DELFT UNIVERSITY OF TECHNOLOGY

MSC GRADUATION PROJECT
ME51035

Burnmeister: Analytical Improvements, Calibration, and Workflow Validation for the Pyrotechnic and Explosive Materials Analysis Device

Toward objective combustion characterization and identification of pyrotechnic mixtures

---

Author:
Nigel Moos (5114462)

Forensic Biomechanical Engineering
Supervisors:
Ir. Kim Hutchinson
Dr. ir. Arjo Loeve

Netherlands Forensic Institute (NFI)
Co-supervisor:
Ing. Jacinta Jansen

November 6, 2025

TU Delft Delft University of Technology

Netherlands Forensic Institute
Ministry of Justice and Security

Preface

When constructing this report, the intention was to present the data and visualizations as clearly and accessibly as possible, ensuring that everything would remain legible in black and white format, even for readers with color vision deficiency. In this specific case, however, such an approach was not entirely feasible, as a large part of the report is focused on color representation and detailed color graphs to evaluate the color of combustion flames. Where possible, color differences were visualized with distinctive symbols or variations in the graphs, to also show the color differences in black and white.

Abstract—The increasing number of explosive-related incidents highlights the need for objective techniques to characterize and identify pyrotechnic materials, including explosives. Current techniques, such as the "hot needle test", rely on subjective human observations of flame color, combustion duration, and intensity, which limits reproducibility and accuracy of the data. In this study, the previously developed Pyrotechnic and Explosive Materials Analysis Device (PEMAD) and its analytical workflow were improved, calibrated, and validated. Combustions of compositions with varying grain sizes, oxidizer/fuel ratios, color compositions, and sample volumes were recorded under controlled conditions. The camera was calibrated against reference values from a spectrometer and a light-meter. Color calibration reduced the deviations between the camera and the spectrometer measurements from 27-35% to 2-3%, enabling accurate flame color characterization. Intensity calibration allowed pixel values to be expressed in Lux, providing interpretable and comparable results across all fibers, although with limitations. Validation confirmed that the PEMAD detects differences in combustion duration, flame color, and intensity for compositions with varying properties. Detection limits were established: slower combustions, such as gunpowder, could be reliably measured, while extremely fast combustions, such as flash powder, exceeded the upper limit for accurate peak intensity estimation. The observed sensitivity to small variations in sample volume underscores the need for strict and consistent sample preparation. With larger datasets, the PEMAD could serve as an objective method for the identification and classification of unidentified explosive materials in forensic applications.

Introduction

Pyrotechnics and improvised explosive devices (IEDs) are often used in criminal activities such as attacks, robberies, riots and acts of vandalism [1]. Incidents involving pyrotechnics or IEDs occur daily [2]. The Petroleum and Explosives Safety Organization (PESO) has prescribed standard formulations for various pyrotechnic mixtures for safe use [3]. However, deviations from safe chemical composition, quantities, application, and packaging methods lead to increasing problems. "In the Netherlands, more than half of the explosive casework at the Netherlands Forensic Institute (NFI) relates to pyrotechnics" [1]. This statement was made in 2014, which is concerning as the number of attacks with explosives continues to rise strongly every year, as visualized in Figure 1. This upward trend highlights the urgent need for new forensic research approaches, as traditional methods are claimed to no longer be sufficient [4].

Forensic research plays a crucial role in criminal investigations by providing scientifically validated evidence for the justice system [8]. When an unexploded IED or unknown material is found at or near a crime scene, the NFI and the Dutch police are mainly interested in two aspects: (1) the chemical composition and component ratios in a mixture and (2) the combustion behavior of the unknown pyrotechnic or explosive device to assess the potential threat level of a full explosion. The NFI uses advanced techniques, such as X-ray diffraction (XRD) and X-ray fluorescence (XRF) to determine the exact chemical and mineral composition of the unknown samples. However, this information is not sufficient to classify an unknown substance or

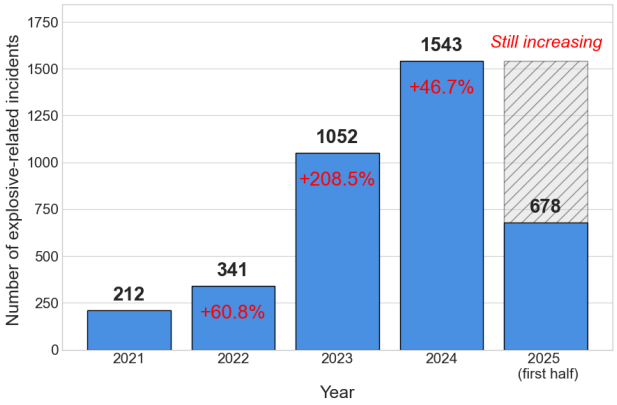


Figure 1: Upward trend for reported explosive-related incidents per year in the Netherlands [4–7]

mixture as pyrotechnic or explosive. That is because these techniques do not provide information about key combustion characteristics, such as flame color, light intensity, and burning rate, which are necessary characteristics to assess the potential threat level [9]. Even when a mixture contains the right components for explosive behavior, it may still fail to detonate at a high rate due to factors such as incorrect ratios between components, which affect the actual burning speed and, hence, the potential danger.

Currently, combustion characteristics such as flame color, light intensity, and combustion duration are determined at the NFI and English forensic laboratories with the so-called "hot needle test" method, where a sample is ignited using a heated needle and the combustion characteristics are observed with the human eye. Although this method is simple and efficient, it has several apparent downsides:

- Human observation of the combustion duration is likely to give varying results between observers or even across repeated measurements, as the assessment is subjective. The duration is currently classified by the observer as: "slow", "relatively slow", "fast", or "extremely fast".
- The perceived color of the combustion may differ between observers, as the observation is man-made and subjective, leading to inconsistent documentation of the combustion color.
- The resulting emitted flame may not be visible to the naked eye or may be too bright for safe and proper observation.
- Human observation of the light intensity is likely to give varying results between observers or even across repeated measurements, as the assessment is subjective. The intensity is currently classified by the observer as: "not bright", "relatively bright", "bright", or "extremely bright".
- The combustion duration, intensity, and color all vary during combustion events, which means that when describing the entire combustion process with a single general characterization, valuable information about its behavior over time is lost.

From the above, it is clear that there is an urgent need for accurate, objective and standardized measurement techniques that minimize the human factor and provide quantified combustion data to strengthen legal evidence and help trace manufacturers or suspects. Therefore, TU Delft and NFI together developed the Pyrotechnical/Explosive Materials Analysis Device (PEMAD) in previous studies [8]. Because previous studies did not perform proper color and intensity calibration or validate the PEMAD workflow to ensure accurate and reproducible combustion analysis, the objective of this study is therefore defined as follows:

Improvement of the design, user-interface, analytical performance, calibration, and validation of the PEMAD workflow to obtain accurate, reproducible, and objective quantification of flame color, combustion intensity, and combustion duration, enabling characterization, differentiation, and identification of a wide range of unknown pyrotechnic and explosive compositions.

Background

2.1 Literature study

To verify and support the design choices of the previous study [8], a prior systematic literature review was conducted to examine objective measurement techniques used to evaluate combustion characteristics in pyrotechnics and explosives, focusing on combustion characteristics such as flame color, light intensity, combustion temperature, and burning rate [9]. The study also highlights how factors such as chemical composition, oxidizer type, metal additives, and packaging significantly affect combustion behavior.

2.2 Prior work and device background

Figure 2 shows the front of the PEMAD, while Figure 3 presents the system with its rear panels removed, revealing the main components. The system uses a standardized and partially automated approach to analyze the combustion of solid (powder or granular) pyrotechnic mixtures. The operator uses a standardized sampling spoon (volume = 0.134 cm^3) to collect a sample and transfers it to the funnel part of the automated sampling system (Figure 2, part 1). The funnel part is automatically raised using a servomotor, minimizing variation in how the material is distributed over the inert ceramic bowl. Samples are ignited by automatically lowering the glow plug into the material. The actuators of the sampling system (servomotor) and for lowering the glow plug (stepper motor) are controlled with an Arduino (Uno R3). The combustion is recorded by two optical systems: a visual camera (Joy-it RB camera with CMOS sensor) equipped with a 3.6 mm lens (VTES-CM3P6MM) for reference footage and a fiber optic camera system (Raspberry Pi HQ camera with an Sony IMX477 sensor) for quantitative analysis of intensity, color, and duration. The cameras are controlled by two Raspberry Pi 4 (RPi) modules configured with a master–slave connection. The fiber camera is connected to six optical fibers (MikroElektronika Duplex Multi Mode Fiber Optic Cable, 1mm core) in a two-rows and three-column configuration, each equipped with a neutral density (ND) filter set (Irix Edge, Gelatin ND Filter Set), ranging in ND factor from 1 to 1/2048, to prevent over- and underexposure across a wide range of light intensities without affecting color.

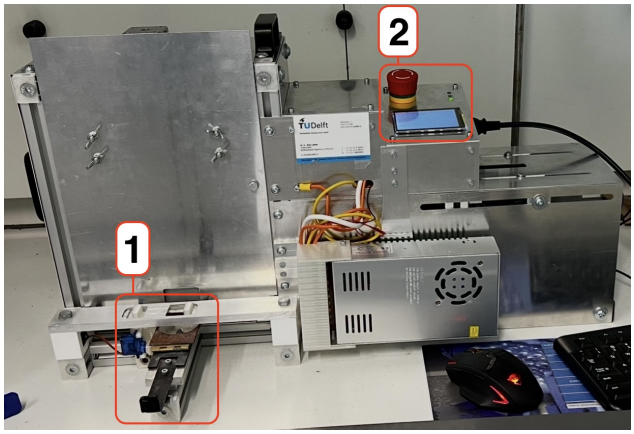


Figure 2: PEMAD system in previous state. [1] Sampling system. [2] LCD screen and emergency off-switch.

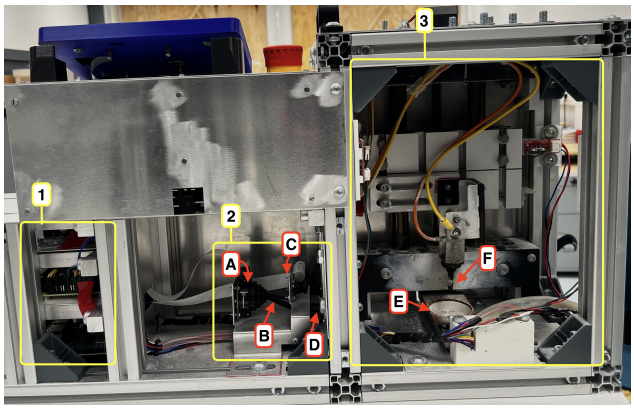


Figure 3: PEMAD system with rear panels removed. [1] Control part containing the two RPi modules and the Arduino. [2] Measurement part with (A) fiber camera, (B) 6 optical fibers, (C) visual camera, and (D) ND filter configuration. [3] Combustion chamber with (E) inert ceramic container and (F) glow plug.

The footage from both cameras is stored in raw h264 format and automatically converted to MP4. Fiber camera footage is processed with a custom OpenCV routine that extracts mean red, green, and blue (RGB) values [0–255] and intensity (from grayscale values) per frame for the region of interest (ROI) drawn around the fibers (see Figure 4).

The system first identifies the optical fiber with the highest measured intensity closest to 225 pixel units as the initial selection for the final result. This value lies slightly below the sensor's saturation point, maximizing

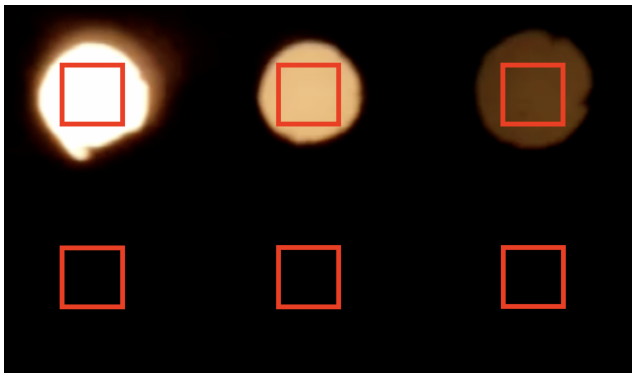


Figure 4: Region of interest (ROI) drawn around fibers (red squares), shown for a combustion event where the intensity reached the first 3 fibers. Within each ROI, the mean RGB and pixel intensity values were calculated over all pixels.

signal strength while avoiding clipping (i.e., loss of detail in overly bright or dark regions). In addition, a lower threshold of 140 pixel units is applied to ensure sufficient signal quality and prevent underexposure, defining an optimal detection range between 140 and 225 pixel units. If the initially selected fiber exceeds 250 pixel units (overexposed) and the subsequent fiber remains below 140, the system automatically switches back to the overexposed fiber, as this prevents excessive signal loss in this case.

The combustion duration is calculated by multiplying the time between frames being captured (1/frame rate) with the number of non-zero intensity frames on the fiber with no ND filter. The results of the combustion characteristics are visualized in one combined graph (Figure 5). The visual camera footage serves as a reference for verifying the results and checking for irregularities during combustion events.

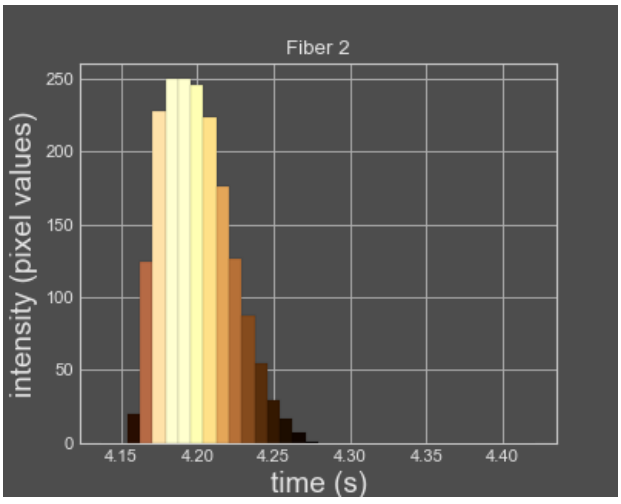


Figure 5: Example PEMAD analysis result of one fiber (fiber 2). The y-axis shows combustion intensity (pixel values), the x-axis shows combustion duration (seconds), and the flame color is represented in each frame (individual bars).

2.3 Camera settings

Accurate imaging is essential for capturing the combustion characteristics. In this application, exposure time/shutter speed, ISO, white balance, and frame rate play a crucial role in ensuring that the recorded data is suitable for accurate and quantitative analysis. Incorrect or variable exposure leads to incorrect/unreliable measurements of color, intensity and combustion duration. Since the automatic camera setting adjustment is not fast enough and fail in this application, the settings were fixed manually as in the previous study [8]:

- **ISO sensitivity:** controls brightness by controlling incoming light on the sensor, where a higher ISO increases brightness but introduces noise.
- **Exposure time:** duration of light collection per frame, where short exposures capture fast moving flame events without motion blur but reduce brightness, while long exposures increase brightness but may cause saturation [9]. The maximum stable exposure time per frame is given by [8]:

$$\text{Exposure time} = \frac{1}{\text{fps}} \times 1,000,000 \quad (\mu s) \qquad (1)$$

- **Frame rate:** frames captured per second, where higher rates improve temporal resolution but may limit exposure time, resolution, and field of view (FOV).
- **White balance:** adjusts red and blue gains to ensure true white.

Methods

Figure 6 provides an overview of the methodological steps taken, which are discussed in detail in the following sections.

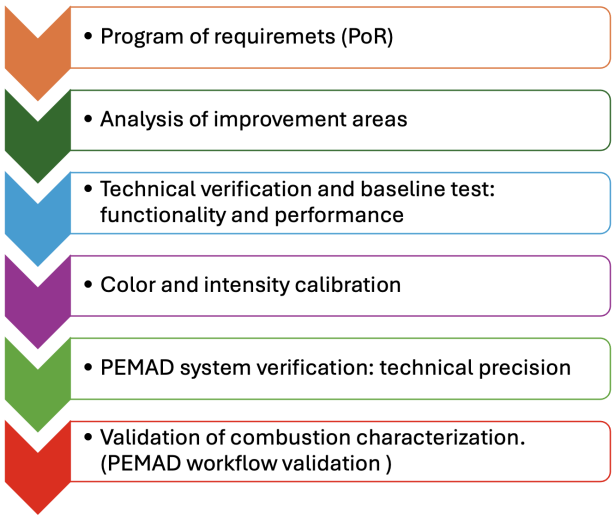


Figure 6: Overview of the step-by-step methodological procedure applied in the PEMAD system development.

3.1 Program of requirements

The main requirements of the PEMAD system, established in agreement with the NFI are listed in Table 1.

Table 1: Main system requirements for the PEMAD, in agreement with the NFI [8]

Category	Requirement
Combustion chamber	Analytical (i.e., cameras, fibers and filters) components protected from heat and debris
	Pressure dissipated to protect components and casing
	Emergency/fast access required (emergency door)
	No significant unburnt residue outside tray
	Automated combustion
	Clean cable management
Light intensity	Measured over time
	Measurement precision/repeatability: ±10%
	Expressed in SI-units
Light color	Wavelength range: 380–700 nm (VIS)
	Wavelength precision: ±10 nm or equivalent
	One spatially averaged color per time frame
Visual camera	Minimum resolution: 360p
	Full area of combustion visible in footage
Optical path	Flame color must not be affected by optics
	Overexposure affects <10% of image surface
	No interference from ambient light
Measurement rate	Full combustion duration must be measured
	At least 5 time frames per combustion event for the fiber camera is required
Sample standardization	Automated and standardized sample provision
	Minimized scatter of the sample in the tray
Igniter	Minimum heat-up temperature: 580°C
	Easy to clean or replace components
	Minimal light contamination
Software/Dataflow	Fixed settings for camera, lighting, and environment
	Control panel interface
Safety	Emergency off-switch included
	No direct view on combustion to protect human eye
	Igniter does not overheat electronics
	Sensor and activator cables connected permanently
User friendliness	Fits in fume hood (1060×690×1560 mm)
	Modular design for easy maintenance
	Easy to clean or replace combustion chamber
	Data export via USB stick or SD card

3.2 Identified areas for improvement

To check the state of the current PEMAD that was delivered by the previous student group, a set of initial inspections and test rounds were conducted. Based on this review and the alignment with the NFI requirements, several improvement areas were identified, including missing aspects such as calibration and validation tests, as listed below:

1) **Overall safety improvements needed:** cable management (exposed/loose cables) and emergency off-switch connection.
2) **Design:** previous student group [8] recommended to improve the graphical user-interface (GUI). The current 3-inch LCD screen limits readability and practicality, as it requires an external keyboard and mouse for operation.
3) **Software:** ensure full compatibility between software and hardware components to guarantee that all components can work together reliably after modifications or upgrades.
4) **Color calibration:** one significant area of improvement that emerged was the current calibration method of the HQ RPi camera. Color calibration was previously performed by matching the perceived colors of the camera with the colors perceived by the human eye. A new, objective, accurate, and reproducible calibration method should be applied.
5) **Intensity calibration:** calibration of light intensity measurements to convert pixel values into standardized units (e.g., Lux), ensuring accurate, comparable, and intuitive/interpretable results across all fibers.
6) **Validation:** the complete combustion characterization workflow (i.e., sample provision and system) still needed to be tested and validated.

To address the identified shortcomings, overall safety was improved in accordance with the NEN-3140 guidelines [10, 11], the user-interface was upgraded with by integrating a 7-inch touchscreen, and the operating system was adjusted to ensure full compatibility. The implementation of these improvements is described in section 4.1: *Device improvements made*.

3.3 Technical verification and baseline test (Pilot)

In order to verify the system's functionality and to have a pre-calibration benchmark, a pilot test was conducted. The test, of which the materials, protocol, and testing conditions are specified in Appendix II, focused on three main aspects:

- **Device performance:** a range of pyrotechnic samples (slow to fast combustions, varying intensities and colors) were tested to verify the device's ability to detect fast intensity changes, capture combustion duration, and distinguish between different flame colors.
- **Sampling and measurement conditions:** samples were collected in the same manner as in the hot needle test, using a fixed-volume sampling spoon rather than precise weighing to mimic the procedure. The functionality of the automated sampling mechanism was assessed visually to confirm correct operation, while the material spread onto the bowl was inspected visually rather than measured quantitatively. ND filters were evaluated to ensure sufficient light attenuation to prevent overexposure, ensuring that at least one fiber consistently provided usable data for analysis. The temporal resolution of the system was assessed by verifying whether at least five non-zero frames were captured during fast combustion events.

- **Functionality and safety:** The device interface, controls, and mechanical components were tested for robustness, user-friendliness, and safe operation.

3.4 Color calibration

Although the PEMAD system operates under fixed conditions (e.g., camera settings and lighting conditions), this alone does not guarantee accurate color representation. Factors such as sensor characteristics and preparation inconsistencies in the optical path (e.g., fiber polishing) can cause deviations between the camera's RGB output and the actual color of the light source.

The calibration process establishes a relationship between the RGB values recorded by the camera and the corresponding spectral reference values measured by a spectrometer. In post-processing, a mapping function (i.e., calibration curve obtained via regression) was derived to minimize deviations from the reference data, thereby improving the measurement accuracy and reliability.

3.4.1 Calibration setup

The previously described PEMAD system was used to collect the color data that needed to be corrected. For reference color data, a StellarNet BLACK-C-SR-50 spectrometer was used in combination with an optical fiber equipped with a CR2 cosine corrector. Cosine correctors are used to collect light uniformly with a 180 degree FOV, to reduce the sensitivity to the direction of light [12]. This ensures that the spectrometer collects the full emission from the light source. According to the Applied lab of the Industrial Design Engineering (IDE) faculty of the TU Delft, the spectrometer was (re-)calibrated by the manufacturer and provided with specific calibration constants, which were applied in the StellarPro V2.2.4 software. A dark reference was applied to remove detector background noise (i.e, dark signal with no light on sensor), and a white reference measurement (i.e., using the light source at full intensity) was taken to establish a baseline for the spectrometer.

As a light source, a Karl Storz endoscope xenon 300 20133120 equipped with an optical fiber was used at multiple intensity levels (i.e. 25, 40, 60 and 80 % of its maximum output) to establish a saturation gradient for each color filter and to examine the influence of luminance (brightness value: Y) on the color representation. The spectrometer-derived tri-stimulus values (X, Y, and Z) were analyzed in both the CIE 1931 XYZ space and converted to the corresponding 2D CIE 1931 chromaticity diagram [9].

Both the PEMAD fiber camera system and the spectrometer were positioned at the same working distance (120 mm) from the light source's optical fiber, corresponding to the distance used in an actual combustion analysis. The optical fiber of the spectrometer and the light source were placed at a height of 65 mm with respect to the base of the PEMAD using custom holders. A Bosch PCL10 line laser was used to ensure that both fibers were at this height and precisely aligned (see Figure 69 in Appendix III.I). The fiber of the light source was aligned with the first fiber of the camera system, using the same laser.

Between the light source and the measurement system, a physical (foil) color filter (140x40 mm) was placed to transmit one specific color toward the measurement system. The filters were positioned 10 mm from the light source, to prevent white light emission past the filter, but also to prevent melting from the heat of the source. Nine physical color filters were used, distributed over the visible spectrum (VIS). The color filter set includes the primary colors red, green and blue and a set of additional colors was used: light blue, light green, orange, purple, yellow and gray. The exact brand and type of the filter set could not be retrieved, but the filters are typically used for camera flash systems.

The calibration experiment was conducted in the PEMAD to mimic the actual combustion environment as closely as possible. The experimental setups for color data acquisition using the spectrometer and the PEMAD are illustrated in Figures 7 and 8, respectively.

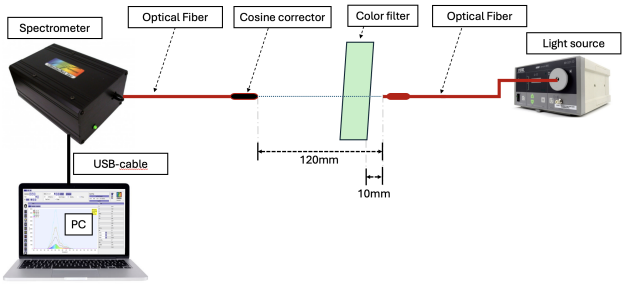


Figure 7: Schematic overview of the experimental setup for color data acquisition with spectrometer

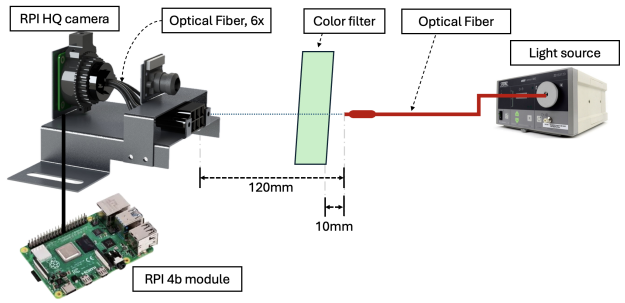


Figure 8: Schematic overview of the experimental setup for color data acquisition with PEMAD system

3.4.2 Spectral measurement and color conversions

Spectral reference data were obtained from the spectrometer as tri-stimulus values (X, Y, and Z) using the CIE 1931 color matching functions (CMFs) with a standard D65 illuminant assumed as reference. Where, the Y coordinate represents luminance, Z represents mostly blue wavelengths, and X represents the remaining wavelengths [9, 13]. From these values, the corresponding chromaticity coordinates (x, y) were calculated to visualize color representation in the CIE 1931 chromaticity diagram. The XYZ values were then

converted into sRGB values using the standard transformation matrix and gamma correction, providing a standardized reference for comparison with the PEMAD camera RGB output. The detailed procedure for spectral data processing, chromaticity conversion, and sRGB transformation is provided in detail in Appendix III.II: *Detailed method for spectral measurement and color conversions*.

3.4.3 PEMAD color data acquisition

Under experimental conditions identical to those used for spectrometer measurements, color data were captured with the RPi HQ camera system in the PEMAD. To not collect more data than necessary and because the light source intensity was too low to be detected through the higher ND filters placed in front of the other fibers, only the fiber with ND = 1 was used. In this way, the camera's sensor is essentially color calibrated, assuming that the ND filters were actually 100% neutral. Images were collected for all color filters and all selected intensity levels of the light source, and the mean RGB values within the ROI were extracted using an OpenCV based Python routine.

3.4.4 Calibration mapping and correction

Calibration correction was implemented entirely in a post-processing routine rather than by adjusting the camera's settings for several reasons:

1) **Consistency:** the camera settings (e.g., exposure time, gain, white balance, and resolution) were already established, fixed and balanced by a previous research group [8] to ensure consistent measurements, frame rates and reliable data acquisition.
2) **Complexity of manual adjusting settings:** matching the camera's RGB output to the spectrometer's RGB reference values is challenging when a wide range of color filters, intensities and camera settings are involved. Manually tuning the settings might improve accuracy for certain colors but reduce it for others, making consistent calibration difficult. Moreover, the camera and spectrometer use different color spaces to quantify color, which means some colors measured by the spectrometer may not be reproducible by the camera [9].
3) **Flexibility:** performing calibration in post-processing allows the use of regression-based methods to systematically use mapping from camera to reference RGB values.

A regression-based calibration model (Python, scikit-learn [14–17]) was used to map the RGB values recorded by the camera to the corresponding RGB reference values derived from spectrometer measurements. Preliminary tests showed that a linear model could not accurately map the camera RGB values to the spectrometer-derived reference RGB values, indicating a non-linear relationship. Therefore, a polynomial regression model was used. To avoid overfitting, ridge regression was used. The polynomial degree and ridge factor α were optimized using a grid search combined with a ten-fold cross-validation. The root mean squared error (RMSE) per color channel (R, G, and B) was used to quantify the accuracy of the calibration, expressing the deviation between corrected camera values and spectrometer references in RGB values. The RMSE was computed per color channel for each filter, averaging the deviations across all five intensity levels. The method is summarized in Figure 9. A more detailed step-by-step procedure is provided in Appendix III.III: *Detailed method for calibration mapping and correction*.

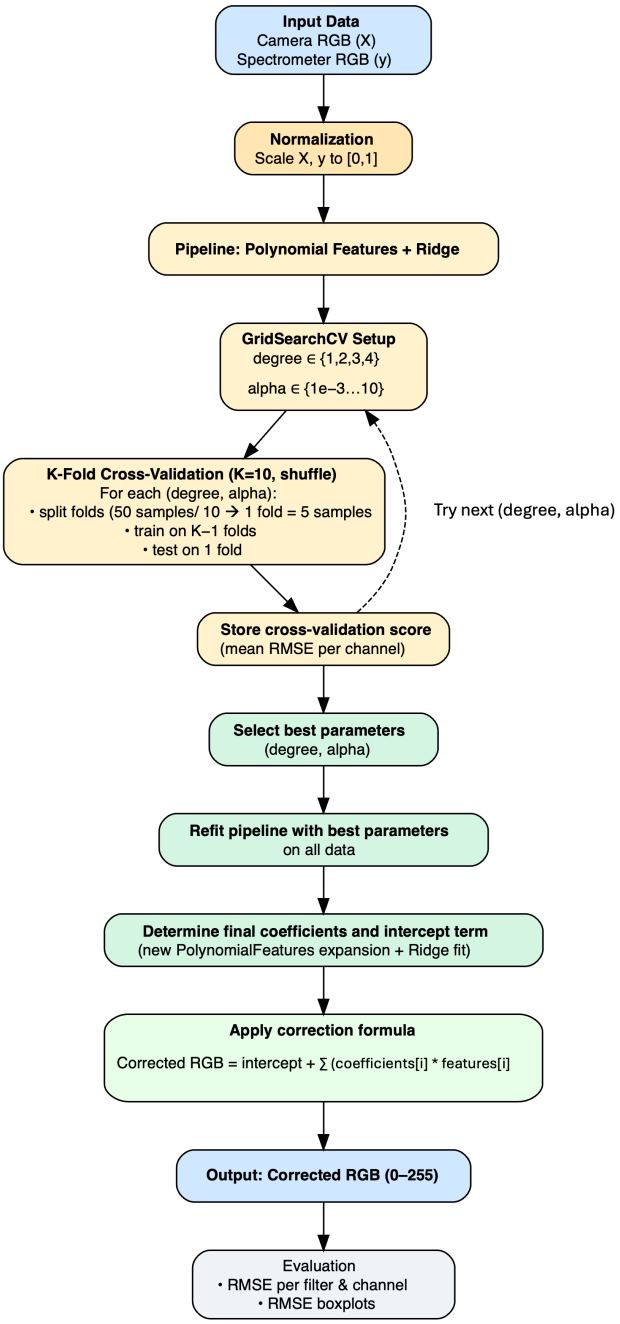


Figure 9: Flow chart of the regression analysis routine: mapping camera RGB values into corrected RGB values that align closely with the spectrometer RGB values

3.5 Intensity calibration

The objective of the light intensity calibration is to establish a quantitative relationship between the pixel

intensity values recorded by the camera's sensor and the corresponding light intensity expressed in SI units (Lux). The pixel values, ranging from 0 to 255, do not have units and depend on the camera's sensor characteristics, settings, and the environment. For instance, a pixel value of 150 may correspond to different intensity values when measured with another sensor, using different camera settings, or using a different fiber or ND filter. Therefore, without calibration, these values do not provide a standardized measure of light intensity. Only relative measure between samples measured by the same system under the same conditions.

3.5.1 Calibration setup

The experimental setup used to capture light intensity data with the PEMAD was almost identical to the color calibration setup (Figure 8), and is shown in Figure 10. The only difference is that the color filter was removed.

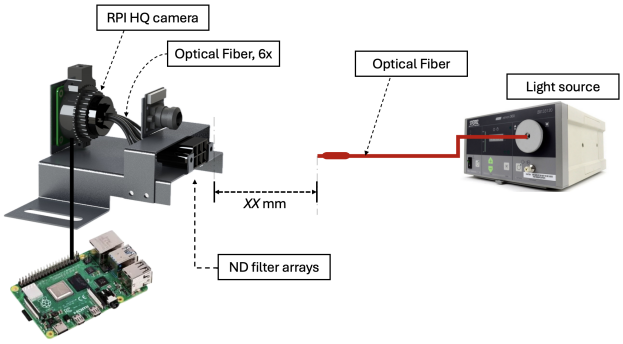


Figure 10: Schematic experimental setup for light intensity data acquisition with PEMAD system

The optical fiber of the light source was precisely aligned with each individual fiber of the camera setup using the same Bosch line laser. Each camera fiber, equipped with a specific ND filter array, was individually calibrated to account for the varying ND transmittance factors. For each fiber, the light source intensity was adjusted to achieve five target pixel intensity values (approximately: 10, 70, 125, 180, and 240), covering the full range of the camera sensor's light sensitivity from near underexposure to almost overexposure, with three intermediate pixel intensities to ensure that the step size is not too large.

Due to the limited power output of the light source, it was not possible to achieve high pixel intensity values at a constant distance, especially for fibers with higher ND transmittance factors. Therefore, the working distance between the light source and the fiber of the camera was varied to achieve the target pixel values. The distances used for this were: $190, 158, 83, 21, 14, 8,$ and $1\ mm$, with decreasing distances used to increase the light intensity reaching the camera's sensor. In Figures 10 and 11, this distance is represented by variable XX.

A custom Python script was created to convert the captured image to grayscale using OpenCV. Within this grayscale image, ROIs corresponding to the fiber positions are drawn, and the mean pixel intensity value is extracted for each ROI.

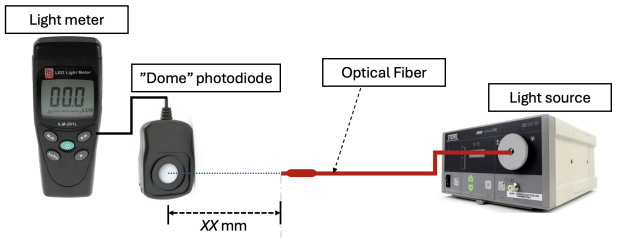


Figure 11: Schematic experimental setup for light intensity data acquisition with Light meter

For the reference intensity measurements, a similar setup was used, illustrated in Figure 11. In this configuration, an RS PRO ILM-201L Lux meter was positioned at the same location as the fibers of the fiber camera system, ensuring the exact same distance $XX\ mm$. The light source's optical fiber and the dome of the photodiode were aligned with the line laser. At each distance and for every corresponding light source intensity level, the illuminance in Lux was recorded. These measurements serve as the reference standard for calibration.

3.5.2 Mapping pixel intensity to illuminance (Lux)

To convert the pixel intensity values of each fiber to illuminance (in Lux), a polynomial regression model was created for each individual fiber. The models were fitted using the measured PEMAD pixel intensity data and the corresponding illuminance measurements from the light meter at multiple light source intensities (data Table 15 in Appendix IV.III). Different polynomial degrees ($d \in \{1, 2, 3, 4\}$) were tested to determine the best fit. Model accuracy was evaluated using the RMSE between the light meter illuminance value and the illuminance predicted by the regression model based on the PEMAD pixel intensity data. The simplest model (i.e., lowest degree) that achieved an error within 1% of the maximum measured illuminance was selected. Therefore, each fiber has its unique calibration curve that maps the measured pixel intensity values to illuminance (Lux), taking into account the ND filters placed in front of each fiber. The detailed conversion method, equations, and model evaluation steps are provided in Appendix IV.II: *Detailed method for mapping pixel intensity to illuminance (Lux).*

3.6 PEMAD system verification: technical precision

The technical precision of the PEMAD was evaluated to assess the repeatability of its measurements under constant conditions. Specifically, the aim was to verify whether the PEMAD produced consistent color and intensity measurements when the light conditions remained unchanged between test repetitions. Twelve test repetitions were conducted using a stable light source (Model: W598B rechargeable work light, with constant white light and optional flashing modes in blue and red), which was connected to an external power supply

to minimize battery discharge. The light source was positioned at the same distance as during the combustion measurements and was assumed to maintain constant output and flickering intervals during testing. To assess the consistency of the duration measurements, the flickering mode was used. The duration of individual flashes and the intervals between flashes were determined from the corresponding number of detected frames. Precision was assessed by calculating the standard deviation (SD) and relative standard deviation (RSD) between repeated measurements.

3.7 PEMAD workflow validation

This practical validation was aimed to check whether the workflow using the PEMAD system can be used to reliably and objectively characterize and distinguish solid pyrotechnic and explosive compositions based on three combustion characteristics: flame color, intensity, and combustion duration. For each combustion characteristic, relevant performance characteristics, such as detection limits, selectivity/specificity, robustness, reproducibility, and repeatability were determined and evaluated against predefined criteria.

3.7.1 Validation approach

According to the "Validatie van Onderzoeksmethode" ("validation of Analytical Methods") guideline document of the NFI [18], *"This is a newly developed method whose performance characteristics have not yet been established. This method requires full validation. This means that all relevant performance characteristics must be determined and the results must be tested against predefined criteria"*. The performance characteristics selected include, detection limit, selectivity/specificity, robustness, repeatability and reproducibility. Table 2 provides an overview of these characteristics, including definitions and relevance for the validation of the PEMAD workflow.

However, for this validation, reproducibility tests related to person dependency were excluded. In agreement with NFI experts, it was concluded that human influence is minimal, as most tasks are automated or standardized. Small variations may occur during sample preparation, but this effect is already covered in the robustness and selectivity tests, making additional person dependency experiments unnecessary.

To systematically evaluate the performance of the PEMAD measurement workflow, predefined acceptance criteria for each performance characteristic were established and are detailed in Table 16 (Appendix V.IV). In addition, a primary objective of this validation is to determine whether the results provided by the PEMAD enable differentiation between different compositions.

3.7.2 Materials and experimental setup

All validation experiments were performed using the PEMAD system, described in section 2.2. All measurements were performed in an NFI laboratory, with constant room temperature (± 21°C), and equipped with available safety provisions, including fume extraction and fire extinguisher. Figure 12 shows the PEMAD in the fume hood.

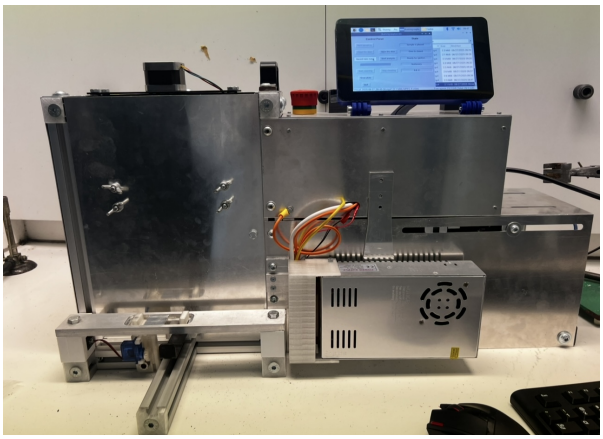


Figure 12: Experimental setup used for validation tests: PEMAD placed in fume hood at the NFI laboratory.

All mixtures, their composition specifications, test IDs (used throughout the rest of this article to refer to specific tests), and the corresponding performance characteristics are summarized in Table 3.

Validation tests were performed using a selection of solid pyrotechnic and explosive mixtures in powder or "flake/granular" form. To evaluate the selectivity/specificity of the PEMAD for three relevant properties, three separate tests were conducted to determine whether the system detects shifts in combustion duration, flame color, and intensity across a range of pyrotechnic mixtures with varying compositions and properties (see Table 3):

- Three oxidizer/fuel ratio variations:
  - Ratio A: 40/60 (m/m%).
  - Ratio B: 50/50 (m/m%).
  - Ratio C: 80/20 (m/m%).
- Three different grain sizes, which were established using scanning electron microscopy (SEM):
  - Grain size "fine" (G-fine) with average diameter and length of 620 μm and 767 μm, respectively (see Figure 13).
  - Grain size "medium" (G-med) with average diameter and length of 709 μm and 1.21 mm, respectively (see Figure 133, Appendix VII.V).
  - Grain size "coarse" (G-coarse) with average diameter and length of 1.07 mm and 1.41 mm, respectively (see Figure 134, Appendix VII.V).
- Five different color emitting compositions, obtained from (category 1) Bengal torches:
  - Red, green, blue, orange, and purple.

To evaluate the robustness for sample volume variations, tests with three starting volumes were performed, including 90%, 100%, and 110% of the standardized sample volume of 0.134 cm^3. Finally, repeatability (precision) of measurements was evaluated for selected compositions, expressed in the standard deviation (SD) and the relative standard deviation (RSD) across repetitions.

Table 2: Overview of performance characteristics defined for the validation of the PEMAD workflow.

Performance characteristic	**Definitions derived from the NFI guideline document [18]**	**Relevance for the PEMAD workflow**
Detection limit	Lowest and highest intensity and combustion duration that can be reliably detected.	Establishes the range of the PEMAD system which can be detected by testing mixtures such as gunpowder (slow, low-intensity) and flash powder (fast, high-intensity).
Selectivity/specificity	The ability of the PEMAD system to distinguish changes in combustion behavior that are caused by the sample composition itself (e.g., oxidizer/fuel ratio, color producing element, or grain size), and not by unrelated factors.	Shows whether PEMAD can reliably identify shifts in combustion duration, flame color, and intensity results when sample composition or material properties change. This helps identifying and distinguish different pyrotechnic compositions.
Robustness	The insensitivity of a method to (intentional or unintentional) variations in the conditions and method variables during implementation. When testing robustness, the effect of small changes in the experimental conditions on the performance characteristics of a method is measured.	Tests whether deviations in the amount of material loaded with the standardized sampling spoon affect the results.
Repeatability	A measure of the spread between measurements obtained using the same method on identical material under the same conditions. The standard deviation is often used as a measure of the spread. The same conditions means: measurements taken simultaneously or shortly after each other by the same researcher, using the same equipment and with the same settings.	Evaluates the precision of the PEMAD workflow under identical experimental conditions (i.e., camera settings, lightning, and environment).
Reproducibility	A measure of the spread between measured values obtained using the same method on identical material under different conditions. As with repeatability, the standard deviation is often used as a measure of the spread. Different conditions include: different researchers, different equipment and/or different times.	Evaluates person dependency: precision of the PEMAD workflow when operated by different researchers. The other different conditions do not apply here.

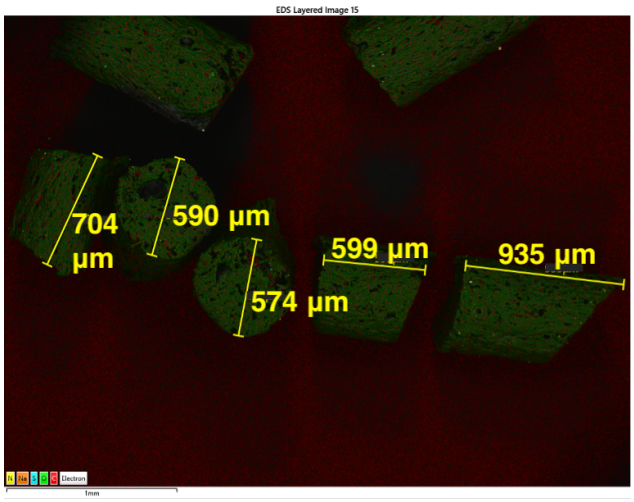


Figure 13: SEM result of material Vihtavuori N310 (G-fine), establishing the dimensions of the grains (yellow measures).

3.7.3 Protocol

The detailed measurement protocol is provided in Appendix VII. The protocol contains instructions on collecting the samples, sample amounts, repetitions, and measurement procedures. All tests were performed eight times (n=8). Where possible, data collected for one performance characteristic were also used to evaluate other characteristics in the post-processing, this increased the efficiency of the process. In addition, a risk analysis was performed prior to the validation tests to evaluate potential operational errors and risks associated with the use of the PEMAD, and to define appropriate preventive or corrective actions in case such situations occurred (Table 17, Appendix VI).

3.7.4 Validation data analysis

The analysis of the validation data consisted of several steps, as described earlier in section 2.2: "Prior work and device background" and summarized in Figure 14. At this stage, the calibration correction functions were applied to the extracted RGB and intensity values. All fiber results were visualized in intensity-time plots, where each frame was shown with its corresponding color (mean RGB value of the ROI for that frame). Data for all validation tests were stored as `.spydata`, and processed with additional python routines. For each measurement, peak intensity (pixel and Lux value), mean R, G, and B color channel values and combustion durations were extracted. To avoid using frames with less useful color information (i.e., clipping: the frames at the start and end of the combustion, with mostly dark or black pixel frames), only frames with intensity $\geq 50\%$ of the peak intensity were selected in the mean color calculation. All the information was collected for all tests and repetitions and stored in a CSV file, and used for visualization, statistical analysis, and interpretation.

Table 3: Composition specifications used in the validation process for each performance characteristic

Performance characteristic	Test ID	Composition specifications	Note
Detection limit:			
Lower limit	DL-1	Gunpowder: Vihtavuori N320	Lower detection limit slow, low intensity combustion Coarse grain size (i.e., flakes)
Upper limit	DL-2	Flash powder *	Upper detection limit fast, high intensity combustion
Selectivity/specificity:			
Oxidizer/fuel ratio	Ratio A	Potassium perchlorate - $KClO_4$ (40%), aluminum - Al (60%)	Determine ratio influence
	Ratio B	Potassium perchlorate - $KClO_4$(50%), aluminum - Al (50%)	Potassium perchlorate, from Sigma-Aldrich
	Ratio C	Potassium perchlorate - $KClO_4$(80%), aluminum - Al (20%)	Aluminum (Pyro 5413H super (<5 μm)), from laboraden.de
Color identification	Red	Strontium Carbonate - $SrCO_3$ (15%), Nitrocellulose - $C_6H_7O_2(ONO_2)_3$ (60%), Ammonium Perchlorate - NH_4CLO_4 (25%)	Identification of red Composition from torch NO.1612
	Green	Potassium Perchlorate — $KClO_4$ (20%), Charcoal — C (10%), Barium Nitrate — $Ba(NO_3)_2$ (20%), Phenolic Aldehyde — $C_{48}H_{42}O_7$ (15%), Magnalium — Al^+Mg Alloy (20%), PVC — $(CH_2 - CHCl-)_n$ (15%)	Identification of green Composition from torch No.1654
	Blue	Aluminum — Al (1%), Potassium Chlorate — $KClO_3$ (25%), Nitrocellulose- $C_6H_7O_2(ONO_2)_3$ (74%)	Identification of blue Composition from torch NO.1613
	Orange	$[C_6H_7O_2(NO_2)_x(OH)_{3-x}]$(60%), Strontium Carbonate — $SrCO_3$ (2%), Cryolite — Na_3AIF_6 (8%) NH_4CIO_4 (30%)	Identification of orange Composition from torch NO.3700
	Purple	Potassium Perchlorate — $KClO_4$ (45%), Magnalium — Al^+Mg Alloy (15%), PVC — $(CH_2 - CHCl-)_n$ (20%) Titanium - Ti (20%)	Identification of purple Composition from torch NO.1651
Grain size	G-fine	Vihtavuori N310	Determine grain size influence Average grain size: diameter of 620 μm and a length of 767 μm
	G-med	Vihtavuori N110	Average grain size: diameter of 709 μm and a length of 1.21 mm
	G-coarse	Vihtavuori N160	Average grain size: diameter of 1.07 mm and a length of 1.41 mm
Robustness:			
Starting volume	V-90	Potassium perchlorate - $KClO_4$ (50%), aluminum - Al (50%)	Determine starting volume influence 90 % volume of standardized spoon
	V-100	Potassium perchlorate - $KClO_4$ (50%), aluminum - Al (50%)	100 % volume of standardized spoon
	V-110	Potassium perchlorate - $KClO_4$ (50%), aluminum - Al (50%)	110 % volume of standardized spoon
Repeatability:			
	R1	Gunpowder: Vihtavuori N320	same composition as DL-1
	R2	Flash powder	same composition as DL-2
	R3	Ratio A	same composition as A
	R4	Ratio B	same composition as B
	R5	Ratio C	same composition as C
	R6	V-90	same composition as V-90
	R7	V-100	same composition as V-100
	R8	V-110	same composition as V-110

*This composition is included for qualitative assessment only, as no validated quantitative information is currently available at the NFI.

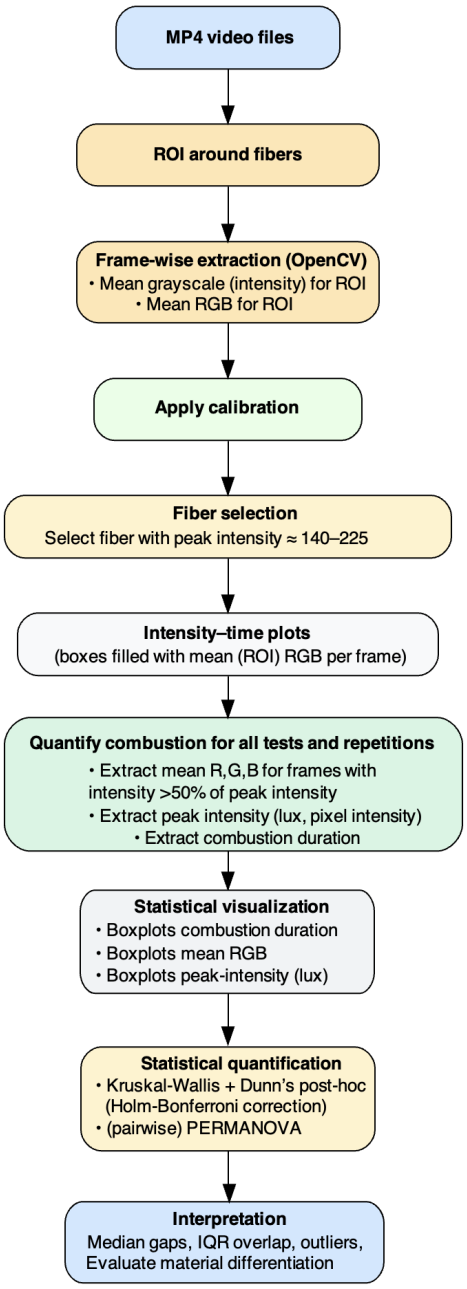


Figure 14: Flow chart of the the data processing workflow applied for the validation data.

3.7.5 Statistics

Because the combustion data were not normally distributed, Kruskal-Wallis tests were used to evaluate the following hypothesis:

- H_0 (null): all compositions within a subset of interest have the same distribution (no differences) for a given metric (i.e., peak intensity, combustion duration or color).
- H_1: at least one composition shows a different distribution for the given metric

When the Kruskal-Wallis test indicated significance, the pairwise group differences were assessed (e.g., composition 1 compared with composition 2 for a single metric, such as combustion duration) using Dunn's post-hoc test for multiple comparisons [19, 20]. With 16 compositions, this resulted in 120 unique pairwise hypotheses (comparison tests). To account for multiple comparisons, the p-values were adjusted using the Holm–Bonferroni correction (α = 0.05). This method effectively controls the family-wise error rate while keeping higher statistical power compared to the classic Bonferroni correction (see Appendix VIII: *choosing multi-comparison correction* for detailed explanation).

Because perceived color is determined by the combination of R, G, and B channels, a multivariate approach was needed to assess whether overall color profiles differed between compositions. Individual channel values alone do not provide enough information to distinguish between the compositions. Therefore, a Permutational Multivariate Analysis of Variance (PERMANOVA) was used on the mean values of all channels combined. This method is also non-parametric [21], and works similarly to the Kruskal-Wallis with Dunn's test. First, a distance matrix was created using the RGB values of the recorded flame colors. Each value in this matrix represents the numerical difference in color between two compositions. The key idea of PERMANOVA is that it assesses the variation between compositions directly from the distance matrix [22]. Euclidean distance was used, as it represents the straight-line distance between RGB values in three-dimensional RGB color space, providing a consistent and interpretable measure of color differences based on the location in that space. In other words, euclidean distance over R, G, B color space is used to define the color similarity [23].

A global PERMANOVA was first applied to the distance matrix to test whether the PEMAD system detected significant differences in the mean RGB color profiles among compositions. After this, post-hoc PERMANOVA tests were performed between all pairs of compositions. Again, this resulted in 120 pairwise comparisons, and the p-values were adjusted using the Holm–Bonferroni correction (α = 0.05) to account for multiple comparisons. The analysis was performed in Python using the `scikit-bio` package for PERMANOVA and `statsmodels` for Holm–Bonferroni correction. The PERMANOVA analysis process is summarized in Figure 16.

Both the uni-variate (Kruskal–Wallis with Dunn) and multivariate (PERMANOVA) analyses were first applied across all 16 compositions as a combined data set to evaluate whether significant differences in combustion behavior could be measured between compositions. In addition, this test assessed whether single or multiple metrics were required for differentiation between compositions. The same statistical methods were applied for the performance characteristics groups separately: robustness (volume variations), selectivity/specificity (color detection, ratio and grain size variations), and detection limit. These comparisons evaluated whether the PEMAD detected significant differences caused by controlled variations in specific material properties (e.g., grain size or ratio).

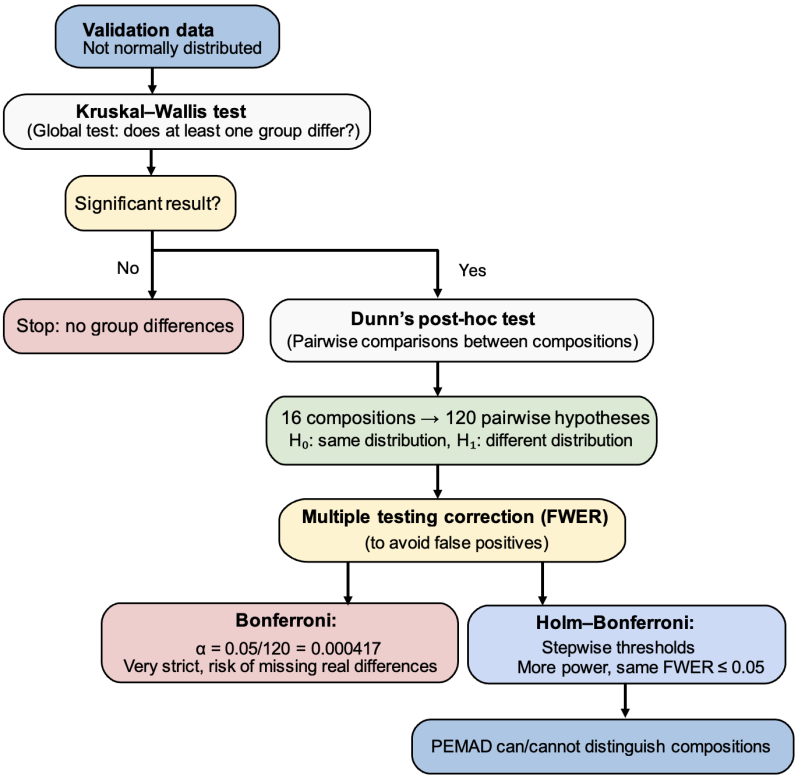


Figure 15: Flow chart of the statistical analysis used for validation (uni-variate)

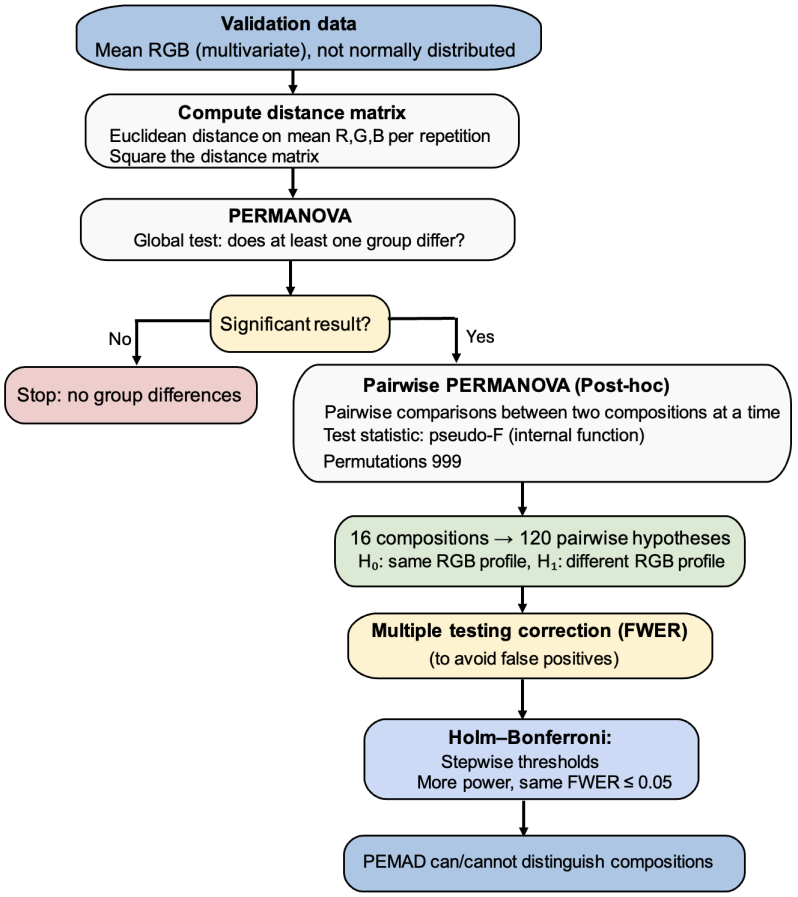


Figure 16: Flow chart of the PERMANOVA statistical analysis used for validation of the multivariate RGB values

Results

4.1 Device improvements made

4.1.1 Overall safety

The inside of the device is now more fireproof as all cables are insulated and permanently connected. The emergency off-switch functions properly and is installed according to NEN-3140, with correct grounding connection via a standard grounded plug, see Figures 60 and 61 in Appendix I.

4.1.2 Design and software

Figure 17 shows the CAD model of the redesigned control unit with the implementation of the new touch display. The screen can be tilted ($\theta \in [20°, 140°]$). The screen is self-locking at all positions within this range, eliminating the need for additional tools. Figure 18, shows the realization of the newly designed control unit. The operating system was adjusted to the 2019-09-26 Buster OS. This setup supports both the touchscreen and the older HQ Raspberry Pi camera hardware, allowing 120 fps recording with the full FOV of the camera sensor, while remaining compatible with all required libraries and system components.

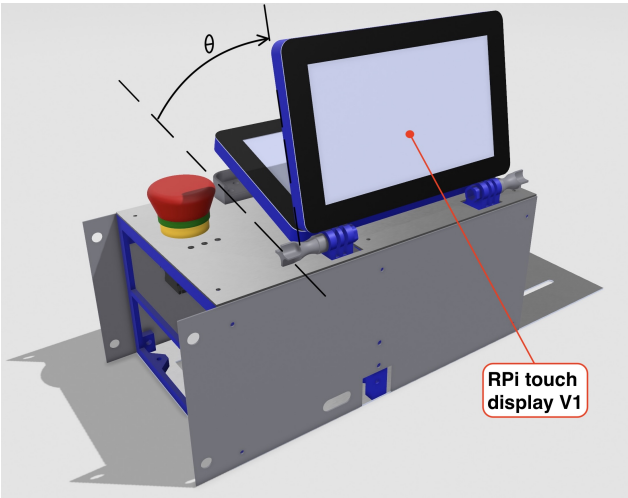


Figure 17: CAD model of the redesigned control box with the 7-inch RPi touch display V1. Where theta (θ) indicates the angle at which the screen can be tilted.

4.2 Technical verification and baseline test (Pilot)

The pilot test conducted at NFI confirmed that PEMAD provides usable, objective, and quantified combustion data for a range of pyrotechnic mixtures and functions almost as intended. However, two main technical issues were identified and addressed. (1) Index errors occasionally occurred, caused by faulty (empty) measurements, which froze the system and required a reboot. This issue was resolved by adding automatic file checks before analysis. In case of a faulty combustion measurement, a pop-up warning now appears, indicating that the measurement must be retaken. (2) Synchronization issues between visual and fiber cameras due to

Figure 18: Implementation of the redesigned control unit with the RPi touchscreen display. Showing the assembled and operational setup corresponding to the CAD model in Figure 17.

timing mismatches in the master–slave connection and insufficient buffering time for file conversion. This was corrected by adding sufficient delay and synchronizing both RPi modules ensuring that recordings start at the same time. For very fast flash powder the number of detected non-zero frames was sometimes fewer than five. Some flame colors also appeared unrealistic at low intensities: this issue was addressed through color calibration, as described in the following sections. The pilot report, including objectives, data, visualization, and discussion is provided in Appendix II.

4.3 Color calibration

The color calibration results presented in this section verify that there is a deviation between the PEMAD color perception and the measured spectral data from the reference spectrometer. In this section, only the results obtained using the blue filter are shown, the additional color data set with the other eight color filters is provided in Appendix III. In addition, the corrected RGB values are provided, showing that the PEMAD was properly calibrated to perceive the colors more accurately, closely matching the spectral reference data.

4.3.1 Spectral measurements

Of the spectral power distribution (SPD) curves obtained with the StellarNet spectrometer across all nine color filters at the five intensity levels of the light source, each SPD showed a unique spectral curve that corresponds to the physical color filter used. The curves confirm that the filters effectively consist in specific

parts of the visible spectrum: each filter has its own unique wavelength range corresponding to the specific color filter. The SPD for the blue filter is shown in Figure 19.

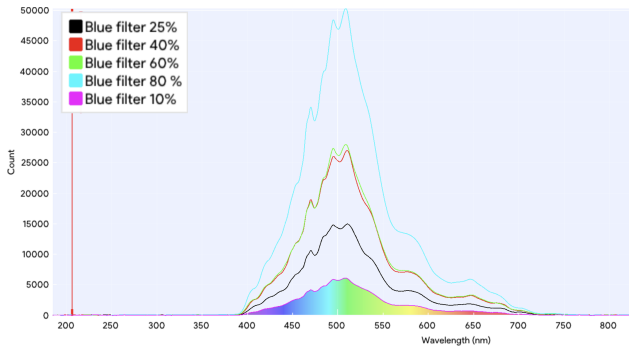


Figure 19: The full visible spectrum measured by the spectrometer through the blue filter for all intensity levels of the used light source. Where the y-axis represents the intensity in counts, and on the x-axis the corresponding wavelength (nm) is shown. Each individual graph represents a different intensity level of the light source, increasing from 10% to 80% of its maximum.

4.3.2 *XYZ color space and CIE 1931 chromaticity diagram visualization*

The obtained X, Y, and Z values for each filter and intensity level were plotted in the standardized CIE XYZ color space (as used by the spectrometer) to visualize the measured colors. Because this color space depends on luminance (Y), the data points are distributed throughout the entire space, illustrating how color saturation changes with increasing light intensity. The XYZ color space with all blue filter color points is shown in Figure 20a, with results for all other filters provided in Appendix III.

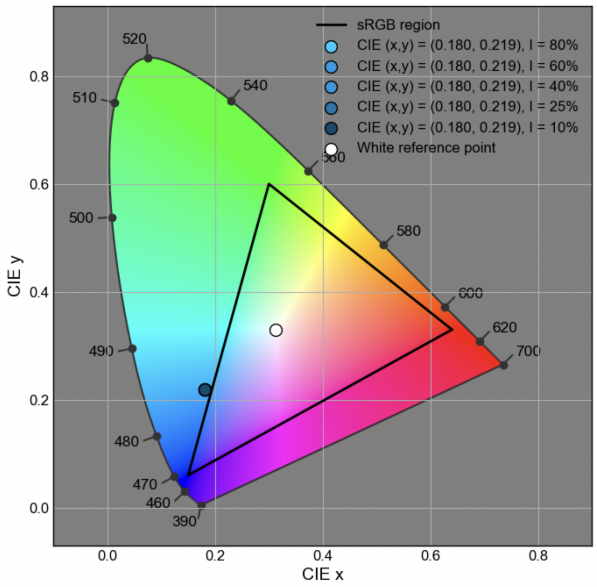
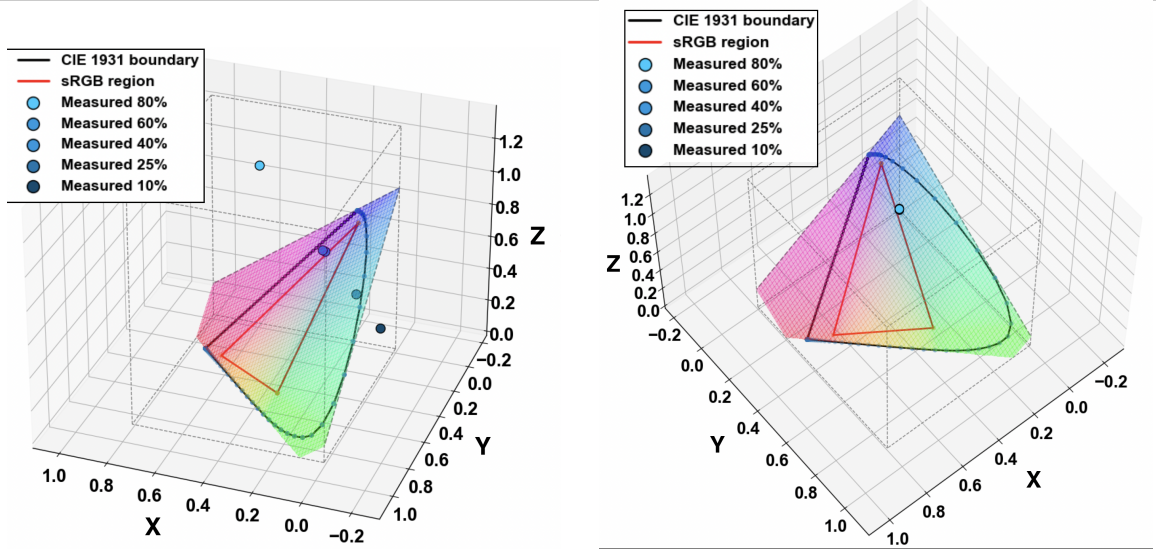


Figure 21: CIE 1931 chromaticity diagram for blue colored light. Where the sRGB region is represented by the black triangle.

When the view is rotated perpendicular to the chromaticity plane ($X + Y + Z = 1$ [9]), differences in luminance are no longer visible as the colors align in a single chromaticity point, as illustrated in Figure 20b. This viewing angle corresponds to the orientation of the CIE 1931 chromaticity diagram in the XYZ space. The CIE 1931 chromaticity diagram was made using the chromaticity coordinates (x,y), indicated with the black "horseshoe" shaped line, in Figure 21. This 2D projection



(a) Color visualization in XYZ color space

(b) Color aligned in one point

Figure 20: Visualization of light transmitted through the blue light filter in the 3D XYZ color space, at light source intensities of 10%, 25%, 40%, 60%, and 80% of the maximum. Where the sRGB region is represented by the red triangle and the CIE 1931 chromaticity diagram with the black "horseshoe" shape.

is less complex and more intuitive [9]. Therefore, this is the most widely used standard for visualizing and evaluating colors [9]. The blue-filter measurements across all intensity levels overlap closely, confirming consistent color detection.

4.3.3 RGB values obtained with spectrometer

The RGB values derived from the spectrometer data using the blue filter are presented in Figure 22.

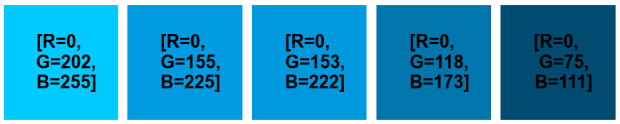


Figure 22: RGB values derived from the spectrometer data for the blue color filter at all light source intensity levels (80, 60, 40, 25, and 10%, from left to right).

4.3.4 RGB values obtained with uncalibrated PEMAD

Figure 23 shows the (uncalibrated) RGB values recorded by the PEMAD camera system for the same blue filter and intensity levels of the light source.

4.3.5 Calibration: color correction

In figure 24, the calibration mapping result for the blue color filter is visualized. It can be seen that the (uncalibrated) camera RGB values deviate strongly from the spectral reference, for the green and blue channels, where the camera underestimates intensity. After applying the correction model, the corrected RGB values align much more closely with the reference values across all intensity levels. This confirms that the correction successfully reduces the deviations in the camera and produces values that better represent the true color.

The RMSE distributions for the red, green, and blue color channels across all filters, before and after calibration, are shown in Figure 25. Before calibration, large

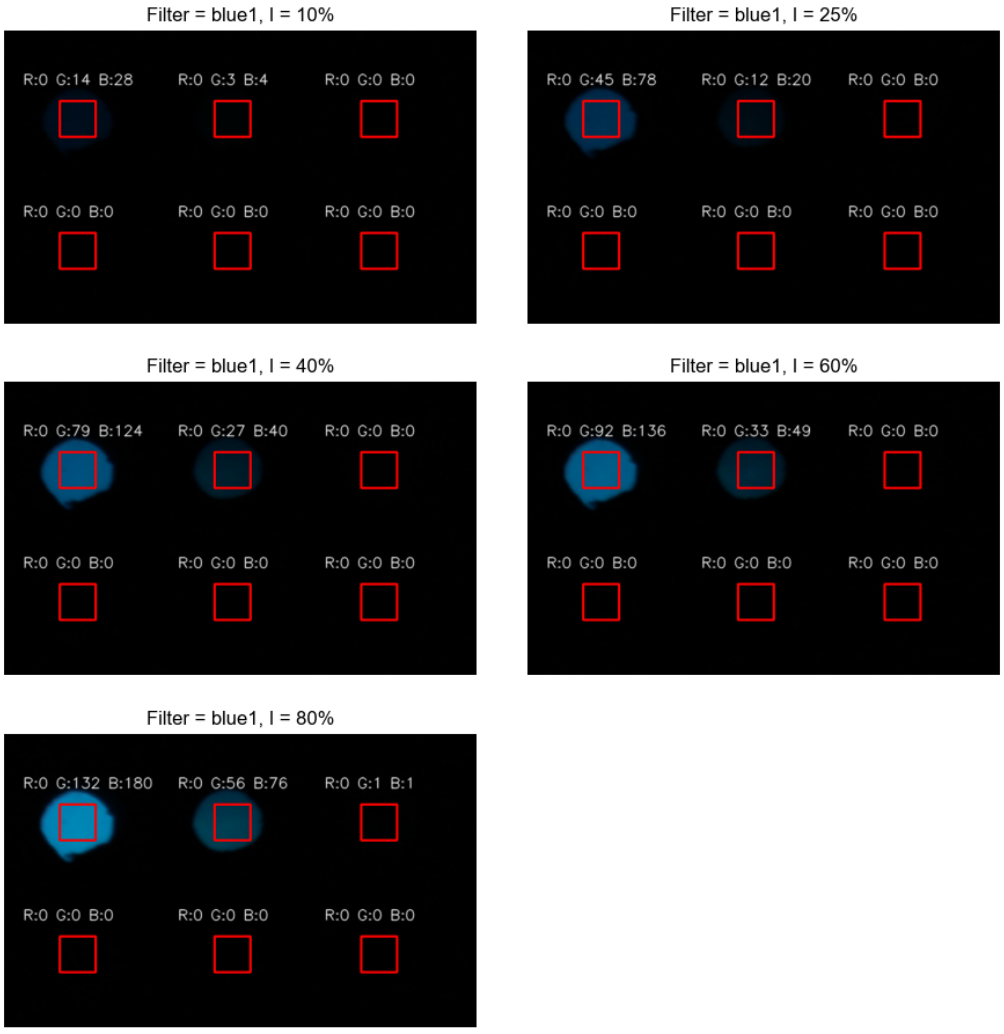


Figure 23: RGB values for blue color filter at all intensity levels of the light source captured with the PEMAD system. In this calibration only the top left fiber (fiber 0) was used for further calibration.

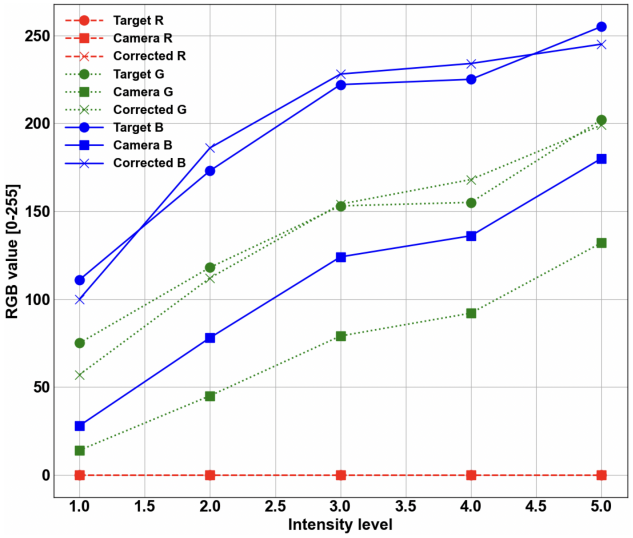


Figure 24: Calibration mapping for blue color filtered light: (PEMAD) camera captured RGB vs spectral reference RGB (spectrometer) vs corrected RGB for all color channels (R, G, and B) at the different intensity levels of the light source (10, 25, 40, 60 and 80 %, indicated on [1-5] range, respectively)

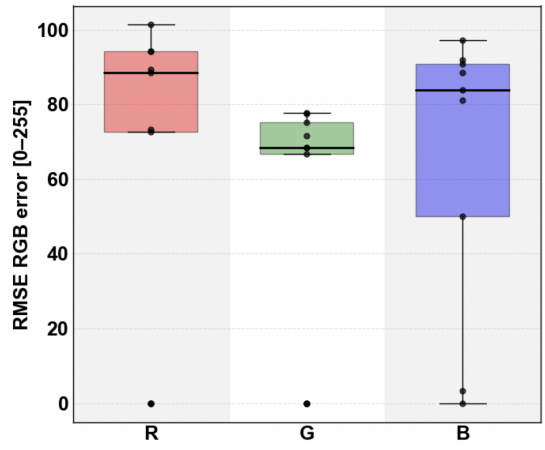
deviations were observed between the spectrometer and PEMAD camera data, with median errors of 88.4 (red), 68.5 (green), and 83.8 (blue) in RGB pixel values. These correspond to errors of approximately 27–35% of the RGB scale. After calibration, the median RMSE values decreased drastically to 5.9 (red), 6.4 (green), and 6.4 (blue), which corresponds to 2–3% of the RGB range (summarized in Table 4).

Table 4: Median RMSE for the color channels across all filters, showing the deviation between PEMAD camera and spectrometer RGB values before and after calibration of the PEMAD. Values are expressed in pixel values on the [0–255] RGB scale, with percentage errors with respect to the full RGB scale shown in parentheses.

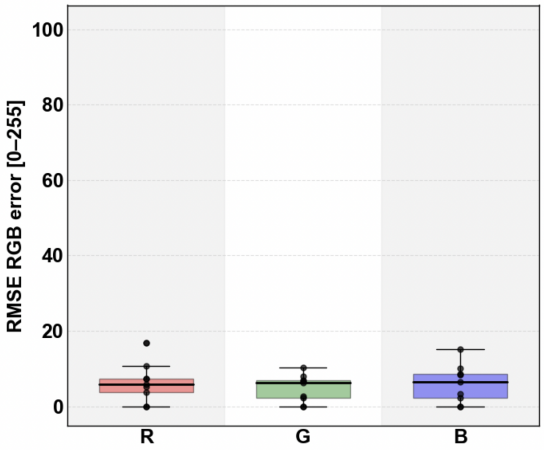
Channel	Before calibration RMSE	After calibration RMSE
Red	88.4 (34.7%)	5.9 (2.3%)
Green	68.5 (26.9%)	6.4 (2.5%)
Blue	83.8 (32.9%)	6.4 (2.5%)

Figure 27, shows the effect of the color calibration. Where Figure 27a and 27b provide the results of the combustion with test ID: Blue (Table 3) for 2 repetitions before calibration and Figure 27c and 27d show the result of the same combustion analysis after applying color calibration.

Moreover, in Figure 27a it can be seen that the low-intensity combustion before calibration appears dark, almost as a black flame, which does not realistically represent the emitted blue color (see Figure 26). After calibration (27b), the output plot captures the blue channel more strongly, resulting in a combustion color that is visually closer to the actual combustion's blue color.



(a) Before correction



(b) After correction

Figure 25: RMSE distribution across all filters for the red (R), green (G), and blue (B) color channels, showing the deviation between PEMAD camera RGB values and spectrometer reference RGB values before (a) and after (b) correction.

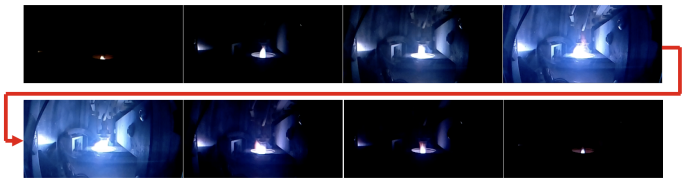
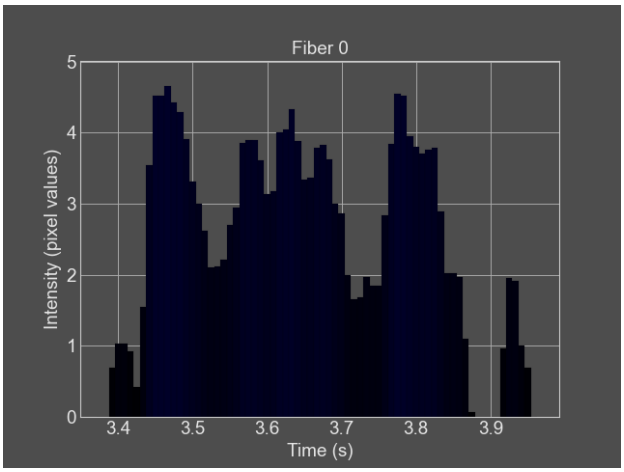


Figure 26: Visual reference footage: still frames extracted from the MP4 recordings of the blue combustion event (Test ID: Blue, repetition 4). The sequence can be read like a film tape and illustrates color and intensity changes over time. Only 8/64 frames were selected to visualize the reference here.

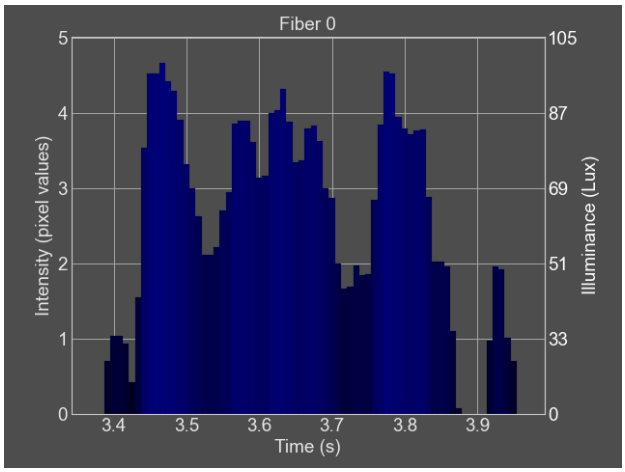
4.4 Intensity calibration

Table 15 in Appendix IV.I provides the measurement data acquired for the intensity calibration, and was used for the regression analysis.

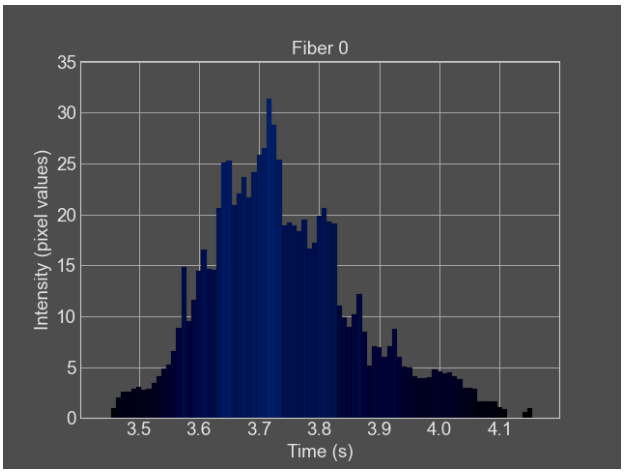
For fiber 1, this resulted in the calibration curve shown in Figure 28. Polynomial functions of different



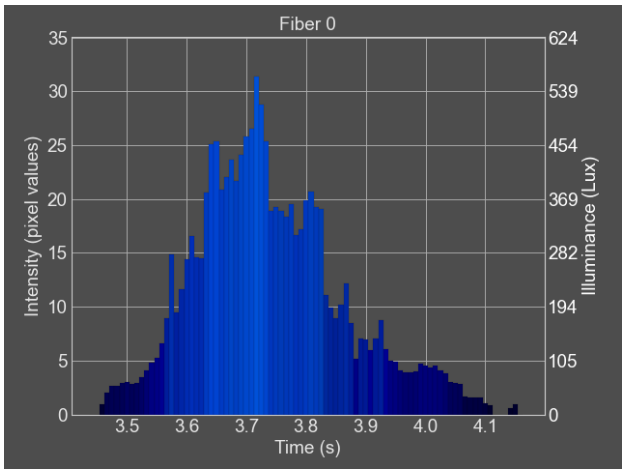
(a) PEMAD outcome of low intensity combustion (peak pixel value: 4.5) with blue flame **before calibration**. (repetition 4)



(b) PEMAD outcome of low intensity combustion (peak pixel value: 4.5) with blue flame **after calibration**. (repetition 4)



(c) PEMAD outcome of higher intensity combustion (peak pixel value: 31.4) with blue flame **before calibration**. (repetition 5)



(d) PEMAD outcome of higher intensity combustion (peak pixel value: 31.4) with blue flame **after calibration**. (repetition 5)

Figure 27: Comparison of PEMAD outcomes (Test ID: Blue, repetitions 4 and 5) before calibration (a,c) and after applying calibration corrections (b,d). In all figures the left y-axis represents the intensity in pixel value, while for the after calibration figures the right y-axis represents the intensity in Lux. On the x-axis it time denoted in seconds. (note: different y-axis scales are used to improve visibility of low-intensity signals).

degrees were evaluated, and the RMSE was computed between the measured illuminance (Lux) values (light meter) and illuminance values predicted by the regression model based on the pixel intensities. The polynomial of degree 4 gave the lowest RMSE (0.0), indicating the best fit for the measured calibration points. This model was therefore selected as the calibration function for fiber 1. The regression models/calibration curves for all other fibers are provided in Appendix IV.III.

Using these calibration models, an additional y-axis in illuminance (Lux) was added to the PEMAD analysis output, which directly corresponds to the pixel intensity values. An example of the calibrated output is given in figure 29. Each pixel value is correctly converted to its corresponding Lux value based on the calibration data available for each fiber individually.

4.5 PEMAD system verification: technical precision

The technical precision of the PEMAD system under constant illumination showed excellent repeatability across all repetitions. For constant white light, both duration and peak intensity showed low RSD values of 0.14 % and 0.91%, respectively. RGB values remained constant at 255 for all channels (white) for all repetitions. For flickering blue light, both duration and peak intensity also showed low RSD values of 0.33 % and 1.25%, respectively. RGB values remained constant for all individual flashes and repetitions. The blue color channel had an RSD of 0.26% across repetitions. For each blue flash, 10–11 frames were detected, corresponding to a duration of 0.083–0.092s, and 9–10 frames between flashes, corresponding to 0.075–0.083s. The PEMAD analysis results are provided in Appendix XIV.

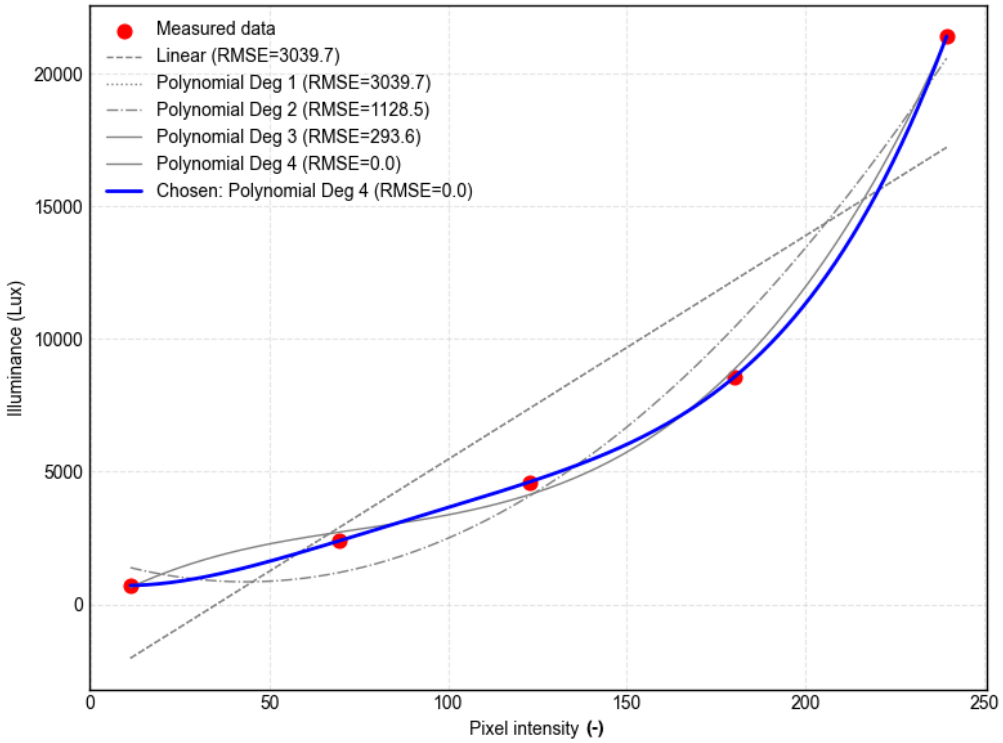


Figure 28: Regression curve selection for fiber 1 intensity data, showing measured data points (pixel value, Lux value) and tested degrees. The polynomial of degree 4 (blue continuous line) provided the best fit and was selected as the calibration curve.

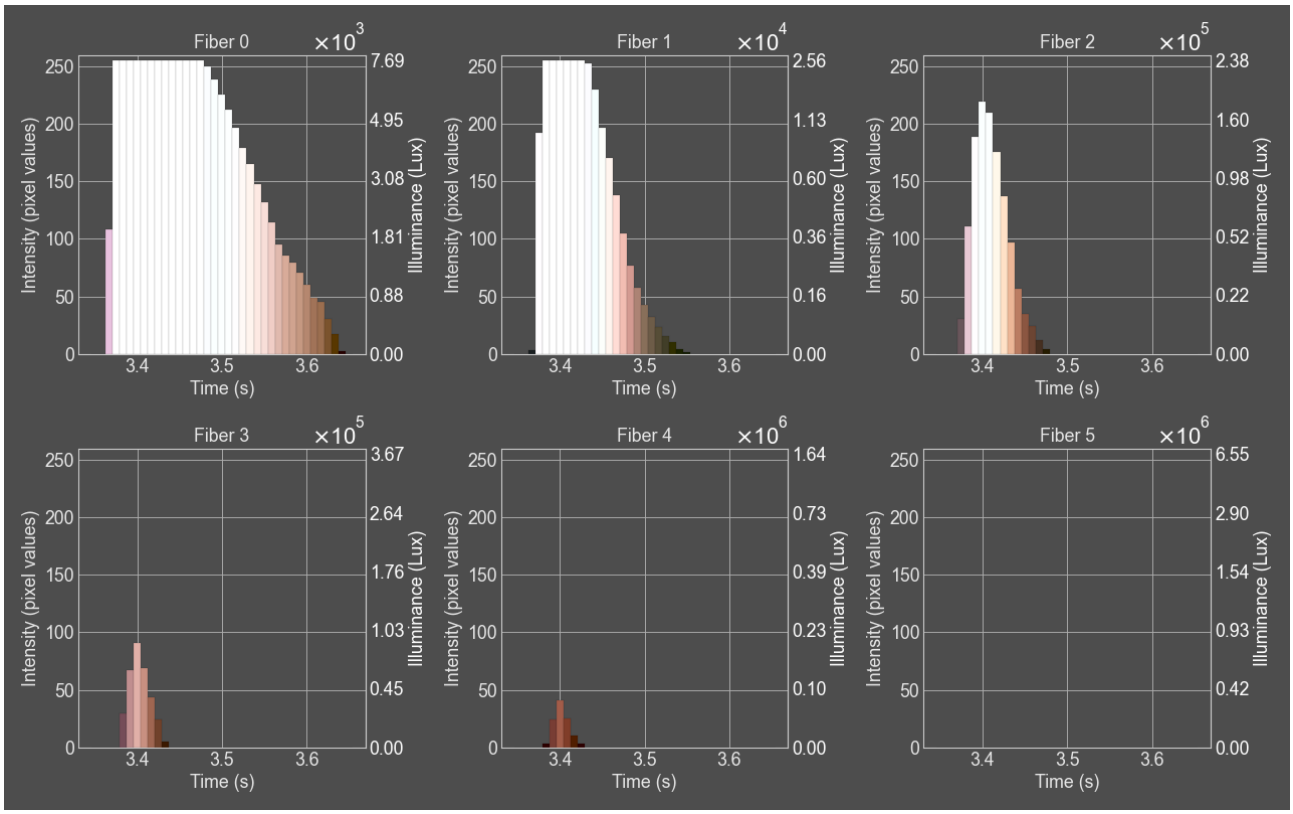


Figure 29: Calibrated intensity values on the right y-axis in illuminance (Lux). Note: keep the multiplying term for the right y-axis on top of each individual plot in mind (scientific notation to keep numbers readable). (Test ID: Ratio A, repetition 4)

4.6 PEMAD workflow validation

A comprehensive overview of all PEMAD analysis results for all relevant fibers, performance criteria, and test repetitions is provided in Appendices IX-XIII.

4.6.1 Detection limit

Figure 30 shows that for both gunpowder and flash powder, the PEMAD system was able to measure and compute the combustion duration based on the non-zero frames detected on the first camera fiber (without ND filter). The results show that gunpowder burns relatively slow (median = 0.33 s), while flash powder burns extremely fast (median = 0.058 s). The spread across repetitions was noticeably larger for gunpowder (IQR: 0.31–0.34 s) than for flash powder (IQR: 0.058–0.063 s), in absolute terms. The corresponding RSD values are shown Table 6 in section 4.6.6.

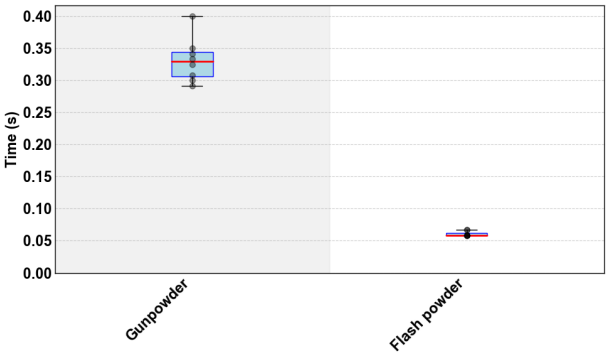


Figure 30: Combustion duration distributions across all repetitions (n=8) of gunpowder and flash powder for detection limit tests. Box-plots show medians (red lines), interquartile ranges (IQR, blue boxes), whiskers (1.5×IQR) indicating variability, and individual repetitions (black dots).

Figure 31 shows the RGB value distributions for gunpowder and flash powder. Together with the PEMAD analysis results in Appendices IX.I and IX.II, flash powder shows a bright white combustion with low dispersion, while gunpowder displays a less bright red flame and longer duration. However, the visual reference in Figure 32 shows that gunpowder combustion appeared orange rather than distinctly red.

The peak intensity measurement was considered reliable for gunpowder, as an average of 37 non-zero frames were recorded across repetitions, exceeding the predefined minimum of five frames on the selected (i.e., well exposed) fiber required for accurate measurement (see Appendix IX. In contrast, for flash powder, fewer than five non-zero frames were detected in most cases, so the peak intensity could not be measured with the same reliability. Figure 33 shows the corresponding peak intensity (Lux) distribution across all repetitions.

4.6.2 Selectivity/specificity: color detection

Figure 34 shows clear differences in combustion duration between the torch compositions. The orange torch

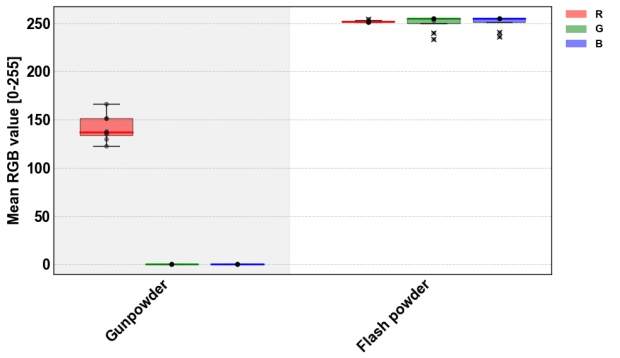


Figure 31: Red (R), green (G), and blue (B) channel value (flame color) distributions across all repetitions (n=8) for gunpowder and flash powder for detection limit tests. For each composition, the boxes are shown in a fixed order: R, G, and B channels.

Figure 32: Visual reference (still) frame for gunpowder combustion. (Test ID: DL-1, repetition 4)

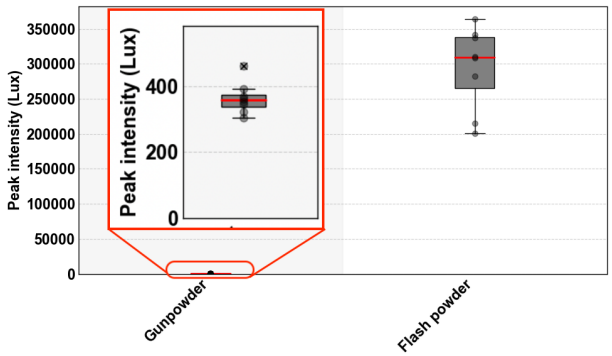


Figure 33: Peak intensity (Lux) distributions (n=8) of gunpowder and flash powder for detection limit tests. (The red square contains a zoomed view of the gunpowder intensity results, showing more details for this low intensity region).

burned the longest and showed greater variability in combustion durations across repetitions, whereas the other torches showed shorter and more consistent durations.

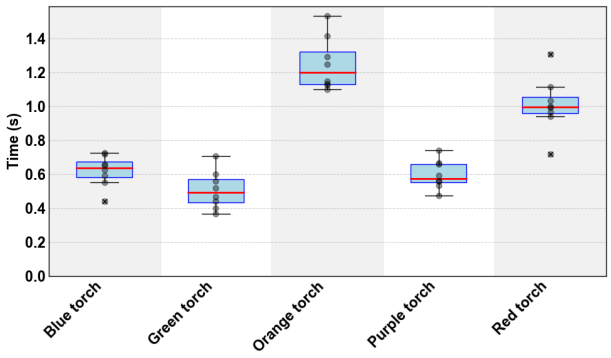


Figure 34: Combustion duration distributions across all repetitions (n=8) of all compositions for selectivity/specificity: color detection tests.

Figure 35 shows the RGB value distributions for the different torch compositions, illustrating that PEMAD detected differences in dominant color channels for each composition. The blue, green, red, and purple torches show RGB profiles in the box-plot that align with their expected flame colors, whereas the orange composition shows only red channel values, appearing as a red flame in the RGB data. To clarify, this is visualized in Figures 38 and 39, and is further detailed in the complete set of PEMAD analysis results in Appendix X. These results correspond well with the observed visual reference footage shown in Figure 26 (section 4.3.5) for the blue torch and in Figure 36 for the remaining colors. In addition, flickering behavior observed in the visual reference videos for most torches corresponds to the intensity fluctuations and multiple peaks visible in Figure 38 and Appendix X.

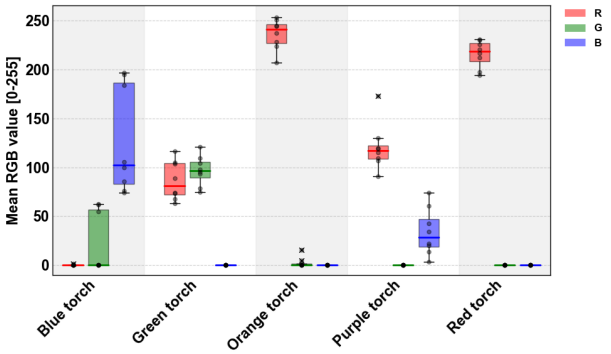
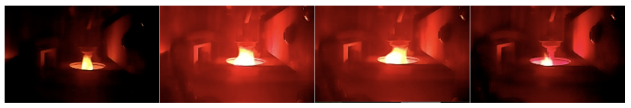


Figure 35: R, G, and B channel value (flame color) distributions across all repetitions (n=8) of all torch compositions for selectivity/specificity: color detection tests.

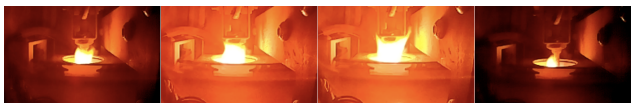
Figure 37 shows distinct differences in peak intensity among the torch compositions. The orange torch produced the highest intensities, while the purple and blue torches remained in the lower intensity range. The PEMAD system also detected low intensity blue compositions, indicating sensitivity across a wide range for this color. The largest variability in peak intensity



(a) Red flame (Test ID: Red, repetition 2)

(b) Green flame (Test ID: Green, repetition 5)

(c) Purple flame (Test ID: Purple, repetition 1)



(d) Lower intensity orange flame, pixel intensity < 80 (Test ID: Orange, repetition 1)

(e) Higher intensity orange flame, pixel intensity > 80 (additional test, orange torch composition mixed with leftover flash powder)

Figure 36: Visual reference footage for results shown in Figures 38 and 39. The figures can be read as a film tape but note that it only shows four selected still frames extracted from the MP4 recording of the corresponding torch composition and test repetition.

across repetitions was observed for the blue and orange torches. Finally, the results showed that at pixel intensities below ±80 (corresponding to ±1413 lux), the orange torch composition was classified as a red flame (Figure 39a). However, once this threshold was exceeded, the system was able to distinguish the flame from red and correctly characterized it as orange (Figure 39b) and for test repetitions 7 and 8 (Figures 189 and 190 in Appendix X.V).

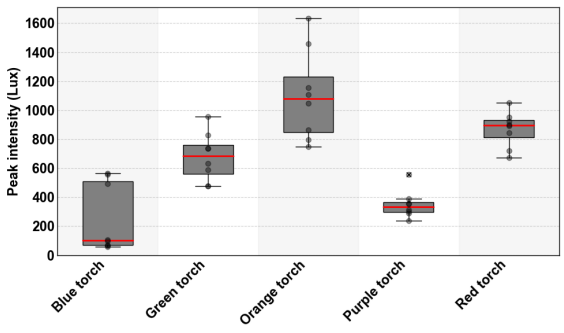
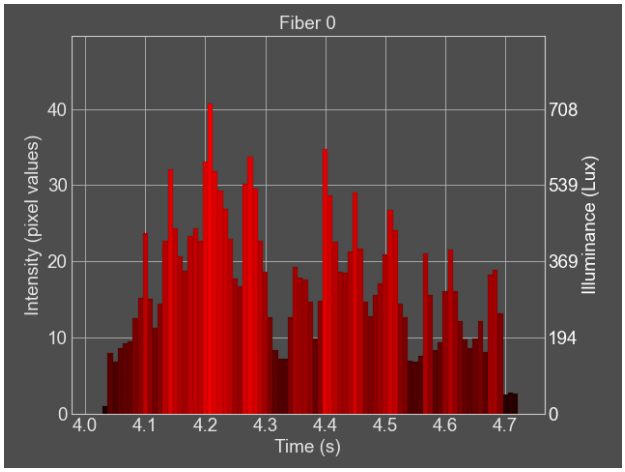
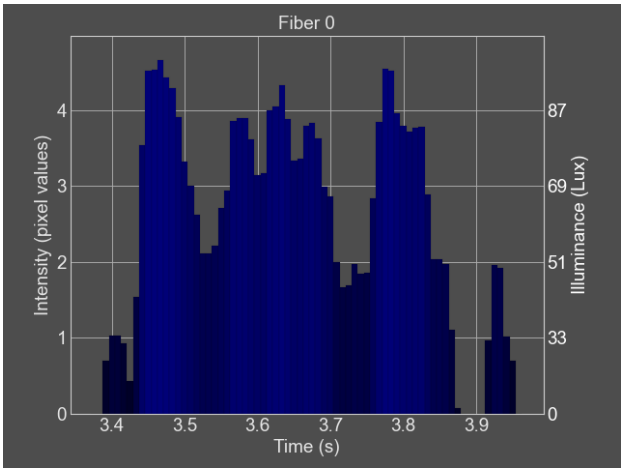


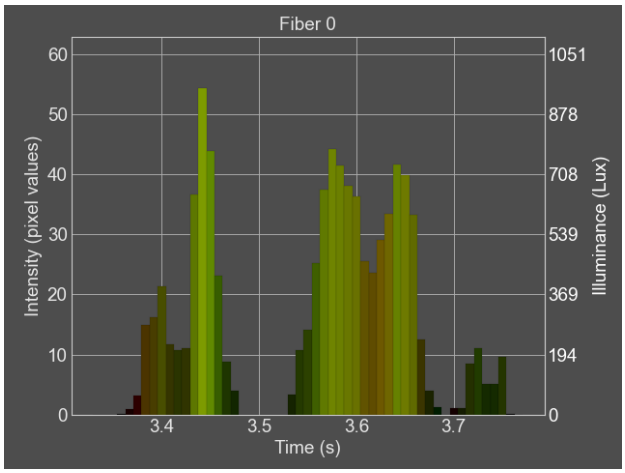
Figure 37: Peak intensity (Lux) distributions across all repetitions (n=8) of all compositions for selectivity/specificity: color detection tests.



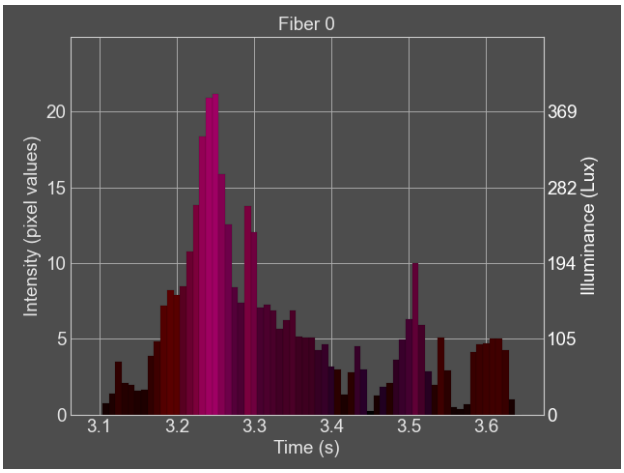
(a) Red flame (Test ID: Red, repetition 2).



(b) Blue flame (Test ID: Blue, repetition 4).

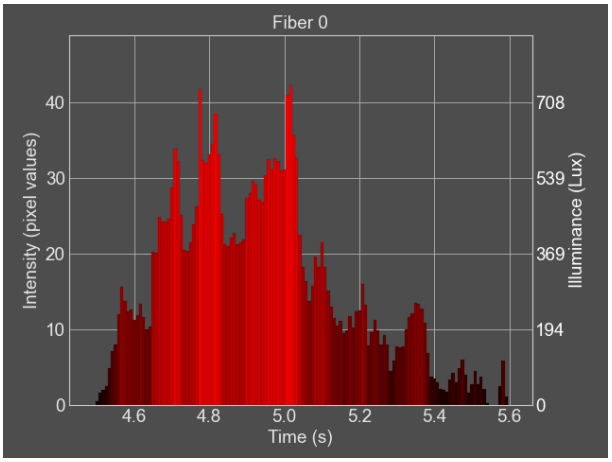


(c) Green flame (Test ID: Green, repetition 5).

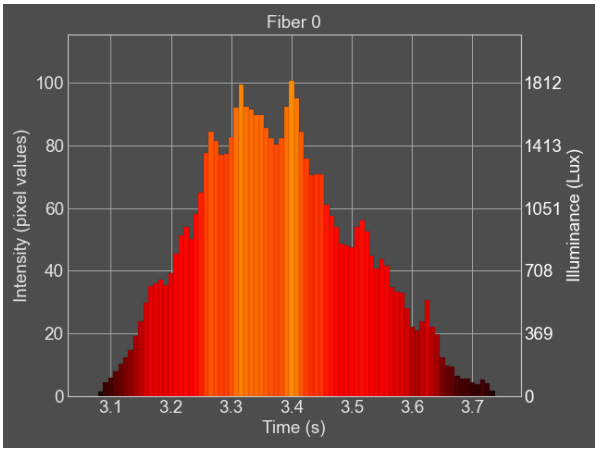


(d) Purple flame (Test ID: Purple, repetition 1).

Figure 38: PEMAD analysis results showing the intensity (pixel value) on the left y-axis and the corresponding illuminance value (Lux) on the right, for (category 1) torches, producing different flame colors. Note the different y-axis scale to improve readability.



(a) Lower intensity orange flame, pixel intensity < 80 (Test ID: Orange, repetition 1).



(b) Higher intensity orange flame, pixel intensity > 80 (additional test, orange torch composition mixed with leftover flash powder).

Figure 39: PEMAD analysis results showing the intensity (pixel value) on the left y-axis and the corresponding illuminance value (Lux) on the right, for orange torch compositions (one repetition) at lower intensity (pixel value < 80)(a) and higher intensity (pixel value > 80)(b).

4.6.3 *Selectivity/specificity: oxidizer/fuel ratio*

Figure 40 shows that the combustion duration clearly differs between the compositions with varying oxidizer/fuel ratios. Ratio A composition burns the slowest followed by ratio B, while the 80/20 composition burns extremely fast (comparable to flash powder). This shows that the PEMAD was able to detect changes in combustion duration as the oxidizer content increased, with higher amount of oxidizer content leading to faster combustions.

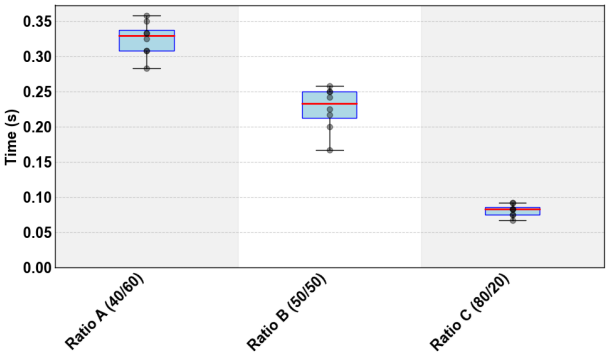


Figure 40: Combustion duration distributions across all repetitions (n=8) for selectivity/specificity: varying oxidizer/fuel ratio tests.

Figure 41 shows that the flame color output changes as the oxidizer/fuel ratio varies. Ratio A shows a large spread in the RGB values, especially in the green and blue channels, while ratio B and C produce more consistent values across repetitions. Ratio C shows more balanced and higher RGB values, corresponding to an overall brighter white flame. This observation is supported by the visual reference footage and PEMAD analysis result for Ratio C (Figure 42), which is a representative example. In contrast, the less balanced RGB values observed for Ratio A are reflected in its visual footage (Figure 43), where the flame shows more color variation instead of a white flame. The complete set of results for all ratio variations and repetitions is provided in Appendix XI.

