.03	3.59	0.33	3.47	0.28
4	0.05	4.94	0.46	5.11	0.42
5	0.1	5.96	0.53	6.19	0.50
6	0.2	7.53	0.62	7.34	0.57
7	0.3	7.62	0.79	8.00	0.71

Due to the closeness of the estimated results and standard deviations of each of them, as well as the subsequent correlations between the optical and visual haze

scales, we proceeded with just 10 observers to speed up the process. After the pilot study, the experiments for the three configurations were performed, and visual scales were estimated.

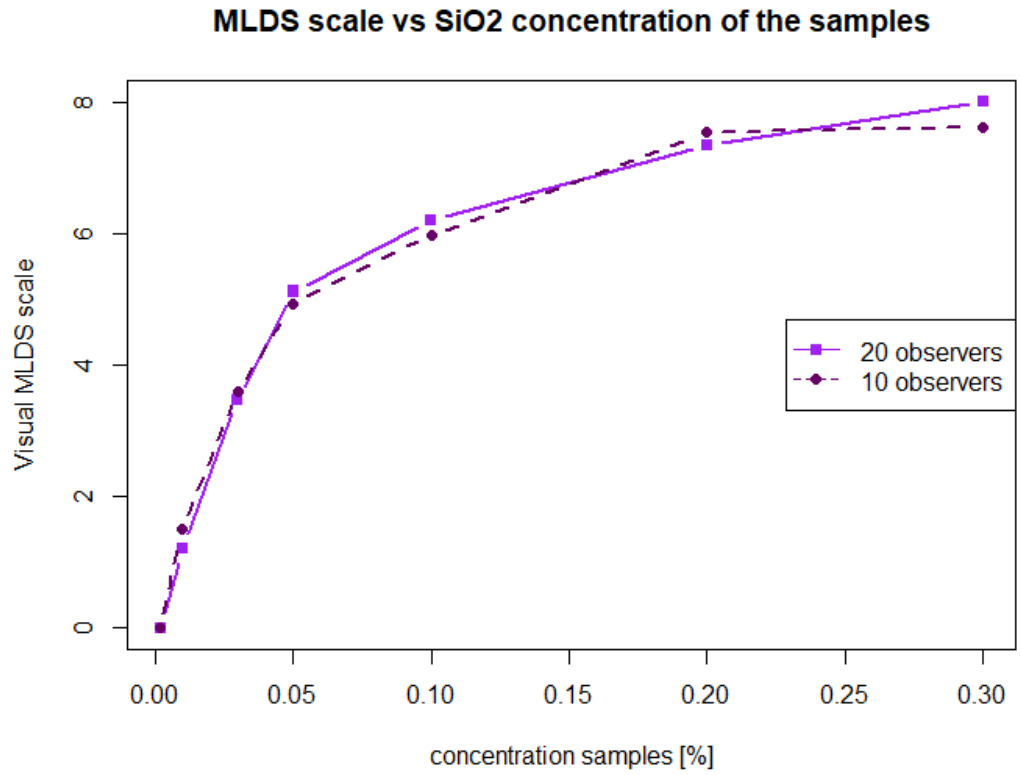


Figure 4.7: *MLDS scales, 10 vs 20 observers*

4.2.2 MLDS scales

Three visual scales were derived from the MLDS procedure for an air gap of 0mm, 4mm, and 8mm, resp. They are presented as a function of SiO_2 concentration of the samples in Figure 4.8. The contrast reduction perceived through the samples increases with the concentration of the contaminant.

Even though this is observed in all the configurations, the 4mm air gap gives the best discrimination between samples with a larger visual scale range. This can also be observed in more detail in Table 4.6 which shows the perceptual scale values associated with each sample used in the experiment and the corresponding SiO_2 concentration. This configuration has a perceptual range from 0 to 7.62 while the others only have a span from 0 to approximately 5.50.

