POLITECNICO DI TORINO

Master's Degree in Communications Engineering

Master's Degree Thesis

Early Warning of Anomalous Events Using Optical Fiber

Supervisor

Roberto Gaudino

Candidate

Dhia El Hak Daamouche

Co-supervisors

Saverio Pellegrini

Giuseppe Rizzelli Martella

Academic Year 2024/2025

Summary

Rockfalls and landslides pose severe risks to infrastructure and public safety in mountainous regions. Detecting these events in real-time is crucial for risk mitigation and early warning systems. This research investigates the use of State of Polarization Angular Speed (SOPAS) data collected from an experimental mountain gully to monitor and detect anomalous events, particularly falling rocks. The study relies on data from four fiber optic circuits, either buried or exposed, deployed along the gully. These circuits continuously record SOPAS values that are derived from the Stokes parameters and computed as the angle variation between consecutive Stokes vectors on the Poincaré sphere. The SOPAS signal is acquired by a polarimeter and then processed using a digital signal processing (DSP) scheme, which includes smoothing filters to minimize environmental noise and a threshold-based detection mechanism to trigger an alarm only when SOPAS values exceed a predefined level for a significant duration.

The main objective of this research is to develop an efficient data processing system capable of analyzing SOPAS data and distinguishing environmental disturbances from rockfall events. The methodology includes filtering the signal using low-pass filters, such as moving average and Savitzky-Golay, to reduce noise and remove system artifacts while preserving key event characteristics. A combination of statistical analysis (including mean, standard deviation, and percentile calculations) and complementary cumulative distribution function (CCDF) analysis is applied to characterize event patterns and examine the distribution tails, emphasizing rare but high-intensity SOPAS fluctuations.

To validate the proposed detection approach, two distinct SOPAS states were analyzed. The first state, called quiet conditions, serves as a baseline and features minimal disturbances. The second state, called event conditions, was recorded during rockfall experiments conducted in the gully, where designed blocks and regular rocks were deliberately dropped onto the fiber circuits to generate controlled SOPAS signals. Moreover, A dual-threshold detection strategy was developed to improve event classification accuracy. This approach is a combination of a SOPAS-based threshold (ω_{th}), to ensure that only significant angular speed variations are considered, and a time-based threshold (d_{th}), to filter short-lived fluctuations and reduce false alarms.

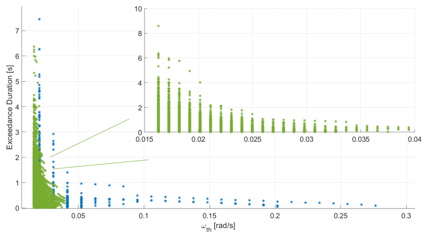


Figure 1: Threshold vs. Exceedance Duration, Circuit-3, for Quiet (green) and Event (blue) Conditions.

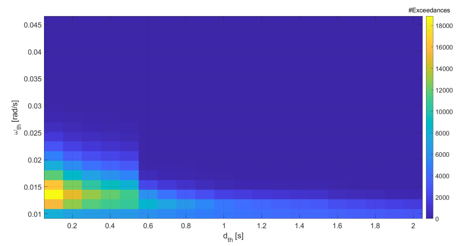


Figure 2: Map of Exceedances for Different ω_{th} and d_{th} Threshold Values (Circuit-2).

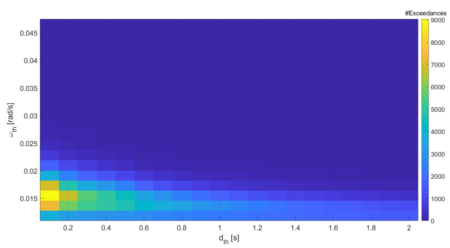


Figure 3: Map of Exceedances for Different ω_{th} and d_{th} Threshold Values (Circuit-3).

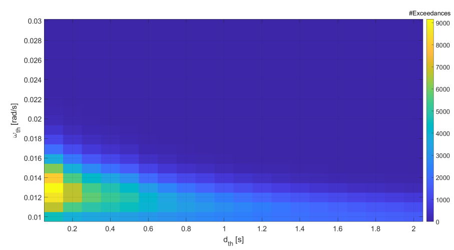


Figure 4: Map of Exceedances for Different ω_{th} and d_{th} Threshold Values (Circuit-4).

Analysis of exceedances (anomalies and events consisting of samples that surpass the threshold ω_{th}) was performed to compare the number and duration of threshold violations across different states. The results revealed clear differences. The quiet conditions state exhibits numerous but short-lived exceedances, often caused by environmental noise. On the other hand, event conditions state generates fewer but longer exceedances, with durations significantly surpassing those of background anomalies. Scatter plots of exceedances durations versus threshold levels, like the one shown in Figure 1, highlight the distinction between background noise and induced events. Moreover, map plots for different circuits were generated (Figure 2, 3, and 4) to visualize the distribution of exceedances that surpass both ω_{th} and d_{th}. These maps provide an intuitive way to assess the impact of different threshold selections. By choosing a specific pair of ω_{th} and d_{th} on the map, one can immediately determine the expected number of false alarms. This enables the identification of an optimal threshold pair where no exceedances meet both conditions, effectively eliminating false alarms. It also offers a clear estimate of the number of false alarms within the observed time window, which aids in fine-tuning the detection parameters.