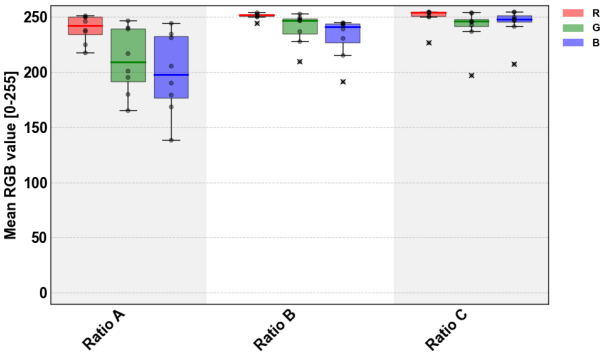


Figure 41: R, G, and B channel value (flame color) distributions across all repetitions (n=8) for selectivity/specificity: varying oxidizer/fuel ratio tests.

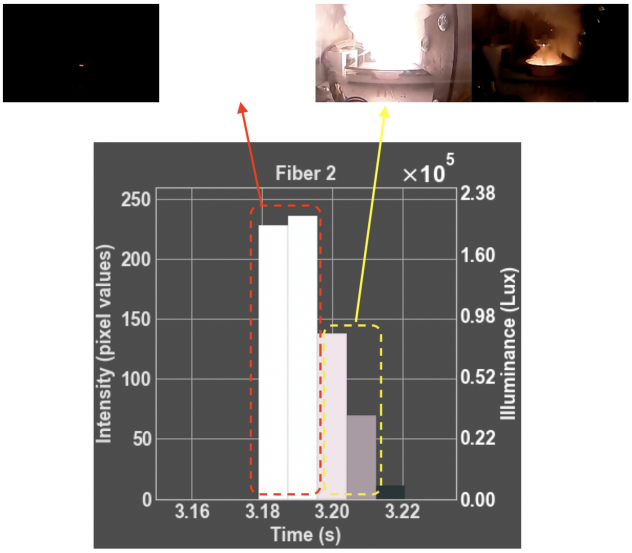


Figure 42: Visual reference footage and corresponding PEMAD analysis result (Test ID: Ratio C, repetition 1). The frames can be read as a film tape, showing four selected still images extracted from the MP4 recording of the combustion. The lower plot presents the corresponding analysis result for the same repetition on the selected fiber for analysis (fiber 2).

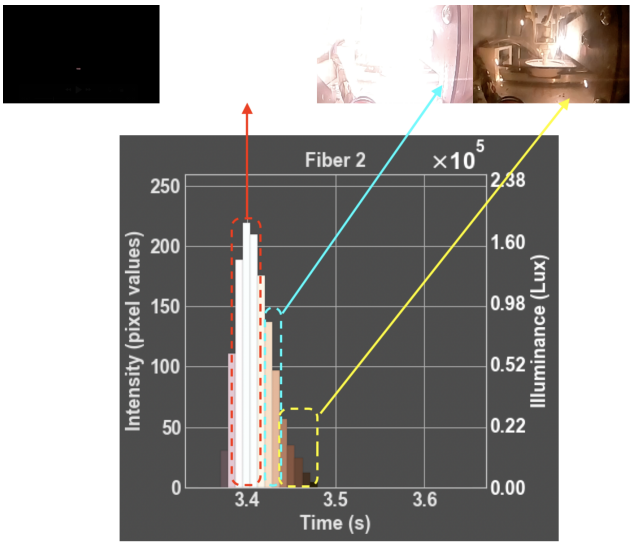


Figure 43: Visual reference footage and corresponding PEMAD analysis result (Test ID: Ratio A, repetition 4). The frames can be read as a film tape, showing four selected still images extracted from the MP4 recording of the combustion. The lower plot presents the corresponding analysis result for the same repetition on the selected fiber for analysis (fiber 2).

Figure 44 shows that the peak intensity varies between the different oxidizer/fuel ratios. Ratio A produced the lowest peak intensities, whereas Ratio B reached the highest values. Ratio C showed intermediate peak intensities that were more consistent across repetitions. Because the visual reference footage for both

ratio A and B was overexposed, intensity differences were evaluated using the fiber camera recordings. As shown in Figure 45, the combustion of Ratio B shows a noticeably higher intensity, as it reaches more fibers (with higher ND filters) compared to the combustion of ratio A. This visual observation aligns with the quantitative results, confirming that the data processing was performed correctly and captured the variations in intensity between the tested oxidizer/fuel ratios.

However, ratio C shows lower intensity because of its very fast combustion, where sometimes less than five non-zero frames were detected for analysis (see Figure 42). As a result, the actual peak intensity for this ratio may be underestimated.

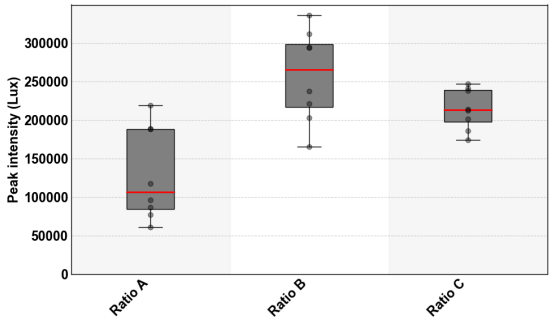


Figure 44: Peak intensity (Lux) distributions across all repetitions (n=8) of all compositions for selectivity/specificity: varying oxidizer/fuel ratio tests.

Figure 45: Comparison of recorded fiber camera frames (before color correction) at peak intensity for (left) Ratio A (repetition 1, fiber 2 selected for analysis) and (right) Ratio B (repetition 3, fiber 3 selected for analysis) illustrating the difference in combustion intensity.

4.6.4 *Selectivity/specificity: grain size*

Figure 46 shows that the combustion duration is influenced by grain size of the particles. Fine grains burn the fastest, medium grain size particles show intermediate durations, and coarse grain size particles burn the slowest. The spread across repetitions (in absolute terms) was largest for coarse grains, while finer powders showed more consistent combustion durations.

The RGB distributions in Figure 47 show that all grain size combustions produced red flames (also seen in PEMAD results in Appendix XII). Fine grains showed the lowest red channel values, indicating less intense combustion, while medium and coarse grains showed progressively higher values, indicating an increase in flame brightness with larger grain sizes. However, the visual reference footage (Figure 48) shows a more orange flame appearance with varying intensities between grain sizes, rather than the distinctly red color indicated by the RGB analysis.

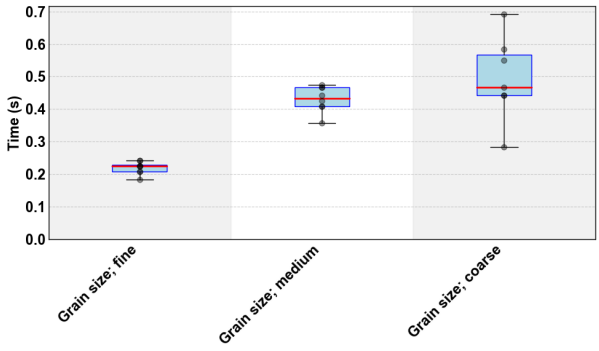


Figure 46: Combustion duration distributions across all repetitions (n=8) of all compositions for selectivity/specificity: grain size variation tests.

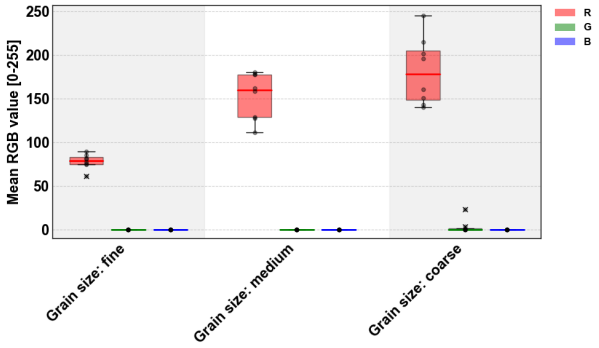


Figure 47: R, G, and B channel value (flame color) distributions across all repetitions (n=8) of all compositions for selectivity/specificity: grain size variation tests.

(a) Grain size "fine". Test ID G-fine, repetition 1

(b) Grain size "medium". Test ID G-med, repetition 1

(c) Grain size "coarse". Test ID G-coarse, repetition 4

Figure 48: Representative visual reference footage for all grain size combustions. Each row shows four selected still frames from the MP4 recording of one repetition, illustrating the combustion progression for (a) fine, (b) medium, and (c) coarse grains.

Finally, Figure 49 shows that the peak intensity increased with grain size, which was confirmed by the visual reference in Figure 48. The spread was highest for coarse grains across repetitions, whereas fine powders showed more consistent intensity behavior. The peak intensity estimates are considered reliable, as each recording contained on average 60 non-zero frames. However, for all tests, the peak intensity was below the threshold of 80 pixel value required to distinguish orange from red (Appendix XII).

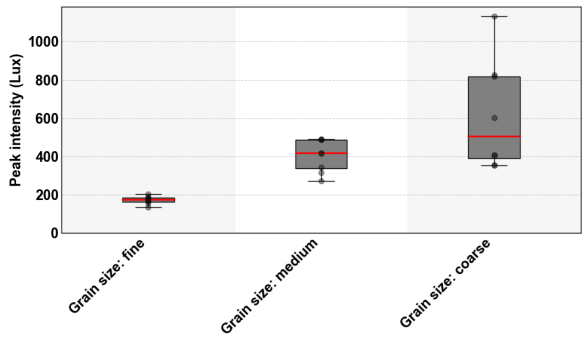


Figure 49: Peak intensity (Lux) distributions across all repetitions (n=8) of all compositions for selectivity/specificity: grain size variation tests.

4.6.5 *Robustness: starting volume*

The predefined criterion requires that the PEMAD system robustly measures combustion duration, flame color, and peak intensity when the sample volume varies by ±10% from the standardized sampling spoon volume ($V = 0.134\ cm^3$). Results must remain within a ±5% deviation from the mean, for the three tested volumes (RSD ≤5%) for all combustion characteristics. Figure 50 shows that combustion duration decreases slightly with increasing volume. The spread between repetitions was relatively low across all three conditions. The RSD across the mean of three starting volumes was 12.97%, which exceeds the predefined robustness criterion.

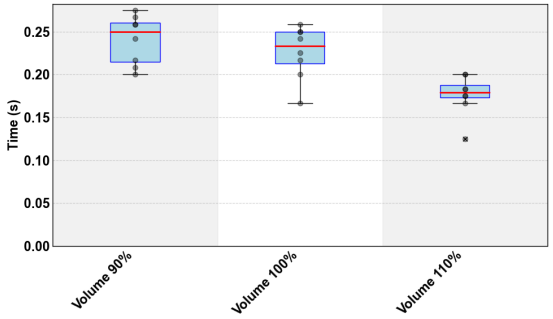


Figure 50: Combustion duration distributions across all repetitions (n=8) of all compositions for robustness: starting volume tests.

The RGB distributions in Figure 51 show that all three volumes produced flames with high red values and slightly lower green and blue channel values, indicating an overall white combustion flame. The 110% volume case showed slightly reduced blue values compared to 90% and 100%. For the 100% and 110% volume samples, larger variability can be observed in the green and blue channels. This is consistent with the visual reference footage. For the 110% volume case, two test repetitions are shown in Figures 52 and 53, where the intensity decreases rapidly (faster combustion, less frames), and the color transitions from bright white to darker colors during fading. In contrast, the 90% volume case (Figure 54) shows a longer lasting and more dominant bright white flame throughout the combustion. The slower intensity decrease is consistent with the longer combustion duration seen in Figure 50.

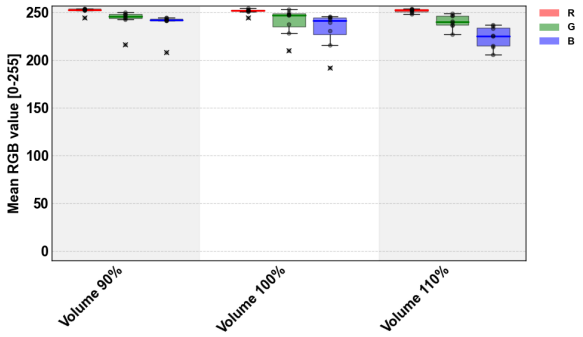


Figure 51: R, G, and B channel value (flame color) distributions across all repetitions (n=8) of all torch compositions for robustness: starting volume tests.

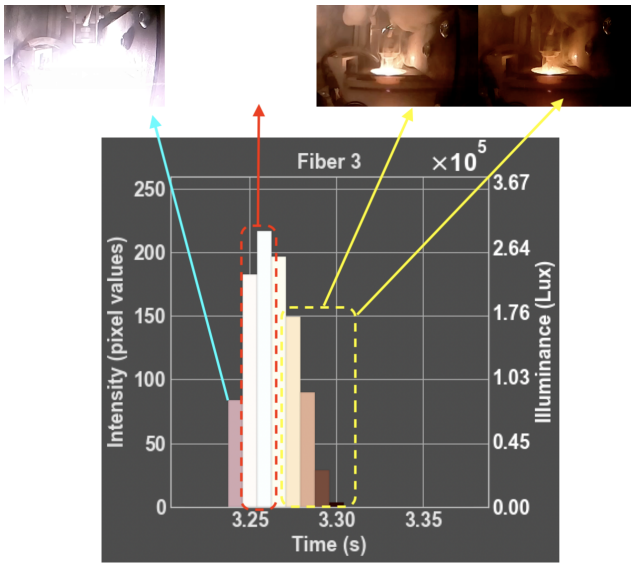


Figure 52: Visual reference footage and corresponding PEMAD analysis result (Test ID: V-110, repetition 5).The frames can be read as a film tape, showing four selected still images extracted from the MP4 recording of the combustion. The lower plot presents the corresponding PEMAD analysis result for the same repetition on the selected fiber used for analysis (fiber 3).

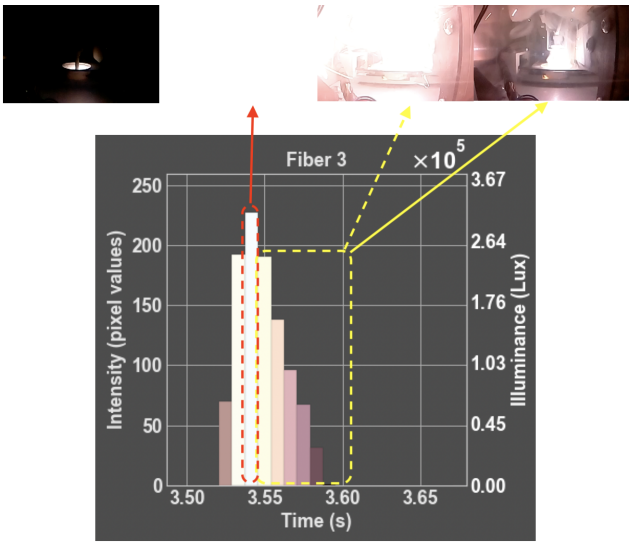


Figure 53: Visual reference footage and corresponding PEMAD analysis result (Test ID: V-110, repetition 6).The frames can be read as a film tape, showing four selected still images extracted from the MP4 recording of the combustion. The lower plot presents the corresponding PEMAD analysis result for the same repetition on the selected fiber used for analysis (fiber 3).

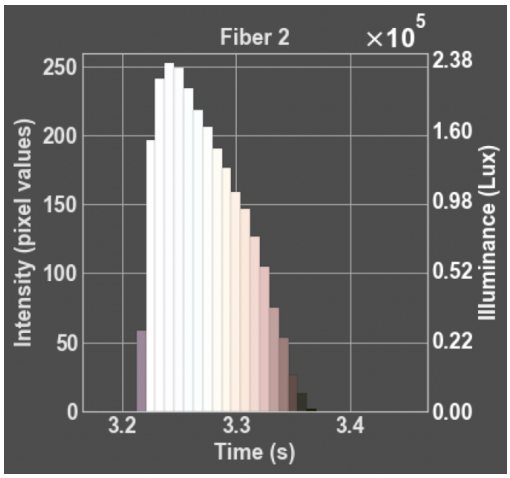


Figure 54: PEMAD analysis results (Test ID: V-90, repetition 6) showing the intensity (pixel value) on the left y-axis and the corresponding illuminance value (Lux) on the right. Color is shown in the frames (boxes). Note: analysis is performed on fiber 2.

As shown in Table 5, the RSD values for all channels remained well below the predefined threshold (red: 0.11%, green: 0.47%, blue: 2.56%), confirming that the robustness criterion for RGB variation was met.

Figure 55 shows that the peak intensity increased with sample volume. The lowest intensities were observed for the 90% volume case, while the 110% volume showed the highest overall intensities. The 100% volume case showed the largest variability across repetitions, indicating lower reproducibility compared to the other two volumes. Because the visual reference footage

Table 5: Robustness results for RGB channels across starting volumes (RSD < 5% required).

Color channel	Overall mean (pixel value)	SD	RSD (%)
Red	251	0.27	0.11
Green	241	1.14	0.47
Blue	231	5.91	2.56

was overexposed, intensity differences were evaluated using the fiber camera recordings. As shown in Figure 56, this visual reference also showed noticeably higher intensity as the volume increased, as it reaches more fibers (with higher ND filters). Finally, the RSD for peak intensity between the three volume conditions was 13.7%.

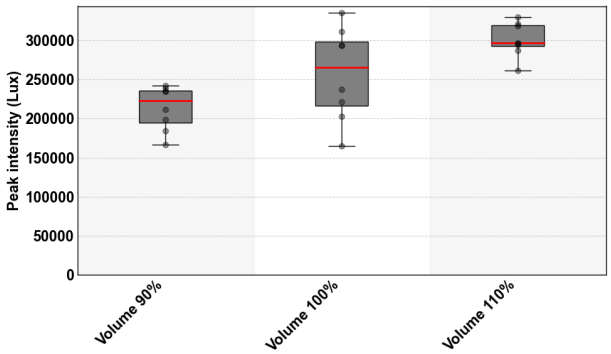


Figure 55: Peak intensity (Lux) distributions across all repetitions (n=8) of all compositions for robustness: starting volume test.

Figure 56: Comparison of recorded fiber camera frames (before color correction) at peak intensity for (left) V-90 (repetition 1, fiber 2 selected for analysis), (middle) V-100 (repetition 4, fiber 3 selected for analysis), and (right) V-110 (repetition 6, fiber 3 selected for analysis) illustrating the difference in combustion intensity when volume increased.

4.6.6 PEMAD workflow repeatability (precision)

In Table 6, the measurement precision for combustion duration is summarized. The RSD values for combustion duration ranged from 6.0% to 12.8% across the tested compositions and all repetitions. Flash powder showed the lowest variability between repetitions, while ratio B and the volume tests showed the highest. All volume tests used the same composition with only small variations in starting volume, so their RSD values are expected to be approximately the same, which is the case here.

Table 6: Precision of combustion duration measurements for selected compositions (SD = standard deviation, RSD = relative standard deviation).

Test ID	Mean (s)	SD (s)	RSD (%)
DL-1 (gunpowder)	0.33	0.03	9.72
DL-2 (Flash powder)	0.060	0.004	6.0
Ratio A (40/60)	0.33	0.02	7.0
Ratio B (50/50)	0.23	0.03	12.8
Ratio C (80/20)	0.081	0.008	9.9
Volume-90	0.24	0.03	11.2
Volume-100	0.23	0.04	12.8
Volume-110	0.18	0.02	12.6

In Table 7-9, the measurements precision for individual color channels are summarized. Channel R values were the most consistent overall. Flash powder showed the lowest variation, while Gunpowder gave the highest. Channel G values were more variable, with RSD values up to 13.4% for ratio A. Most other compositions were within 10%, and four of seven tests were within the required 5%. Channel B values showed the largest spread.

Table 7: Precision of color channel R measurements for selected compositions.

Test ID	Mean	SD	RSD (%)
DL-1 (gunpowder)	141	13.1	9.26
DL-2 (Flash powder)	252	1.18	**0.47**
Ratio A (40/60)	240	11.7	**4.86**
Ratio B (50/50)	251	2.73	**1.09**
Ratio C (80/20)	250	8.86	**3.55**
Volume-90	252	2.91	**1.16**
Volume-100	251	2.73	**1.09**
Volume-110	252	1.82	**0.72**

Table 8: Precision of color channel G measurements for selected compositions.

Test ID	Mean	SD	RSD (%)
DL-1 (gunpowder)	0	-	-
DL-2 (Flash powder)	250	7.95	**3.18**
Ratio A (40/60)	211	28.2	13.4
Ratio B (50/50)	240	13.7	5.70
Ratio C (80/20)	241	17.2	7.17
Volume-90	243	10.3	**4.25**
Volume-100	240	13.7	5.70
Volume-110	240	6.70	**2.79**

Peak intensity measurements (Table 10) showed variation between compositions and across repetitions within the same test. Gunpowder, the 90% and 110% volume cases, and Ratio C exhibited the lowest variability. Flash powder, Ratio B, and the 100% volume case showed intermediate variability, while Ratio A displayed the highest variation in peak intensity between repetitions.

Table 9: Precision of color channel B measurements for selected compositions.

Test ID	Mean	SD	RSD (%)
DL-1 (gunpowder)	0	–	–
DL-2 (Flash powder)	251	7.34	**2.93**
Ratio A (40/60)	199	34.4	17.3
Ratio B (50/50)	232	17.9	7.73
Ratio C (80/20)	244	14.4	5.90
Volume-90	238	11.4	**4.77**
Volume-100	232	17.9	7.73
Volume-110	224	10.7	**4.78**

Table 10: Precision of peak intensity ($kLux$) measurements for selected compositions. (Note: keep units in mind ($kLux$) not (Lux)

Test ID	Mean ($kLux$)	SD ($kLux$)	RSD (%)
DL-1 (gunpowder)	0.364	0.045	12.5
DL-2 (Flash powder)	295	55.4	18.8
Ratio A (40/60)	129	56.4	43.7
Ratio B (50/50)	258	55.9	21.7
Ratio C (80/20)	214	24.8	11.6
Volume-90	214	26.6	12.4
Volume-100	258	55.9	21.7
Volume-110	301	20.5	6.82

4.6.7 *Characterization and distinguishing pyrotechnic compositions*

All tested compositions were compared with each other for each combustion characteristic: combustion duration, flame color, and peak intensity, to assess whether the PEMAD system detected measurable differences between pyrotechnic/explosive compositions with different material properties.

Figure 57 shows that the combustion duration differs strongly between certain compositions: flash powder and ratio C composition burned extremely fast, while torch compositions showed noticeably longer combustion durations. Compositions with variations in grain size, oxidizer/fuel ratio (A and B), and volume showed smaller measured differences in combustion duration. The mean RGB distributions in Figure 58 show clear differences in flame color between groups. However, within testing groups, it is sometimes difficult to see clear variations.

Finally, Figure 59 shows the variation in peak intensities across the compositions. Flash powder, volume variation tests, and ratio compositions reach very high intensity levels, while torches, gunpowder and grain size samples remain in a lower region. The zoomed-in view (Figure 59b) clarifies the differences in the low intensity region. Visually, the PEMAD system shows differences in combustion behavior between compositions with different material properties through measurable variations in combustion duration, flame color, and intensity.

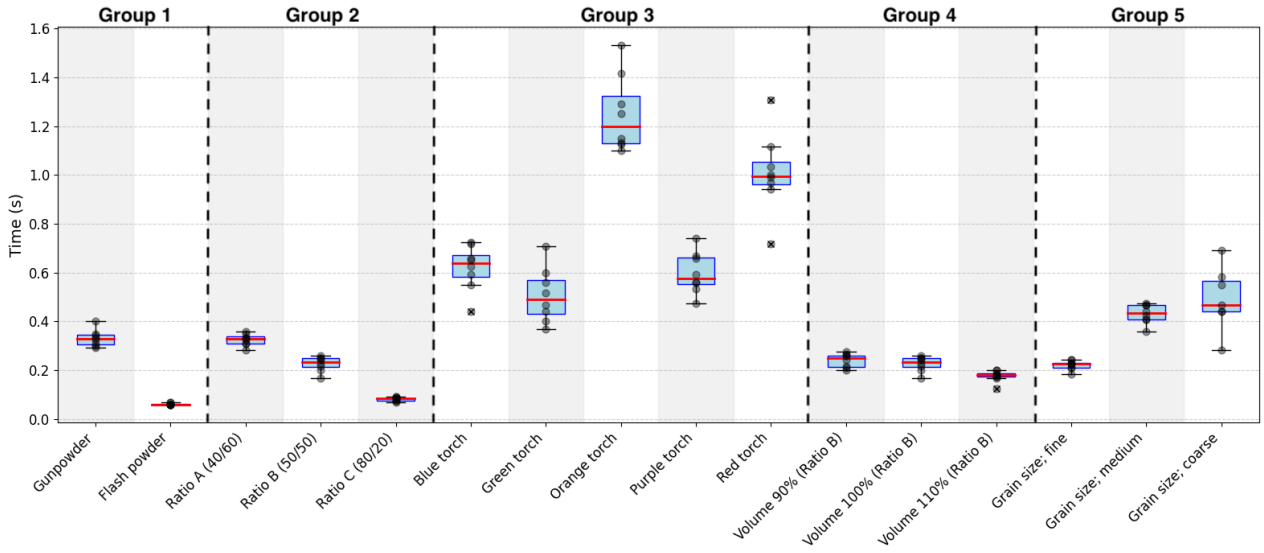


Figure 57: Combustion duration distributions of different materials (n=8, for each material). Box-plots show medians (red lines), interquartile ranges (IQR), whiskers (1.5×IQR), and individual repetitions (black dots). Dashed lines separate the different test groups: (1) detection limit, (2) oxidizer/fuel ratios variations, (3) torch compositions, (4) starting volume variations, and (5) grain size variations.

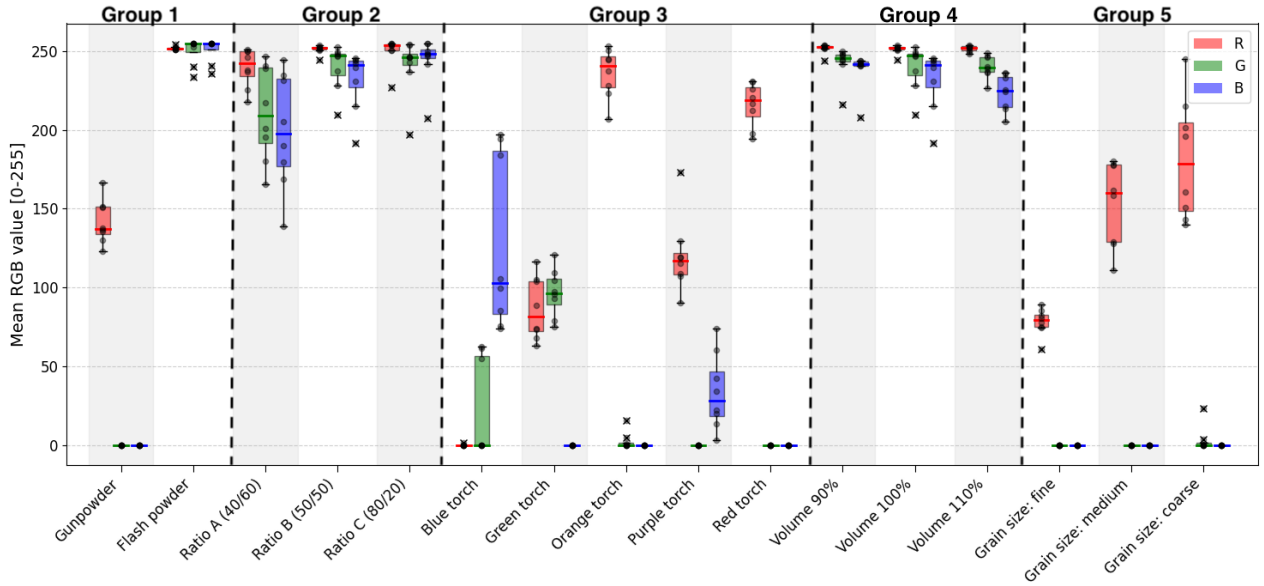
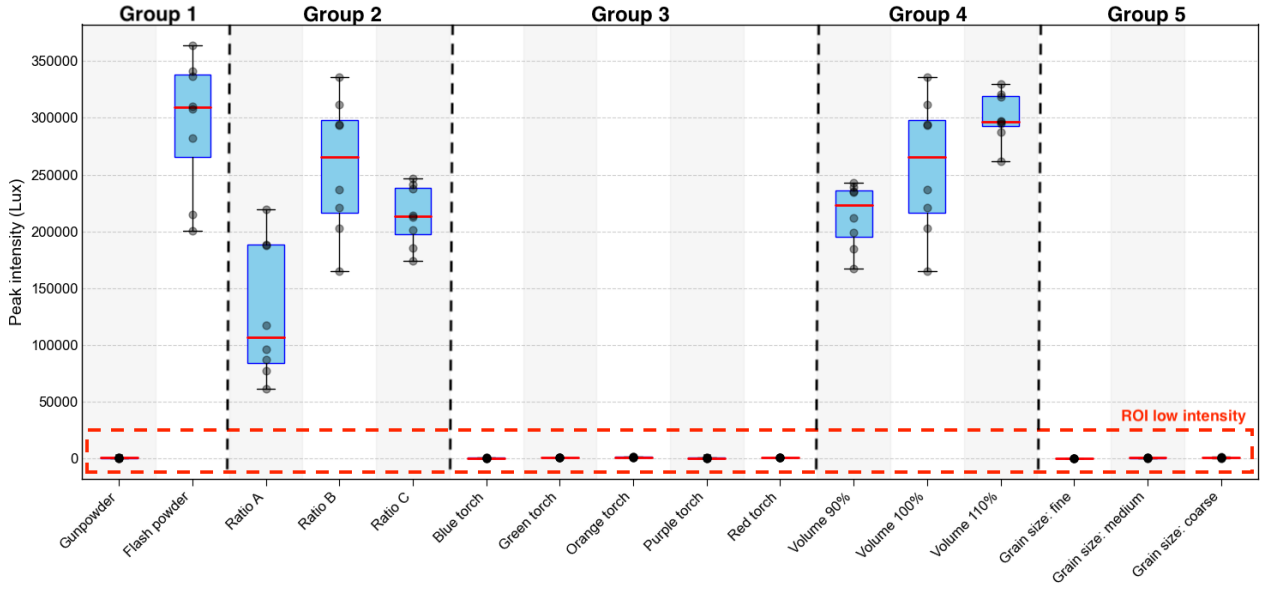
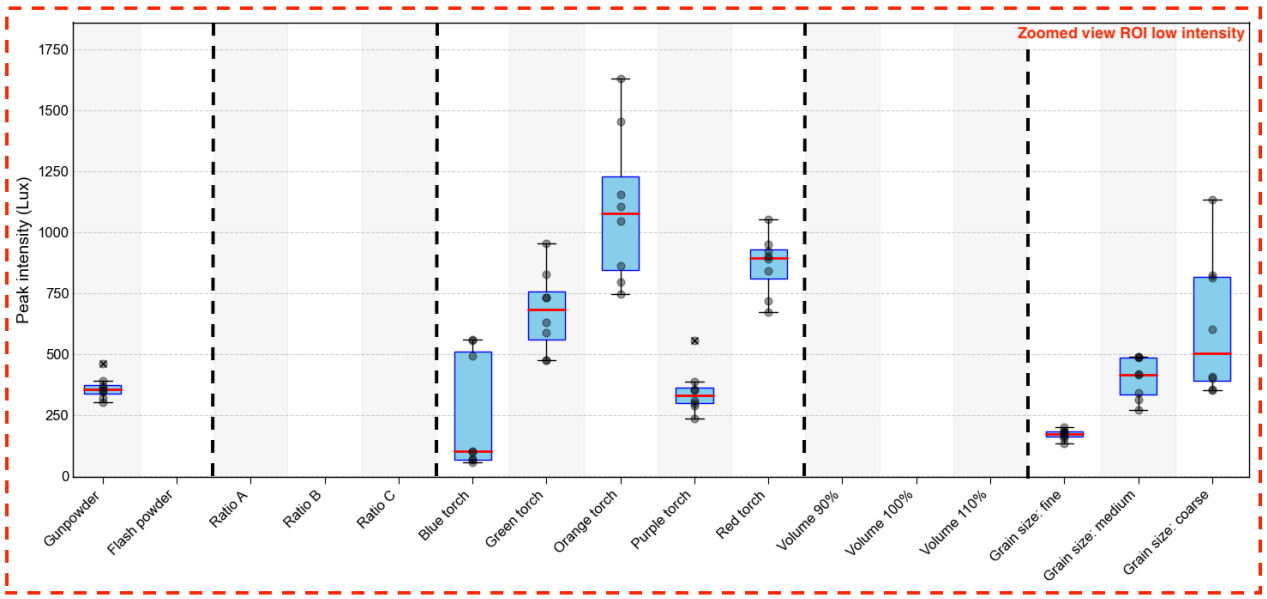


Figure 58: Mean RGB value distributions for different materials (n=8, for each material). Red, green, and blue boxes correspond to the R, G, and B color channels, respectively. Box-plots show medians (red lines), interquartile ranges (IQR), whiskers (1.5×IQR), and individual repetitions (black dots). Dashed lines separate the different test groups: (1) detection limit, (2) oxidizer/fuel ratios variations, (3) torch compositions, (4) starting volume variations, and (5) grain size variations.



(a) Full-range PEMAD outcome showing peak intensity distributions across all tested compositions.



(b) Zoomed view of the ROI from figure (a), showing more detailed peak intensity distributions for the low-intensity compositions

Figure 59: Peak intensity distributions (n=8, for each material) in Lux units for different pyrotechnic compositions. Figure (a) Shows the full range of values, where large differences between compositions are visible. The dashed rectangle highlights the low-intensity region of interest (ROI). Figure (b) Provides a zoomed view of this ROI, enabling clearer comparison of low-intensity combustions. Box-plots indicate medians (red lines), interquartile ranges (IQR), whiskers (1.5×IQR), and individual repetitions (black dots). Vertical dashed lines separate the different test groups: (1) detection limit, (2) oxidizer/fuel ratios variations, (3) torch compositions, (4) starting volume variations, and (5) grain size variations.

Table 11: Dunn's post-hoc test results (Holm-Bonferroni (FWER control) adjusted p-values) for combustion duration. Where ns = not significant, * indicates p < 0.05 and ** indicates p < 0.001

	Blue torch	Flash powder	Grain size: coarse	Grain size: fine	Grain size: medium	Green torch	Gunpowder	Orange torch	Purple torch	Ratio A (40/60)	Ratio B (50/50)	Ratio C (80/20)	Red torch	Volume 100%	Volume 110%	Volume 90%
Blue torch	-															
Flash powder	**	-														
Grain size: coarse	ns	**	-													
Grain size: fine	*	ns	ns	-												
Grain size: medium	ns	**	ns	ns	-											
Green torch	ns	**	ns	ns	ns	-										
Gunpowder	ns	ns	ns	ns	ns	ns	-									
Orange torch	ns	**	ns	**	ns	ns	ns	-								
Purple torch	ns	**	ns	ns	ns	ns	ns	ns	-							
Ratio A (40/60)	ns	ns	ns	ns	ns	ns	ns	ns	ns	-						
Ratio B (50/50)	ns	ns	ns	ns	ns	ns	ns	**	ns	ns	-					
Ratio C (80/20)	**	ns	**	ns	*	**	ns	**	**	ns	ns	-				
Red torch	ns	**	ns	**	ns	ns	ns	ns	ns	ns	**	**	-			
Volume 100%	ns	ns	ns	ns	ns	ns	ns	**	ns	ns	ns	ns	**	-		
Volume 110%	**	ns	*	ns	ns	*	ns	**	**	ns	ns	ns	**	ns	-	
Volume 90%	ns	ns	ns	ns	ns	ns	ns	**	ns	ns	ns	ns	**	ns	ns	-

4.6.8 Statistics

The Kruskal–Wallis tests showed multiple significant differences between compositions for both combustion duration and peak intensity (Lux) (See tables 11 and 12). In addition, Dunn's tests were also applied to the red, green, and blue channels, separately. However, these tables do not provide useful insight in all cases where it was expected, because color perception is determined by the combination of the RGB channel values rather than each channel separately, this will be evaluated more in the discussion.

Pairwise PERMANOVA results for flame color (RGB value) (Table 13) showed significant differences between most compositions after Holm–Bonferroni correction, indicating that multivariate RGB color profiles differed across most compositions. Exceptions included some closely related groups, such as ratio B vs ratio C and volume variations, where no significant differences were detected.

Table 14 summarizes the pairwise statistical results within each performance characteristic group. By applying comparisons to related compositions, this indicates whether PEMAD detects the systematic effects of controlled variations rather than just large differences between unrelated compositions. Significant values highlight where combustion duration, peak intensity, or RGB color profiles were different enough to capture these effects. These results show that PEMAD detected the effects of controlled variations on the combustion behavior in multiple cases. For robustness, significant differences were detected between starting volume conditions, demonstrating that the system is sensitive to these variations. However, the statistical results also showed significant differences between orange and red ($p = 0.019$), although in practice the PEMAD system classified orange flames as red at lower intensities.

Table 12: Dunn's post-hoc test (Holm-Bonferroni (FWER control) adjusted p-values) for peak intensity (Lux). Where ns = not significant, * indicates p < 0.05 and ** indicates p < 0.001

	Blue torch	Flash powder	Grain size: coarse	Grain size: fine	Grain size: medium	Green torch	Gunpowder	Orange torch	Purple torch	Ratio A (40/60)	Ratio B (50/50)	Ratio C (80/20)	Red torch	Volume 100%	Volume 110%	Volume 90%
Blue torch	-															
Flash powder	**	-														
Grain size: coarse	ns	*	-													
Grain size: fine	ns	**	ns	-												
Grain size: medium	ns	**	ns	ns	-											
Green torch	ns	ns	ns	ns	ns	-										
Gunpowder	ns	**	ns	ns	ns	ns	-									
Orange torch	ns	ns	ns	ns	ns	ns	ns	-								
Purple torch	ns	**	ns	ns	ns	ns	ns	ns	-							
Ratio A (40/60)	ns	ns	ns	*	ns	ns	ns	ns	ns	-						
Ratio B (50/50)	**	ns	ns	**	**	ns	**	ns	**	ns	-					
Ratio C (80/20)	**	ns	ns	**	ns	ns	*	ns	*	ns	ns	-				
Red torch	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	-			
Volume 100%	**	ns	ns	**	**	ns	**	ns	**	ns	ns	ns	ns	-		
Volume 110%	**	ns	**	**	**	*	**	ns	**	ns	ns	ns	ns	ns	-	
Volume 90%	**	ns	ns	**	ns	ns	*	ns	*	ns	ns	ns	ns	ns	ns	-

Table 13: pairwise PERMANOVA test results (Holm-Bonferroni (FWER control) adjusted p-values) for RGB values. Where ns = not significant, * indicates p < 0.05 and ** indicates p < 0.001

	Blue torch	Flash powder	Grain size: coarse	Grain size: fine	Grain size: medium	Green torch	Gunpowder	Orange torch	Purple torch	Ratio A (40/60)	Ratio B (50/50)	Ratio C (80/20)	Red torch	Volume 100%	Volume 110%	Volume 90%
Blue torch	-															
Flash powder	*	-														
Grain size: coarse	*	*	-													
Grain size: fine	*	*	*	-												
Grain size: medium	*	*	ns	*	-											
Green torch	*	*	*	*	*	-										
Gunpowder	*	*	ns	*	ns	*	-									
Orange torch	*	*	ns	*	*	*	*	-								
Purple torch	*	*	*	*	*	*	*	*	-							
Ratio A (40/60)	*	*	*	*	*	*	*	*	*	-						
Ratio B (50/50)	*	ns	*	*	*	*	*	*	*	ns	-					
Ratio C (80/20)	*	ns	*	*	*	*	*	*	*	ns	ns	-				
Red torch	*	*	ns	*	*	*	*	ns	*	*	*	*	-			
Volume 100%	*	ns	*	*	*	*	*	*	*	ns	ns	ns	*	-		
Volume 110%	*	*	*	*	*	*	*	*	*	ns	ns	ns	*	ns	-	
Volume 90%	*	ns	*	*	*	*	*	*	*	ns	ns	ns	*	ns	ns	-

Table 14: Pairwise comparisons across all performance criteria . Significant differences are indicated in bold: values below 0.001 are shown as **<0.001**. Combustion duration and peak intensity p-values are determined with Dunn's post-hoc test (Holm-Bonferroni (FWER control) adjusted p-values. The p-values for the combined color channels (RGB) are determined with the pairwise PERMANOVA test, also Holm-Bonferroni corrected)

Comparison	Peak intensity (Lux)	Combustion duration (s)	Flame color (RGB)
Detection limit			
Flash powder vs Gunpowder	**<0.001**	**<0.001**	**<0.001**
Selectivity/specificity: color detection			
Blue vs Green	0.119	0.843	**0.002**
Blue vs Orange	**<0.001**	**0.004**	**0.002**
Blue vs Purple	0.883	0.843	**0.002**
Blue vs Red	**0.003**	0.101	**0.001**
Green vs Orange	0.211	**<0.001**	**0.002**
Green vs Purple	0.211	0.843	**0.002**
Green vs Red	0.621	**0.003**	**0.002**
Orange vs Purple	**<0.001**	**0.002**	**0.002**
Orange vs Red	0.883	0.843	**0.019**
Purple vs Red	**0.007**	0.063	**0.002**
Selectivity/specificity: oxidizer/fuel Ratio			
Ratio A (40/60) vs Ratio B (50/50)	**0.003**	**0.047**	0.070
Ratio A (40/60) vs Ratio C (80/20)	0.062	**<0.001**	**0.037**
Ratio B (50/50) vs Ratio C (80/20)	0.273	**0.047**	0.332
Selectivity/specificity: grain size			
Coarse vs Fine	**<0.001**	**<0.001**›	**<0.001**
Coarse vs Medium	0.396	0.287	0.109
Fine vs Medium	**0.006**	**0.008**	**<0.001**
Robustness: starting volume			
Volume 100% vs Volume 110%	0.166	**0.025**	0.881
Volume 100% vs Volume 90%	0.166	0.365	0.881
Volume 110% vs Volume 90%	**0.003**	**0.002**	0.209

Discussion

5.1 Design and software

The implementation of the touchscreen improved the user-friendliness of the system by removing the need for an external mouse and keyboard. However, it is important to note that the Raspberry Pi operating system must remain on operating system version Buster (2020, legacy) rather than being upgraded to the newer versions Bullseye or Bookworm. This ensures compatibility between hardware, software, and the required camera settings. Specifically, achieving 120 fps at full FOV with the older Raspberry Pi HQ camera is only possible with the legacy camera stack from Buster OS. Newer OS (Bullseye, Bookworm) work with libcamera and picamera2, which no longer support the legacy camera stack. As a result, camera sensor modes on these systems are limited, failing to reach 120 fps or can only be achieved with a cropped FOV. It is recommended to maintain the current setup for full compatibility. However, if higher frame rates are required in future work, both the camera and software should be upgraded to achieve higher fps on newer operating systems without constraints in FOV or sensor modes.

5.2 Technical verification and baseline test (Pilot)

The pilot test showed that the device is capable of detecting slow and fast combustions and provided quantitative color, combustion duration, and intensity data. Most of the criteria for evaluating functionality were satisfied and consistently at least one fiber was available for analysis. However, several bugs emerged, including synchronization issues with the visual camera and fiber camera, occasional random errors in fiber recordings, calibration, and challenges in accurately showing low-intensity colors of the flame. The pilot was useful in finding areas for improvement, setting a baseline for the system at the start of the project. These issues were resolved prior to the calibration and validation phases.

5.3 Color calibration

The calibration results showed significant improvements in the accuracy of detecting RGB values with the PEMAD system with respect to the spectrometer's reference, reducing the median RMSE from 27–35% to just 2–3% of the RGB range [0-255] (Table 4, Figure 25). Before calibration, deviations (errors) were large, which were very noticeable in the lower-intensity regions, where colors were often displayed as black (Figure 27). After calibration, these deviations were corrected, resulting in visibly more accurate flame colors and

less "black" flames in the low-intensity regions, this improvement can be observed in Figures 27b and 27d compared to Figures 27a and 27c, respectively.

However, there are also limitations and assumptions that have to be taken into account. First, the calibration was performed only on the first fiber, which does not have a neutral density (ND) filter. In this way, the camera sensor is practically calibrated and it was further assumed that the ND filters are 100% neutral (i.e., the ND filter does not affect the color of the light). If the ND filters are not perfectly neutral, slight color shifts between fibers could occur, causing (small) systematic deviations in the measured RGB values.

A set of nine color filters was used to represent the full color spectrum. For future work, using more filters will result in more accurate analysis. For instance, this was seen for the orange torch composition, this flame color was detected as red at low intensity (pixel value <80). Implementing more orange filters with slightly different color values would increase the detection accuracy for this color and improves the system's capability of distinguishing orange from red.

Furthermore, it was assumed that all fibers behaved identically in terms of light transmission and potential losses. In practice, deviations/variations may occur as a result of fiber quality or polishing effects. Finally, the angular sensitivity (i.e., alignment of fiber with respect to the light source) of the fibers was taken into account in calibration by aligning the light source and spectrometer using a laser. However, in actual combustion measurements, the fibers are arranged in a [2x3] geometry, which means that not all fibers are perfectly aligned with respect to the flame. Although these effects are minimal, as these distances between the fibers are very small, they could be further minimized in future work by implementing cosine correctors on all fibers or by arranging the fibers in a slight angle toward the flame. This becomes more important when the set of fibers is expanded and the distance between them becomes greater.

5.4 Intensity calibration

Intensity calibration was performed successfully on the available data, and the chosen method was effective. This approach made it possible to convert pixel values into corresponding Lux values, representing "real-world" units for brightness levels rather than relative sensor output. This allows for more intuitive interpretation and direct comparison of intensity, even when results are not evaluated on the same fiber.

However, even after applying the calibration models, the peak values of consecutive fibers did not always match. To clarify (Figure 29), fiber 2 reached a maximum of 219 pixel value, corresponding to $\pm$187,000 Lux, which is in line with its calibration curve. However, fiber 3 reached a peak intensity of 91 pixel value for that same combustion, corresponding to $\pm$91,500 Lux, also consistent with its calibration data. Based on calibration theory, it would be expected that fiber 3 shows the same peak intensity in Lux as fiber 2, but at a lower pixel intensity due to its stronger ND filter. According to its calibration data, fiber 3 would need to reach a pixel value of 135–140 to represent the 187,000 Lux of fiber 2. The pixel values themselves were correctly converted to Lux according to each fiber's calibration data. This difference likely occurred because fiber 3 did not capture enough light during this combustion, and therefore did not reach the optimal detection range (140–225 pixel value), leading to an underestimated pixel intensity. Consequently, the computed Lux value reflects a correct conversion, but of an underestimated input signal. This indicates that the peak mismatch is not caused by errors in the calibration model but rather by system-related limitations, primarily the large step size between ND filters. In general, when one fiber captures multiple frames within the optimal detection range without overexposure, the subsequent fiber is likely to fall outside this range, resulting in mismatched peak intensities. Variations in fiber and ND filter quality, including polishing or manufacturing effects, may further contribute to light detection losses.

A limitation of the intensity calibration is that fibers 4 and 5 could not be calibrated directly because their ND filters were too strong for the light source and intensity meter available to provide sufficient data points. Therefore, calibration coefficients of fiber 1, which had complete data and the best RMSE for the calibration curve, were used and scaled with their ND factor.

These findings indicate that the system can be further improved by using additional intermediate and higher-quality ND filters and optical fibers. Furthermore, employing a Lux-meter with a larger measurement range and increasing the number of calibration data points (currently five, selected to cover the full range from near underexposure to near overexposure) would further improve this calibration.

5.5 PEMAD system verification: technical precision

The technical precision tests demonstrated that the PEMAD system has excellent repeatability. Both color and intensity measurements showed very low RSD values across repeated tests, confirming high measurement precision. This was observed for both constant white light and flickering blue light conditions, indicating good repeatability for static (white mode) and dynamic (flickering mode) illumination. However, it was assumed that the light source maintained constant output during tests and that the flickering mode produced consistent flash intensities, colors, and intervals, as no official documentation could be retrieved. Therefore, it cannot be determined with certainty whether the very small observed variability originated from the light source itself or from the measurement system.

5.6 PEMAD workflow validation

5.6.1 Detection limit

The PEMAD system successfully measured combustion duration and intensity for slow burning compositions

such as gunpowder. In contrast, for the fastest flash powder combustions (approx. $0.06s$), fewer than five frames were detected, marking the upper limit of the system. As a result, peak intensity for these very fast combustions could not yet be determined as reliably as for slower combustion. To improve the upper detection limit for intensity measurements in future work, a camera with higher frame rates should be implemented. Additionally, gunpowder showed that the system could not reliably distinguish orange from red, representing another detection limit in terms of color detection.

Furthermore, the results in section 4.6.7 show that the lower detection limit of the PEMAD is not limited to gunpowder detection. As the torches unexpectedly extended this range even more: the intensity of the blue torch was lower than gunpowder, which indicates that the PEMAD detects blue flames at very low intensity (peak intensity: 2-3 pixel value (50-69 Lux), Appendix X.III), comparable to a dimmed living room environment.

5.6.2 Selectivity/specificity: color detection

The color detection tests showed that the PEMAD system can distinguish between most torch compositions based on their RGB values alone (Table 14). The system clearly identified red, green, blue, and purple torch compositions, each showing its own characteristic color. These results align well with the expected visual flame color. The spread in measured intensity and RGB channels across repetitions differs per composition. This variability indicates inconsistent combustion behavior, which was also observed in the visual reference camera (i.e., flickering behavior). The combustion of the orange torch composition was initially detected as a red flame (Figure 39a). This effect can be explained by the fact that orange flames require a higher intensity to be reliably distinguished from red by the PEMAD system. As shown in Figure 39, the system cannot distinguish orange from red below a threshold of approximately 80 pixel value (= 1400 lux). Additional tests were performed by mixing the orange torch composition with flash powder to achieve a higher intensity. These tests in combination with the initial tests established to threshold of approximately 80 pixel value. Once this intensity threshold is exceeded, orange is correctly detected. As indicated in the discussion section of color calibration, extending the number of filters used in calibration may also improve this limitation.

5.6.3 Selectivity/specificity: oxidizer/fuel ratio

The results showed that the PEMAD system detects measurable changes in combustion characteristics when the oxidizer/fuel ratio is varied (Table 14). Significant differences in combustion duration were observed between all ratio variations. Combustion duration decreased systematically with increasing oxidizer content (as expected [9]), demonstrating the system's ability to capture differences in combustion duration.

In terms of flame color (RGB distribution), ratio A showed higher variability across repetitions, indicating a less uniform mixture, while Ratio B and C produced more consistent results (i.e., less spread across repetitions). A significant effect of the oxidizer/fuel ratio on flame color was observed, with only ratios A and C showing a clear difference. This was expected because ratios B and C were visually white, while ratio A produced a darker, less intense flame. The materials were tested as normally received by the NFI, without additional crushing, to replicate the conditions of the hot needle test. Consequently, the grain size was less homogeneous in the mixture, contributing to the observed variability in measured peak intensity across repetitions. Therefore, future work should consider using uniformly crushed mixtures to improve repeatability and reduce this source of variation. Ratio C burned extremely fast (comparable to flash powder), resulting in fewer than five frames being detected for the selected fiber, which limited the accuracy of peak intensity identification and confirms the indicated upper limit.

Overall, the PEMAD detects shifts in the combustion characteristics between mixtures with varying oxidizer/fuel ratios, showing its potential for characterizing and classifying unknown compositions once sufficient reference data and camera with higher frame rates are available. In this study, only one composition was tested with these varying ratio conditions.

5.6.4 Selectivity/specificity: grain size

The grain size tests showed that the combustion behavior is influenced by variations in particle grain size, and the PEMAD detected corresponding shifts in combustion duration, color, and peak intensity across fine, medium, and coarse grain sizes. For the tested compositions it showed that fine grains consistently burned the fastest, with low peak intensities and compositions medium grains showed intermediate burning behavior in both duration and intensity, while coarse grains burned the slowest but reached the highest peak intensities. This behavior is explained by differences in surface area[9]: fine grains ignite quickly and burn fast due to their high area-to-volume ratio, but release less energy per grain, resulting in lower peak intensity. In contrast, coarse grains ignite slower but burn longer and brighter, producing higher peak intensities.

Coarse grains showed the largest spread in both combustion duration and peak intensity, indicating less uniform mixtures (i.e., variations of grain sizes, shapes, or mixing in the coarser powders) with lower reproducibility. In contrast, fine powders showed more reproducible results across repetitions, suggesting that smaller particle size have more uniform mixtures. The RGB distributions further confirmed these findings. Fine powders consistently showed low red channel values, indicating weaker combustion with darker red color. As grain size increased, the red channel had higher values, indicating intenser combustions with brighter red color.

In this test, variations of gunpowder were used, as these were the only materials available at the NFI with clear different grain sizes. The low combustion intensity again caused orange flames to be detected as red, con-

firming the previously established color detection limit of the PEMAD system.

For future work, additional validation tests should be performed using a wider range of composition types beyond gunpowder to confirm the findings and strengthen the trends regarding grain size effects on the combustion characteristics.

5.6.5 Robustness: volume variations

After reassessment, the predefined RSD criterion ($\leq$5%) should be interpreted here as a sensitivity threshold rather than a strict performance requirement. If the RSD $\leq$5%, the system can be considered robust to small volume deviations and do not affect the results, indicating good reproducibility. When RSD>5% between the volume conditions, it indicates that accurate and precise sample preparation is required to obtain reproducible results.

Combustion duration and peak intensity measurements both exceeded the RSD threshold between volume cases, suggesting that these characteristics are influenced by small variations in sample volume ($\pm$ 10%). In contrast, the RGB color channels remained relatively stable, indicating that color detection is less sensitive to volume differences. However, this robustness test was based on a single white combustion composition type with high and stable RGB values. Colored compositions would likely show greater variability. The larger spread in blue and green for the higher volume case can be explained by fewer bright frames above the 50% peak-intensity threshold (faster, so fewer frames), whereas the V-90% case produced more white frames due to longer combustion and thus more stable and higher RGB values.

To improve reproducibility, it is recommended to reduce operator dependence by further automating sample provision (e.g., integrating an automated weighing tool) and extending robustness testing to a wider set of different compositions.

Finally, results showed that even small variations in volume produce measurable differences in combustion behavior, indicating that the hot-needle test is similarly volume sensitive, which underscores the need for a more objective alternative method. However, it also provides an opportunity for future work. By performing a series of combustions with multiple standardized sample volumes (e.g., from 80 to 120% of the standardized sampling spoon) in the PEMAD and plotting the resulting burning-rate/peak-intensity vs. standardized volumes to quantify how rapidly combustion severity increases with volume. This can show how dangerous a composition is in larger volumes and can help the NFI to strengthen the scientific information used in legal court cases.

5.6.6 PEMAD workflow repeatability (precision)

The technical precision tests showed high measurement repeatability of the PEMAD system under constant and controlled illumination. For all measured combustion characteristics, the RSD values between repetitions remained well below the predefined 5% criterion.

Grain size test compositions and colored torches were excluded from the precision analysis for actual combustion events. SEM analysis (Figures 132–134) confirmed variations in particle dimensions within grain size groups, resulting in inconsistent combustion behavior for coarser grains (evaluated in section 5.6.4). The torches were also excluded due to inconsistent combustion behavior in both duration and intensity, with flickering.

The repeatability of the entire PEMAD workflow, including sample preparation, was assessed in this validation precision test (section 4.6.6). The tested mixtures were not uniformly crushed or precisely weighed, in order to mimic NFI casework conditions. As a result, small variations in grain size (section 4.6.4), the sample volume (section 4.6.5, and the way the sample falls and spreads onto the tray during loading all influence the combustion behavior. Although partial sampling automation helped minimize some of these effects, complete elimination of variability is not realistic without weighing (or equivalent) or precisely controlling each sample (i.e., uniformly mixed with homogeneous grain sizes). For these reasons, the RSD values between repeated tests reflect the variability present in the entire PEMAD workflow, primarily caused by sample preparation rather than by the measurement system itself. Furthermore, this underscores the need to replace the "hot needle test" once again, since two scoops of the same composition can show considerable differences in combustion behavior.

5.6.7 Characterization and distinguishing pyrotechnic compositions

The results show that the PEMAD system has the potential to distinguish between different pyrotechnic and explosive compositions based on combustion characteristics (Figures 57-59). The large variation in combustion durations between the different groups (Figures 57), from short combustions such as flash powder and ratio C to longer durations of torch compositions, shows the differences between the compositions which are detected by the system. However, combustion durations of fine grain size, ratio A/B, and volume variations show less clear differences, making them harder to distinguish. Therefore, not every composition can be distinguished based on one single metric (i.e., combustion duration). Additional measures, such as flame color and intensity are needed to further distinguish the compositions that are closer related in terms of combustion characteristics. For instance, grain size tests produced red flames with increasing brightness, while ratio and volume variation tests resulted in white flame colors due to values across all color channels. These differences confirm that flame color can serve as an additional distinguishing metric when the duration is not sufficient. Finally, the peak intensity data also proved to be distinct for different compositions.

Conclusion

This study showed that the PEMAD system is calibrated and validated for the detection and characterization of a selection of solid pyrotechnic and explosive compositions. The calibration, validation and system performance showed both strengths and limitations.

- **Design:** User-friendliness was improved by implementing a larger touchscreen, and overall safety was improved.
- **Color calibration:** reduced the error of the PEMAD measurements compared to the reference spectrometer from 27–35% to 2–3%.
- **Intensity calibration:** enabled pixel intensity values to be converted into Lux, providing interpretable and directly comparable data between fibers, with some limitations.
- **Detection limits:** reliable measurements for slower, lower intensity combustions. Only partially succeeded to capture sufficient frames for very fast combustions, establishing a current upper limit for detection. As a result, the peak intensity of such fast combustions could not yet be determined as reliably as for lower speeds.
- **Selectivity/specificity (color detection):** clear differences were measured between red, green, blue, and purple torch compositions, while closely related mixtures were more difficult to distinguish. Orange flames are, for low intensities, still challenging to distinguish from red flames.
- **Selectivity/specificity (grain size variation):** the system clearly detected considerable effects of grain size on combustion behavior. Finer grains burned fast with small variability and low intensity. Coarser grains burned more slowly, less consistently and with brighter flames.
- **Selectivity/specificity (oxidizer/fuel ratio variations):** the system detected shifts in combustion behavior as component ratios changed. An increasing oxidizer content showed increased speeds and intensities.
- **Robustness (volume variations):** even small variations in sample volume showed measurable differences in combustion behavior, underlining the importance of strict standardization of the samples preparation process for reproducible results.
- **Repeatability/precision:** the reproducibility was influenced by the material properties and preparation of the compositions. Variations in composition and preparation caused measurable differences in combustion behavior, whereas the technical precision tests confirmed excellent repeatability of the PEMAD measurements itself.
- **Overall composition differentiation:** each composition was characterized by quantified combustion characteristics. Comparing compositions based on these metrics demonstrates the potential to distinguish between a wide range of pyrotechnic/explosive compositions with the PEMAD's results. Individual metrics were sometimes insufficient alone to differentiate two compositions when they show relatable combustion behavior.

In summary, calibration showed substantial improvements, and validation demonstrated that the PEMAD system produced quantitative and objective data for combustion characteristics. The results show both the current limitations and its strong potential as an objective method for the characterization and identification of unknown pyrotechnic/explosive compositions in forensic applications, providing and strengthening information/evidence that can support legal court cases.

Appendix I: device improvements made

This appendix shows the main hardware and safety improvements made to the PEMAD system, including wiring adjustments for the emergency switch.

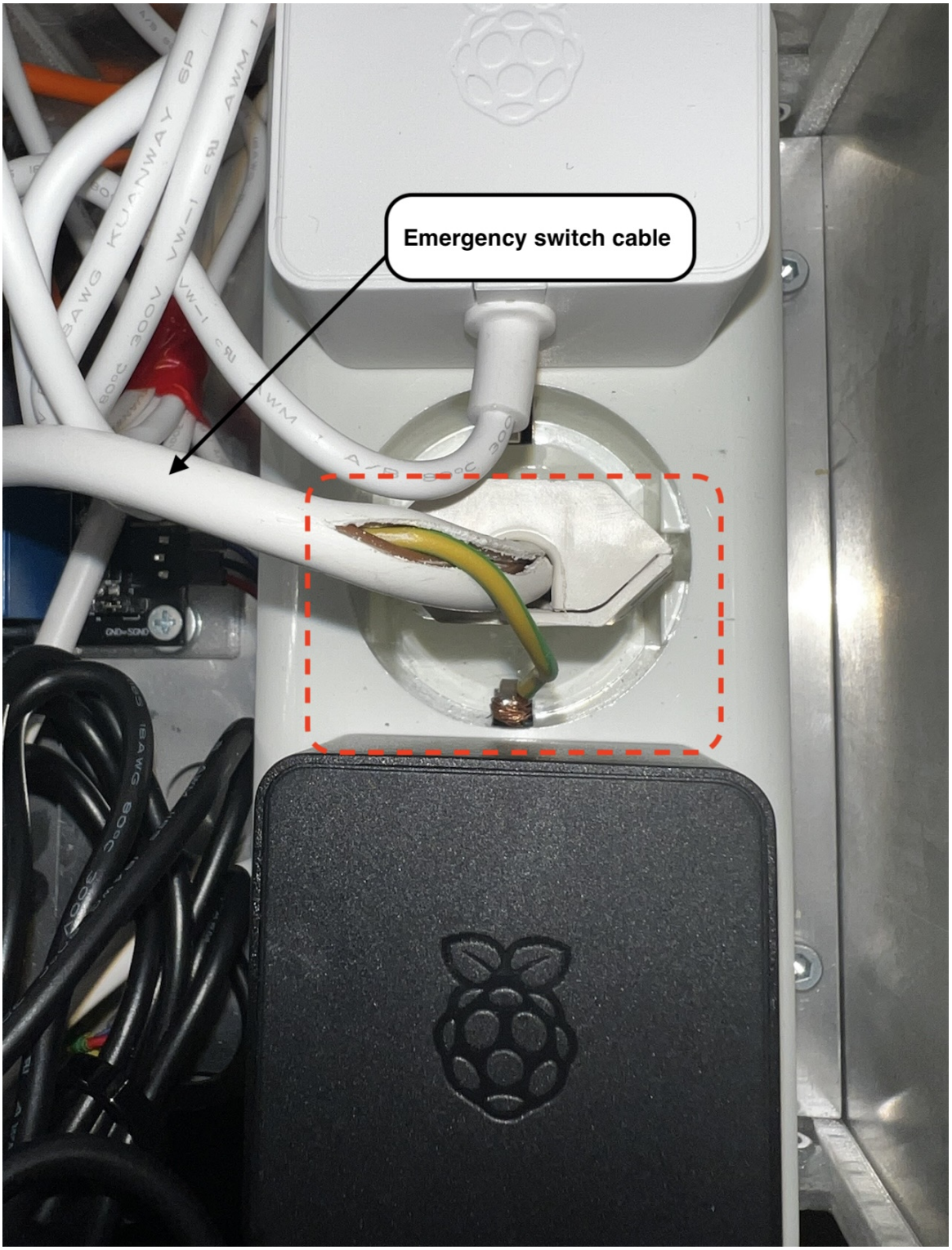


Figure 60: Internal view of the PEMAD control unit showing the previous unsafe grounding connection of the emergency switch cable. The ground wire was only wrapped around the plug base as a temporary solution (red dotted square), with a risk of overheating and short circuit.

Figure 61: Corrected emergency switch connection and grounding setup inside the PEMAD control unit. The cables are now properly insulated and permanently connected in accordance with the NEN-3140 safety standard. The grounding wire is securely fixed using a standardized grounded plug, minimizing the risk of overheating or short circuits.

Appendix II: pilot test report - 04/04/2025

This appendix contains a short pilot test plan for small initial tests. The plan shows proposal to the NFI. In short, the objective of this pilot test is to evaluate the functionality and current state of the PEMAD for measuring different types of pyrotechnic combustions. The test will assess whether the device can distinguish between slow and fast combustions and if the device itself is functional, safe and robust. This test should reveal potential areas of failure or where the device could be improved. In addition, this test is used to verify the functions and capability to quantify combustion characteristics data.

This test is expected to take one day at the NFI, with an approximate duration of 2-3 hours. Additionally, if possible, access to a variety of pyrotechnic mixtures with different kinds of behavior (see Appendix II.II for the specific materials required) and a testing environment with fume hood will be required.

II.I Objectives

1) **Testing the ability to capture combustion characteristics of a wide range of different pyrotechnics.**
   - **Is the device able to capture and distinguish slow and fast combustions?**
     Criteria for success:
     - The device must clearly distinguish between slow and fast combustions. The result is plotted in corresponding figures (for each fiber) and the distinction is easy to observe.
     - The device must detect and plot fast intensity changes accurately. Fast combustions have a steeper rate of change, reaching peak intensity faster. The device's temporal resolution is sufficient if the results show distinct, sharp peaks in the bar plots that correspond to rapid intensity changes (i.e., no smoothening of the graph due to missing fast intensity peaks).
     - The device should provide the duration of combustion for both slow and very fast pyrotechnics.
   - **Does it give sufficient and logical data/results?**
     Criteria for success:
     - The device should be able to follow expected patterns (i.e., if there are two bright flashes, it is expected that there are two intensity peaks with in between low intensity.) This should be confirmed by comparing visual footage with the analysis outcomes.
     - The device should not miss any data points (i.e., gaps in the bar plot, which indicate a missing data point)
   - **Is the array of ND filters sufficient?**
     Criteria for success:
     - The device should be able to analyze at least one fiber that is not overexposed for all combustion types. Otherwise, a higher fractional transmittance has to be applied or better quality ND filters (e.g., glass ND filters with more accurate and exact known transmittance values.)
     - The device should not show any color haze on the captured videos
     - The device should be able to analyze captured videos in the optimal intensity window (i.e., pixel value is between 140 and 225, This fiber selection should be done automatically as well).
   - **Is the device able to capture enough non-zero frames for analysis?**
     Criteria for success:
     - The device should be able to capture at least five non-zero frames for analysis to determine the peak intensity accurately
   - **Is there any interference of reflection on the cameras' sensor?**
     Criteria for success:
     - There is no or minimal reflection from the protection screen on the sensor, with no significant interference on the captured frames
     - There is no or minimal reflection from the sheet metal environment on the sensor, with no significant interference on the captured frames
2) **Testing functionality and user-friendliness**
   - **Do all the functions and buttons work without incorrect responses or errors?**
   - **Do all the steps that have to be taken make sense?**
     Criteria for success:
     - No redundant steps that slow down the process
     - The steps that have to be taken are in a logical order
   - **Check what can be better (e.g., does it make sense to implement a (bigger) touchscreen)**
   - **Is the device safe to use and robust?**
     Criteria for success:
     - There should be no loose parts or cables after multiple tests (>10)

  - There should be no broken parts (tray, ignition module, limit switches, etc etc) or cables after multiple tests (>10)
  - There should be no fire damage on the electrical parts/mechanical parts and the camera's protection shield
  - There should be no exposed cables or other electrical parts that could harm the user's safety
3) **Sampling mechanism**
   Criteria for success:
   - It is possible to provide the sample onto the tray without spilling
   - The sample is provided in identical way every time

II.II Materials and sample preparation

The mixtures were prepared and mixed by an NFI researcher prior to testing. Each sample was provided to the PEMAD using a standardized sampling spoon with a fixed volume of V=0.134 cm^3, minimizing variation in sample amount. The sample was then provided into the ceramic (inert) bowl via an automated sampling mechanism with funnel part and operated by a servo motor, ensuring consistent and uniform material provision while minimizing influence of the researcher. A range of pyrotechnic mixtures, varying from slow to fast combustions with different colors and intensities, was tested. This variety allowed evaluation of the PEMAD's ability to capture different combustion behaviors.

Materials supplied by NFI:
1) Variety of fast and slow pyrotechnic mixtures ($\pm$ 4 samples all with volume of 0.134 cm^3). Samples included, gunpowder and flash powder variations.
2) Variety of pyrotechnic mixtures with different kinds of behaviors, such as flickering, one flash, bright flash, different colors, etc. (all samples with volume of 0.134 cm^3). The specific composition of the samples was not relevant for this stage of testing, as the main objective of this pilot test was to evaluate the PEMAD's functionality across different combustion behaviors. Therefore, an NFI researcher selected available in-stock pyrotechnic mixtures, which were not all documented in detail.
3) 3 samples with volume of 0.134 cm^3 with known ratios/concentrations. Al/Barium nitrate, with ratios 30/70, 20/80, and 10/90 (m/m%).
4) Safety clothing (e.g., glasses, lab coat)

Materials supplied by Nigel:
1) PEMAD and power cable
2) Mouse and keyboard
3) Laptop or pen/paper
4) External flash drive or hard drive
5) Dossier and user-manual
6) Funnel part
7) Cleaning equipment
8) USB-hub

II.III Methods

The pilot test was performed at the NFI to evaluate the functionality, performance, and robustness of the PEMAD system. The test aimed to verify the system's ability to record and distinguish various combustion behaviors and identify potential technical limitations. All tests were conducted under identical environmental conditions, including (laboratory) room temperature, lightning conditions and the PEMAD placed in a fume hood.

- Device performance: A range of pyrotechnic mixtures, from slow to fast combustions with varying intensities and colors, was tested. Each sample was analyzed to verify whether the PEMAD could detect fast intensity changes, capture combustion duration, and differentiate between slow and fast combustions.
- Sensitivity: The results of mixtures with different compositions and behaviors were compared to assess whether the device was sensitive enough to detect differences in combustion characteristics.
- Temporal resolution and ND filter evaluation: The system's temporal resolution was evaluated by verifying whether at least five non-zero frames were captured during fast combustions, allowing accurate representation of fast intensity combustions. The array of ND filters was also checked to confirm that at least one fiber remained within the optimal pixel intensity range (140–225), ensuring the availability of a usable fiber outcome for analysis.
- If the mixtures and their ratios/concentrations are known, it is possible to compare the results, and to see if the PEMAD sensitive enough, such that it is able to provide results which show shifts in combustion characteristics when the ratios change.

II.IV Results and discussion

This part will reflect on the objectives mentioned in section; Objectives.

1) **Testing the ability to capture a wide range of different pyrotechnics.**
   - **Is the device able to capture and distinguish slow and fast combustions?**
     Criteria for success:
     - ✓The device must clearly distinguish between slow and fast combustions. The result is plotted in corresponding figures (for each fiber) and the distinction is easy to observe.
     - ✓The device must detect and plot fast intensity changes accurately. Fast combustions have a steeper rate of change, reaching peak intensity faster. The device's temporal resolution is sufficient if the results show distinct, sharp peaks in the bar plots that correspond to rapid intensity changes (i.e., no smoothening of the graph due to missing fast intensity peaks).
     - ✓The device should provide the duration of combustion for both slow and very fast pyrotechnics.

     The PEMAD successfully captured both slow and fast combustions, with distinct differences in rate of intensity change visible in the results (Figures 62 and 63). Slow combustions displayed a gradual increase and decrease in intensity, while fast combustions showed sharp peaks with much shorter durations.

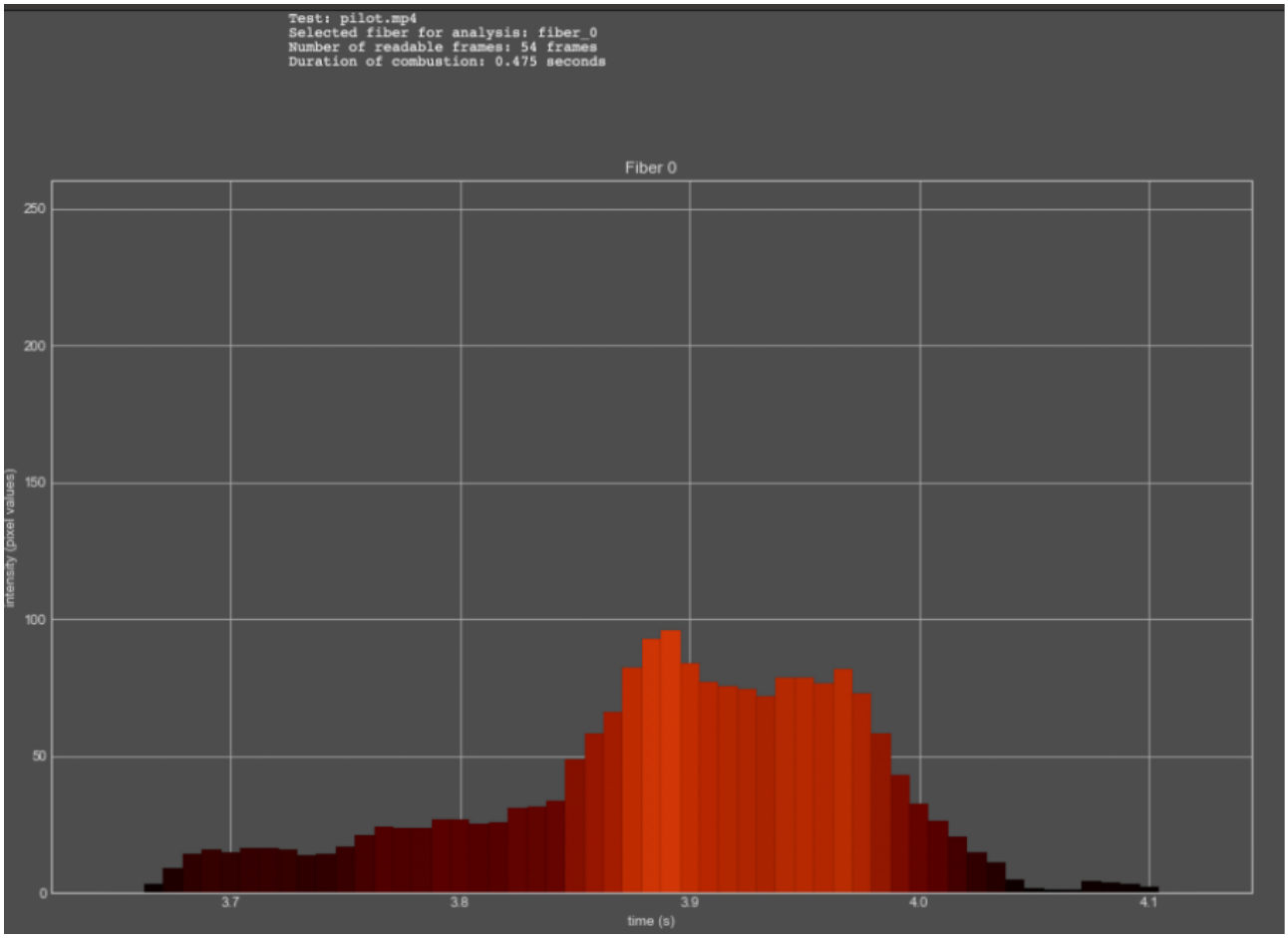


Figure 62: Slow and low intensity combustion captured with Fiber0

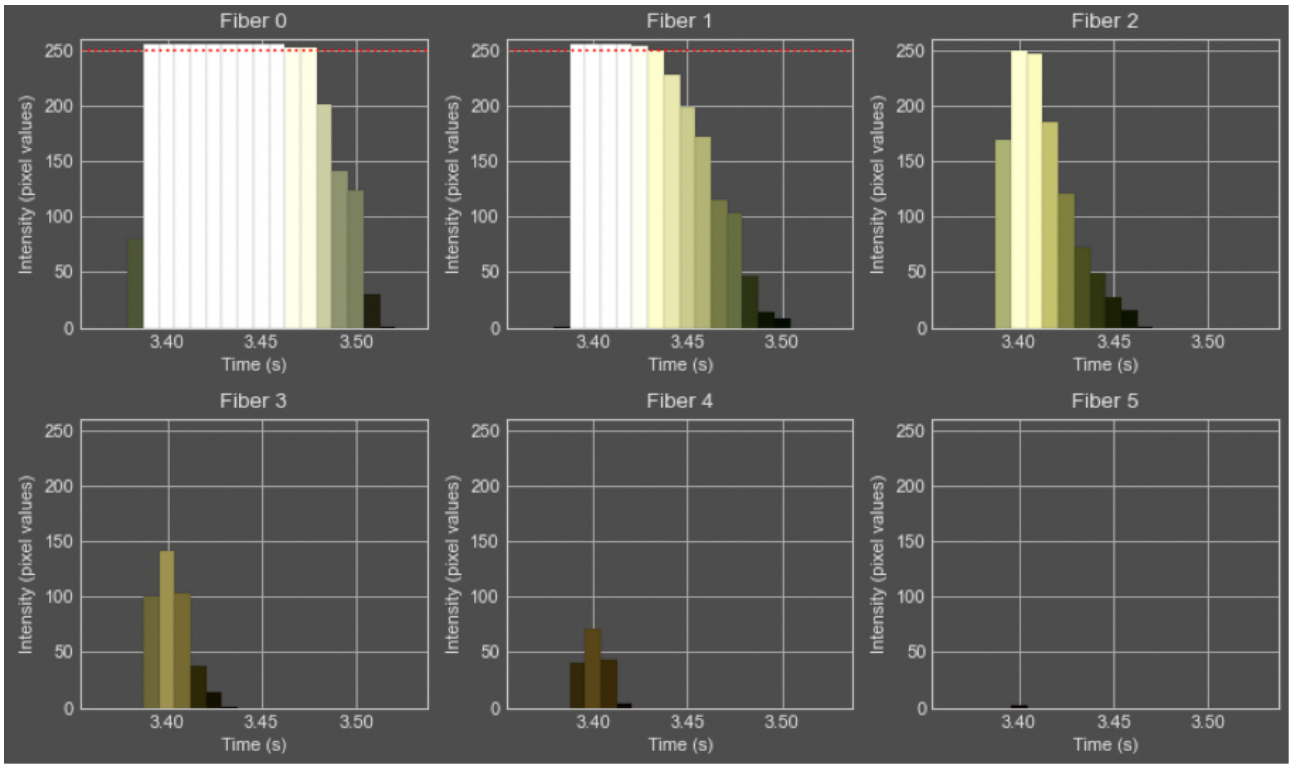


Figure 63: Fast and high intensity combustion captured with multiple fibers

- **Does it give sufficient and logical data/results?**
  Criteria for success:
  - ✓The device should be able to follow expected patterns (i.e., if there are two bright flashes, it is expected that there are two intensity peaks with in between low intensity.) This should be validated by comparing visual footage with the analysis outcomes.

    This is also successful. Most of the test results matched the expected results which were observed during the hot needle test. However, not in all cases the visual camera was synchronized well, in these cases it was not possible to see if the visuals matched the combustion results. In Figure 64 can be observed that there is a pattern, with first a real bright flash, and secondly a lower intensity burining period

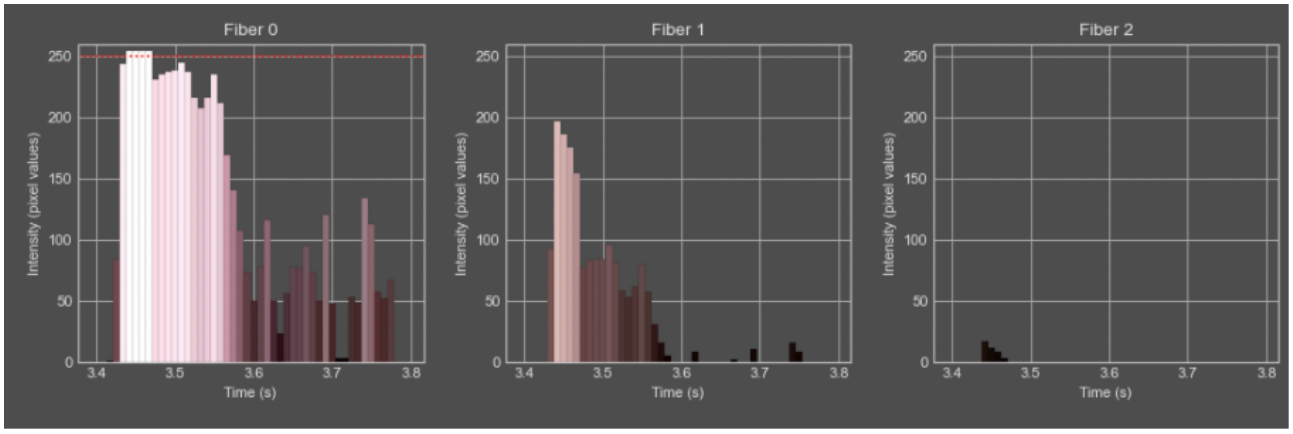


Figure 64: Combustion with high intensity and low intensity behavior

  - ✗/✓The device should not miss any data points (i.e., gaps in the bar plot, which indicate a missing data point)
    No missing data points were observed in the analyzed graphs. However, random errors emerged: in some cases the fiber camera only recorded a dark view (no data), where the visual camera did observe a combustion. In those cases, when analyzing the video, a index error raised in python, which froze the system and had to be rebooted.

- **Is the array of ND filters sufficient?**
  Criteria for success:
  - ✓The device should be able to analyze at least one fiber that is not overexposed for all combustion types. Otherwise, a higher fractional transmittance has to be applied or better quality ND filters (e.g., glass ND filters with more accurate and exact known transmittance values.)
  - ✓The device should not show any color haze on the captured videos
  - ✓The device should be able to analyze captured videos in the optimal intensity window (i.e., pixel value is between 140 and 250, This fiber selection should be done automatically as well), otherwise more ND filters need to be added to minimize the step-size for fractional transmittance of the ND filters

  Each test without the random index error was able to analyze at least one fiber that was not overexposed. In addition, all fibers that were captured for more than one fiber (not only Fiber0), were analyzed in the desired and optimal pixel intensity region (140-225). Examples can be seen in Figures 63, 65, 66.

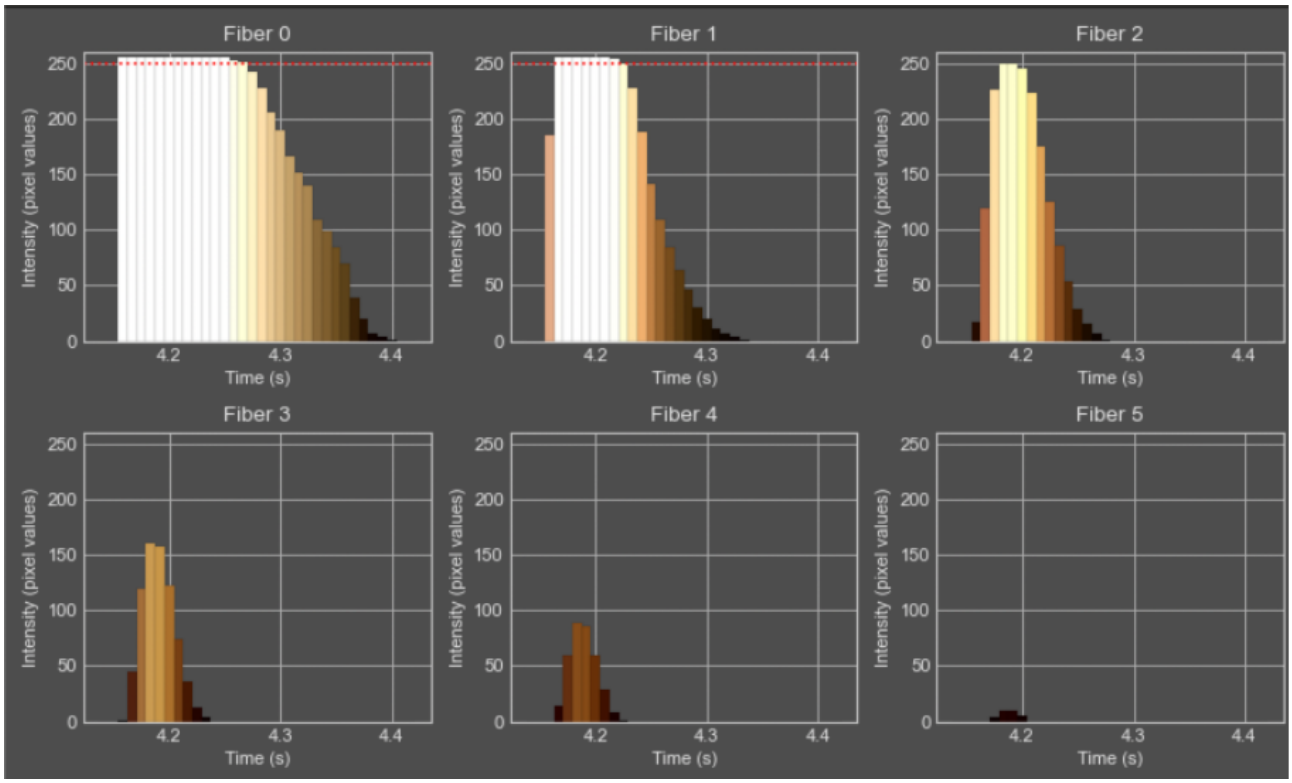


Figure 65: Combustion with high intensity; at least one fiber is not overexposed, and analysis is between 140 and 250 pixel value

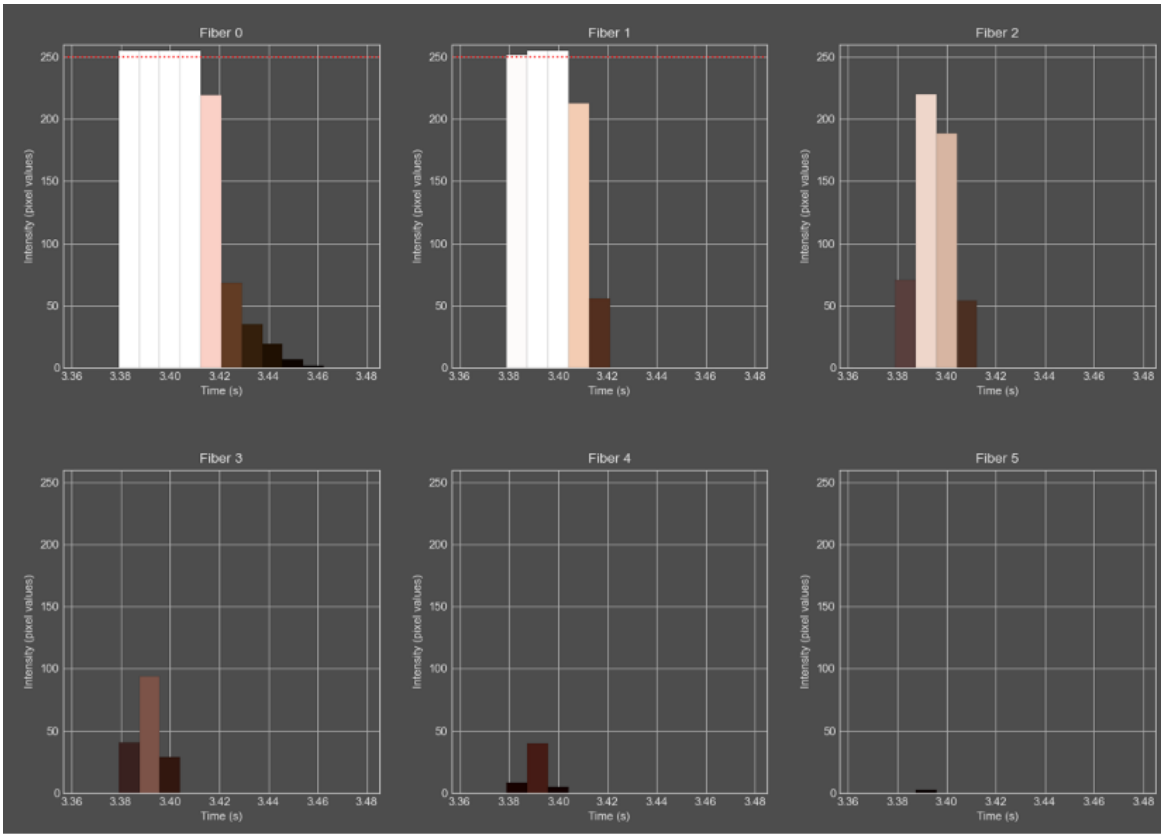


Figure 66: Combustion with high intensity; at least one fiber is not overexposed, and analysis is between 140 and 250 pixel value

- **Is the device able to capture enough non-zero frames for analysis?**
  Criteria for success:
  - ✗/✓The device should be able to capture at least 5 non-zero frames for analysis

    For almost all test the minimum of five non-zero frames were achieved. However, sometimes for flash powder, 3 to 5 frames were captured for the selected fiber (not always). In addition, The last overexposed fiber always give more than five frames. To clarify, if it is an high intensity combustion it always has five non-zero frames, but not always for the selected fiber, see Figure 67. More examples are shown in Figure 66 and 68. An explanation for this observation could be that the tests with fewer than five non-zero frames involved mixtures that were manually prepared to explore the red coloration of strontium. These mixtures were sometimes less reactive, resulting in weaker or shorter combustions.

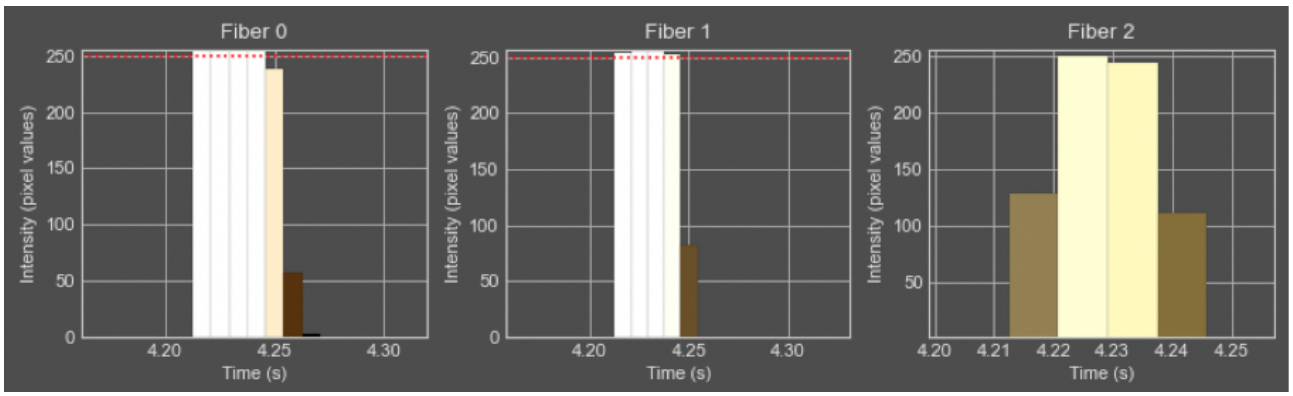


Figure 67: Very fast combustion with less than 5 non-zero frames at the selected fiber

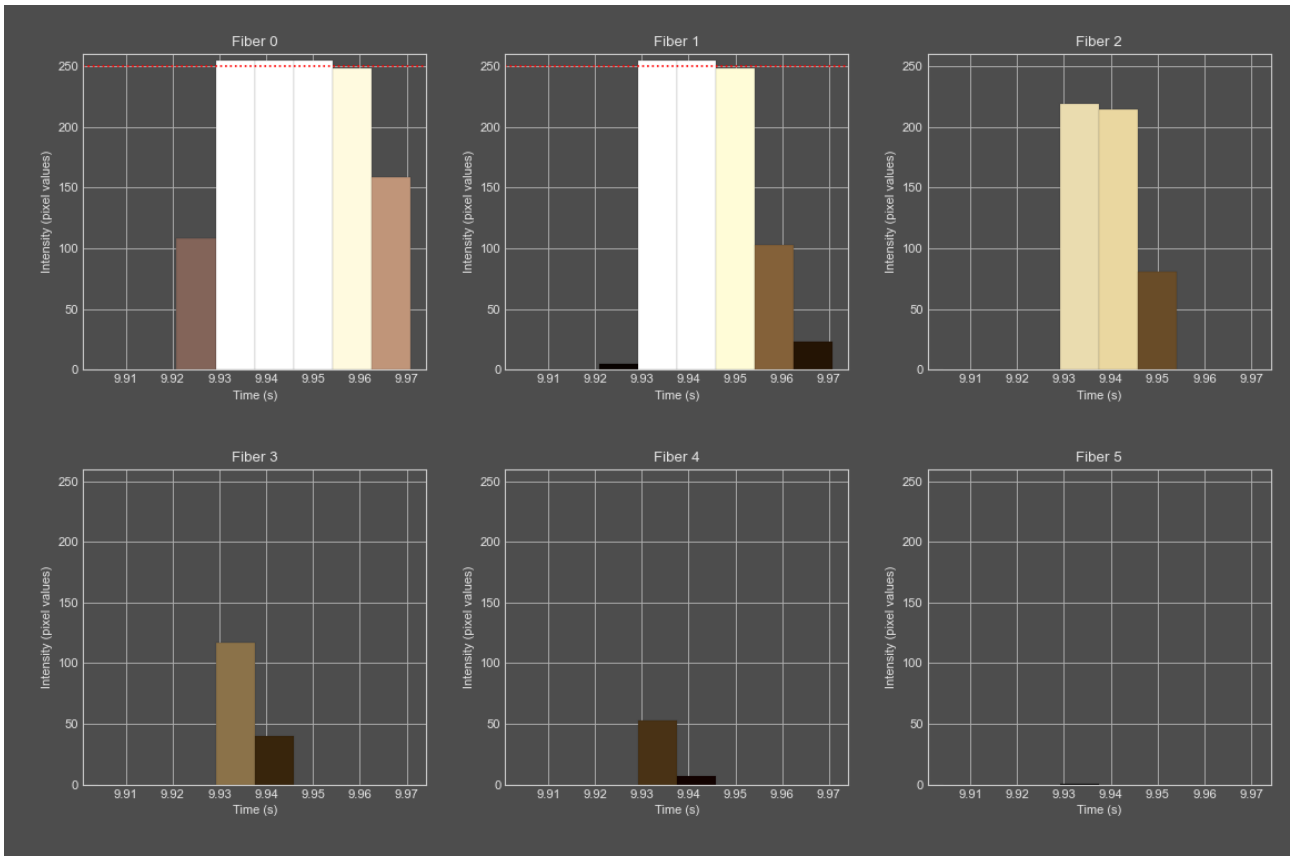


Figure 68: Combustion with less than 5 non-zero frames at the selected fiber

- **Is there any interference of reflection on the cameras' sensor?**
  Criteria for success:
  - ✓There is no or minimal reflection from the protection screen on the sensor, with no significant interference on the captured frames
  - ✓There is no or minimal reflection from the sheet metal environment on the sensor, with no significant interference on the captured frames

    No or minimal reflection from the surroundings was seen on the visual camera or fiber camera, as the focus of the lens was directly on the combustion place. Indicating a proper placed ROI.

2) **Functionality and User-friendliness**

- ✓**Do all the functions and buttons work without incorrect responses or errors?**
  No errors regarding buttons and moving parts, sometimes the door was a bit stuck, push it further with hand
- **Do all the steps that have to be taken make sense?**
  Criteria for success:
  - ✓No redundant steps that slow down the process

    No redundant steps in this process. However, the door opening and closing for sampling is a bit slow, but ensures safety.
  - ✓The steps that have to be taken are in a logical order
    The process for a test has a logical order
- ✗**Check what can be better (e.g., does it make sense to implement a (bigger) touchscreen)**
  - *The sampling mechanism slows down the process, but is necessary. The funnel part did not fit properly. The automated sampling mechanism worked correctly and the spread of the material was more consistent between samples than providing it by hand to the bowl.*
  - *Jacinta and the other NFI researcher agreed that the screen was too small. In addition, it would save space and leave out the risk of forgetting a mouse and keyboard*
- **Is the device safe to use and robust?**
  Criteria for success:

  - ✓There should be no loose parts or cables after multiple tests (>10)
  - ✓There should be no broken parts (tray, ignition module, limit switches, etc etc) or cables after multiple tests (>10)
  - ✓There should be no fire damage on the electrical parts/mechanical parts and the camera's protection shield
  - ✓There should be no exposed cables or other electrical parts that could harm the user's safety

3) **Sampling mechanism**
   Criteria for success:
   - ✓It is possible to provide the sample onto the tray without spilling
   - ✓The sample is provided in identical way every time

II.V Conclusion

In general this test was successful. Most of the criteria were met, confirming that the system was capable of quantifying combustion characteristics and showing differentiation between different kind of combustion behaviors. In addition, the system was robust over the multiple tests conducted. The device was able to capture fast and slow pyrotechnics with different kind of behaviors. At least one fiber was not overexposed for analysis. In some cases the selected fiber for analysis, has less than 5 non-zero frames, but the fiber before the selected fiber has more than 5 non-zero frames in most cases. Besides that the device meets most criteria, there are also areas for improvement.

Several improvement areas emerged from the pilot test:

- Synchronization of the visual camera: the visual camera started to early and ended to early, which resulted in missed combustions. This problematic when results of the fiber camera and the visual frames wanted to be compared.
- Random errors: in some cases the fiber camera only recorded a dark view, where the visual camera did observe a combustion. The video is recorded with the fiber camera and is successfully transferred to a MP4 file. However, when analyzing this video, a index error emerged. The captured frames were all zero. The error resulted in a frozen system, which caused the system to reboot. The RPi can only receive a certain amount of data from the camera, so if the frame rate is increased, the resolution of the footage will drop, this can be a explanation for the random errors. Thus, trying a lower fps can be a solution, keeping in mind the requirement of the minimum of 5 non-zero frames. Or setting longer sleeps in between to give the system longer time to convert the videos.
- Establishing the minimum number of 5 non-zero frames. To see if the non-zero frames <5 result from fast mixtures. More tests have to be performed using fast flash powder.
- Color calibration needs to be performed to ensure accurate color representation. Colors were dark (black) in many cases at low intensity regions. In addition, color were sometimes off compared to the observed visual reference.
- Intensity calibration needs to be performed to ensure accurate, intuitive and standardized (SI-units) data.

Appendix III: Color calibration

This appendix contains additional materials used for color calibration, including clarification sections (Detailed method for spectral measurement and color conversions and Detailed method for calibration mapping and correction), images of experimental setup, graphs, spectrometer data (SPDs), visualization plots (XYZ space and CIE 1931 chromaticity), PEMAD analysis results, and regression analysis plots (correction).

III.I Experimental setup

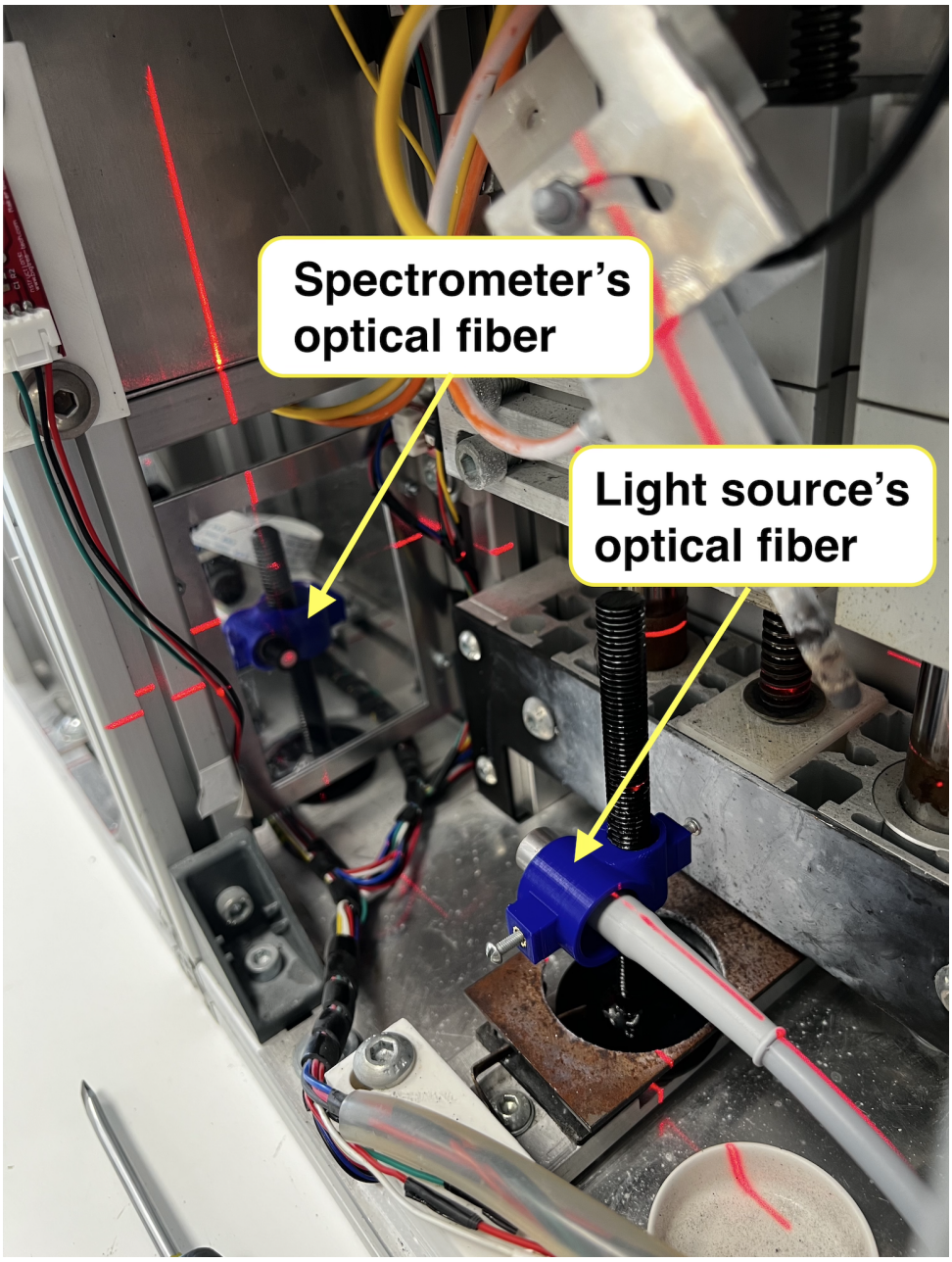


Figure 69: Alignment of the optical fiber of the spectrometer with the optical fiber of the light source

III.II Detailed method for spectral measurement and color conversions

As explained in prior literature study [9], it is possible to evaluate color in multiple color spaces, including the 3D XYZ, 2D CIE 1931 chromaticity, and the RGB space. RGB color spaces are hardware dependent, and thus not standardized, while color spaces like XYZ are standardized. For this reason, it was chosen to calibrate the PEMAD system with a spectrometer measuring tri-stimulus values X, Y, and Z. According to StellarNet Inc. [13], the human eye has three types of cones that respond to different wavelengths of light, which gives the color perception. Colors can be quantified based on the amount to which they stimulate the three cone types, described by the LMS tri-stimulus values. Because the wavelength range of each LMS channel (cones) overlap, direct color calculations are complex. To address this, CIE defined a new set of XYZ tri-stimulus values that are derived from the LMS values, ensuring non-overlapping, and linearly independent from each other. Here, Y represents luminance (brightness), Z primarily corresponds to blue wavelengths, and X covers the remaining wavelenghts. StellarPro software provides these values for a measured spectral power distribution (SPD) of the emitted light. In this study, the CIE 1931 standard observer was assumed, as this correlates color with human perception the best. This standard is based on experimental measurements of human visual response to different wavelengths and provides a standardized model that closely approximates how an average human perceives color [9]. In addition, a standard D65 illuminant was assumed as a reference in post-processing because this corresponds to the sRGB color space and the CIE 1931 standard [9]. This choice aligns with the output of the RPi HQ camera used in the PEMAD system, which provides sRGB images via OpenCV. The sRGB color space is still the most widely used standard in RGB image data. From the full SPD, it is possible to integrate the standard color matching functions (CMF) for the CIE 1931 observer over the VIS region. The color matching functions are multiplied with the obtained SPD, this results in the tri-stimuli value X, Y, and Z, summarized in Equations 2-4

$$X = \int S(\lambda) \cdot \overline{x}(\lambda)\, d\lambda, \tag{2}$$

$$Y = \int S(\lambda) \cdot \overline{y}(\lambda)\, d\lambda, \tag{3}$$

$$Z = \int S(\lambda) \cdot \overline{z}(\lambda)\, d\lambda \tag{4}$$

Where:

- $S(\lambda)$ is the measured spectral power distribution (SPD) of the light.
- $\overline{x}(\lambda), \overline{y}(\lambda), \overline{z}(\lambda)$ are the CIE 1931 CMFs representing the average human eye response at each wavelength.
- X, Y, Z are the resulting tri-stimulus values in the CIE XYZ color space, which quantify the perceived color.

The CMFs are visualized in Figure 70, with the corresponding color representation.

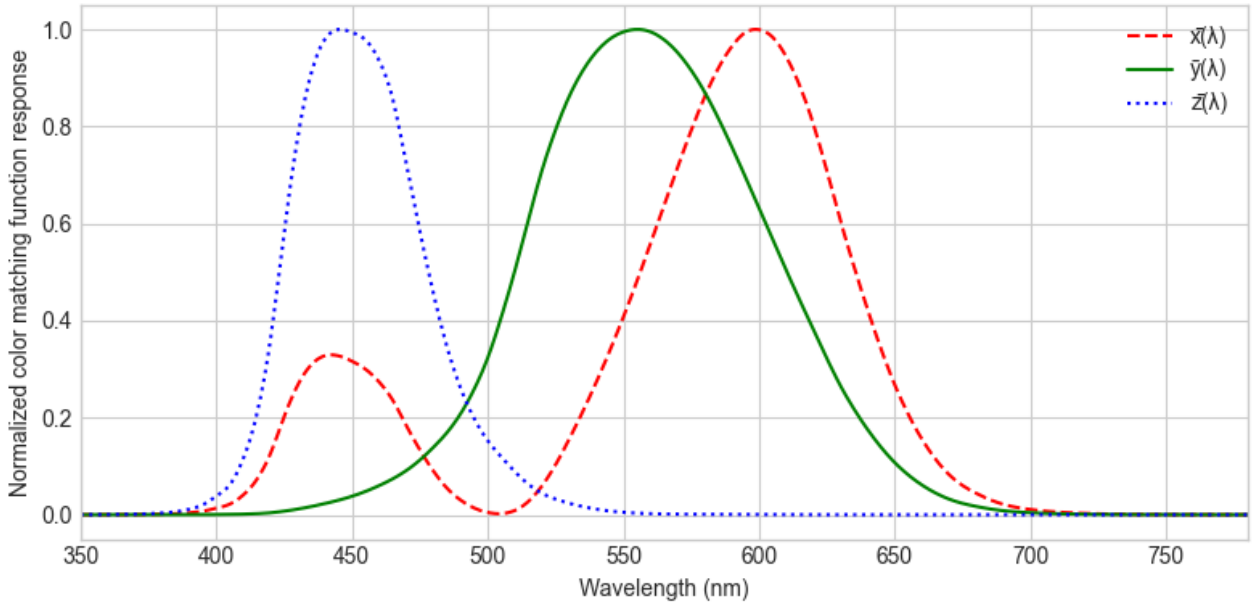


Figure 70: visualization of the CIE 1931 standard color matching functions

With the measured X, Y, and Z values, the corresponding CIE 1931 chromaticity coordinates (x,y), which represent the color appearance, were calculated using the following standard equations:

$$x = \frac{X}{X+Y+Z}, \tag{5}$$

$$y = \frac{Y}{X+Y+Z} \tag{6}$$

These coordinates enable visualization of the color in the CIE 1931 chromaticity diagram and how they relate to the sRGB space used in the PEMAD system and the related post-processing routine with OpenCV.