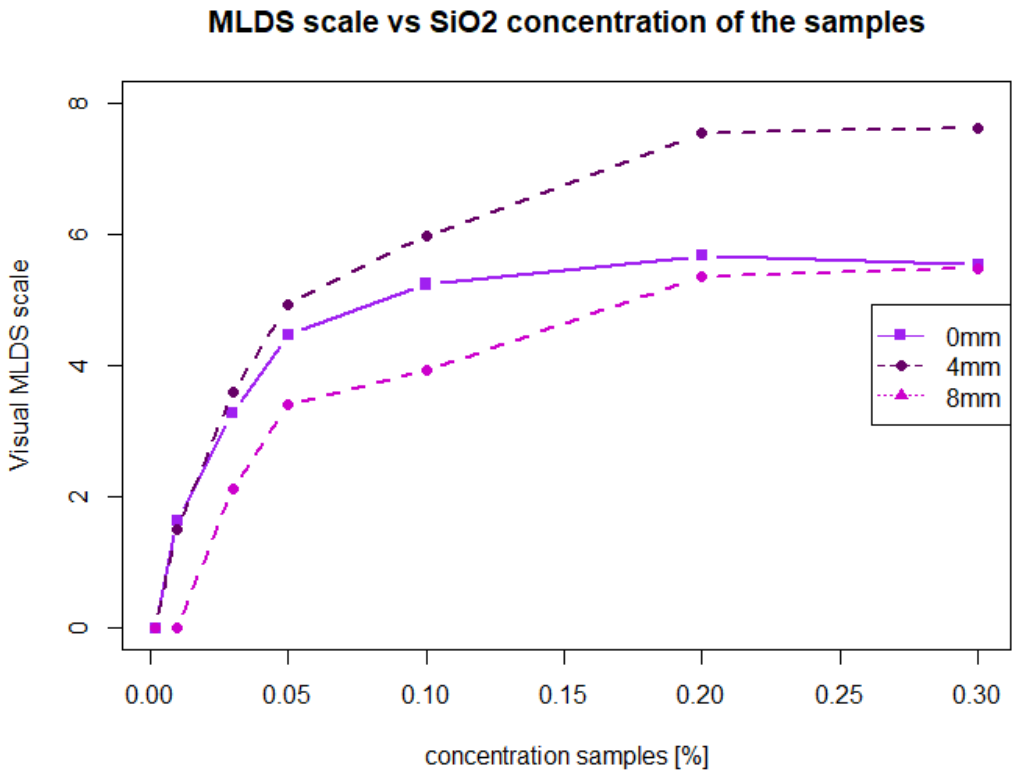


Figure 4.8: *MLDS scales for the three air gap configurations*

Moreover, on the lower end of the scales, a small increase in SiO_2 concentration seems to lead to a significant increase in visual haze. On the contrary, for larger SiO_2 concentrations (from 0.20% SiO_2 on), a further increase of contaminant concentration does not reflect an additional difference in visual haze, a saturation point is observed.

Table 4.5: *Visual perceptual scale estimated through MLDS method and standard deviation for 0mm air gap*

Sample number	SiO_2 concentration	Estimated scale	Standard deviation
1	0.002	0.00	0.00
2	0.01	1.63	0.20
3	0.03	3.27	0.37
4	0.05	4.46	0.45
5	0.1	5.24	0.52
6	0.2	5.67	0.57
7	0.3	5.52	0.75

Information on the scale coefficients derived from the MLDS model for each sample, taking into account setups with an air gap ranging from 0 to 8 mm, is summarised in the accompanying Tables 4.5, 4.6 and 4.7, resp. They include the estimated scale value and the standard deviation.

The estimated scale in Table 4.5 does not follow the progressive increment as the other configurations. It jumps around from samples 6 to 7, the value for sample 7 being lower than for sample 6. This unexpected change could be the result of an unidentified error during the experiments, or because the estimated value of the perceived difference corresponding to these samples, based on the responses given by the observers, has led to an inaccurate result. That could be because the presentation of pairs containing these samples in this specific setup may lead to confusion or ambiguous differences. In addition, a small amount of data can lead to a small bias in MLDS estimates. Since we only have 7 samples and not many repetitions of each pair comparison, that could be another reason for this anomaly.

The standard deviation associated with each estimated value for the three configurations is greater for the haziest samples, indicating a higher uncertainty of these estimations.

Table 4.6: *Visual perceptual scale estimated through MLDS method and standard deviation for 4mm air gap*

Sample number	SiO_2 concentration	Estimated scale	Standard deviation
1	0.002	0.00	0.00
2	0.01	1.49	0.15
3	0.03	3.59	0.33
4	0.05	4.94	0.46
5	0.1	5.96	0.53
6	0.2	7.53	0.62
7	0.3	7.62	0.79

Table 4.7: *Visual perceptual scale estimated through MLDS method and standard deviation for 8mm air gap*

Sample number	SiO_2 concentration	Estimated scale	Standard deviation
1	0.002	0.00	0.00
2	0.01	0.01	0.07
3	0.03	2.11	0.24
4	0.05	3.39	0.36
5	0.1	3.94	0.41
6	0.2	5.34	0.51
7	0.3	5.48	0.67

4.2.3 Observer variability

There is no established method to evaluate the variability among and between observers for psychophysical investigations. Revising the evaluation of different studies, it can be found that these methods may change from experiment to experiment depending on the type of data and analysis method used. Due to this, one of the straightforward approaches based on the Standard Error of Measurement stated in Popovic and Thomas (2017), is employed in this case.