The findings demonstrate that SOPAS analysis is a promising tool for rockfall detection. The study identified the averaging window W and ω_{th} as key parameters for event detection, requiring careful optimization. While the tested averaging windows performed well, setting ω_{th} is more crucial, as it depends on event intensity, fiber placement, and background noise conditions. Future work could focus on optimizing detection thresholds, integrating machine learning techniques for events classification, and expanding fiber optic deployments to cover larger geographic regions. This would enhance monitoring capabilities and improve real-time hazard detection in mountainous environments.

Contents

Acknowledgements

I would like to deeply thank my supervisors, Prof. Roberto Gaudino, Dr. Saverio Pellegrini, and Dr. Giuseppe Rizzelli, for guiding me during my research. Their insights and advice have been invaluable. I also extend my sincere gratitude to the EDISU organization for providing me with a scholarship, without which it would have been difficult to complete my degree. Finally, I would like to express my heartfelt gratitude to my family and friends for their encouragement and support throughout my studies.

Chapter 1

Introduction

1.1 Context

Natural hazards pose significant threats to human life, infrastructure, and economic stability, particularly in mountainous regions. Rockfalls, avalanches, landslides, and debris flows are especially dangerous due to their unpredictability and destructive potential, which are further exacerbated by climate change. Increased extreme weather events, such as heavy rainfall and rapid snowmelt, weaken slopes and raise the likelihood of hazardous movements [1]. In the Italian Alps, these risks are prevalent, with over 93% of municipalities facing threats from landslides and floods leading to substantial annual damages [2]. Urban expansion into high-risk areas further amplifies potential losses, highlighting the urgent need for reliable early warning systems.

Traditional monitoring methods, including ground-based sensors, visual inspections, and satellite imagery, provide valuable insights but fall short in real-time detection over large areas. Many require costly infrastructure, suffer from environmental interference, or demand significant human involvement. For instance, ground-based sensors are often limited in coverage, while satellite-based systems can be obstructed by cloud cover or low temporal resolution [3]. Optical fiber sensing technology emerges as a promising alternative, offering real-time ground movement detection through variations in the state of polarization (SOP) of light. Unlike conventional sensors, optical fibers provide wide-area coverage, cost-effectiveness, and resilience to harsh environmental conditions, making them suitable for continuous, large-scale monitoring.

Developing effective early warning systems presents several challenges, including the high cost and logistical difficulties of deploying conventional sensors in remote mountain gullies. Many existing methods struggle to differentiate between genuine hazards and environmental noise, leading to false alarms or missed detections, which can compromise public trust and safety. Additionally, extreme weather conditions can degrade sensor performance, reducing their reliability when needed most. Traditional detection systems also rely on centralized data analysis, causing delays that limit their effectiveness in fast-moving hazards like rockfalls and avalanches.

This research aims to overcome these limitations by leveraging optical fiber sensing technology. By focusing on the angular speed variations of SOP (SOPAS), the system can detect ground movements and vibrations that may signal impending natural hazards. This approach offers several advantages: it is cost-effective, as fibers serve both as sensors and data transmission channels; is highly durable in extreme conditions; and enables real-time event detection without human intervention. By exploiting these unique features, this study seeks to develop an innovative early warning system that enhances safety and resilience in mountain environments.

1.2 History

Understanding the history behind any technology is just as important as studying its principles. That is why I want to begin this Chapter by providing a brief history of the evolution of optical fiber. However, this won't be a comprehensive one—the focus will be only on the moments that I believe are most relevant. What started as simple light bending in water experiments has evolved into a technology that connects the world, enabling rapid and reliable communication, and transforming the way we communicate, work, and live. While the fiber optic history can be traced back to the mid-19th century when John Tyndall showed total internal reflection in 1854, explaining how light propagates in different media, scientists began using bent glass rods for medical purposes in the early 20th century. In 1930, Heinrich Lamm transmitted an image through a fiber bundle, which heavily contributed to the development of the contemporary optical fiber [4].

The tipping point arrived in the 1950s and 60s when important developments in the field were made. Bram van Heel and Harold Hopkins enhanced image transmission using cladded optical fibers that greatly reduced light losses. Narinder Singh Kapany, who popularized the term "fiber optics" in 1960, helped in giving popularity to the field. Single-mode fibers (SMF) were proposed by Elias Snitzer in 1961, demonstrating their ability for efficient data transmission. In 1964, Charles Kao, who would eventually receive a Nobel prize in 2009, along with George Hockham, asserted that by minimizing imperfections in the glass, low attenuation through the fiber could be achieved. Their work then encouraged Corning Glass Works to develop low-loss optical fibers in 1970, attaining the important milestone of 20 dB/km attenuation limit, which made long-distance applications of fiber communication possible [4] [5].