Although all of the above processing steps are performed with the StellarPro software, understanding each step is essential for verification and reproducibility.

After computing the chromaticity coordinates for each intensity levels of the light source and color filter, the resulting values were visualized in the 3D XYZ color space. By visualizing the chromaticity coordinates in the 3D XYZ color space, each measured color can be represented as a point in the 3D space. When this space is rotated or viewed from a specific angle, it should become clear that changes in intensity affect only the luminance (Y) value. As a result, the data points corresponding to different intensity levels for one color filter should align in one point. This point represents the color in the CIE 1931 chromaticity diagram, which is independent of the luminance.

Finally, the measured X, Y, and Z values were then converted into the sRGB color space using standardized color conversion matrices and gamma correction functions. This conversion provides the reference RGB values for the calibration procedure. All data processing and conversions were performed using custom Python scripts, and evaluated here shortly [24, 25]:

First, the sRGB transformation matrix was applied to the X, Y, and Z values, which gives linear sRGB values:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = [M] \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{7}$$

Where, R', G', and B' are the linear RGB components, M is the standard sRGB transformation matrix corresponding to the CIE 1931 observer and looks like:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 3.2406 & -1.5372 & -0.4986 \\ -0.9689 & 1.8758 & 0.0415 \\ 0.0557 & -0.2040 & 1.0570 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{8}$$

Second, gamma correction was applied to the spectrometer-derived values to enable a valid comparison with the RPi HQ camera's RGB output values. The camera data in the PEMAD system is recorded as .MP4 files and analyzed using OpenCV, which provides gamma corrected sRGB values representing color in a non-linear format, optimized for human perception and displays [9, 26]. On the other hand, the initial RGB values computed from the SPD measurements via the CIE XYZ color matching functions and the sRGB transformation matrix are linear. Without gamma correction, comparing these linear RGB values to the camera's non-linear sRGB output would be invalid. Therefore, gamma correction was applied individually to each color channel, using the following equation: [9, 24, 25].

$$V = \begin{cases} 12.92 \cdot V', & V' \leq 0.0031308 \\ 1.055 \cdot V'^{1/2.4} - 0.055, & V' > 0.0031308 \end{cases} \tag{9}$$

Where:

- V' is the linear RGB component (R', G' and B') obtained from the CIE XYZ to RGB transformation.
- V is the gamma-corrected (non-linear) sRGB component (R, G and B).

Finally, the resulting non-linear sRGB values in the [0, 1] range were scaled to [0, 255] range to match the format used by OpenCV and the PEMAD system.

III.III Detailed method for calibration mapping and correction

Calibration mapping/correction was implemented entirely in a post-processing routine rather than by adjusting the camera's settings for several reasons:

1) **Fixed settings:** The camera settings (e.g., exposure time, gain, white balance, and resolution) were already established, fixed and balanced by a previous research group [8] to ensure consistent measurements, frame rates and reliable data acquisition.
2) **Complexity of manual adjusting settings:** Matching the camera's RGB output to the spectrometer's RGB reference values is challenging when the data includes a wide range of colors and intensities. Manual tuning might improve accuracy for certain colors but can possibly lowers it for other colors, making it difficult to achieve a consistent outcome for all colors. Moreover, The camera and spectrometer use different color spaces to quantify color and have different spectral sensitivities [9]. Therefore, the camera cannot capture all colors with the same accuracy as the spectrometer and some colors measured by the spectrometer may not be reproducible by the camera. The large number of camera settings (e.g., exposure, white balance, gain) further makes it difficult to find a single configuration that achieves consistent accuracy across all colors.
3) **Flexibility:** Performing calibration in post-processing allows the use of regression-based methods to systematically use mapping from camera to spectrometer RGB values. This approach allows to make a reusable correction model that can be applied to new captured data without needing hardware adjustments or checking the settings every time.

The camera-derived RGB values were compared directly with the reference RGB values computed from the spectrometer data. Deviations were observed, which confirms the necessity of calibration. To improve the accuracy of camera RGB measurements and match them more closely with the spectral reference data, a polynomial regression model from scikit-learn library was tried.

Color variations in camera sensors often have non-linear color response due to sensor characteristics, lighting, and settings. The HQ RPi camera uses a CMOS sensor with a non-linear response to light. Therefore, a simple linear model cannot capture these effects and leads to large deviations/errors. Polynomial regression was therefore used, as it can better represent these relationships.

High-order polynomial models have the risk of overfitting, producing unrealistic colors as a result. This was noticed while creating the correction model. The use of a second-order polynomial regression model worked fine but had a considerable large RMSE error for the corrected RGB values. A third-order model was tried, but it changed the colors unrealistically, indicating overfitting (e.g., white color turning purple). Because of this, a new method was sought and ridge regression was used. Ridge regression adds a penalty term (parameter alpha) on the size of coefficients, excluding extreme values which reduces overfitting and improving the model's results [14]. To find the best polynomial degree and penalty (alpha), a grid search was used for degree 1-4, and multiple alpha values (1e-3, 1e-2, 1e-1, 1, 10).

Overfitting can also occur when a model is trained and tested on the same data [15]. This results in good results on known data, but fails on new data. For this reason, ten-fold cross-validation is used as well, where the dataset is divided into ten folds, using nine folds (45 samples) for training and one fold (5 samples) for testing in each split. Shuffling ensures that training and test data is random selected across the folds. The cross-validation is schematically visualized in Figure 71.

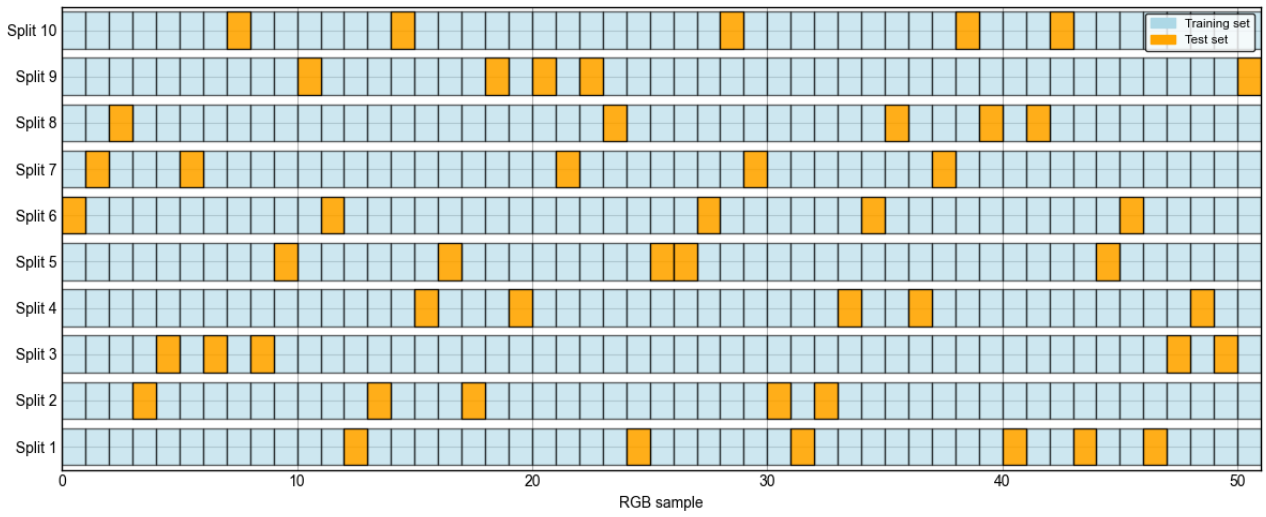


Figure 71: Ten-fold cross-validation visualization: split between training and test data in random shuffle order

The Root Mean Squared Error (RMSE) per color channel and per filter was calculated to quantify calibration accuracy. Plots comparing raw camera, corrected output, and target spectrometer colors demonstrated improvements or errors. RMSE was chosen because it clearly shows how much the corrected colors deviate from the true RGB values. RMSE expresses the error in the same units as the data (RGB values), making the results more interpretable [16]. The Mean Square Error (MSE) calculates errors in squared units.

When the best model was found, it was refitted on the full dataset. The resulting intercept, coefficients, and the the polynomial feature expansion, were then used to map camera RGB values to the spectrometer reference RGB values following the process based on scikit-learn Polynomial Features [17]:

1) **Normalization:** The measured mean RGB values [0-255] are normalized to a [0-1] scale to avoid large numbers in calculations.
2) **Polynomial feature computation:** a set of 20 polynomial features is computed (internally of the function), as the model is of degree 3. To clarify:
   - **Degree 0:** 1 feature
   - **Degree 1:** R, G, B (3 features)
   - **Degree 2:** $R^2, G^2, B^2, R \cdot G, R \cdot B, G \cdot B$ (6 features)
   - **Degree 3:** $R^3, G^3, B^3, R^2 \cdot G, R^2 \cdot B, R \cdot G^2, R \cdot B^2, G^2 \cdot B, G \cdot B^2, R \cdot G \cdot B$ (10 features)

   Thus, each feature is a combination of powers of R, G, and B as specified in the `powers` array in the code. For instance:
   - R^2: captures non-linear effects of red intensity.
   - $R \cdot G$: captures interactions between red and green channels (e.g., how the red color influences green in the camera sensor)
   - B^3: captures higher-order non-linear effects in the blue channel.
3) **Corrected RGB values:** each polynomial feature is multiplied by its corresponding coefficient from the `coef` array obtained from the model, all values are summed together with the intercept from `intercept` to obtain the corrected RGB value:

$$corrected_{RGB} = intercept + \sum_{i=1}^{20} coef[i] \cdot feature[i] \tag{10}$$

   where:
   - `intercept` is a vector representing the baseline correction for the R, G, and B channels (obtained from model).
   - `coef[i]` is a vector of the learned weights (ridge model) for the i-th polynomial feature.
   - `feature[i]` is the computed value of the i-th polynomial feature computed from the combinations in `powers` for actual RGB values.
   - The sum iterates over all 20 polynomial features, capturing non-linear relationships and interactions between the color channels.
   - This process is done internally of the scikit-learn functions.
4) **Output:** The final corrected RGB values replace the raw camera measurements in the analysis.

III.IV Spectral measurements: spectral power distributions

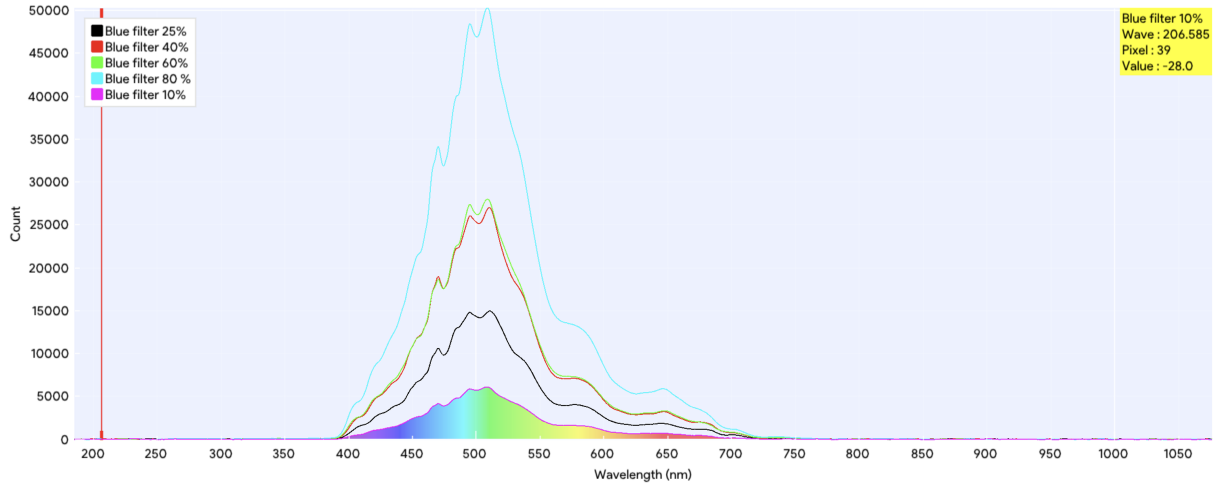


Figure 72: Spectral power distribution of light transmitted through blue filter at light source intensities of 10%, 25%, 40%, 60%, and 80%

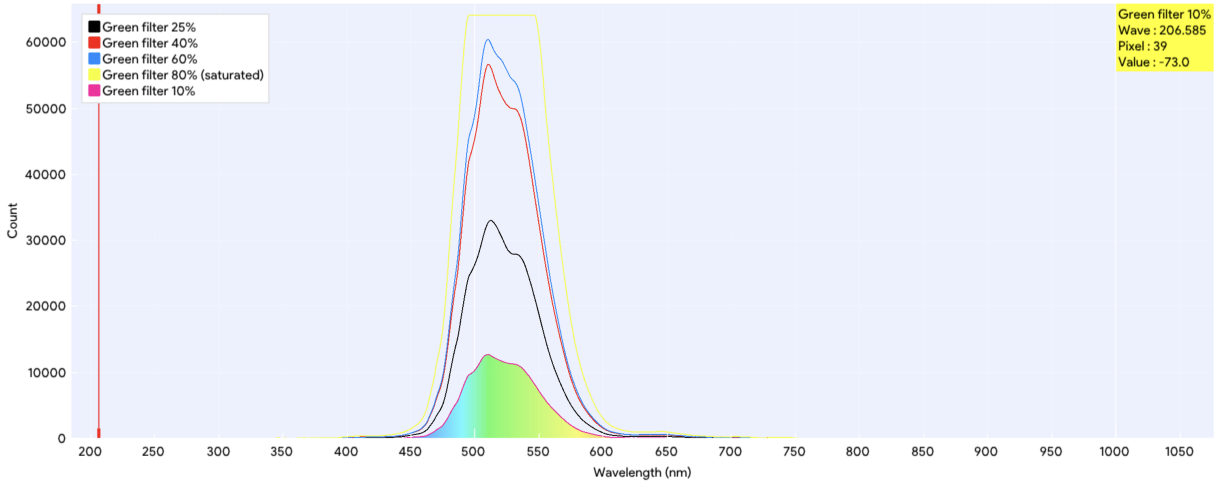


Figure 73: Spectral power distribution of light transmitted through green filter at light source intensities of 10%, 25%, 40%, 60%, and 80%

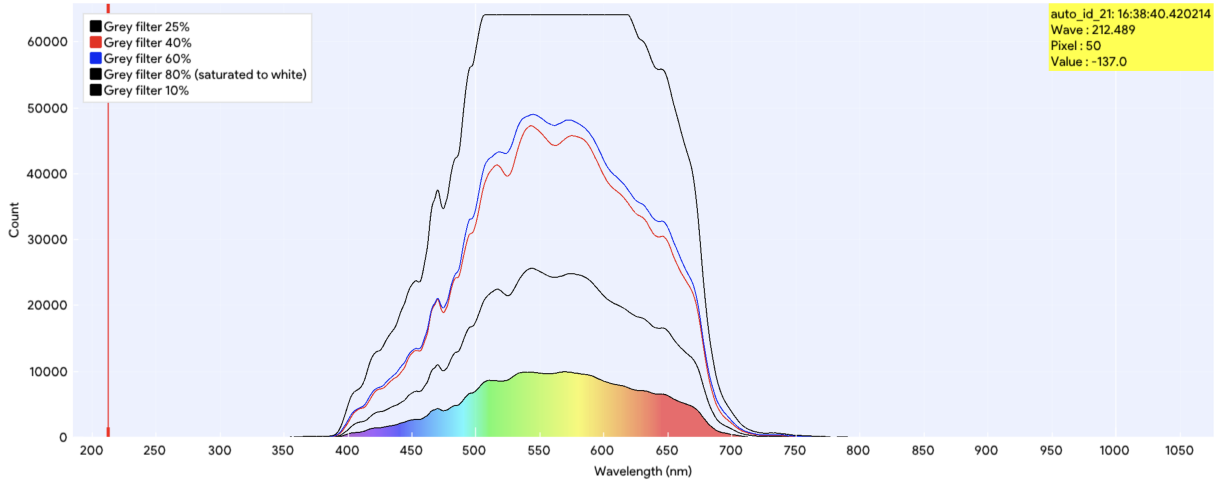


Figure 74: Spectral power distribution of light transmitted through grey filter at light source intensities of 10%, 25%, 40%, 60%, and 80%

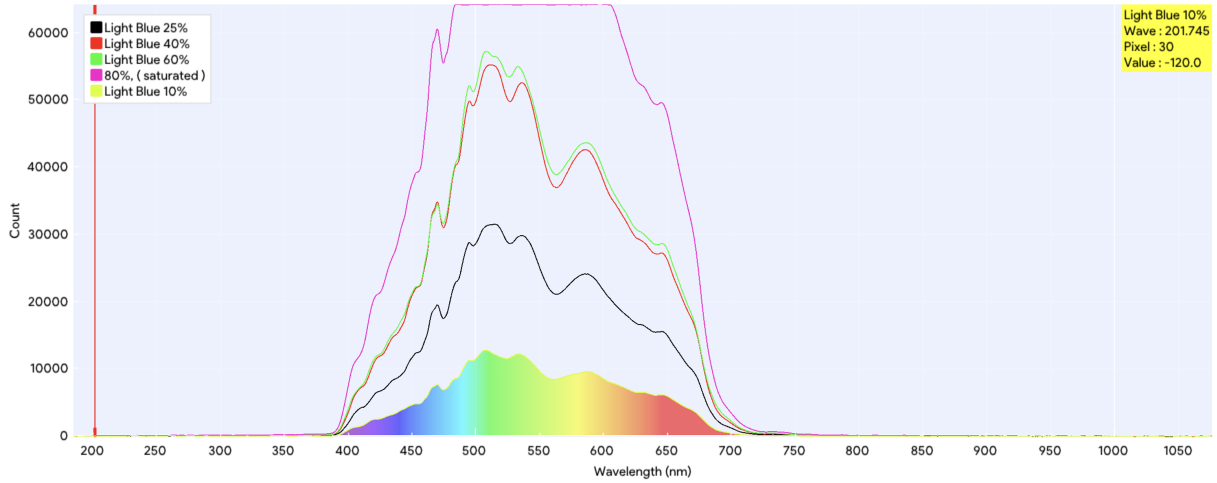


Figure 75: Spectral power distribution of light transmitted through light blue filter at light source intensities of 10%, 25%, 40%, 60%, and 80%

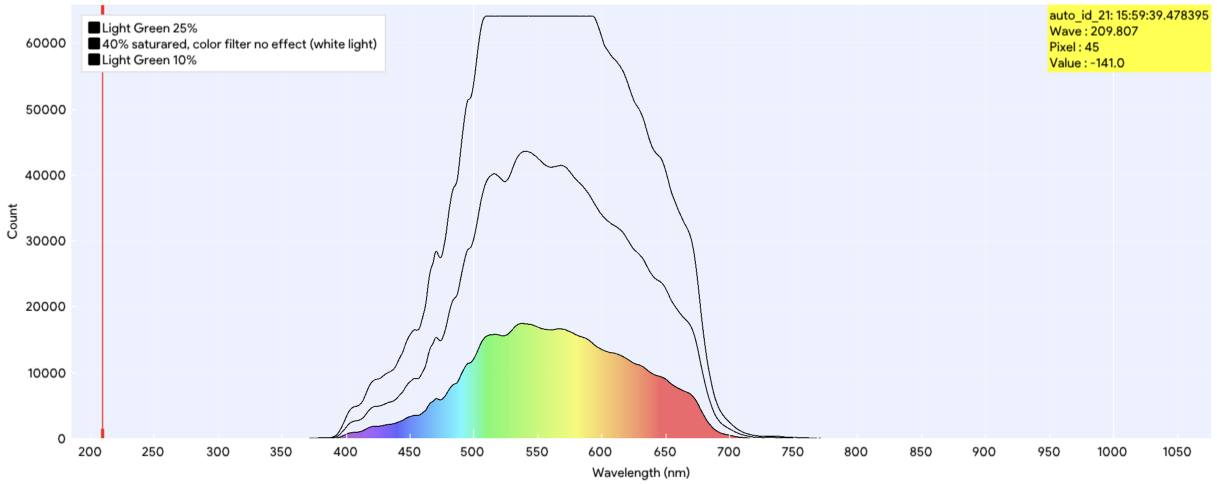


Figure 76: Spectral power distribution of light transmitted through light green filter at light source intensities of 10%, 25%, 40%, 60%, and 80%

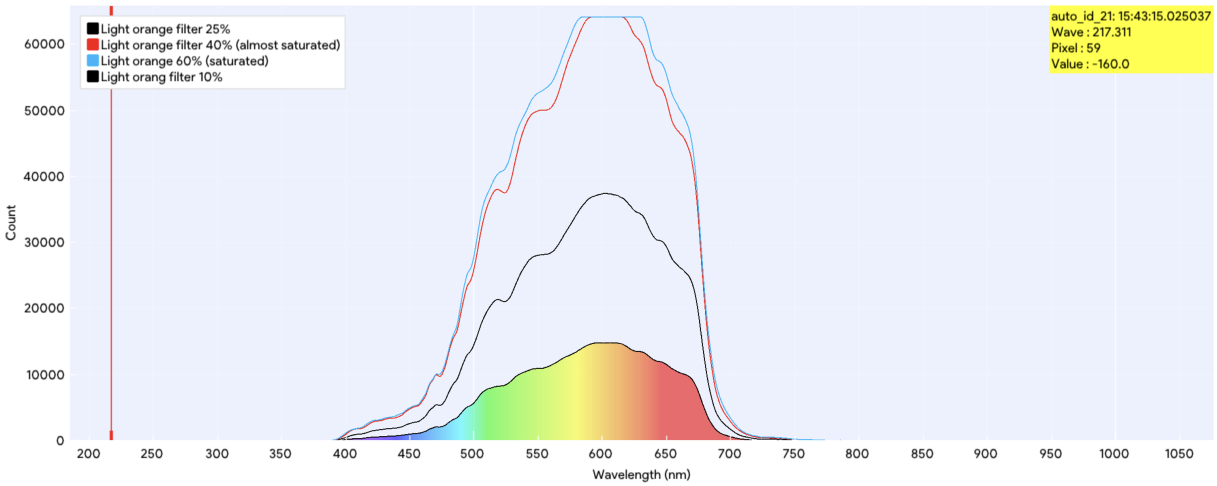


Figure 77: Spectral power distribution of light transmitted through (light) orange filter at light source intensities of 10%, 25%, 40%, 60%, and 80%

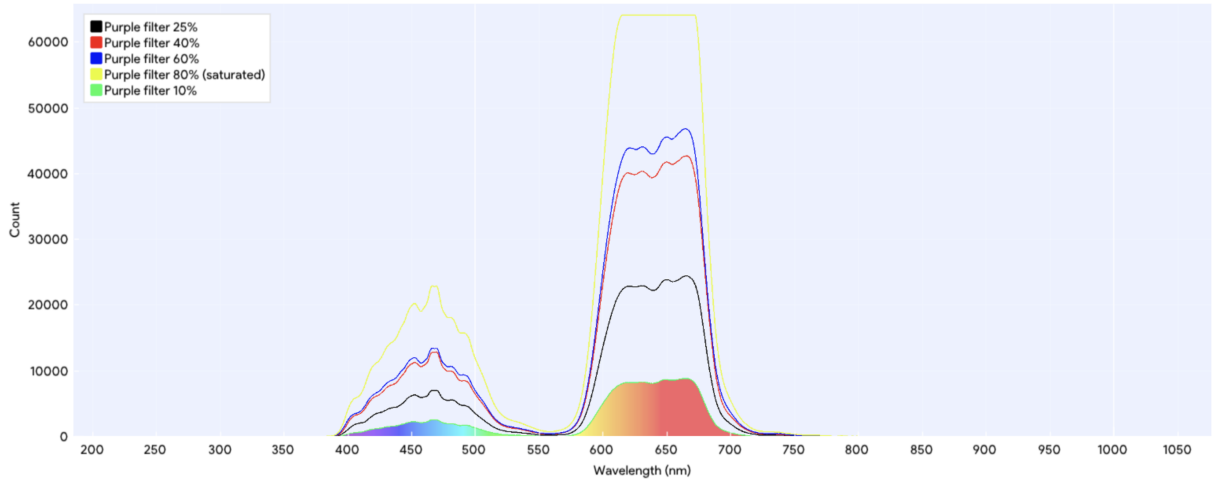


Figure 78: Spectral power distribution of light transmitted through purple filter at light source intensities of 10%, 25%, 40%, 60%, and 80%

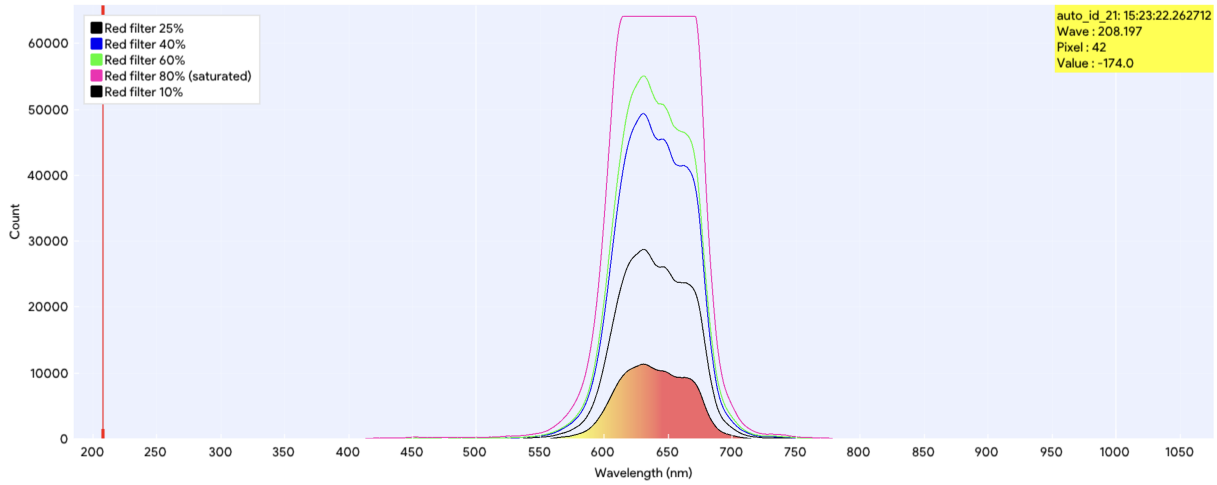


Figure 79: Spectral power distribution of light transmitted through red filter at light source intensities of 10%, 25%, 40%, 60%, and 80%

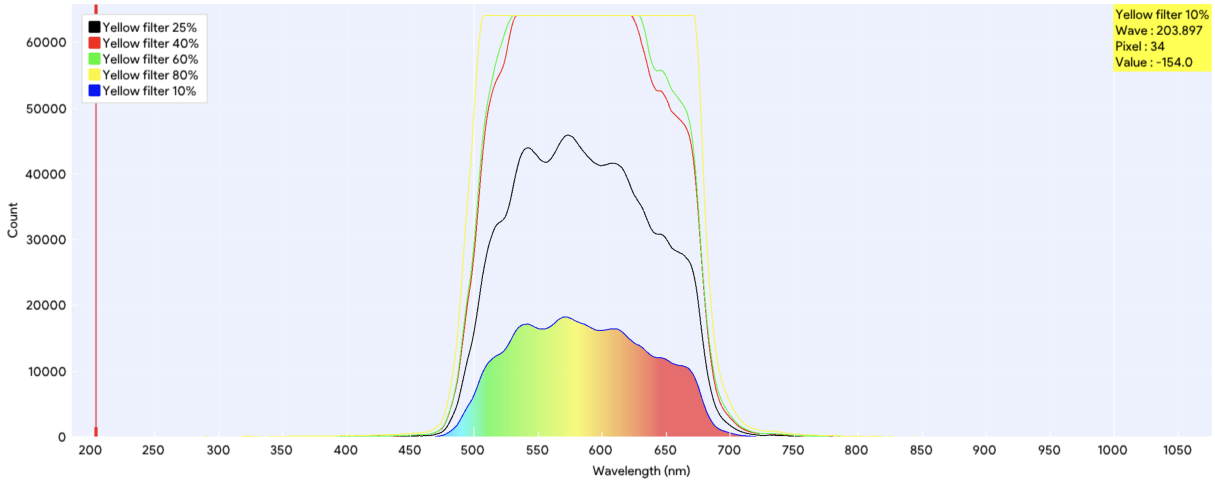


Figure 80: Spectral power distribution of light transmitted through yellow filter at light source intensities of 10%, 25%, 40%, 60%, and 80%

III.V XYZ color space visualization plots

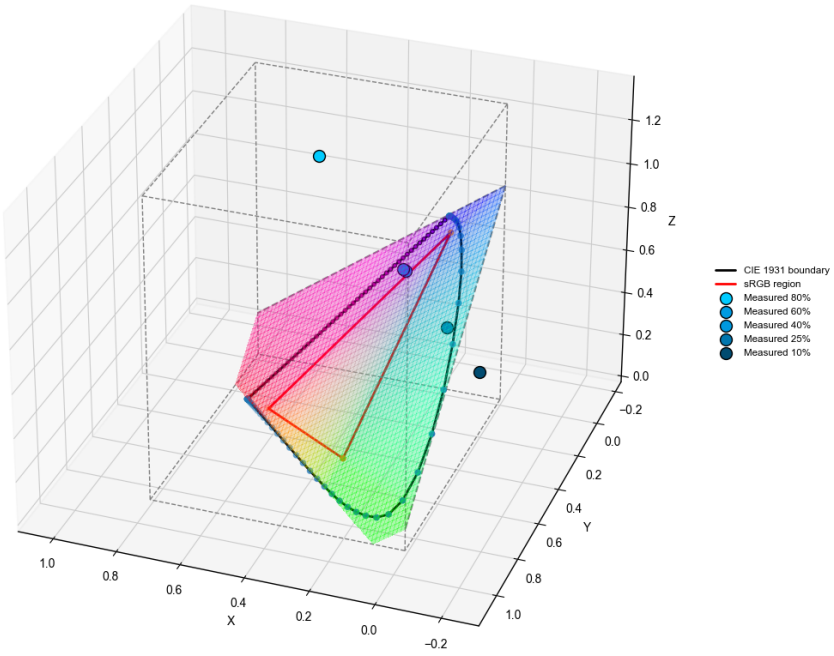


Figure 81: 3D XYZ color space visualization of light transmitted through the blue filter, at light source intensities of 10%, 25%, 40%, 60%, and 80% of the maximum

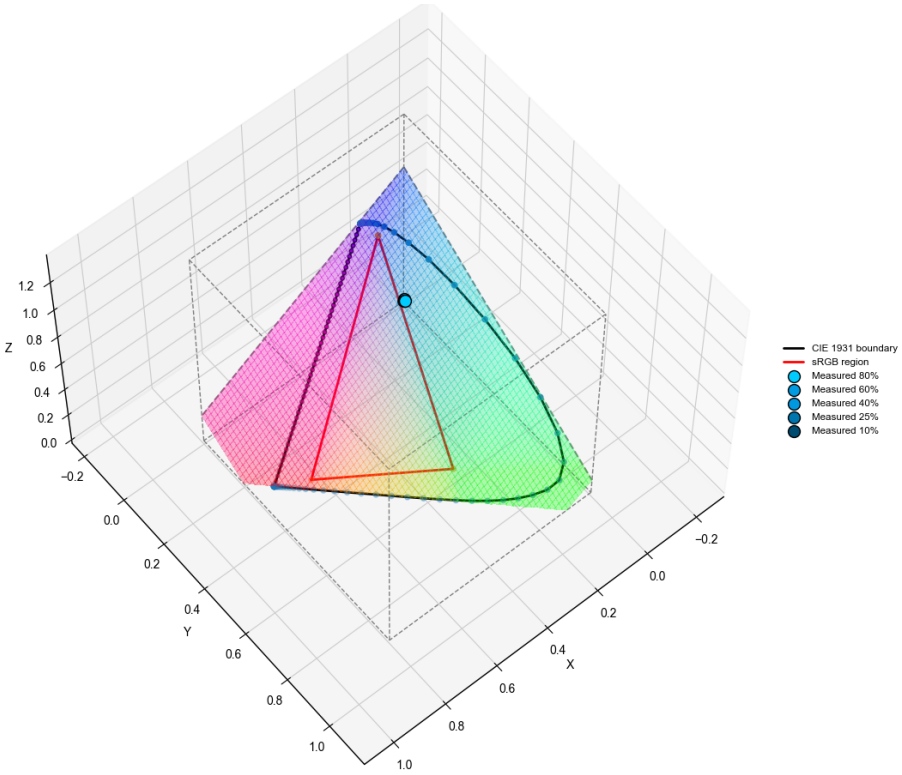


Figure 82: 3D XYZ color space visualization of light transmitted through the blue filter aligned in one XYZ coordinate, for all light source intensities (10%, 25%, 40%, 60%, and 80% of the maximum)

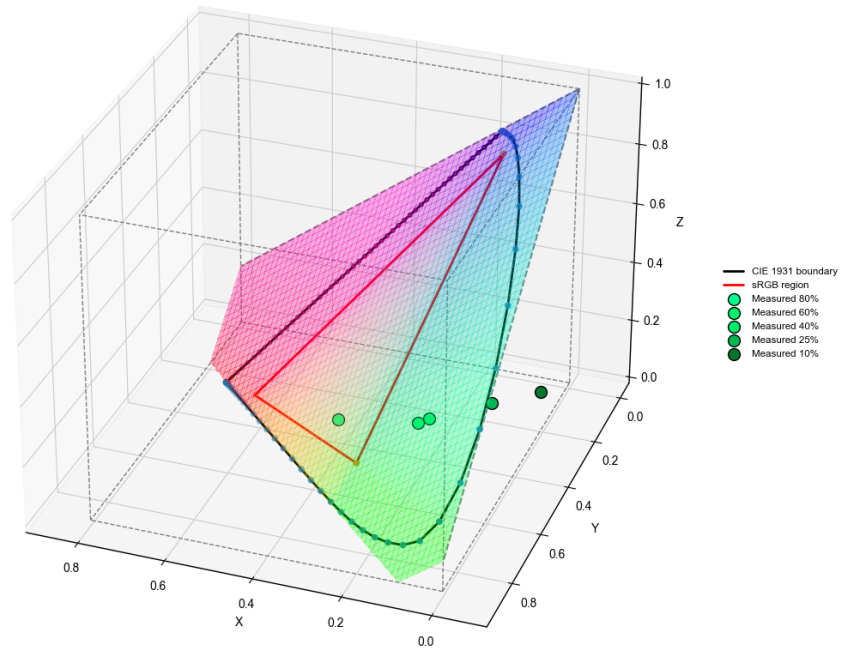


Figure 83: 3D XYZ color space visualization of light transmitted through the green filter, at light source intensities of 10%, 25%, 40%, 60%, and 80% of the maximum

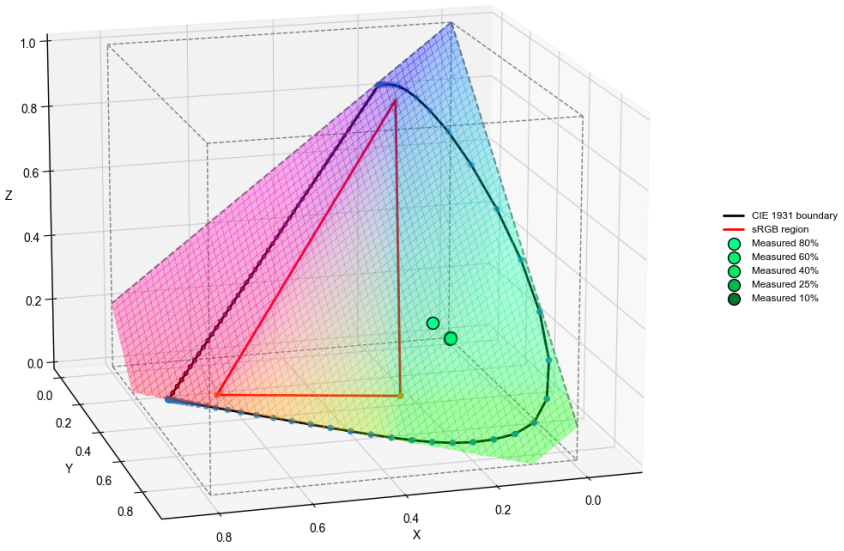


Figure 84: 3D XYZ color space visualization of light transmitted through the green filter aligned in one XYZ coordinate, for all light source intensities (10%, 25%, 40%, 60%, and 80% of the maximum)

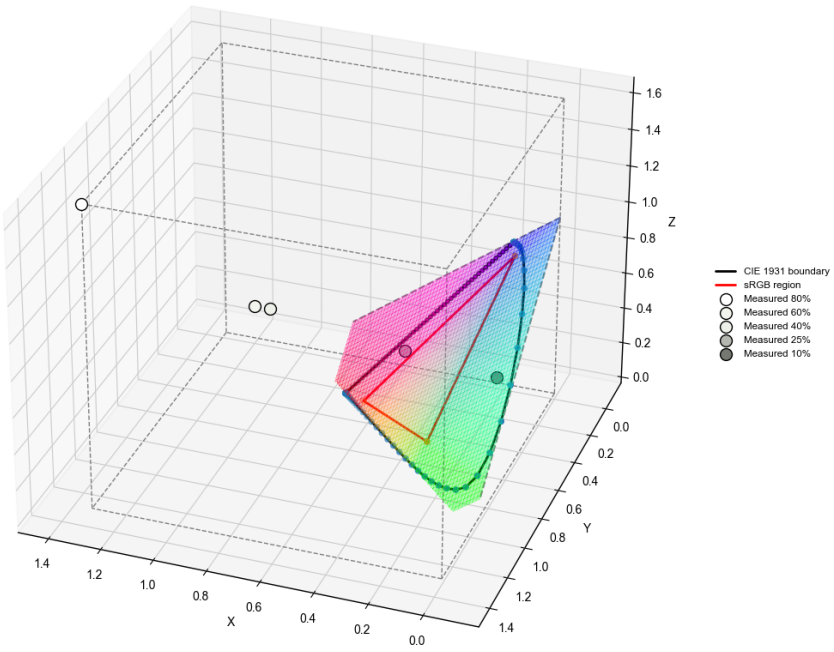


Figure 85: 3D XYZ color space visualization of light transmitted through the grey filter, at light source intensities of 10%, 25%, 40%, 60%, and 80% of the maximum

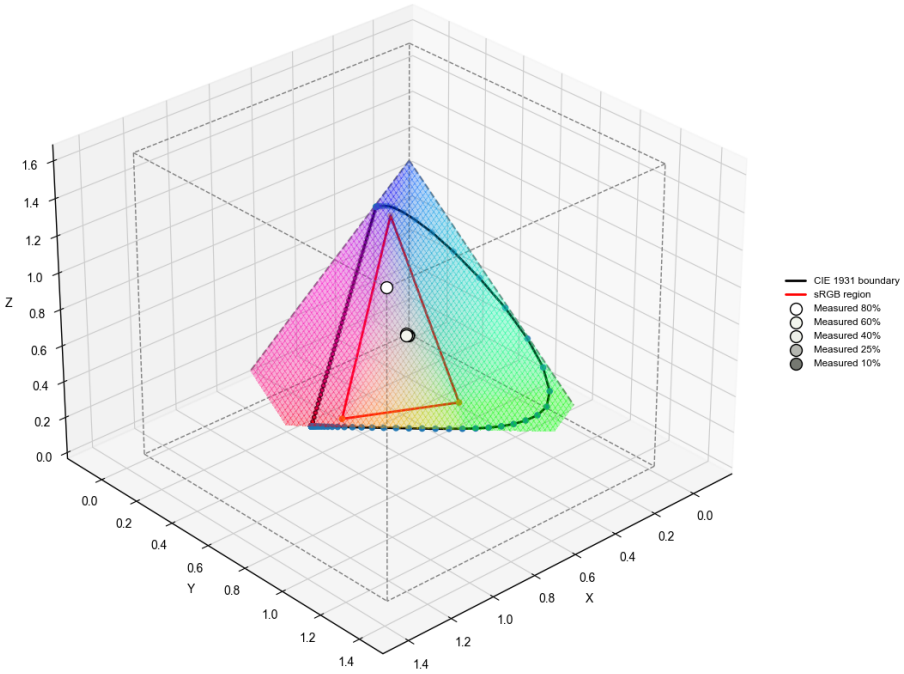


Figure 86: 3D XYZ color space visualization of light transmitted through the grey filter aligned in one XYZ coordinate, for all light source intensities (10%, 25%, 40%, 60%, and 80% of the maximum)

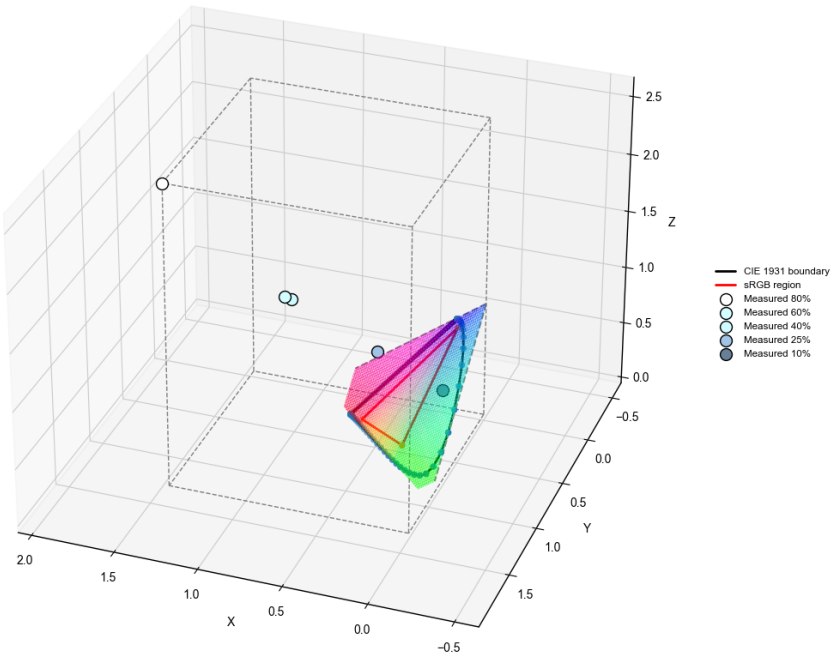


Figure 87: 3D XYZ color space visualization of light transmitted through the light blue filter, at light source intensities of 10%, 25%, 40%, 60%, and 80% of the maximum

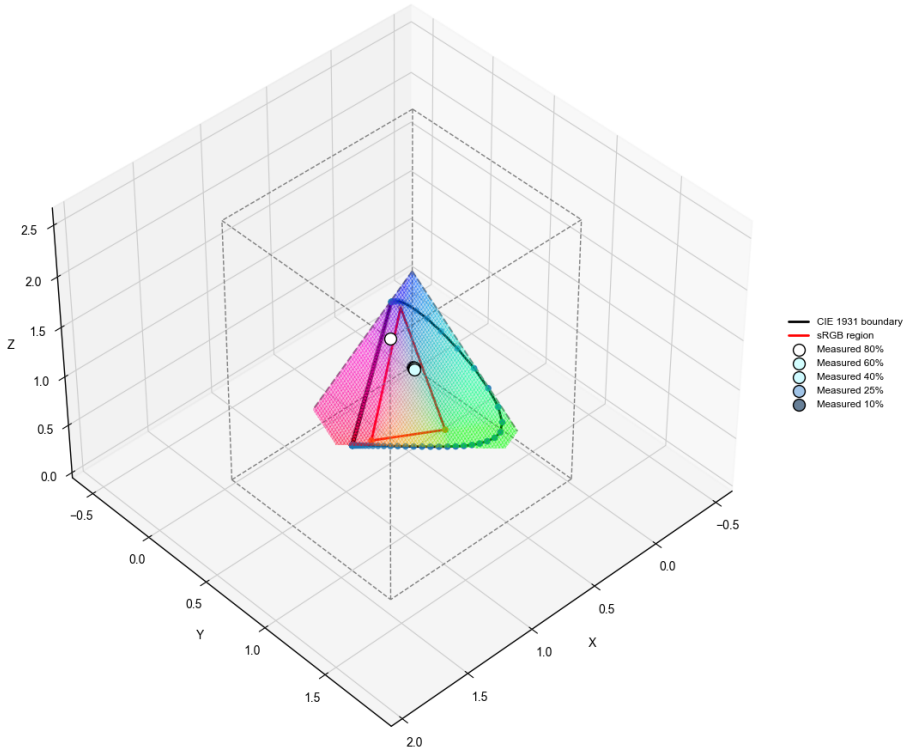


Figure 88: 3D XYZ color space visualization of light transmitted through the light blue filter aligned in one XYZ coordinate, for all light source intensities (10%, 25%, 40%, 60%, and 80% of the maximum)

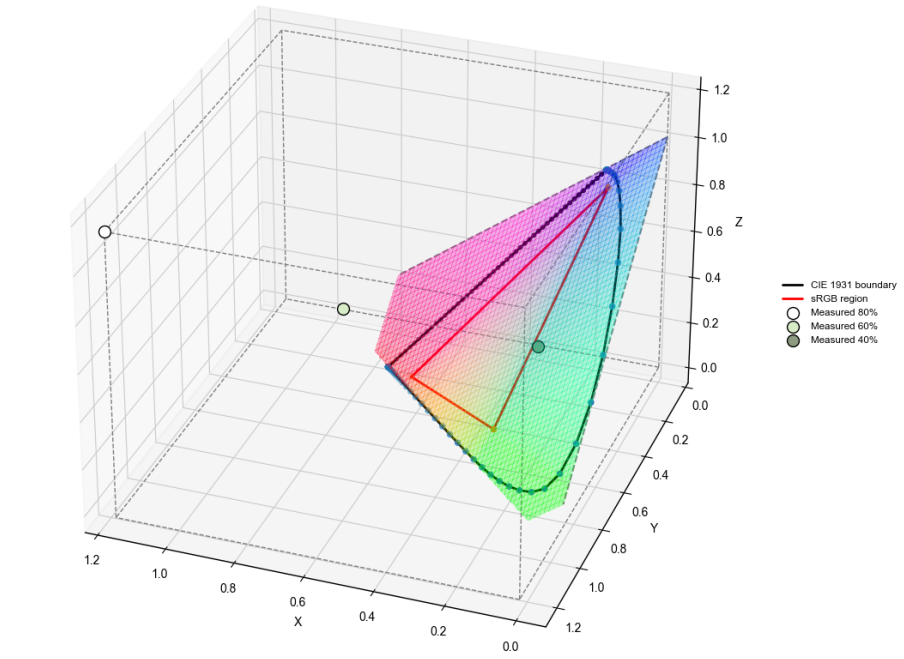


Figure 89: 3D XYZ color space visualization of light transmitted through the light green filter, at light source intensities of 10%, 25%, 40%, 60%, and 80% of the maximum

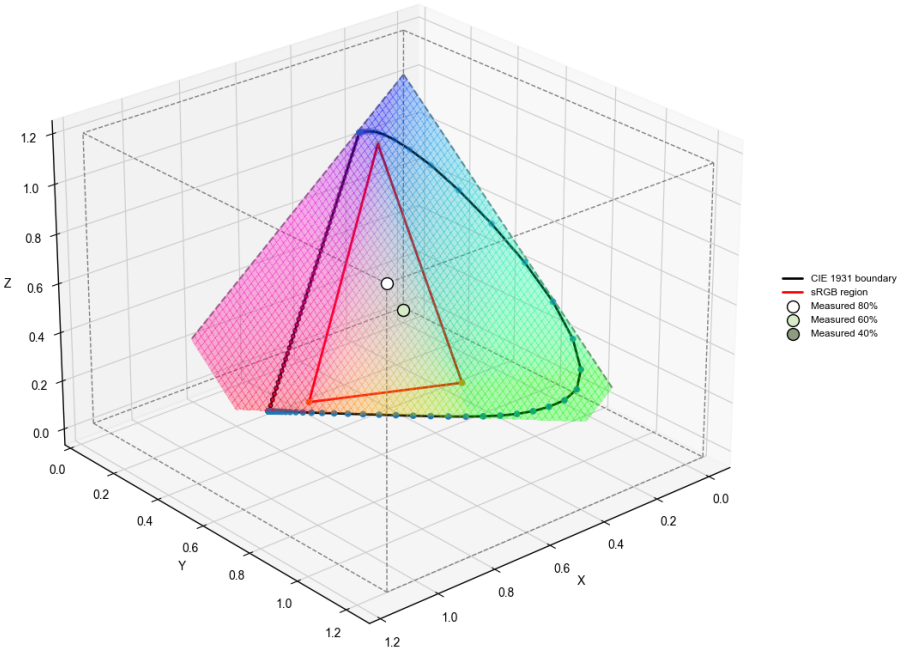


Figure 90: 3D XYZ color space visualization of light transmitted through the light green filter aligned in one XYZ coordinate, for all light source intensities (10%, 25%, 40%, 60%, and 80% of the maximum)

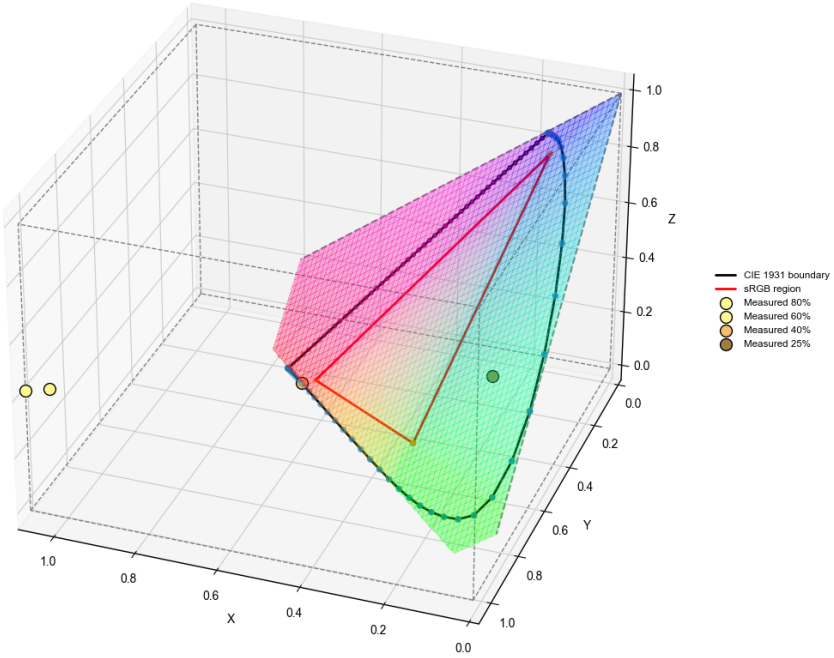


Figure 91: 3D XYZ color space visualization of light transmitted through the (light) orange filter, at light source intensities of 10%, 25%, 40%, 60%, and 80% of the maximum

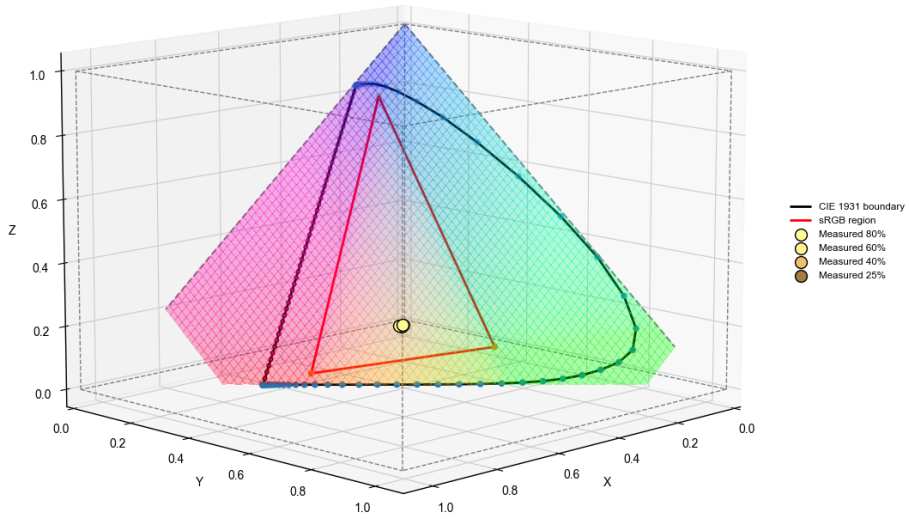


Figure 92: 3D XYZ color space visualization of light transmitted through the (light) orange filter aligned in one XYZ coordinate, for all light source intensities (10%, 25%, 40%, 60%, and 80% of the maximum)

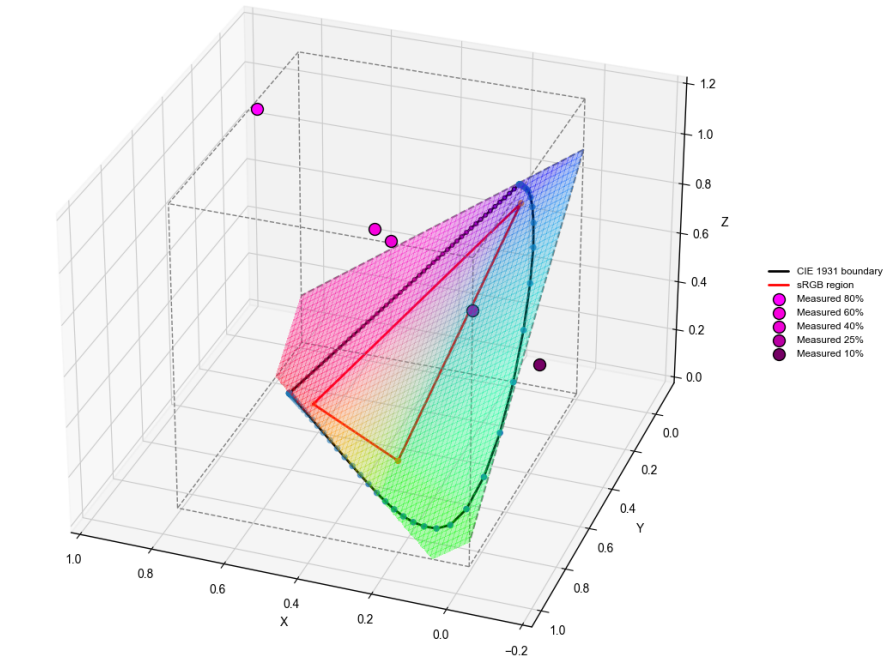


Figure 93: 3D XYZ color space visualization of light transmitted through the purple filter, at light source intensities of 10%, 25%, 40%, 60%, and 80% of the maximum

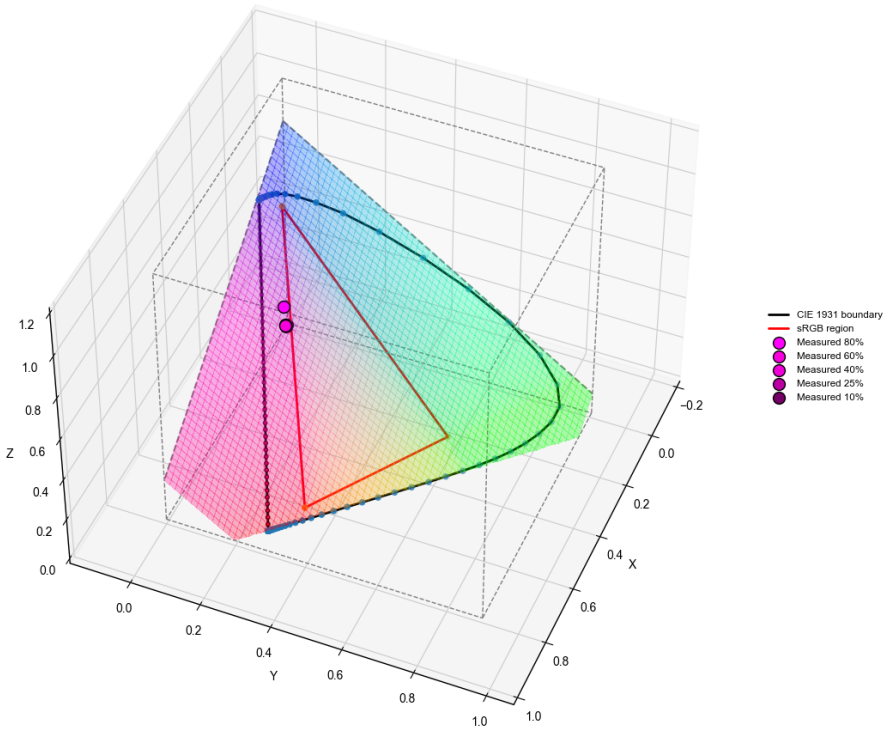


Figure 94: 3D XYZ color space visualization of light transmitted through the purple filter aligned in one XYZ coordinate, for all light source intensities (10%, 25%, 40%, 60%, and 80% of the maximum)

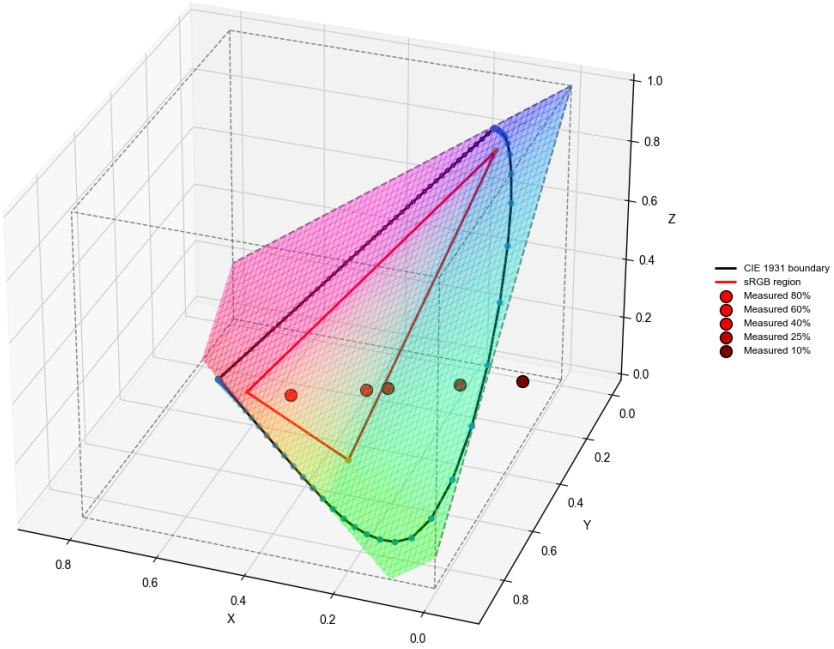


Figure 95: 3D XYZ color space visualization of light transmitted through the red filter, at light source intensities of 10%, 25%, 40%, 60%, and 80% of the maximum

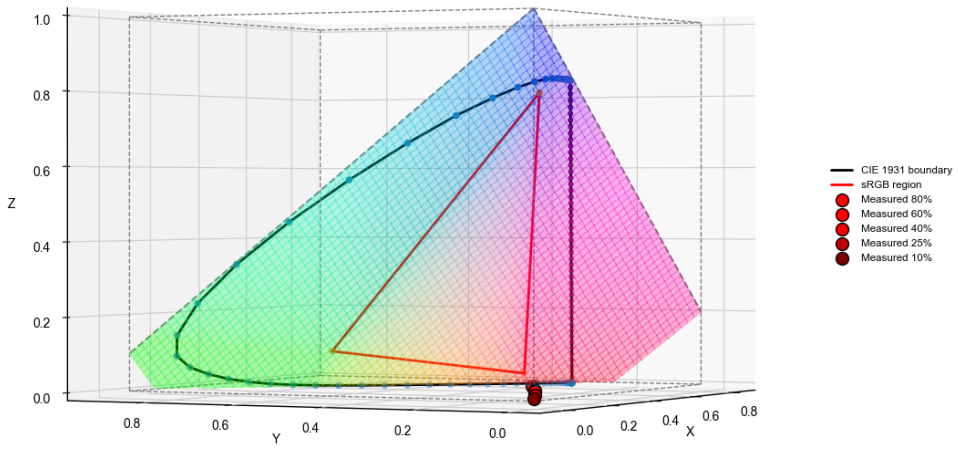


Figure 96: 3D XYZ color space visualization of light transmitted through the red filter aligned in one XYZ coordinate, for all light source intensities (10%, 25%, 40%, 60%, and 80% of the maximum)

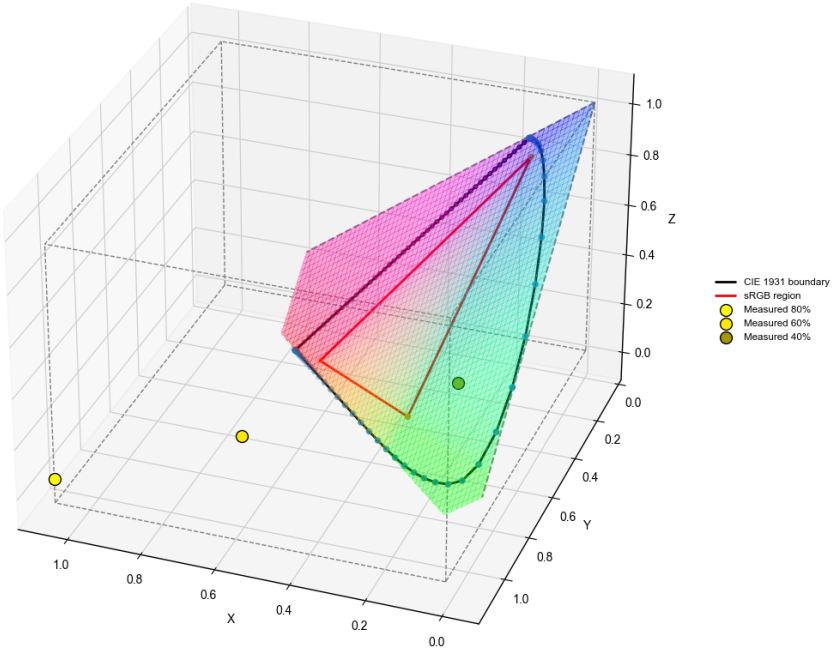


Figure 97: 3D XYZ color space visualization of light transmitted through the yellow filter, at light source intensities of 10%, 25%, 40%, 60%, and 80% of the maximum

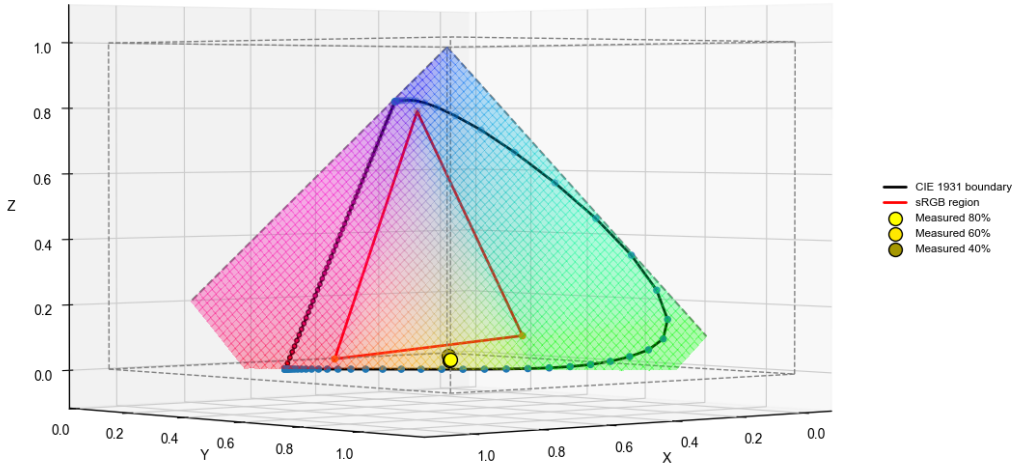


Figure 98: 3D XYZ color space visualization of light transmitted through the yellow filter aligned in one XYZ coordinate, for all light source intensities (10%, 25%, 40%, 60%, and 80% of the maximum)

III.VI CIE 1931 chromaticity diagrams

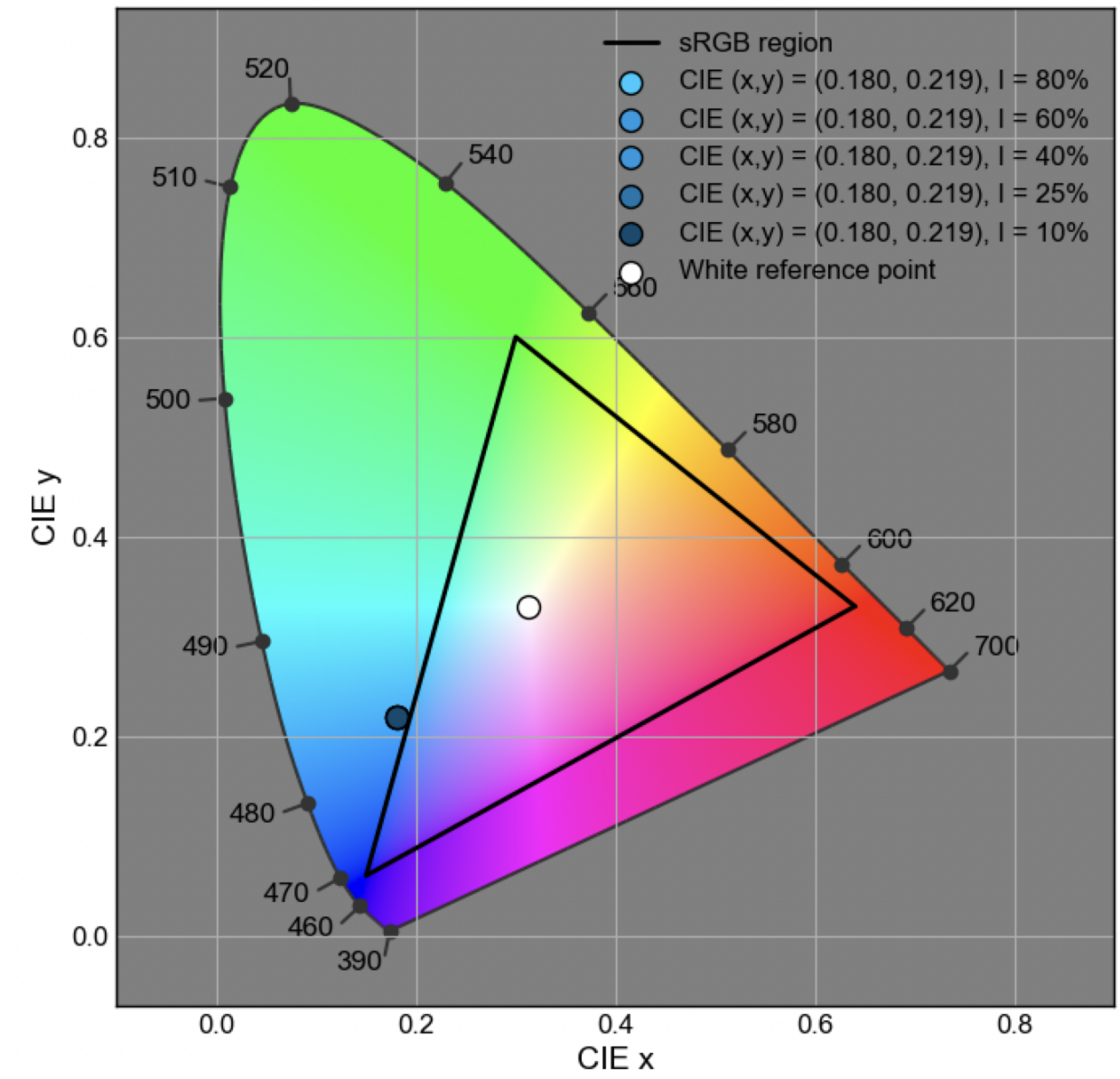


Figure 99: CIE 1931 chromaticity diagram of light transmitted through the blue filter. On the right, the color gradient corresponds to light source intensities of 10%, 25%, 40%, 60%, and 80% (bottom to top), with the resulting RGB values from the spectrometer shown in brackets

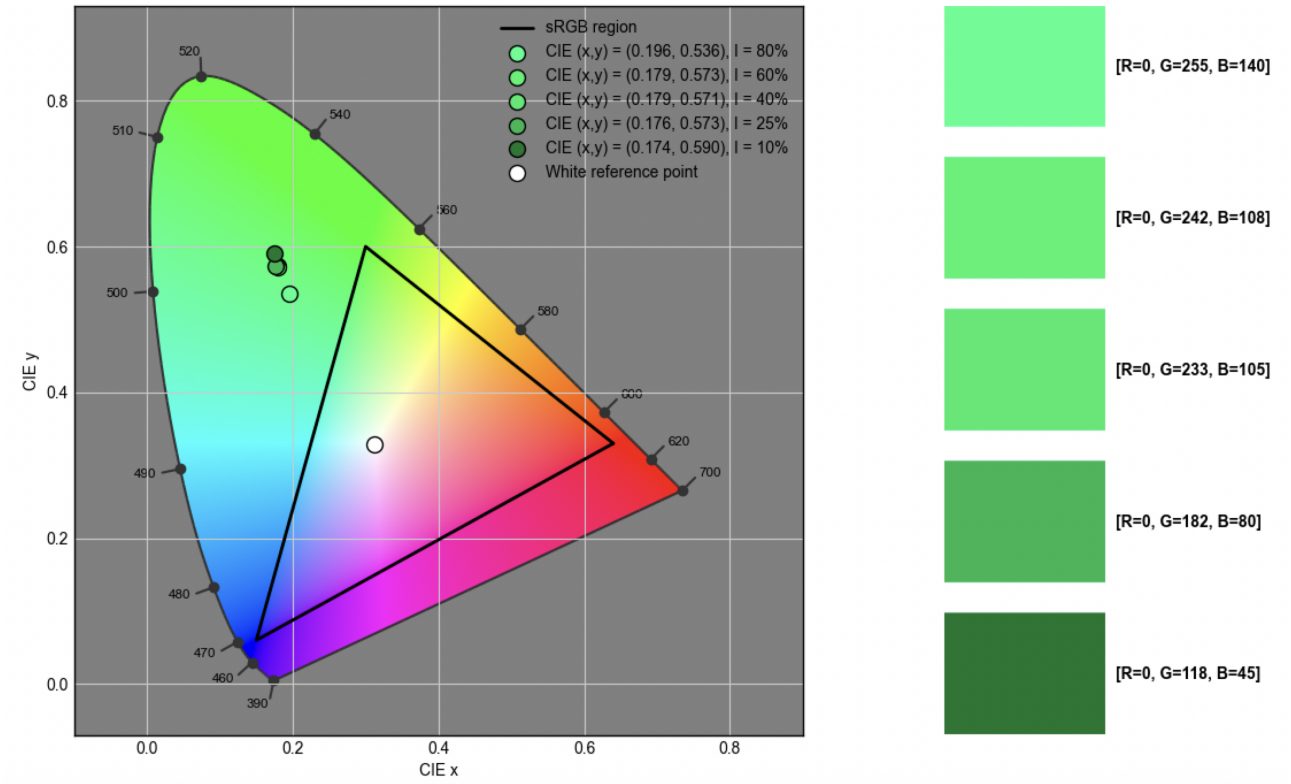


Figure 100: CIE 1931 chromaticity diagram of light transmitted through the green filter. On the right, the color gradient corresponds to light source intensities of 10%, 25%, 40%, 60%, and 80% (bottom to top), with the resulting RGB values from the spectrometer shown in brackets

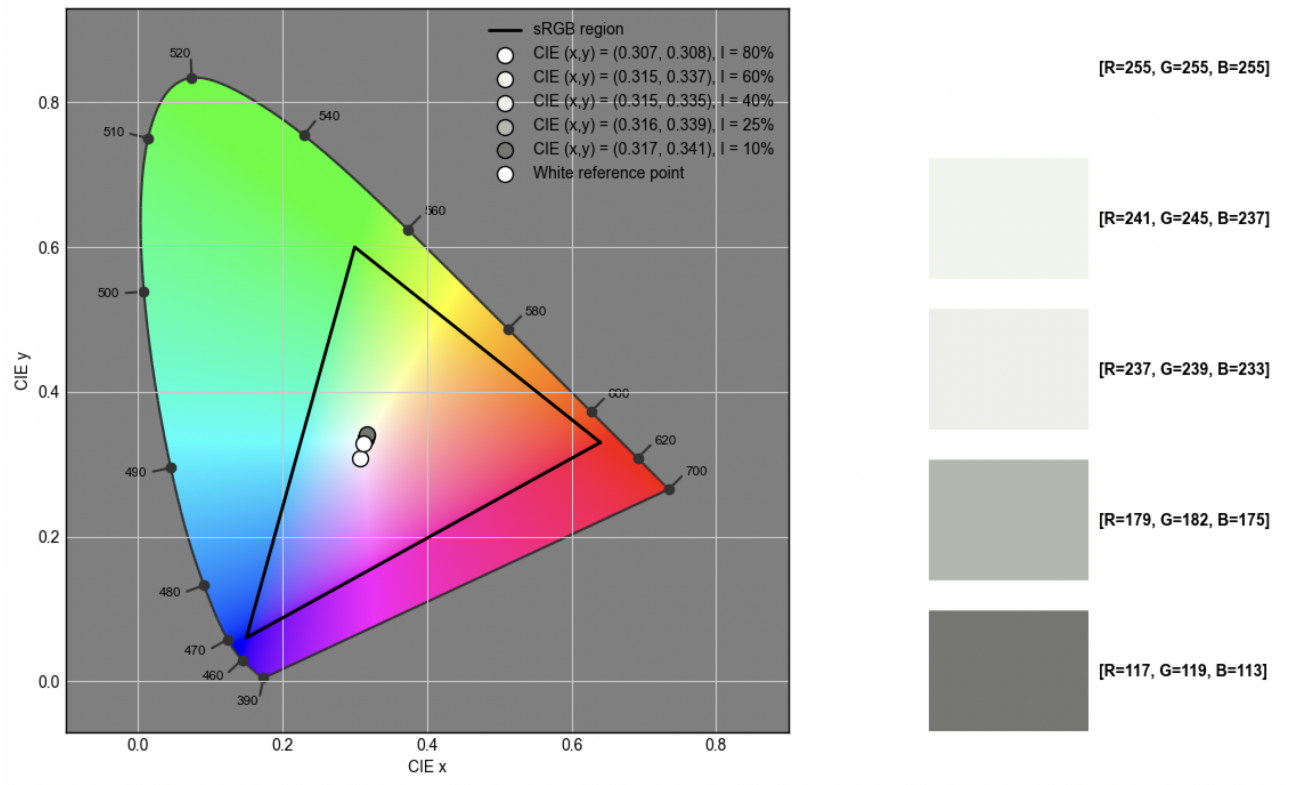


Figure 101: CIE 1931 chromaticity diagram of light transmitted through the grey filter. On the right, the color gradient corresponds to light source intensities of 10%, 25%, 40%, 60%, and 80% (bottom to top), with the resulting RGB values from the spectrometer shown in brackets

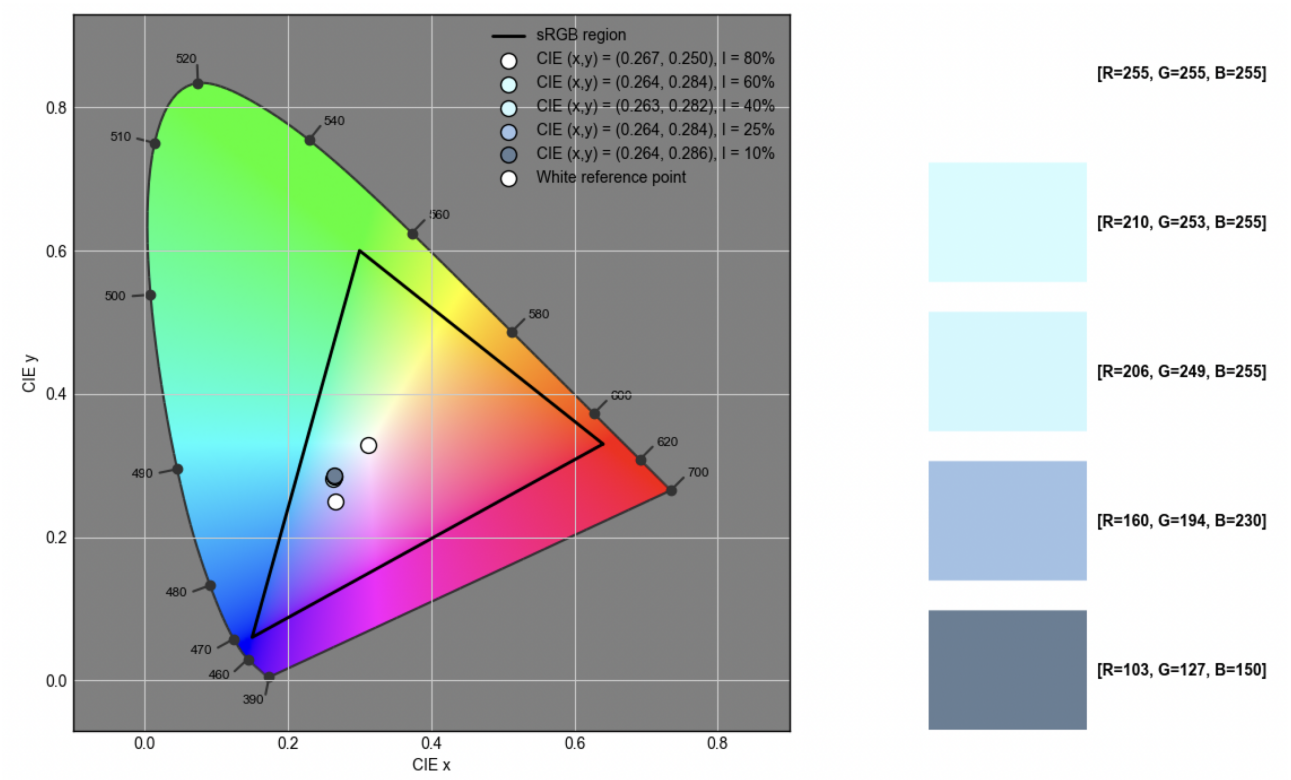


Figure 102: CIE 1931 chromaticity diagram of light transmitted through the light blue filter. On the right, the color gradient corresponds to light source intensities of 10%, 25%, 40%, 60%, and 80% (bottom to top), with the resulting RGB values from the spectrometer shown in brackets

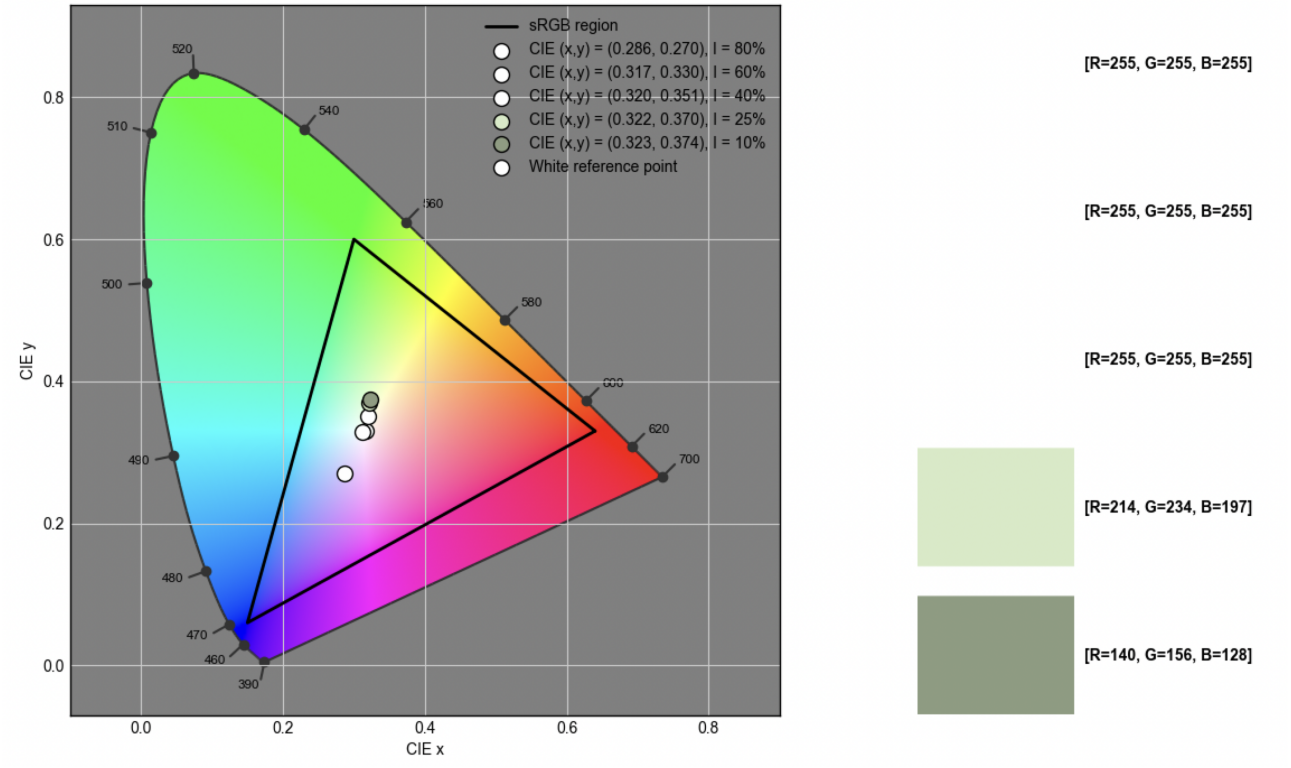


Figure 103: CIE 1931 chromaticity diagram of light transmitted through the light green filter. On the right, the color gradient corresponds to light source intensities of 10%, 25%, 40%, 60%, and 80% (bottom to top), with the resulting RGB values from the spectrometer shown in brackets

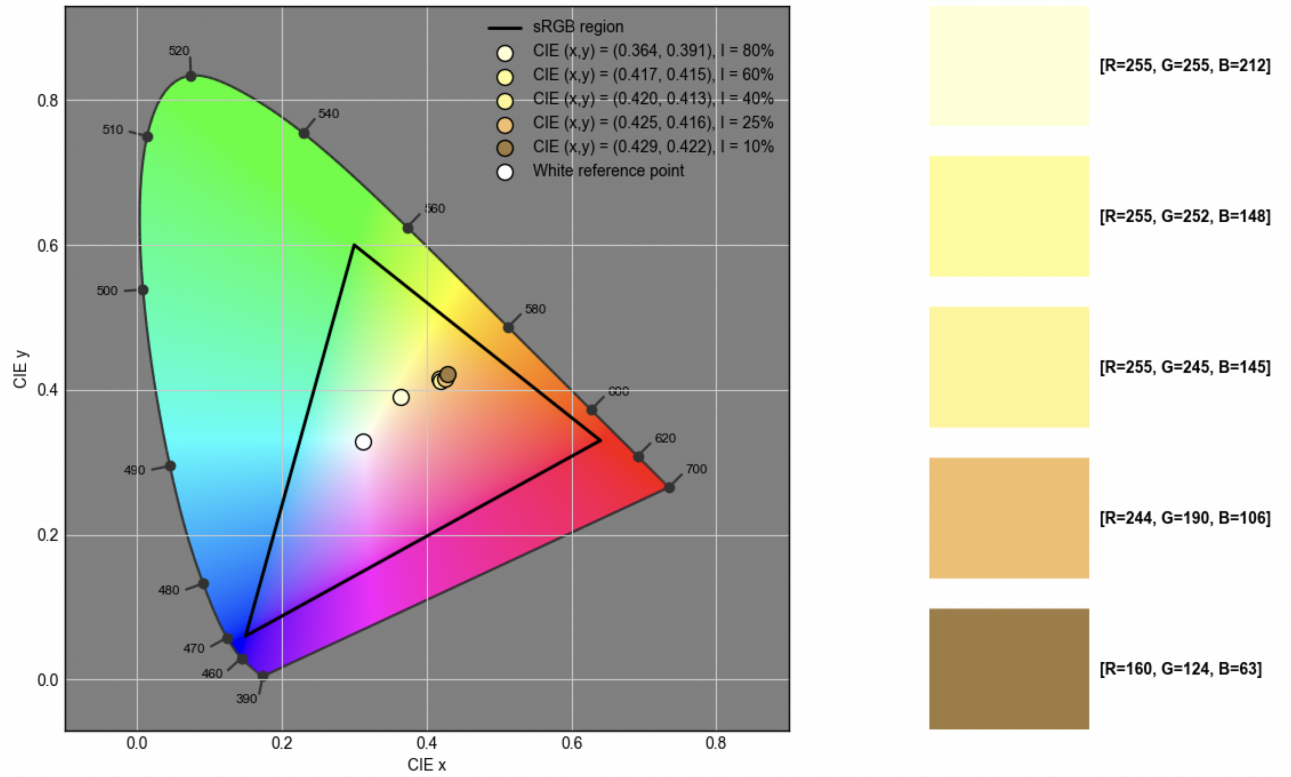


Figure 104: CIE 1931 chromaticity diagram of light transmitted through the (light) orange filter. On the right, the color gradient corresponds to light source intensities of 10%, 25%, 40%, 60%, and 80% (bottom to top), with the resulting RGB values from the spectrometer shown in brackets

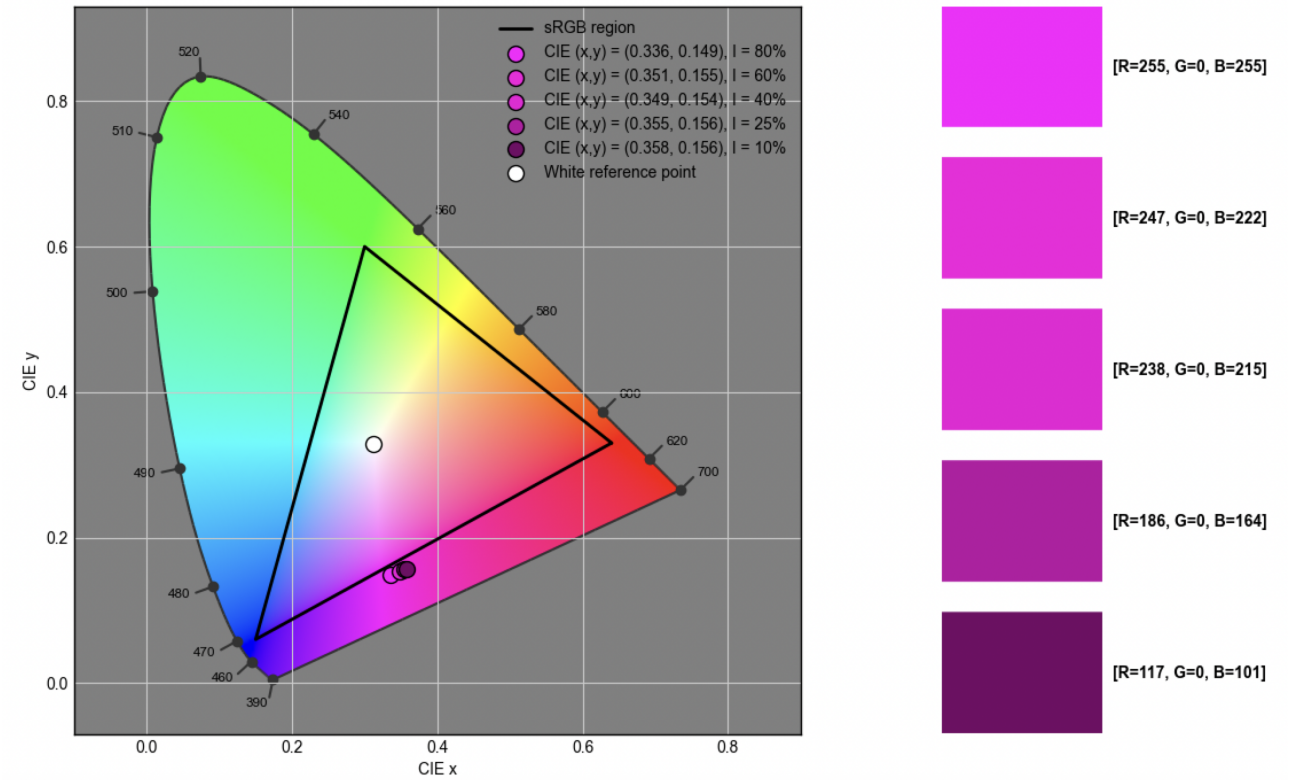


Figure 105: CIE 1931 chromaticity diagram of light transmitted through the purple filter. On the right, the color gradient corresponds to light source intensities of 10%, 25%, 40%, 60%, and 80% (bottom to top), with the resulting RGB values from the spectrometer shown in brackets

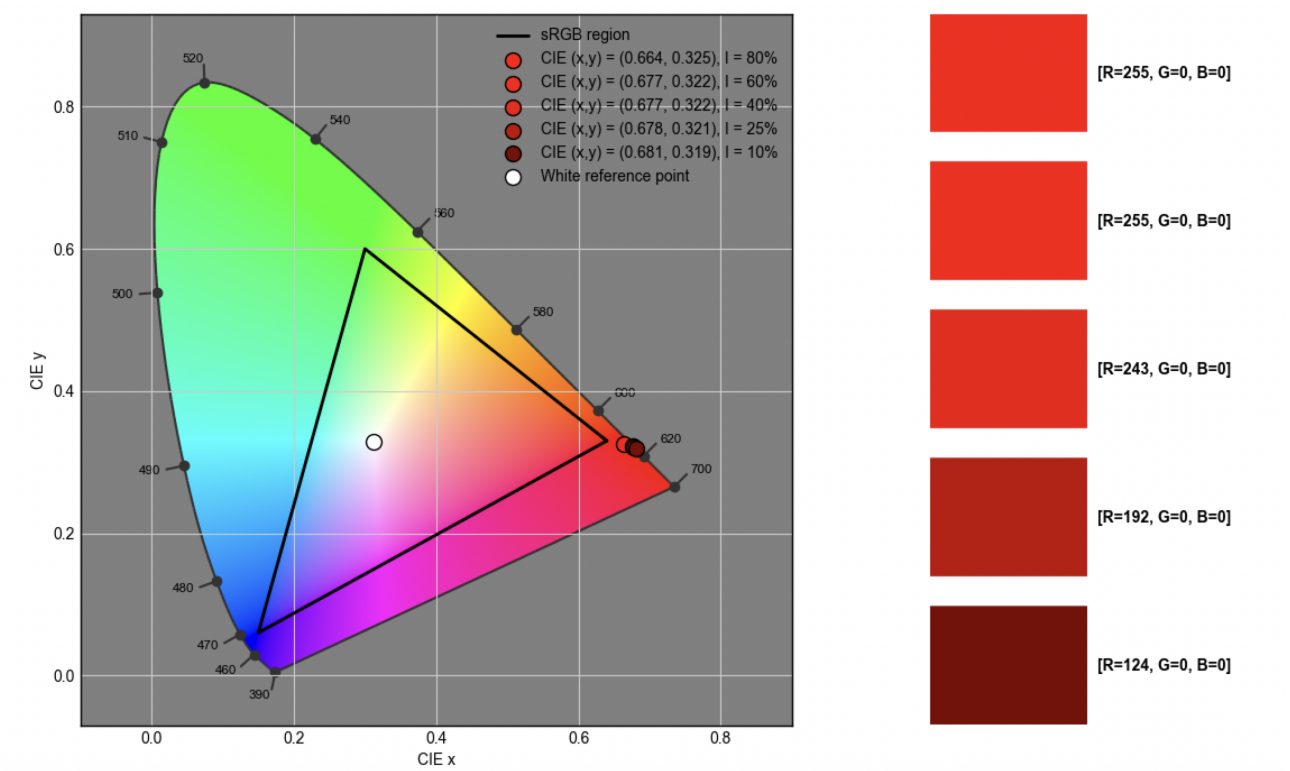


Figure 106: CIE 1931 chromaticity diagram of light transmitted through the red filter. On the right, the color gradient corresponds to light source intensities of 10%, 25%, 40%, 60%, and 80% (bottom to top), with the resulting RGB values from the spectrometer shown in brackets

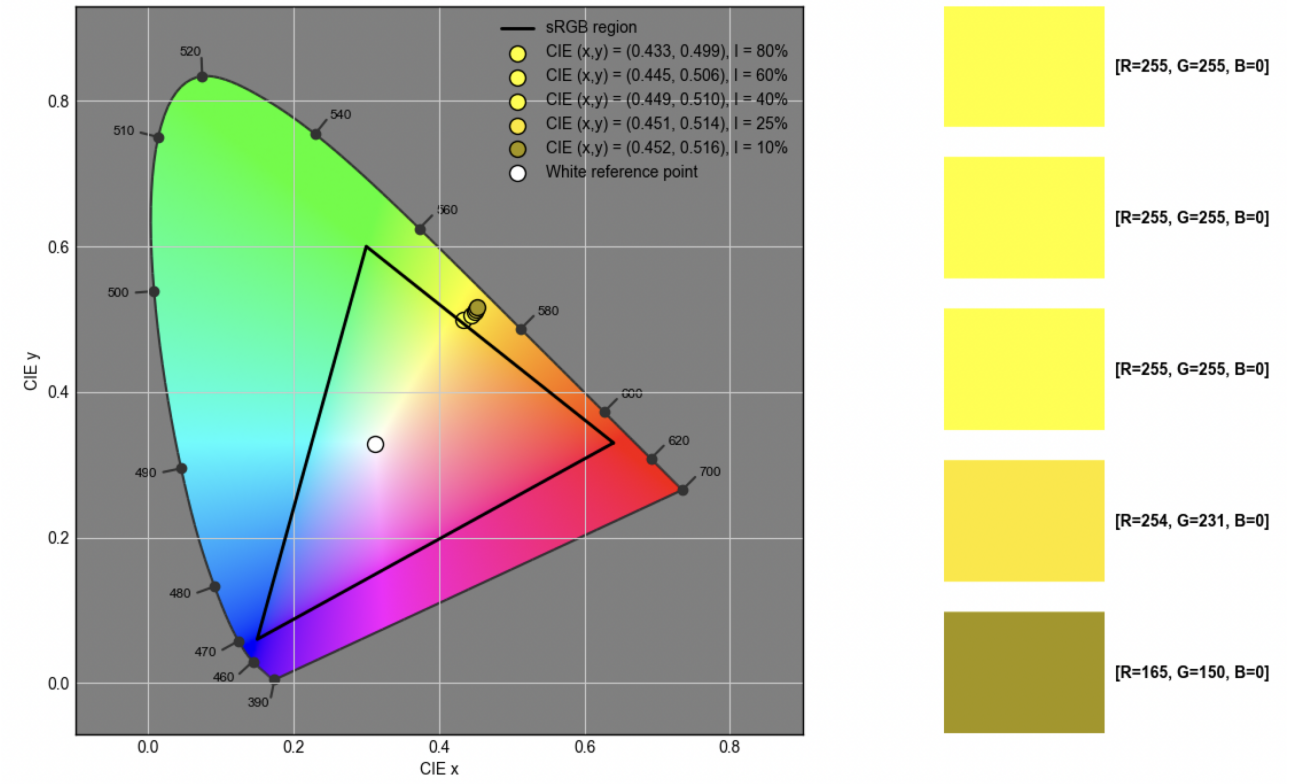


Figure 107: CIE 1931 chromaticity diagram of light transmitted through the yellow filter. On the right, the color gradient corresponds to light source intensities of 10%, 25%, 40%, 60%, and 80% (bottom to top), with the resulting RGB values from the spectrometer shown in brackets

III.VII RGB values obtained with PEMAD system per filter

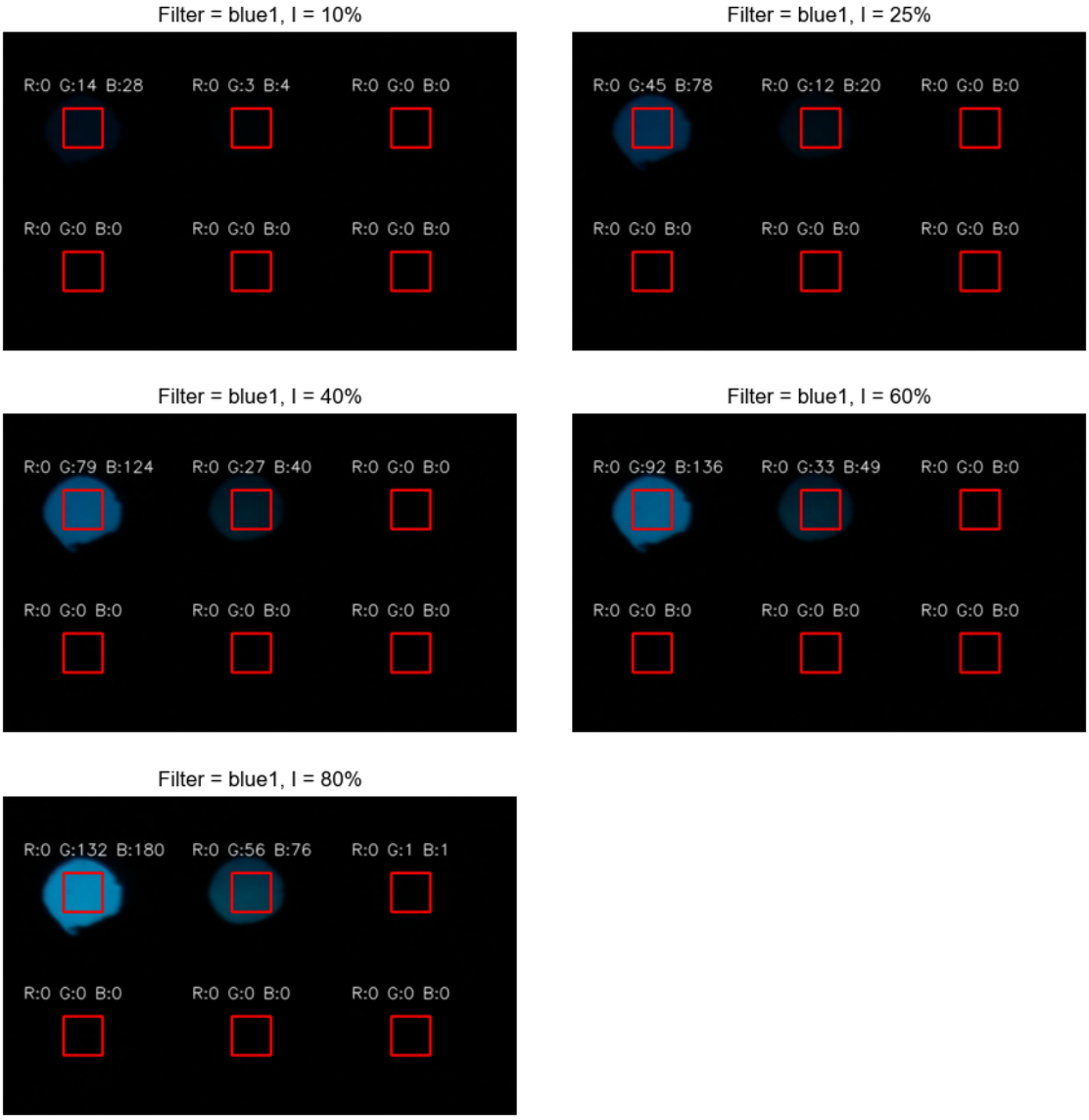


Figure 108: RGB values for light through the blue color filter at all intensity levels of the light source (10%, 25%, 40%, 60%, and 80%) captured with the PEMAD system

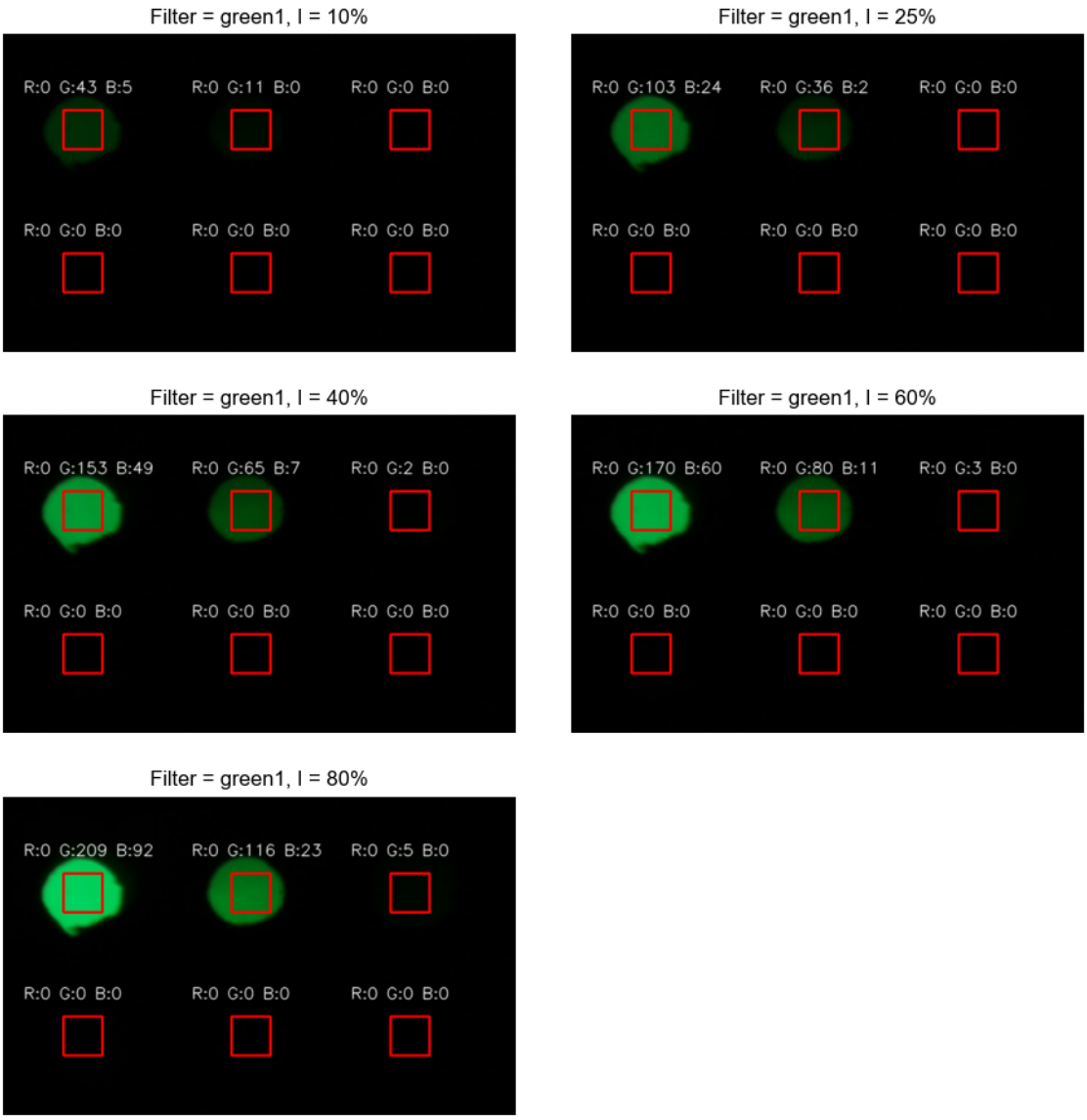


Figure 109: RGB values for light through the green color filter at all intensity levels of the light source (10%, 25%, 40%, 60%, and 80%) captured with the PEMAD system

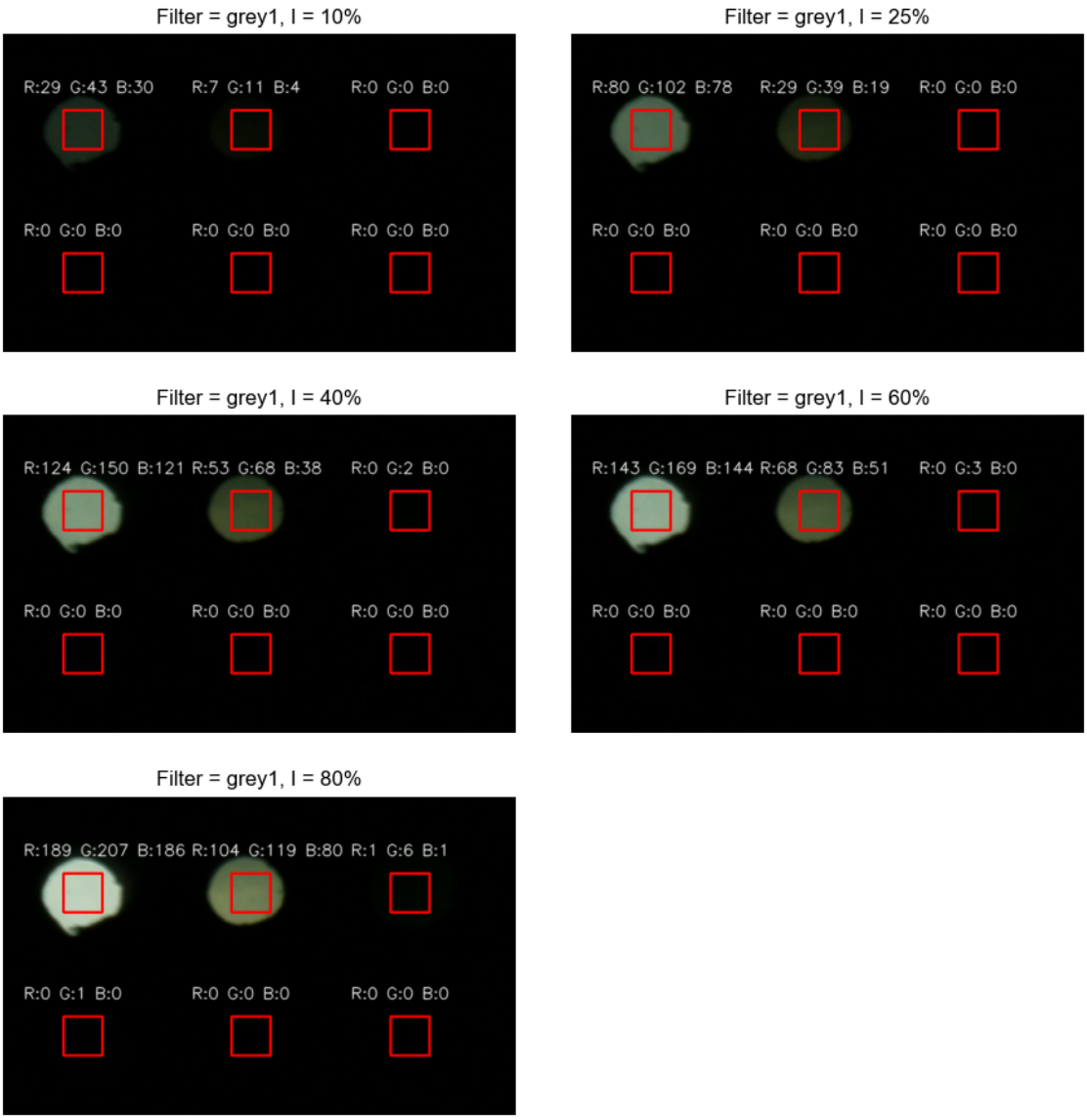


Figure 110: RGB values for light through the grey color filter at all intensity levels of the light source (10%, 25%, 40%, 60%, and 80%) captured with the PEMAD system