Two measurements of precision, repeatability, and reproducibility, must be defined in order to comprehend observer variability. Repeatability related to intra-observer variability is the ability of the same observer to obtain the exact or similar result when the same sample is presented for the second time. The ability of various observers to reach the same answer is known as reproducibility, related to inter-observer variability.

In this study, these metrics are obtained by computing the standard deviation of each pair of repeated measurements to determine repeatability and the standard deviation of the responses from all the observers for the same presented quadruple in the case of reproducibility. Afterwards, the mean of these standard deviations is calculated to obtain intra- and inter-observer variability values.

The values are reported in Table 4.8. It can be observed, as expected, that in every case the intra-observer variability is smaller than the inter-observer variability. Furthermore, the variability grows as the air gap widens, meaning that the consistency of judgment between observers decreases. Therefore, the best option to evaluate haze with this setup would be 0mm or 4mm, in the end, considering all the facts, the best configuration will be proposed.

Table 4.8: *Intra-observer and Inter-observer variability*

Air gap configuration	Intra-observer	Inter-observer
0mm	0.05	0.16
4mm	0.08	0.20
8mm	0.13	0.25

4.2.4 Impressions and judgments of the observers

Here, the opinions and preferences of the observers based on their judgments and comments during the experiments are described.

The most considerable remark was with the presentation of the pairs featuring samples 6 and 7 in the 8mm configuration. The observers always claimed that the pair with the greater contrast difference was the opposite because samples 6 and 7 in this arrangement held such a high level of haze that they were unable to discern any contrast difference. Checking back the perceptual simulation of the LMK images for this particular scenario in Figure 4.4c, it can be seen that the difference perceived between samples 6 and 7 is barely perceptible, which is consistent with the opinions of the observers. This could imply that samples 6 and 7 no longer yield a suprathreshold difference in the 8mm air gap configuration, reaching the saturation point previously plotted, see Figure 4.8. Therefore, it could be stated that for this arrangement, the MLDS approach to estimate the perceptual scale is not the most effective choice for evaluating perceptual haze.

However, when pairs containing samples 5 and 7 or 4 and 6 were presented against other pairs, the pairs with the haziest samples were consistently chosen to have the higher difference in contrast reduction across the three configurations.

On the other hand, when the pair including the less hazy samples was shown (containing samples 1 and 2), the opposite pair was selected in all configurations as the one with the biggest difference in contrast reduction because samples 1 and 2 do not significantly alter the contrast pattern. The only case where this is not fulfilled is for the 8mm configuration where the haziest samples (samples 6 and 7) create more confusion and less capacity to analyse the contrast difference than samples 1 and 2.

There was a tendency to select the haziest pair of samples as the one starring the greatest difference in contrast reduction in the configuration of 0mm air gap. The exclusions were made mostly in instances when non-consecutive sample pairs were evaluated (i.e. pairs with samples 1 and 4, 1 and 5, 2 and 5, and 3 and 5, against the pair with samples 6 and 7, resp.).

4.3 Correlation between optical and visual haze

Finally, the correlation between the results of optical measurements and of the psychophysical testing, as described in previous sections, is evaluated.

4.3.1 Commercial instruments haze

Figure 4.9 shows the observed correlation between the optical measurements of transmission haze made using BYK and Rhopoint commercial instruments (Instru-

ment haze [%]) and the derived visual haze scales in arbitrary units [a.u.]. With different air gaps, a similar linear correlation between the instrument readings and the visual scales is obtained, it can be observed in Table 4.9. The 4 mm air gap configuration provides the highest correlation to both commercial instrument results, $R^2 = 0.9941$ for Rhopoint, and $R^2 = 0.9893$ for BYK.