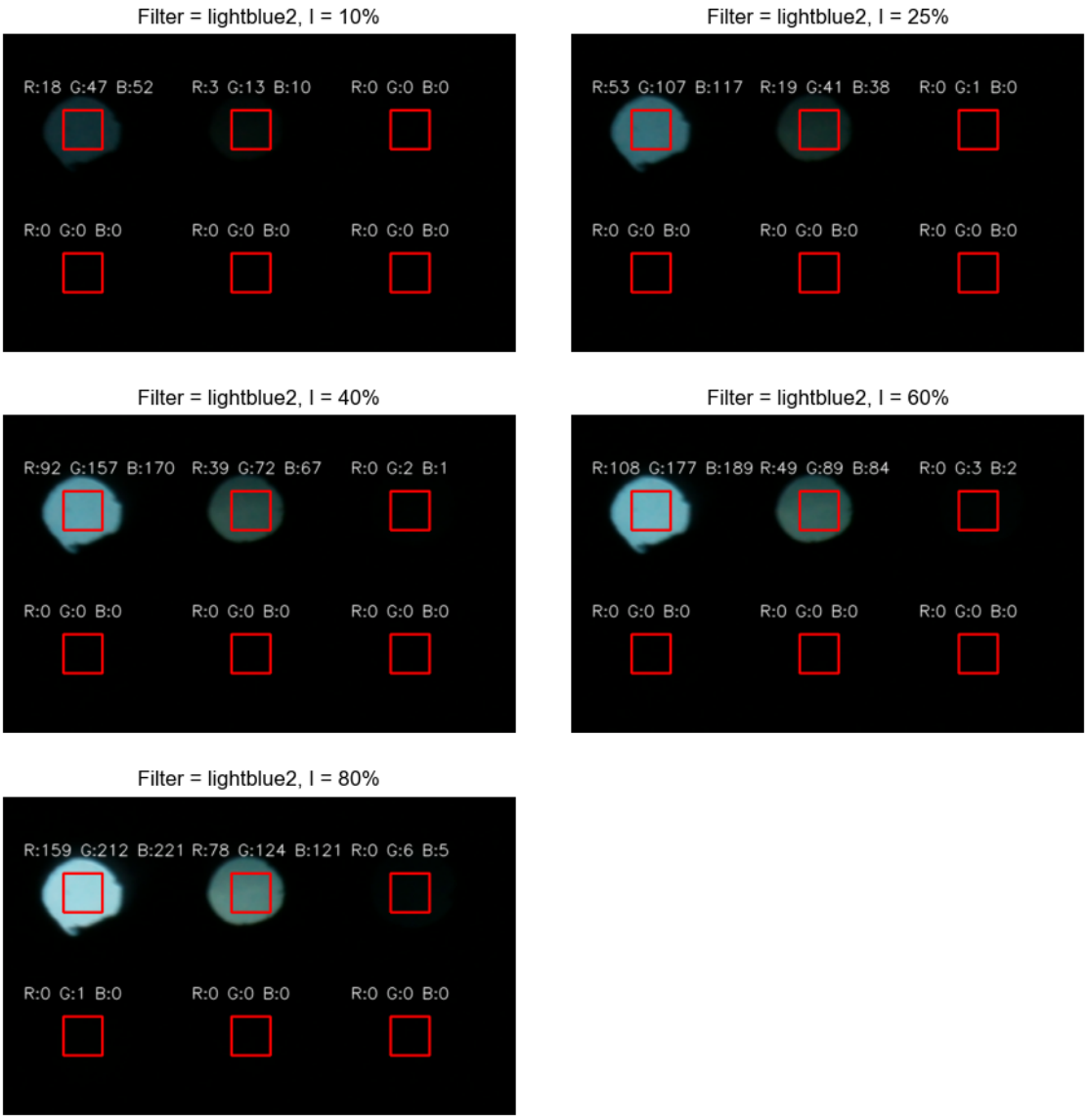


Figure 111: RGB values for light through the light blue color filter at all intensity levels of the light source (10%, 25%, 40%, 60%, and 80%) captured with the PEMAD system

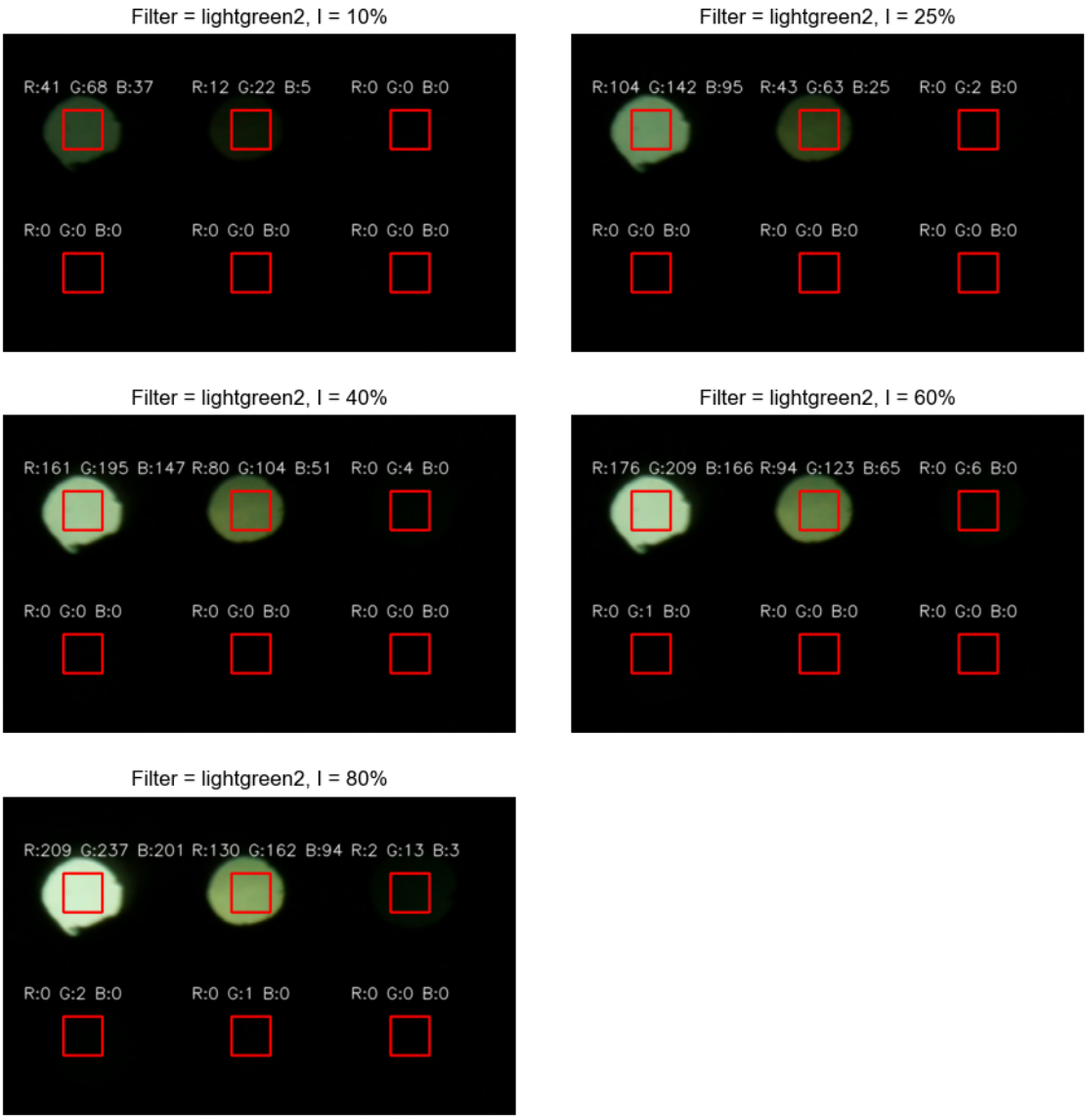


Figure 112: RGB values for light through the light green color filter at all intensity levels of the light source (10%, 25%, 40%, 60%, and 80%) captured with the PEMAD system

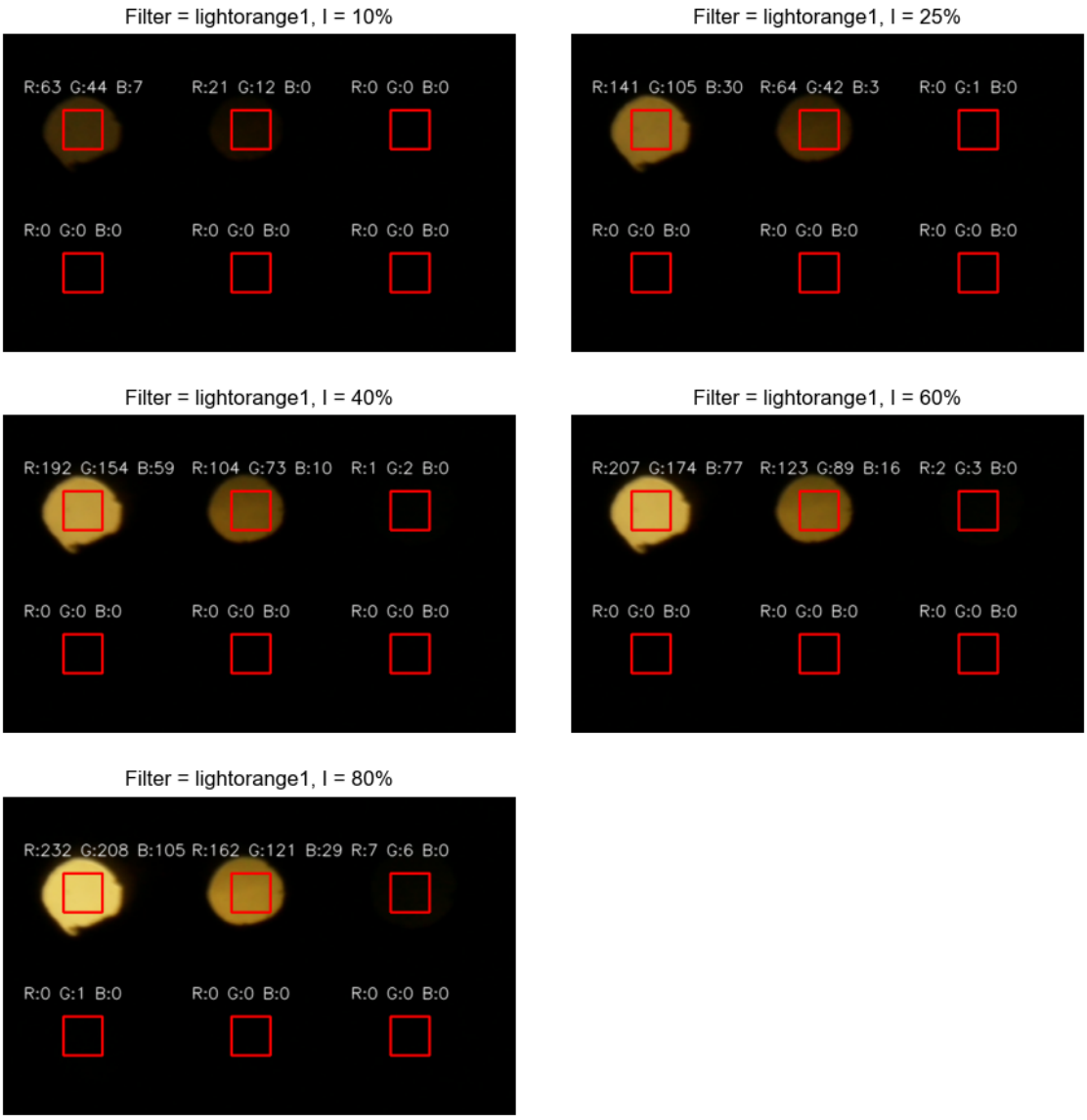


Figure 113: RGB values for light through the (light) orange color filter at all intensity levels of the light source (10%, 25%, 40%, 60%, and 80%) captured with the PEMAD system

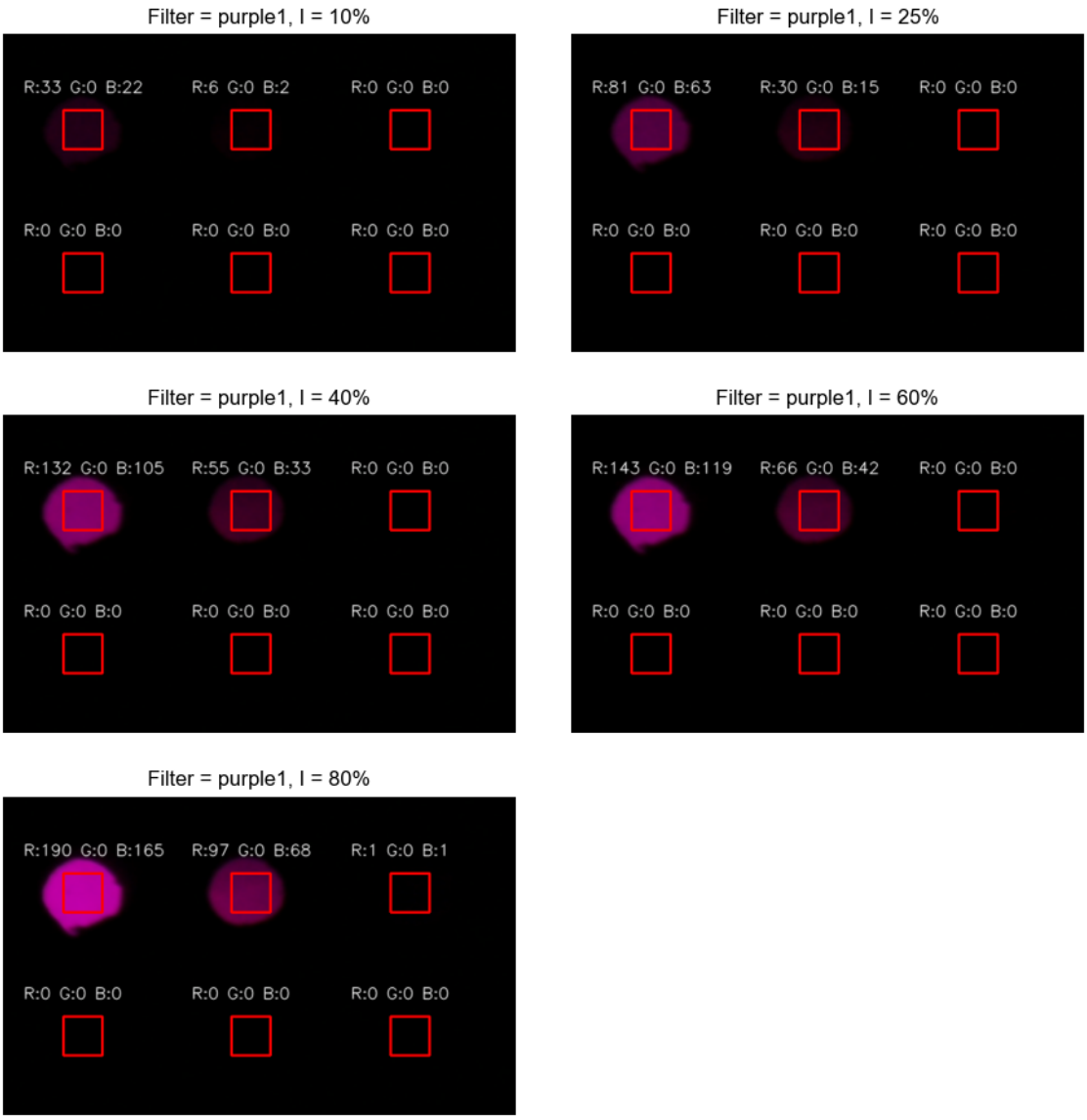


Figure 114: RGB values for light through the purple color filter at all intensity levels of the light source (10%, 25%, 40%, 60%, and 80%) captured with the PEMAD system

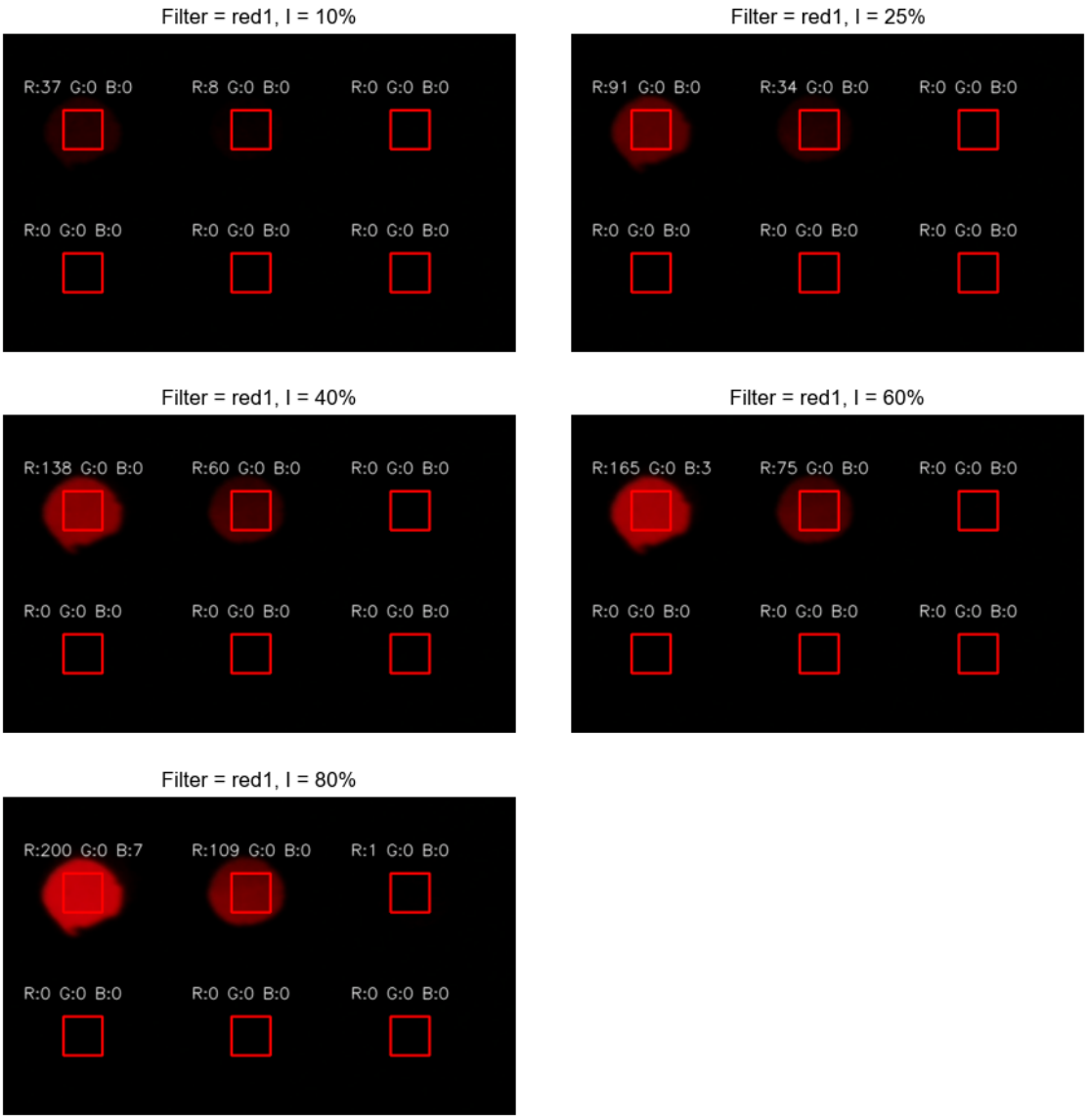


Figure 115: RGB values for light through the red color filter at all intensity levels of the light source (10%, 25%, 40%, 60%, and 80%) captured with the PEMAD system

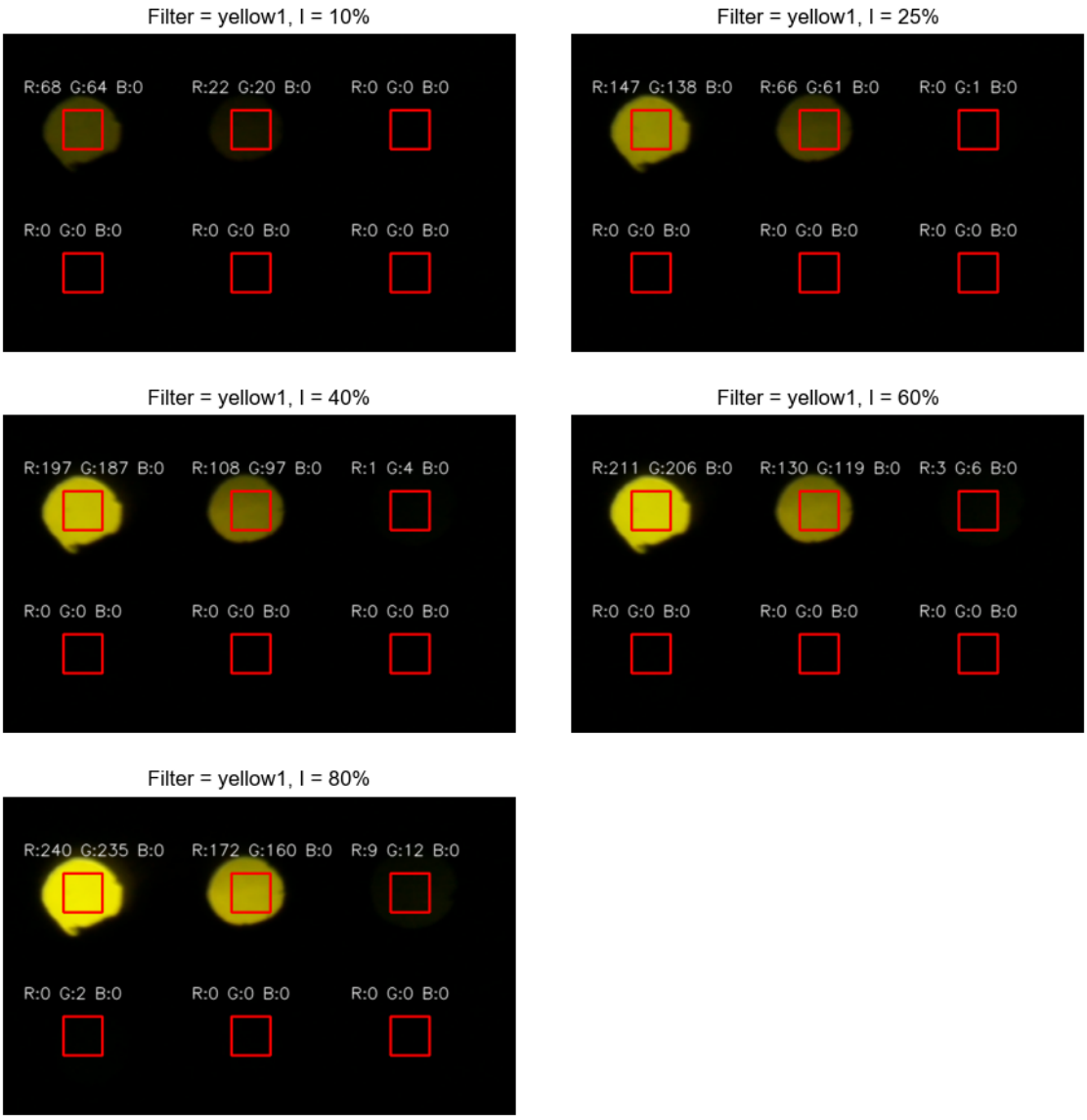


Figure 116: RGB values for light through the yellow color filter at all intensity levels of the light source (10%, 25%, 40%, 60%, and 80%) captured with the PEMAD system

III.VIII Regression plots transformation: camera RGB vs spectral reference RGB vs corrected RGB

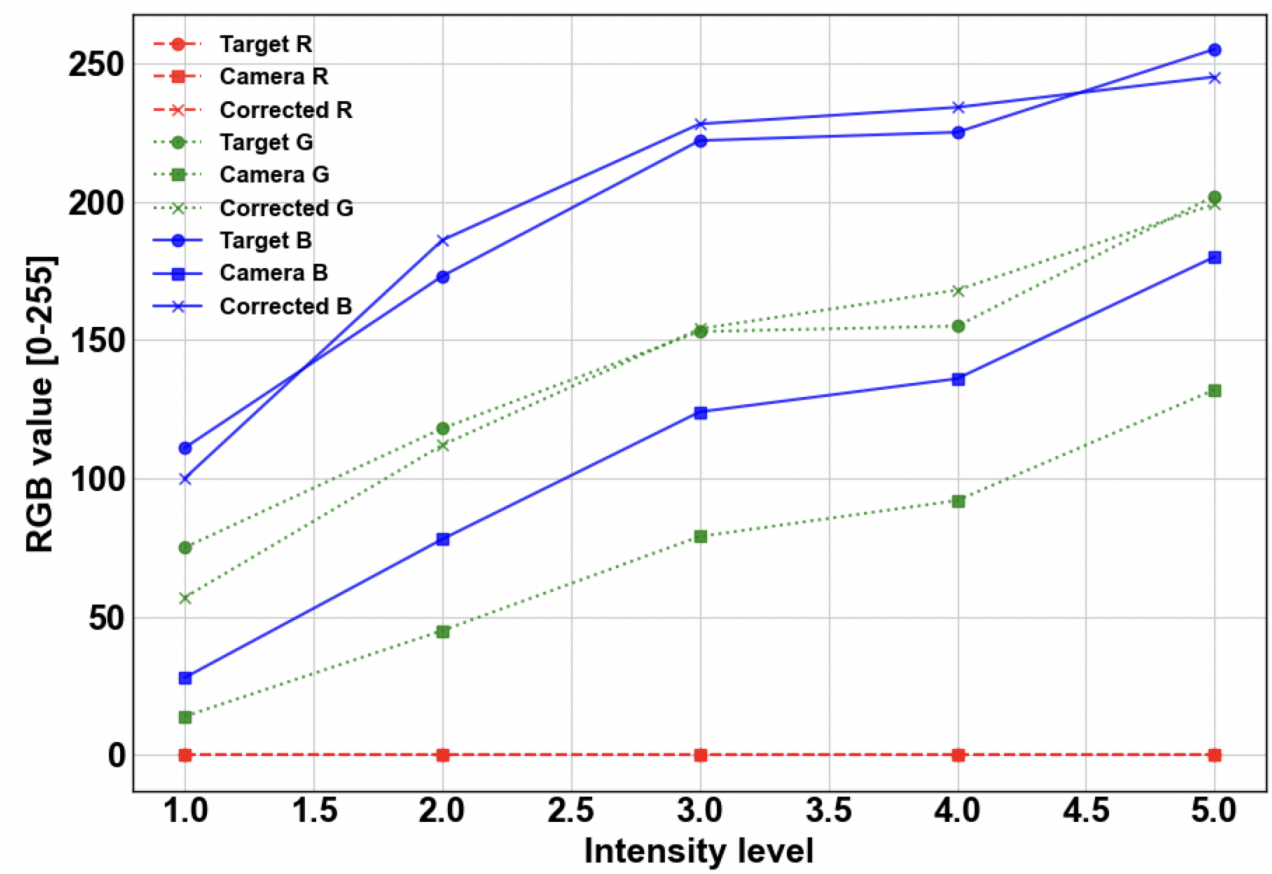


Figure 117: Calibration mapping for blue color filter: camera captured RGB vs spectral reference RGB vs corrected RGB for all color channels (R, G and B)

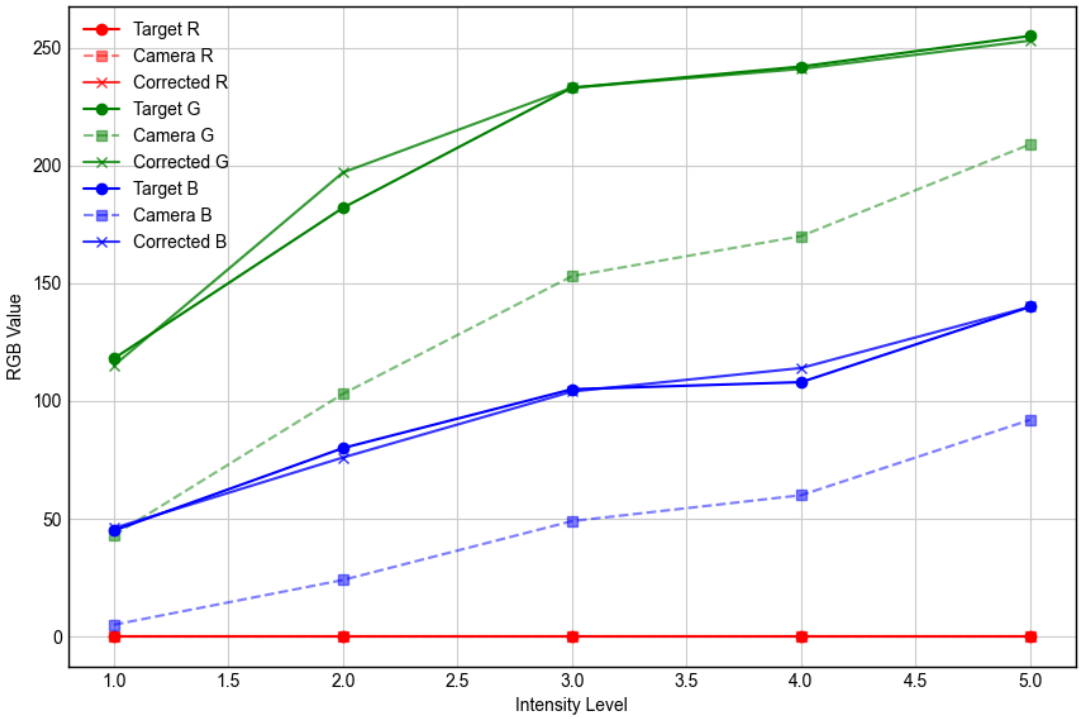


Figure 118: Calibration mapping for light through the green color filter: camera captured RGB vs spectral reference RGB vs corrected RGB for all color channels (R, G and B)

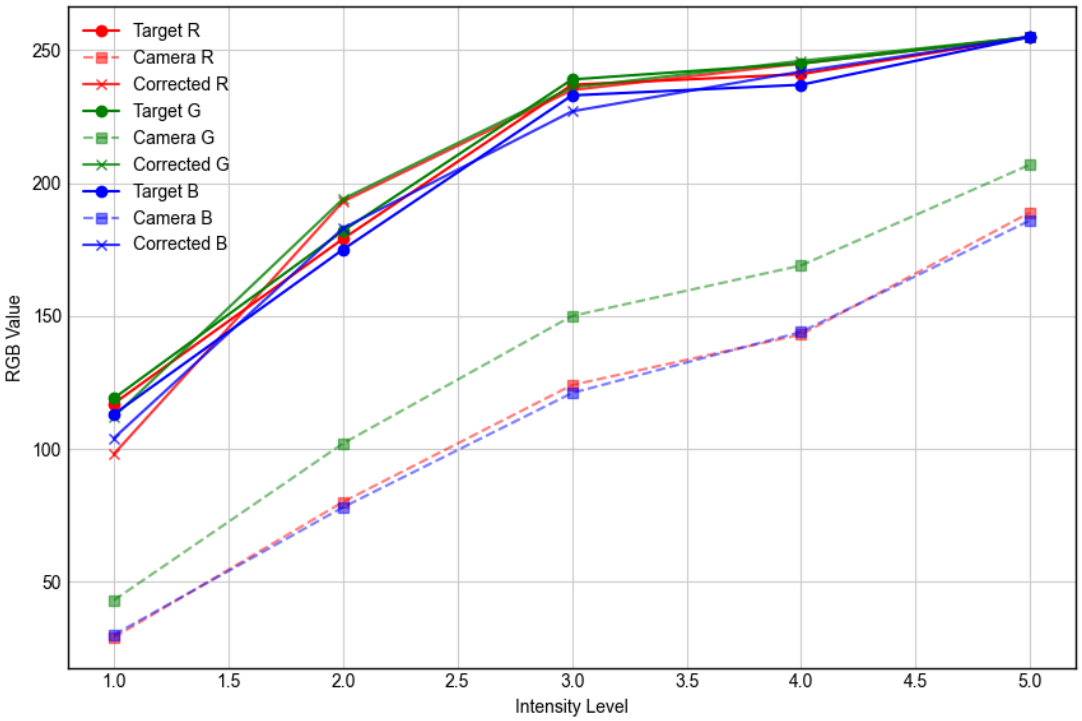


Figure 119: Calibration mapping for light through the grey color filter: camera captured RGB vs spectral reference RGB vs corrected RGB for all color channels (R, G and B)

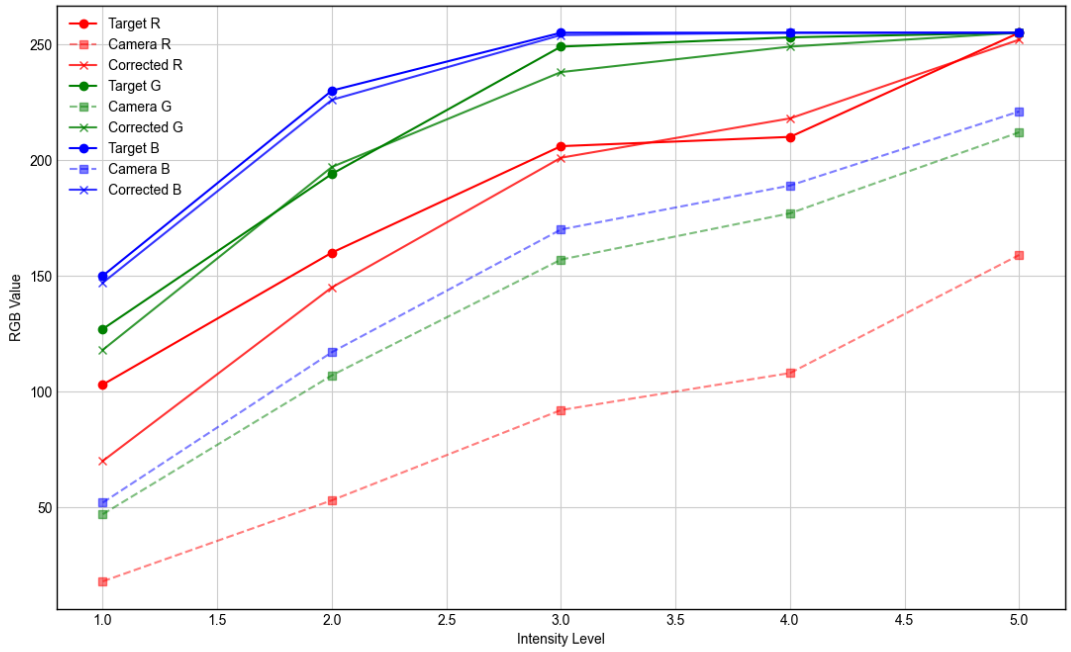


Figure 120: Calibration mapping for light through the light blue color filter: camera captured RGB vs spectral reference RGB vs corrected RGB for all color channels (R, G and B)

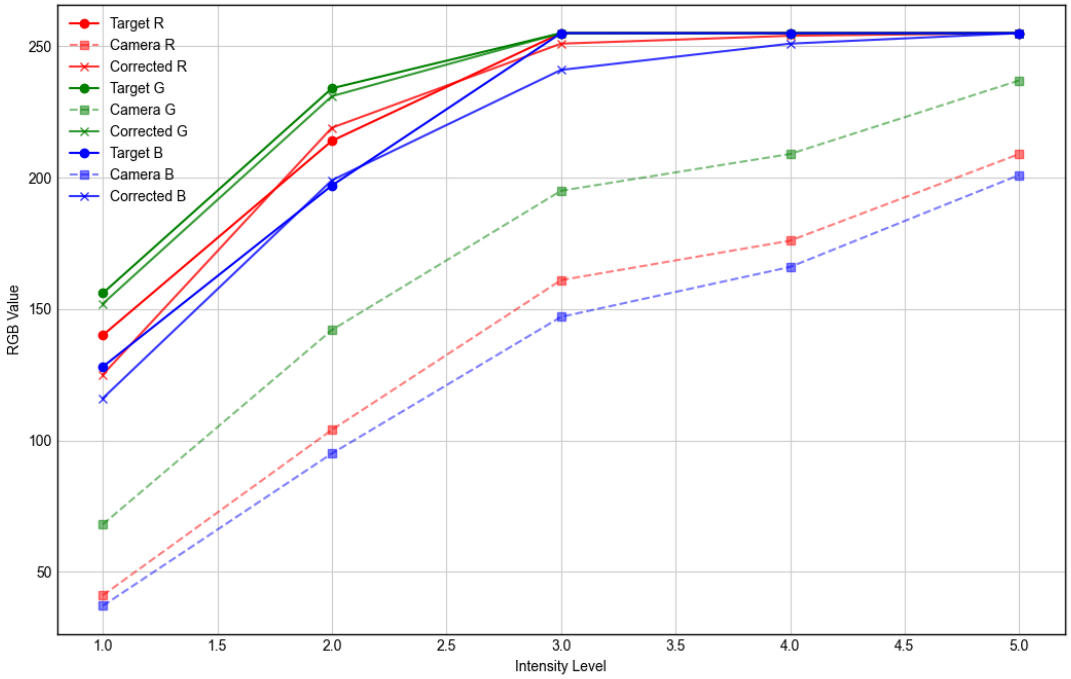


Figure 121: Calibration mapping for light through the light green color filter: camera captured RGB vs spectral reference RGB vs corrected RGB for all color channels (R, G and B)

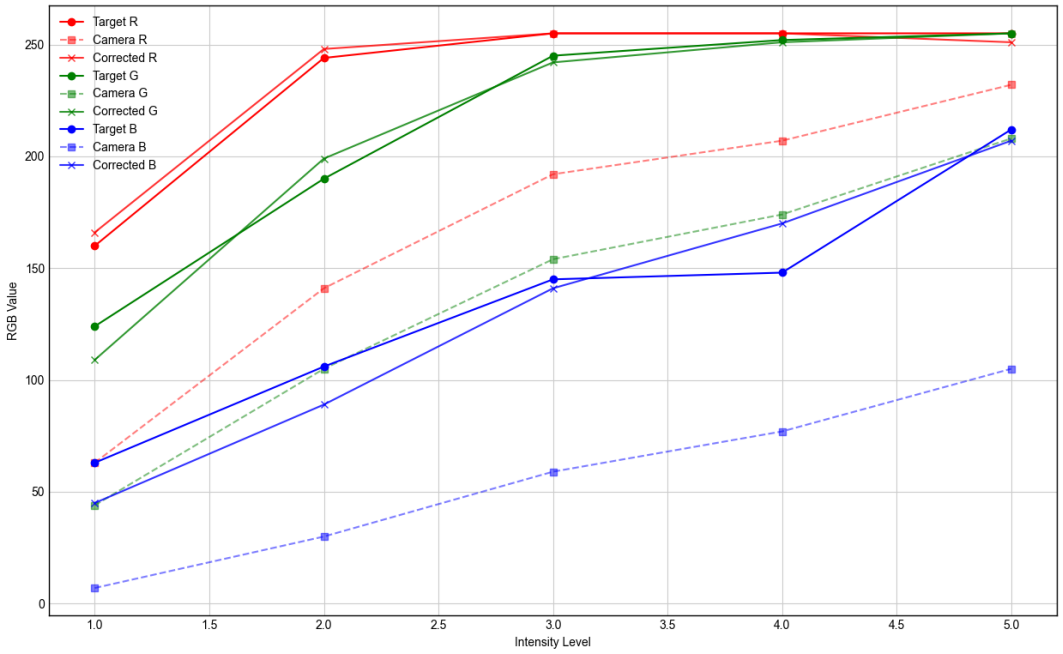


Figure 122: Calibration mapping for light through the (light) orange color filter: camera captured RGB vs spectral reference RGB vs corrected RGB for all color channels (R, G and B)

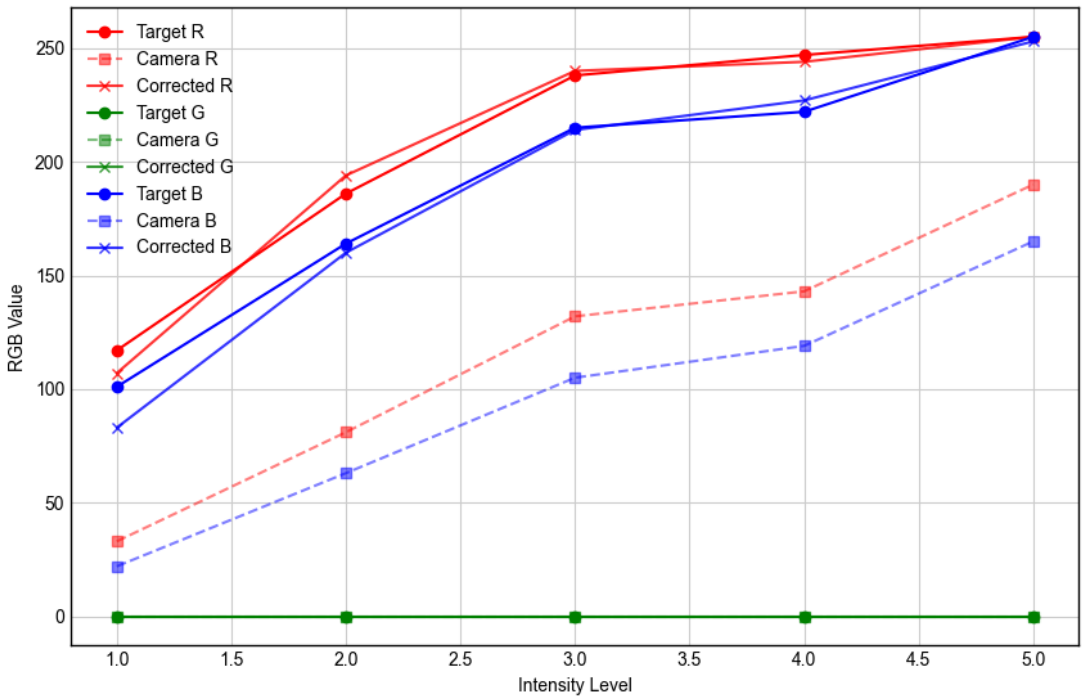


Figure 123: Calibration mapping for light through the purple color filter: camera captured RGB vs spectral reference RGB vs corrected RGB for all color channels (R, G and B)

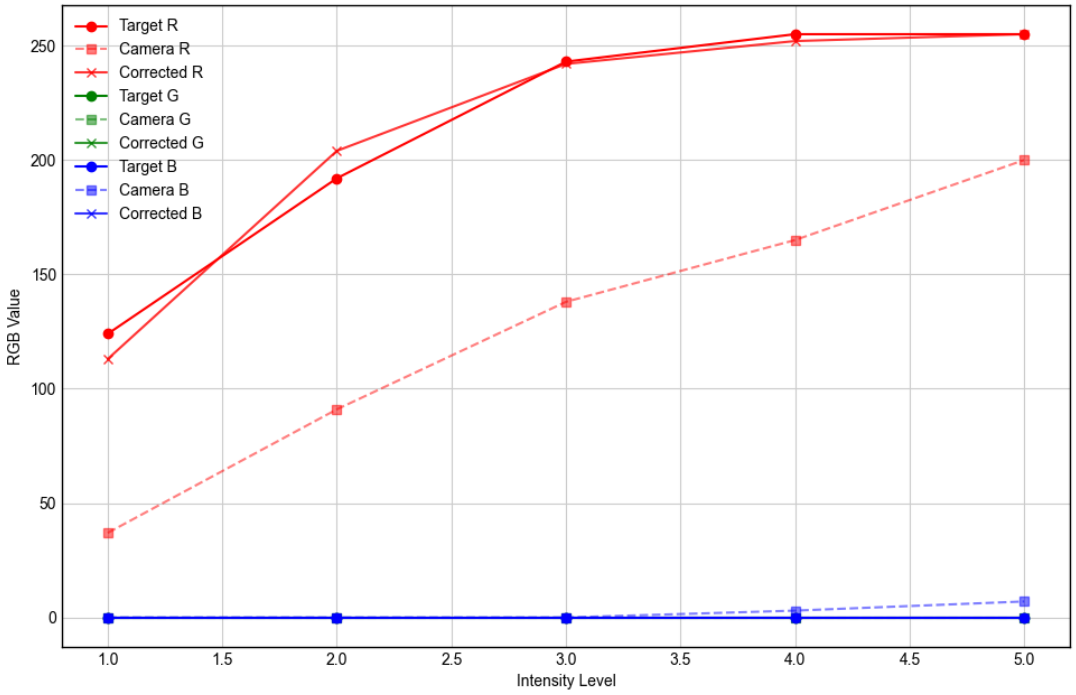


Figure 124: Calibration mapping for light through the red color filter: camera captured RGB vs spectral reference RGB vs corrected RGB for all color channels (R, G and B)

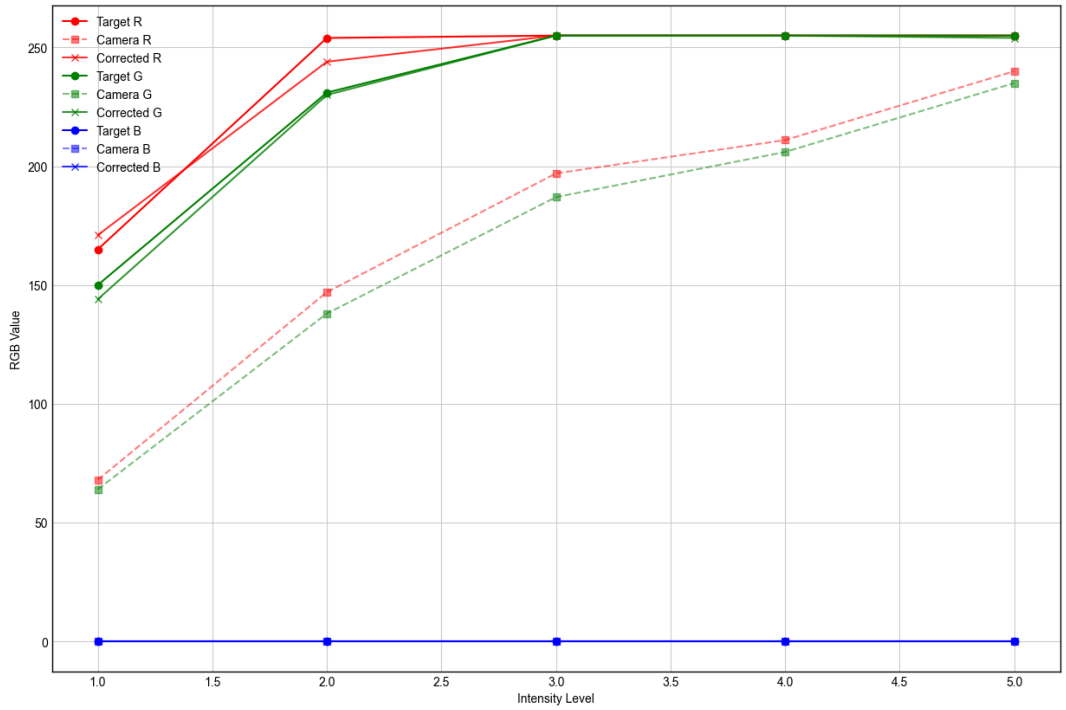


Figure 125: Calibration mapping for light through the yellow color filter: camera captured RGB vs spectral reference RGB vs corrected RGB for all color channels (R, G and B)

Appendix IV: Intensity calibration

This appendix contains additional materials used for intensity calibration, including clarification section (Detailed method for mapping pixel intensity to illuminance (Lux)), data table for the light-meter measurement results and actual pixel intensity (PEMAD), and regression analysis plots (correction).

IV.I Table: calibration data: camera intensity per fiber and the light meter intensity

Table 15 presents all the data obtained during the calibration process, as described in the methodology section.

Table 15: Intensity calibration measurements: pixel intensity vs illuminance (Lux) across all fibers and ND filters

Fiber	ND Factor	Light Source Intensity (%)	Distance (cm)	Pixel Intensity (Target / Actual)	Illuminance (Lux)
0	1	5	19.0	10.0 / 12.2	260
		22	19.0	70.0 / 73.2	1252
		35	19.0	125.0 / 124.6	2420
		78	19.0	180.0 / 179.7	4080
		26	8.3	240.0 / 237.6	6920
1	1/8	9	15.8	10.0 / 11.2	714
		9	8.3	70.0 / 69.7	2400
		19	8.3	125.0 / 122.7	4610
		30	8.3	180.0 / 180.2	8580
		93	8.3	240.0 / 239.4	21400
2	1/32	31	8.3	10.0 / 10.3	9220
		21	2.1	70.0 / 72.4	35100
		36	2.1	125.0 / 126.9	73500
		79	2.1	180.0 / 180.3	134400
		100	1.4	240.0 / 235.1	Out of scope
3	1/128	5	0.8	10.0/ 14.2	13400
		22	0.8	70.0 / 69.6	66100
		37	0.8	125.0 / 124.9	137300
		88	0.8	180.0 / 177.5	Out of scope
		100	0.1	240.0 / 226.5	Out of scope
4	1/512	6	0.1	10.0 / 11.5	28600
		27	0.1	70.0 / 69.2	127200
		71	0.1	125.0 / 121.4	Out of scope
		100	0.1	180.0 / 165.0	Out of scope
		–	0.1	240.0 / –	Out of scope
5	1/2048	35	0.1	10.0 / 10.3	196300
		100	0.1	70.0 / 32.2	Out of scope
		–	0.1	125.0 / –	Out of scope
		–	0.1	180.0 / –	Out of scope
		–	0.1	240.0 / –	Out of scope

IV.II Detailed method for mapping pixel intensity to illuminance (Lux)

To establish the relationship between the mean pixel value (called Y here) of the ROI of each fiber with the measured illuminance E, a regression model was created for each individual fiber. Polynomial models of increasing degree $d \in \{1, 2, 3, 4\}$ were tried to find the best possible fit, using `numpy.polyfit` in Python.

The general form of the polynomial model is:

$$\hat{E}(Y_m) = a_0 + a_1 Y + a_2 Y^2 + \cdots + a_k Y^k \tag{11}$$

where:

- $\hat{E}(Y)$ is the predicted illuminance (Lux),
- Y is the mean pixel intensity,
- k is the degree of the polynomial,
- $\{a_0, a_1, \ldots, a_k\}$ are the model coefficients determined with ordinary least squares (OLS) internally in np.polyfit.

The predicted illuminance $(\hat{E}(Y))$ is the estimated light level corresponding to a given pixel intensity Y, predicted from the corresponding calibration curve. The regression model is created based on the measured pixel intensity and the measured illuminance values, which can be found in Table 15 in Appendix IV.III. In this way, it is possible to estimate new lux values that are not measured. For instance, if a pixel value of 180 typically corresponds to around 4080 lux for fiber 0, then the regression model can estimate the lux value for 190 pixel intensity for the same fiber.

The accuracy of the model was computed with the RMSE. The best fit minimizes the RMSE value between the measured Lux value and the predicted illuminance value, the equation used is noted as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{m=1}^{n} \left(E_m - \hat{E}(Y_m)\right)^2} \tag{12}$$

Where:

- E_m is the measured illuminance (Lux),
- Y is the mean pixel intensity,

The simplest model (i.e., lowest degree) that achieved an RMSE with error $\varepsilon \leq 1\%$ ("nice to have" requirement [8]) of the maximum measured illuminance for each fiber is selected.

$$\text{RMSE}_{\max} = \varepsilon \cdot \max(E) \tag{13}$$

Where:

- $\max(E)$ is the maximum measured illuminance,
- ε is the acceptable error requirement.

Each fiber has its own unique calibration curve with specific (polynomial) coefficients, which maps the measured pixel intensity values to illuminance values. These calibration models take the ND filters applied to each fiber into account. In the main code, these coefficients are stored into a dictionary of calibration models, per fiber. Using these models, the measured grayscale mean pixel intensities for each fiber are converted to lux values. These lux values are then stored next to the pixel intensities in the main dictionary. Finally, the lux values corresponding to the selected fiber are displayed on an extra y-axis in the resulting plots.

IV.III Regression (calibration) curve for each fiber

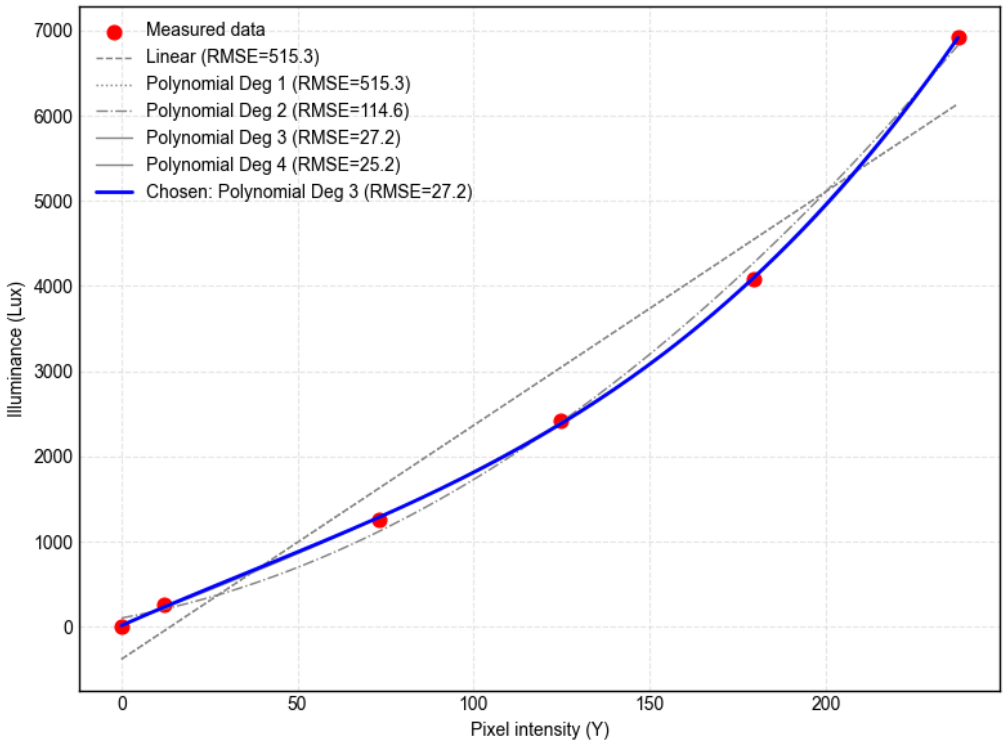


Figure 126: Calibration (regression) curve selection for fiber 0 intensity data

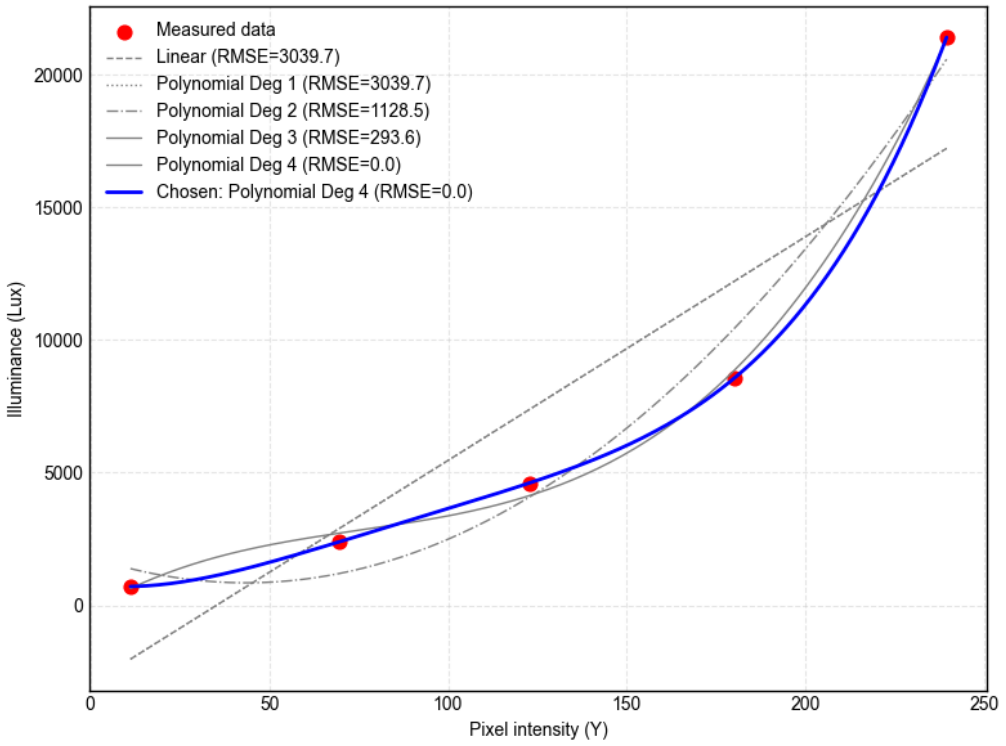


Figure 127: Calibration (regression) curve selection for fiber 1 intensity data

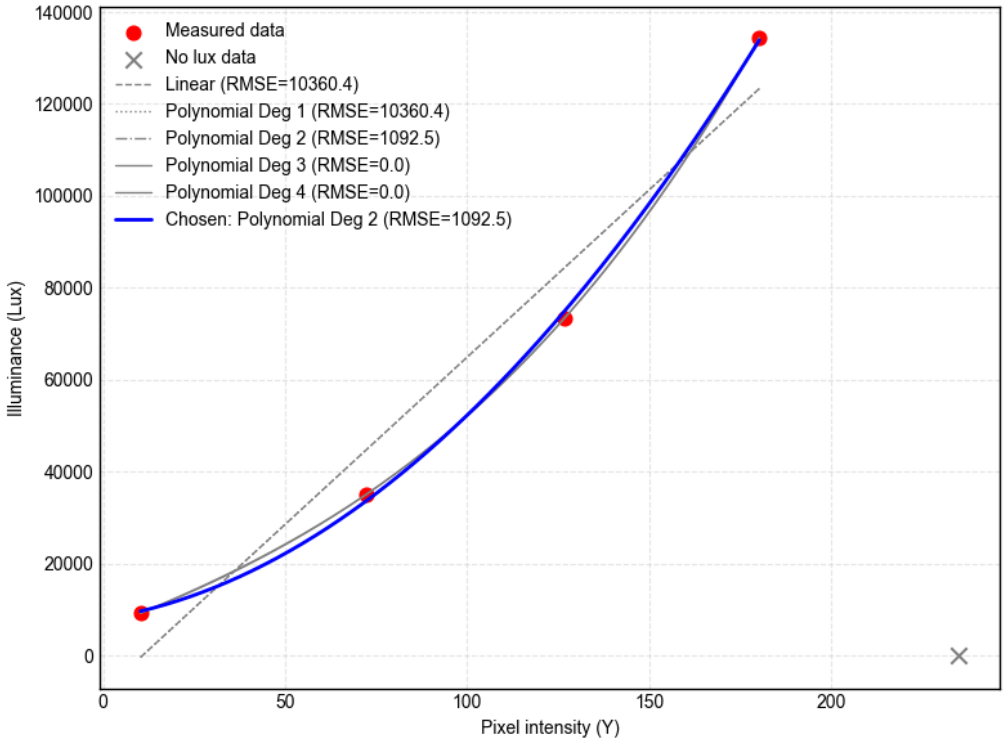


Figure 128: Calibration (regression) curve selection for fiber 2 intensity data

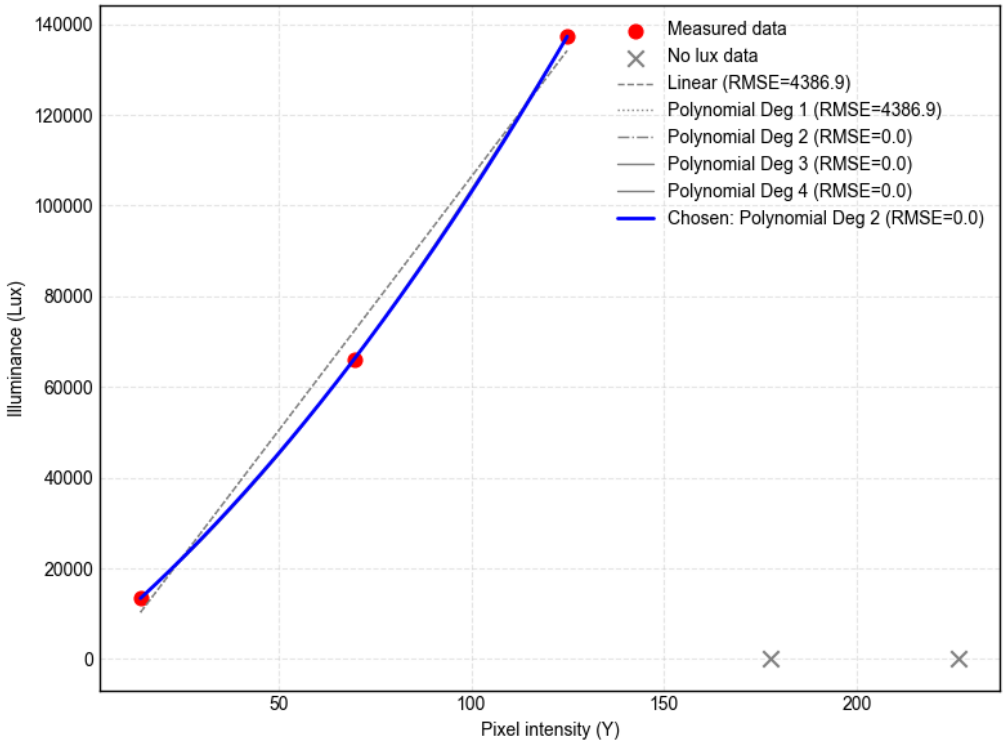


Figure 129: Calibration (regression) curve selection for fiber 3 intensity data

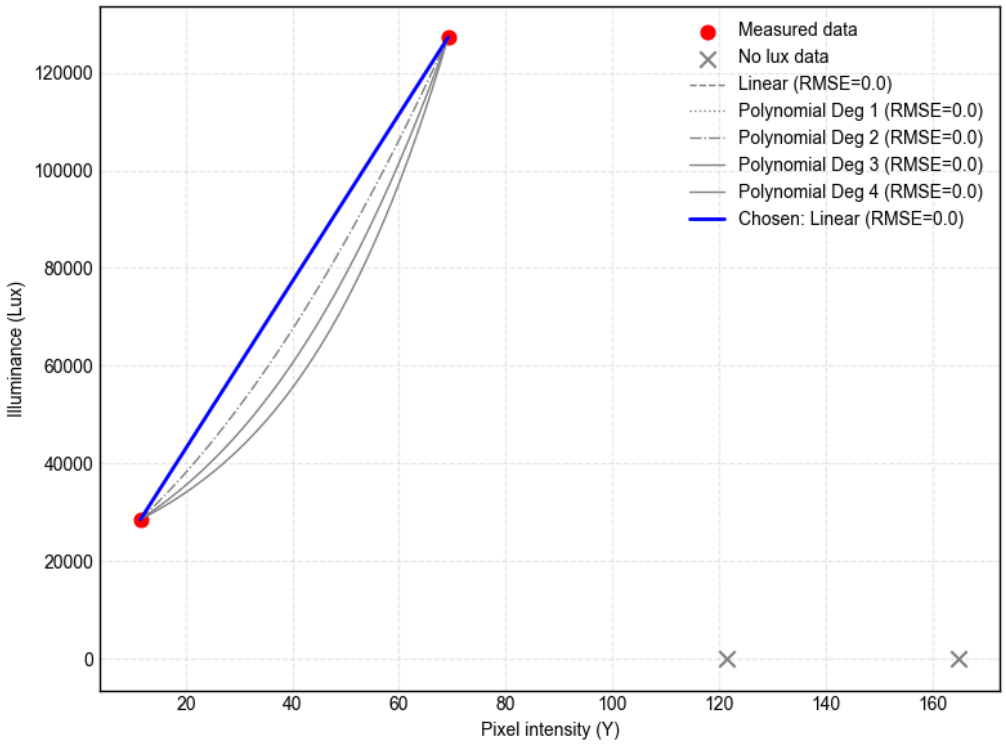


Figure 130: Calibration (regression) curve selection for fiber 4 intensity data (Not useful due to lack of illuminance (lux) data)

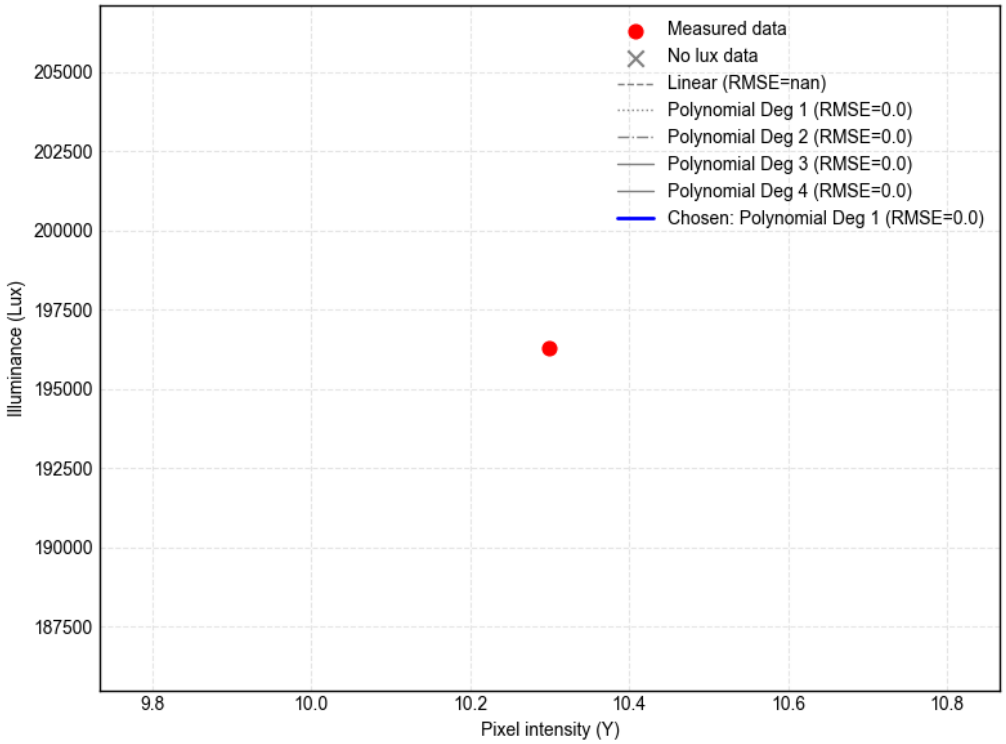


Figure 131: Calibration (regression) curve selection for fiber 5 intensity data (Not useful due to lack of illuminance (lux) data)

Appendix V: Validation plan

In this appendix, the initial validation plan is provided as a proposal to the NFI. This plan is constructed in accordance with the outline (guidelines) from the NFI. Providing all relevant sections needed according to the validation guideline, including introduction, materials, method, and predefined performance characteristics, with criteria.

Validation plan identification (e.g., NFI project number, version number, date, division, team, etc.)	Version 1 November 6, 2025 Team ..
Involved personnel (project owner, project leader, etc.)	Jacinta Jansen Ouke de Jonge
Author (submitter in Inception)	Nigel Moos
"Shadow" (controller in Inception)	..
Authorizer (process owner in Inception)	..
Executing researcher(s)	Nigel Moos Jacinta Jansen Ouke de Jonge

Name and signature

Name and signature

Name and signature

Name and signature

V.I Introduction

The Pyrotechnical/Explosive Materials Analysis Device (PEMAD) for the Burnmeister project is designed to objectively measure and analyze three combustion characteristics to identify the combustion behavior of a pyrotechnic or explosive mixture:

- Flame color,
- Combustion intensity (brightness of the light output),
- Burning rate (speed of combustion)

In short, this system is based on a dual Raspberry Pi (RPi) setup with a high-quality RPi HQ camera connected to six optical fibers and ND filter arrays. A visual camera is included as a visual reference for the researcher. Using a custom Python analysis routine based on openCV, the three combustion characteristics can be objectively analyzed.

The objective of this validation is to assess the detection range, accuracy, and precision of the PEMAD system. For each combustion characteristic, relevant performance characteristics, such as detection limits, selectivity/specificity, robustness, reproducibility and repeatability will be determined and evaluated against predefined criteria. Solid pyrotechnic and explosive materials will be used. When the measured performance characteristics meet the predefined criteria, as described in this validation plan, it would be possible to replace the so-called "hot needle test" and use the PEMAD system for the identification and characterization of pyrotechnic and explosive materials.

V.II Method and materials

The PEMAD system comes with a user manual [27] and an instruction guide [28], in which the method of using the device is described step by step. However, the system can only ignite and analyze certain materials with specific preparation, which is not specifically mentioned in the manual provided by previous study [8, 27, 28]:

- Only solid materials can be tested.
- The samples are only provided to the PEMAD system using a standard sampling spoon with a radius of $4mm$ and a fixed volume of $V = 1.34E^{-7}m^3 = 0.134cm^3$.
- The sample needs to be placed in an inert ceramic container using automatic sampling equipment and the sampling spoon.
- The sample is ignited with a glow plug heated to at least 580 degrees Celsius [8].
- The PEMAD system must be capable of testing the range of grain sizes encountered in practice.

V.III Validation approach

According to the "Validatie van Onderzoeksmethode" ("validation of Analytical Methods") guideline document of the NFI [18], "*a newly developed method whose performance characteristics have not yet been established. This method requires full validation. This means that all relevant performance characteristics must be determined and the results must be tested against predefined criteria*". The performance characteristics selected include, detection limit, selectivity/specificity, robustness, repeatability and reproducibility. Table 2 provides an overview of the performance characteristics, including definitions and relevance for the validation of the PEMAD system.

V.IV Performance characteristics

Table 16 provides an overview of the performance characteristics, the corresponding acceptance criteria, the materials required, and the test methods for each combustion characteristic. These are defined according to the combined qualitative and quantitative validation method according to the NEN 17025 standard.

Table 16: Overview of performance characteristics for each combustion characteristics in a qualitative and partial quantitative validation method (NEN 17025)

Performance characteristic	**Materials and execution**	**Burning rate criteria**	**Color detection criteria**	**Intensity detection criteria**
Detection limit	Combustion of 2 different mixtures (slow to fast rates and low to high light intensity combustions). Gunpowder (slow) and flash powder (fast) *	As the lower limit, the PEMAD must be able to detect slow combustions (in the range of seconds) such as gunpowder	N/A	The PEMAD must measure and plot a minimum of 5 frames (5 bars in the output plots) per combustion for fiber 0 to provide an accurate detection of the peak intensity
		In addition; the PEMAD must be able to detect fast combustions (in the range of milliseconds) such as flash powder	N/A	In addition; at least be able to detect gunpowder *(or other low light intensity emitting pyrotechnic)* and flash powder (high intensity emitting pyrotechnic)
Selectivity/ specificity	3 different oxidizer/fuel ratios 20/80 (m/m%), 50/50 (m/m%), 80/20 (m/m%) *	The PEMAD must be able to measure changes in burning rate when the sample ratios change, and the researcher is able to identify these changes	The PEMAD must be able to measure color shifts when the sample ratios change, and the researcher is able to identify these color shifts	The PEMAD must be able to measure intensity shifts when the sample ratios change, and the researcher is able to identify these (peak) intensity shifts
	5 different color emitting samples (e.g., Sr = red, Ba = green, Cu = blue, etc.) *	N/A	The PEMAD must clearly distinct between color emitting components in sample mixtures (i.e., Cu mixtures must be indicated as blue and Ba mixtures as green), and the researcher is able to identify these color outcomes	N/A
	Small variations in grain size: average of xx μm, xx μm and, xx μm (of the same mixture) *	The PEMAD must clearly distinct between the same sample but with variations in grain size. (i.e., finer grain sizes of the same sample burn faster) **	The outcome must not exceed a deviation of ±5% per RGB channel from the mean of the measurements of the 3 different grain sizes ($RSD \leq 5\%$) **	The PEMAD must clearly distinct between the same sample but with variations in grain size. (i.e., finer grain sizes of the same sample have higher combustion intensity) **

continued on next page

Performance characteristic	Materials and execution	Burning rate criteria	Color detection criteria	Intensity detection criteria
Robustness	**Small variations in the starting volume ($\pm$ 10 % w.r.t the sampling spoon volume of $V = 0.134cm^3$) of the same sample** (that meet the performance criteria for selectivity and detection limit) *	The outcome must not exceed a time (s) deviation of ±5% from the mean of the measurements of the 3 different starting volumes (relative standard deviation; $RSD \leq 5\%$) **	The outcome must not exceed a deviation of ±5% per RGB channel from the mean of the measurements of the 3 different starting volumes ($RSD \leq 5\%$) **	The outcome must not exceed a peak intensity deviation of ±5% from the mean of the measurements of the 3 different starting volumes ($RSD \leq 5\%$) **
	Repetition of same measurement with the same mixture (that meet the performance criteria for selectivity and detection limit) performed at 3 **different room temperatures** *	The outcome must not exceed a time (s) deviation of ±5% from the mean of the measurements of the 3 room temperatures ($RSD \leq 5\%$) **	The outcome must not exceed a deviation of ±5% per RGB channel from the mean of the measurements of the 3 room temperatures ($RSD \leq 5\%$) **	The outcome must not exceed a peak intensity deviation of ±5% from the mean of the measurements of the 3 room temperatures ($RSD \leq 5\%$) **
Repeatability	8 repetitions of the same measurements with the same mixture (that meet the performance criteria for selectivity and detection limit) performed by the **same researcher: device dependency**	The outcome must not exceed a time (s) deviation of ±5% from the mean of the 3 measurements ($RSD \leq 5\%$) **	The outcome must not exceed a deviation of ±5% per RGB channel from the mean of the 3 measurements ($RSD \leq 5\%$) **	The outcome must not exceed a peak intensity deviation of ±5% from the mean of the 3 measurements ($RSD \leq 5\%$) **
Reproducibility	Repetition of same measurement with the same mixture (that meet the performance criteria for selectivity and detection limit) performed by 3 **different researchers: person dependency** *	The outcome must not exceed a time (s) deviation of ±5% from the mean of the 3 measurements ($RSD \leq 5\%$) **	The outcome must not exceed a deviation of ±5% per RGB channel from the mean of the 3 measurements ($RSD \leq 5\%$) **	The outcome must not exceed a peak intensity deviation of ±5% from the mean of the 3 measurements ($RSD \leq 5\%$) **

* 8 Measurements will be performed for each sample/mixture

** (And/or:) Based on the outcome, the researcher(s) draws the same conclusion all the times

Appendix VI: Validation; Risk analysis

In this section the risks regarding the use of the PEMAD system in practice, after validation were evaluated. For each risk, the likelihood of occurrence, detectability, and potential impact on casework are assessed using a scale of low, medium, or high (see clarification below Table 17). In addition, the necessary preventative measures or actions to overcome these risks are described. Table 17 provides a comprehensive overview of this risk analysis.

Table 17: Risk analysis for the PEMAD system in practice

Risk	**Likelihood**	**Detectability**	**Impact on casework**	**Preventive measures or actions**
Loss of video footage	Medium	High	High	Back-up of storage on external drives after 100+ measurements and cleaning the PEMAD internal storage. Action when it happens; perform measurement again after cleaning the storage space.
Freezing of the system	Low	High	Medium	Error message is implemented which warns for incorrect measurements and prevent freezing of the system. Prevents system rebooting.
Color or intensity calibration is disrupted	Low	Medium	High	Keep the settings and device identical. Clean inside of PEMAD and protection shield regularly. When parts or software is adjusted; calibration needs to be redone.
Variation during sample preparation (due to user input or pre-processing)	Medium/ High	Low	Medium/ High	Use standardized sample protocols: fixed volume, consistent positioning, fixed pre-processing to obtain the same grain size each time, training of users.
Failure to detect weak combustions (low light output or wrong measurement)	Medium	High	High	System provides error of detection and no measurement, perform measurement again.
Software error during analysis (e.g., wrong fiber selected)	Low	High	High	Validation of software output by verification with visual camera footage. System provides warning when wrong measurement or error arises. Plots of all fibers are provided to fall back to manual selection of the fiber.
Incorrect color detection due to over- or underexposure	Low	High	High	Use of ND filters ensure well exposed frames, checks with the visual footage of visual camera.
Fire in unintended areas of the PEMAD	Low	High	High	Perform measurements in fire-safe environment with fume extraction, and fire extinguishing equipment. Use emergency button. Check for broken cables.
Glow plug does not ignite mixture due to accumulation of residue	Medium	High	Medium	Clean glow plug regularly.
Automatic door for sampling the mixture gets stuck	Low	High	Low	Keep inside mechanism clean. Manually push the door down.

Scale clarification:

- **Likelihood:**
  - Low: unlikely to occur (rough approximation: <5% chance)
  - Medium: possible under certain conditions/with lack of certain actions (rough approximation: 5–50%)
  - High: more likely to occur (rough approximation: >50%)
- **Detectability:**
  - Low: difficult to notice
  - Medium: can be noticed, by checking twice
  - High: easily detected by system warnings or visual checks
- **Impact on casework:**
  - Low: no significant effect on results or conclusions
  - Medium: may require new measurement
  - High: could provide invalid results, need new measurement

Appendix VII: Validation; measurement protocol

Before any measurements are performed, a detailed plan is needed to list all required materials, determine the correct quantities, and cover each performance characteristic. By combining performance characteristic tests, such as the detection limit and repeatability experiments, the number of test cycles can be reduced. Eight replicates per condition were used.

All tests use, when not indicated otherwise:

- **Sample volume:** 0.134 cm^3 (obtained with a standardized sampling spoon)
- **Setup:** PEMAD system with sample provided in an inert ceramic cup
- **Environment:** (normally used) room temperature (*i.e., 21°C*)
- **Repetitions:** 8 per condition
- **Executing researcher(s):** sample provider/maker (Ouke de Jonge (NFI researcher)), auditor (Jacinta Jansen (NFI researcher)) and performing measurements (Nigel Moos (student))

VII.I Detection limit

Materials:

- 2 different kinds of pyrotechnic mixtures; gunpowder and flash powder are used to determine the lower and upper detection limits, respectively.
- For each mixture, prepare one big batch (at least 8* 0.134cm^3= 1.072cm^3)
- From the 2 big batches prepare 8 replicates (with standardized volume (V=0.134cm^3))

Table 18: Detection limit test supplies

Test ID	Mixture name	Volume	Reps	Total volume	Notes
DL-1	Gunpowder	0.134 cm^3	8	1.072 cm^3	Slow burn, low intensity
DL-2	Flash powder	0.134 cm^3	8	1.072 cm^3	Very fast, very high intensity
		Total samples	**16**	**2.144 cm^3**	

In this way, the detection limit of the PEMAD system can be assessed, while simultaneously evaluating the repeatability (i.e., device dependency) of these pyrotechnic mixtures, which represent the outer limits the PEMAD must detect. Eight repetitions were used to remove potential outliers.

VII.II Device repeatability

Materials:

- To ensure repeatability and minimize variability in the mixtures, smokeless powder is used to evaluate device dependency. The samples are prepared using premixing techniques to make them as uniform as possible. The use of smokeless powder reduces measurement disturbances and ensures known compositions.
- For each mixture, prepare one big batch (at least 8* 0.134cm^3= 1.072cm^3)
- From the 2 big batches prepare 8 replicates (with standardized volume (V=0.134cm^3))

Table 19: Repeatability test supplies

Test ID	Mixture name	Volume	Reps	Total volume	Notes
R-1	Smokeless powder 1	0.134 cm^3	8	1.072 cm^3	To validate the device dependency between multiple measurement tests
R-2	Smokeless powder 2	0.134 cm^3	8	1.072 cm^3	
		Total samples	**16**	**2.144 cm^3**	

In addition, if necessary, the test results of the 'Detection limit' and 'Selectivity/specificity: color emitting components' experiments can be used to supplement the data for more reliable results.

VII.III Selectivity/specificity: oxidizer/fuel ratio

Materials:

- 1 type of uniformly premixed pyrotechnic mixture (e.g., flash powder*) will be prepared with three different oxidizer-to-fuel ratios. Achieving a uniform mixture is important to ensure the accuracy of the intended ratios and to minimize variations that could affect combustion behavior and outcomes. The three oxidizer-to-fuel ratios are;
  - A=40/60 (m/m%), (lower oxidizer ratio will most likely not ignite (seen in pilot tests))
  - B=50/50 (m/m%),
  - C=80/20 (m/m%)
- For each ratio, prepare one big batch (at least 8* 0.134cm^3= 1.072cm^3)
- From the 3 big batches prepare 8 replicates (with standardized volume (V=0.134cm^3))

Table 20: Selectivity / specificity test supplies

Test ID	Mixture name	Oxidizer/fuel ratio	Volume	Reps	Total volume	Notes
A	Flash powder	40/60	0.134 cm^3	8	1.072 cm^3	To validate that with the PEMAD system expected color and intensity shifts across varying component ratios can be identified
B	Flash powder	50/50	0.134 cm^3	8	1.072 cm^3	
C	Flash powder	80/20	0.134 cm^3	8	1.072 cm^3	
			Total samples	**24**	**3.216 cm^3**	

*Flash powder was chosen because it's one of the hardest mixtures to detect in terms of combustion duration. If the system works well with flash powder, it means it will likely work with other, easier-to-detect mixtures too. The flash powder composition consist of: Potassium perchlorate $KClO_4$(m%) with aluminum Al (m%)

VII.IV Selectivity/specificity: color emitting components

Materials:

- 5 different kinds of pyrotechnic mixtures with one characteristic color emitting component (exact mixture composition will follow for the Bengal flares)
- For each mixture with one specific color emitting component, prepare one big batch (at least 8* 0.134cm^3= 1.072cm^3)
- From the 5 big batches prepare 8 replicates (with standardized volume (V=0.134cm^3)) for each specific color emitting pyrotechnic mixture

Table 21: Color selectivity test supplies

Test ID	Mixture composition weight (%)	Color	Volume	Reps	Total volume	Notes
Red	See below[1]	Red	0.134 cm^3	8	1.072 cm^3	Validate identification of red
Green	See below[2]	Green	0.134 cm^3	8	1.072 cm^3	Validate identification of green
Blue	See below[3]	Blue	0.134 cm^3	8	1.072 cm^3	Validate identification of blue
Orange	See below[4]	Orange	0.134 cm^3	8	1.072 cm^3	Validate identification of Orange
Pruple	See below[5]	Purple	0.134 cm^3	8	1.072 cm^3	Validate identification of purple
			Total samples	**40**	**5.36 cm^3**	

This test was used to evaluate the ability of color identification with the PEMAD system to distinguish between different color emitting components in mixtures.

The specific Mixture compositions with corresponding weight percentages (%) are:

[1] **Red:**

- Total of 20 grams of effect powder with:
- Strontium Carbonate — $SrCO_3$ (15%)
- Nitrocellulose- $C_6H_7O_2(ONO_2)_3$ (60%)
- Ammonium Perchlorate - NH_4CLO_4 (25%)

[2] **Green:**

- Total of 20 grams of effect powder with:
- Potassium Perchlorate — $KClO_4$ (20%)
- Charcoal — C (10%)
- Barium Nitrate — $Ba(NO_3)_2$ (20%)
- Phenolic Aldehyde — $C_{48}H_{42}O_7$ (15%)
- Magnalium — Al^+Mg Alloy (20%)
- PVC — $(CH_2 - CHCl-)_n$ (15%)

[3] **Blue:**

- Total of 20 grams of effect powder with:
- Aluminum — Al(1%)
- Potassium Chlorate — $KClO_3$ (25%)
- Nitrocellulose- $C_6H_7O_2(ONO_2)_3$ (74%)

[4] **Orange:**

- Total of 19 grams of effect powder with:
- $[C_6H_7O_2(NO_2)_x(OH)_{3-x}]$(60%)
- Strontium Carbonate — $SrCO_3$ (2%)
- Cryolite — Na_3AIF_6 (8%)
- NH_4CIO_4 (30%)

[5] **Purple:**

- Total of 20 grams of effect powder with:
- Potassium Perchlorate — $KClO_4$ (45%)
- Magnalium — Al^+Mg Alloy (15%)
- PVC — $(CH_2 - CHCl-)_n$ (20%)
- Titanium - Ti (20%)

VII.V Selectivity/specificity: grain size

Materials:

- 1 kind of pyrotechnic mixture (i.e., .. *the mixtures obtained by Ouke (name?)*?) with 3 different grain sizes (fine, medium, coarse, defined numerically after further analysis of the compounds). This mixture ensures that the grain size is the only variable.
- Prepare 3 big batches of the same mixture with 3 different grain sizes.
- From the 3 big batches, prepare 8 replicates (with standardized volume (V=0.134cm^3)) for each grain size batch.

Table 22: Selectivity/specificity test supplies for grain size

Test ID	Mixture name	Grain size	Volume	Reps	Total volume	Notes
G-fine	Vihtavuori N310	Fine	0.134 cm^3	8	1.072 cm^3	Finer grain size effect
G-med	Vihtavuori N110	Medium	0.134 cm^3	8	1.072 cm^3	Medium (standard) grain size effect
G-coarse	Vihtavuori N160	Coarse	0.134 cm^3	8	1.072 cm^3	Coarser grain size effect
			Total samples	**24**	**3.216 cm^3**	

This test assesses the sensitivity of the PEMAD system against variations in grain size.

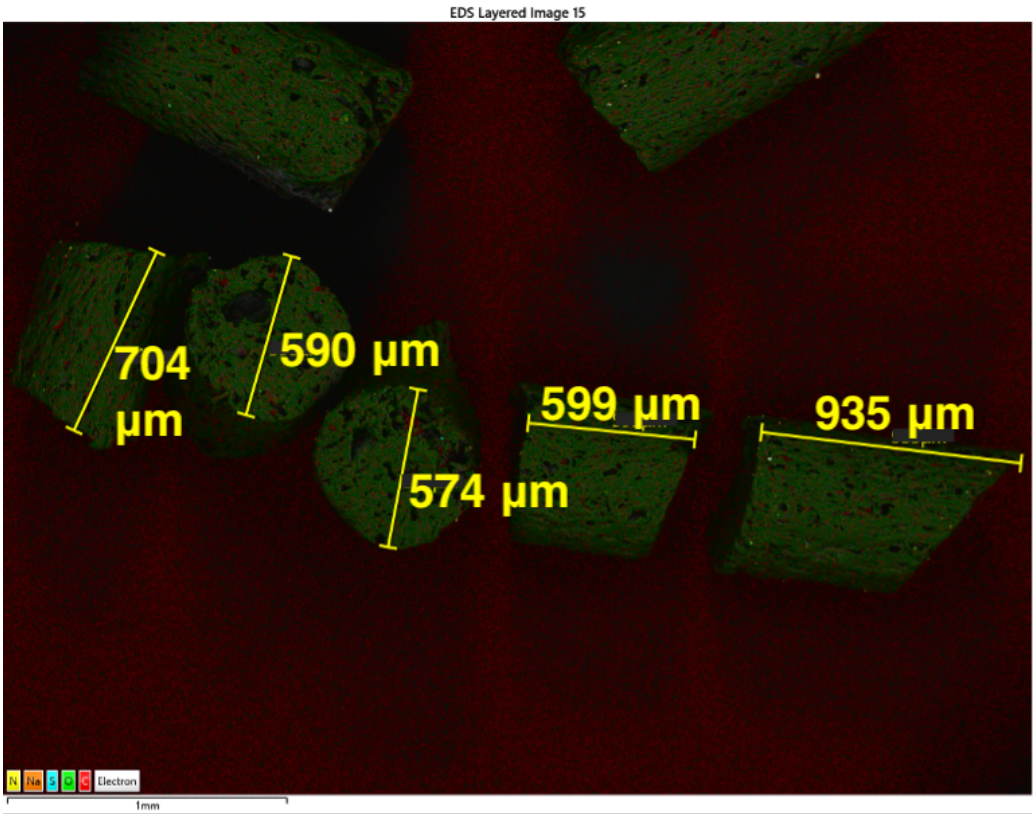


Figure 132: SEM result of material Vihtavuori N310 (fine), establishing the dimensions of the grain. Dimensions are shown with yellow measures

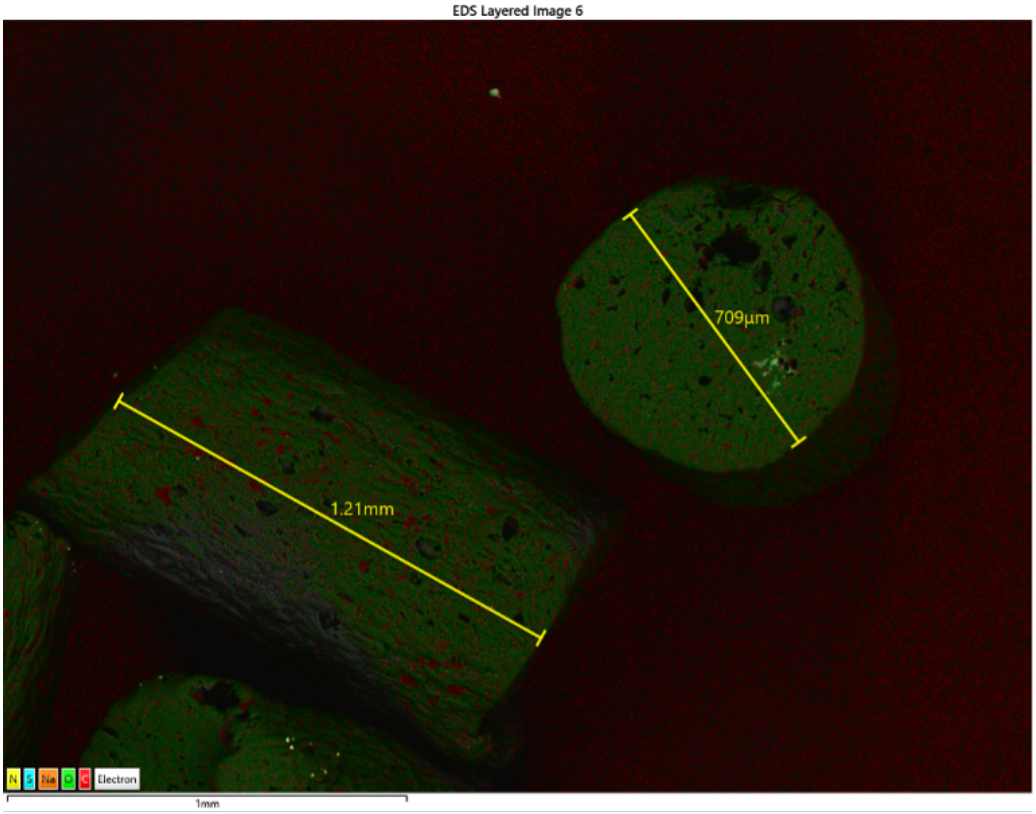


Figure 133: SEM result of material Vihtavuori N110 (medium), establishing the dimensions of the grain. Dimensions are shown with yellow measures

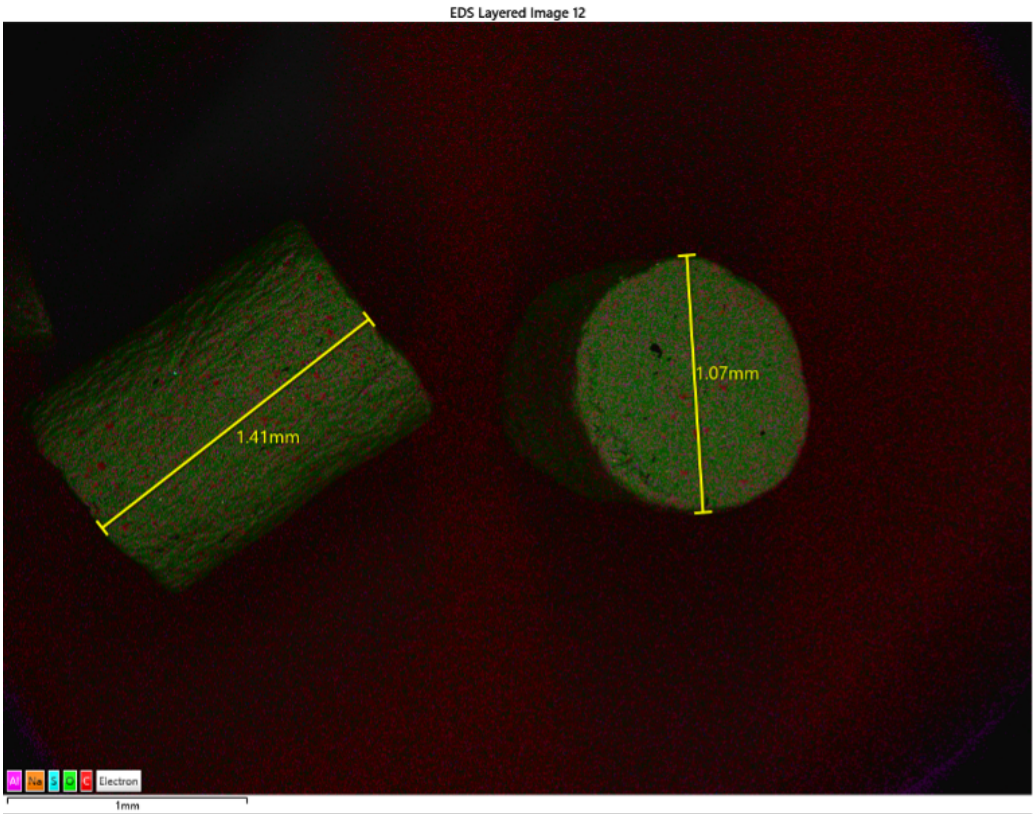


Figure 134: SEM result of material Vihtavuori N160 (coarse), establishing the dimensions of the grain. Dimensions are shown with yellow measures

VII.VI Robustness: starting volume

Materials:

- 1 kind of uniformly premixed pyrotechnic mixture with 3 different starting volumes with respect to the standard sampling spoon volume (90%, 100%, 110%)
- Prepare one big batch of the same mixture (at least 24 * 0.134cm^3= 3.2162 cm^3
- From the big batch, prepare 8 replicates (with standardized volume (V=0.134cm^3)) for each starting volume, with the provided standard spoons for 90%, 100% and 110% of the volume.

Table 23: Robustness test for starting volume supplies

Test ID	Mixture name	Volume portion	Volume	Reps	Total volume	Notes
V-90	Ratio B mixture*	90%	0.1206 cm^3	8	0.9648 cm^3	Lower volume effect
V-100	Ratio B mixture*	100%	0.134 cm^3	8	1.072 cm^3	Standard
V-110	Ratio B mixture*	110%	0.1474 cm^3	8	1.1792 cm^3	Higher volume effect
			Total samples	**24**	**3.216 cm^3**	

*Potassium perchlorate $KClO_4$(50%) with aluminum Al (50%)
This test evaluates the robustness of the PEMAD system against small deviations in volume during sample preparation.

The volume of a half-sphere is given by:

$$V = \frac{2}{3}\pi r^3$$

Given the original radius $r_0 = 0.4\,\text{cm}$, the volume is:

$$V_0 = \frac{2}{3}\pi (r_0)^3 = 0.134\,\text{cm}^3$$

To achieve 110% and 90% of the original volume:

$$V_{110} = 1.10 \cdot V_0 = 0.1474\,\text{cm}^3, \quad V_{90} = 0.90 \cdot V_0 = 0.1206\,\text{cm}^3$$

Solving for the required radius:

$$r = \left(\frac{3V}{2\pi}\right)^{1/3}$$

Therefore:

$$r_{110} = \left(\frac{3 \cdot 0.1474}{2\pi}\right)^{1/3} = 0.4124\,\text{cm} = 4.124\,\text{mm}$$

$$r_{90} = \left(\frac{3 \cdot 0.1206}{2\pi}\right)^{1/3} = 0.3836\,\text{cm} = 3.836\,\text{mm}$$

VII.VII Reproducibility: person dependency

changed with respect to the initial validation plan: no extra tests needed, researcher does not have to do specific tasks which can influence the measurements significantly, concluded during the meeting at NFI (17-07-2025). Interpretation of the outcomes is based on the resulting graphs (color) and numerical outcomes (intensity and burning time), which is objective. In addition, the influence in small volume variations is tested in the validation process as a performance criteria. In this way the deviation per researcher is taken into account regarding small variations in sample preparation.

Appendix VIII: statistics, choosing multi-comparison correction

This appendix contains the statistical reasoning and methods used to control false positives when performing multiple pairwise comparisons between compositions. With 16 compositions, this resulted in:

$$\binom{16}{2} = \frac{16!}{2!(16-2)!} = 120 \tag{14}$$

unique pairwise hypotheses (comparison tests) in total.

Testing each pair with $\alpha = 0.05$ (i.e., the usual significance level) without correction would lead to approximately six false positives by chance ($120 \times 0.05 = 6$), and the probability of obtaining at least one false positive across all 120 tests would be 99.4%. To clarify [29]:

$$1 - (1 - 0.05)^{120} \approx 0.994 \tag{15}$$

To avoid this, all pairwise comparisons were treated as a family of hypotheses with a more general question: can the PEMAD system distinguish compositions based on the combustion characteristics? This means that instead of each test having a $\leq 5\%$ chance of a false positive, now the entire set of 120 tests together has a $\leq 5\%$ chance of having one single false positive [29].

To control for false positives (detecting a difference when there is none) across the 120 tests, p-values were adjusted using the Holm-Bonferroni method, which controls the Family-wise Error Rate (FWER) with $\alpha = 0.05$. Holm-Bonferroni method controls false positives the same as the classic Bonferroni method, but with more statistical power to detect true differences. To clarify, this can be seen in the following equations [29, 30]:

Bonferonni divides α by the number of tests. So in this case for 120 tests:

$$\alpha = \frac{0.05}{120} \approx 0.000417 \tag{16}$$

In other words, every test needs a $p < 0.000417$ to be significantly different. This is very strict and reduces false positives, but also makes it likely to miss differences (false negatives).

The Holm-Bonferroni method orders p-values from smallest to largest ($p_1 \leq p_2 \leq \cdots \leq p_n$) and applies it stepwise:

$$p_i < \frac{\alpha}{(n - i + 1)} \tag{17}$$

where n is the total number of tests and i is the rank of the hypothesis (rank of the p-value of the comparison between two compositions).

- The smallest p-value (p_1) is tested at α/n (same as Bonferroni).
- Each next p-value is tested at a slightly less strict α (e.g., $\alpha/119$, $\alpha/118$, etc).
- Testing stops at the first non-significant result.

In short: Bonferroni uses the same (strict) threshold to every test, while Holm adjusts step by step, with strict for the smallest p-values, but less strict for the larger ones. This gives Holm more power to find real differences, while still keeping the overall risk of false positives at $\leq$5% [31]. This statistical analysis process is shown schematically in a flow-chart in Figure 15.

Other correction methods were also considered. The Benjamini-Hochberg method controls the false discovery rate (FDR). FDR is more sensitive (generally gives more significant results) and has more statistical power. However, it is not possible to be sure about the outcome of the test (comparing two groups to see if they are different on that metric). When many comparisons are performed, this method has a much higher chance of false positives, which was not suitable in this validation [29, 31].

Appendix IX: validation results for detection limit

This appendix presents the validation results of the detection limit tests performed with gunpowder and flash powder compositions. The following figures show the PEMAD analysis results for each test replicate (n = 8), illustrating the combustion intensity profiles over time and the corresponding color behavior for both compositions.