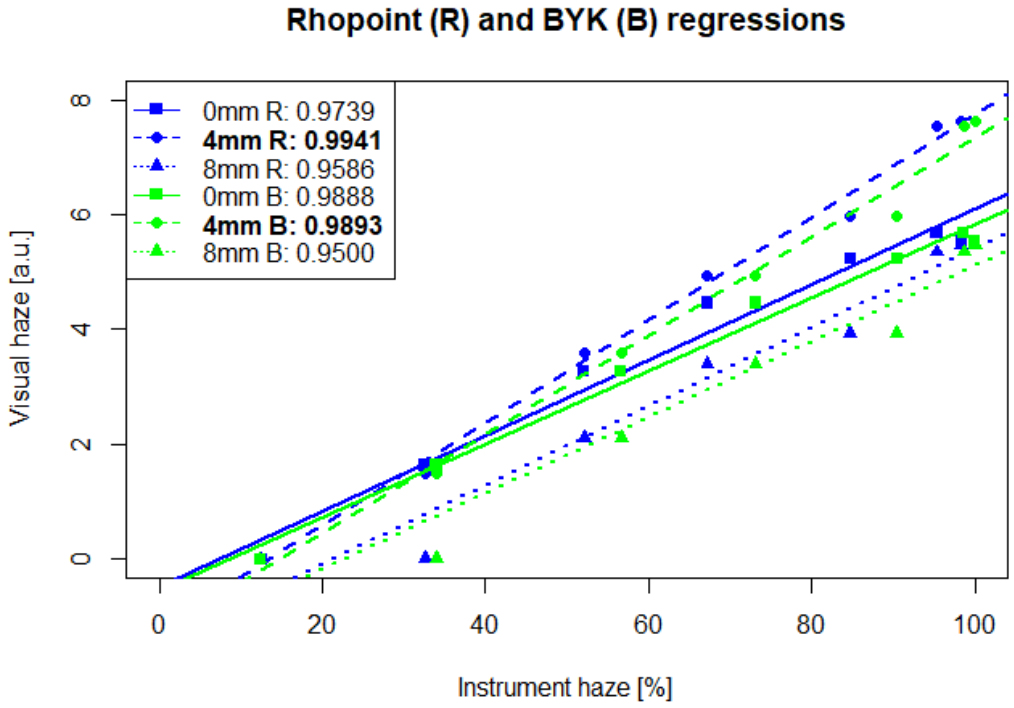


Figure 4.9: *Linear correlation between the results of optical measurements of transmission haze, performed with the two commercial instruments - Rhopoint (blue) and BYK (green) - and the derived visual haze scales in 3 configurations (0mm, 4mm and 8mm air gap, resp.)*

Table 4.9: *Linear regression values from commercial instruments*

Air gap configuration	Rhopoint	BYK
0mm	0.9739	0.9888
4mm	**0.9941**	**0.9893**
8mm	0.9586	0.9500

4.3.2 BTDF haze

Similar to Figure 4.9, the linear relationship between the visual scales under the three assessment conditions (0 mm, 4 mm, and 8 mm air gap, resp) and the haze values determined from BTDF measurements (BTDF haze [%]) for three simulated exit port apertures (1.5^o, 3.0^o, and 4.5^o, resp.) is illustrated in Figure 4.10.

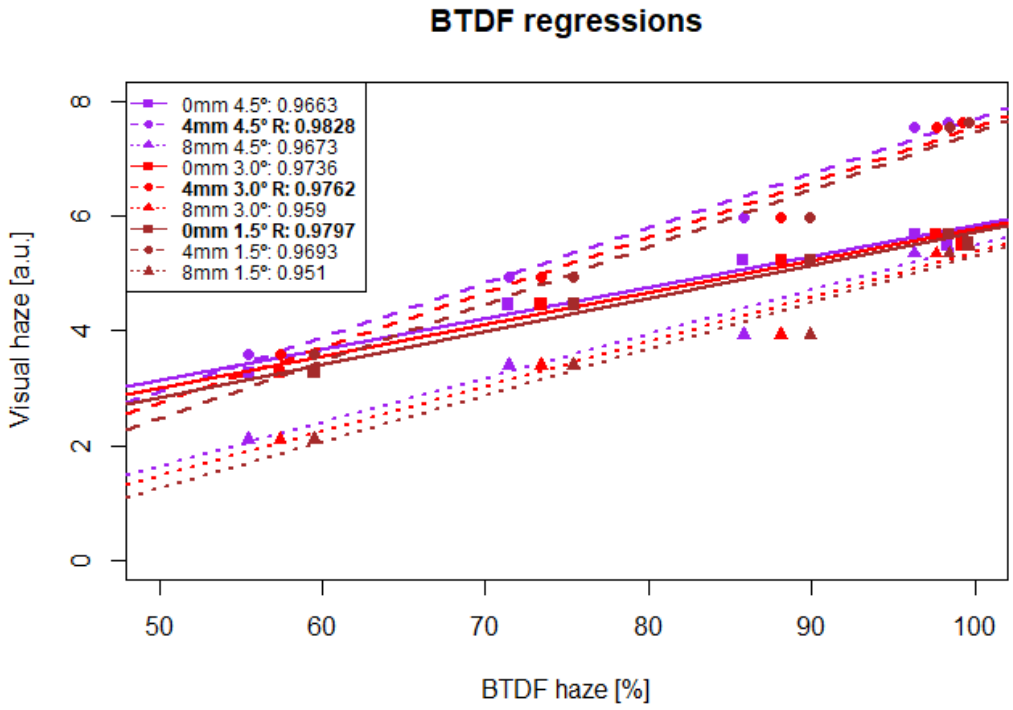


Figure 4.10: *Linear correlation between the haze, calculated from BTDF measurements for 3 simulated exit port apertures - 1.5^o (brown), 3.0^o (red), 4.5^o (purple) - and the visual haze scales obtained under 3 assessment conditions (0mm, 4mm and 8mm air gap, resp.)*

In general, the same results can be drawn. For 2 of the 3 haze scales produced from the BTDF measurements, the visual haze scale of the 4 mm gap configuration has the highest R^2 values, see Table 4.10. The BTDF haze data with the 1.5^o exit port aperture constitutes an exception, for which the correlation is higher for the 0 mm air gap configuration, although the difference with the 4 mm air gap correlation value is small.

Table 4.10: *Linear regression values from BTDF measurements*

Air gap configuration	4.5º	3.0º	1.5º
0mm	0.9663	0.9736	**0.9797**
4mm	**0.9828**	**0.9762**	0.9693
8mm	0.9673	0.9590	0.9510

4.3.3 LMK haze

The linear regression for each of the setup configurations (0mm, 4mm and 8mm air gap, resp.) measuring the contrast loss with the LMK camera to afterwards calculate the haze value can be seen in Figure 4.11 and Table 4.11.

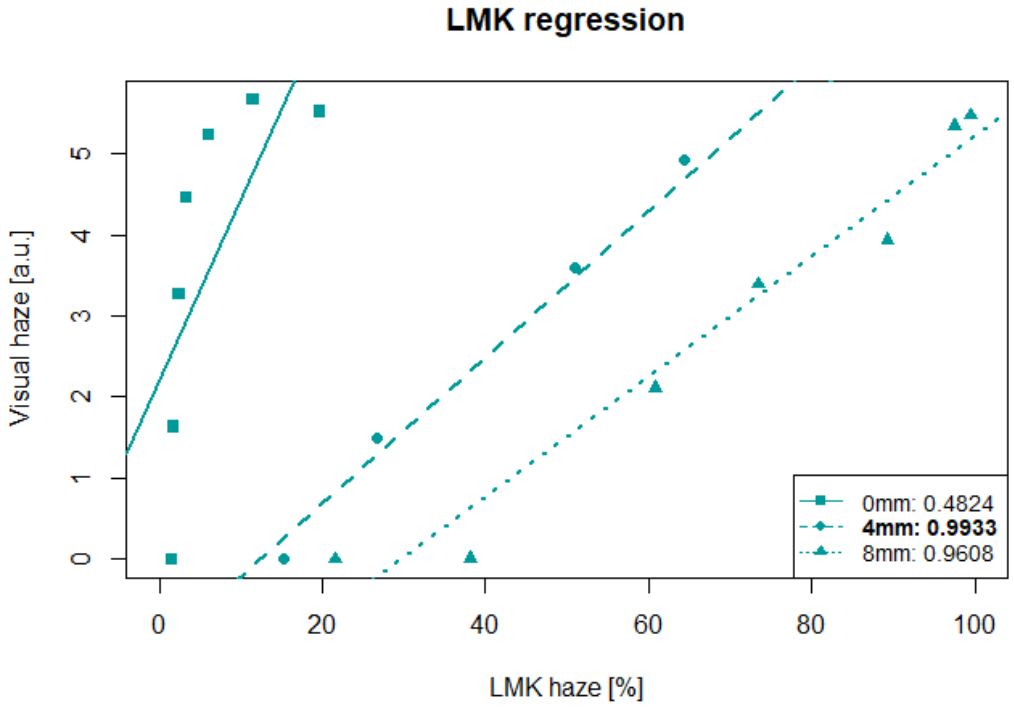


Figure 4.11: *Linear correlation between haze values calculated from LMK measurements and the derived visual haze scales in 3 configurations (0mm, 4mm and 8mm air gap, resp.)*

The 4mm configuration has the highest R^2 correlation value. The condition getting the lowest $R^2 = 0.4824$ is the scenario of 0mm. Because calculated values are so low and distinct from the other methods of assessment, the LMK haze

estimates for this situation do not correspond adequately with the perceptual scale. This different pattern from configuration 4mm and 8mm air gap can be observed in Figure 4.11.

Table 4.11: *Linear regression values from LMK measurements*

Air gap configuration	LMK
0mm	0.4824
4mm	**0.9933**
8mm	0.9608

4.3.4 Comparison between linear regressions

In this section, all the linear regression values between optical and visual haze scales are presented. Table 4.12 shows the values of the R^2 coefficient obtained in each case for the three air gap configurations.

Table 4.12: *Linear regression values for all the cases*

Air gap	Rhopoint	BYK	BTDF 4.5º	BTDF 3.0º	BTDF 1.5º	LMK
0mm	0.9739	0.9888	0.9663	0.9736	**0.9797**	0.4824
4mm	**0.9941**	**0.9893**	**0.9828**	**0.9762**	0.9693	**0.9933**
8mm	0.9586	0.9500	0.9673	0.9590	0.9510	0.9608

Even though different instruments use different techniques, for almost all of them, except from the BTDF measurements simulating a 1.5º exit port aperture, the best correlation value is derived from the 4mm air gap between the sample and the contrast pattern. Rhopoint instrument correlated to the visual scale of 4mm air gap gets the highest value, followed by LMK measurement at 4mm air gap that uses the same methodology to calculate the haze value.