IX.I PEMAD analysis results for detection limit: gunpowder (Test ID: DL-1)

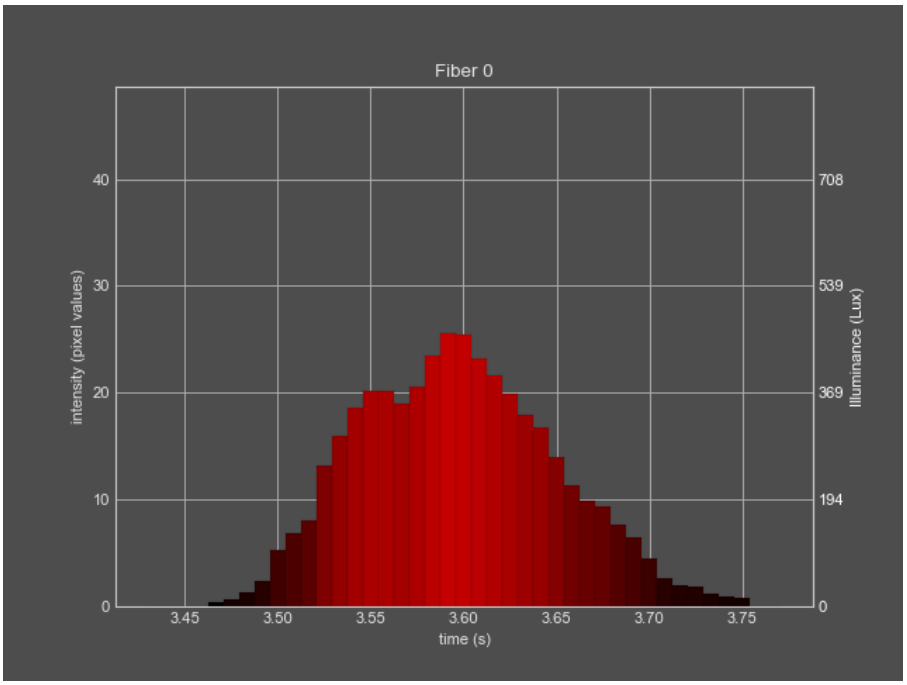


Figure 135: Result detection limit: Gunpowder test 1

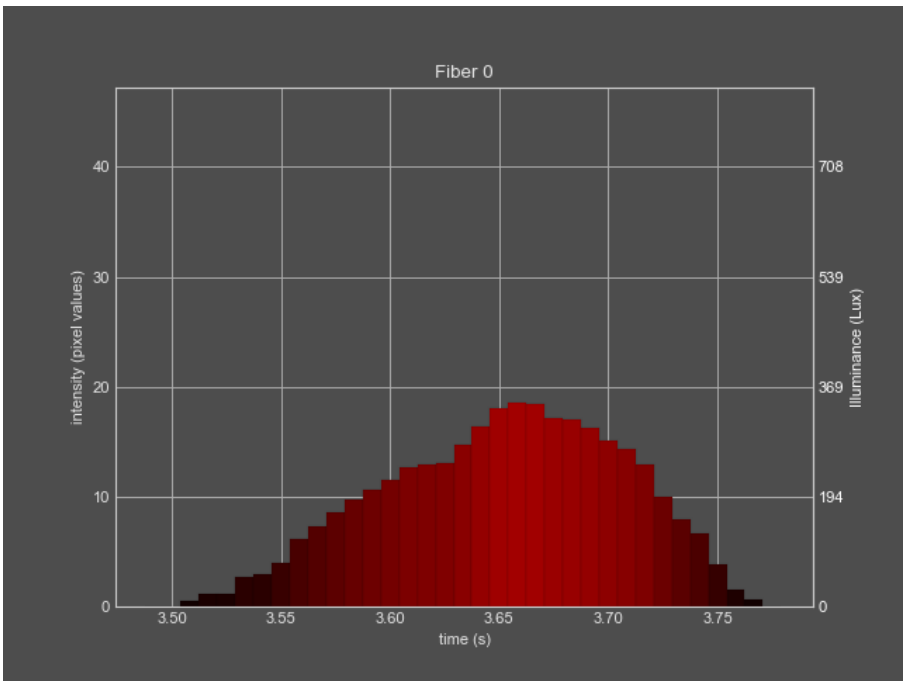


Figure 136: Result detection limit: Gunpowder test 2

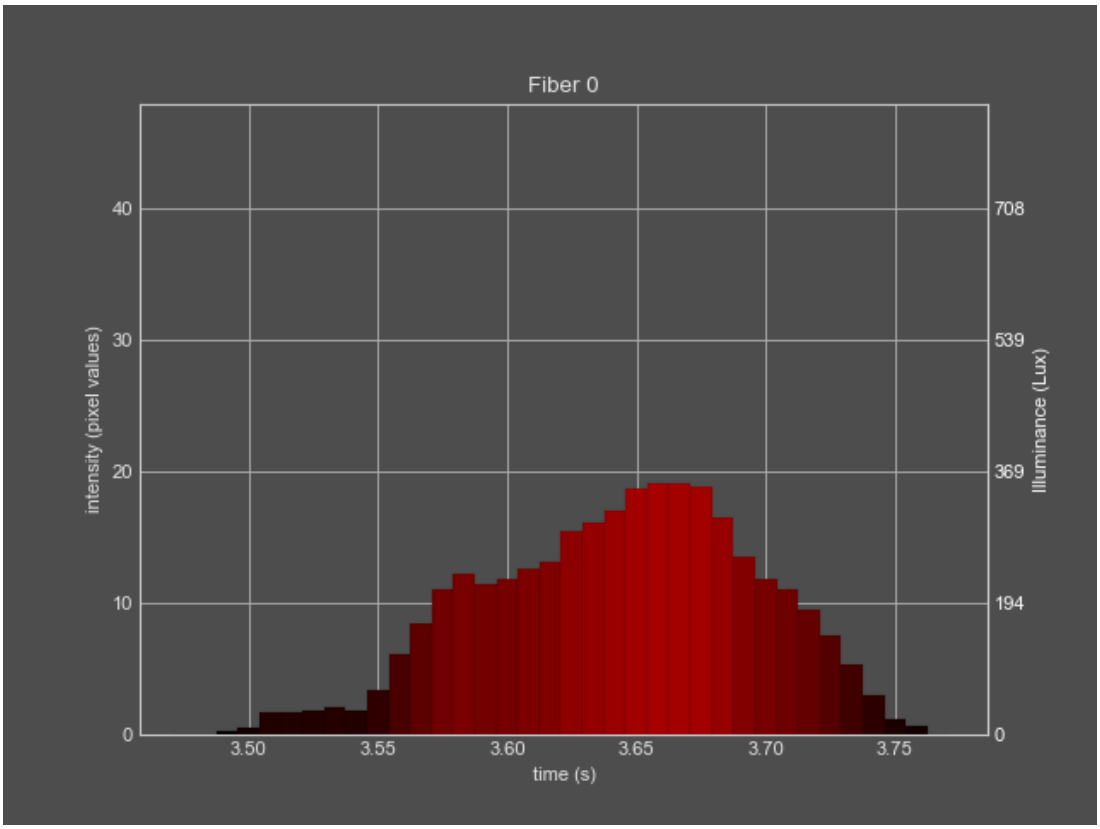


Figure 137: Result detection limit: Gunpowder test 3

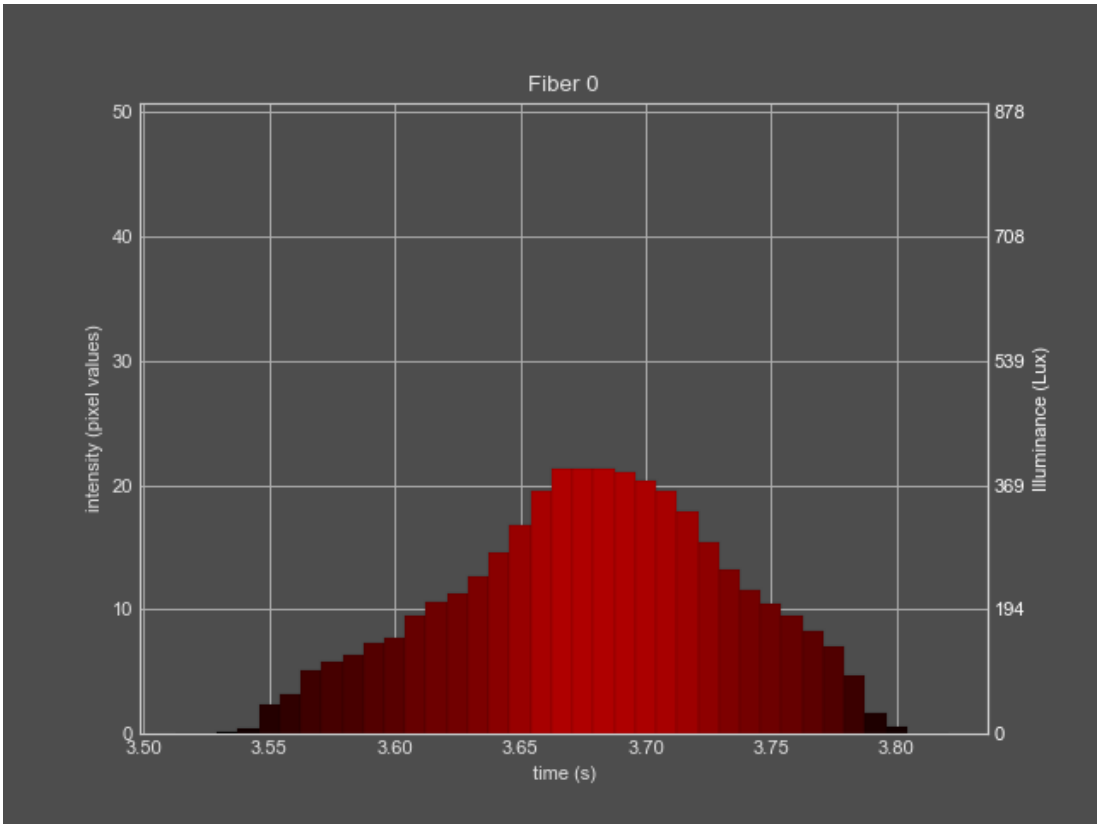


Figure 138: Result detection limit: Gunpowder test 4

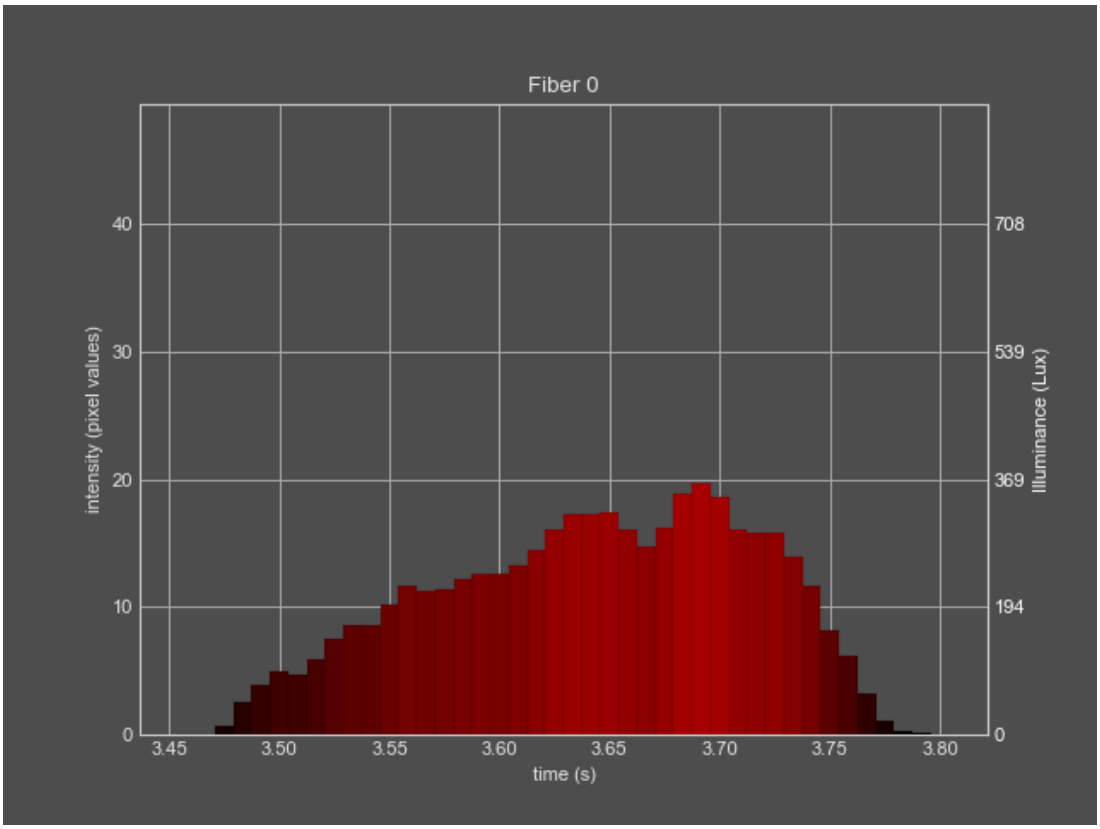


Figure 139: Result detection limit: Gunpowder test 5

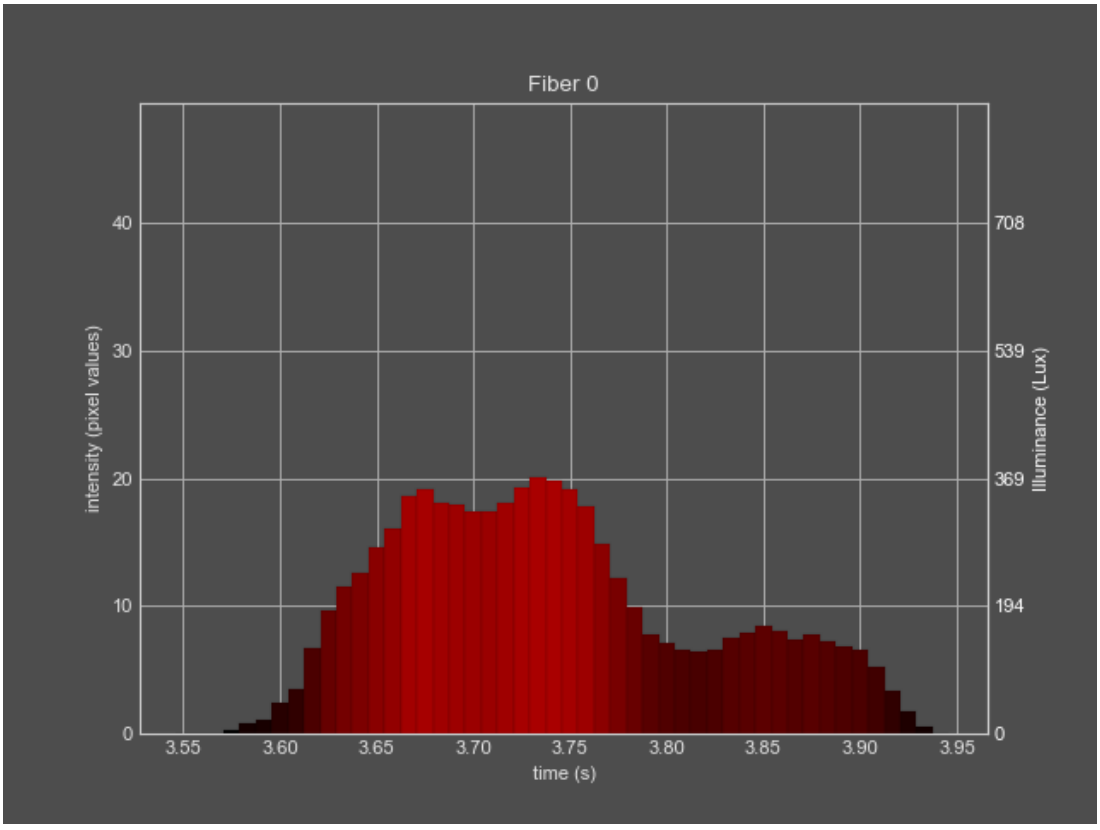


Figure 140: Result detection limit: Gunpowder test 6

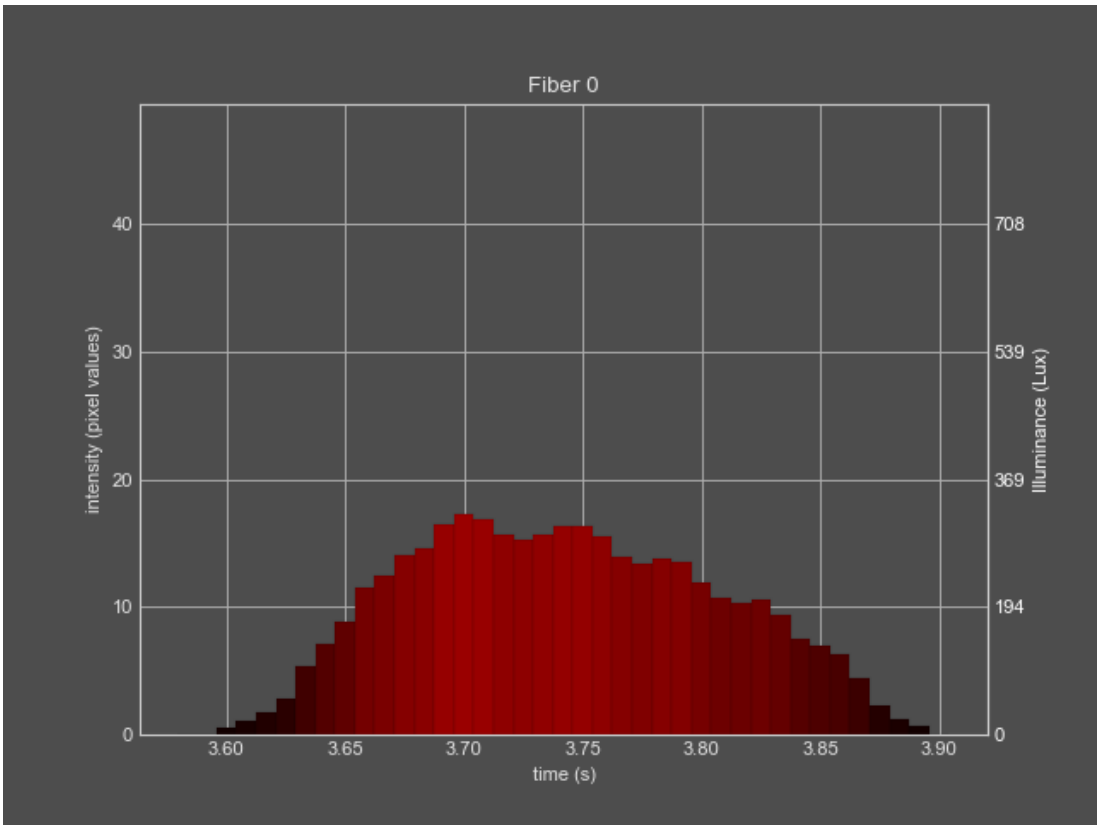


Figure 141: Result detection limit: Gunpowder test 7

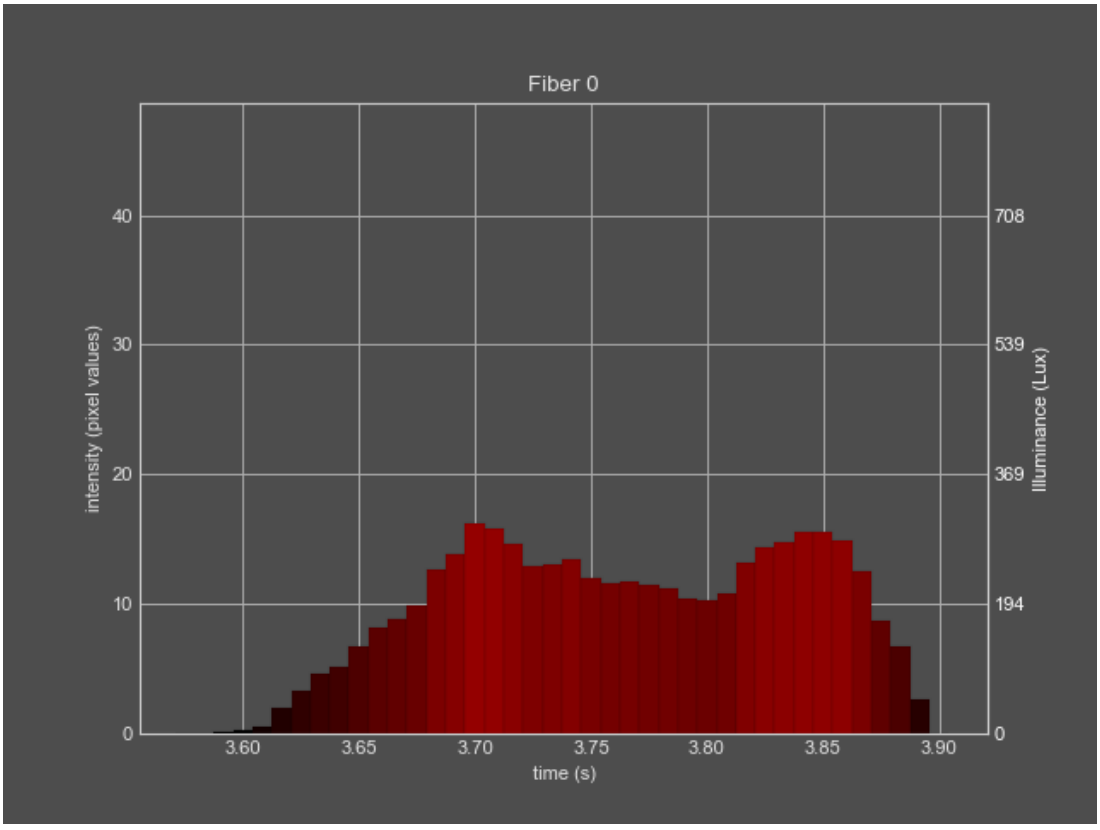


Figure 142: Result detection limit: Gunpowder test 8

IX.II PEMAD analysis results for detection limit: Flash powder (Test ID: DL-2)

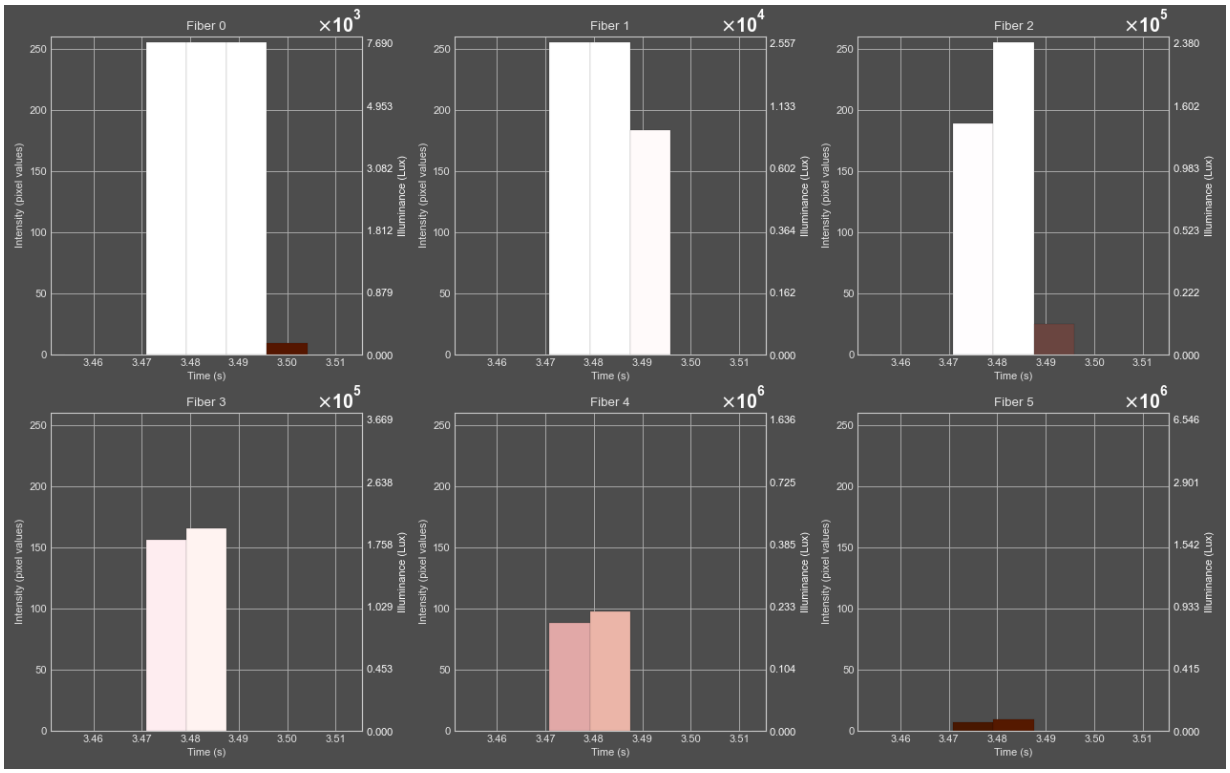


Figure 143: Result detection limit: flash powder test 1

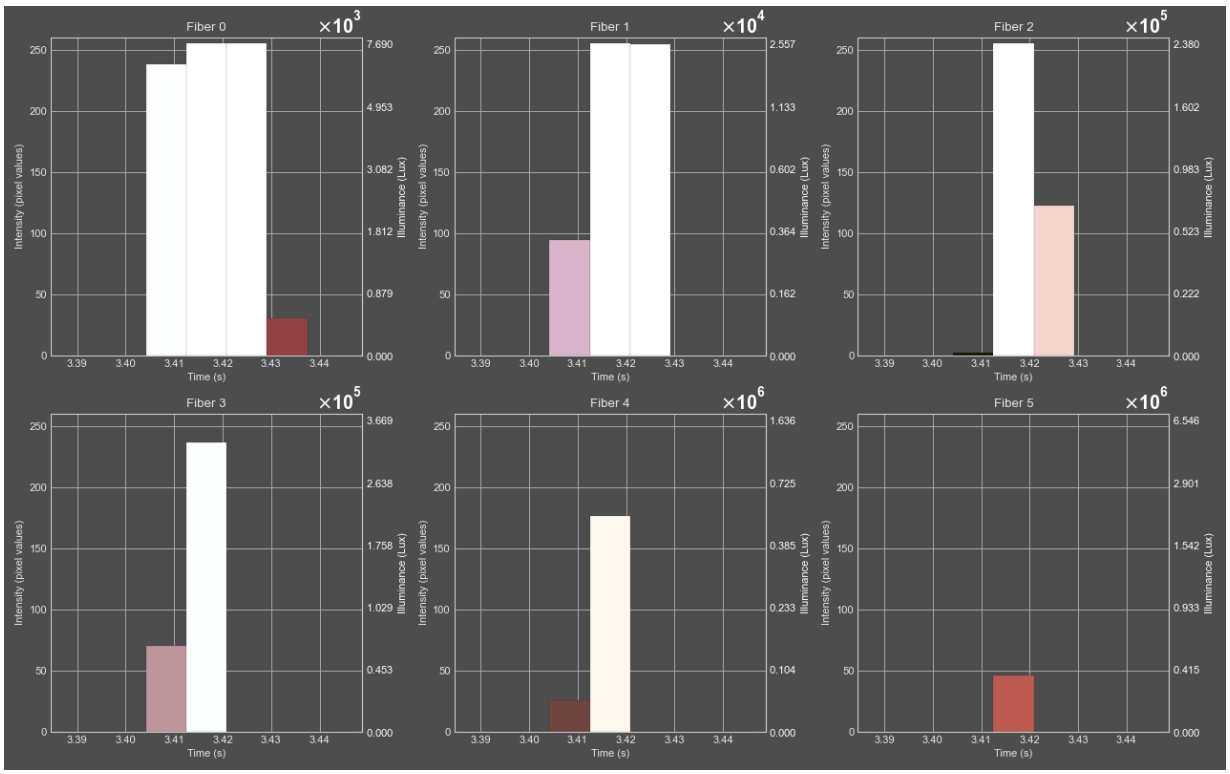


Figure 144: Result detection limit: flash powder test 2

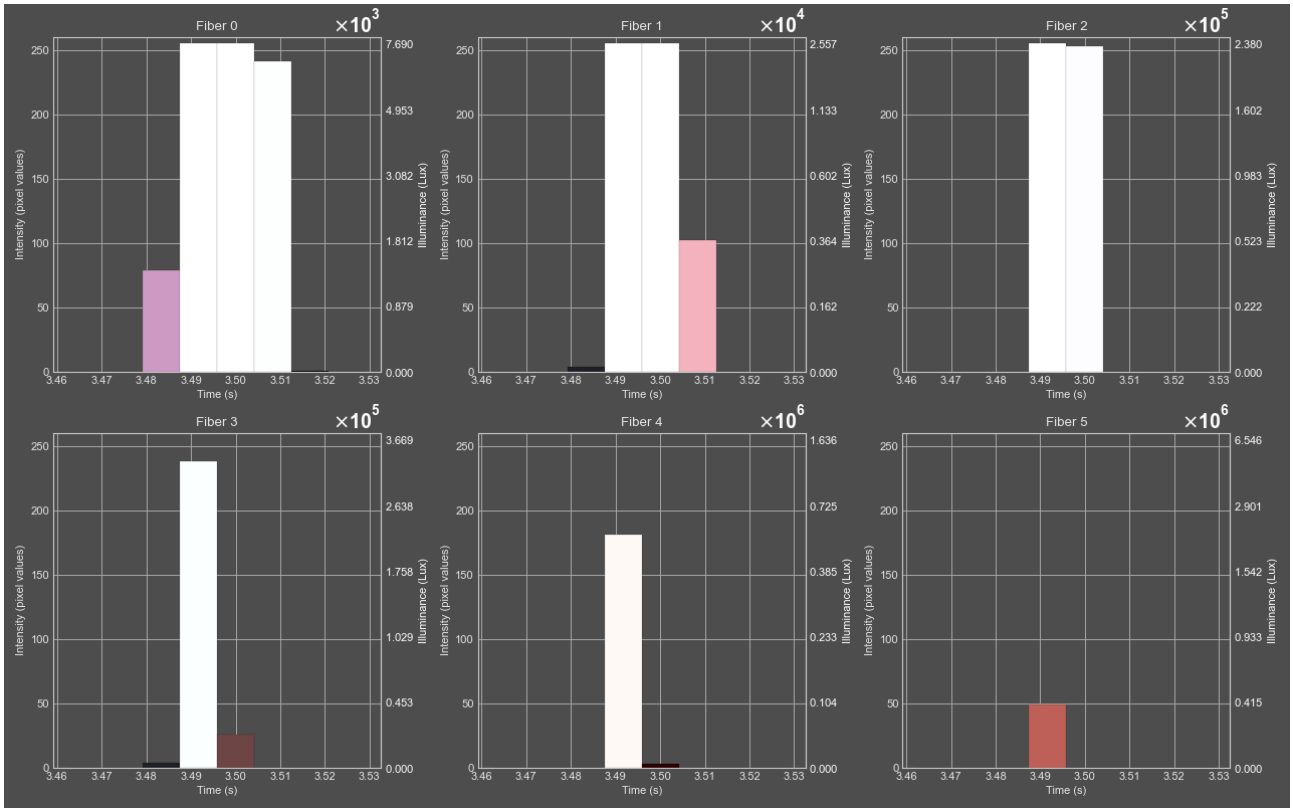


Figure 145: Result detection limit: flash powder test 3

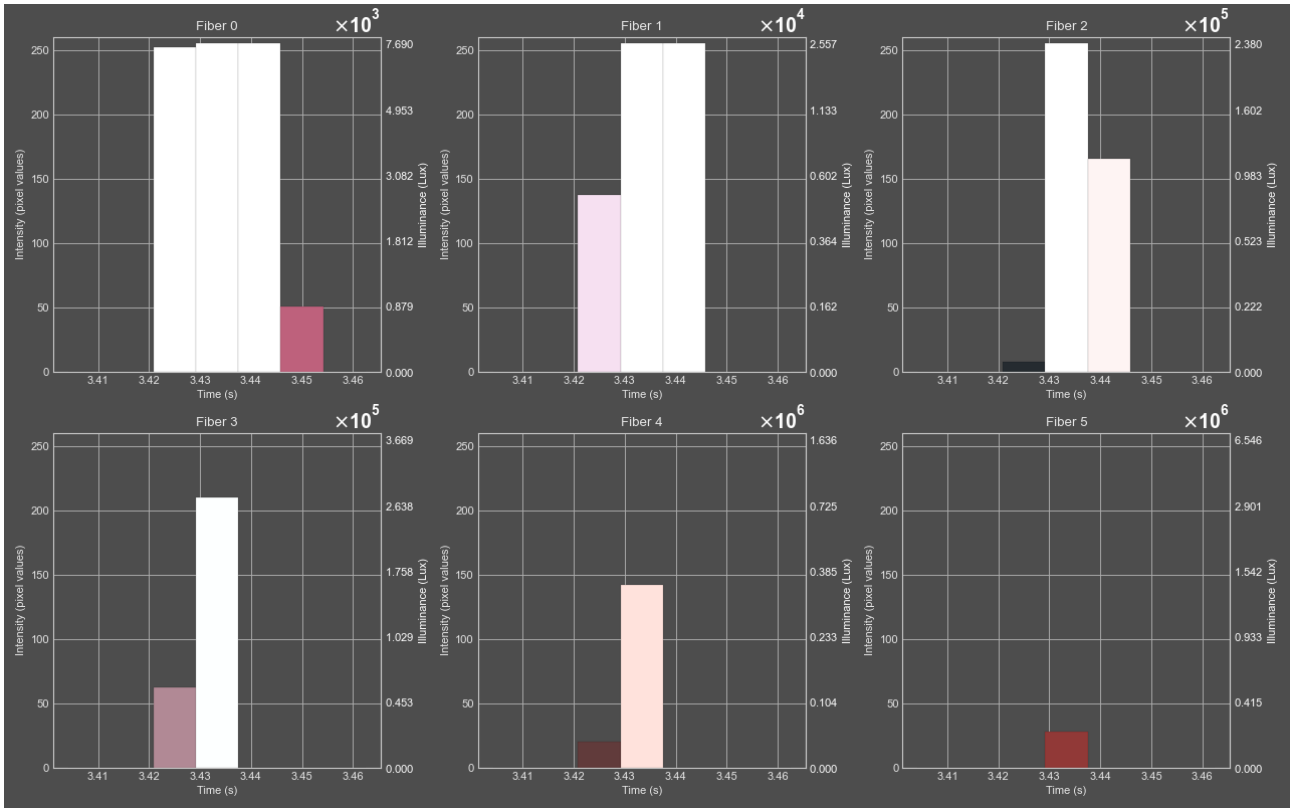


Figure 146: Result detection limit: flash powder test 4

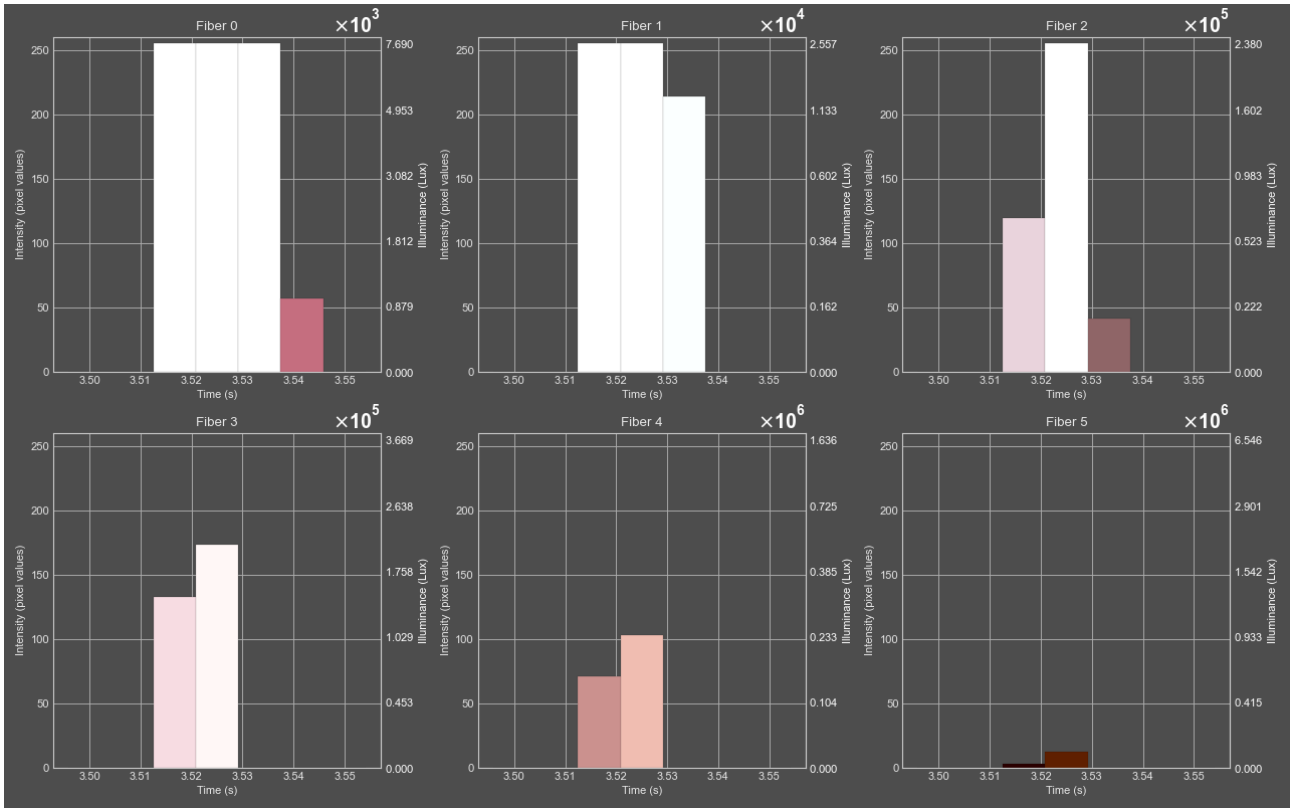


Figure 147: Result detection limit: flash powder test 5

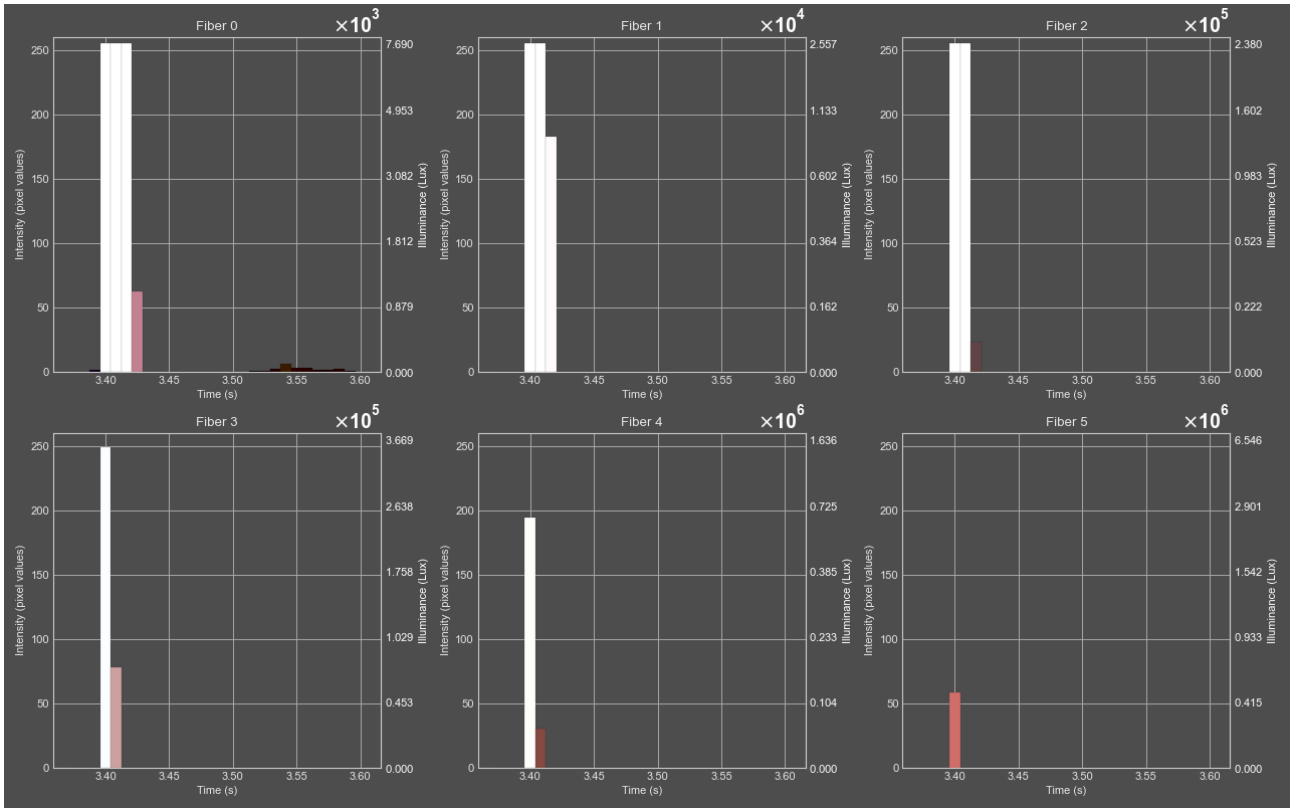


Figure 148: Result detection limit: flash powder test 6

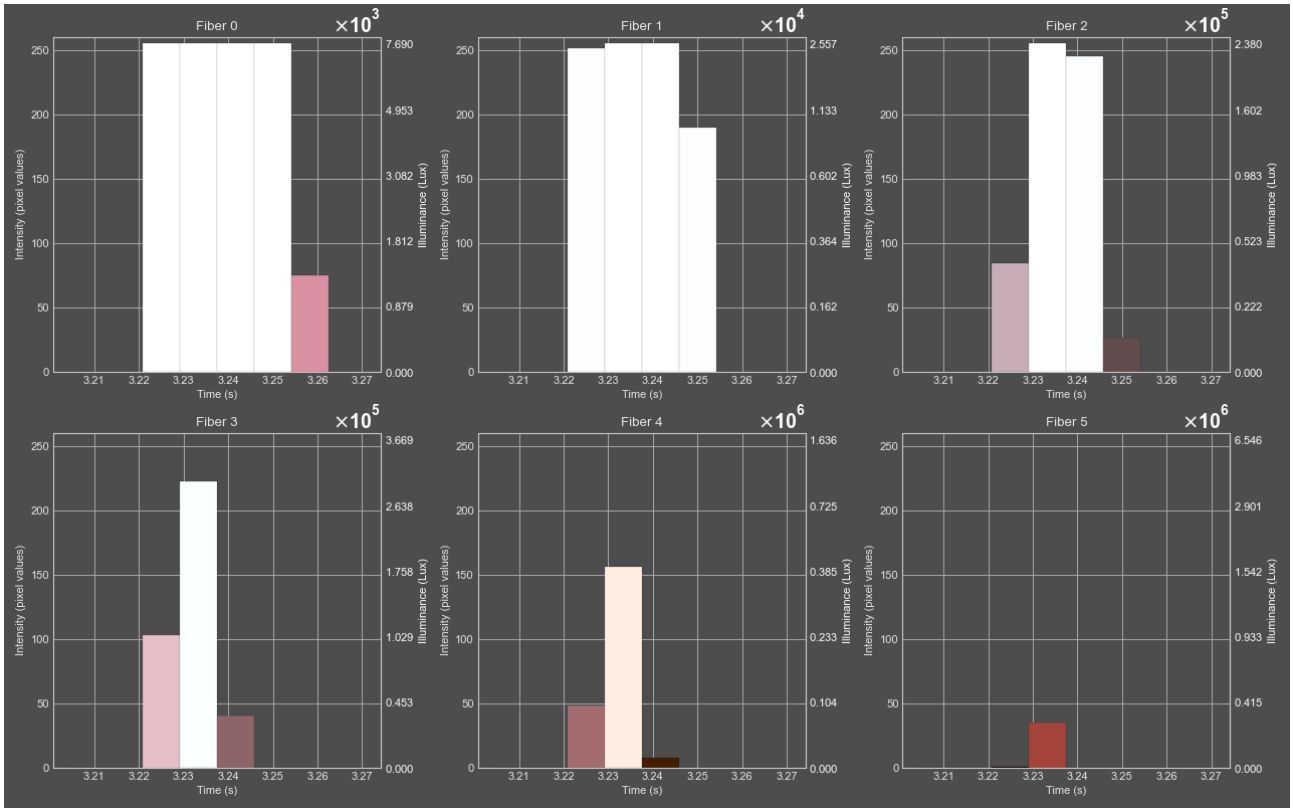


Figure 149: Result detection limit: flash powder test 7

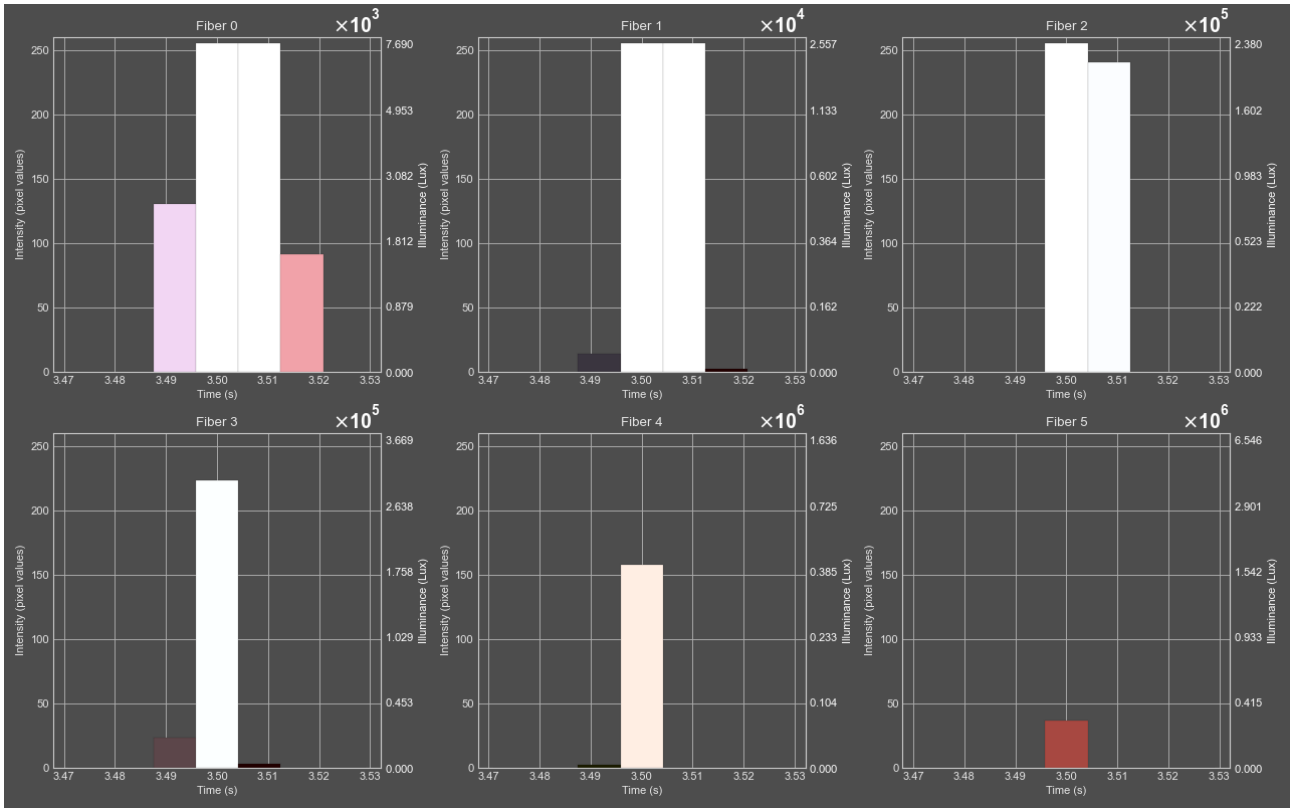


Figure 150: Result detection limit: flash powder test 8

Appendix X: validation results for specificity/selectivity: color detection

This appendix presents the validation results of the color detection tests performed with a variety of torch compositions. The following figures show the PEMAD analysis results for each test replicate (n = 8), illustrating the combustion intensity profiles over time and the corresponding color behavior for both compositions.

X.I selectivity/specificity: red flame color identification

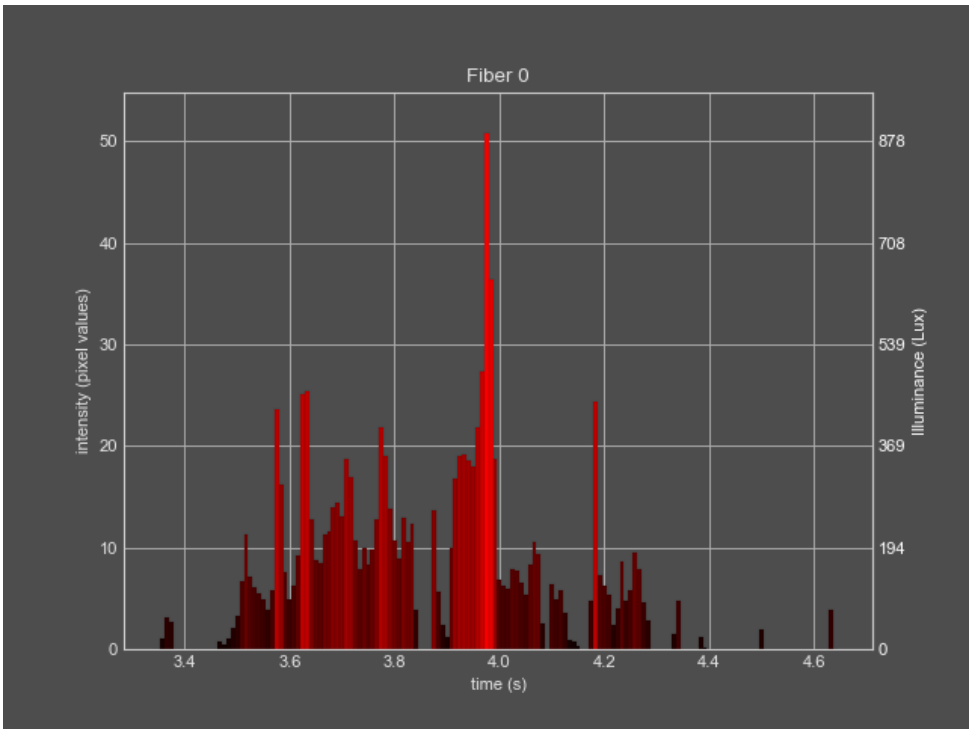


Figure 151: Selectivity/specificity results: red flame color identification test 1

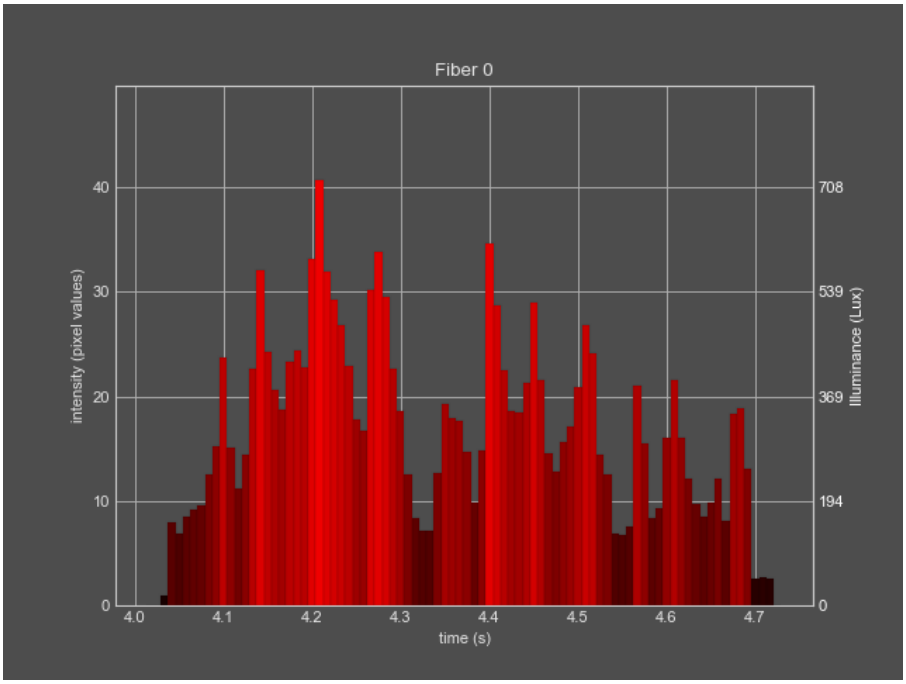


Figure 152: Selectivity/specificity results: red flame color identification test 2

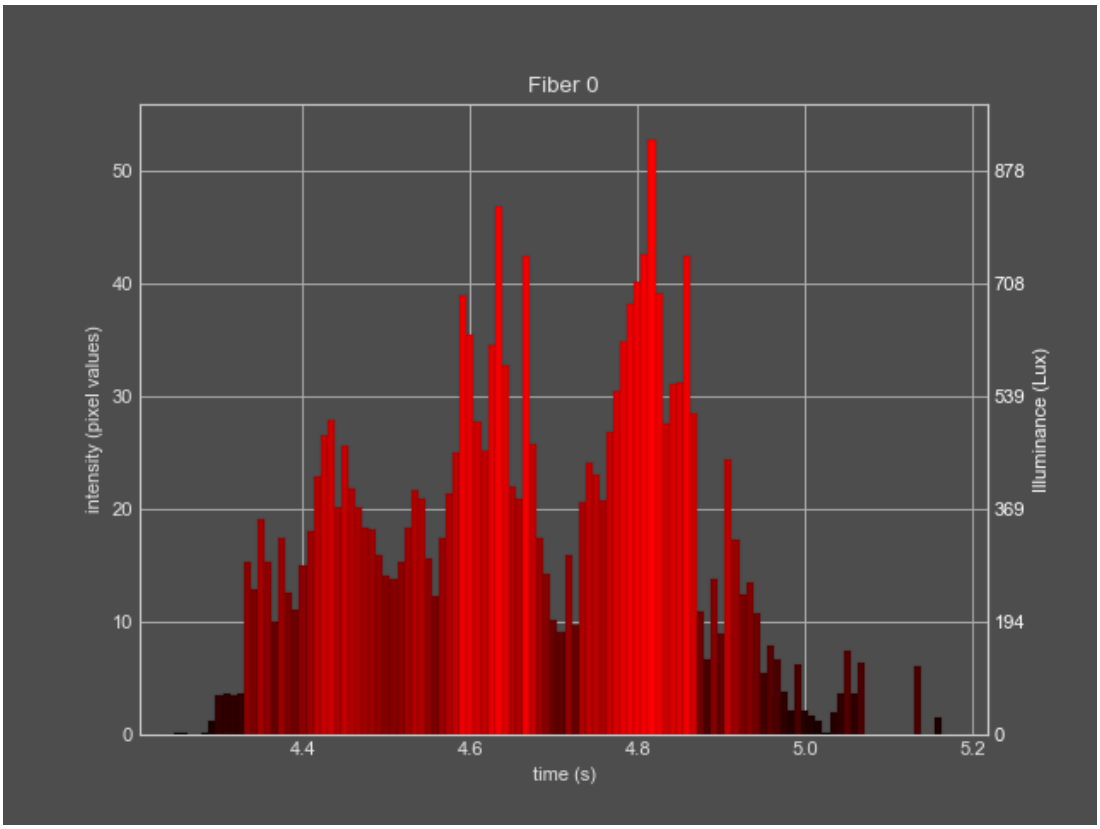


Figure 153: Selectivity/specificity results: red flame color identification test 3

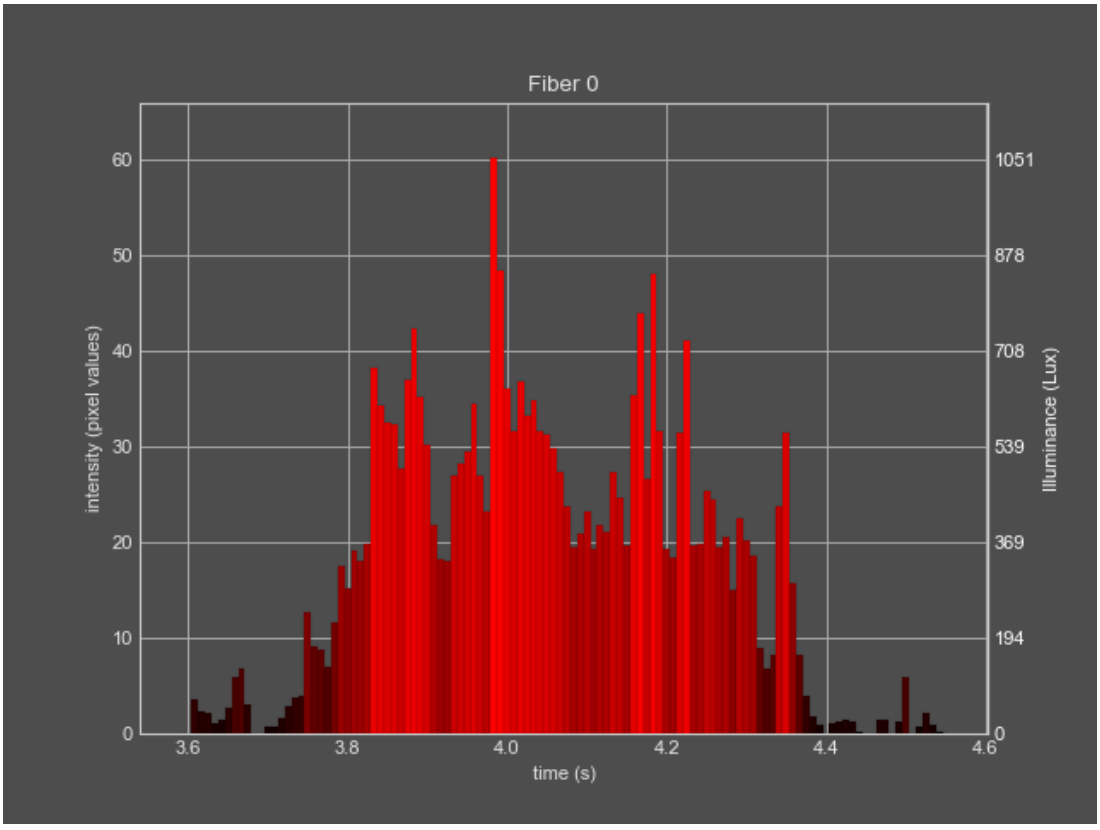


Figure 154: Selectivity/specificity results: red flame color identification test 4

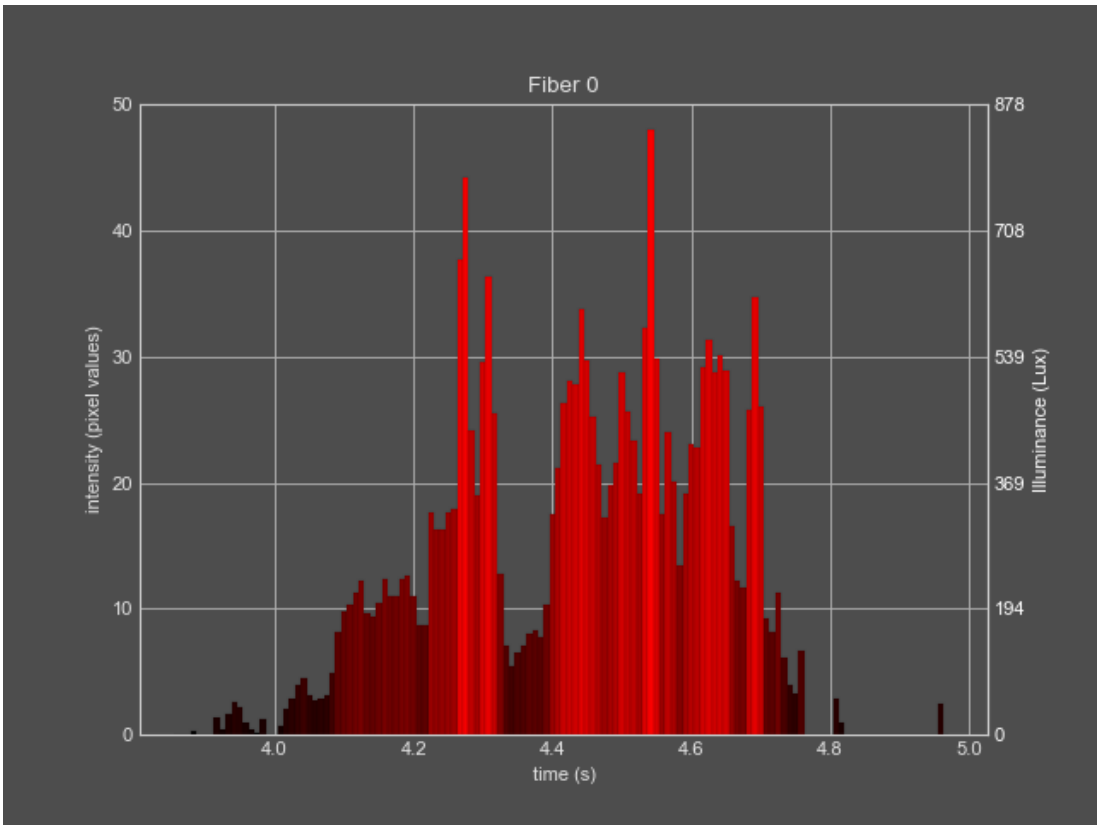


Figure 155: Selectivity/specificity results: red flame color identification test 5

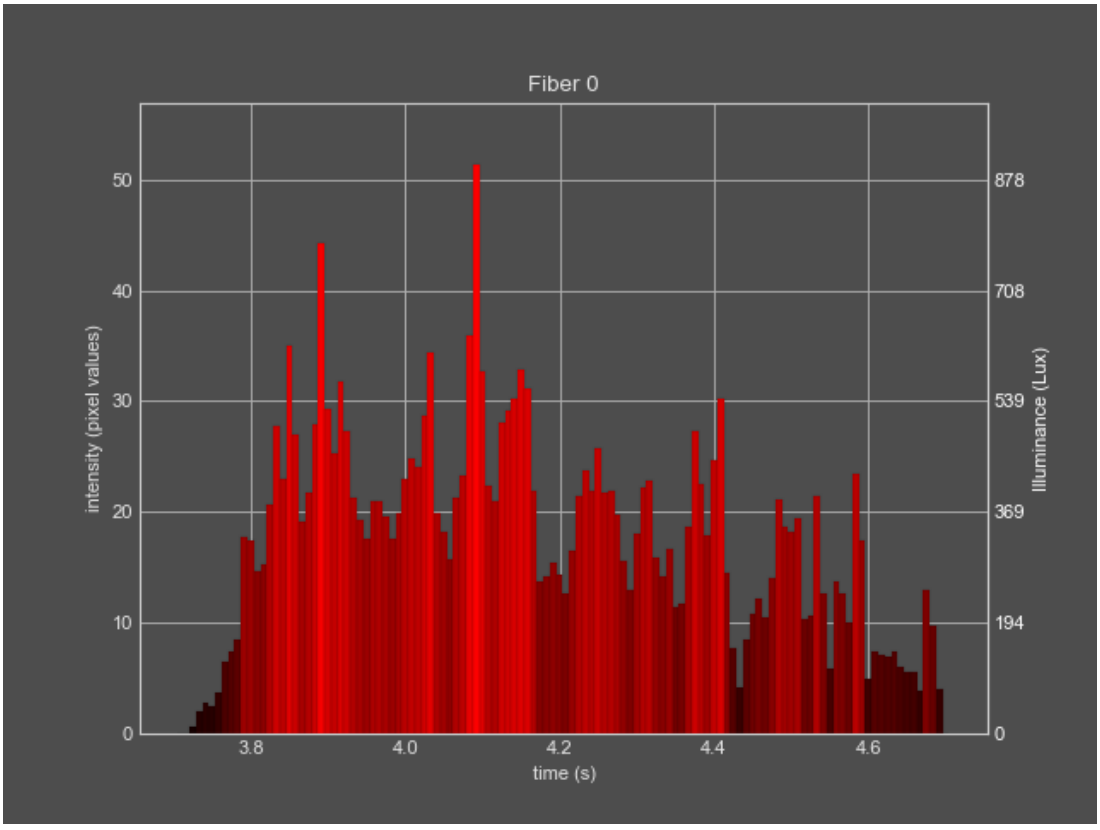


Figure 156: Selectivity/specificity results: red flame color identification test 6

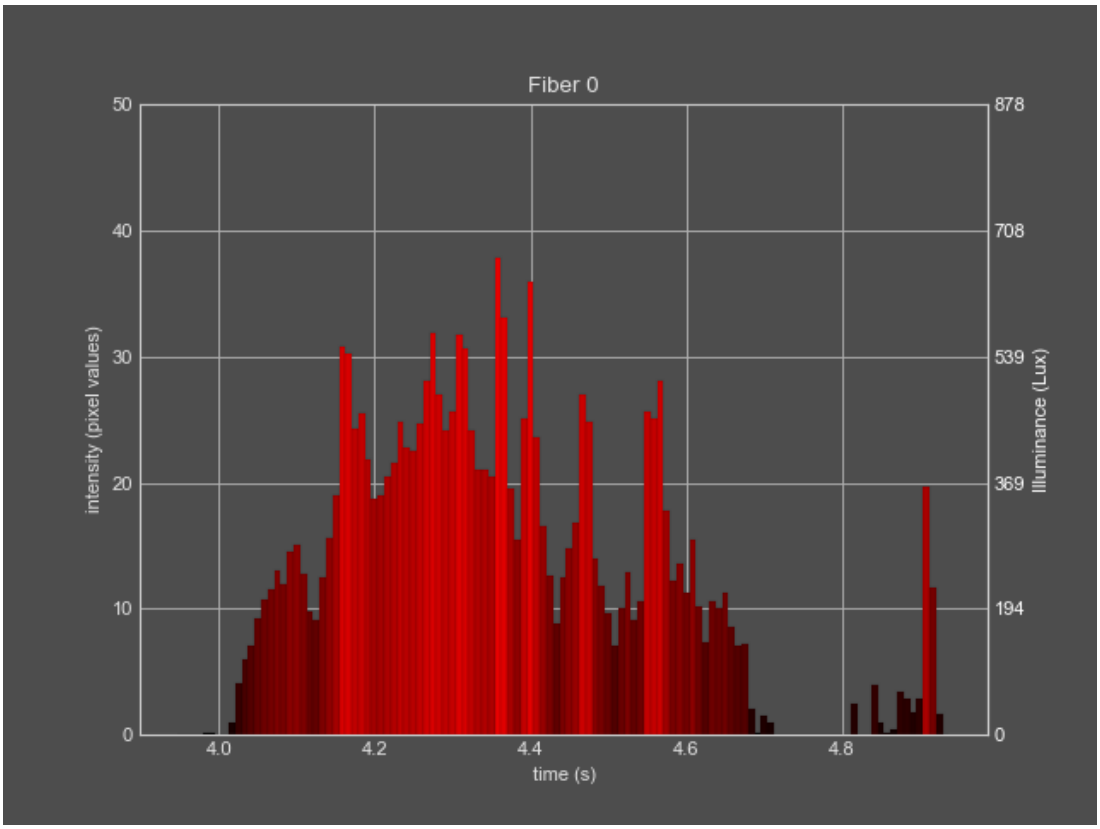


Figure 157: Selectivity/specificity results: red flame color identification test 7

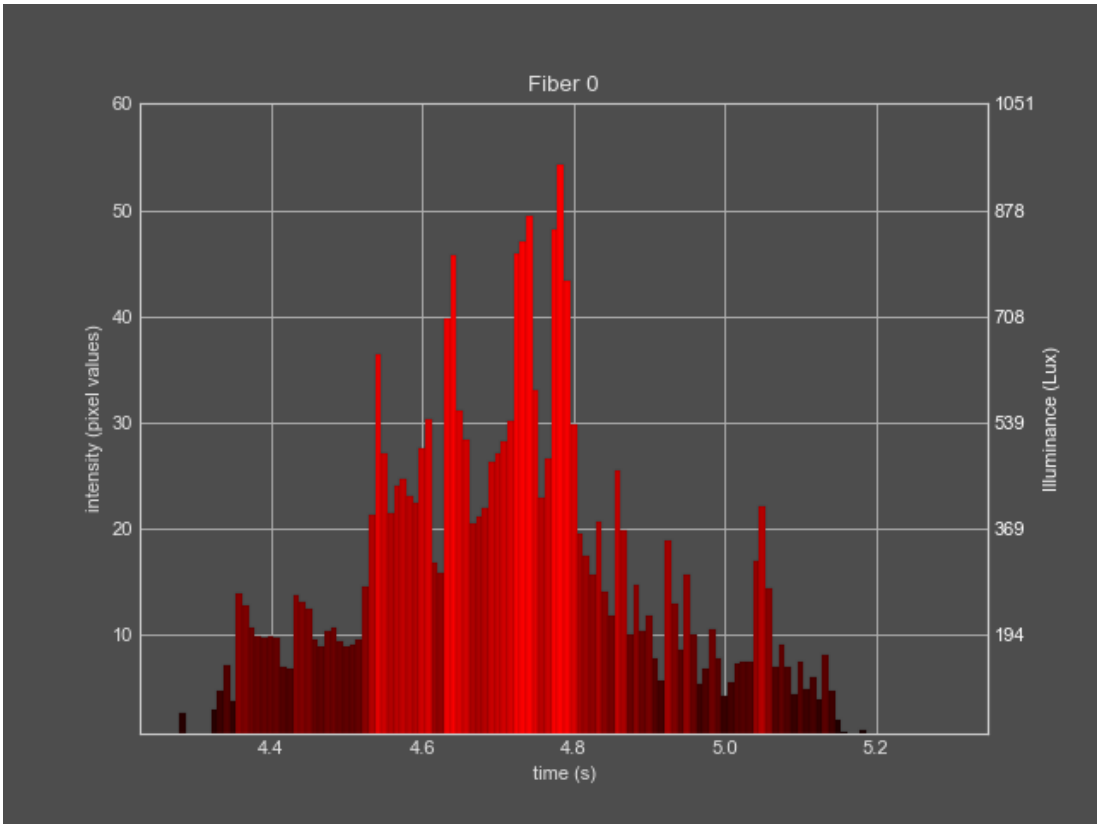


Figure 158: Selectivity/specificity results: red flame color identification test 8

X.II selectivity/specificity: green flame

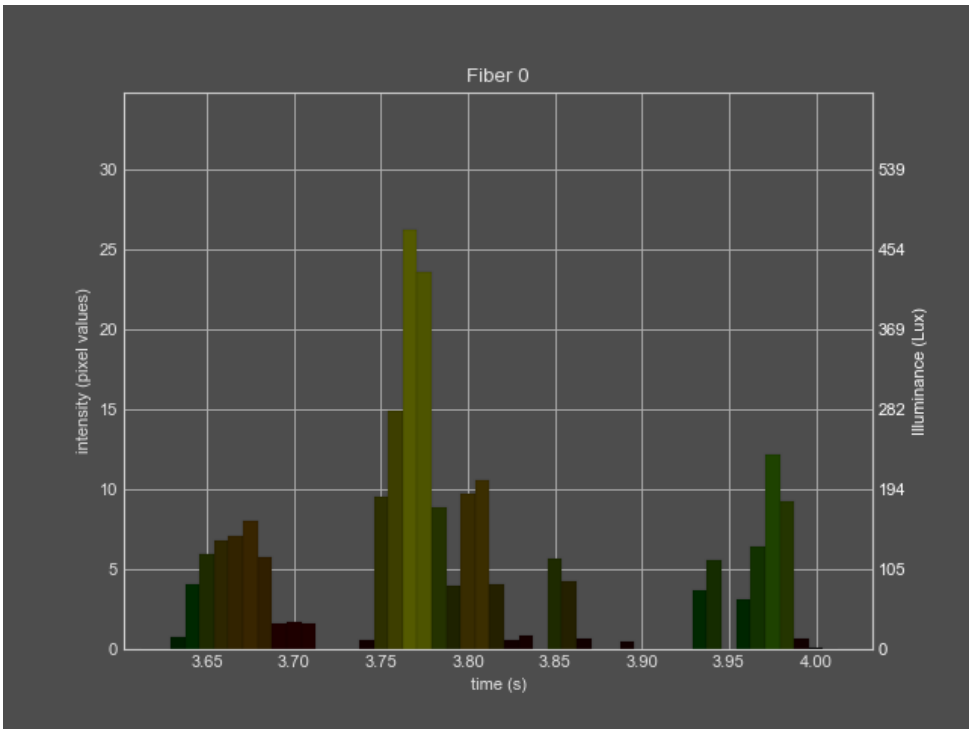


Figure 159: Selectivity/specificity results: green flame color identification test 1

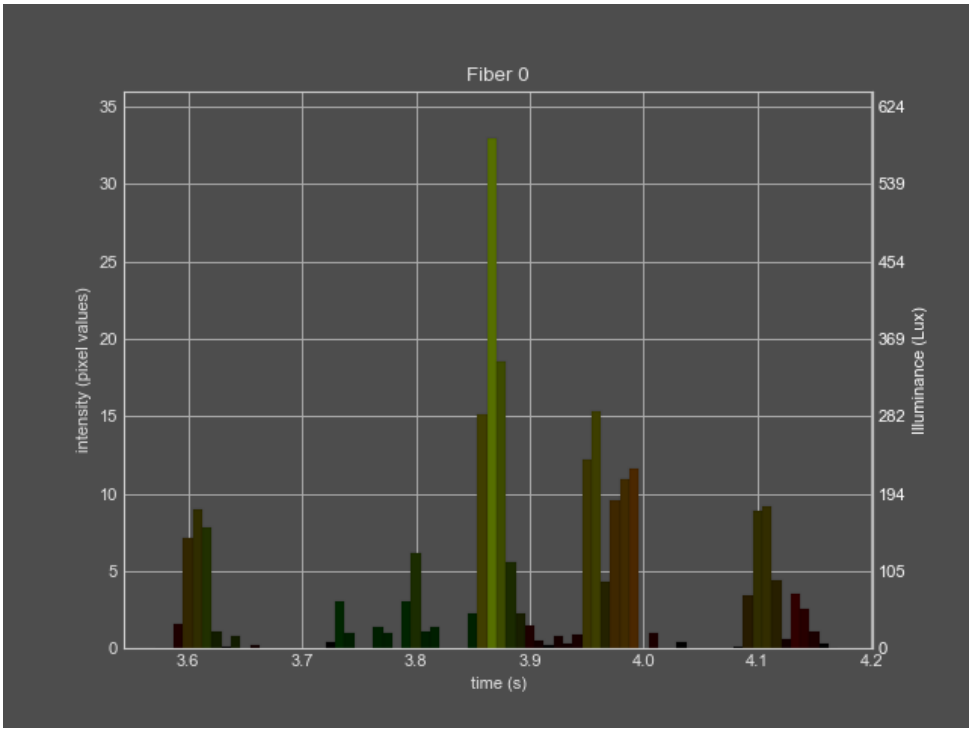


Figure 160: Selectivity/specificity results: green flame color identification test 2

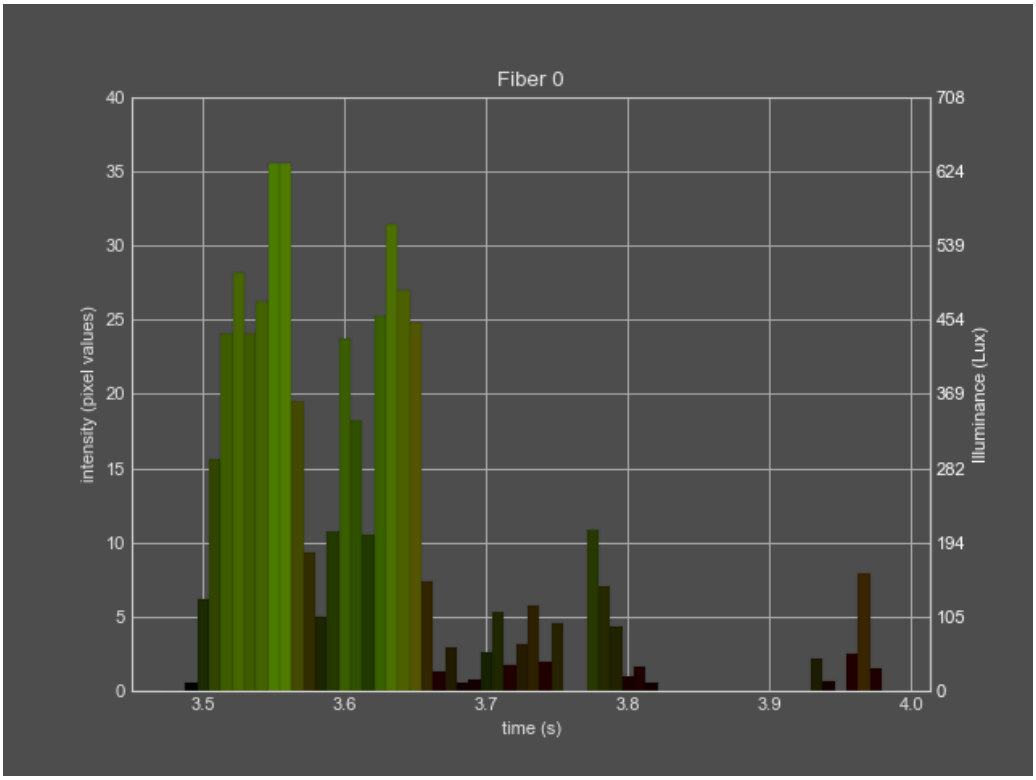


Figure 161: Selectivity/specificity results: green flame color identification test 3

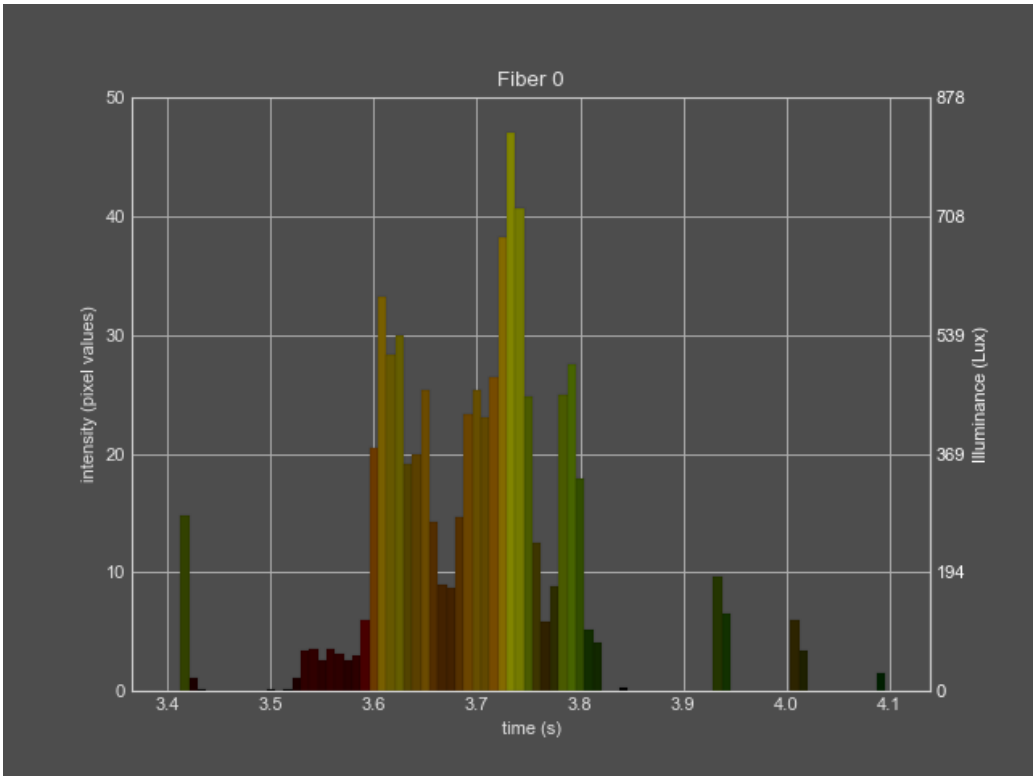


Figure 162: Selectivity/specificity results: green flame color identification test 4

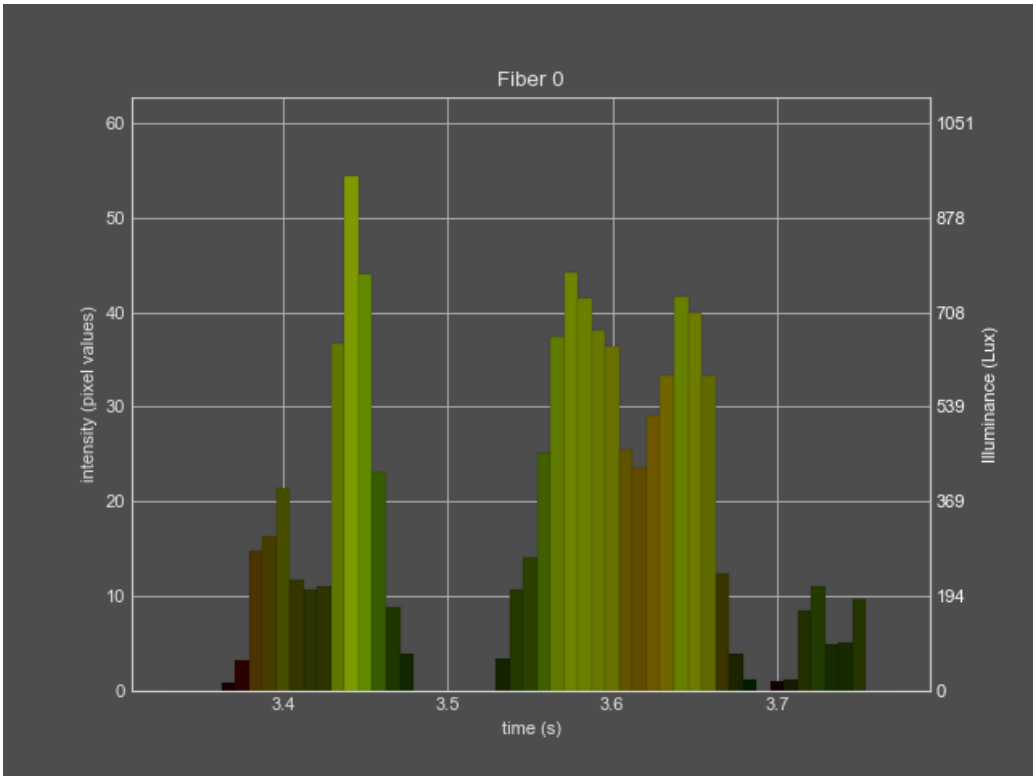


Figure 163: Selectivity/specificity results: green flame color identification test 5

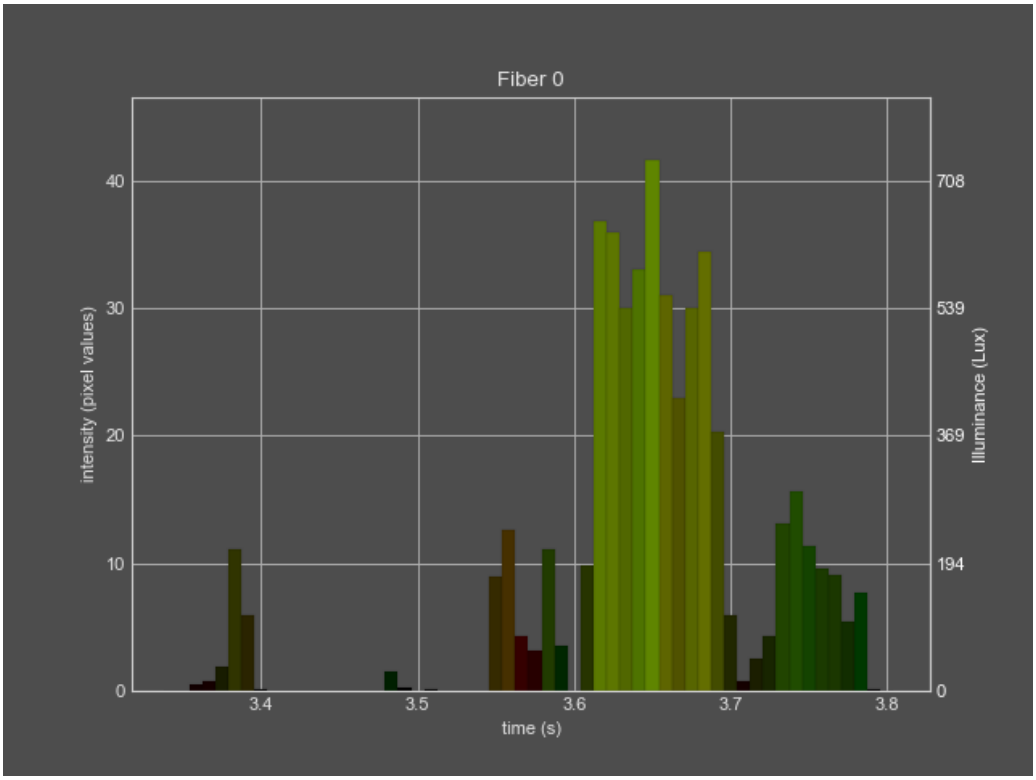


Figure 164: Selectivity/specificity results: green flame color identification test 6

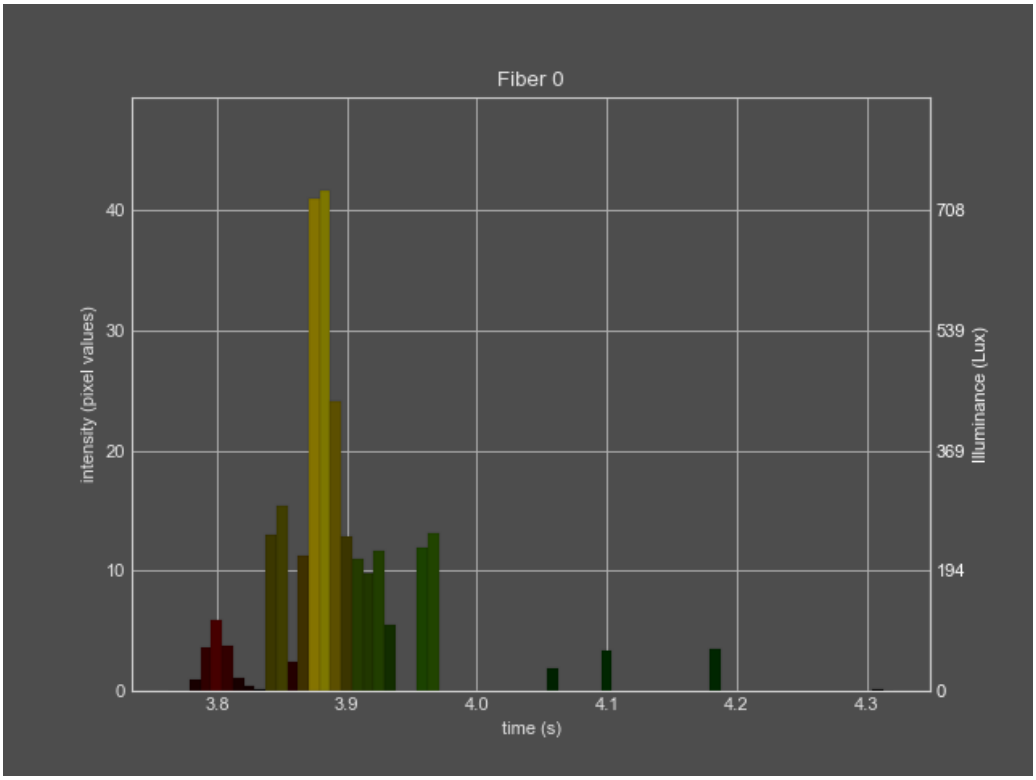


Figure 165: Selectivity/specificity results: green flame color identification test 7

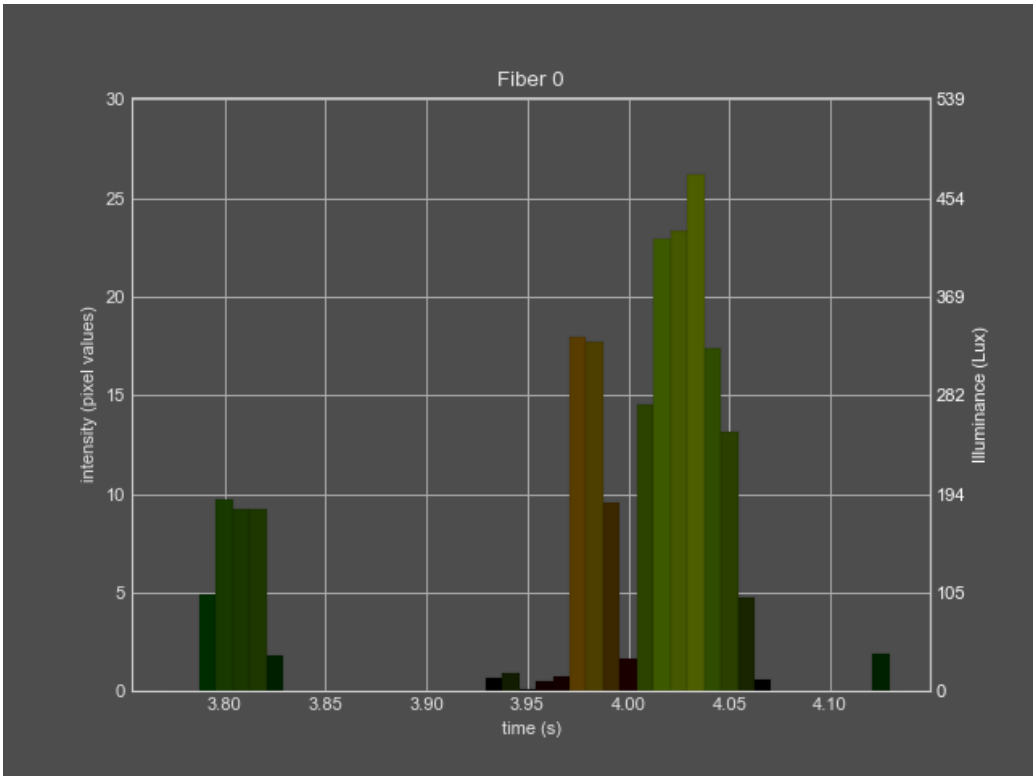


Figure 166: Selectivity/specificity results: green flame color identification test 8

X.III selectivity/specificity: blue flame

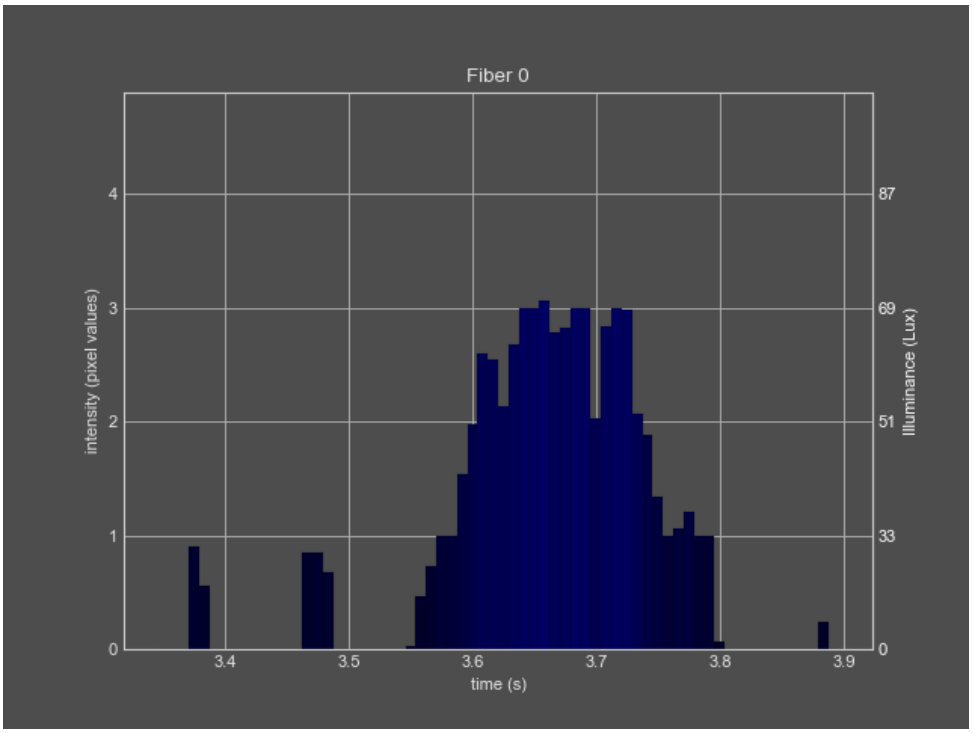


Figure 167: Selectivity/specificity results: blue flame color identification test 1

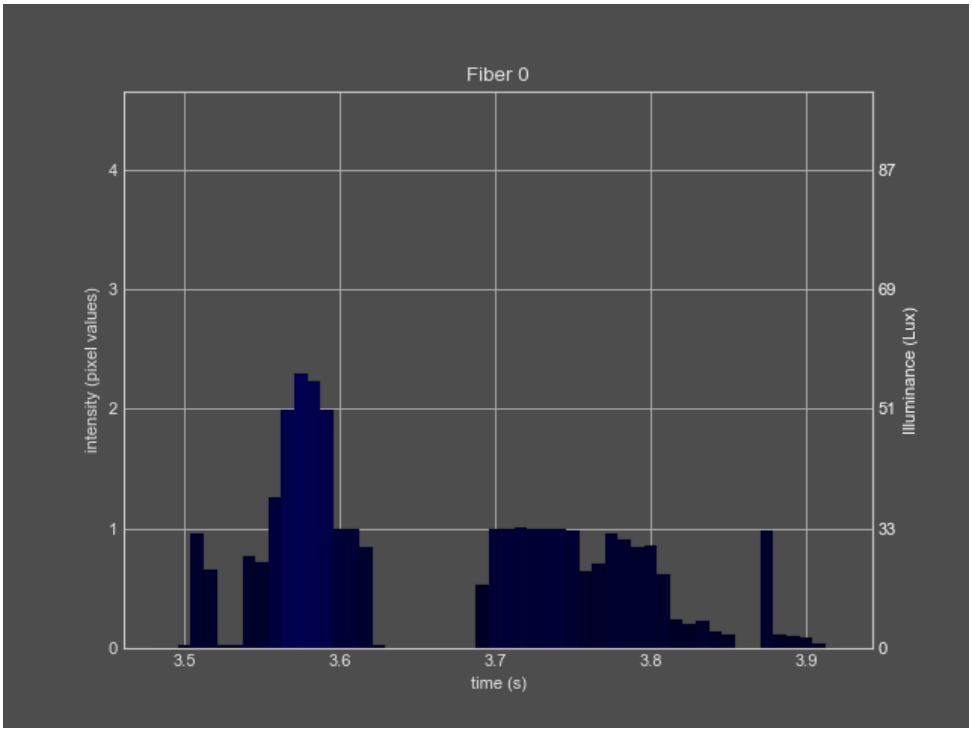


Figure 168: Selectivity/specificity results: blue flame color identification test 2

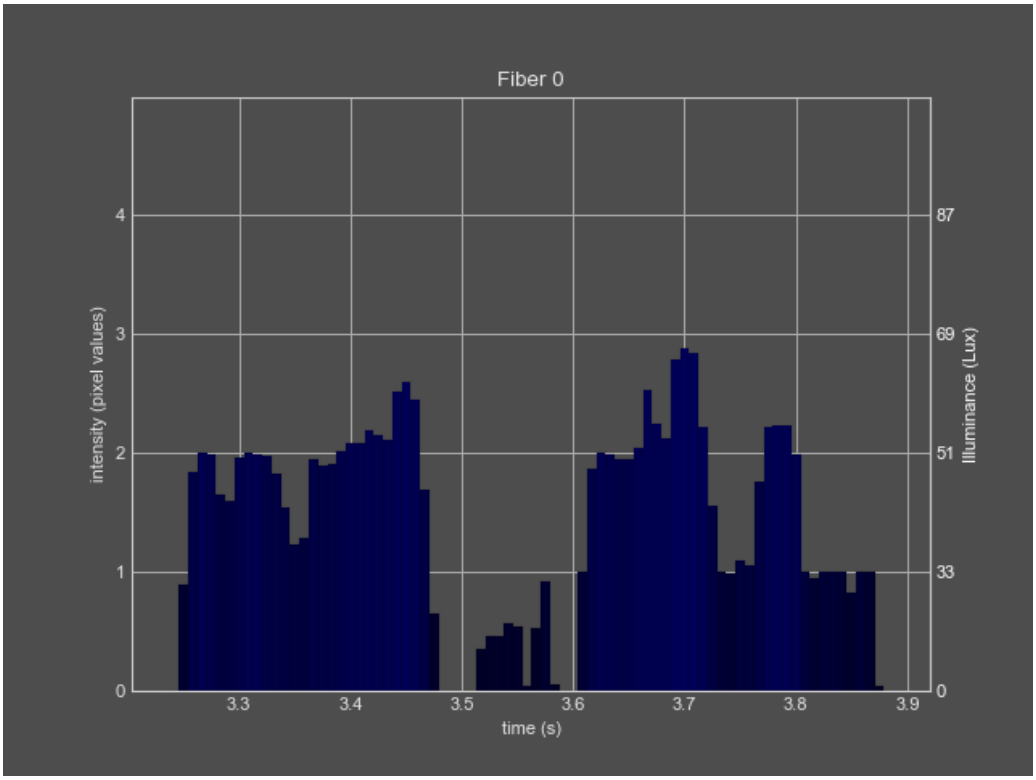


Figure 169: Selectivity/specificity results: blue flame color identification test 3

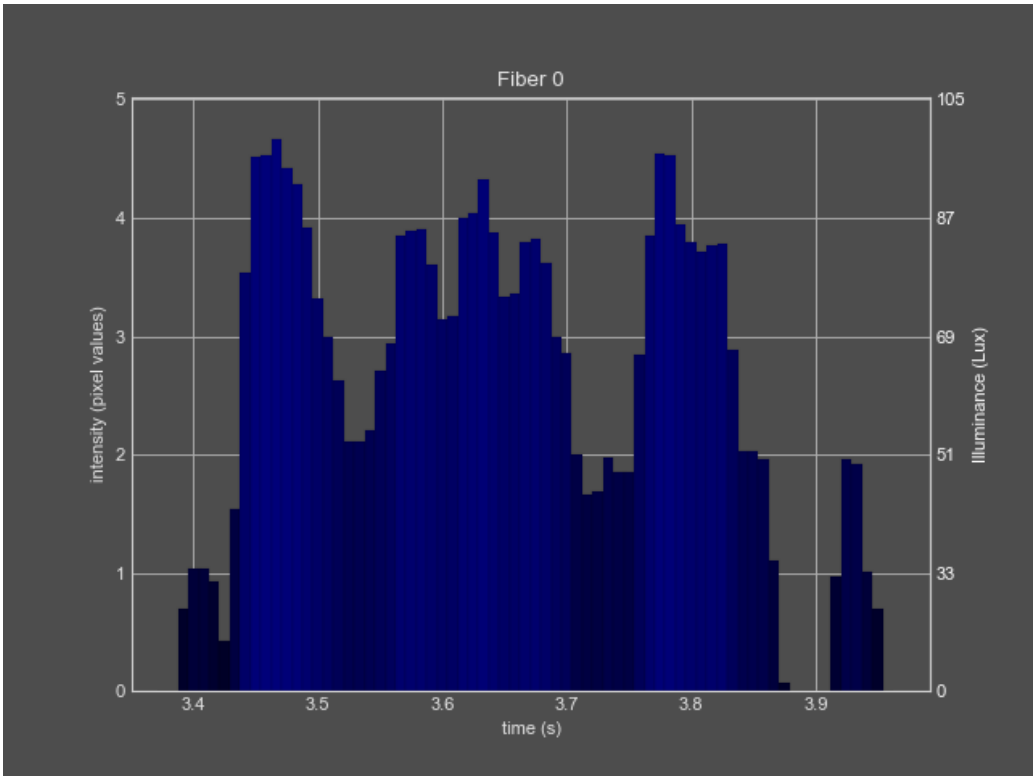


Figure 170: Selectivity/specificity results: blue flame color identification test 4

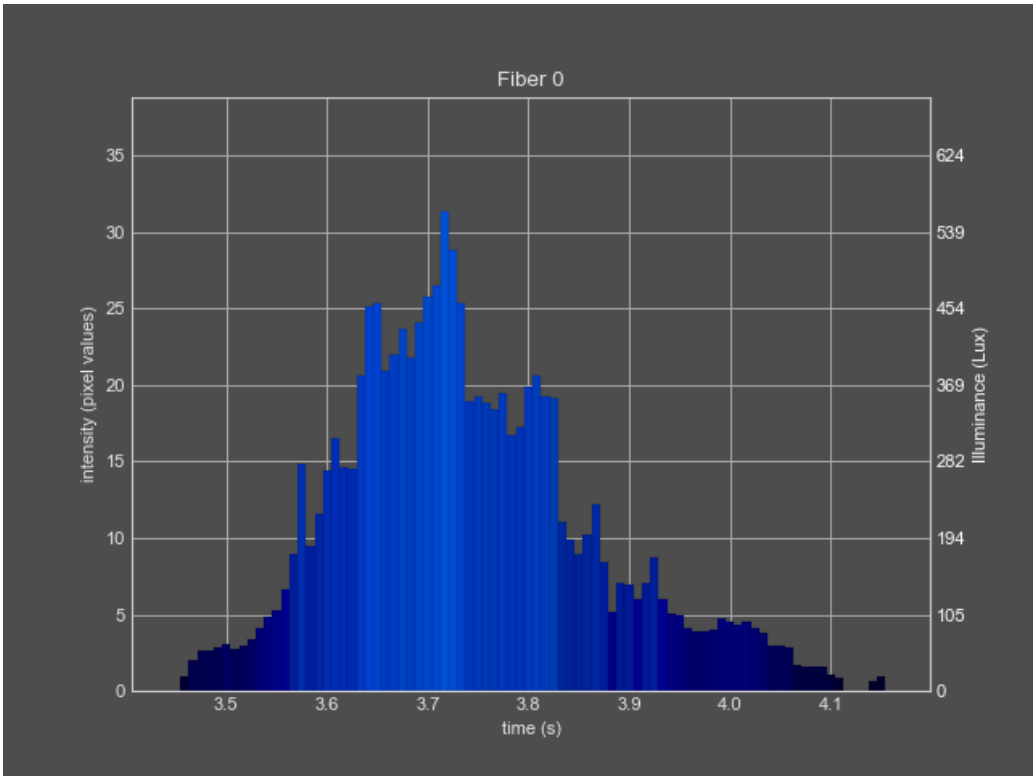


Figure 171: Selectivity/specificity results: blue flame color identification test 5

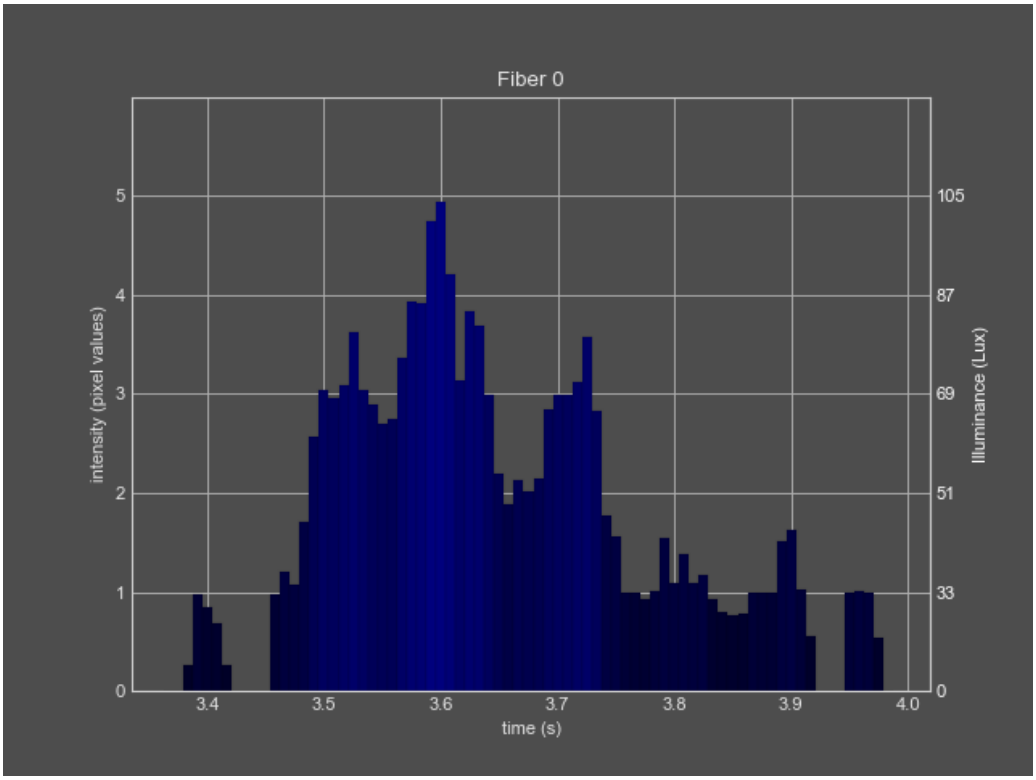


Figure 172: Selectivity/specificity results: blue flame color identification test 6

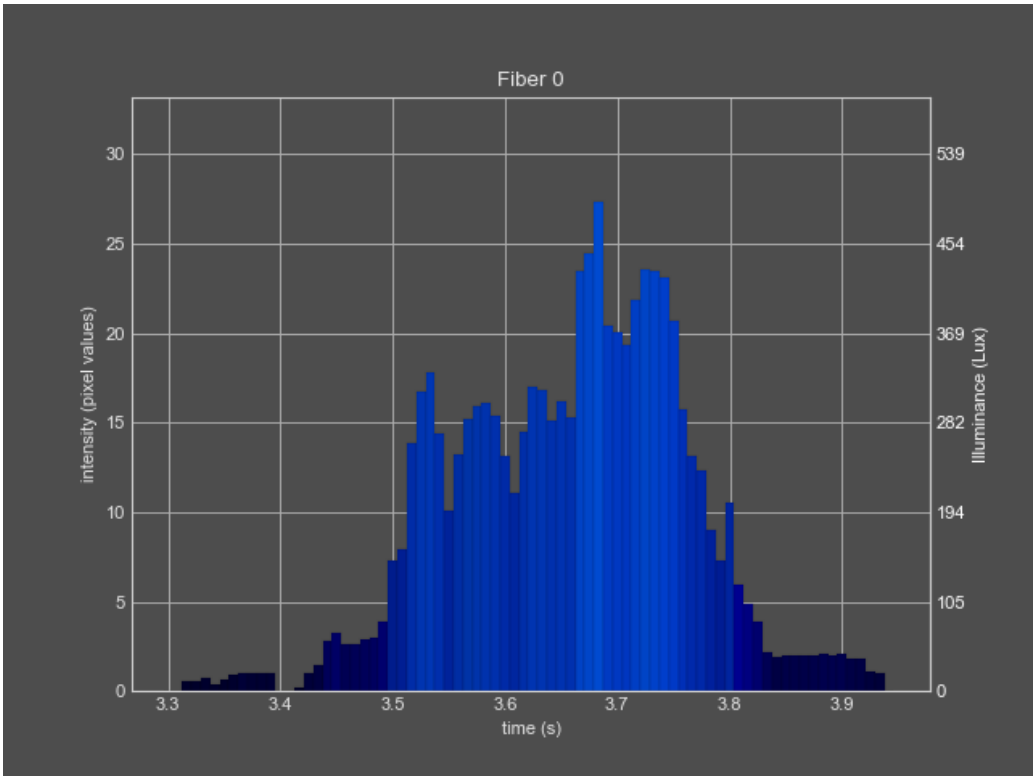


Figure 173: Selectivity/specificity results: blue flame color identification test 7

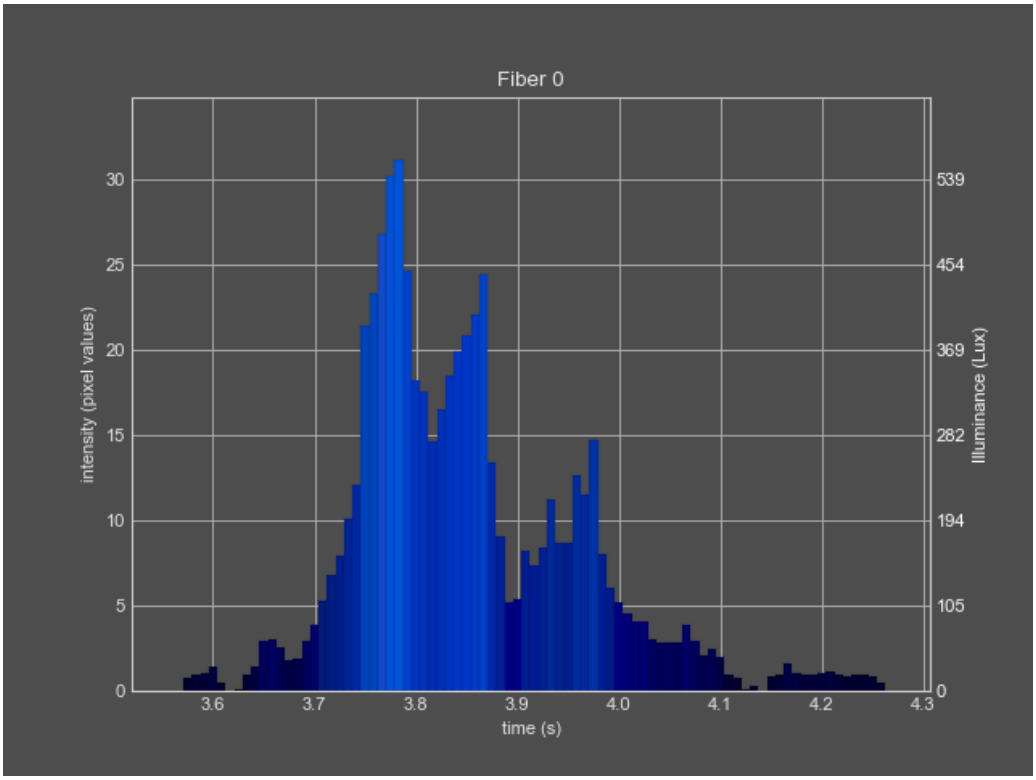


Figure 174: Selectivity/specificity results: blue flame color identification test 8

X.IV selectivity/specificity: purple flame

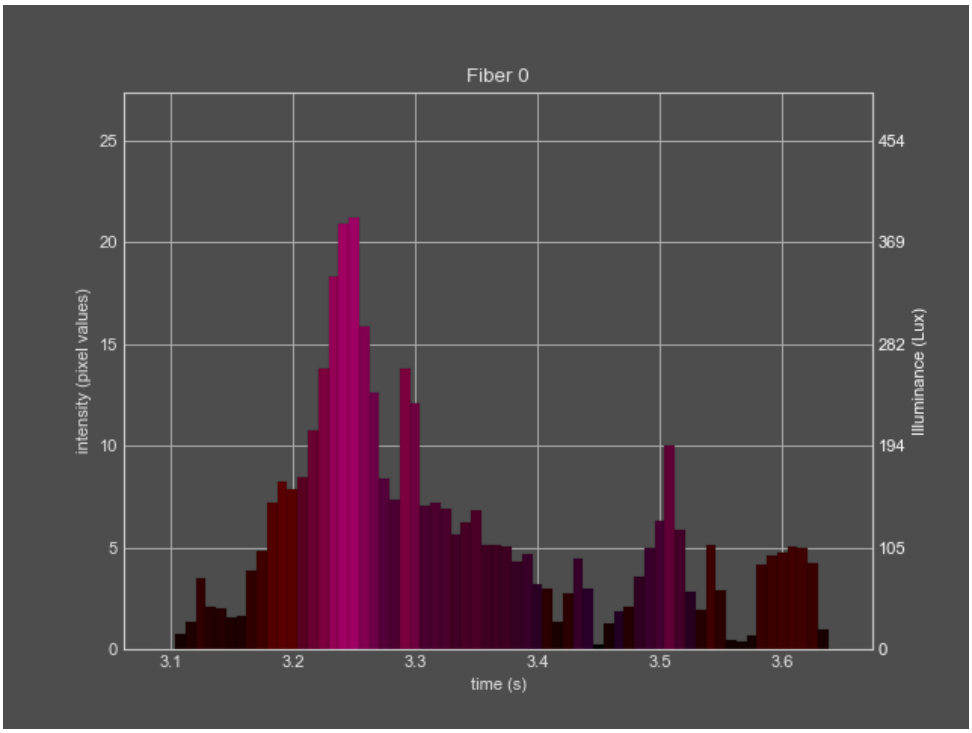


Figure 175: Selectivity/specificity results: purple flame color identification test 1