Furthermore, it can be observed that there is not a significant variation between R^2 values of 0mm and 4mm air gap, but the values obtained for the 8mm air gap are always lower (except in the LMK case).

According to the psychophysical analysis of the visual scales, the 0mm air gap does not have a progressive scale, and from the images that mimic the visual perception of haze, the difference between samples is more noticeable in the 4mm air gap configuration. Furthermore, the 8mm air gap configuration seems to not have a suprathreshold difference between samples 6 and 7.

From the observer variability, it was concluded that 0mm or 4mm air gap configurations were the options with higher consistency of judgments among and between observers.

This, together with the findings of correlations between visual and optical data, may lead to infer that the 4mm air gap configuration is preferable to the other configurations for the analysis of haze.

5 Conclusions

This study evaluated the relationship between optical measuring techniques for transmission haze characterisation and visual perception of transmission haze. A sample set comprising 7 polymer samples at different SiO_2 concentrations was developed for this purpose. Haze was measured from an optical perspective with the use of two commercial devices, through BTDF measurements, and deriving the contrast and haze metrics directly from luminance images obtained with a luminance camera.

Even though the measuring techniques employed by the two commercial devices and the BTDF measurements are entirely different, the observed variation in haze for the two methods may be connected to a different aperture size of the simulated exit port. Moreover, the air gap between the sample and the contrast pattern fixed when the luminance images were taken seems to be related to the aperture of the exit port. Results for a larger air gap correspond to a smaller simulated aperture size of the exit port, relatively resulting in a higher amount of scattered light and producing higher haze levels. Consequently, more contrast loss ought to be perceived, as is the case with an 8mm air gap.

Using the MLDS method, three visual haze scales were created from the conducted psychophysical studies. Each scale pertains to a different configuration, corresponding to an established air gap distance of 0mm, 4mm, and 8mm between the specimen and the seen contrast pattern, resp. The best sample discrimination and a larger perceptual scale are produced by a 4mm air gap. For the following reasons, it is also possible to not consider the other configurations as a good way to evaluate haze perception. The configuration of 0mm does not have a correlative scale, and the suprathreshold between samples 6 and 7 appears to be missed in the visual scale of the 8mm air gap configuration.

Linear correlations between the outcomes of the optical measurements and the derived visual haze scales were demonstrated. So far, it appears that a 4 mm air gap provides the best correlation for all the measurement techniques, with the exception of the data from the BTDF measurement at 1.5^{o}. The best correlation values are for Rhopoint ($R^2 = 0.9941$) and the values calculated from the LMK camera measurements ($R^2 = 0.9933$). Both are based on the Michelson contrast

calculation. This could reveal that this image-based straightforward technique is rather helpful for haze evaluation.

Currently, there are no standardised visual evaluation methods for polymers, since the visual perception of haze is a novel field of study that has not been extensively investigated. However, the method suggested in this work is the first attempt at standardisation. It appears to operate correctly and enables to propose of a visual assessment method and protocol that are reliable.

There is a dearth of information and analysis on haze perceptual assessment. Therefore, many additional experiments might be conducted to supplement this study or can be evaluated as future work.

A clear distinction between clarity and transparency in the main standard organizations and evaluation approaches still needs to be defined, along with all visual perceptions of it and the correlation between assessments. Clarity is another transparency feature that could be evaluated in a similar manner as haze. However, due to the lack of clarification in the definition and the time limitation of the master's thesis, it was left out.

The contrast pattern employed in the setup and the colour of the hazy samples are two variables that could also be assessed. The visual performance may alter when the type of target is changed, as indicated in Marasco and Task (2001). A different contrast pattern such as the knife-edge used in Busato et al. (2021) or a simpler one with two plain squares, one black and one white, might affect the perceptual evaluation. In such situations, the Mach bands effect may have an influence by making a sharp step edge's contrast appear uneven and exaggerated as soon as the edges come into contact. On the other side, Carrasco and Siebert (1999) specified that different colours of liquids change the perceived turbidity. Therefore, the same evaluation could be done with solid polymer samples. Although it is more closely tied to chemical composition, it was found in this study and others like Fratini et al. (2006) that the particle size, the materials used, and the fabrication techniques may all affect the perception of haze through the samples. In consequence, different contaminants in the samples may produce different results. In order to get a good performance for the intended usage, it is important to take the composition of the materials into account.