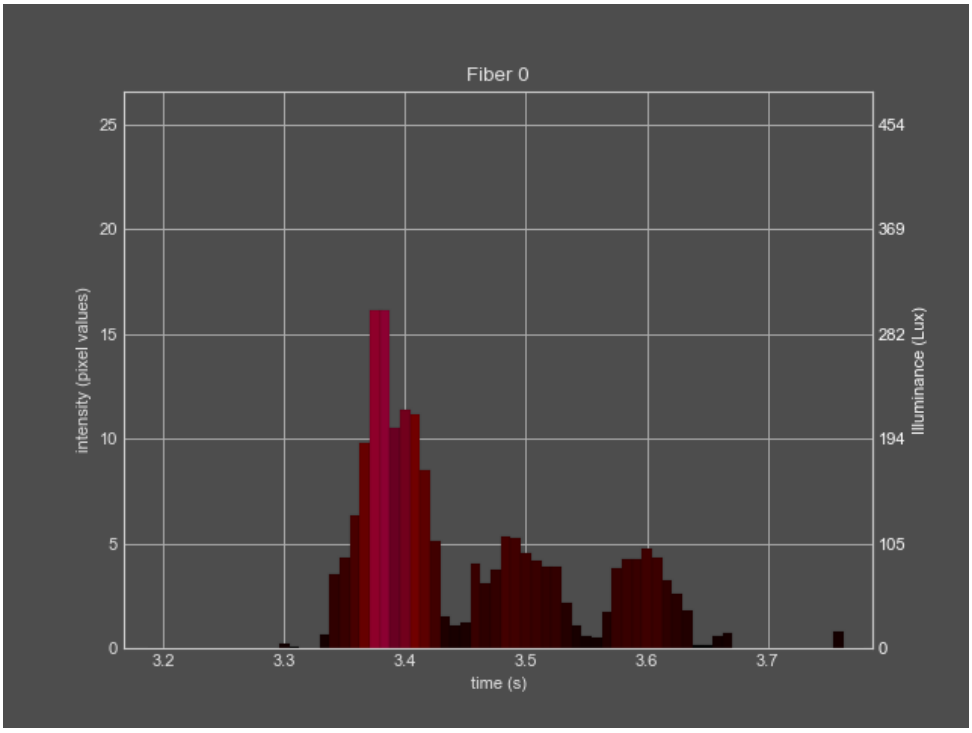


Figure 176: Selectivity/specificity results: purple flame color identification test 2

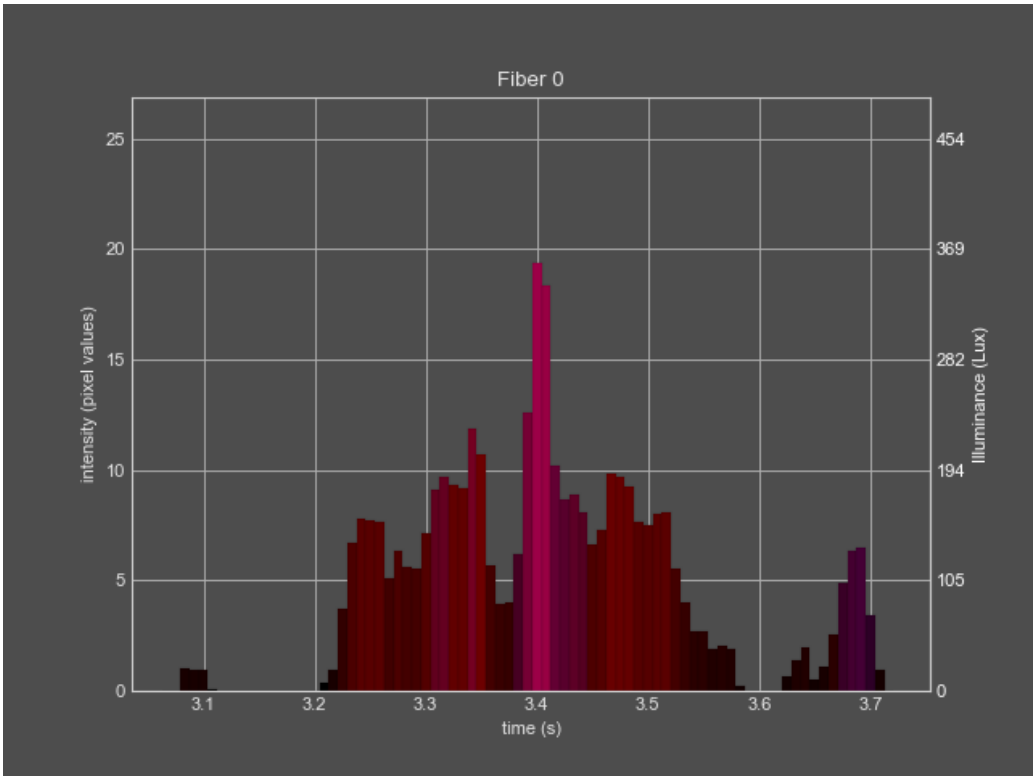


Figure 177: Selectivity/specificity results: purple flame color identification test 3

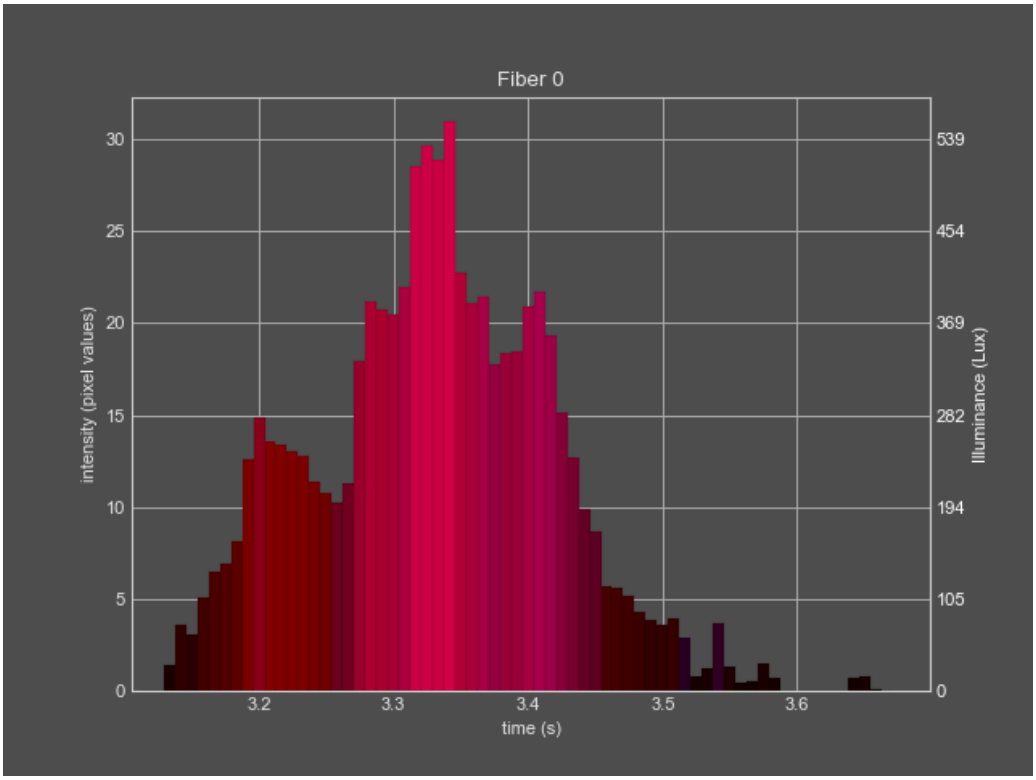


Figure 178: Selectivity/specificity results: purple flame color identification test 4

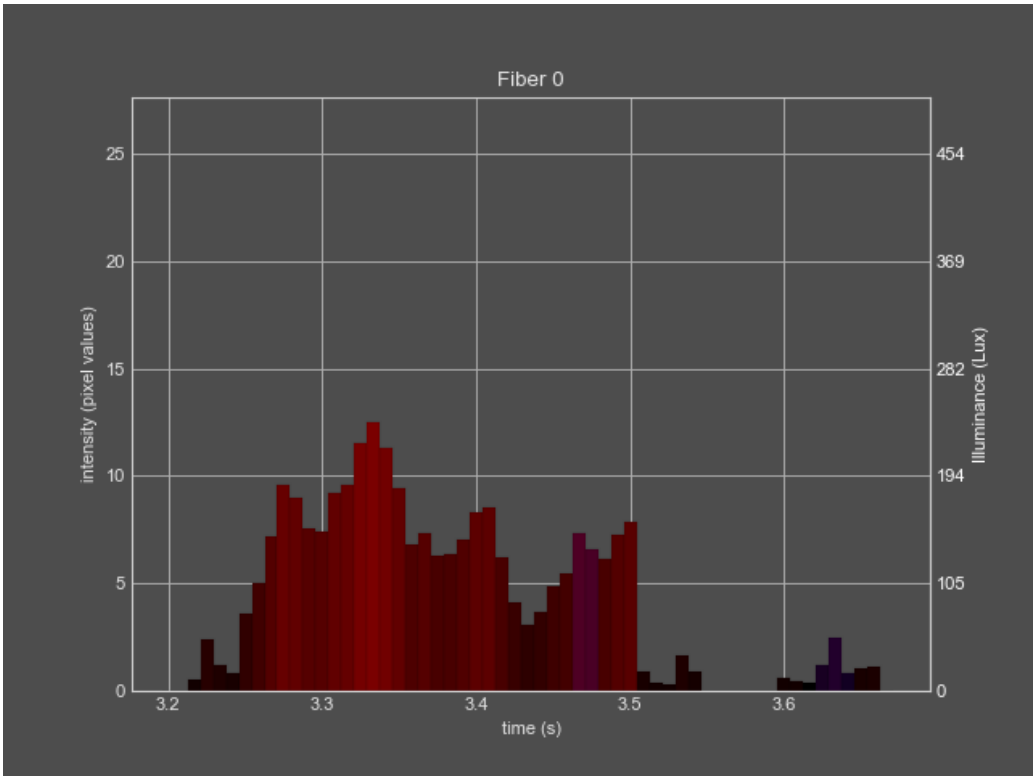


Figure 179: Selectivity/specificity results: purple flame color identification test 5

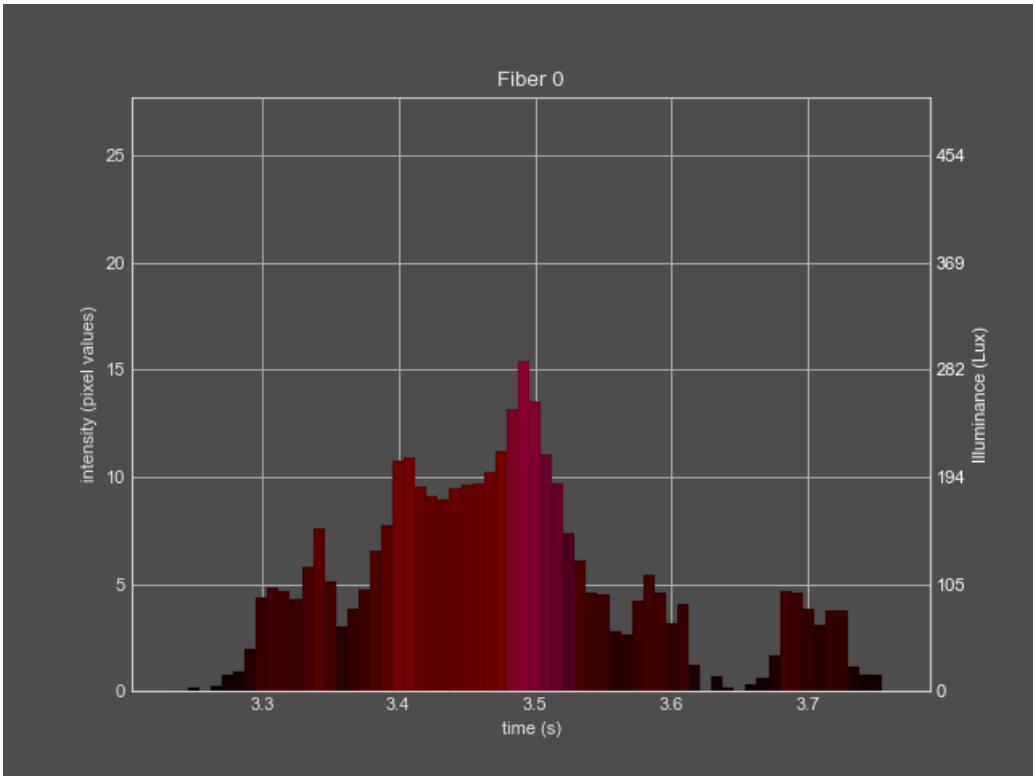


Figure 180: Selectivity/specificity results: purple flame color identification test 6

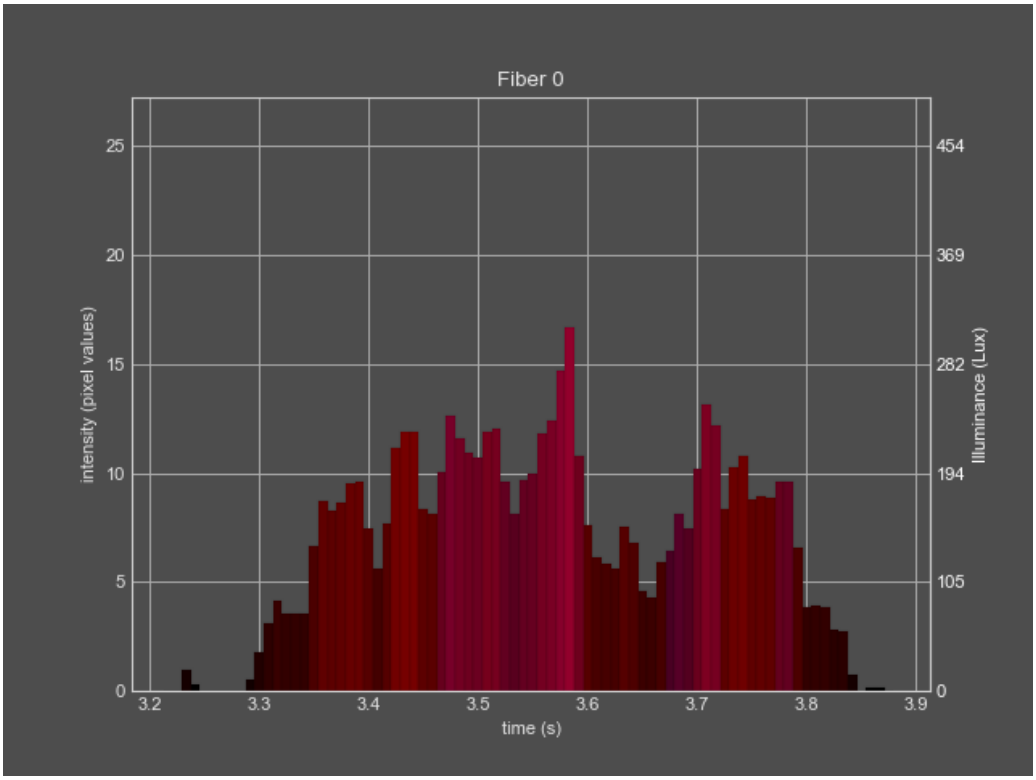


Figure 181: Selectivity/specificity results: purple flame color identification test 7

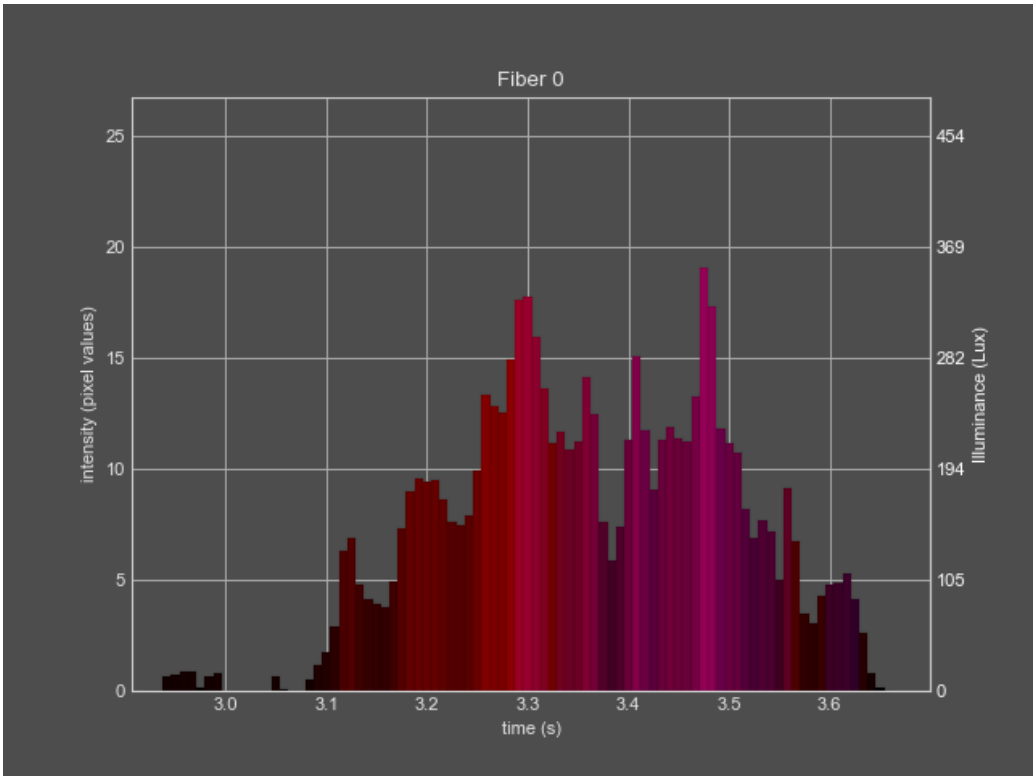


Figure 182: Selectivity/specificity results: purple flame color identification test 8

X.V selectivity/specificity: orange flame

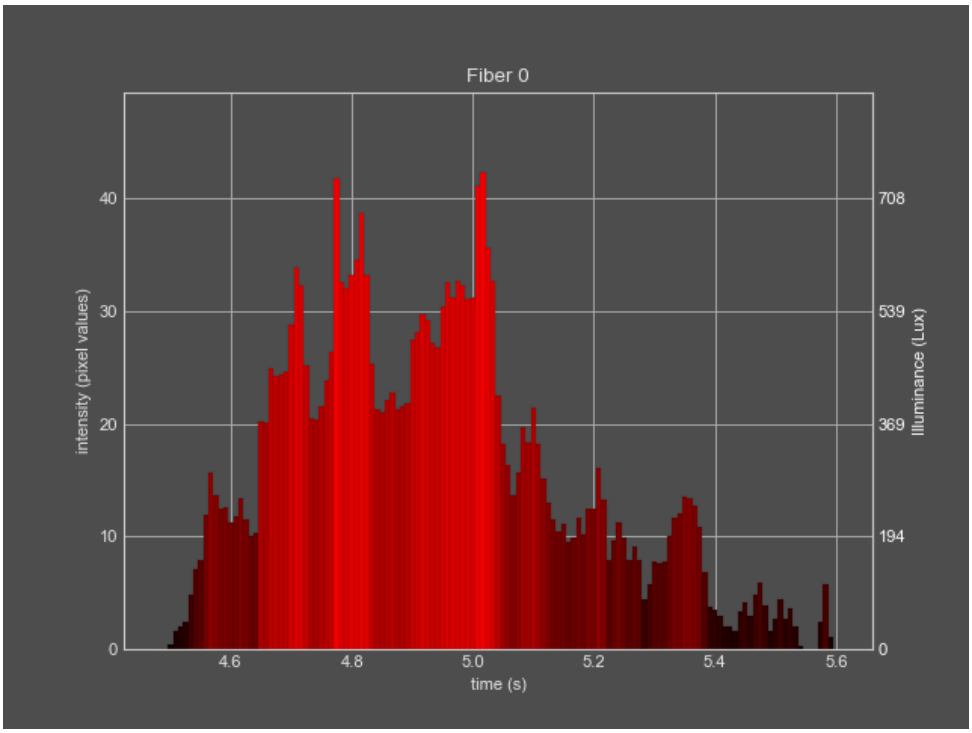


Figure 183: Selectivity/specificity results: orange flame color identification test 1

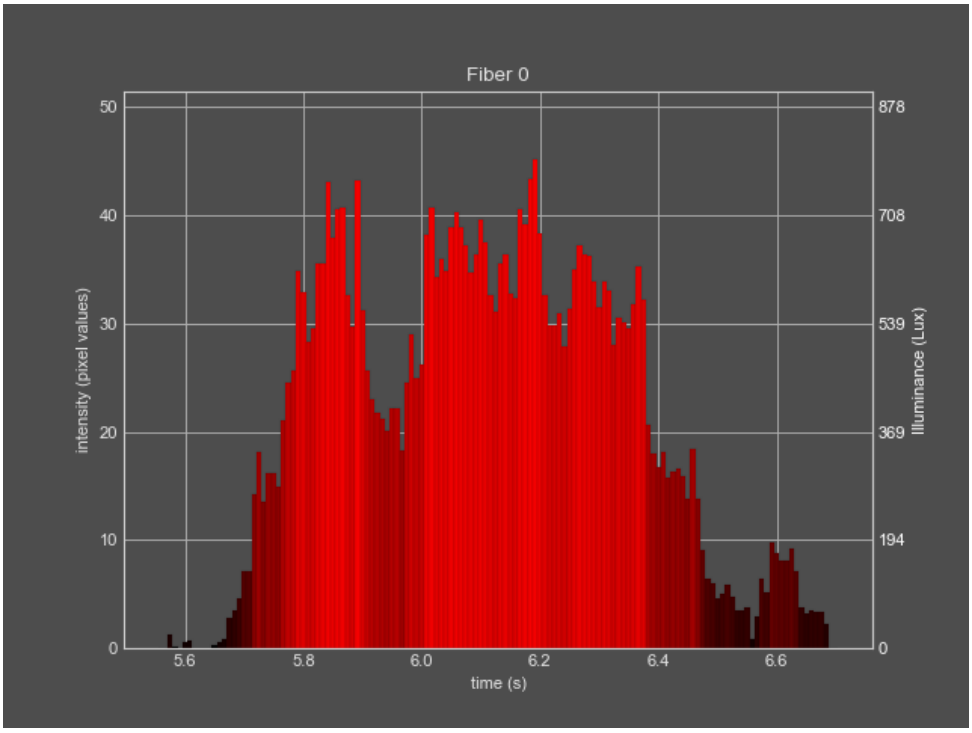


Figure 184: Selectivity/specificity results: orange flame color identification test 2

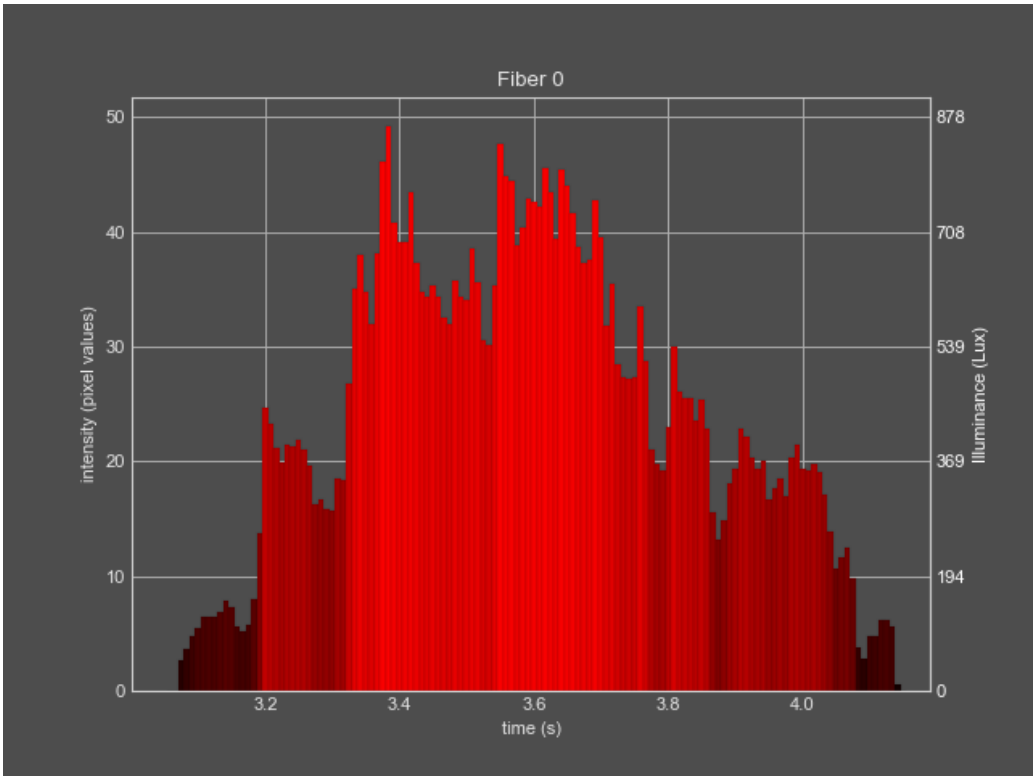


Figure 185: Selectivity/specificity results: orange flame color identification test 3

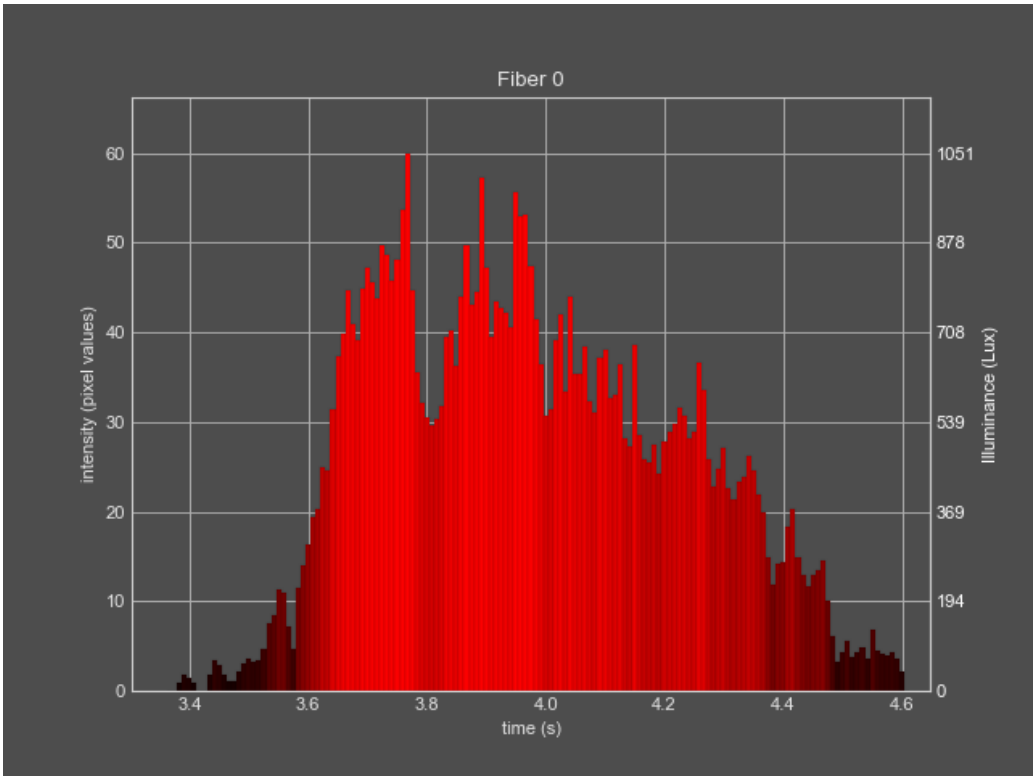


Figure 186: Selectivity/specificity results: orange flame color identification test 4

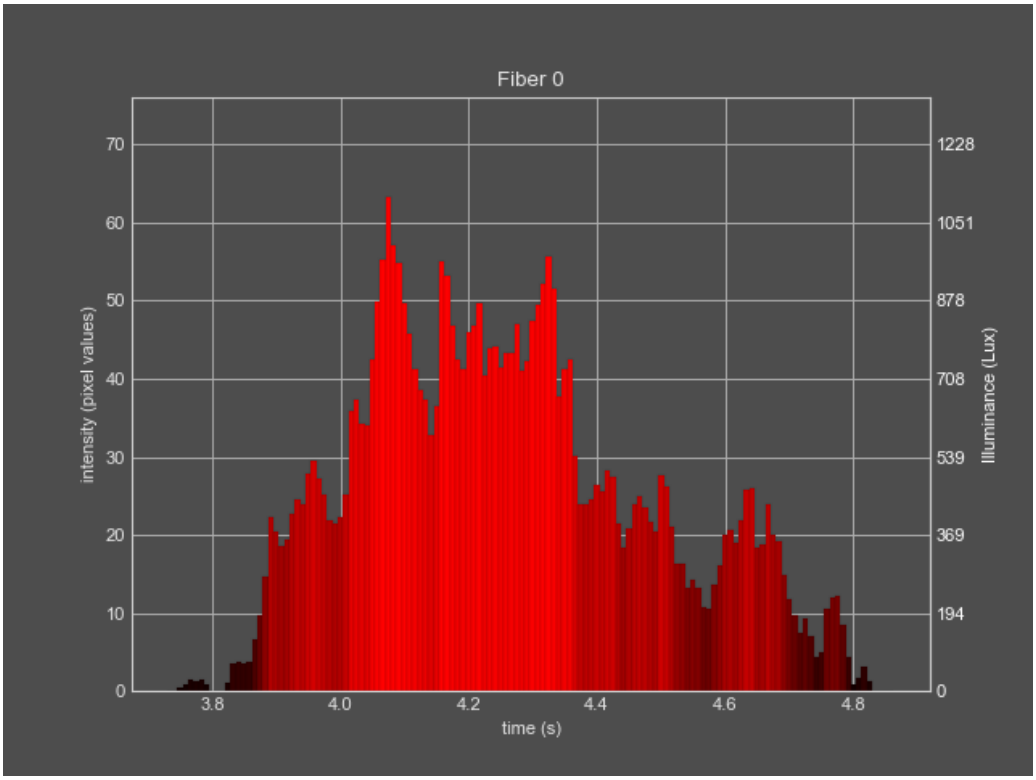


Figure 187: Selectivity/specificity results: orange flame color identification test 5

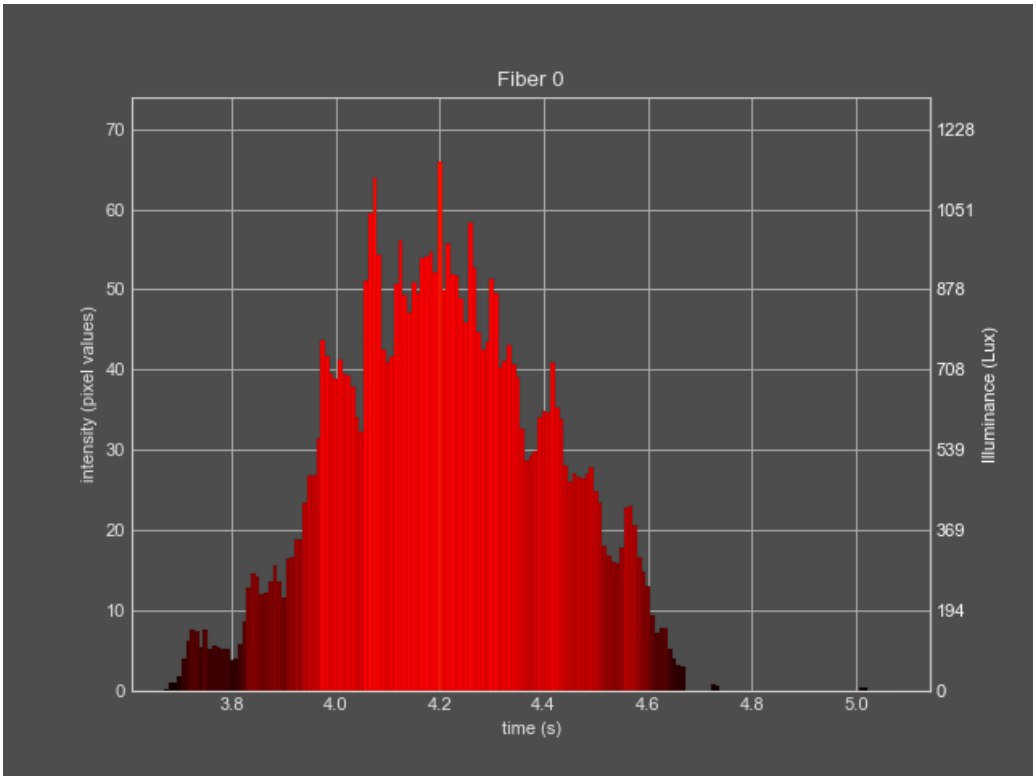


Figure 188: Selectivity/specificity results: orange flame color identification test 6

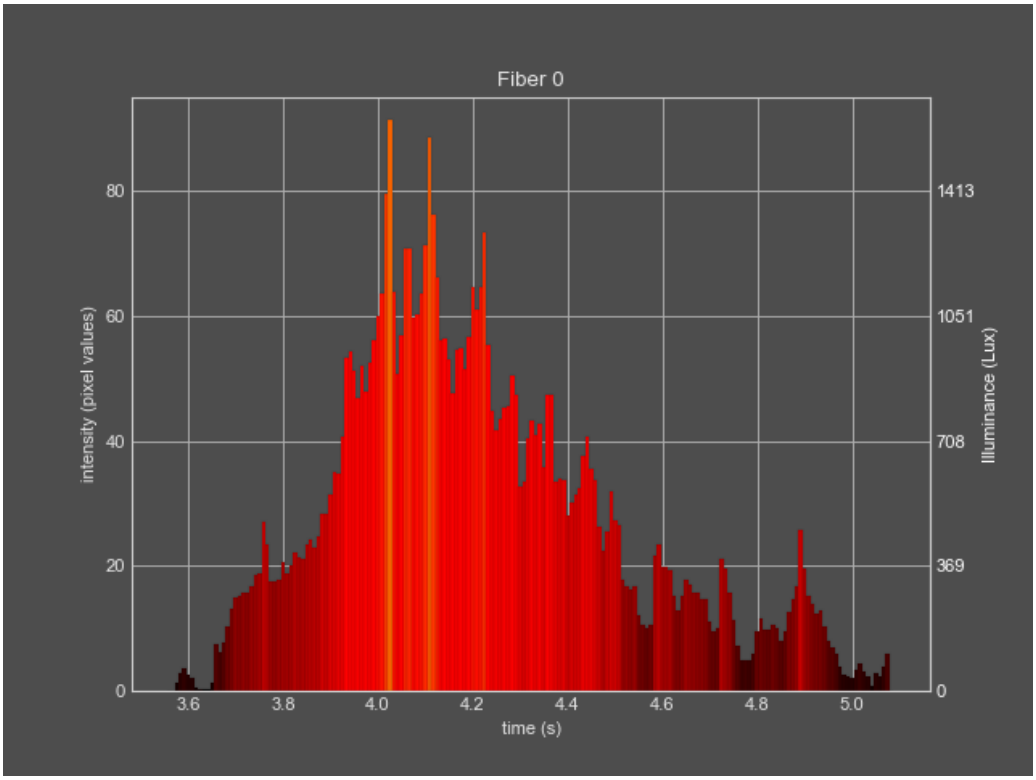


Figure 189: Selectivity/specificity results: orange flame color identification test 7

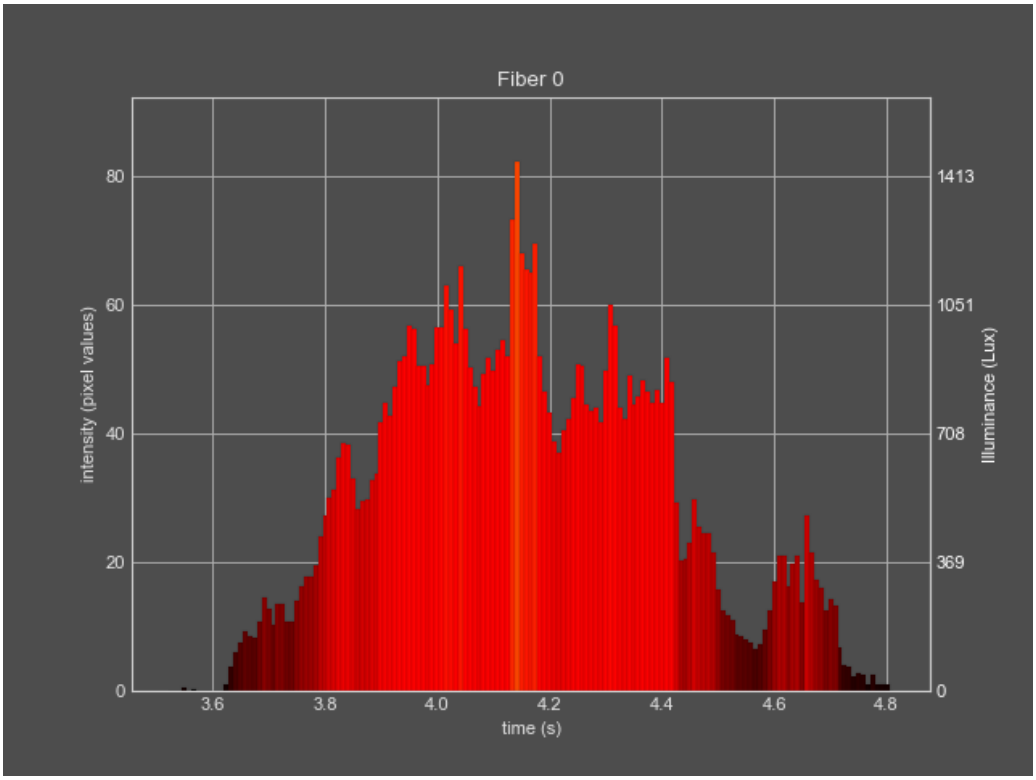


Figure 190: Selectivity/specificity results: orange flame color identification test 8

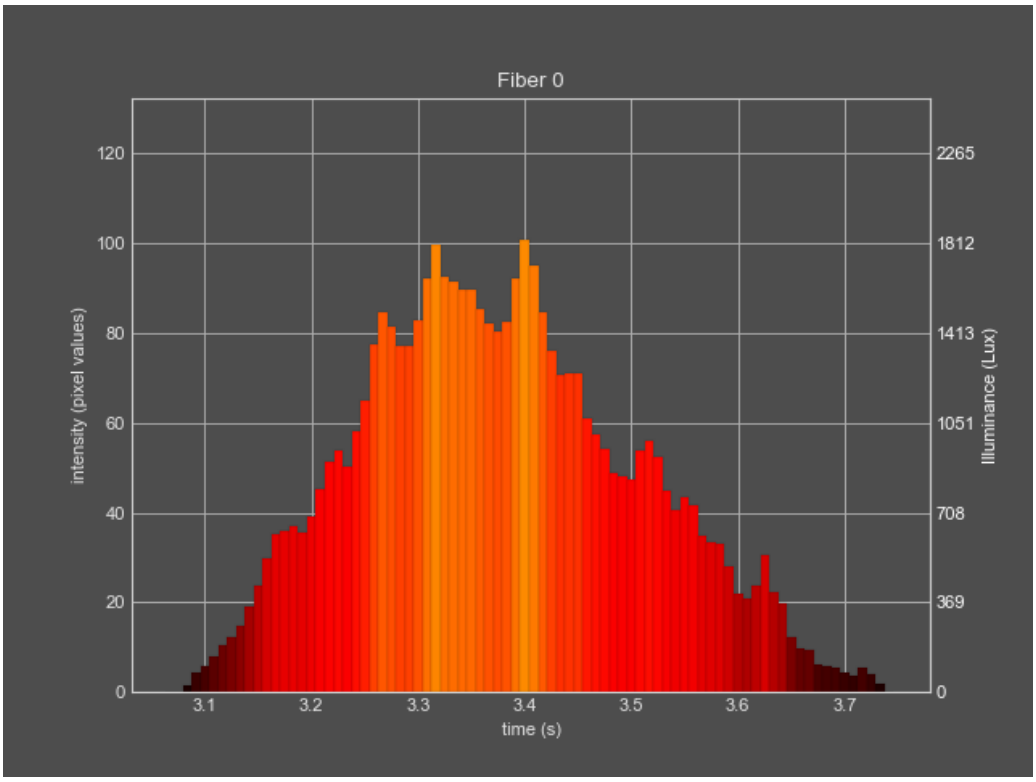


Figure 191: Selectivity/specificity results: orange flame color identification test 9, in combination with flash powder to achieve high intensity

Appendix XI: validation results for specificity/selectivity: oxidizer/fuel ratio variations

This appendix presents the validation results of the combustions of compositions with variation in oxidizer/fuel ratios tests. The following figures show the PEMAD analysis results for each test replicate (n = 8), illustrating the combustion intensity profiles over time and the corresponding color behavior for both compositions.

XI.I Ratio A: potassium perchlorate (40%)/aluminum (60%)

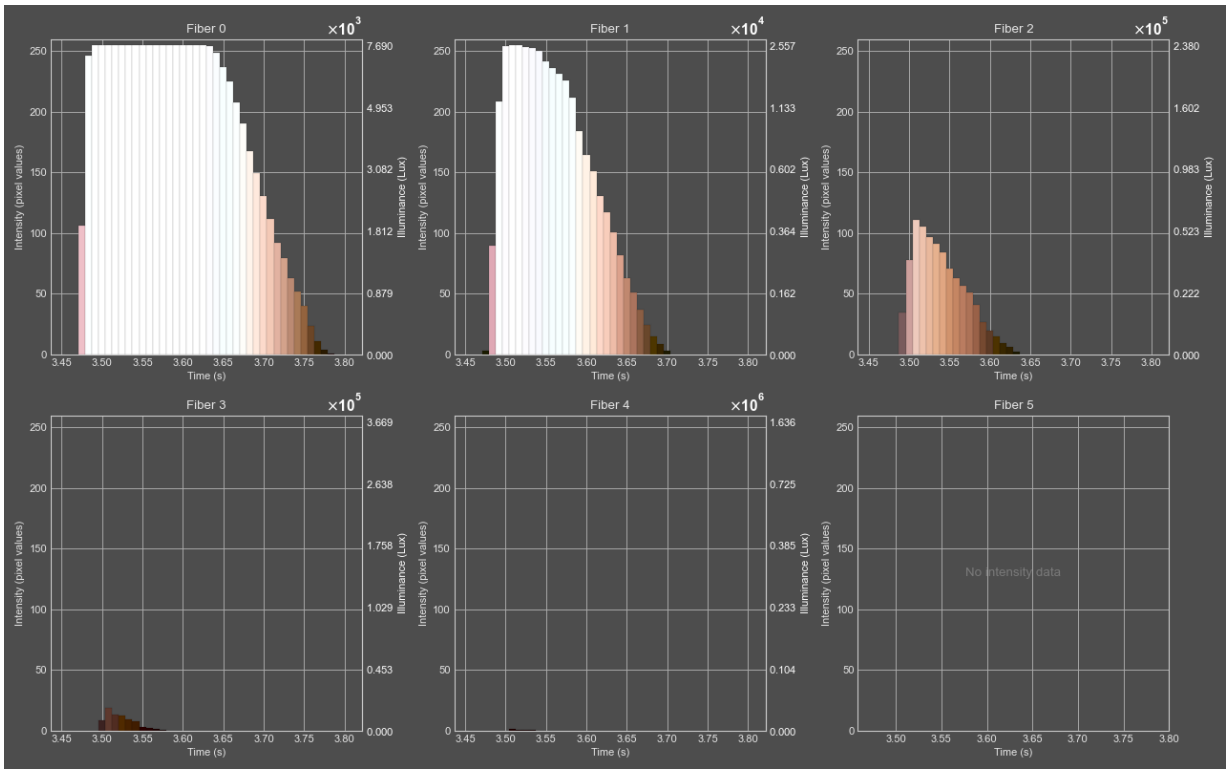


Figure 192: Selectivity/specificity: Ratio (40/60, m/m%) test 1

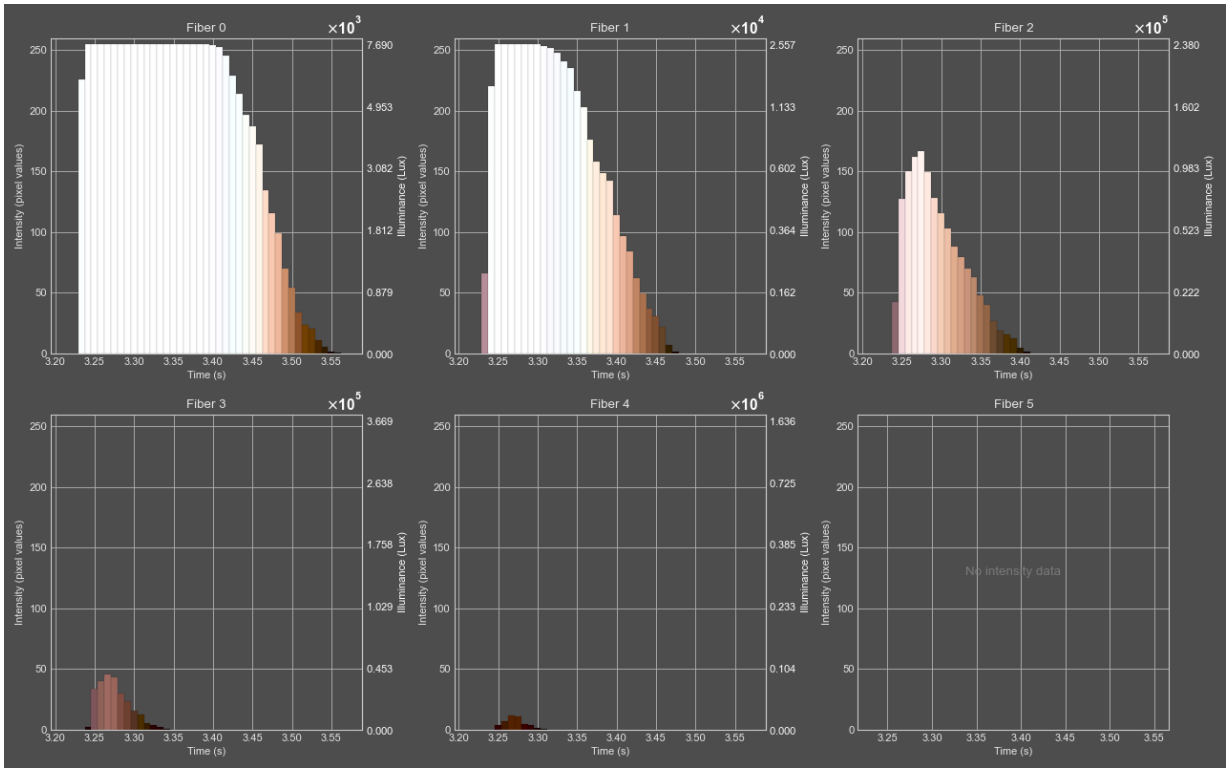


Figure 193: Selectivity/specificity: Ratio (40/60, m/m%) test 2

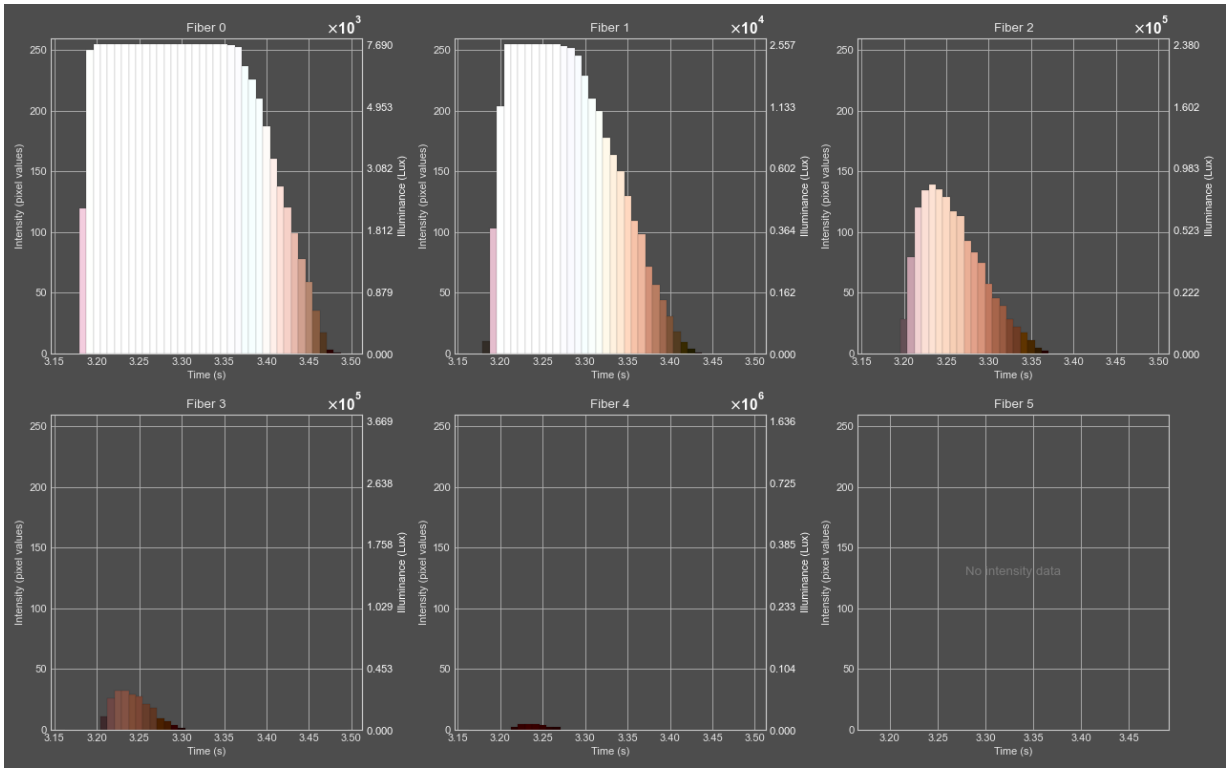


Figure 194: Selectivity/specificity: Ratio (40/60, m/m%) test 3

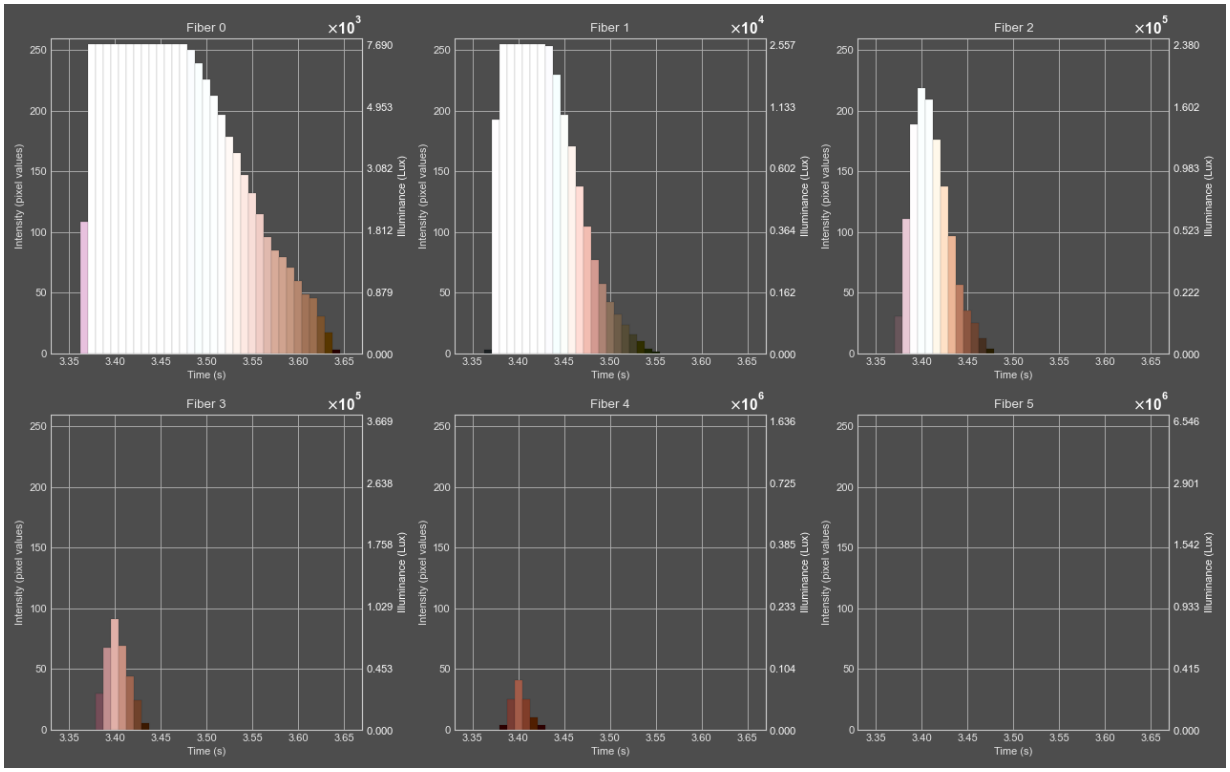


Figure 195: Selectivity/specificity: Ratio (40/60, m/m%) test 4

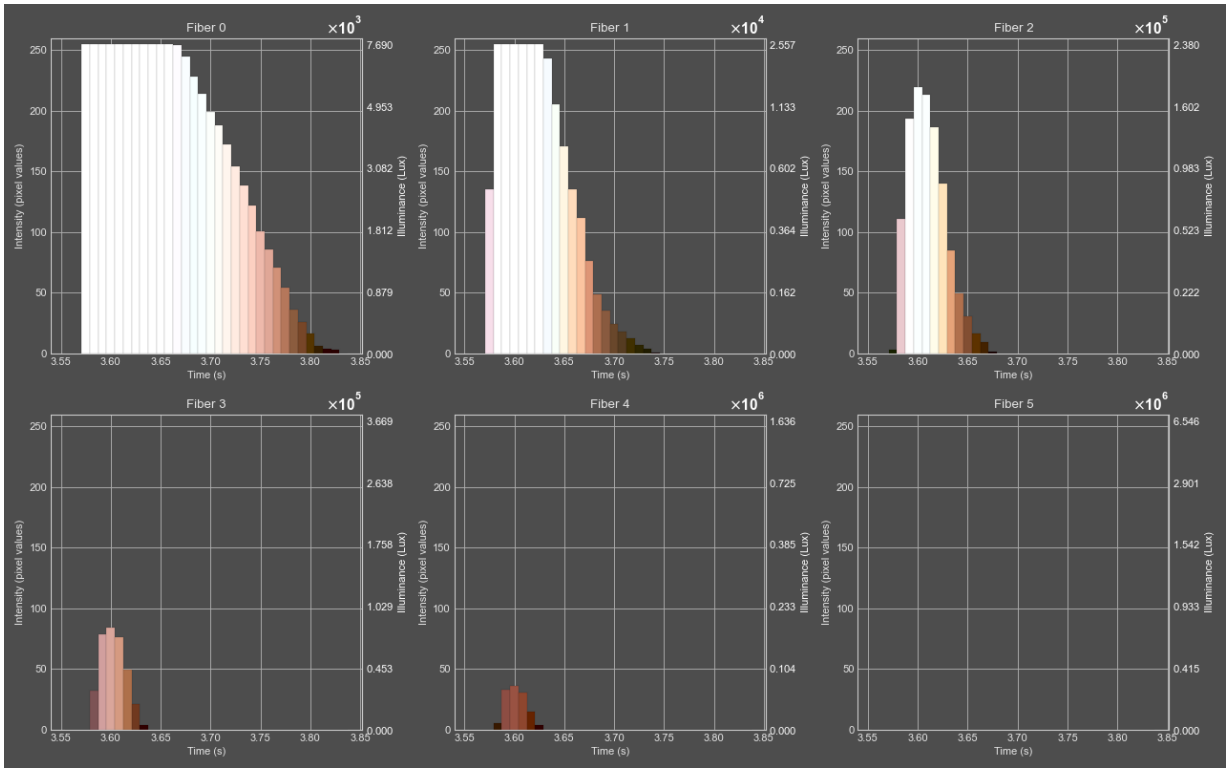


Figure 196: Selectivity/specificity: Ratio (40/60, m/m%) test 5

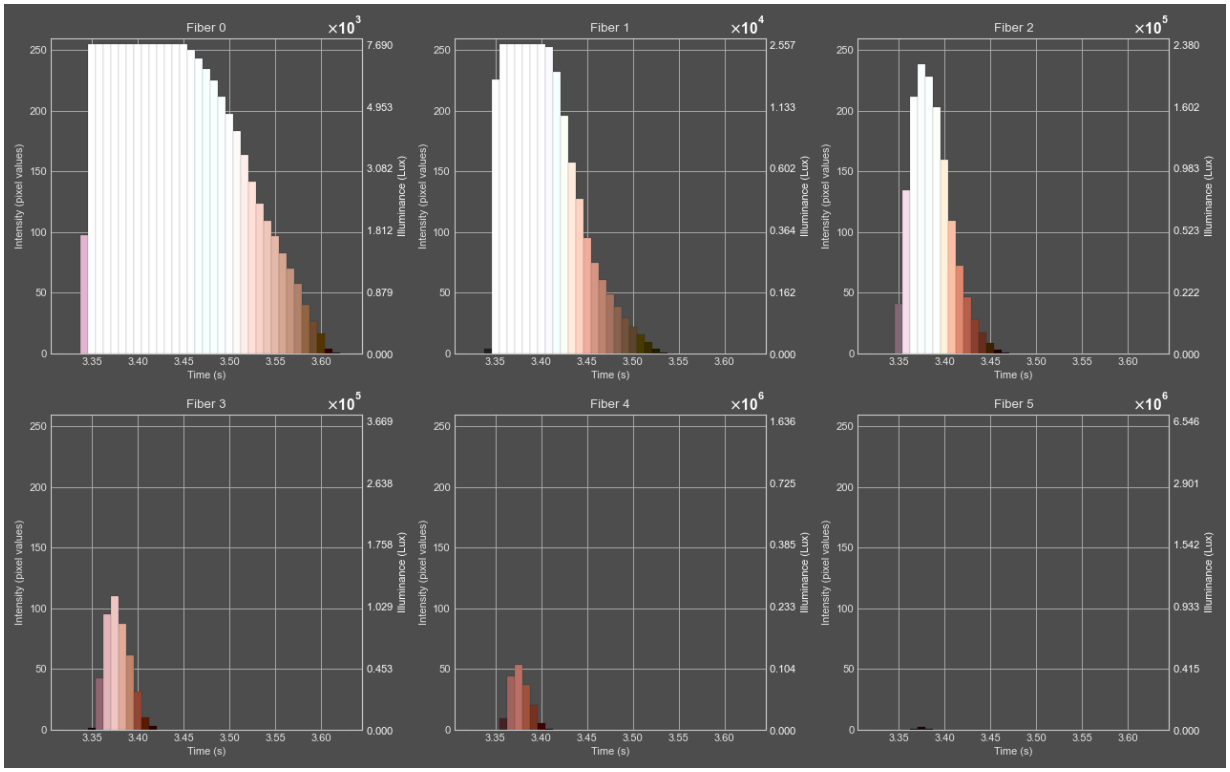


Figure 197: Selectivity/specificity: Ratio (40/60, m/m%) test 6

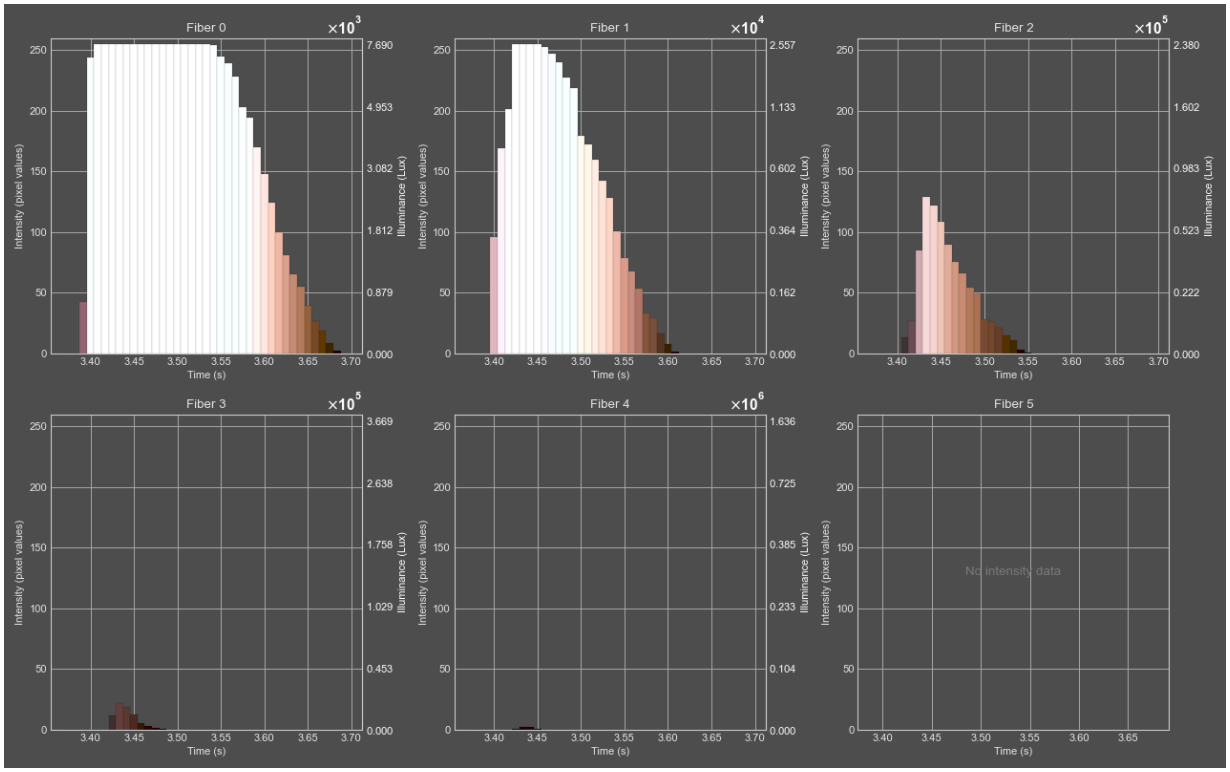


Figure 198: Selectivity/specificity: Ratio (40/60, m/m%) test 7

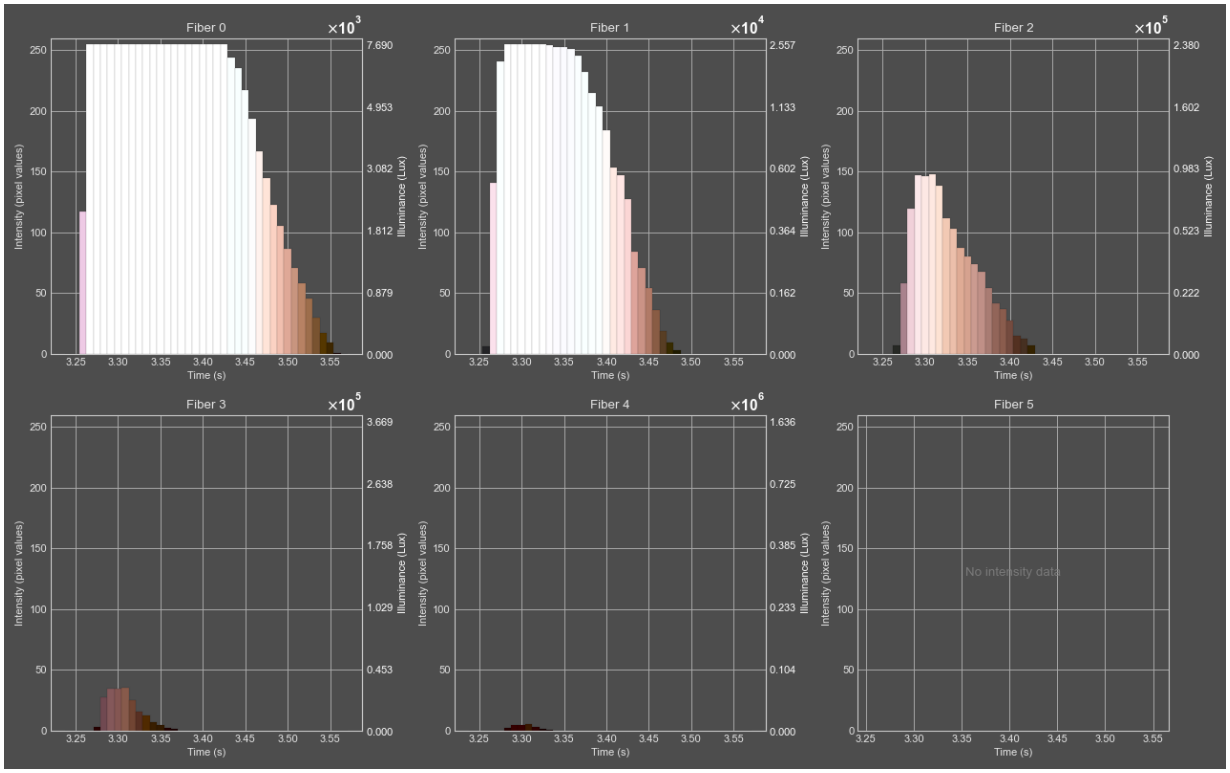


Figure 199: Selectivity/specificity: Ratio (40/60, m/m%) test 8

XI.II Ratio B: potassium perchlorate (50%)/aluminum (50%)

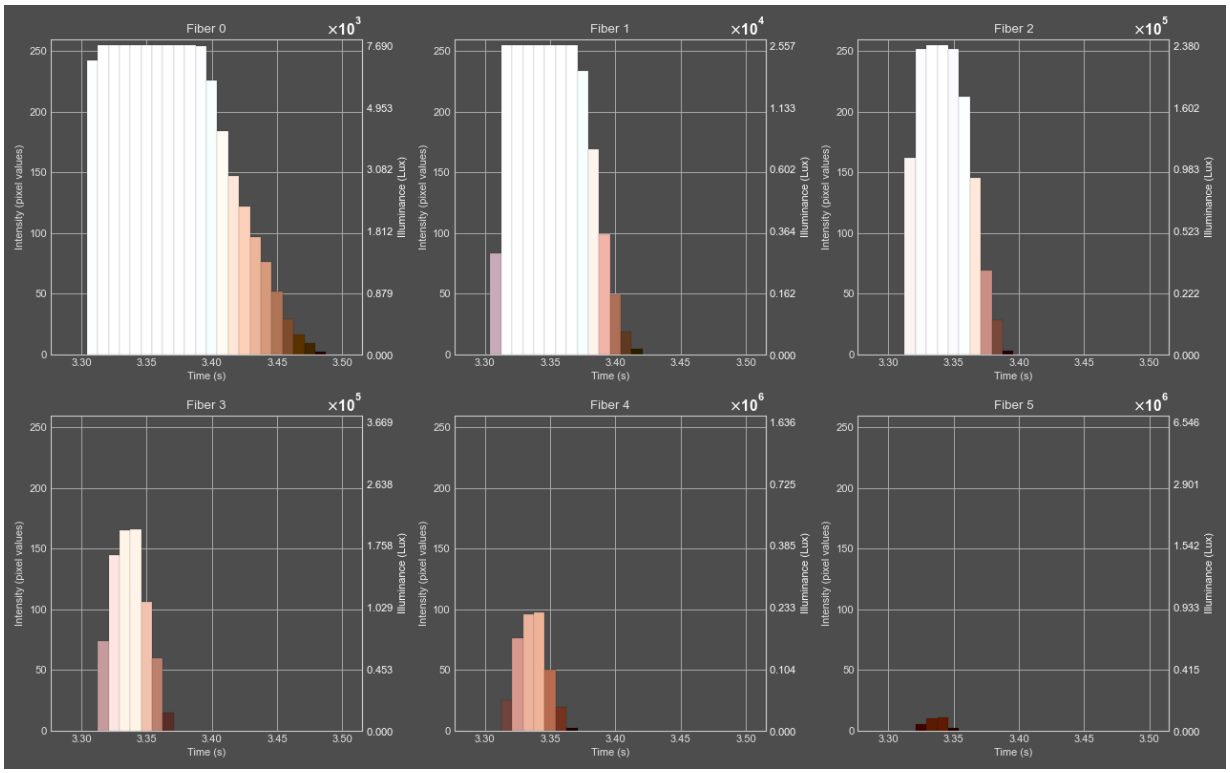


Figure 200: Selectivity/specificity: Ratio (50/50, m/m%) test 1

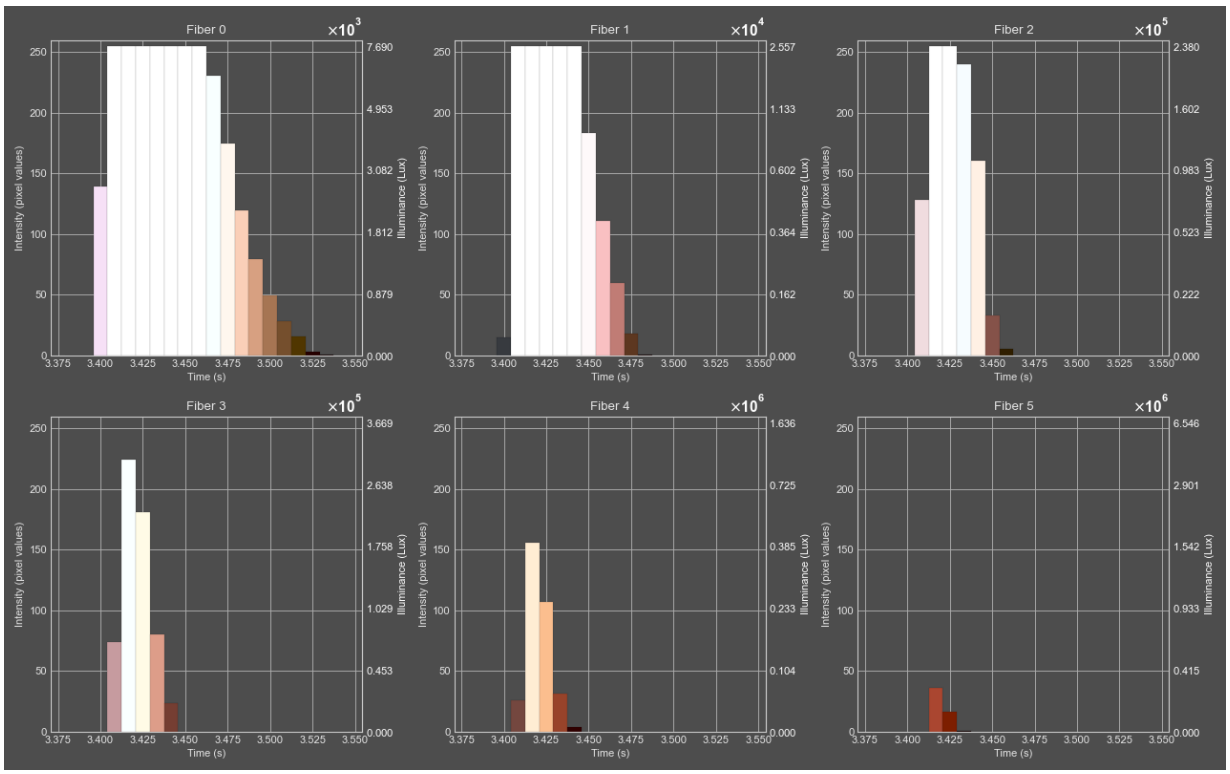


Figure 201: Selectivity/specificity: Ratio (50/50, m/m%) test 2

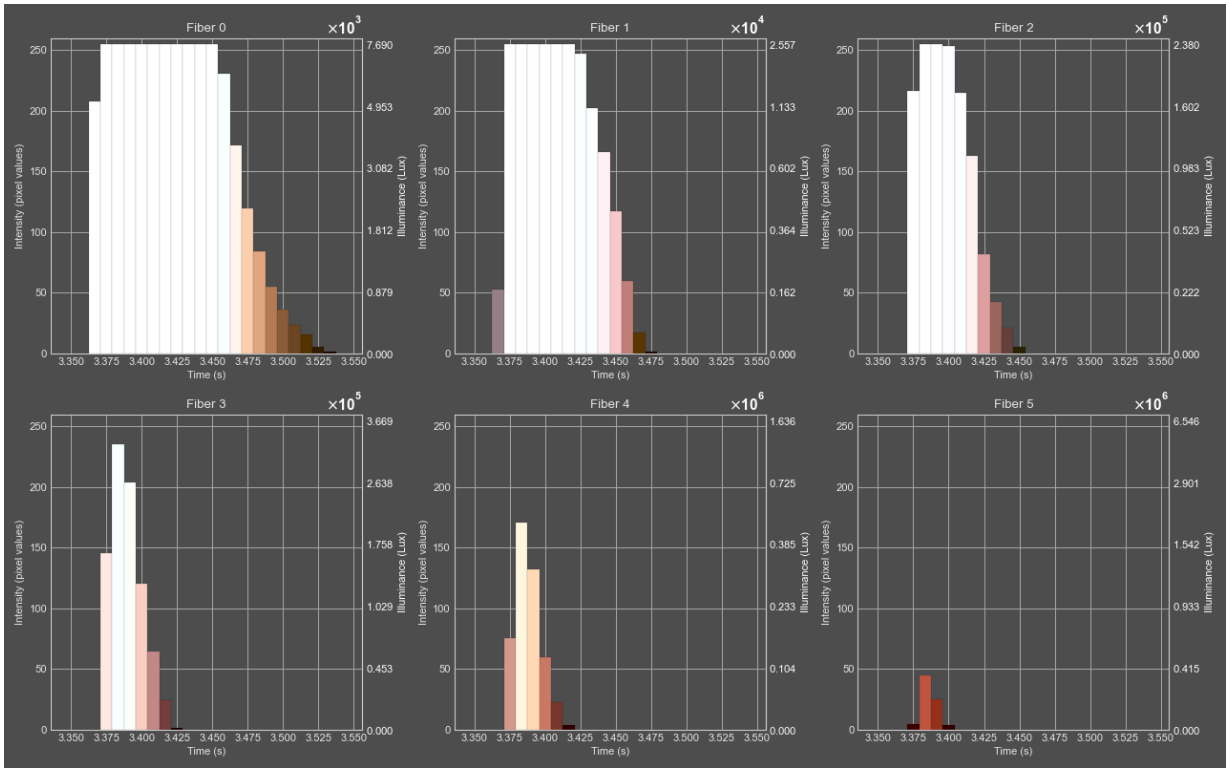


Figure 202: Selectivity/specificity: Ratio (50/50, m/m%) test 3

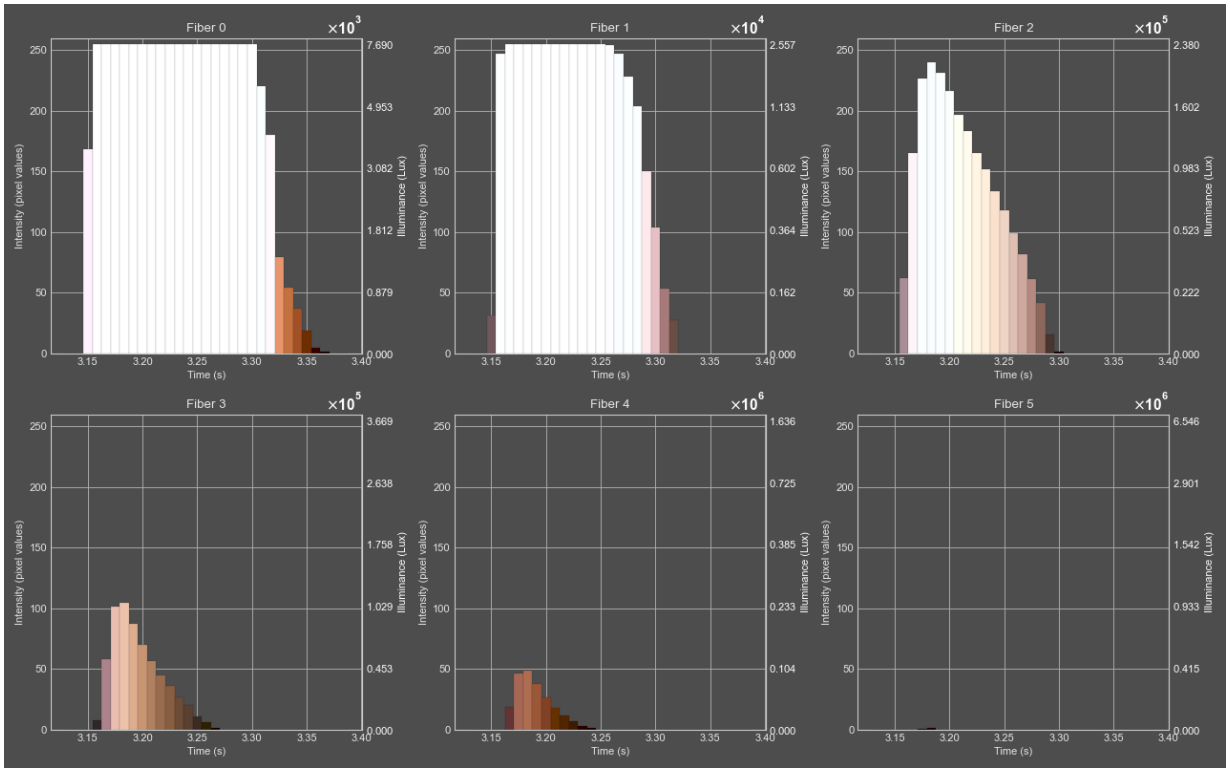


Figure 203: Selectivity/specificity: Ratio (50/50, m/m%) test 4

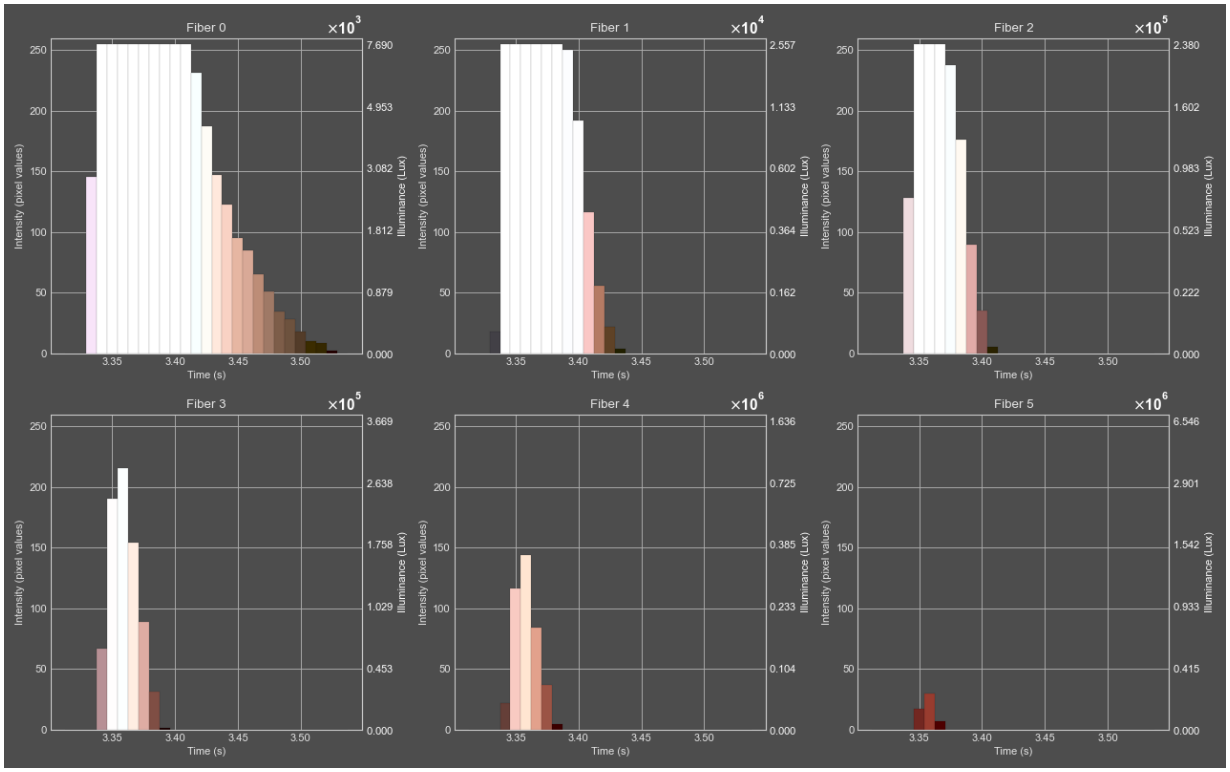


Figure 204: Selectivity/specificity: Ratio (50/50, m/m%) test 5

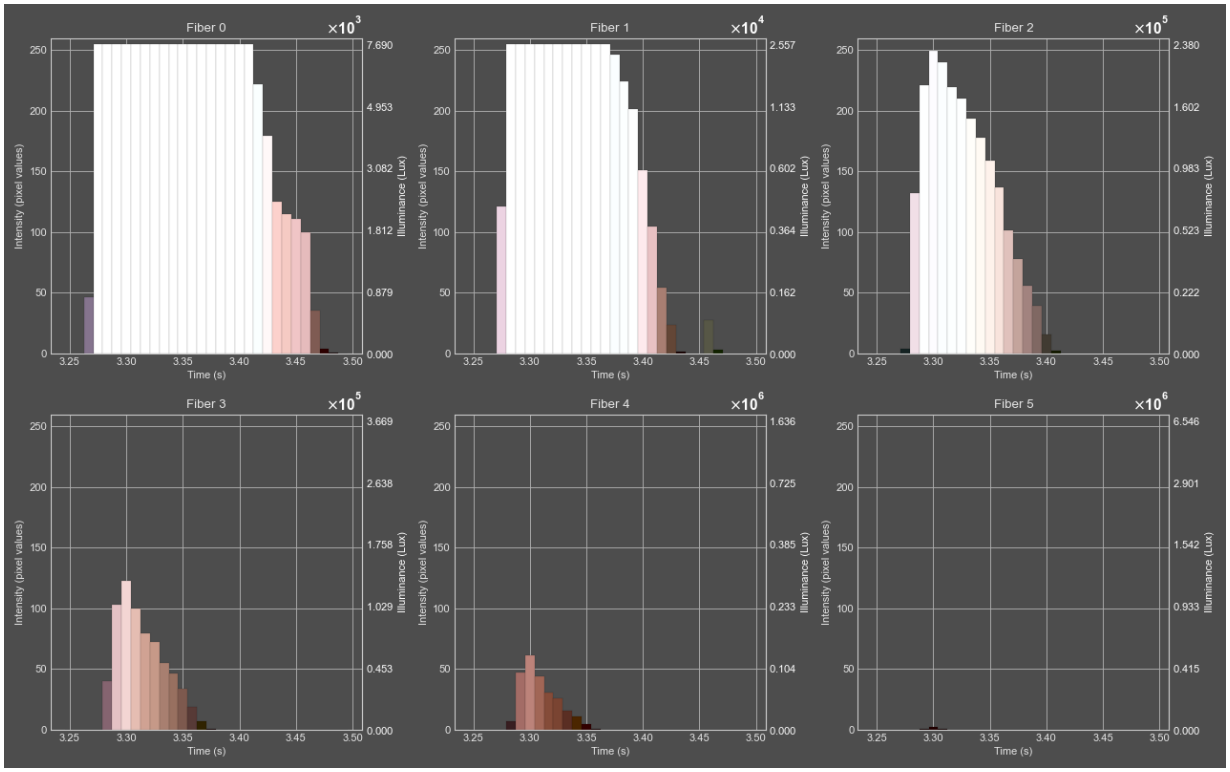


Figure 205: Selectivity/specificity: Ratio (50/50, m/m%) test 6

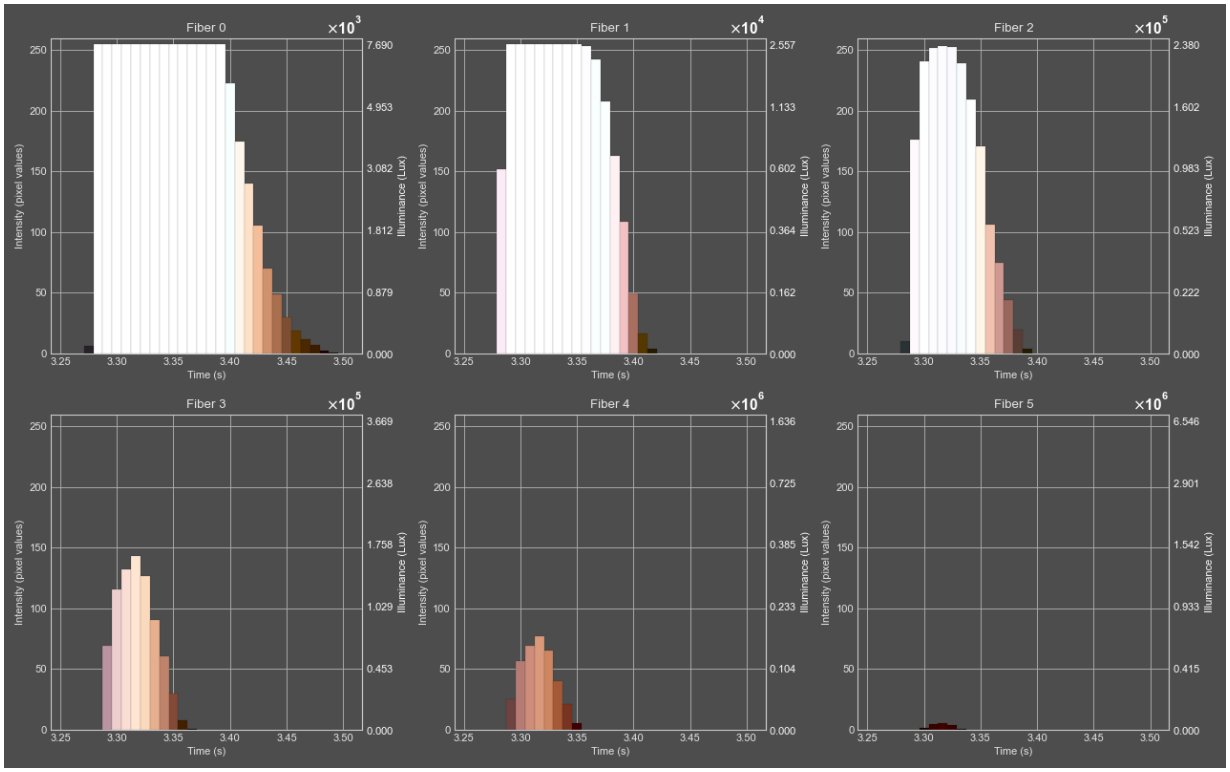


Figure 206: Selectivity/specificity: Ratio (50/50, m/m%) test 7

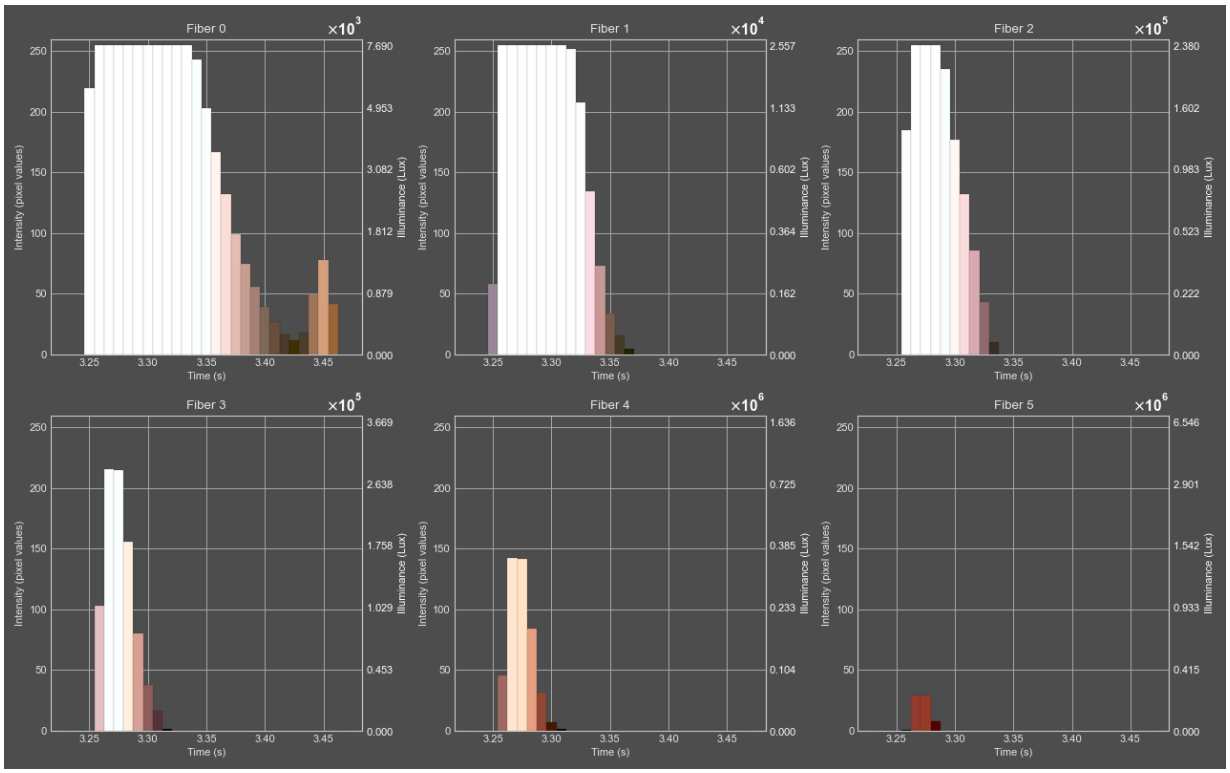


Figure 207: Selectivity/specificity: Ratio (50/50, m/m%) test 8

XI.III Ratio C: potassium perchlorate (80%)/aluminum (20%)

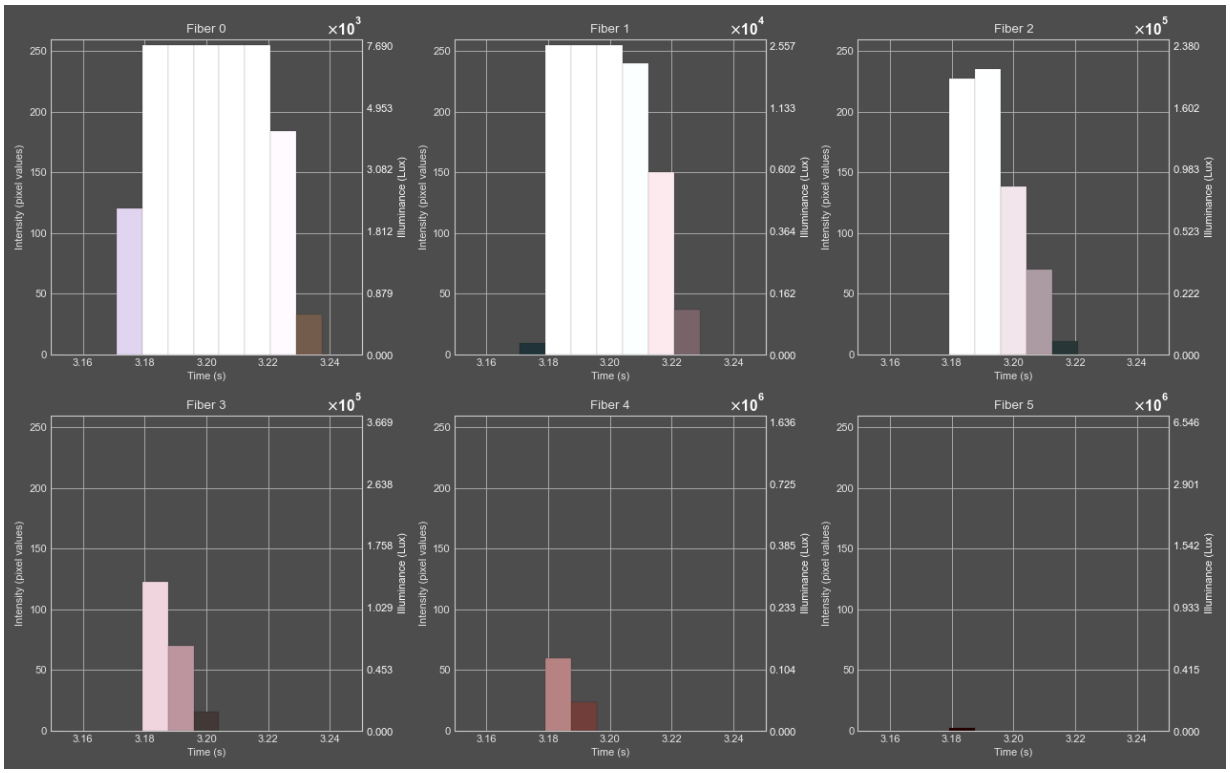


Figure 208: Selectivity/specificity: Ratio (80/20, m/m%) test 1

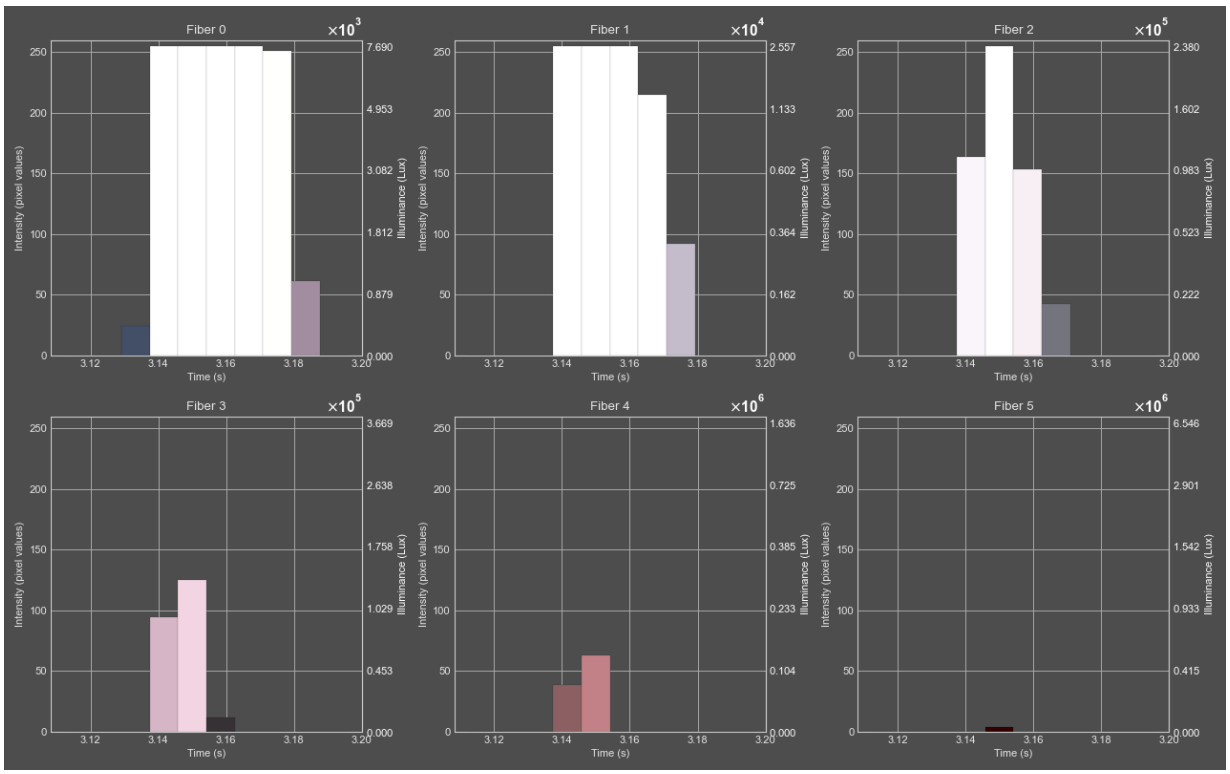


Figure 209: Selectivity/specificity: Ratio (80/20, m/m%) test 2

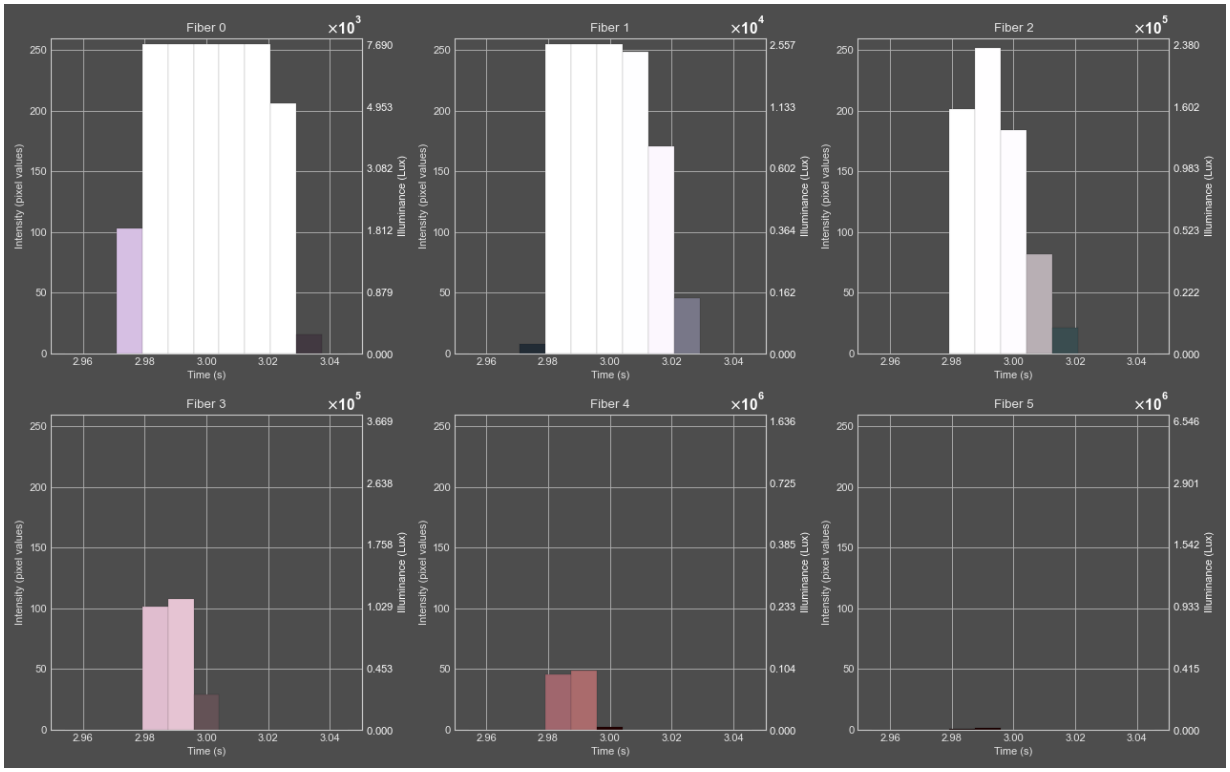


Figure 210: Selectivity/specificity: Ratio (80/20, m/m%) test 3

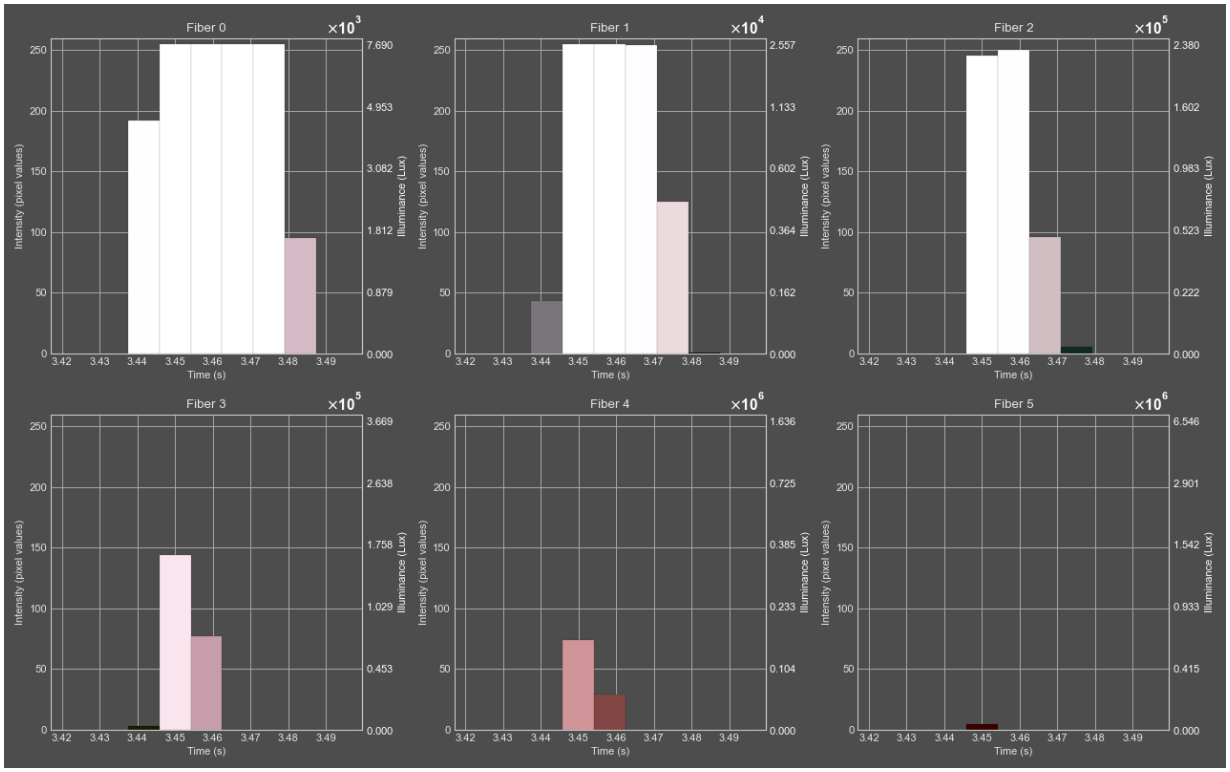


Figure 211: Selectivity/specificity: Ratio (80/20, m/m%) test 4

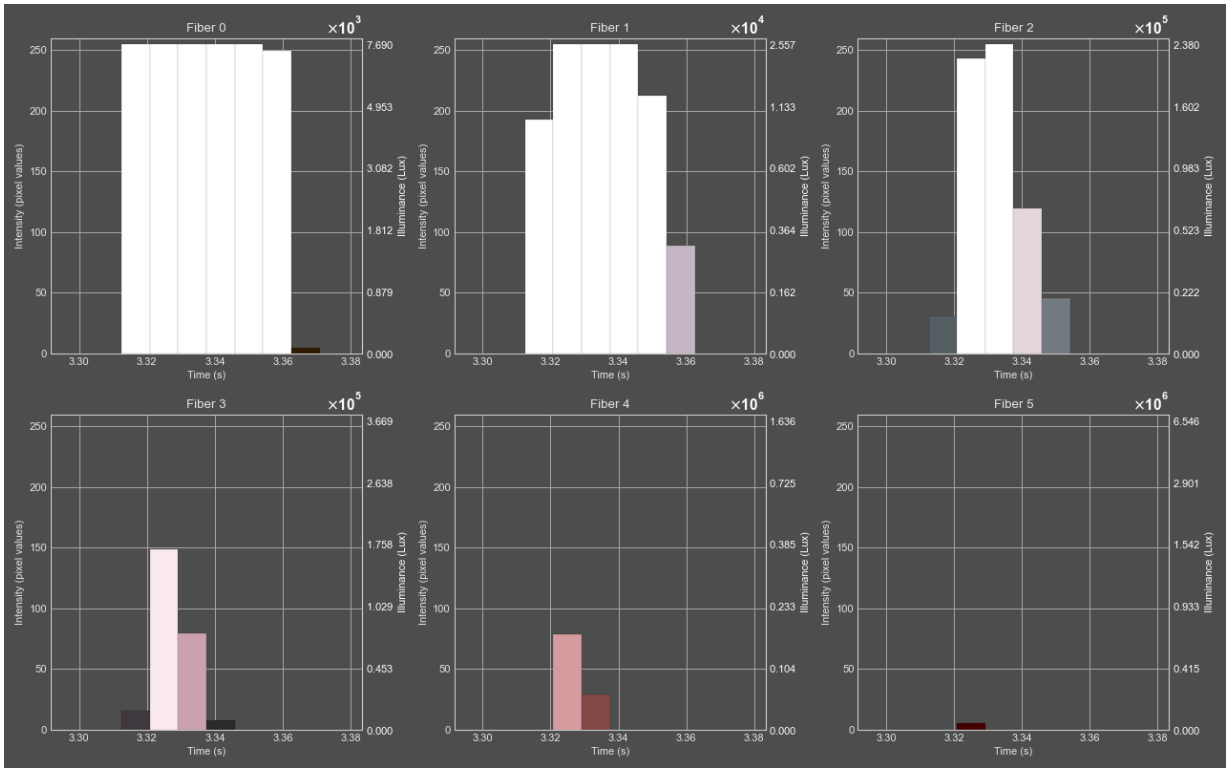


Figure 212: Selectivity/specificity: Ratio (80/20, m/m%) test 5

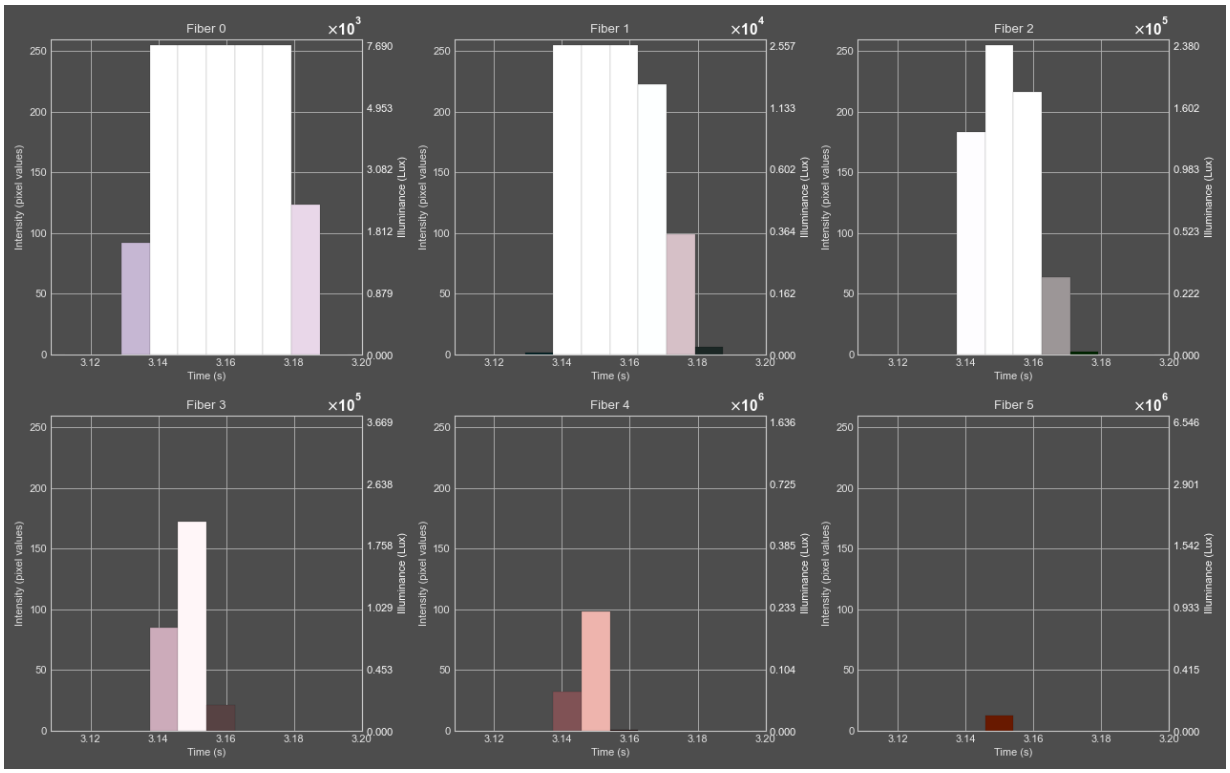


Figure 213: Selectivity/specificity: Ratio (80/20, m/m%) test 6

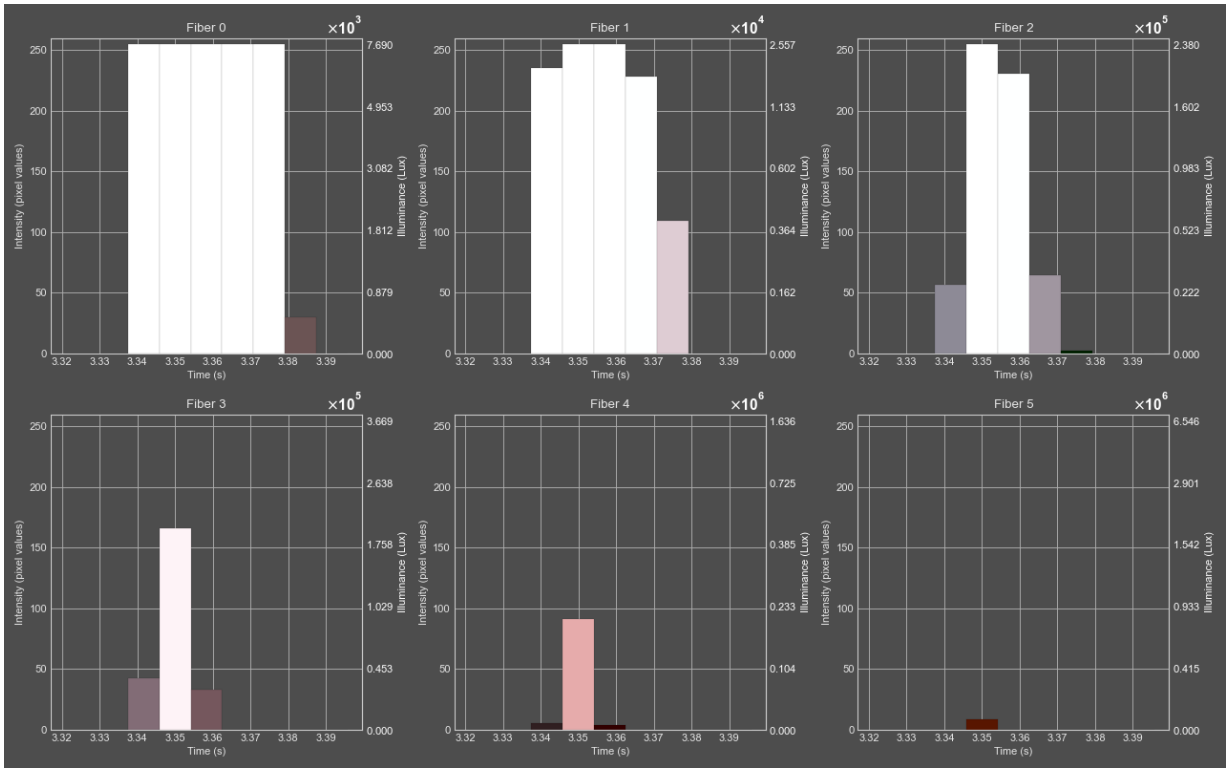


Figure 214: Selectivity/specificity: Ratio (80/20, m/m%) test 7

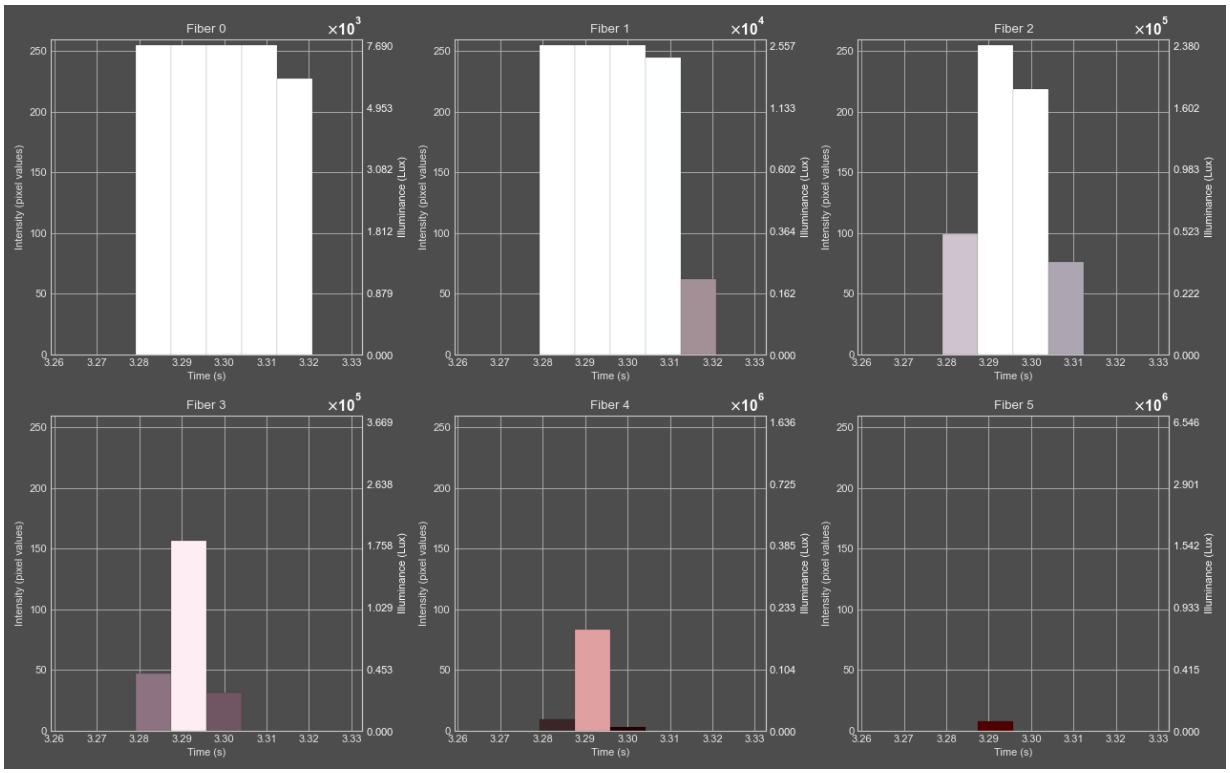


Figure 215: Selectivity/specificity: Ratio (80/20, m/m%) test 8

Appendix XII: validation results for specificity/selectivity: grain size variations

This appendix presents the validation results of the combustions of compositions with variation in grain sizes tests. The following figures show the PEMAD analysis results for each test replicate (n = 8), illustrating the combustion intensity profiles over time and the corresponding color behavior for both compositions.

XII.I Grain size fine (composition: N310)

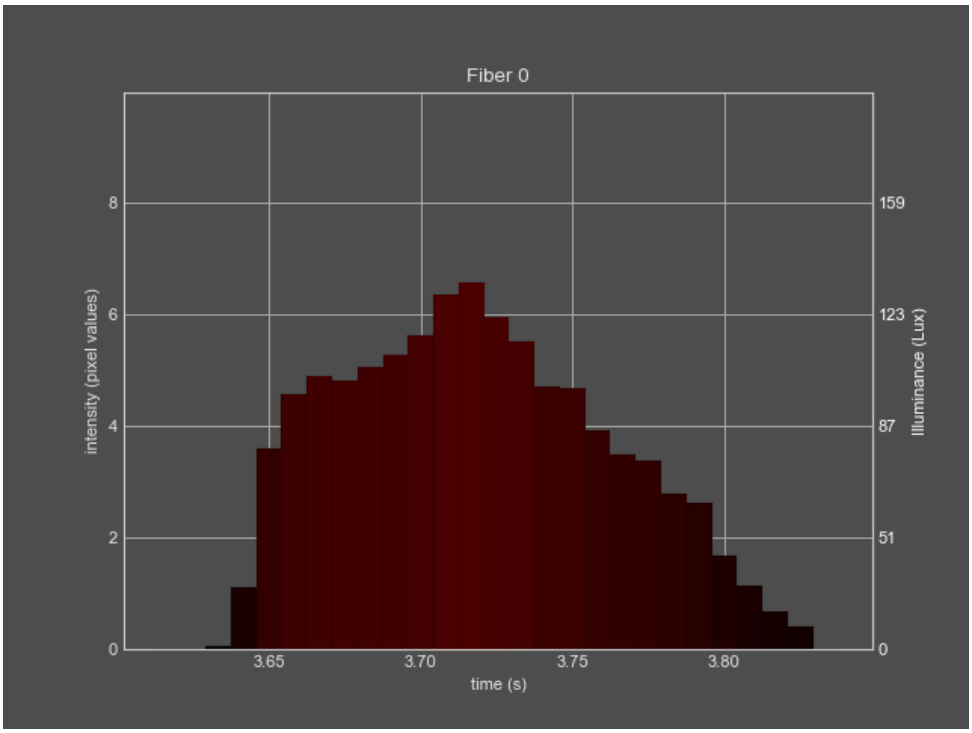


Figure 216: Selectivity/specificity results: grain size (composition: N310) "fine" test 1 (2)

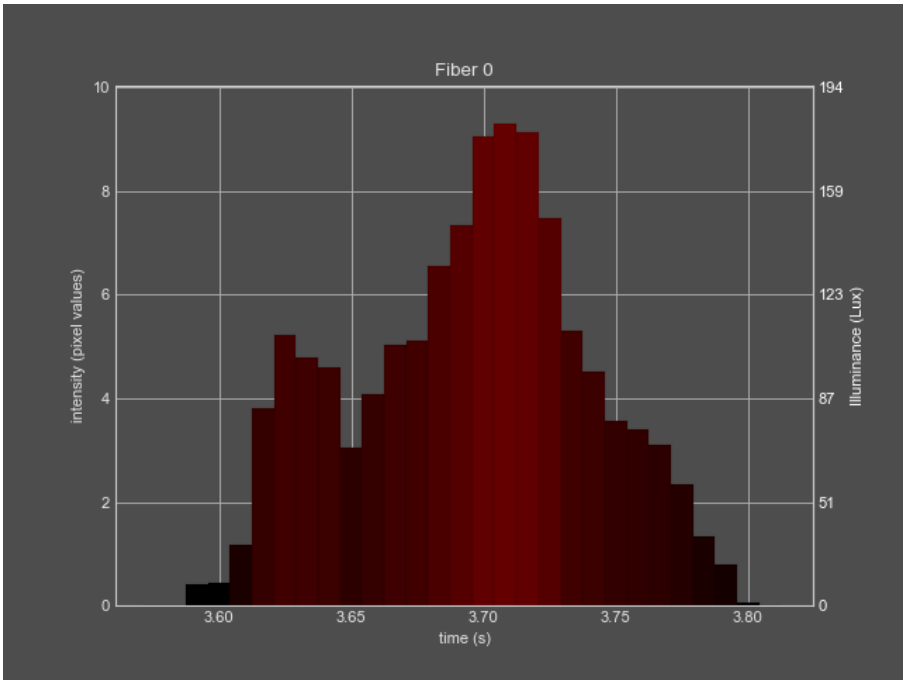


Figure 217: Selectivity/specificity results: grain size (composition: N310) "fine" test 2 (2)

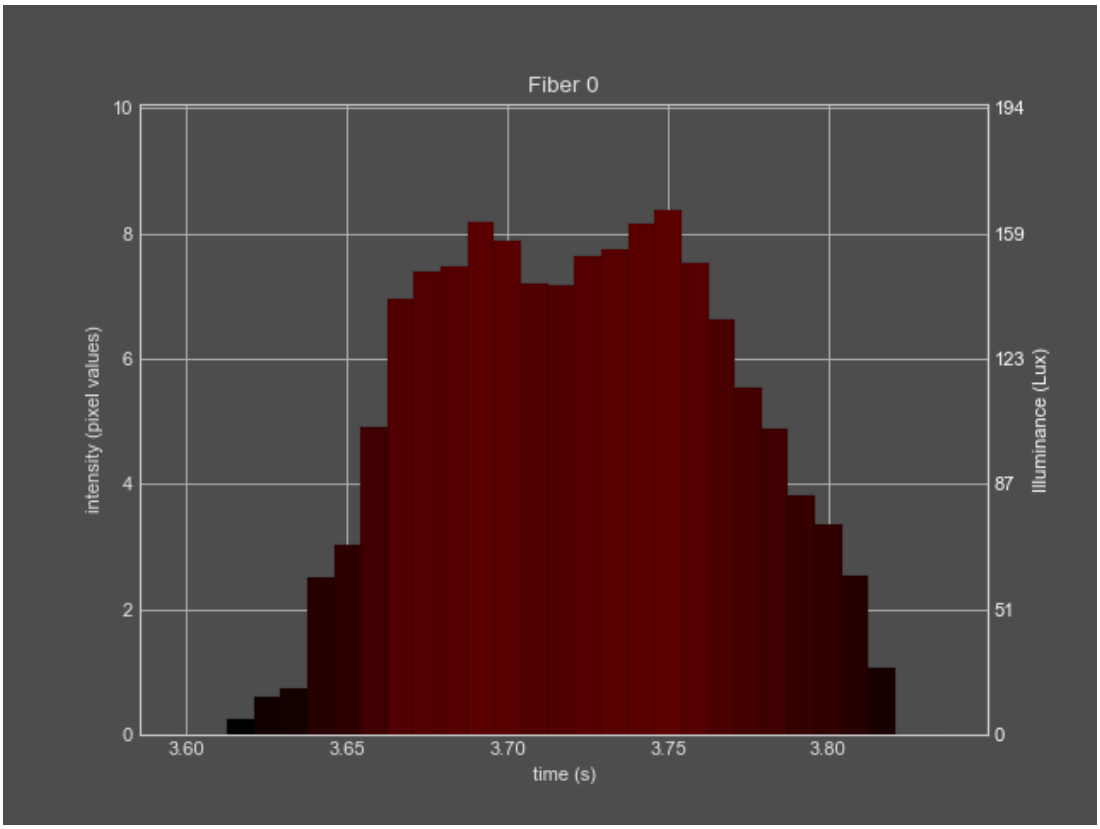


Figure 218: Selectivity/specificity results: grain size (composition: N310) "fine" test 3 (2)

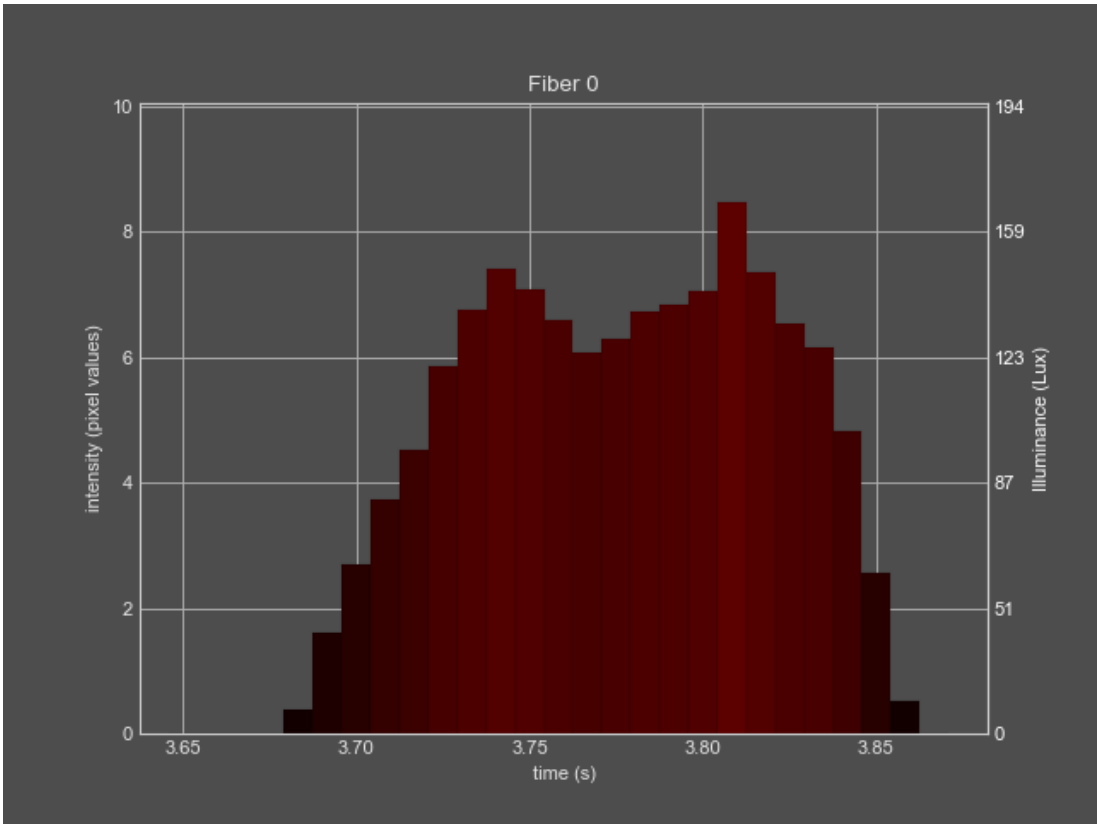


Figure 219: Selectivity/specificity results: grain size (composition: N310) "fine" test 4 (2)

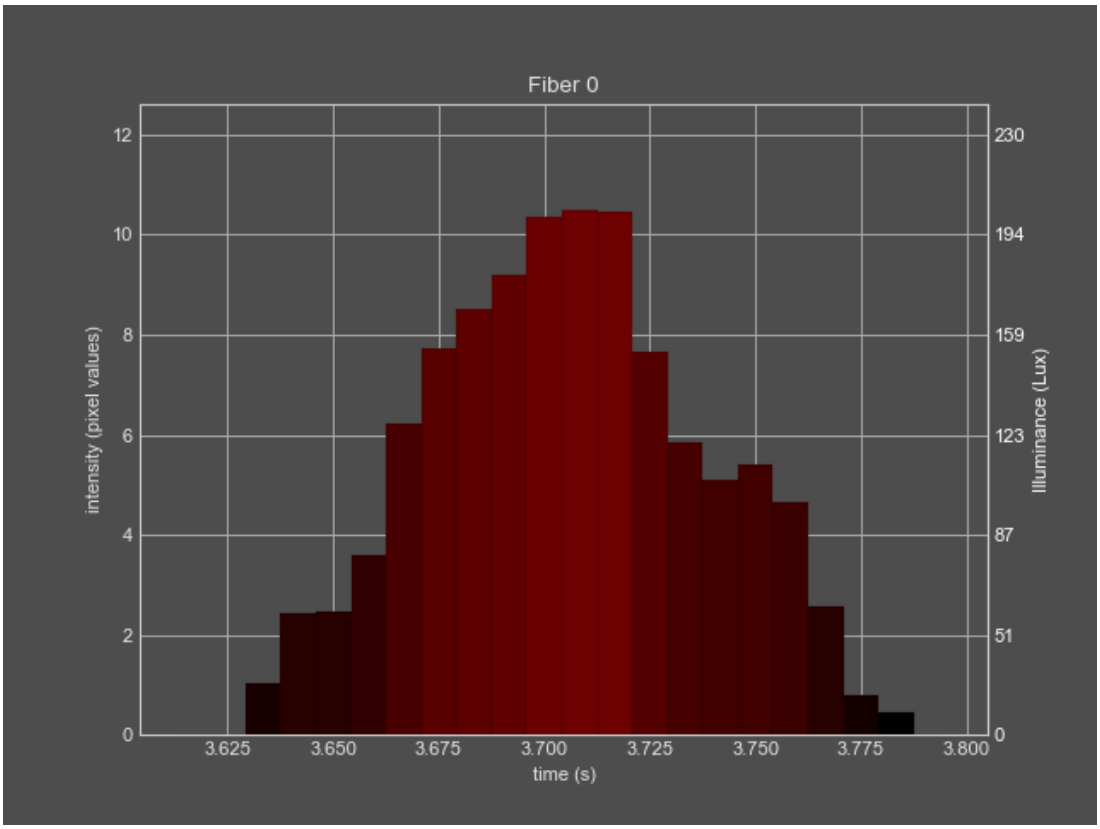


Figure 220: Selectivity/specificity results: grain size (composition: N310) "fine" test 5 (2)

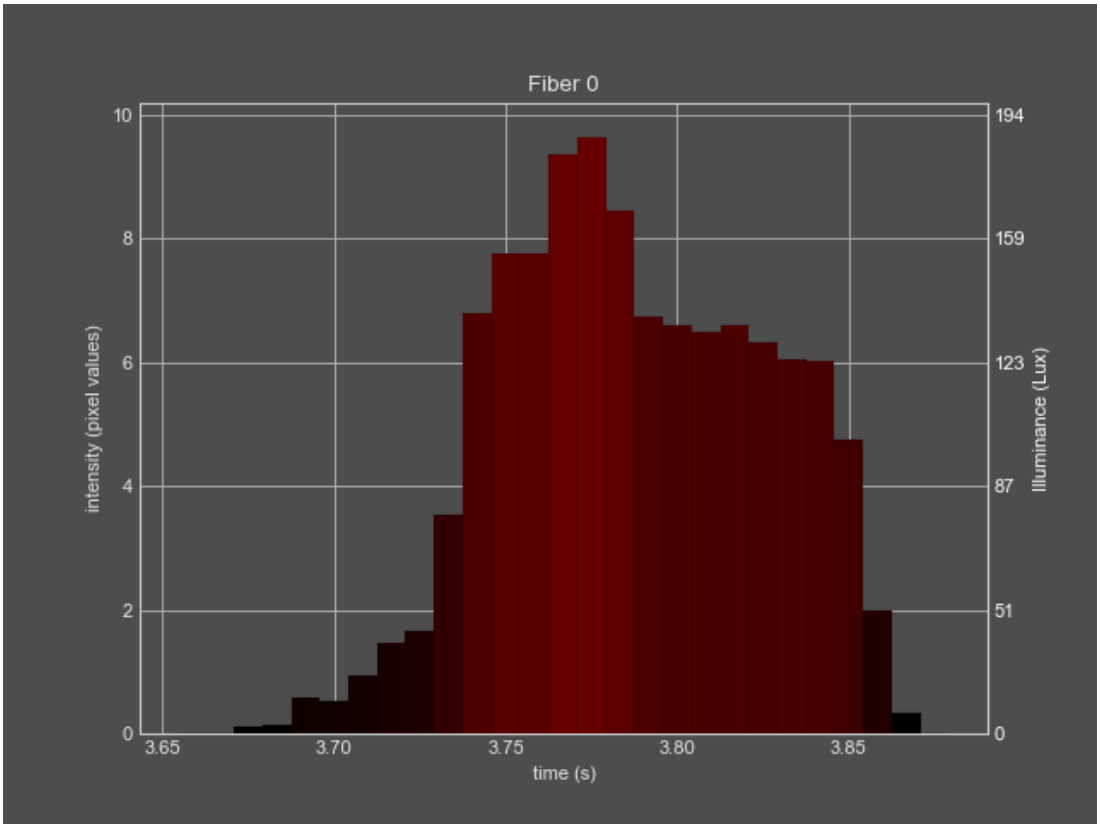


Figure 221: Selectivity/specificity results: grain size (composition: N310) "fine" test 6 (2)

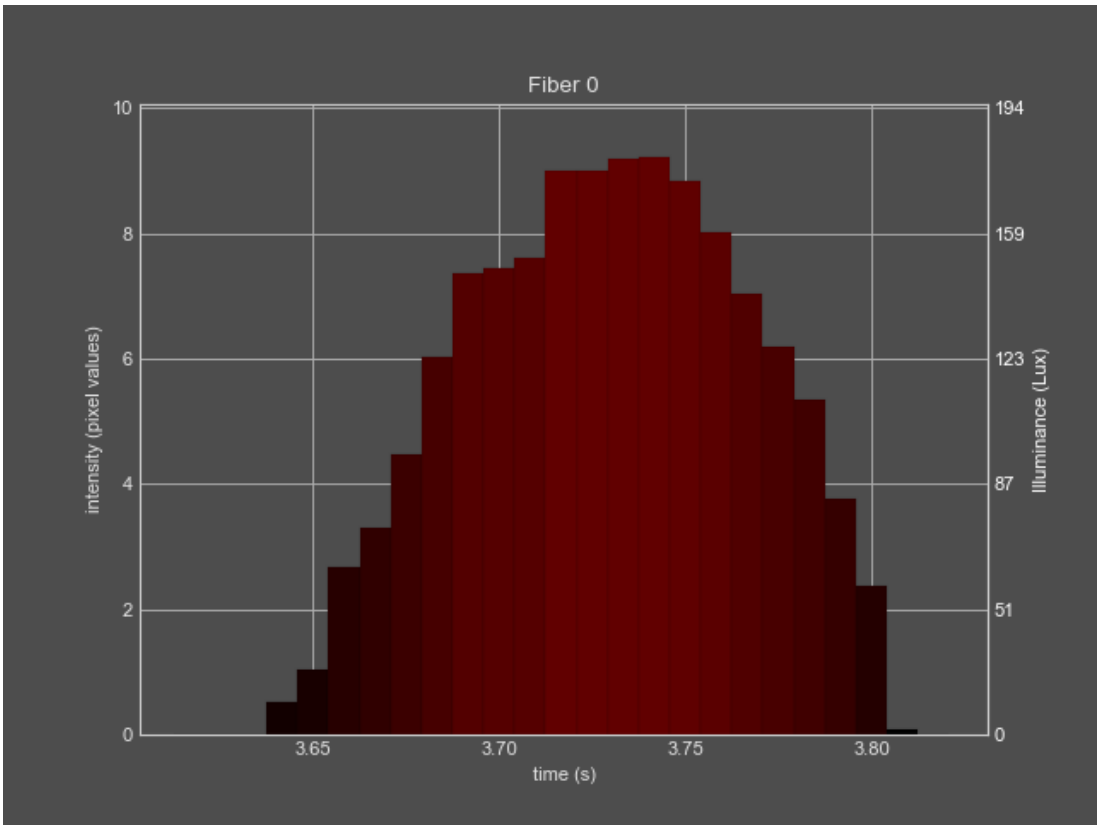


Figure 222: Selectivity/specificity results: grain size (composition: N310) "fine" test 7 (2)

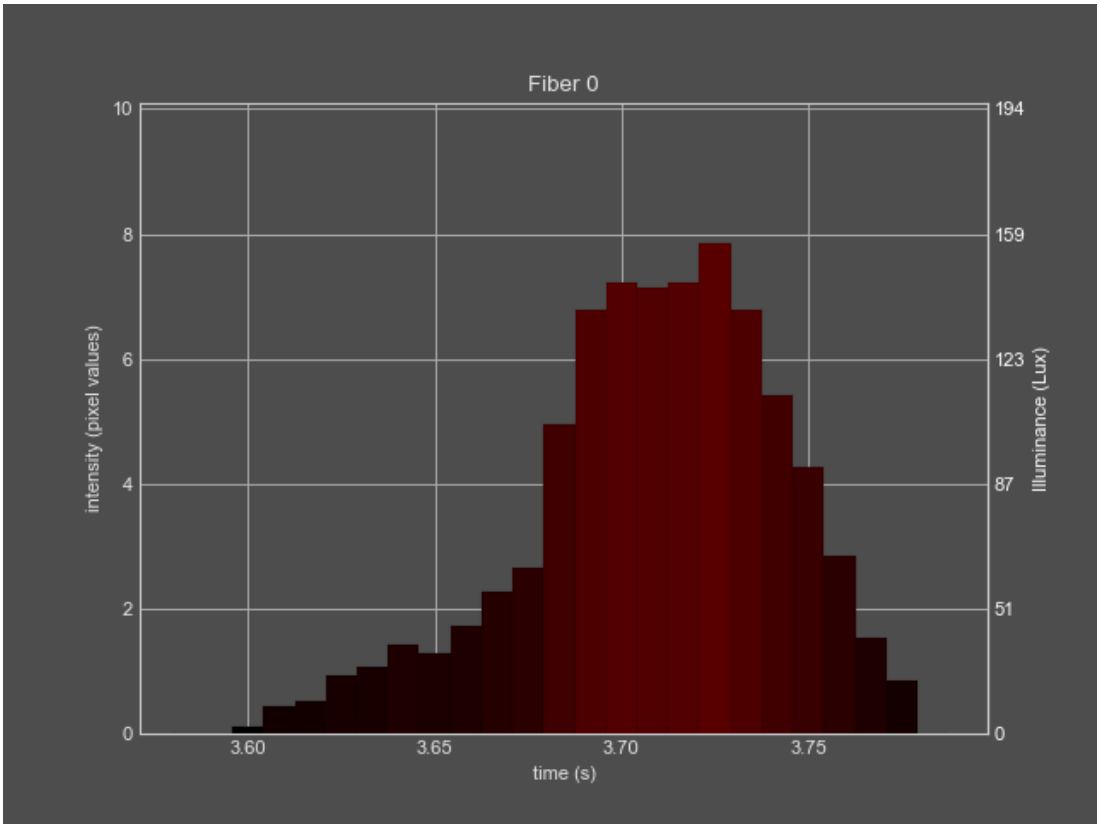


Figure 223: Selectivity/specificity results: grain size (composition: N310) "fine" test 8 (2)

XII.II Grain size medium (composition: N110)

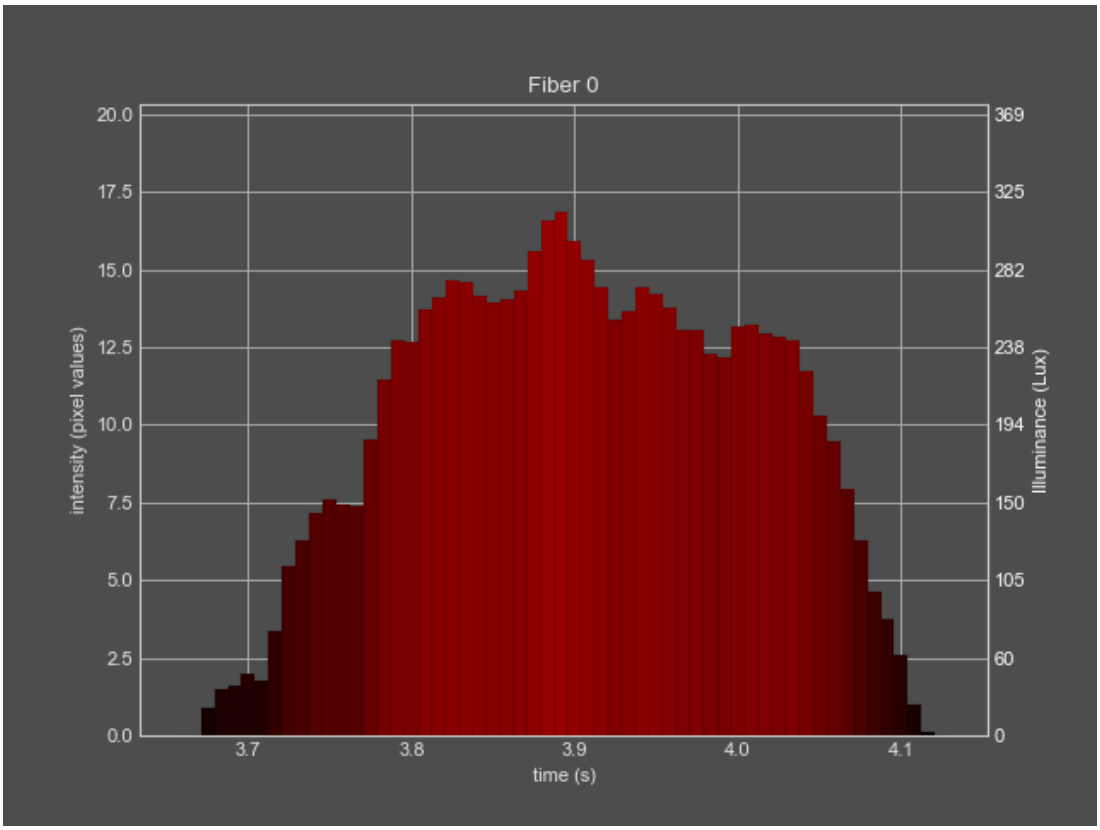


Figure 224: Selectivity/specificity results: grain size (composition: N110) "medium" test 1 (2)

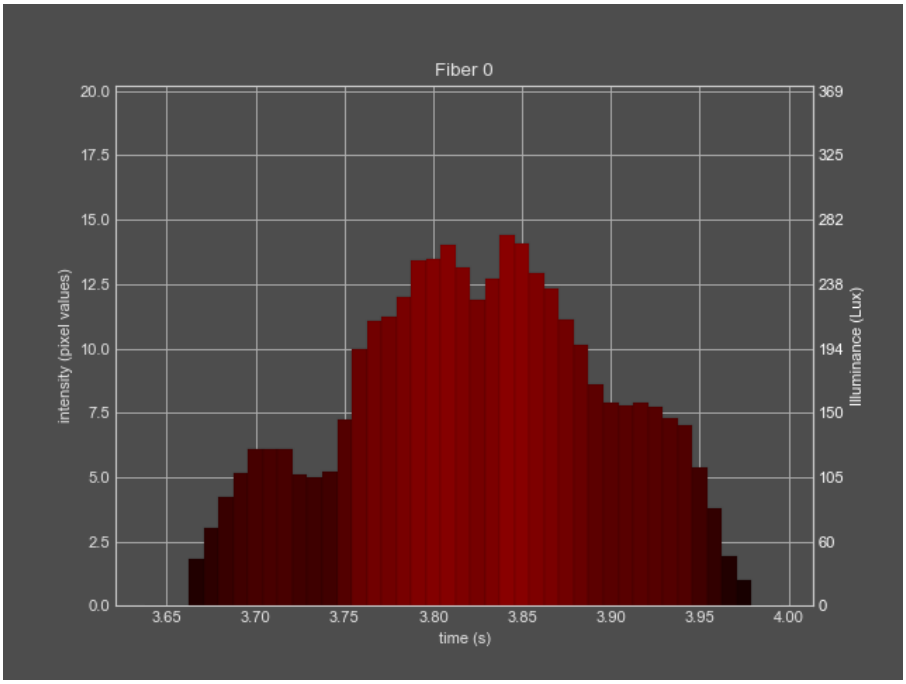


Figure 225: Selectivity/specificity results: grain size (composition: N110) "medium" test 2 (2)

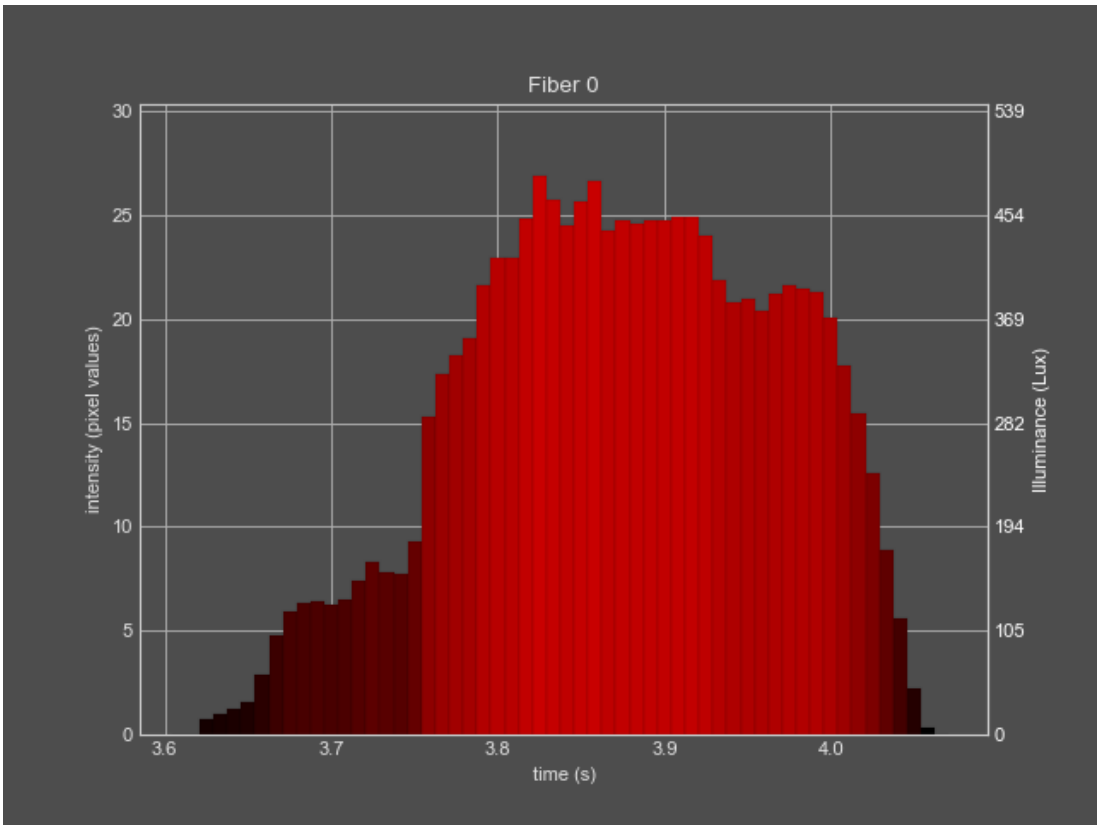


Figure 226: Selectivity/specificity results: grain size (composition: N110) "medium" test 3 (2)

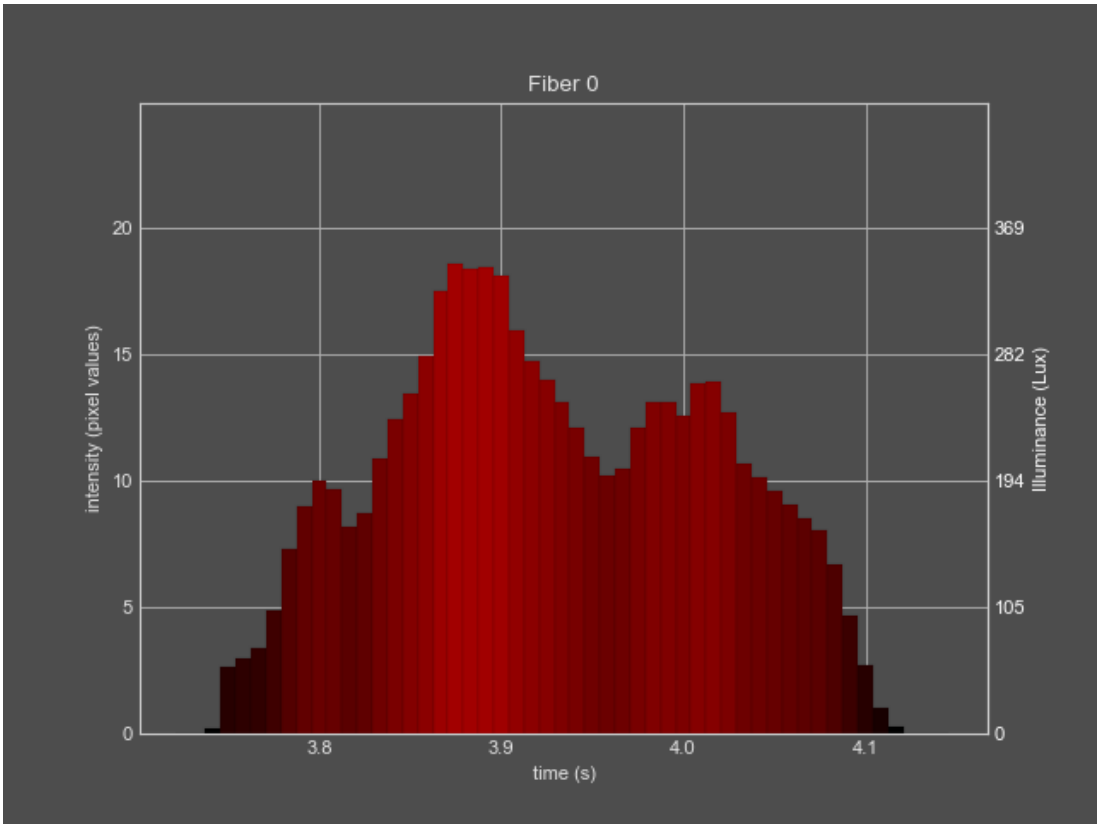


Figure 227: Selectivity/specificity results: grain size (composition: N110) "medium" test 4 (2)

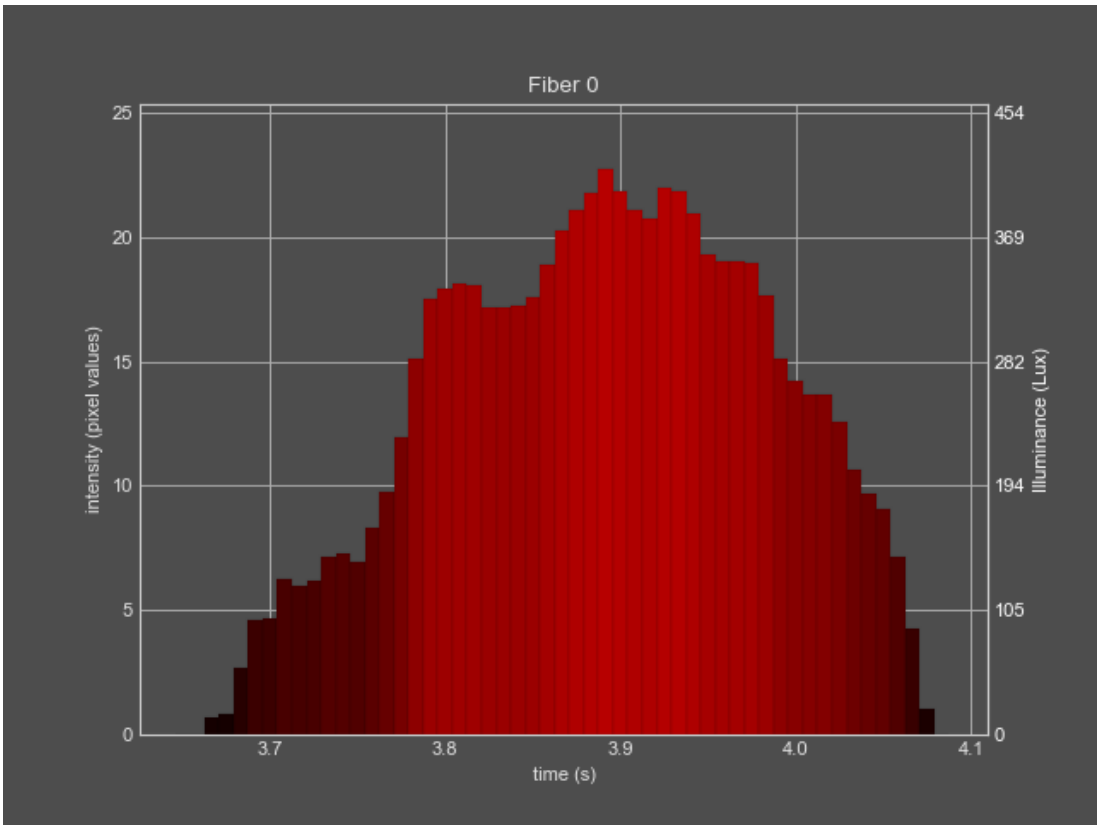


Figure 228: Selectivity/specificity results: grain size (composition: N110) "medium" test 5 (2)

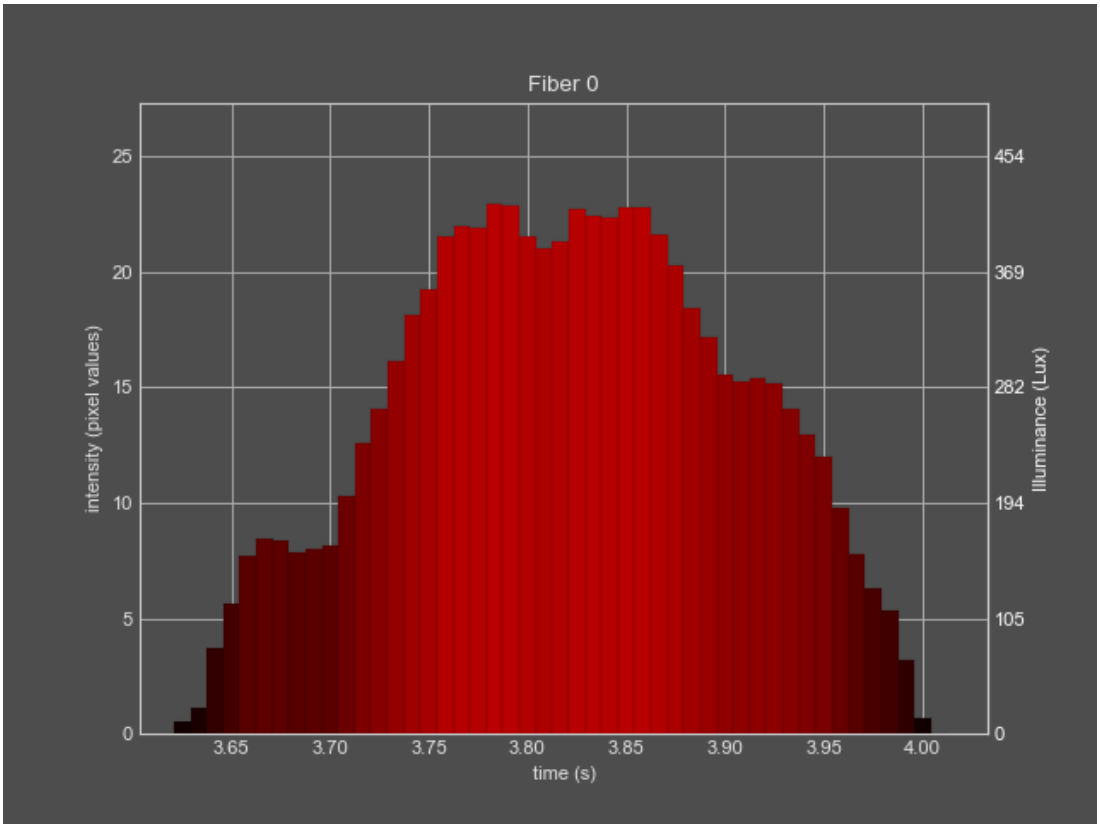


Figure 229: Selectivity/specificity results: grain size (composition: N110) "medium" test 6 (2)

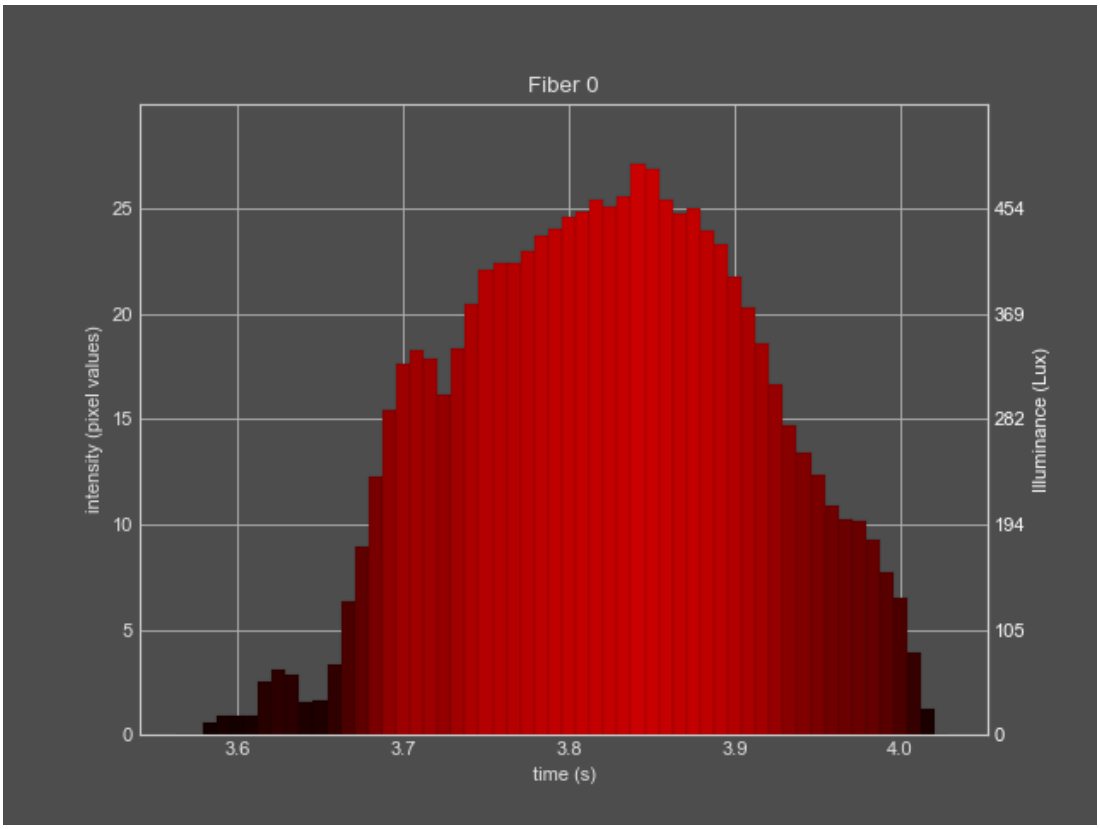


Figure 230: Selectivity/specificity results: grain size (composition: N110) "medium" test 7 (2)

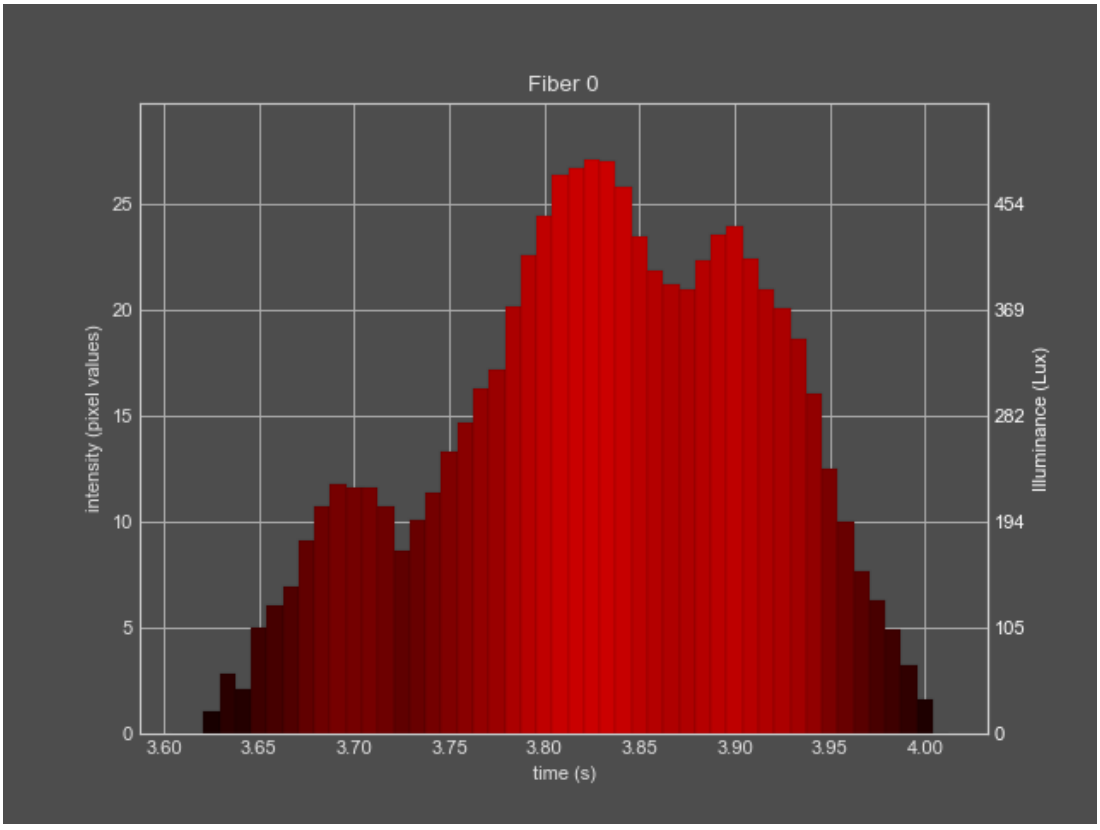


Figure 231: Selectivity/specificity results: grain size (composition: N110) "medium" test 8 (2)

XII.III Grain size coarse (composition: N160)

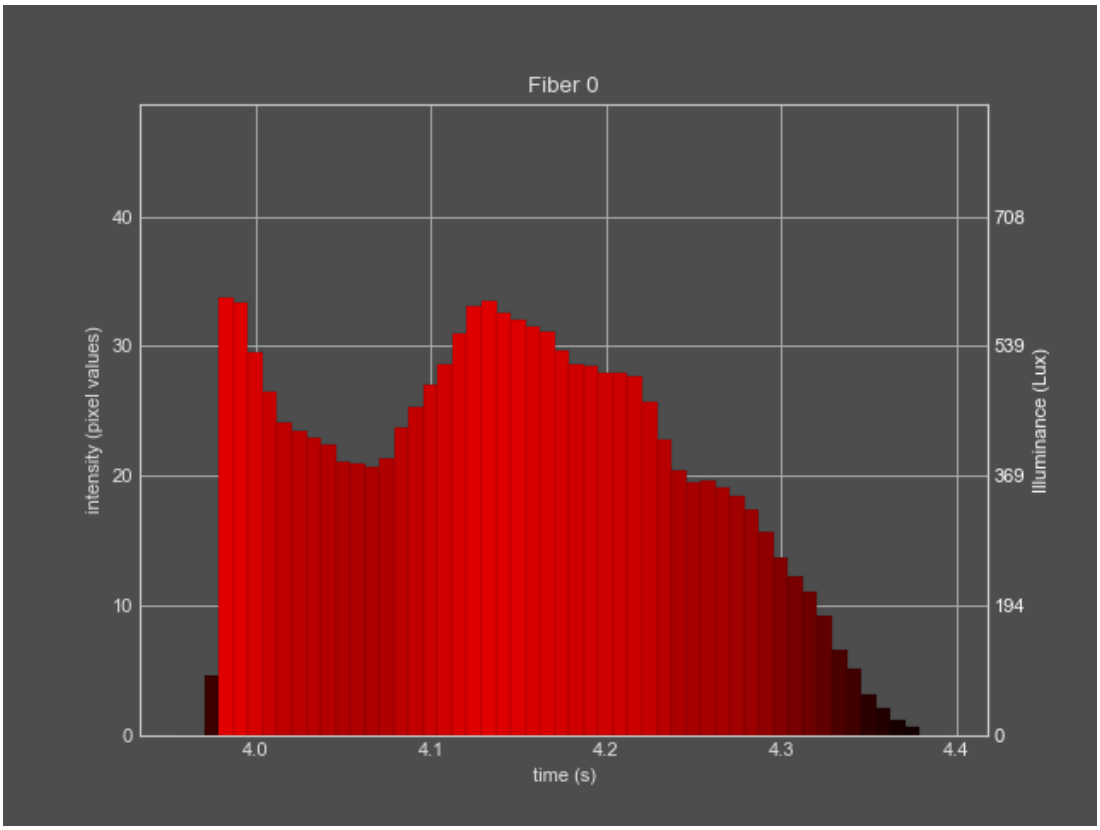


Figure 232: Selectivity/specificity results: grain size (composition: N160) "coarse" test 1 (2)

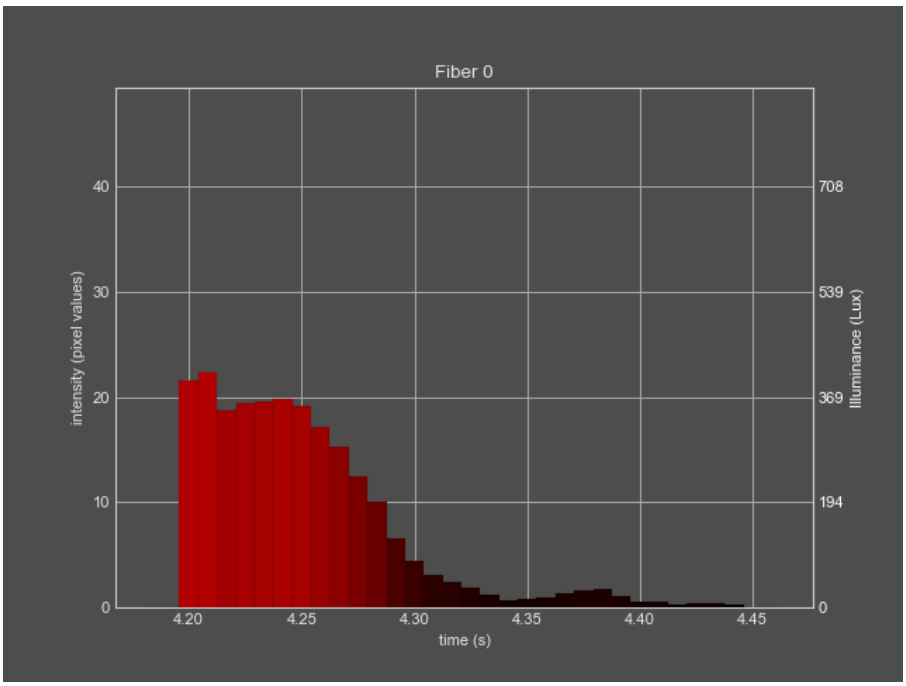


Figure 233: Selectivity/specificity results: grain size (composition: N160) "coarse" test 2 (2)

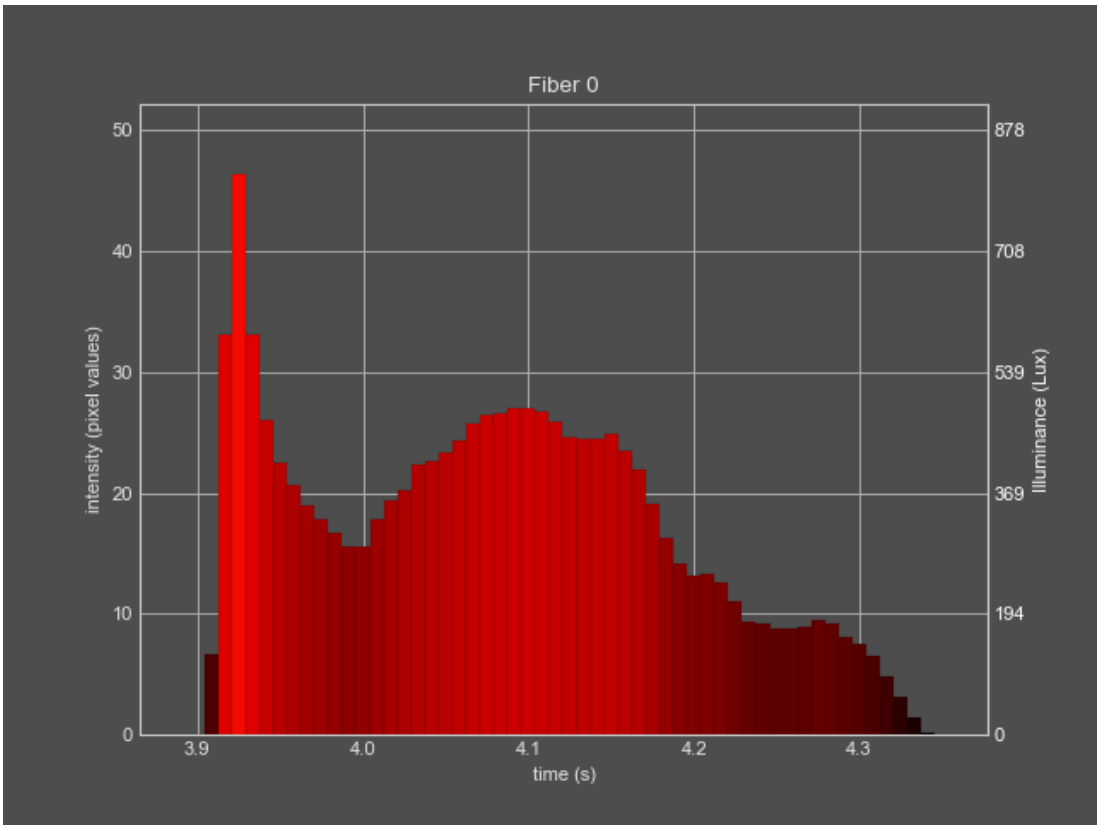


Figure 234: Selectivity/specificity results: grain size (composition: N160) "coarse" test 3 (2)

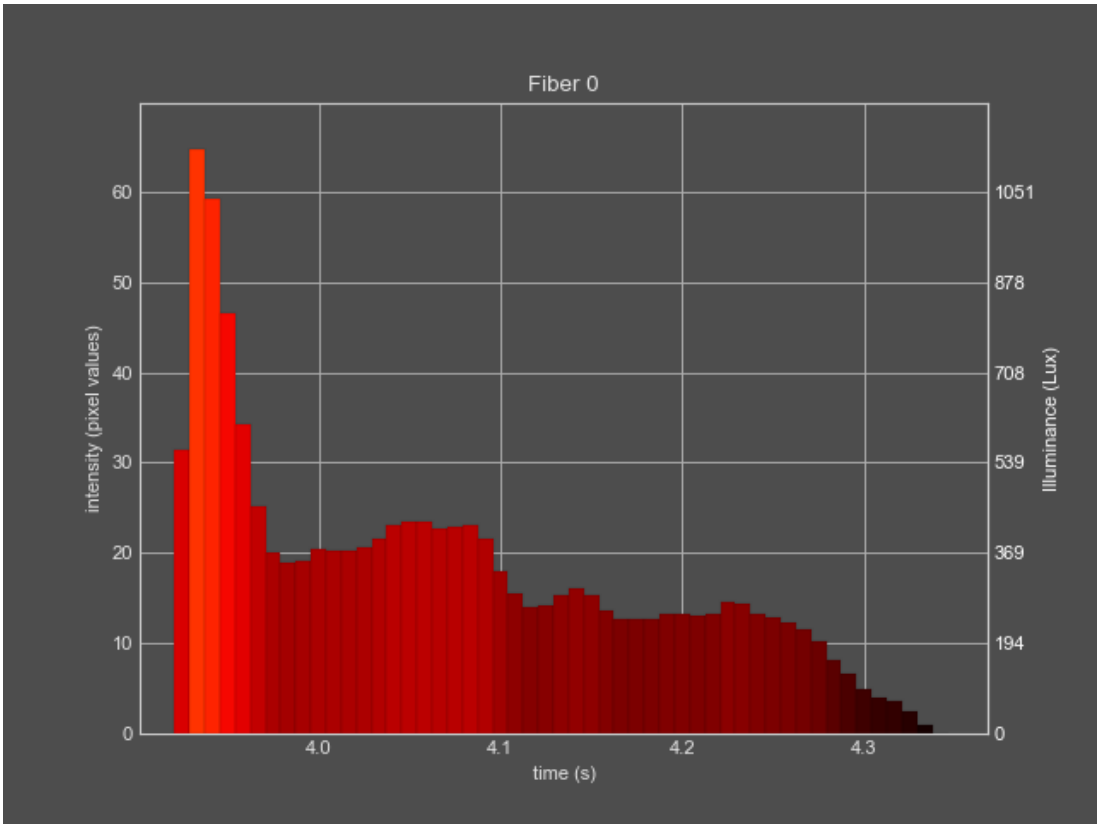


Figure 235: Selectivity/specificity results: grain size (composition: N160) "coarse" test 4 (2)

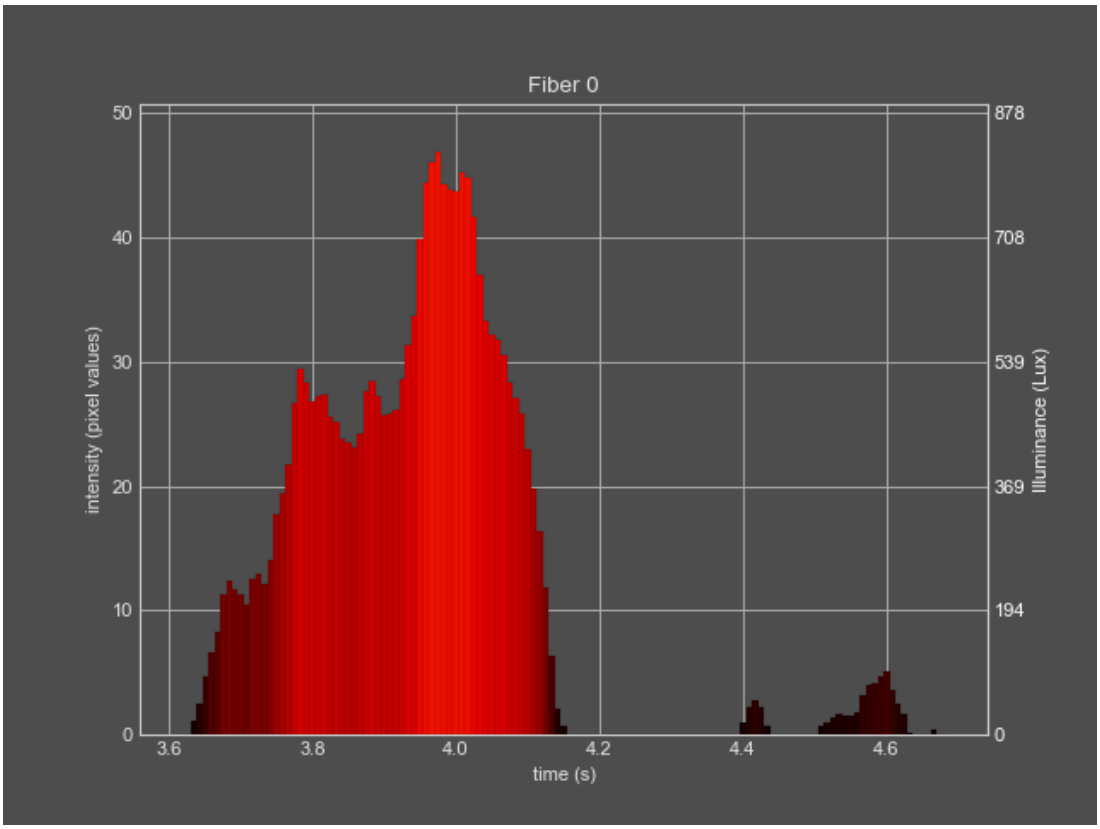


Figure 236: Selectivity/specificity results: grain size (composition: N160) "coarse" test 5 (2)

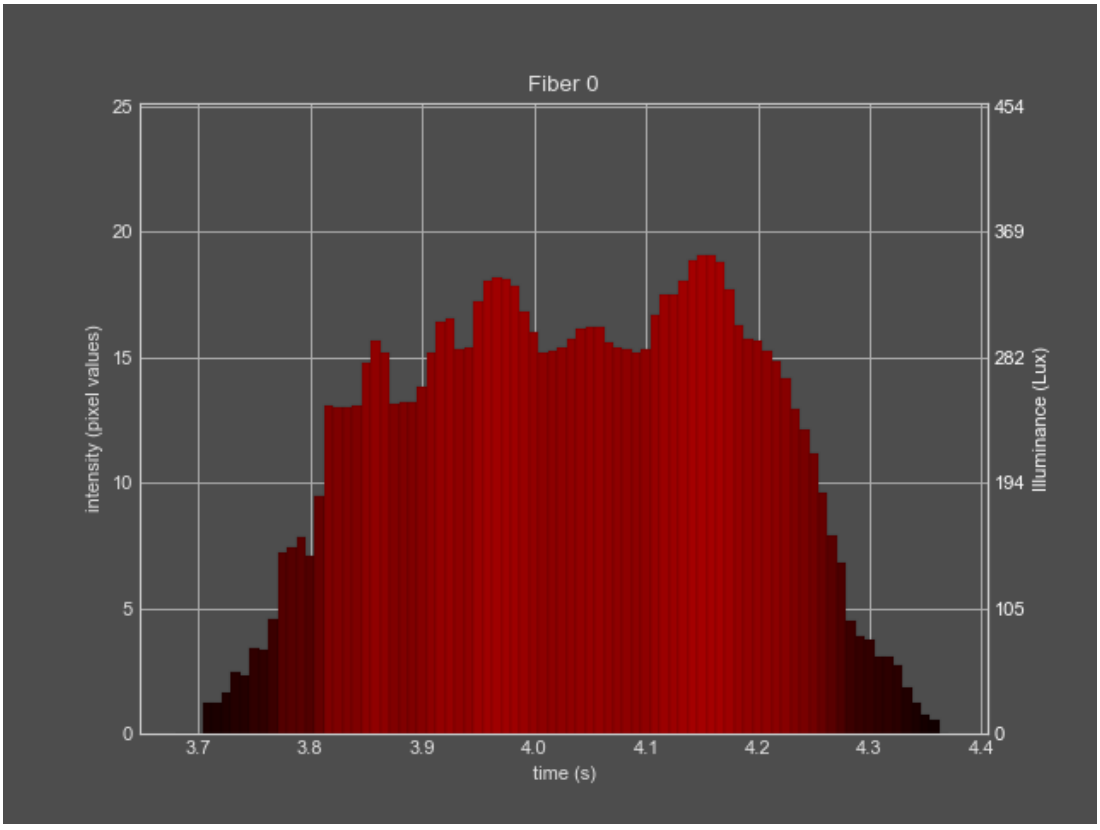


Figure 237: Selectivity/specificity results: grain size (composition: N160) "coarse" test 6 (2)

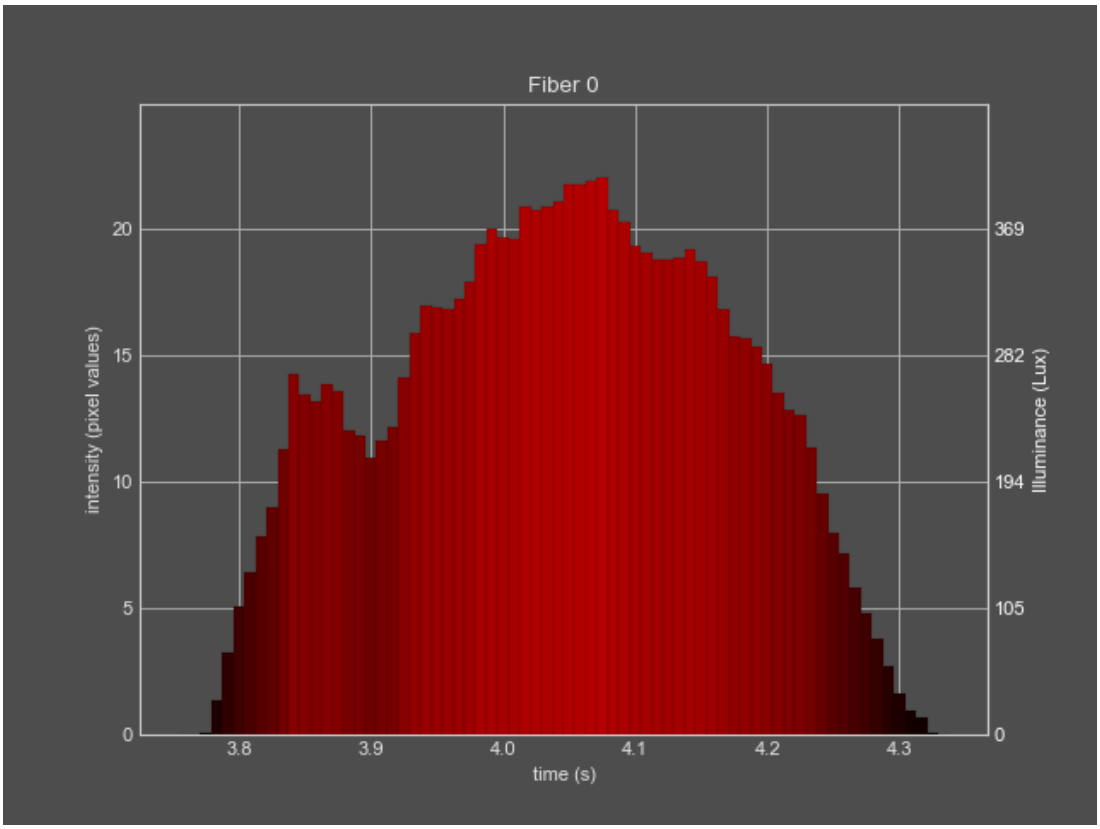


Figure 238: Selectivity/specificity results: grain size (composition: N160) "coarse" test 7 (2)

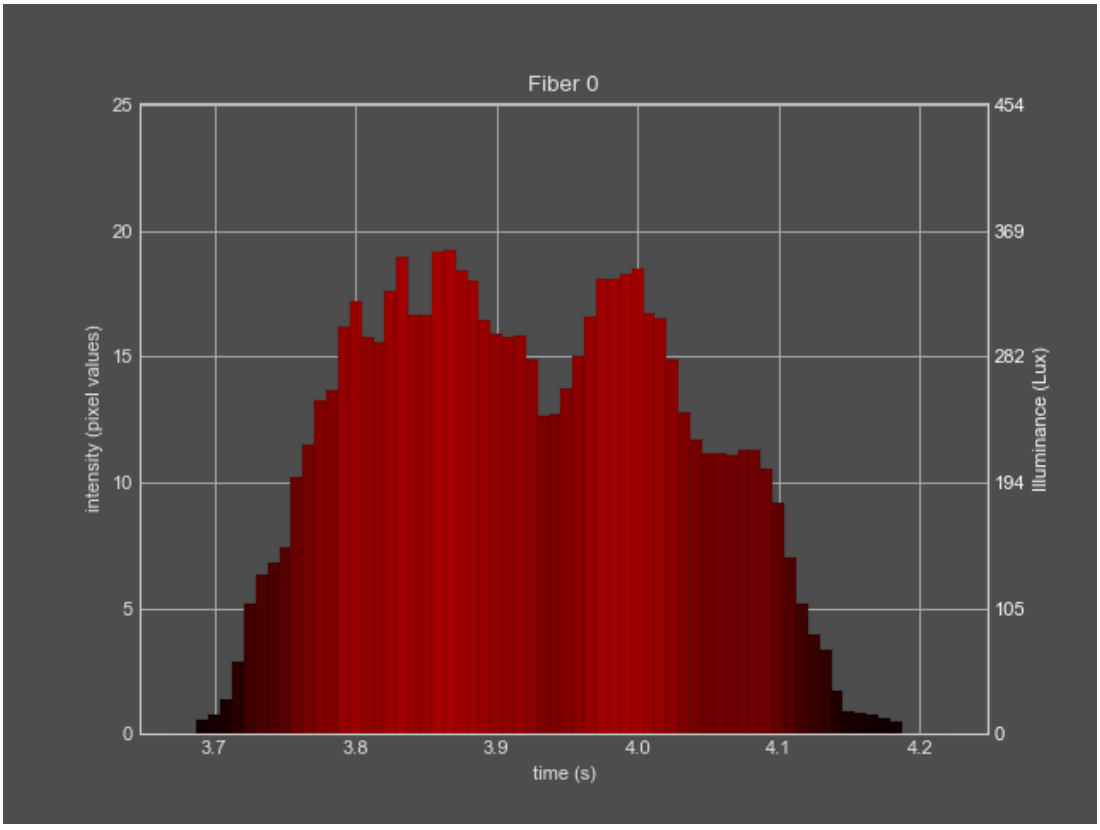


Figure 239: Selectivity/specificity results: grain size (composition: N160) "coarse" test 8 (2)

Appendix XIII: validation results for robustness: sample volume variations

This appendix presents the validation results of the combustions of compositions with variations in sample volume tests. The following figures show the PEMAD analysis results for each test replicate (n = 8), illustrating the combustion intensity profiles over time and the corresponding color behavior for both compositions.