Moreover, a different visual evaluation technique can also yield dissimilar outcomes and relationships between the optical and perceptual scales of haze. Different perceptual scaling methods could also be evaluated and compared to determine the most optimal.

In our experiments, haze is evaluated as a loss of contrast in the visual performance, but other visual processes, including masking and accommodative trapping, which are not discussed here but are defined in Marasco and Task (2001), may also impair this performance.

According to Marasco and Task (2001), veiling luminance, rather than the sample haze values, has a stronger correlation with visual performance. This is due to the fact that the veiling luminance which is scattered from a sample takes into account the illumination and observation geometry, but measurements following the conventional ASTM D1003, do not consider these factors. Then, as additional work, this assertion may also be verified.

Bibliography

American Society for Testing and Materials (2019). ASTM E2387-19, Standard Practice for Goniometric Optical Scatter Measurements. Technical report. (cited on page 9)

American Society for Testing and Materials (2021). ASTM D1003-21, Standard test method for haze and luminous transmittance of transparent plastics. Technical report. (cited on pages 7, 8, and 63)

Bartell, F. O., Dereniak, E. L., and Wolfe, W. L. (1981). The Theory And Measurement Of Bidirectional Reflectance Distribution Function (Brdf) And Bidirectional Transmittance Distribution Function (BTDF). volume 257, pages 154–160. SPIE. (cited on page 19)

Binsbergen, F. L. and Van Duijn, J. (1967). Transparency measurement of plastic sheet and film. *Journal of Applied Polymer Science*, 11(10):1915–1929. (cited on pages 2, 3, and 14)

Busato, S., Kremer, D., and Perevedentsev, A. (2021). Imaging-Based Metrics Drawn from Visual Perception of Haze and Clarity of Materials. I. Method, Analysis, and Distance-Dependent Transparency. *Macromolecular Materials and Engineering*, 306(5):2100045 (1–13). (cited on pages 10, 12, 17, 18, 21, 24, 32, 35, 39, 44, 56, and 63)

BYK-Gardner (2010). Transmission Haze Visual Perception object sample observer. Technical report. (cited on pages 2 and 63)

BYK-Gardner GmbH, . (2015). haze-gard i | Transparency and haze meter. (cited on page 18)

Carrasco, A. and Siebert, K. J. (1999). Human visual perception of haze and relationships with instrumental measurements of turbidity. Thresholds, magnitude estimation and sensory descriptive analysis of haze in model systems. *Food Quality and Preference*, 10(6):421–436. (cited on pages 15, 32, and 56)

CIE (2020). CIE S 017/E:2020. ILV: International Lighting Vocabulary, 2nd Edition. Technical report. (cited on pages 1, 4, 12, 13, and 15)

Dövencioğlu, D. N., van Doorn, A., Koenderink, J., and Doerschner, K. (2018). Seeing through transparent layers. *Journal of Vision*, 18(9):25, 1–19. (cited on page 16)

Fratini, C. M., Esker, A. R., Gibson, H. W., Ward, T. C., and Wilkes, G. L. (2006). Study of the morphology and optical properties of propylene/ethylene copolymer films. Technical report, Faculty of the Virginia Polytechnic Institute and State University. (cited on page 56)

Goos, P. and Großmann, H. (2011). Optimal design of factorial paired comparison experiments in the presence of within-pair order effects. *Food Quality and Preference*, 22(2):198–204. (cited on page 32)

International Organisation for Standards (2023). ISO 12233:2023, Photography — Electronic still picture imaging — Resolution and spatial frequency responses. Technical report. (cited on pages 10 and 32)

International Organization for Standardization (2021). ISO 14782:2021, Plastics - Determination of haze for transparent materials. Technical report. (cited on pages 7, 9, and 10)

Kendall, M. and Babington Smith, B. (1940). On the method of paired comparisons. *Biometrika*, 31(3/4):324–345. (cited on page 33)

Knoblauch, K. and Maloney, L. T. (2008). MLDS: Maximum likelihood difference scaling in R. *Journal of Statistical Software*, 25(2):1–26. (cited on page 30)

Koenderink, J., Van Doorn, A., Valsecchi, M., Wagemans, J., and Gegenfurtner, K. (2017). Eidolons capricious local sign. *IST International Symposium on Electronic Imaging Science and Technology*, pages 24–35. (cited on pages 16 and 31)

Lim, Y.-W., Jin, J., and Bae, B.-S. (2020). Optically Transparent Multiscale Composite Films for Flexible and Wearable Electronics. *Advanced Materials*, 32(35):1907143, 1–15. (cited on page 4)

Loebich, C., Wueller, D., Klingen, B., and Jaeger, A. (2007). Digital camera resolution measurement using sinusoidal siemens stars. *SPIE*, 6502. (cited on page 25)