XIII.I Robustness: starting volume 90% of 0.134cm^3

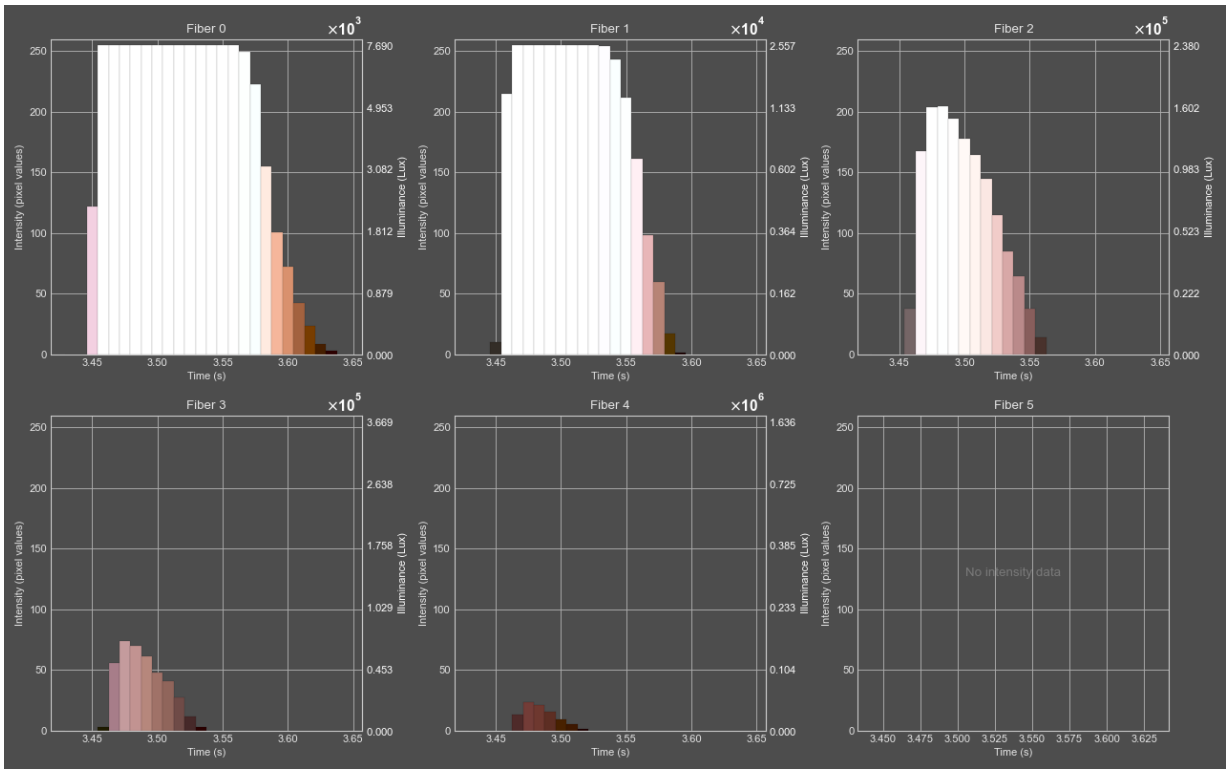


Figure 240: Robustness results: starting volume 90% of 0.134cm^3 test 1

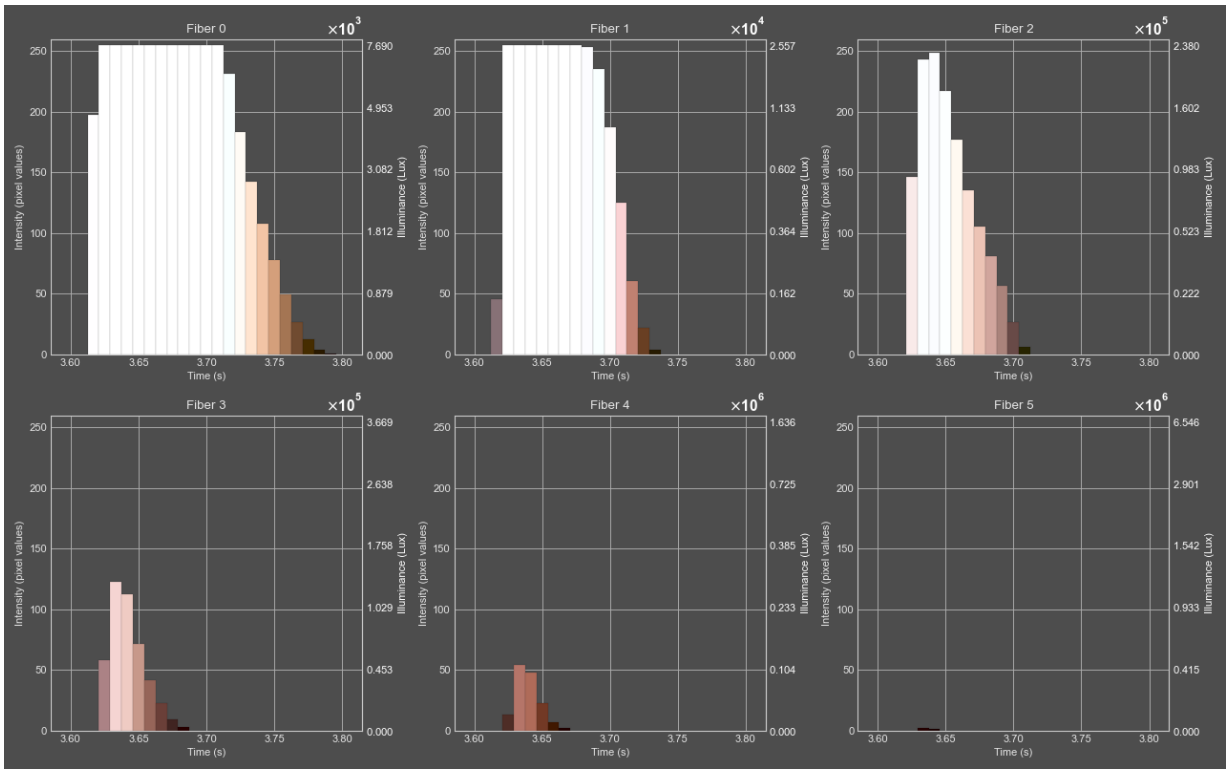


Figure 241: Robustness results: starting volume 90% of 0.134cm^3 test 2

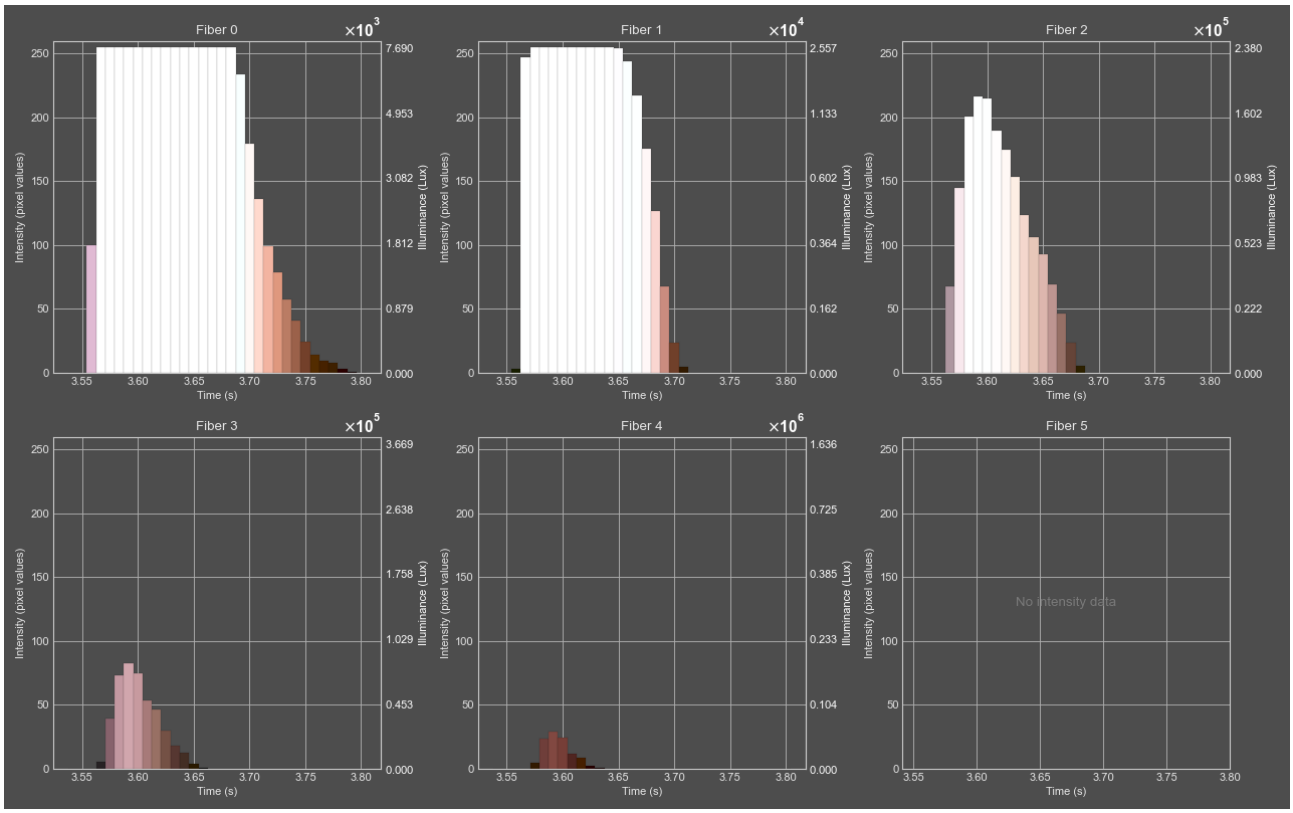


Figure 242: Robustness results: starting volume 90% of 0.134cm^3 test 3

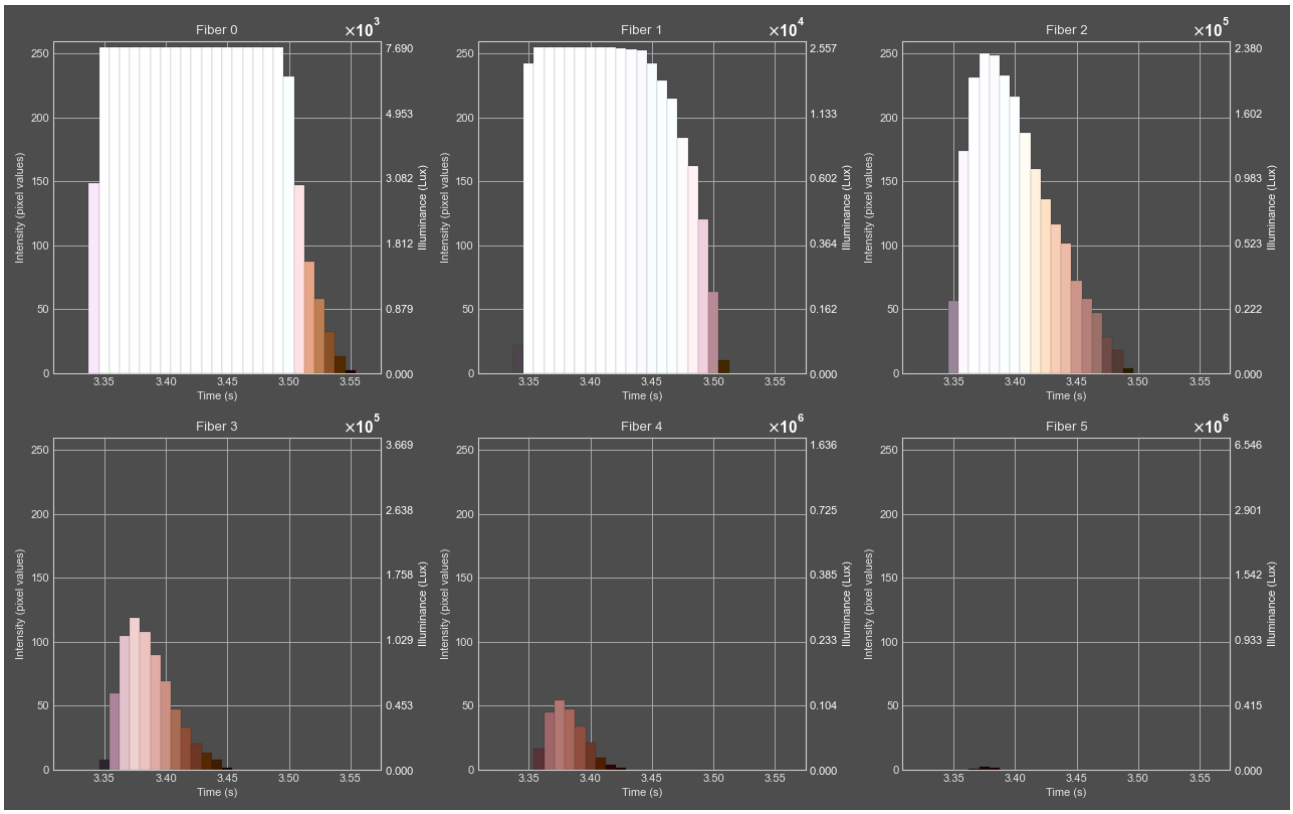


Figure 243: Robustness results: starting volume 90% of 0.134cm^3 test 4

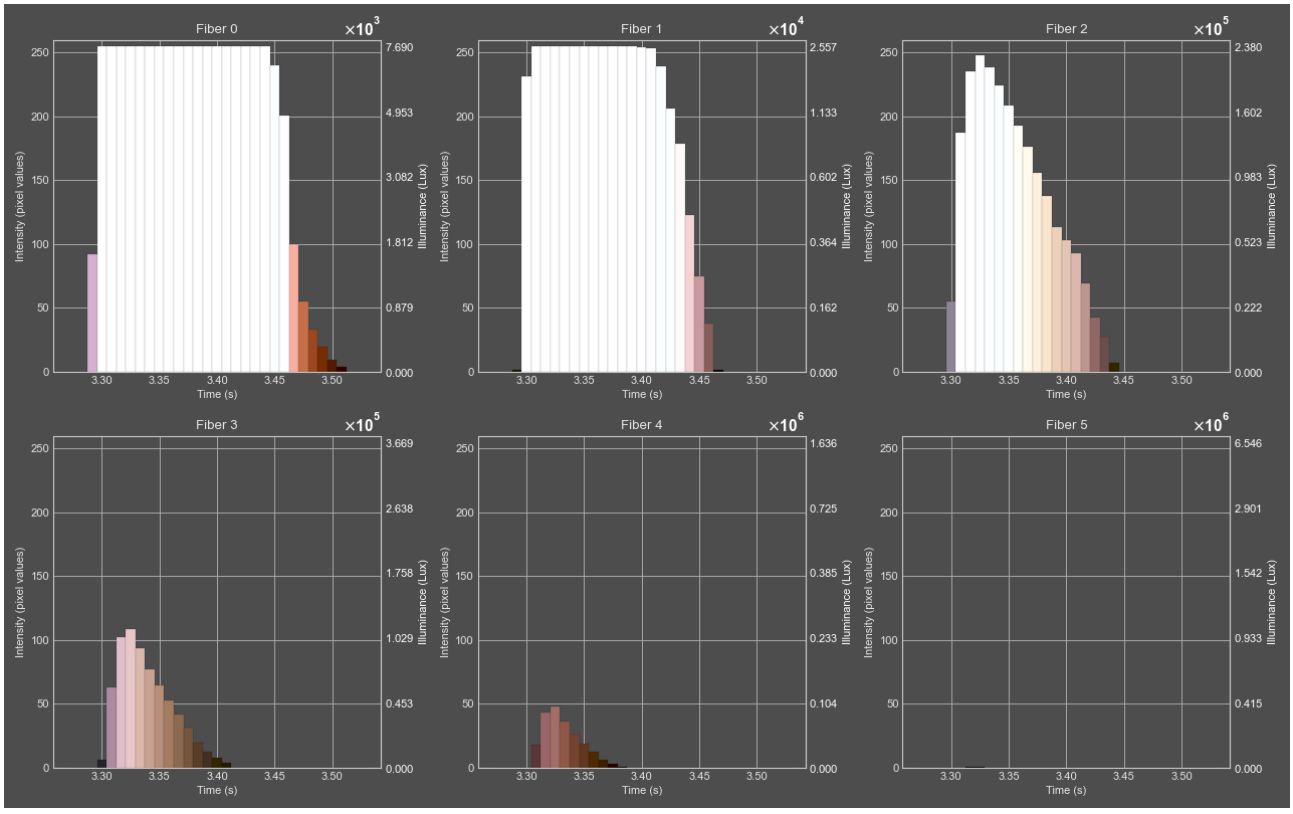


Figure 244: Robustness results: starting volume 90% of 0.134cm^3 test 5

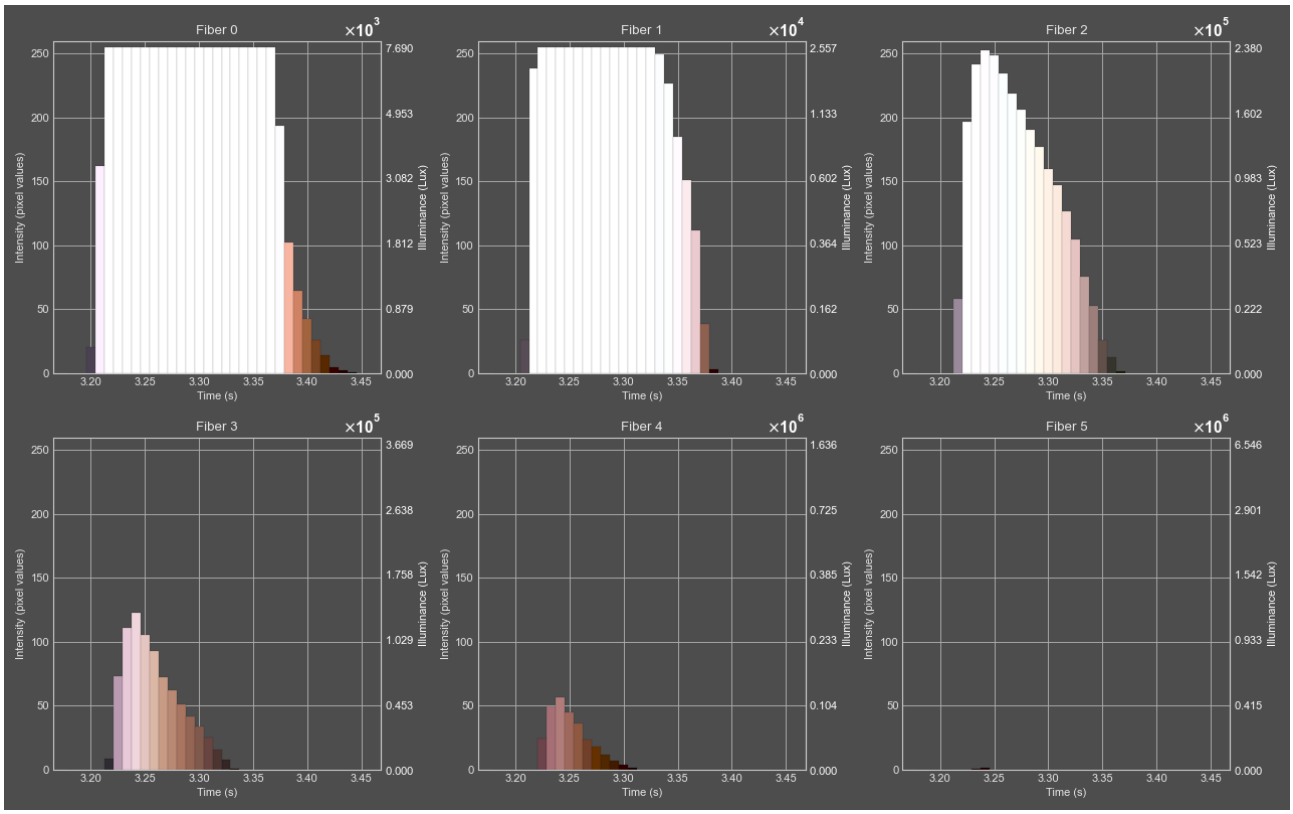


Figure 245: Robustness results: starting volume 90% of 0.134cm^3 test 6

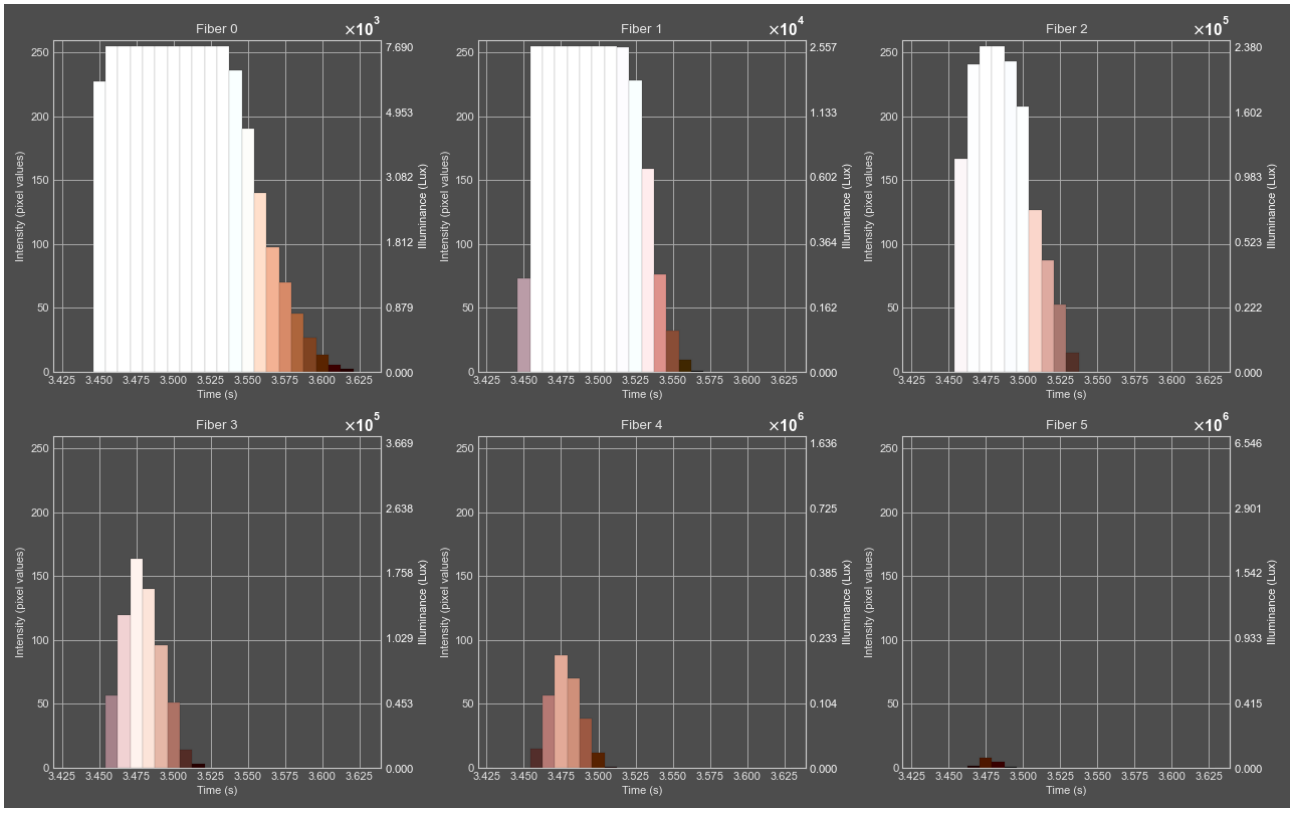


Figure 246: Robustness results: starting volume 90% of 0.134cm^3 test 7

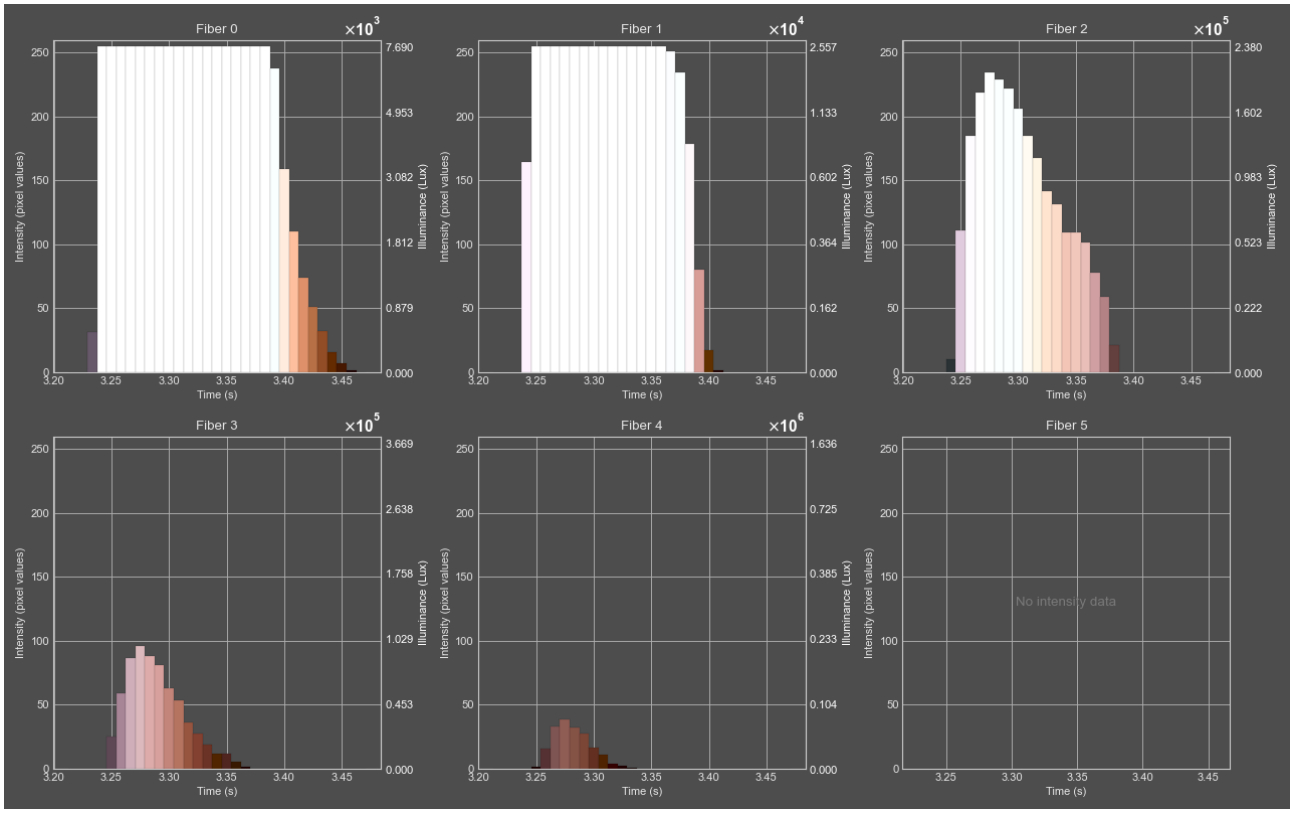


Figure 247: Robustness results: starting volume 90% of $0.134cm^3$ test 8

XIII.II Robustness: starting volume 100% of 0.134cm^3

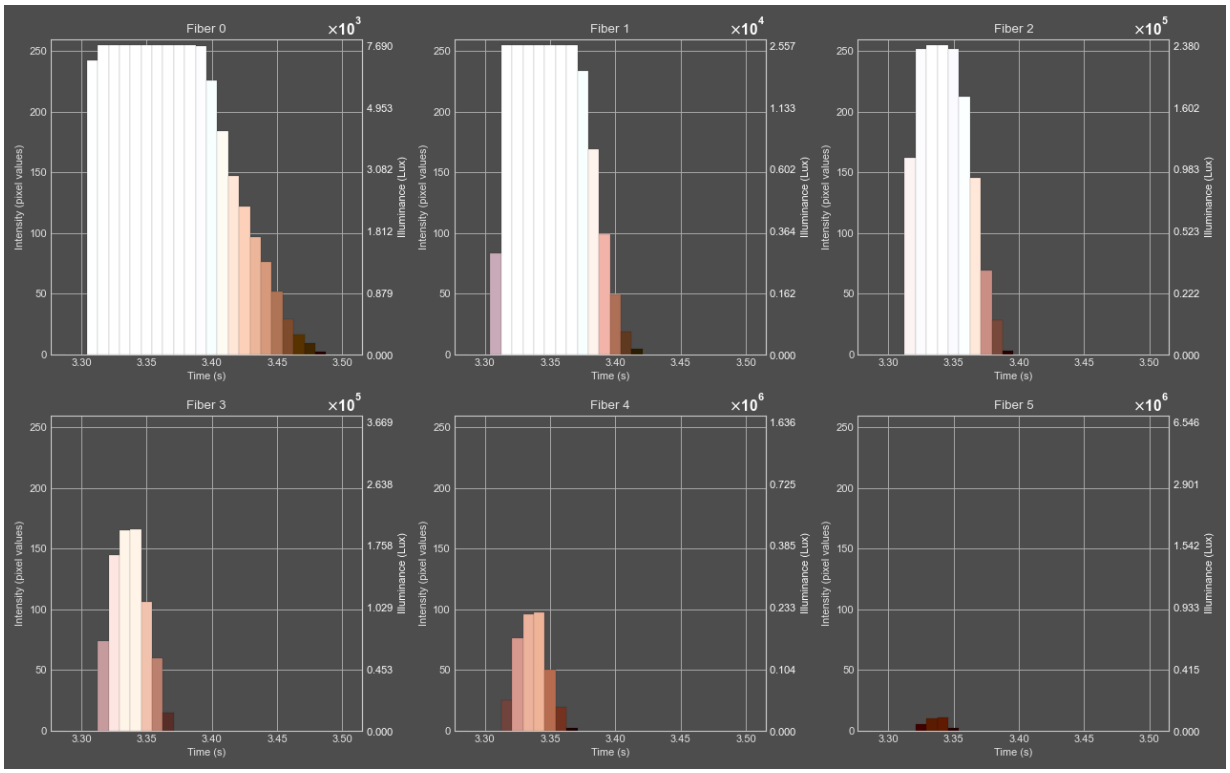


Figure 248: Robustness results: starting volume 100% of 0.134cm^3 test 1

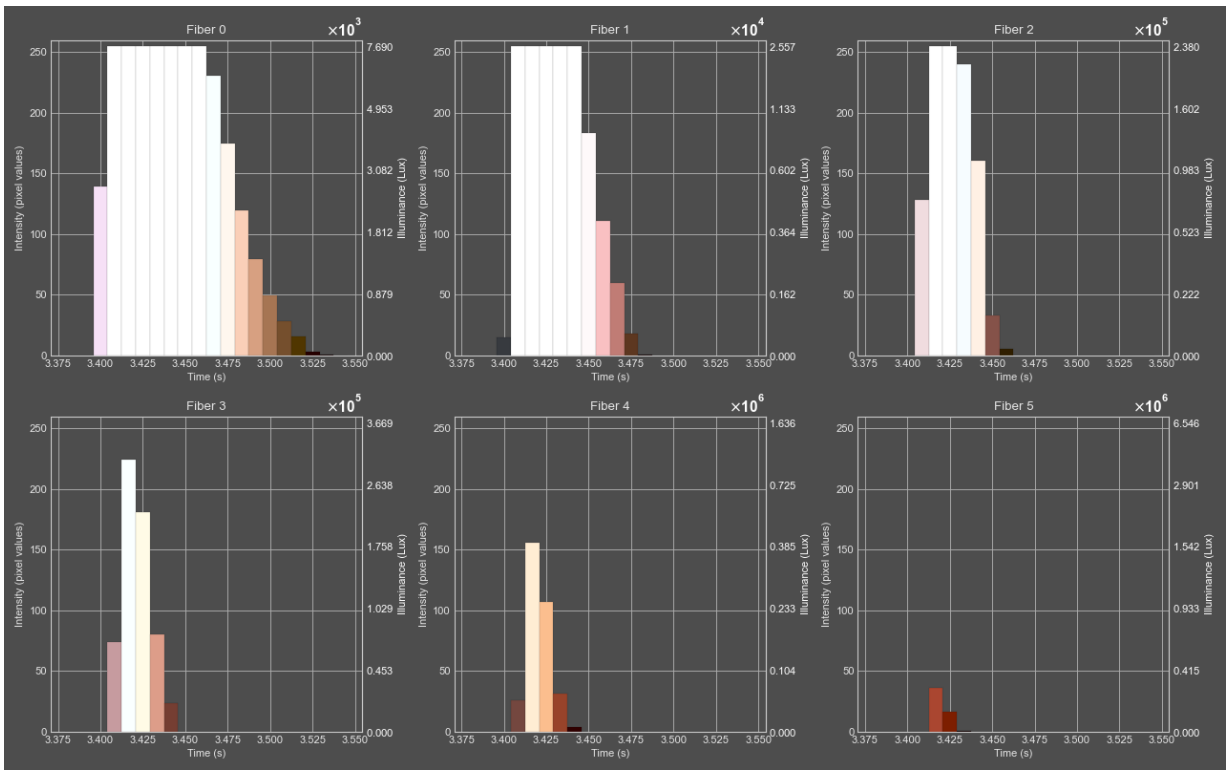


Figure 249: Robustness results: starting volume 100% of 0.134cm^3 test 2

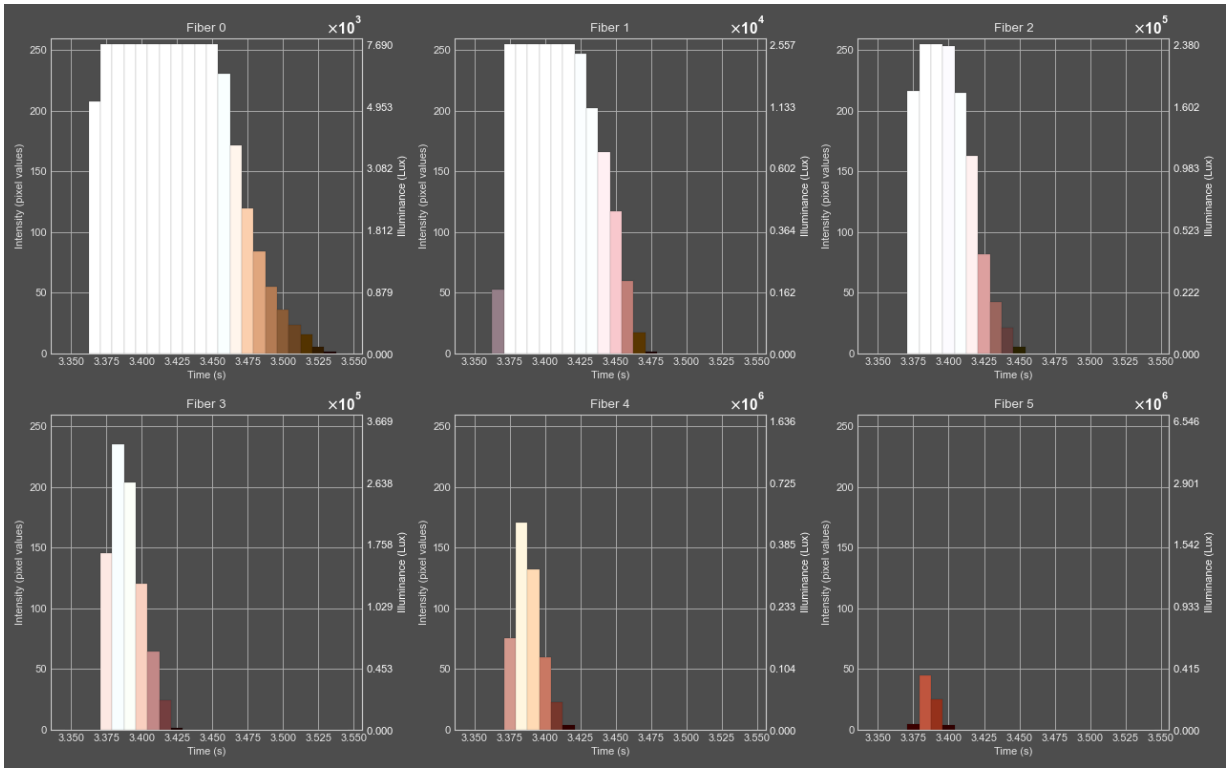


Figure 250: Robustness results: starting volume 100% of 0.134cm^3 test 3

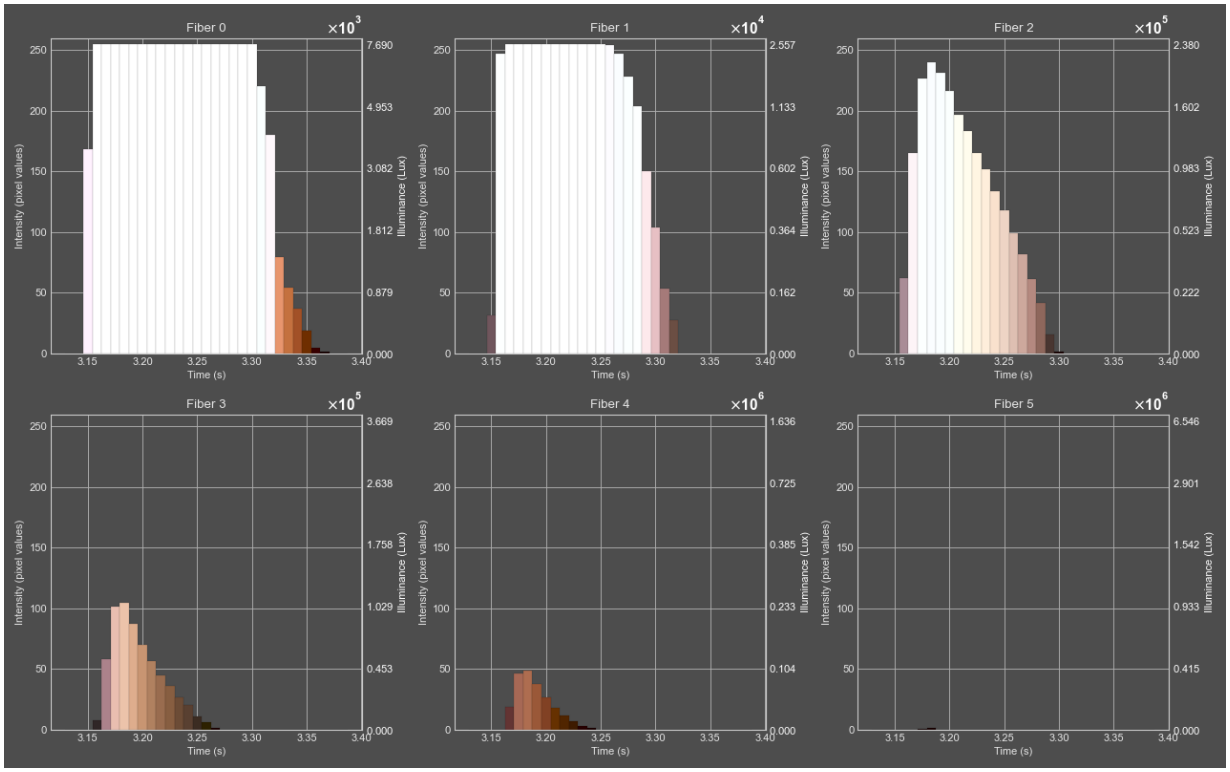


Figure 251: Robustness results: starting volume 100% of 0.134cm^3 test 4

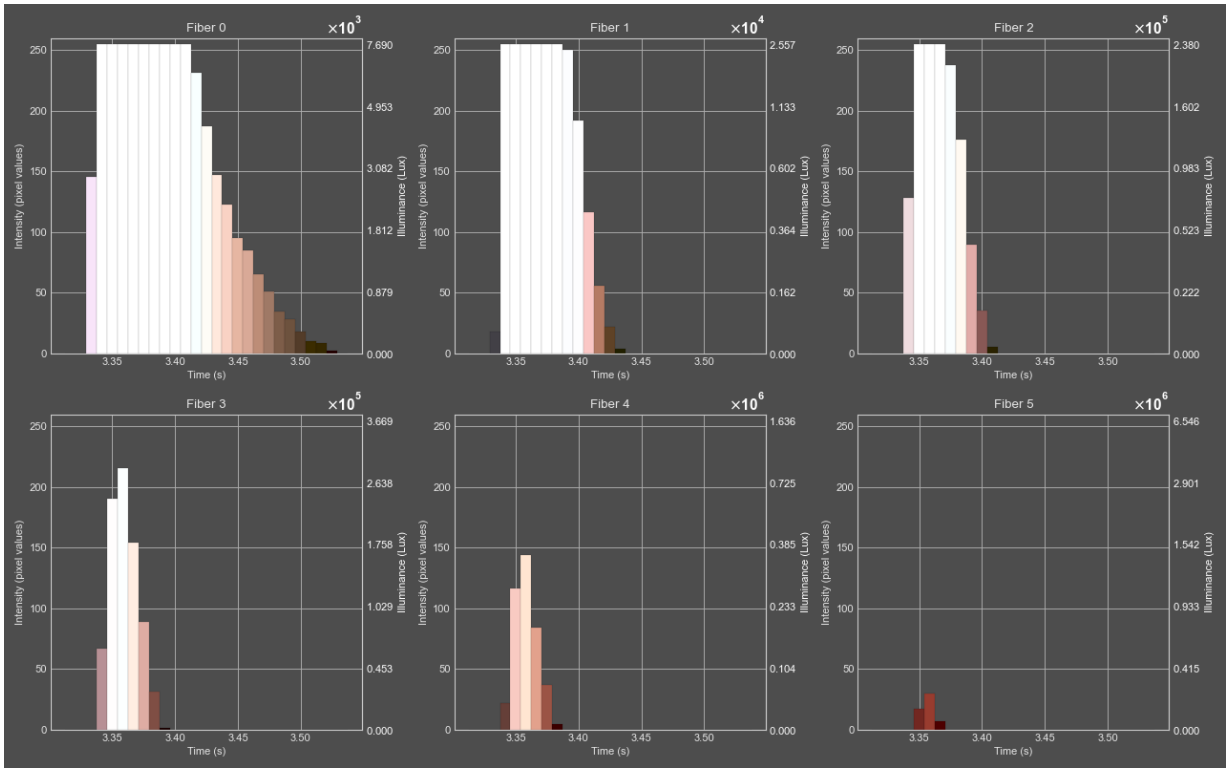


Figure 252: Robustness results: starting volume 100% of 0.134cm^3 test 5

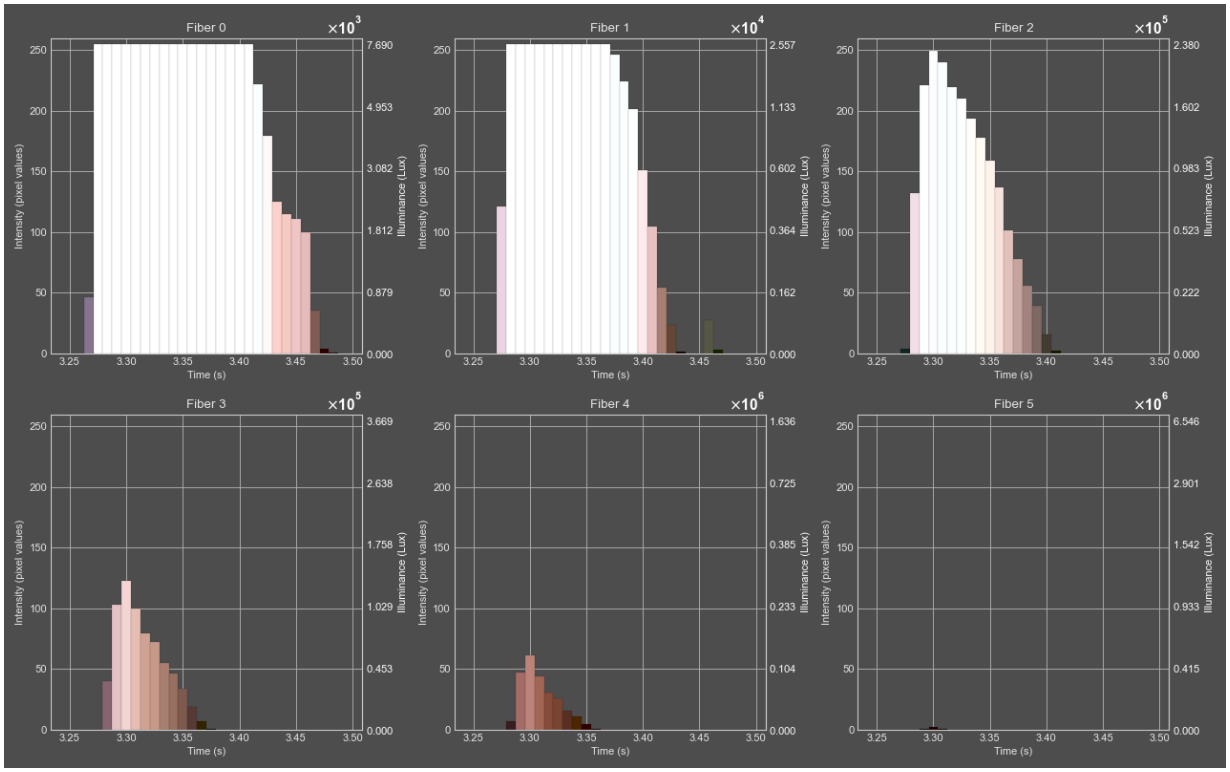


Figure 253: Robustness results: starting volume 100% of 0.134cm^3 test 6

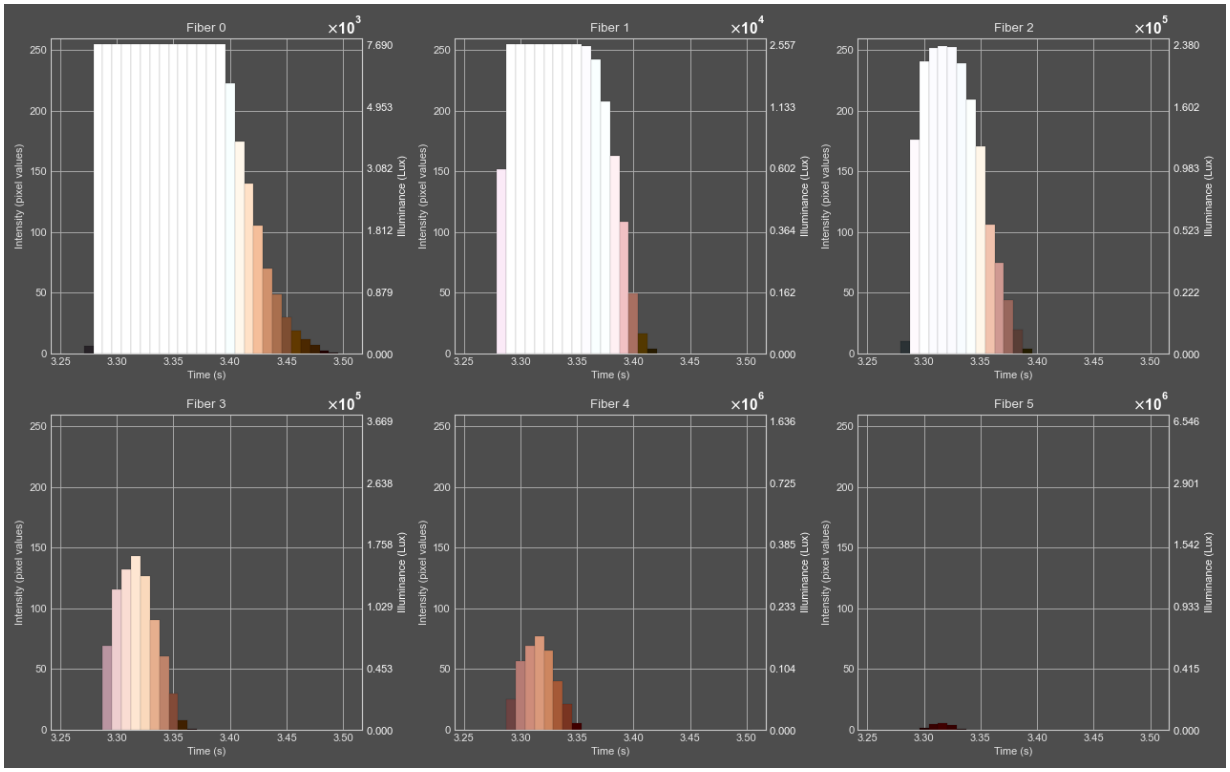


Figure 254: Robustness results: starting volume 100% of 0.134cm^3 test 7

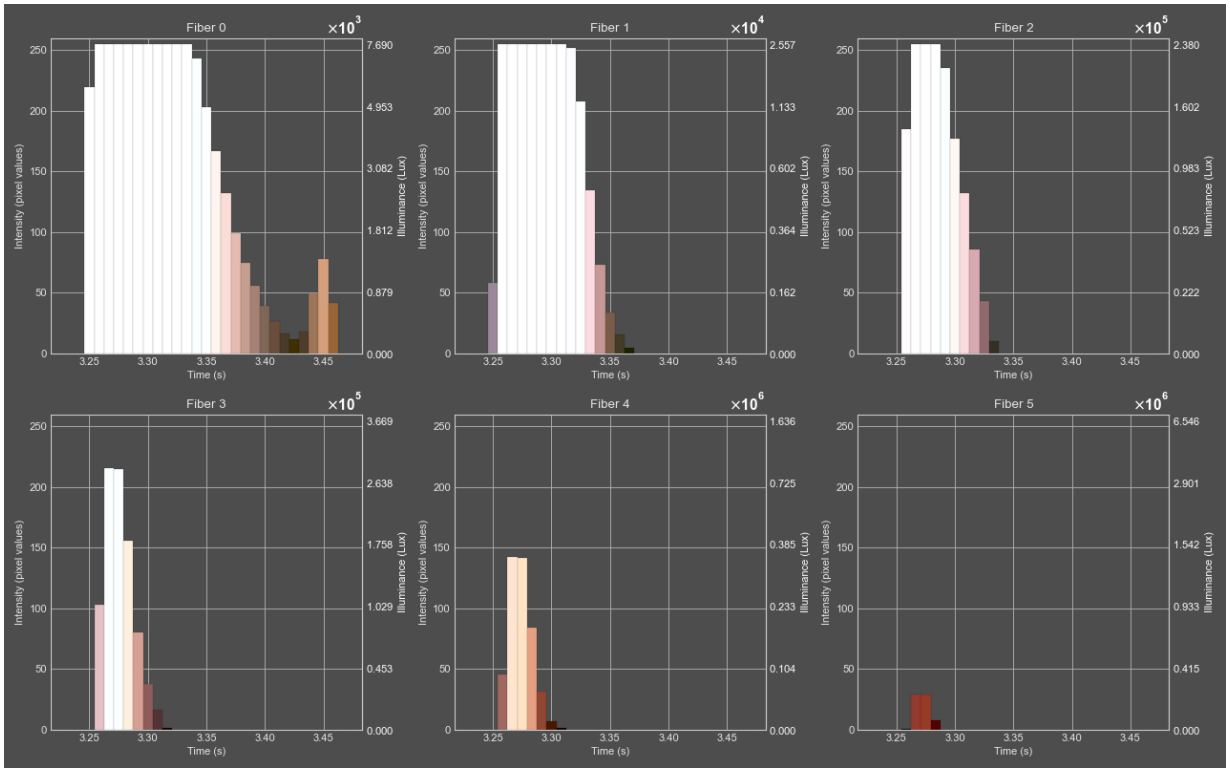


Figure 255: Robustness results: starting volume 100% of 0.134cm^3 test 8

XIII.III Robustness: starting volume 110% of 0.134cm^3

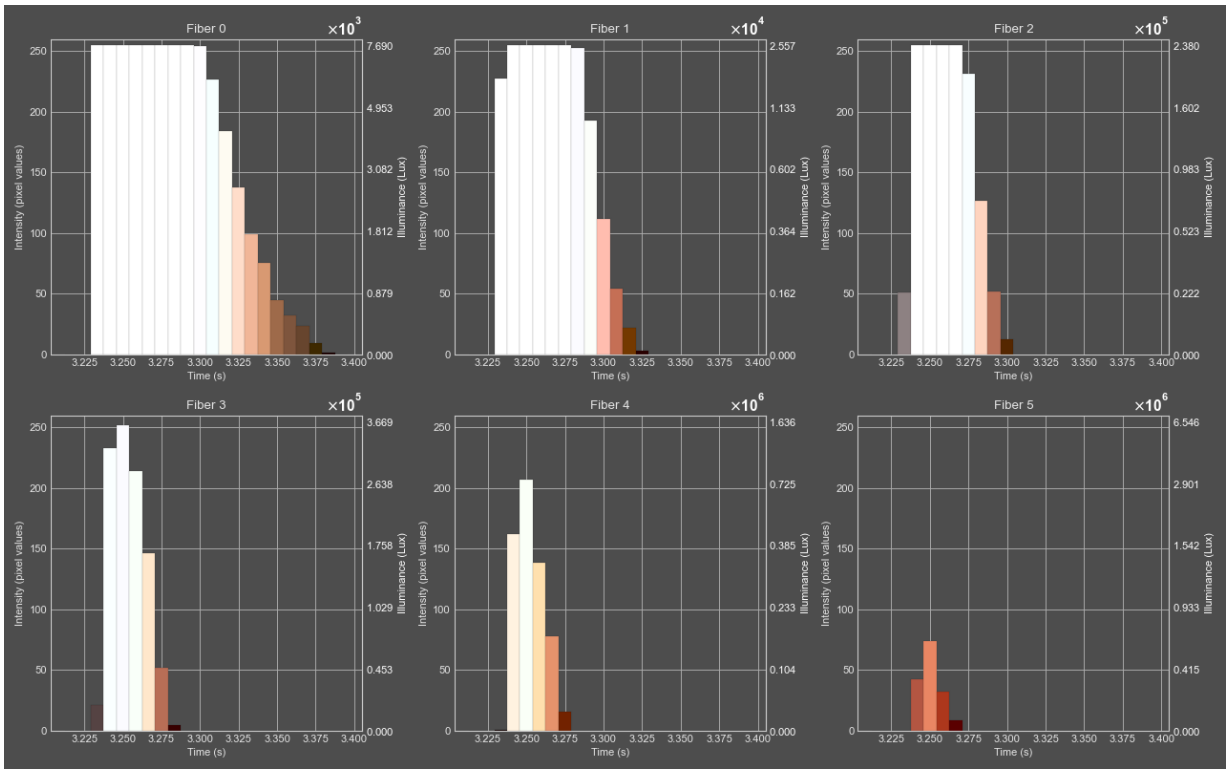


Figure 256: Robustness results: starting volume 110% of 0.134cm^3 test 1

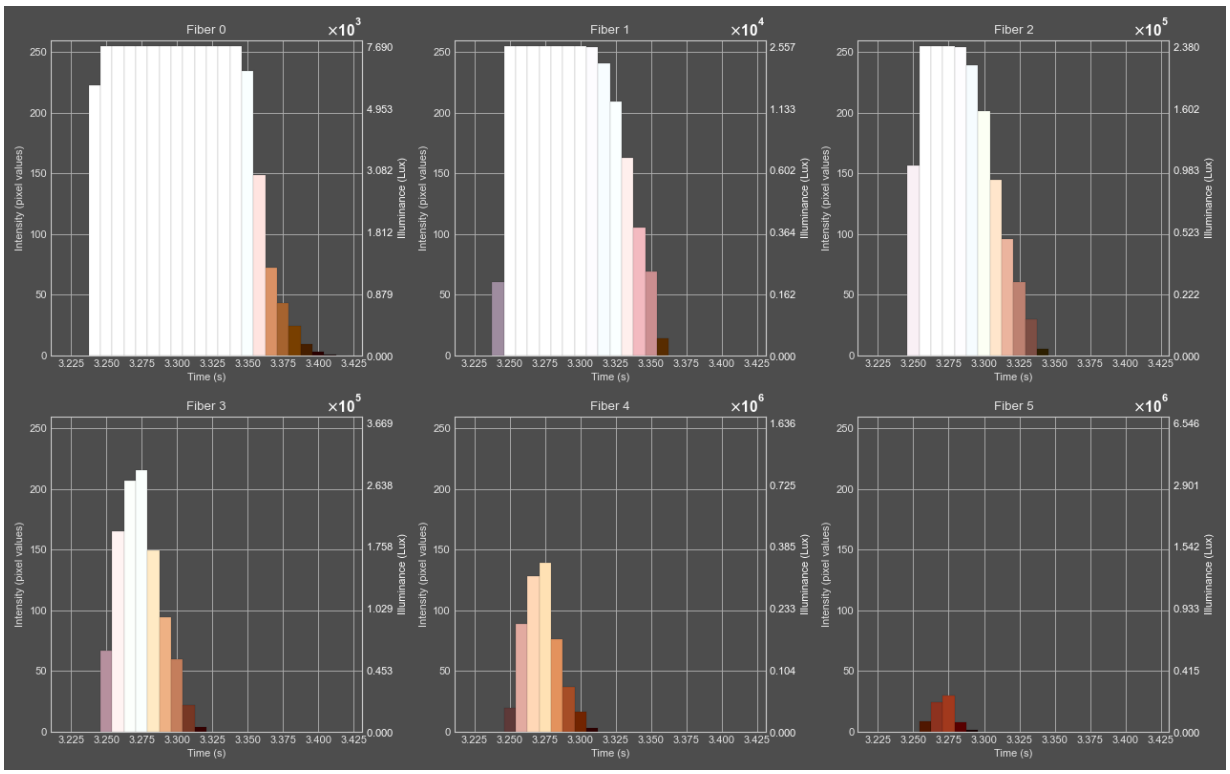


Figure 257: Robustness results: starting volume 110% of 0.134cm^3 test 2

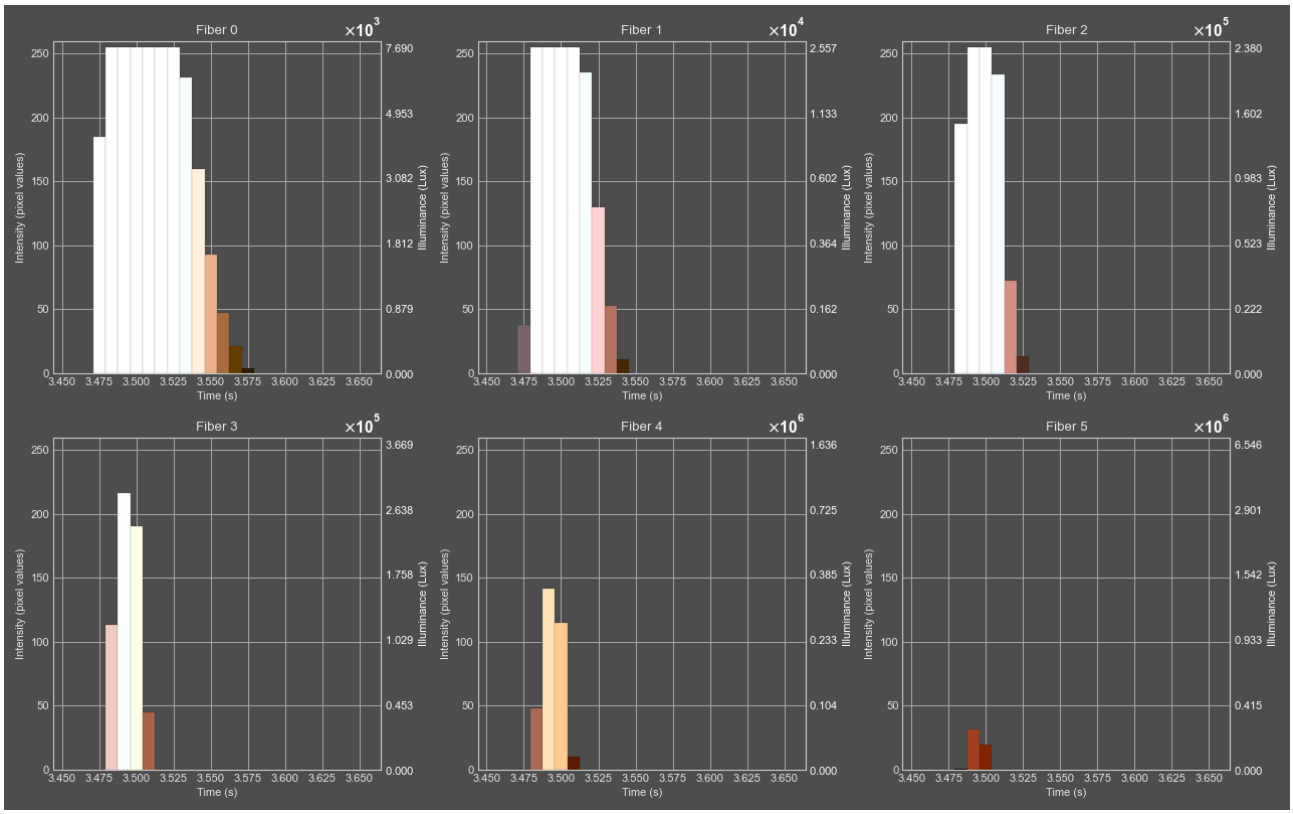


Figure 258: Robustness results: starting volume 110% of 0.134cm^3 test 3

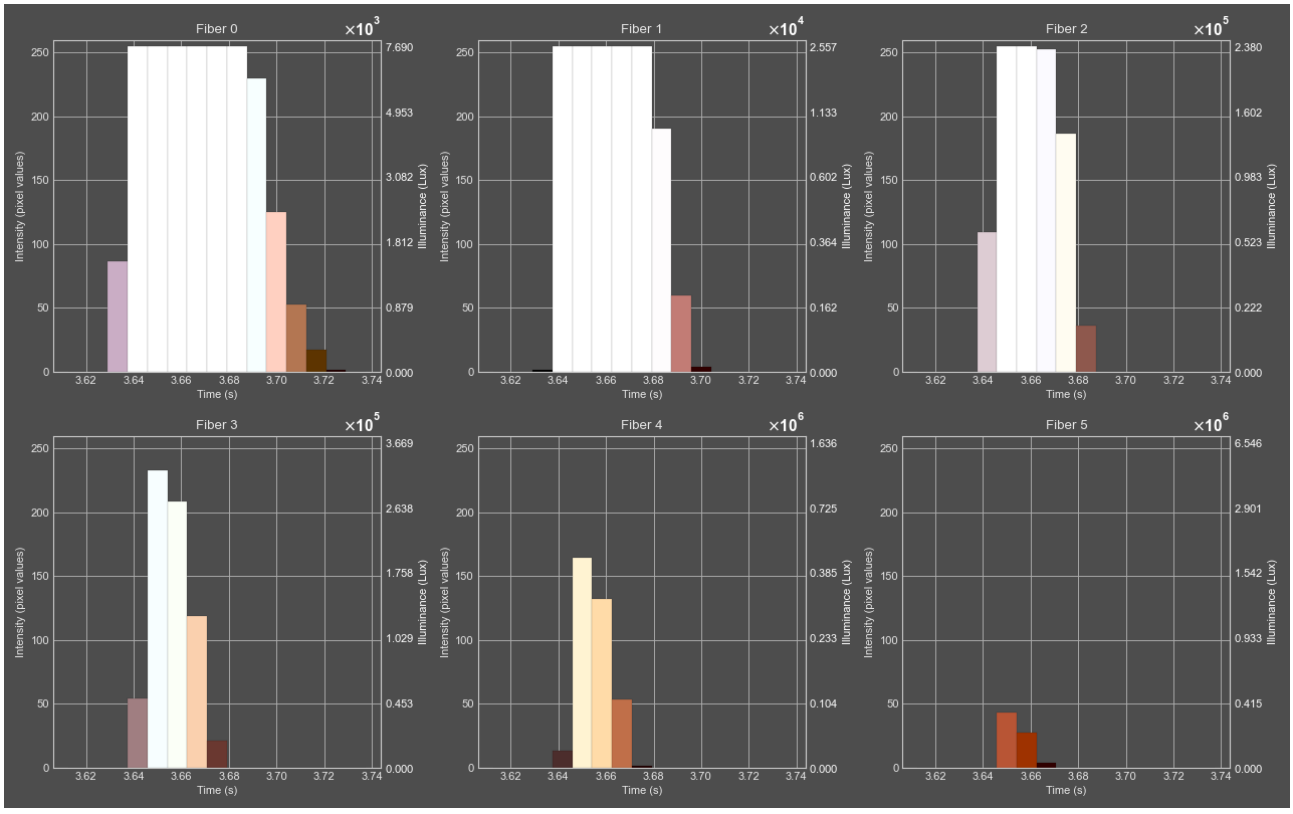


Figure 259: Robustness results: starting volume 110% of 0.134cm^3 test 4

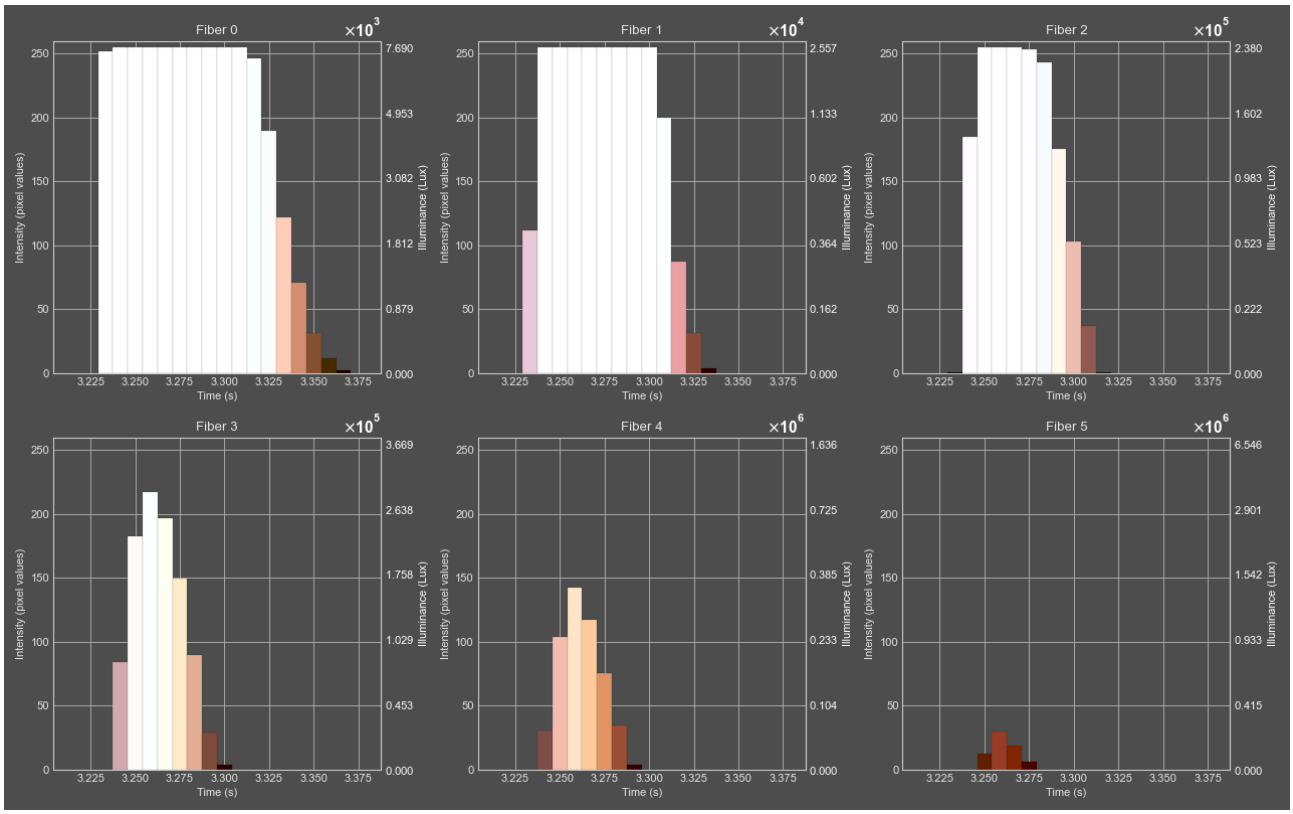


Figure 260: Robustness results: starting volume 110% of 0.134cm^3 test 5

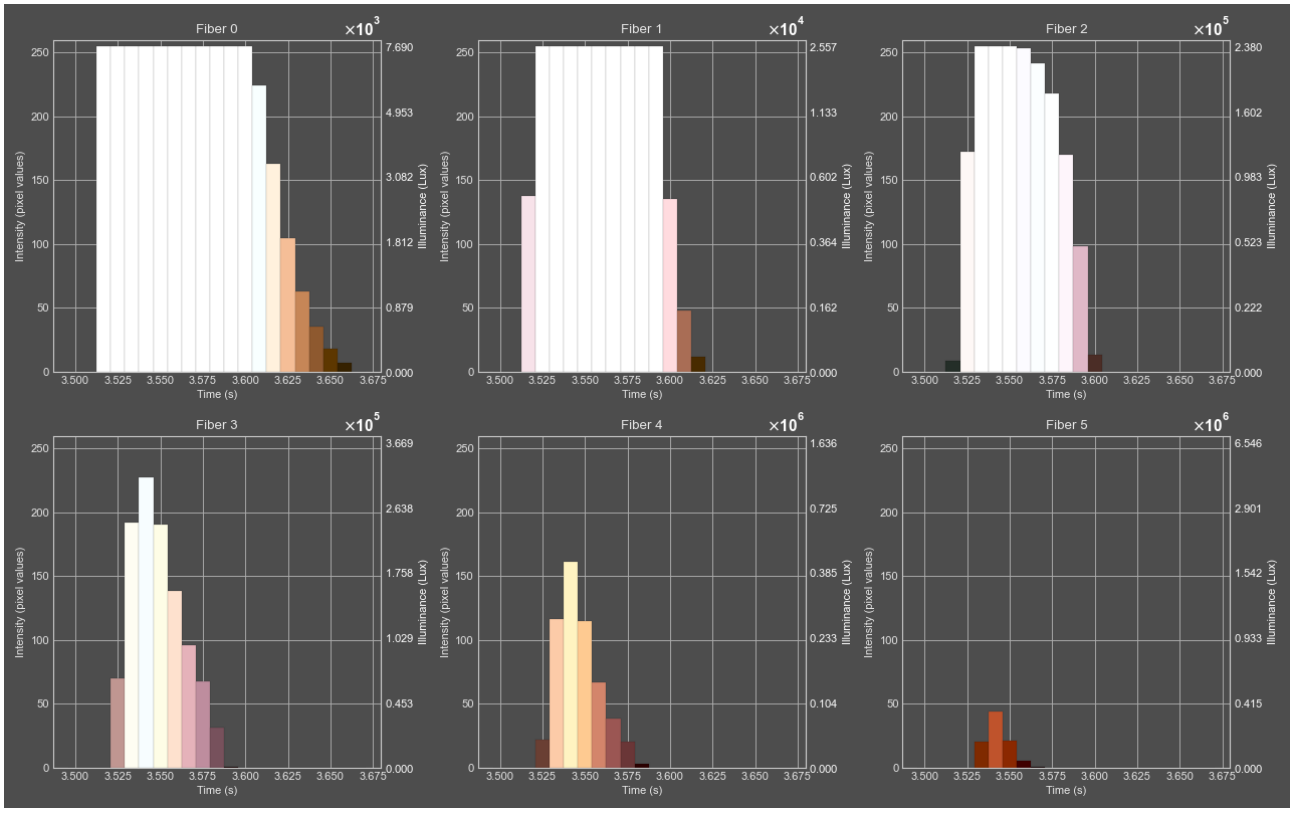


Figure 261: Robustness results: starting volume 110% of 0.134cm^3 test 6

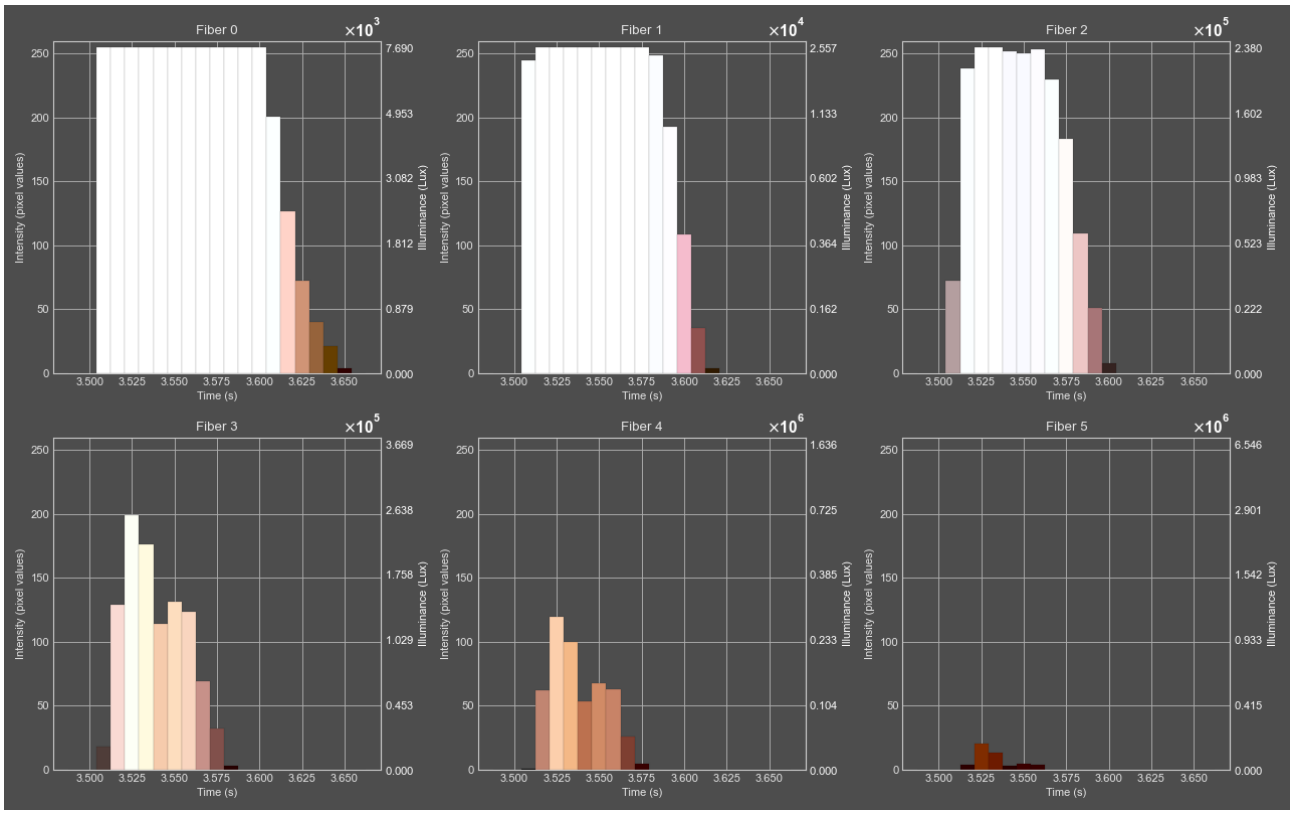


Figure 262: Robustness results: starting volume 110% of 0.134cm^3 test 7

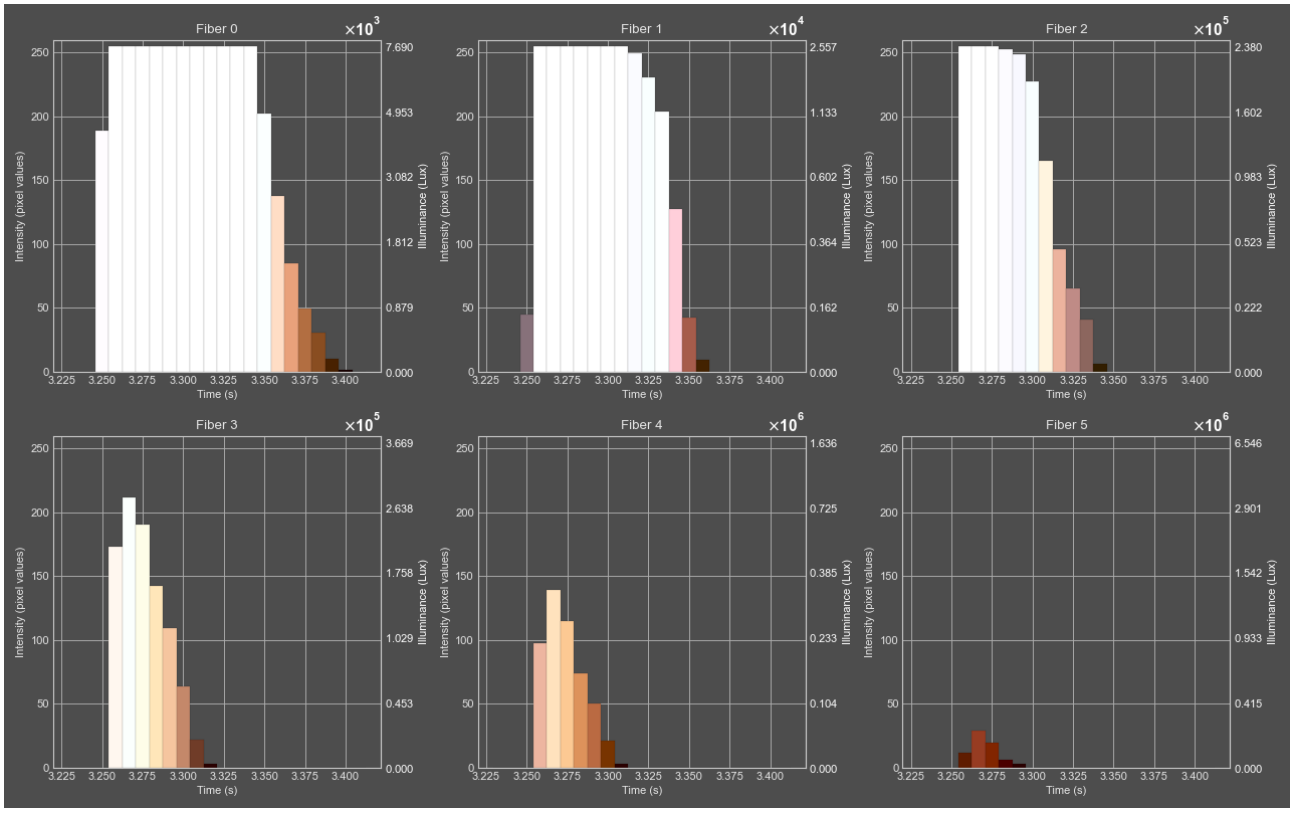


Figure 263: Robustness results: starting volume 110% of 0.134cm^3 test 8

Appendix XIV: PEMAD results for technical precision measurements

XIV.I Technical precision test results for constant white light

The color appears gray due to the large number of frames, as the shaded framework of the box gives the image a grayish tone. The second figure shows a zoomed-in view revealing the true color appearance (white).

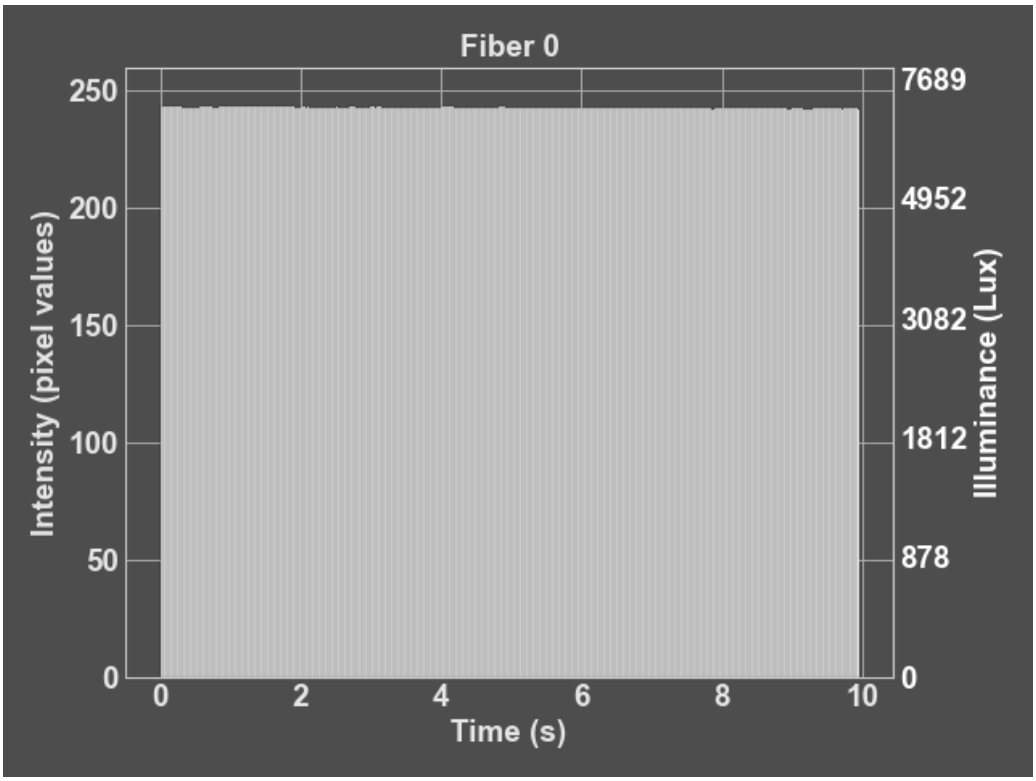


Figure 264: Technical precision test results for constant white light test 1. The color looks gray because of the number of frames, the framework of the box is shaded so it appears gray.

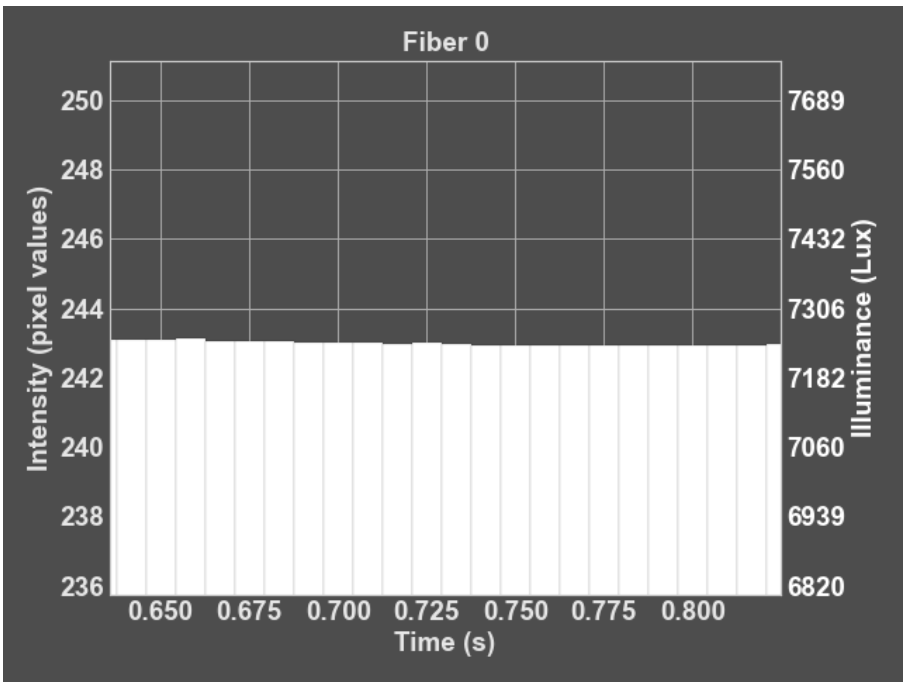


Figure 265: Technical precision test results for constant white light test 1: zoomed in view to show the true color.

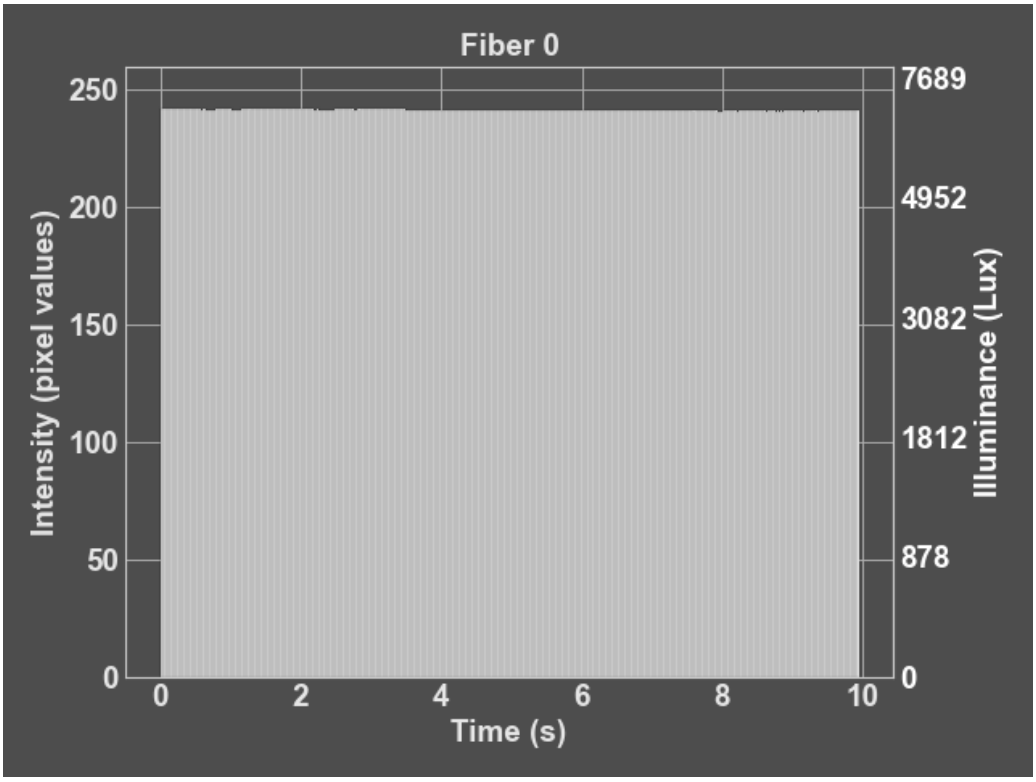


Figure 266: Technical precision test results for constant white light test 2. The color looks gray because of the number of frames, the framework of the box is shaded so it appears gray.

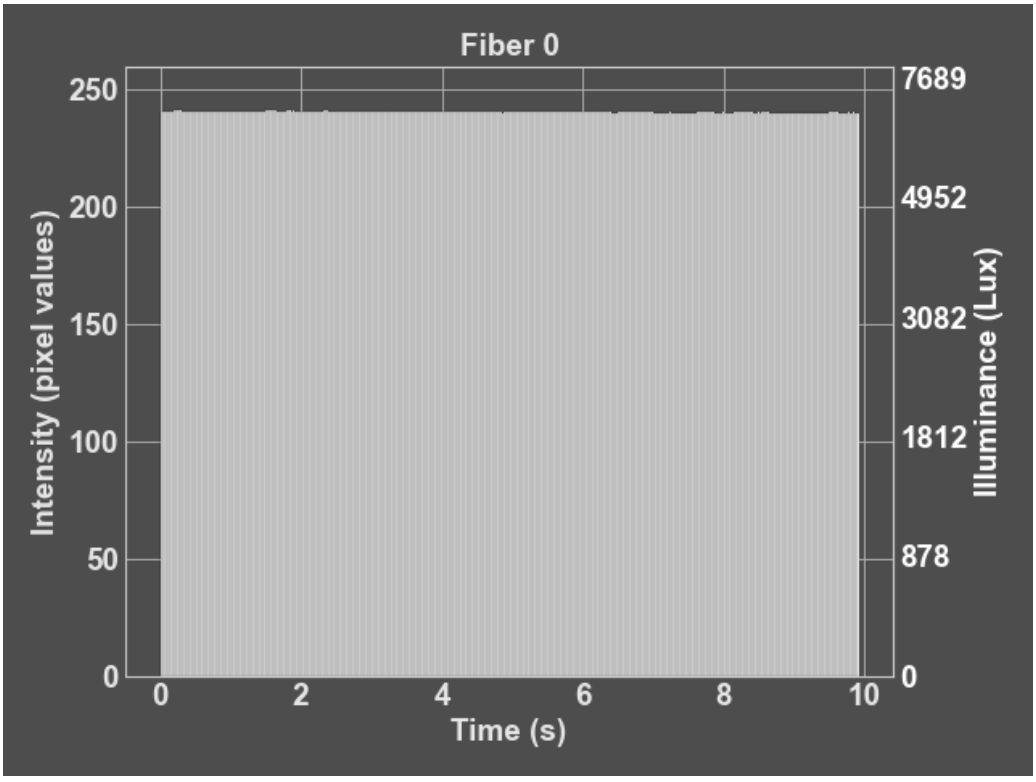


Figure 267: Technical precision test results for constant white light test 3. The color looks gray because of the number of frames, the framework of the box is shaded so it appears gray.

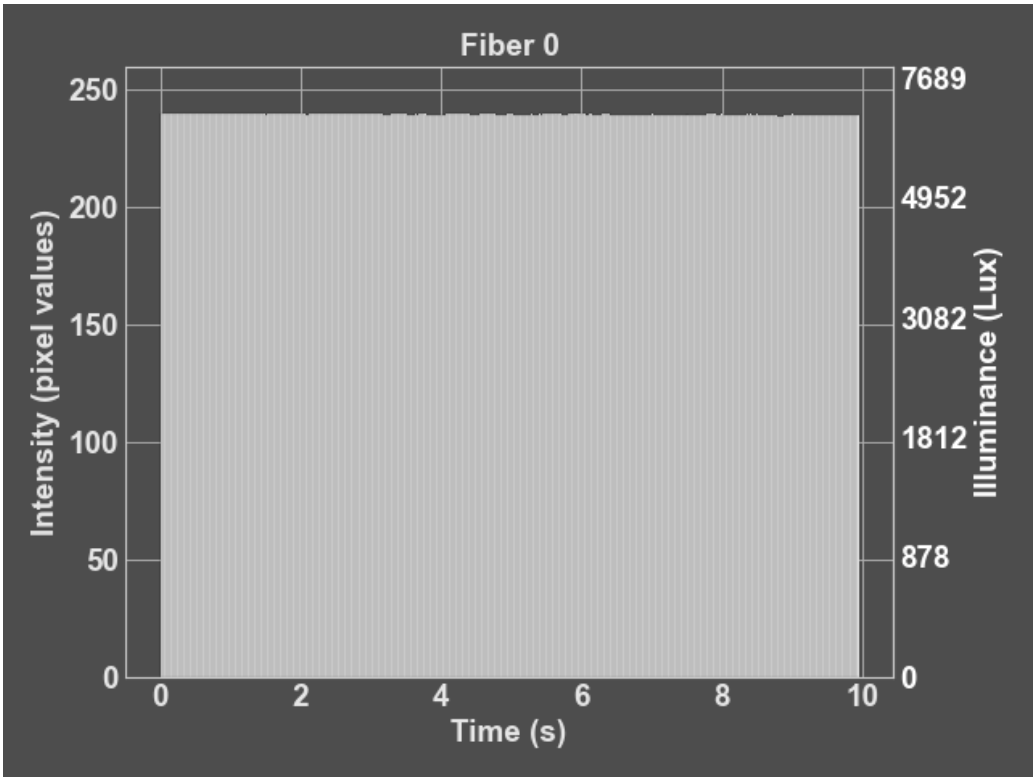


Figure 268: Technical precision test results for constant white light test 4. The color looks gray because of the number of frames; the framework of the box is shaded, so it appears gray.

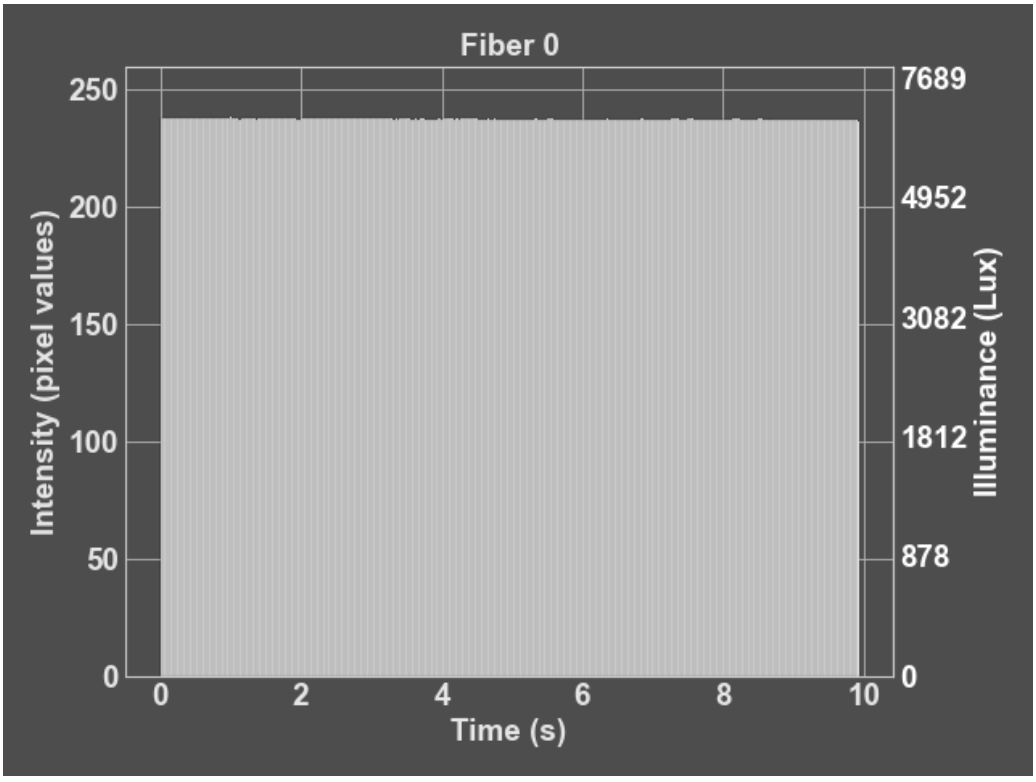


Figure 269: Technical precision test results for constant white light test 5. The color looks gray because of the number of frames; the framework of the box is shaded, so it appears gray.

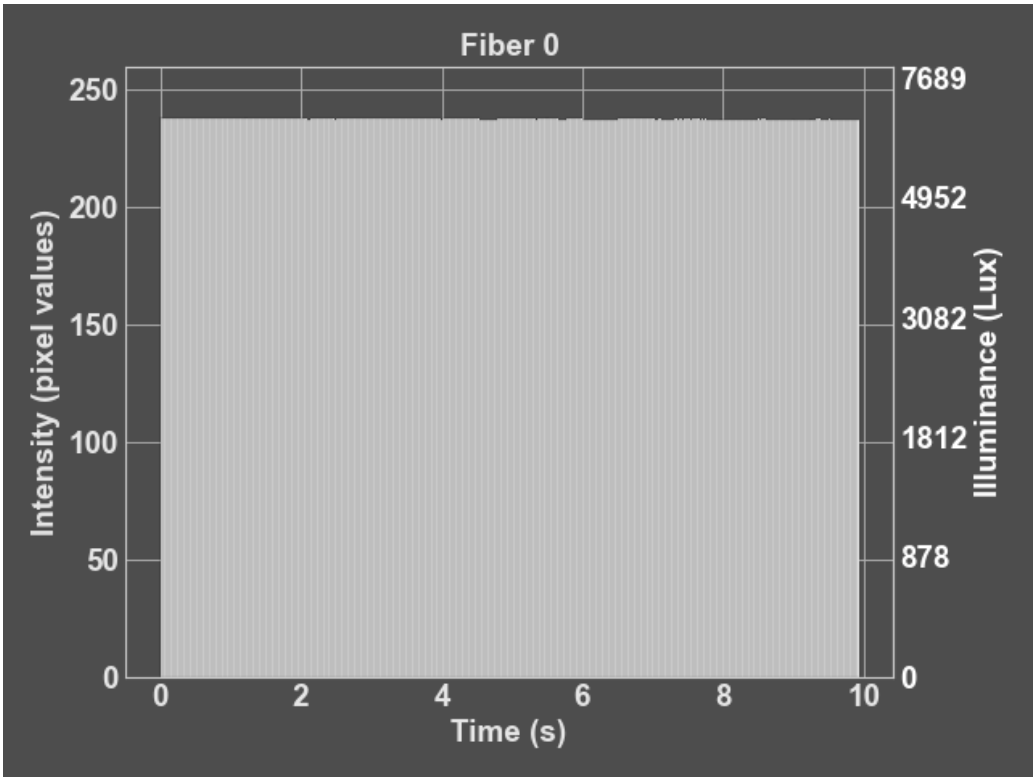


Figure 270: Technical precision test results for constant white light test 6. The color looks gray because of the number of frames; the framework of the box is shaded, so it appears gray.

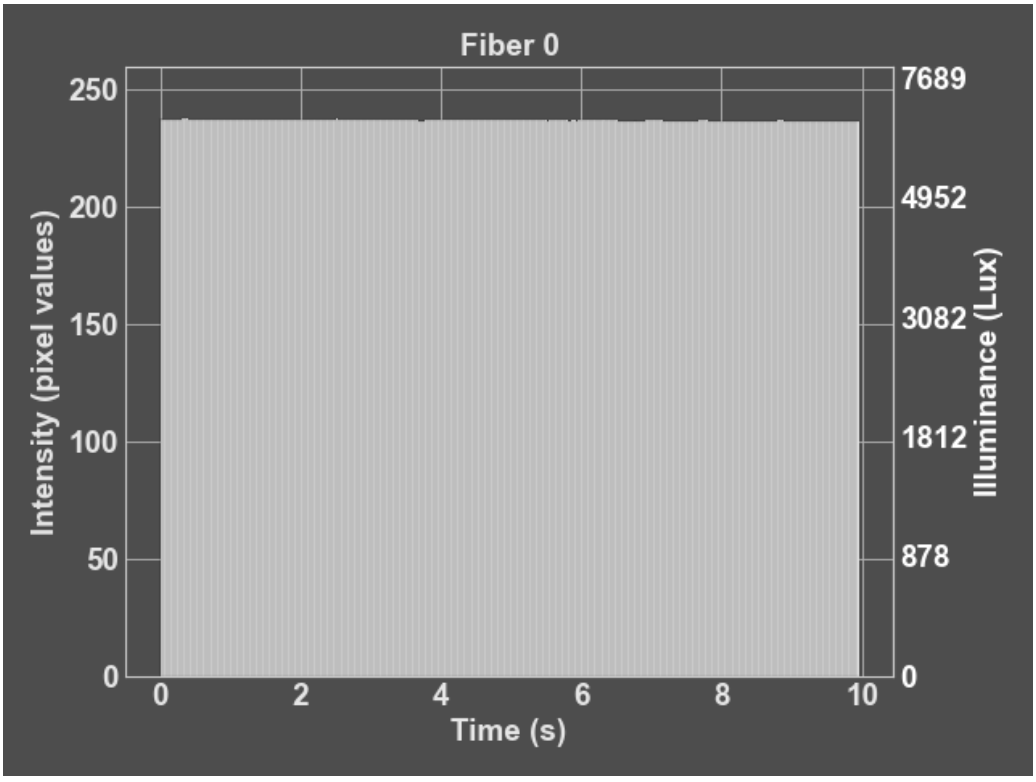


Figure 271: Technical precision test results for constant white light test 7. The color looks gray because of the number of frames; the framework of the box is shaded, so it appears gray.

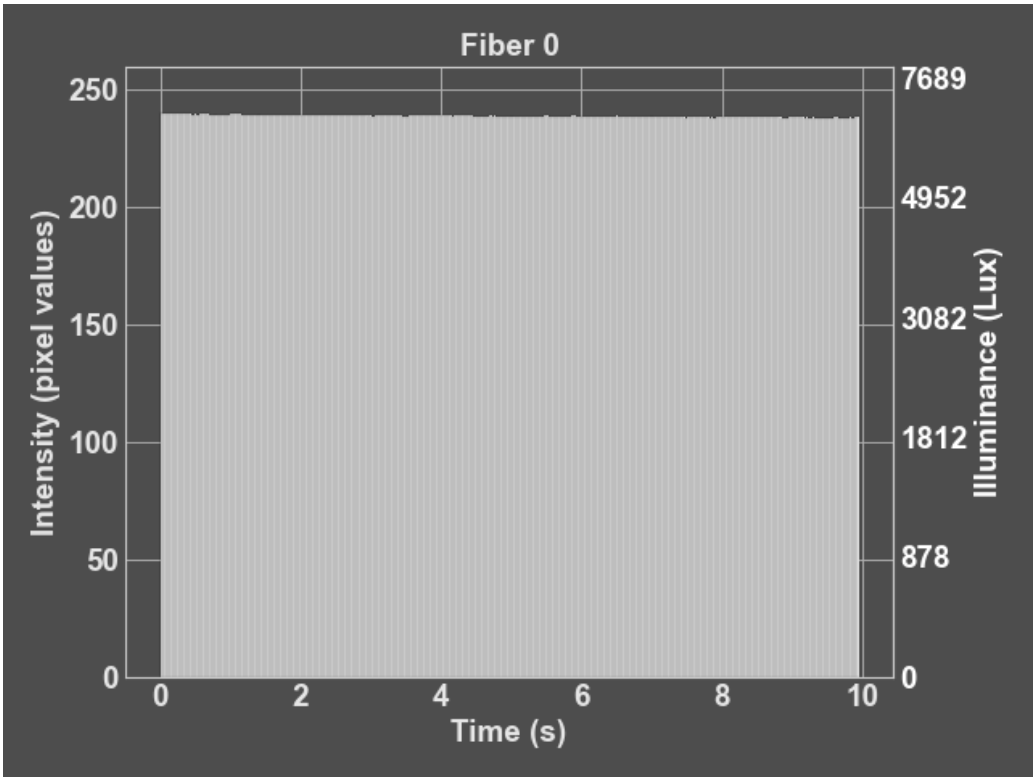


Figure 272: Technical precision test results for constant white light test 8. The color looks gray because of the number of frames; the framework of the box is shaded, so it appears gray.

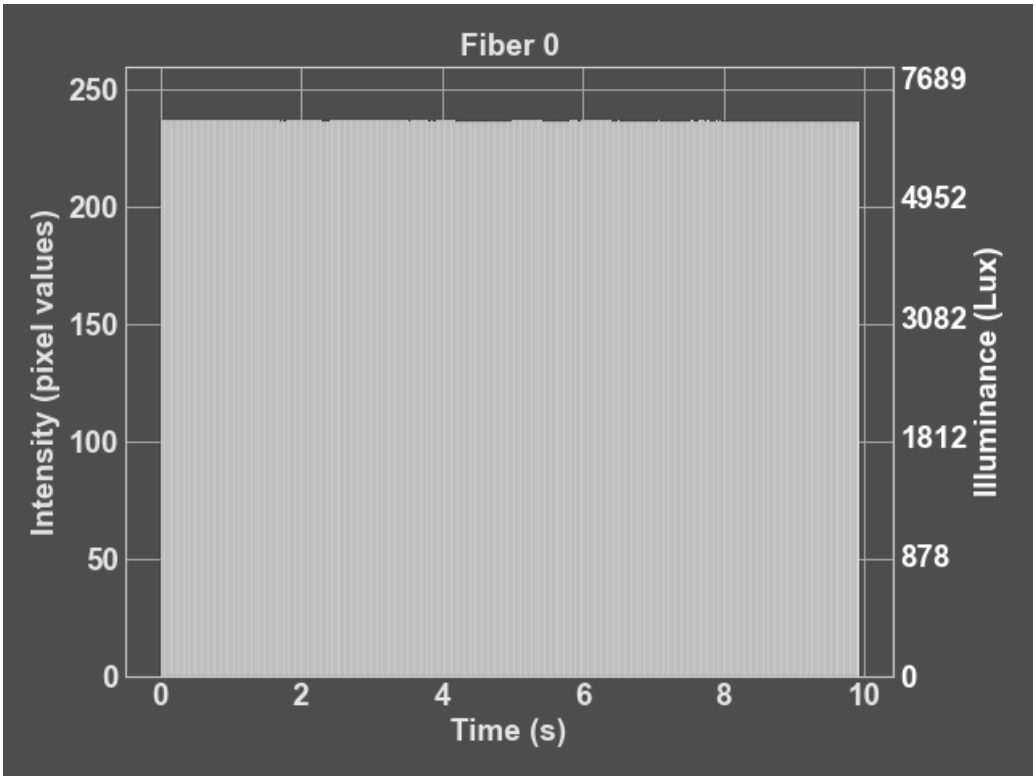


Figure 273: Technical precision test results for constant white light test 9. The color looks gray because of the number of frames; the framework of the box is shaded, so it appears gray.

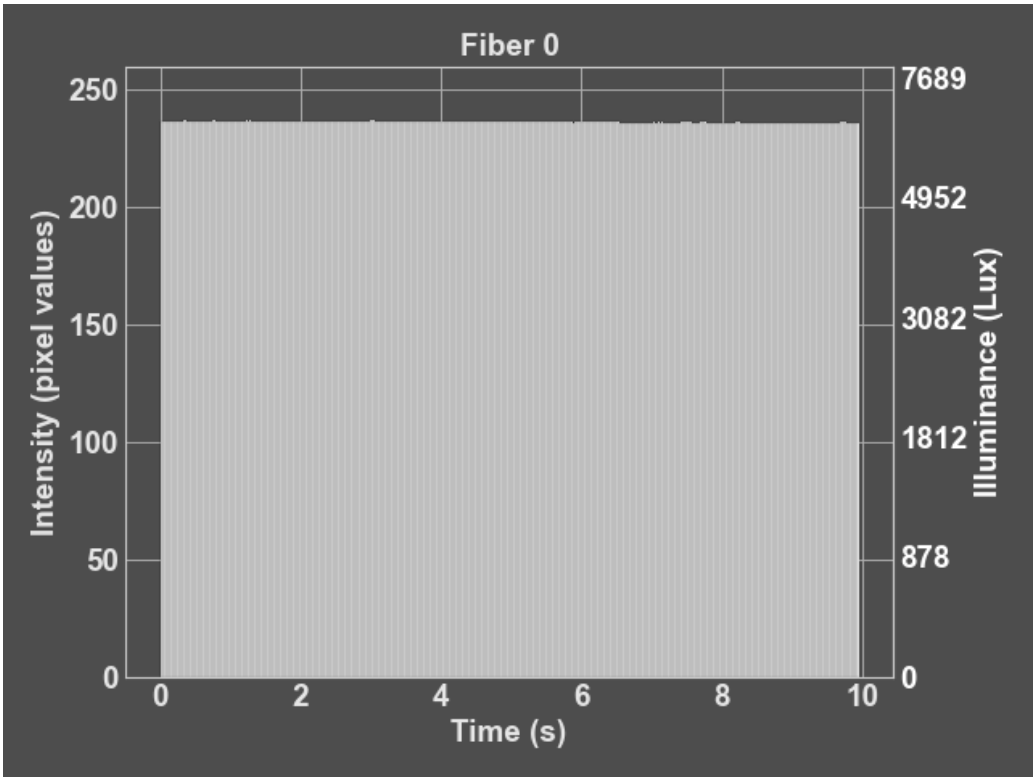


Figure 274: Technical precision test results for constant white light test 10. The color looks gray because of the number of frames; the framework of the box is shaded, so it appears gray.

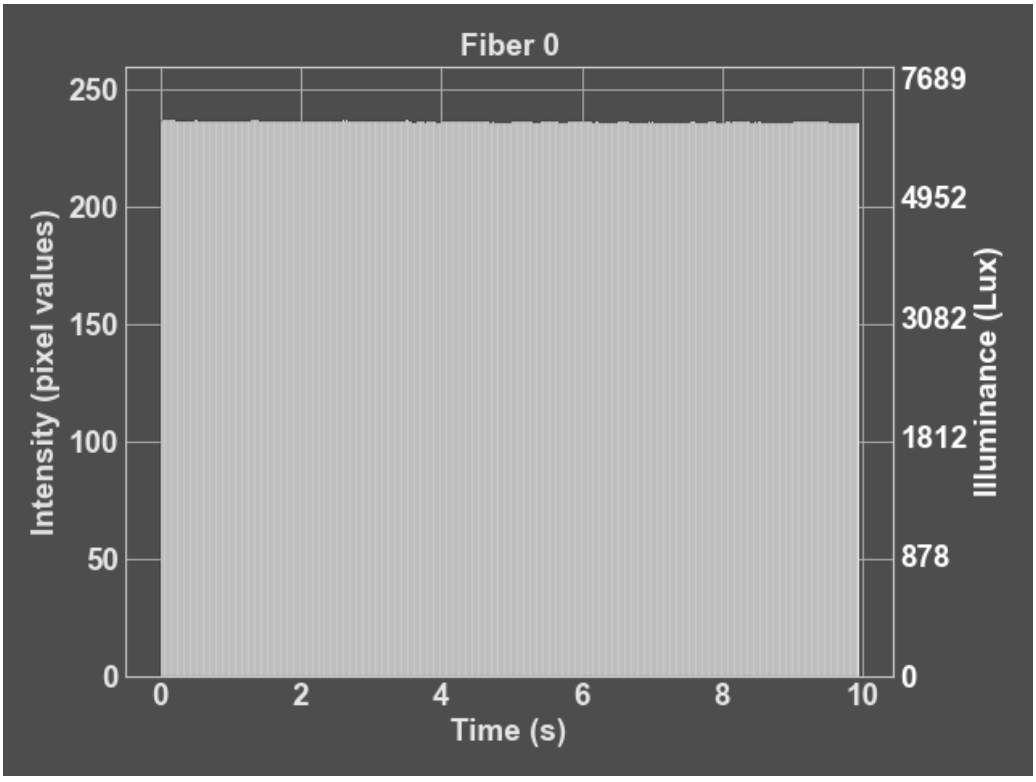


Figure 275: Technical precision test results for constant white light test 11. The color looks gray because of the number of frames; the framework of the box is shaded, so it appears gray.

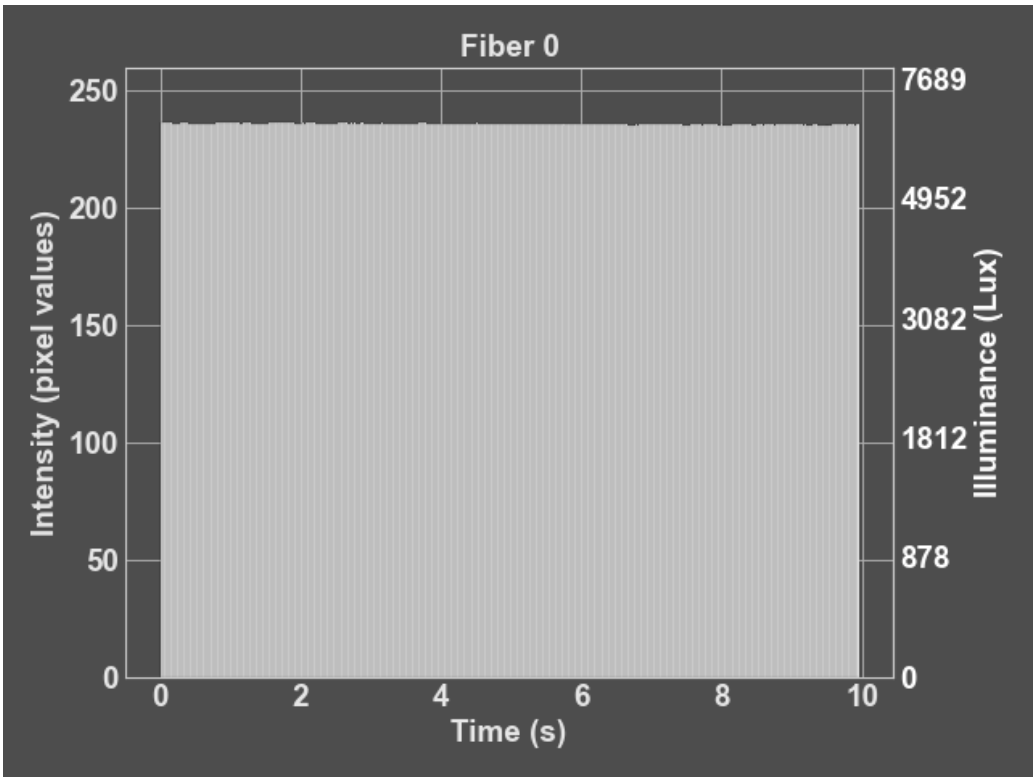


Figure 276: Technical precision test results for constant white light test 12. The color looks gray because of the number of frames; the framework of the box is shaded, so it appears gray.

XIV.II Technical precision test results for constant Blue flickering light

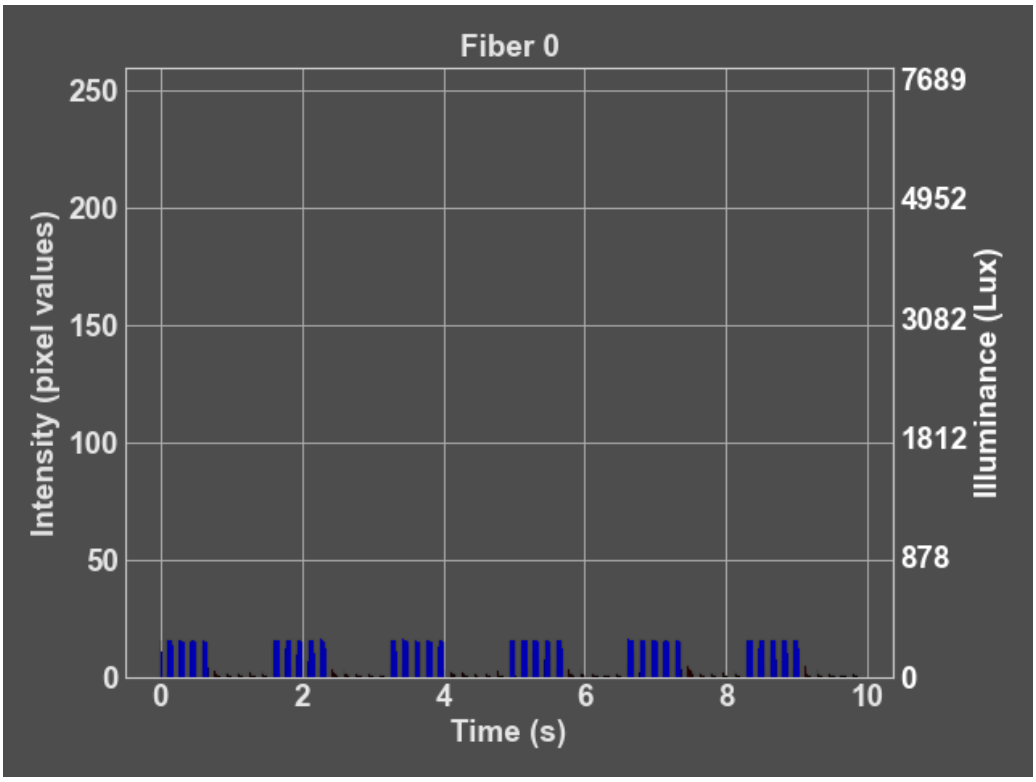


Figure 277: Technical precision test results for blue flickering light test 1.

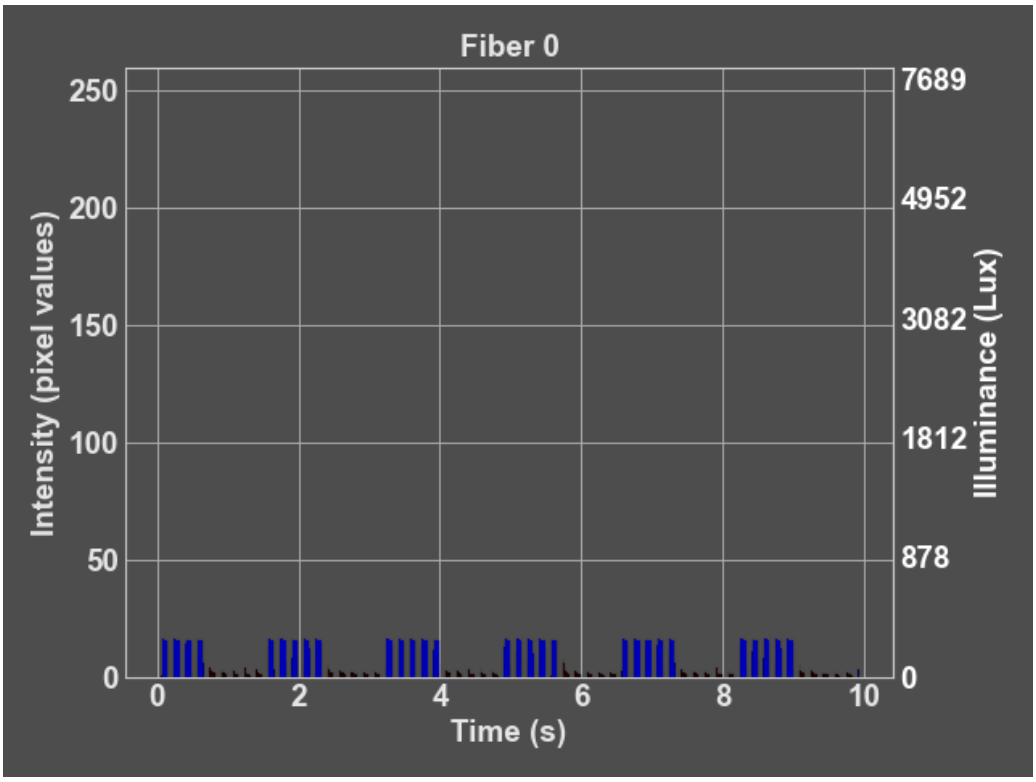


Figure 278: Technical precision test results for blue flickering light test 2.

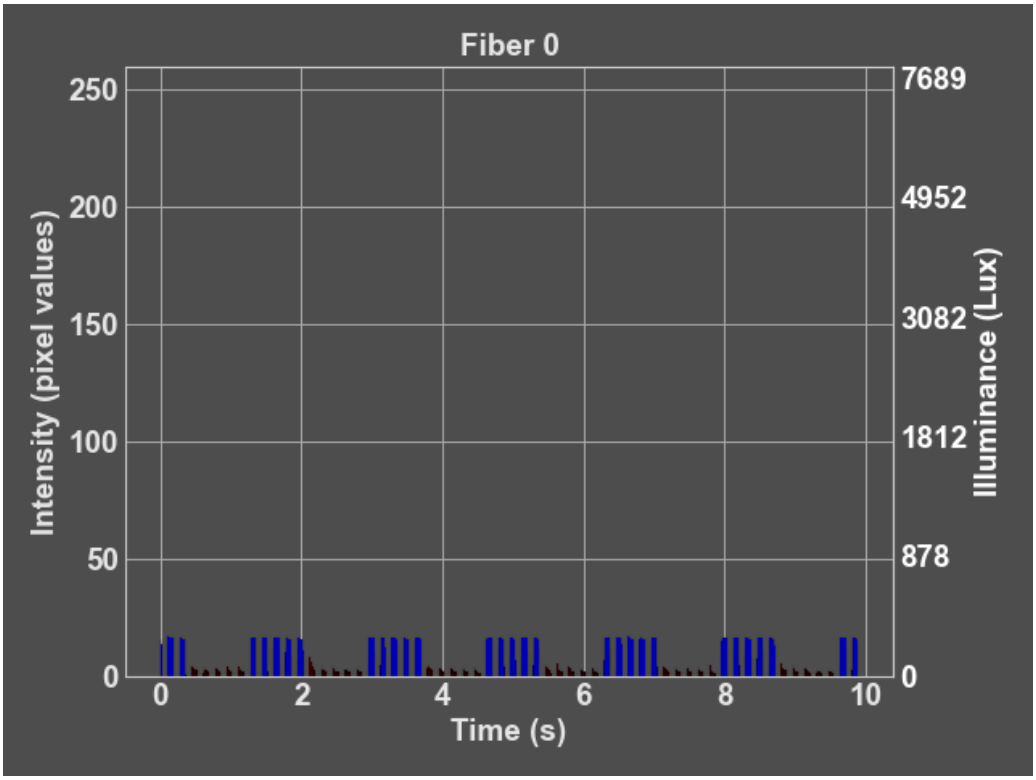


Figure 279: Technical precision test results for blue flickering light test 3.

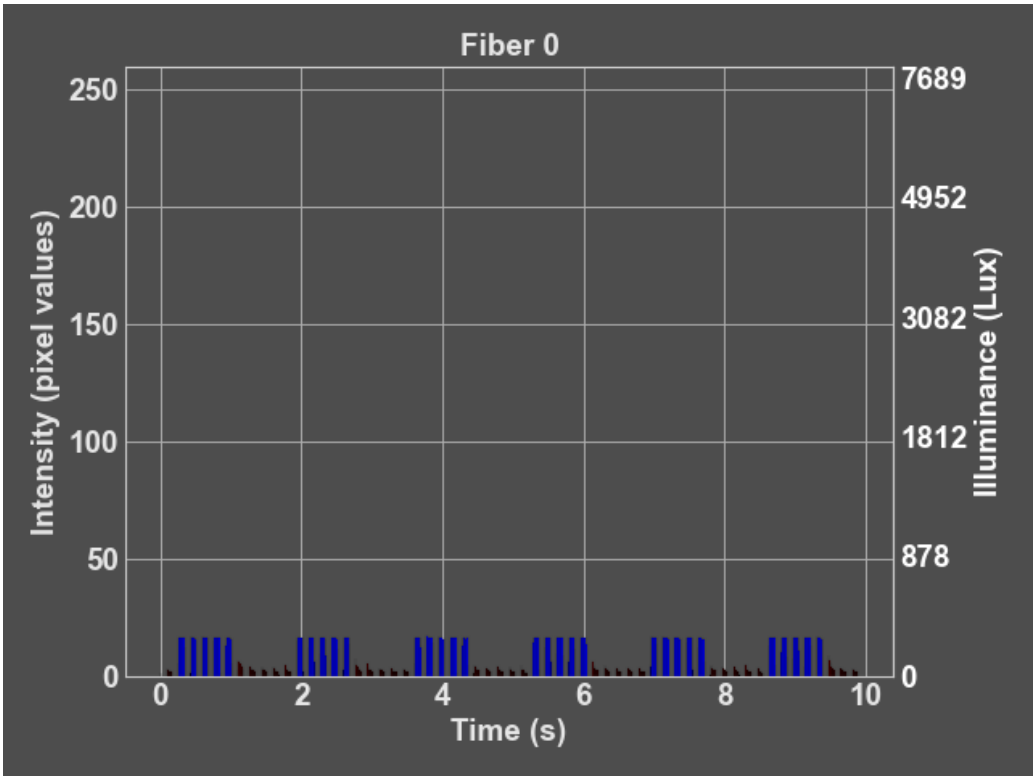


Figure 280: Technical precision test results for blue flickering light test 4.

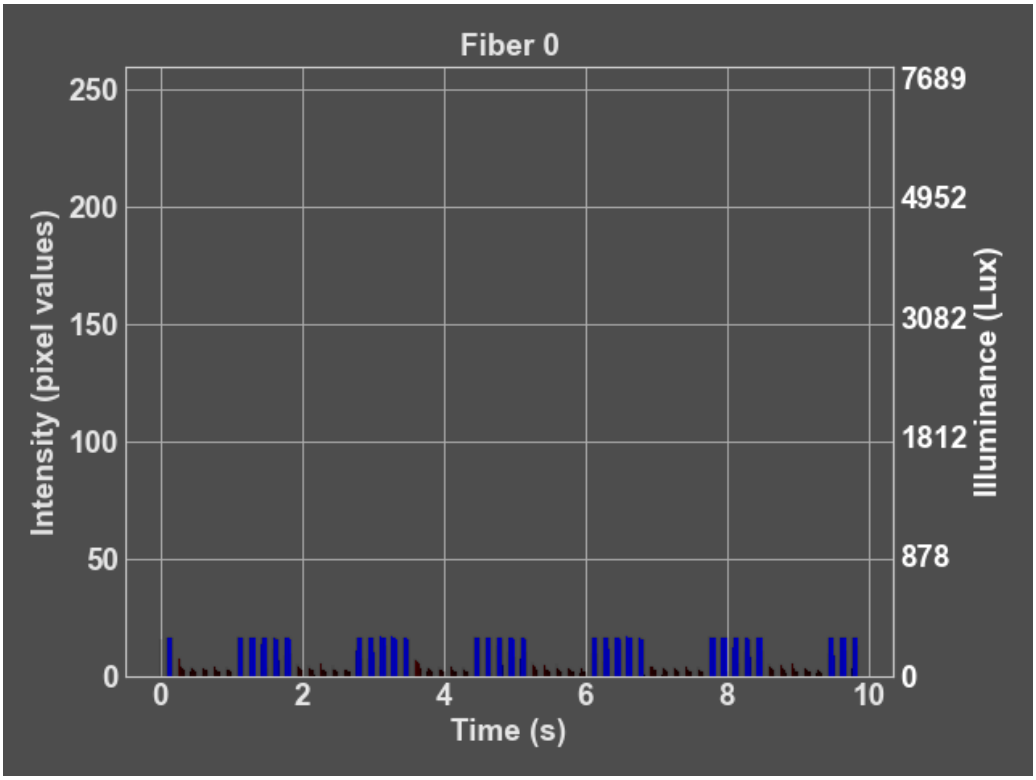


Figure 281: Technical precision test results for blue flickering light test 5.

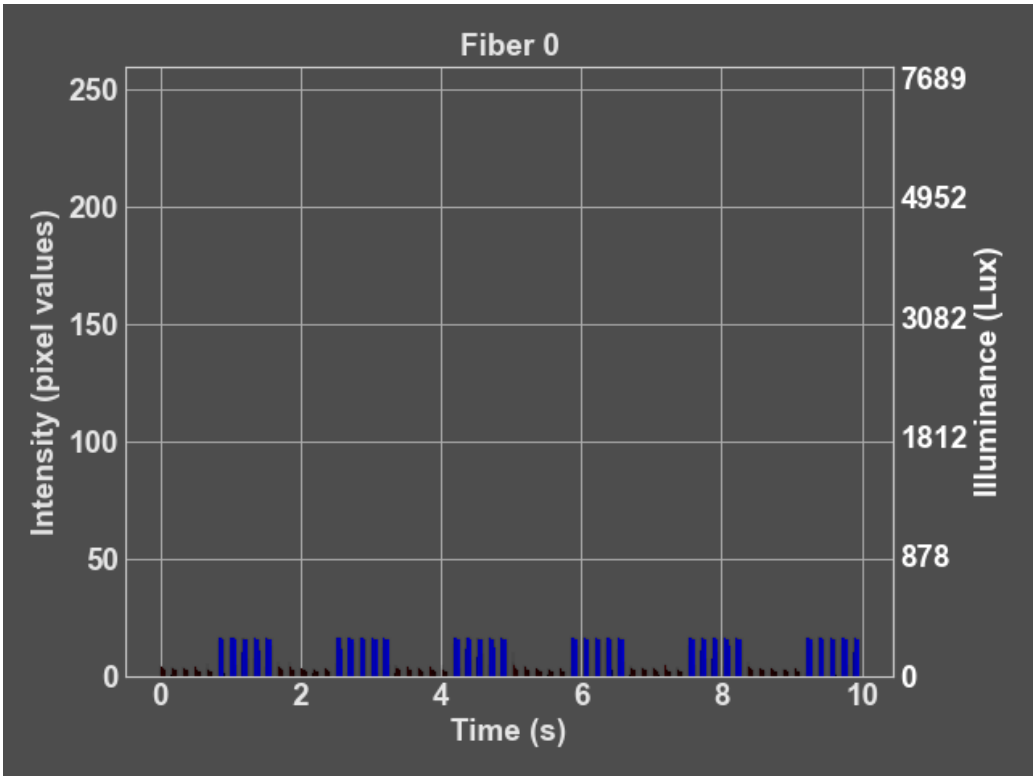


Figure 282: Technical precision test results for blue flickering light test 6.

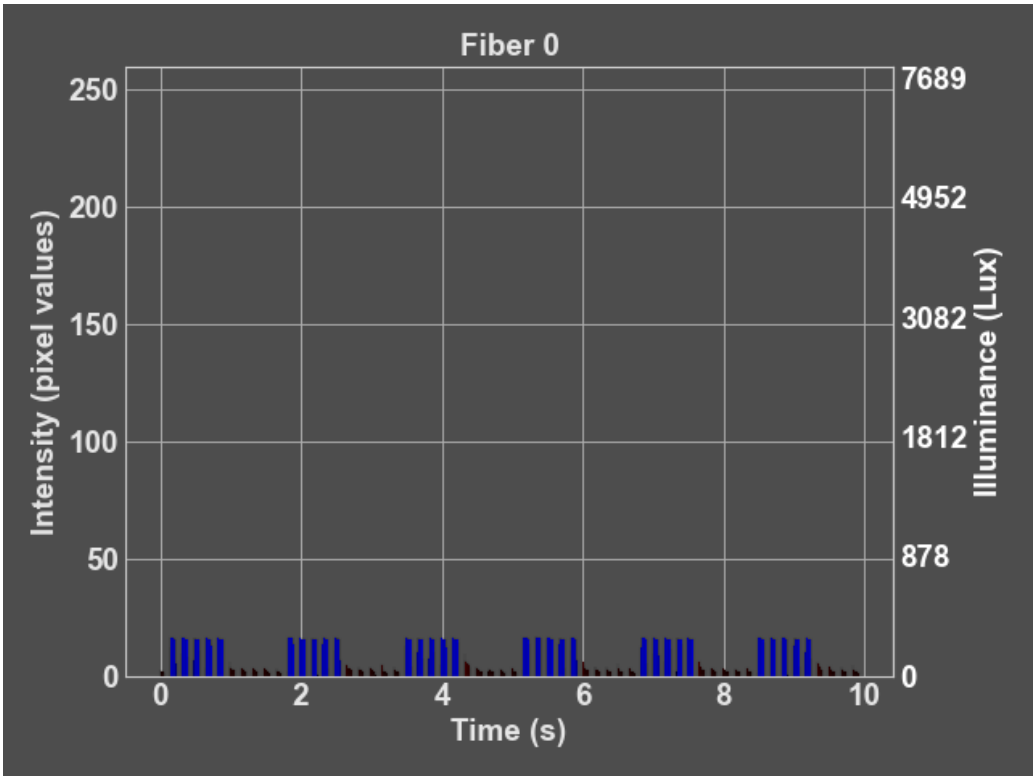


Figure 283: Technical precision test results for blue flickering light test 7.

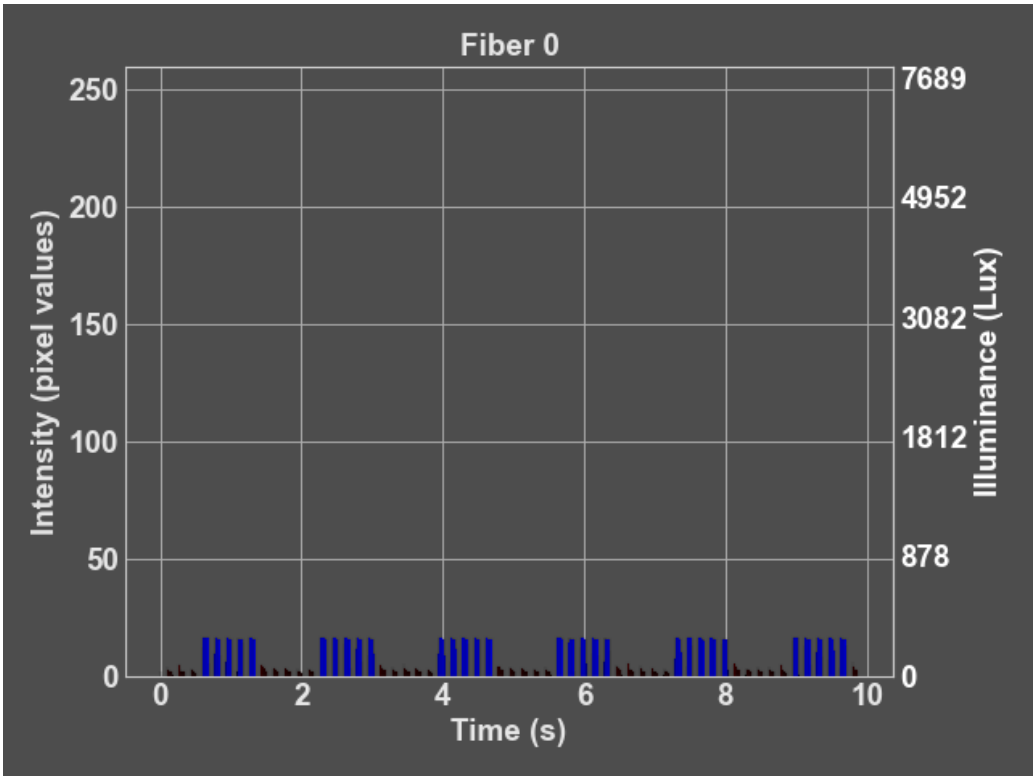


Figure 284: Technical precision test results for blue flickering light test 8.

References

[1] C. Martín-Alberca and C. García-Ruiz, "Analytical techniques for the analysis of consumer fireworks," *TrAC Trends in Analytical Chemistry*, vol. 56, pp. 27–36, 2014.

[2] M. van Justitie en Veiligheid, "Ai model trained to recognise fireworks more quickly after explosion," *News item | Netherlands Forensic Institute*, Aug 2024.

[3] A. Arumugachamy, N. S. Russel, and A. Rajamanickam, "A preliminary study of combustion flame by digital image processing and residue for fireworks flash powder," *ISA Transactions*, vol. 153, pp. 482–489, 2024.

[4] NOS, "2024 was het jaar van de explosies: 'klassiek opsporingswerk niet meer voldoende'," Dec 2024.

[5] Rijksoverheid, "Offensief presenteert aanpak tegen aanslagen met explosieven," Apr 2025.

[6] Politie, "Meer aanslagen met explosieven; zwaar vuurwerk meest gebruikt," Oct 2024.

[7] Nu.nl, "Aantal aanslagen met explosieven in eerste half jaar opnieuw gestegen," Jul 2025.

[8] K. Annot, J. Roza, M. Lokerse, F. Bos, and S. Kassels, "Burning explosives - dossier," *TU Delft*, 2021.

[9] N. Moos, "Identifying pyrotechnic and explosive mixtures: systematic review on techniques for measuring combustion characteristics," *TU Delft, NFI*, Jan 2025.

[10] Elektrolas, "Nen 3140: Safety inspection of electrical tools." https://www.elektrolas.com/en/service/nen-3140.

[11] HTT, "Nen 3140 – electric equipment." https://httbv.nl/en/nen-3140-electric-equipment/, 2025.

[12] Avantes, "Cosine correctors." https://www.avantes.com/products/fiber-optics/fiber-optic-accessories/cosine-correctors/, Jul 2024.

[13] Stellarnet, "What are the cie xyz tristimulus values?." https://www.stellarnet.us/cie-xyz-tristimulus-values/, Dec 2017.

[14] ScikitLearn, "Ridge regression." https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Ridge.html, 2007-2025.

[15] ScikitLearn, "cross-validation: Evaluating estimator performance." https://scikit-learn.org/stable/modules/cross_validation.html, 2007-2025.

[16] Z. Bobbit, "Mse vs. rmse: Which metric should you use?." https://www.statology.org/mse-vs-rmse/, Sep 2021.

[17] ScikitLearn, "Polynomial features." https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.PolynomialFeatures.html, 2007-2025.

[18] NFI, "Validatie van onderzoeksmethoden, qol-00337," aug 2024.

[19] M. Jesussek and H. Volk-Jesussek, "Kruskal-wallis-test." https://numiqo.com/tutorial/kruskal-wallis-test, 2025.

[20] A. Rajesh, "A post-hoc test for kruskal-wallis." https://www.theanalysisfactor.com/dunns-test-post-hoc-test-after-kruskal-wallis/, May 2023.

[21] Scikit, "Permanova: skbio.stats.distance.permanova." https://scikit.bio/docs/dev/generated/skbio.stats.distance.permanova.html, 2014-2025.

[22] J. D. Bakker, "Permanova." https://uw.pressbooks.pub/appliedmultivariatestatistics/chapter/permanova/, Jan 2024.

[23] H. Cheng, X. Jiang, and J. Wang, "Color image segmentation: Advances and prospects," Dec 2001.

[24] Image.Engineering, "Home." https://www.image-engineering.de/library/technotes/958-how-to-convert-between-srgb-and-ciexyz, Aug 2017.

[25] OpenCV and Doxygen, "Color conversions." https://docs.opencv.org/3.4/de/d25/imgproc_color_conversions.html, Apr 2025.

[26] D. Wüller and H. Gabele, "The usage of digital cameras as luminance meters," *Proc. SPIE*, vol. 6502, 03 2007.

[27] K. Annot, J. Roza, M. Lokerse, F. Bos, and S. Kassels, "Dossier - user manual – pyrotechnical/explosive materials analysis device (pemad)," *TU Delft*, 2021.

[28] K. Annot, J. Roza, M. Lokerse, F. Bos, and S. Kassels, "Dossier- user instruction guide," *TU Delft*, 2021.

[29] E. Slavich, "Four strategies for dealing with multiple comparisons." https://analytical.unsw.edu.au/sites/default/files/document_related_files/2018-05-10_multiple_comparisons_statscentral_0.pdf.

[30] S. Glen and A. Leonardo, "Holm-bonferroni method: Step by step." https://www.statisticshowto.com/holm-bonferroni-method/.

[31] Statsig-Team, "Holm-bonferroni correction: Controlling false positives." https://www.statsig.com/perspectives/holm-bonferroni-correction-false-positives, Nov 2024.