Maloney, L. T. and Yang, J. N. (2003). Maximum likelihood difference scaling. *Journal of Vision*, 3(8):573–585. (cited on pages 27, 28, 29, 30, 31, and 32)

Marasco, P. L. and Task, H. L. (1999). The effect on vision of light scatter from HMD visors and aircraft windscreens. *Defense Technical Information Center*. (cited on pages 3, 19, 32, 38, and 63)

Marasco, P. L. and Task, H. L. (2001). Measurement of visual performance through scattering visors and aerospace transparencies. In Lewandowski, R. J., Haworth, L. A., Girolamo, H. J., and Rash, C. E., editors, *Helmet- and Head-Mounted Displays VI*, volume 4361, pages 188 – 197. International Society for Optics and Photonics, SPIE. (cited on pages 14, 56, and 57)

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models, 2nd Edn.* (cited on page 30)

Morris, B. A. (2017). *The Science and Technology of Flexible Packaging Multilayer Films from Resin and Process to End Use*, chapter 12. Frictional and Optical Properties, pages 435–462. (cited on pages 4 and 12)

Nicodemus, F. E., Richmond, J. C., Hsia, J. J., Ginsberg, I. W., and Limperis, T. (1977). Geometrical considerations and nomenclature for reflectance. (cited on page 19)

Nomura, S., Ishida, A., Ishigo, E., Okamoto, T., Mizushima, Y., and Nagata, N. (2011). Lace curtain: Modeling and rendering of woven cloth using microfacet BSDF - Production of a catalog of curtain animations. In *ACM SIGGRAPH 2011 Posters, SIGGRAPH'11*, number August, pages 10–11. (cited on page 20)

Pastor, A., Krasula, L., Zhu, X., Li, Z., and Le Callet, P. (2022). Improving Maximum Likelihood Difference Scaling Method To Measure Inter Content Scale. *arXiv:2203.13186*. (cited on pages 31 and 32)

Popovic, Z. B. and Thomas, J. D. (2017). Assessing observer variability: A user's guide. *Cardiovascular Diagnosis and Therapy*, 7:317–324. (cited on page 48)

Rhopoint Instruments Ltd, . (2018). Rhopoint ID Imaging Transmission Appearance Meter. (cited on page 18)

Ross, R. T. (1934). Optimum order for the presentation of pairs in the method of paired comparisons. *Journal of Educational Psychology*, 25(5):375–382. (cited on page 27)

Ross, R. T. (1939). Discussion: Optimal orders in the method of paired comparisons. *Journal of Experimental Psychology*, 25:414–424. (cited on page 27)

Steinberg, S., Sen, P., and Yan, L.-Q. (2022). Towards practical physical-optics rendering. *ACM Trans. Graph.*, 41(4). (cited on page 20)

Task, H. and Genco, L. V. (1985). The measurements of aircraft windscreen haze and its effet on visual performance. Technical report, Air Force Aerospace Medical Research Laboratory, Virginia. (cited on pages 12, 13, 14, 37, and 63)

TechnoTeam Bildverarbeitung GmbH, . (2016). Goniophotometer RiGO801 – 600. (cited on page 20)

TechnoTeam Bildverarbeitung GmbH, . (2017). LMK 5-5 Color. (cited on page 21)

Thorne, B. R. S. W. and Nannestad, I. (1959). SOME CONSIDERATIONS ON THE PHYSICAL SIGNIFICANCE OF TURBIDITY ESTIMATES. *Journal of the Institute of Brewing*, 66(2):175–188. (cited on page 15)

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34:273–286. (cited on page 32)

Wang, H., Zheng, X., and Du, W. (2015). Automatic measurement of spectral bidirectional transmittance distribution function on translucent optical materials. *Measurement : journal of the International Measurement Confederation*, 69:126–133. (cited on pages 19 and 63)

Webber, A. C. (1957). Method for the Measurement of Transparency of Sheet Materials. *Journal of the Optical Society of America*, 47(9):785–789. (cited on pages 3, 4, 12, 14, and 63)

Wimmer, S. and Schwarz, P. (2014). Apparatus and method for measuring optical properties of transparent materials (US Patent No. 8,749,791 B2). (cited on pages 8 and 10)

World Meteorological Organization (2019). *Manual on Codes - International Codes*, volume I.1. Geneva, Switzerland. (cited on page 16)

Wu, G., Li, W., Ni, W., Chen, L., Sun, G., and You, B. (2019). Preparation of Nonspherical Fluorinated Acrylate Polymer Particles by a "Surface Tension Controlling" Method and Their Applications in Light-Diffusing Films. *Macromolecular Materials and Engineering*, 304(8):1900174, 1–12. (cited on page 4)

List of Figures

List of Tables