By the 1980s, the fiber optics began to supplant copper wires for communication. In 1977, the honor for installing the first metropolitan fiber network was taken by Turin, Italy, demonstrating urban data transmission capabilities. In America, Sprint also undertook the building of a nationwide fiber-optic network, while 1988 witnessed the installation of the first transatlantic fiber-optic line with erbium-doped fiber amplifiers (EDFAs) for signal amplification. In 1991, photonic crystal fiber was introduced for enhanced fiber performance and, in the late 1990s, the start of the construction of large-scale global networks like TPC-5 Pacific cable and FLAG (Fiber Link Around the Globe) was seen [4] [5].

Optical fiber is now the backbone of global communication, acting in a range of sectors, from telecommunications to healthcare and beyond. Its unique capacity to transport data at fast speeds with minimum loss has allowed for the creation of huge networks that support everything from internet access to real-time medical imaging. With ongoing technological advancements, optical fibers have reached new milestones, achieving record-breaking transfer speeds of 402 terabits per second (Tbps) [6]. This demonstrates the critical role of optical fiber in meeting the ever-increasing demand for faster, more reliable data transmission in an increasingly connected world.

1.3 Basic Principles of Optical Fiber

1.3.1 Overview

The optical fiber is a cylindrical, nonconducting waveguide that transmits light along its axis. It is a thin strand of very purified glass. Its standard diameter is an eighth of a millimeter (125 μm), which is roughly the thickness of a human hair. The glass is made of SiO_2 (silicon dioxide), similar to the glass found in window panes or drinking glasses, but much purer.

Optical fibers are made up of layers of dielectric materials, as shown in Figure 1.1. High-performance fibers have an interior portion (the core) with a diameter of approximately (10 μm). The part of the fiber surrounding the core is called the *cladding*, and its diameter is larger than the core, depending on the type of fiber used. To provide mechanical protection and prevent physical damage, a layer surrounding the core and cladding is used; it's called the *buffer* or *protective coating*.

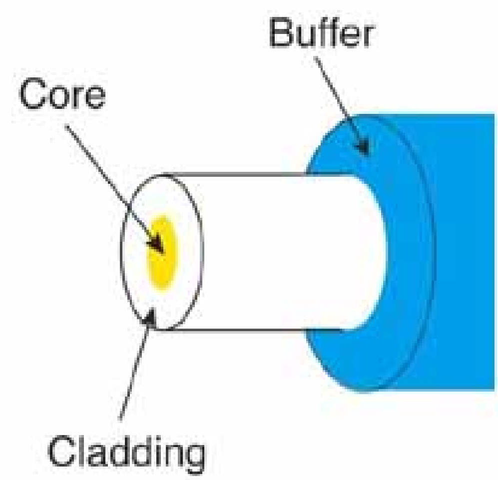


Figure 1.1: Structure of an Optical Fiber. (Image from [7])

1.3.2 Refractive Index

The refractive index is a property of the material that determines how light propagates through it. In a vacuum, such as outer space, light travels at its quickest. The speed of light in vacuum is approximately 300,000 kilometers per second. To compute the refractive index of a medium, we divide the speed of light in vacuum by the speed of light in that medium. Mathematically, this is expressed as:

$$n = \frac{c}{v} \tag{1.1}$$

where:
n = refractive index of the medium
c = speed of light in vacuum
v = speed of light in the medium

In the vacuum, $n = 1$ by definition. A typical single-mode fiber (see next subsection) for telecommunications has a pure silica cladding with an index of approximately 1.444 at wavelength of 1500 nanometre (nm), and a doped silica core with an index of roughly 1.4475. That said, a signal using optical fiber for communication will travel at around 200,000 kilometers per second, two-thirds of the speed of light. For instance, a phone call carried by fiber between Turin and Algiers, a distance of approximately 1000 kilometers, would have a minimum delay of 5 milliseconds between when the first person speaks and the other hears. Mathematically, the delay is calculated as follows:

$$\text{Delay} = \text{Distance} \;/\; \text{Speed of light in fiber}$$

For the Turin-Algiers example:

$$\text{Delay} = 1000 \text{ (km)} \;/\; 200000 \text{ (km/s)} = 0.005 \text{ seconds (5 milliseconds).}$$

1.3.3 Single-Mode Fiber

Optical fibers can be classified into single-mode fiber (SMF) and multi-mode fiber (MMF) based on the mode of light propagation. SMF, which is the focus of this study, is designed for long-distance communication. It has a small core, typically $8 - 10$ μm, allowing only a single mode of light to propagate. This design minimizes signal attenuation—the loss caused by light not traveling in a straight path—and dispersion, where different wavelengths travel at different speeds. As a result, SMF enables high-bandwidth data transmission. It is commonly utilized in situations where reliable performance is required, such as long-distance telecommunications networks or places with harsh weather conditions. In contrast, MMF have a larger core ($50-62.5\ \mu m$), which permits multiple propagation modes. This makes them suitable for shorter distances (up to a few kilometers), but they suffer from higher dispersion and signal loss over long distances.

SMF operates based on the refractive index profile of the core. The most common type is the step-index fiber, where the core has a uniform refractive index that sharply drops at the core-cladding interface, as shown on Figure 1.2. This abrupt transition helps confine the light within the core, ensuring efficient transmission. Step-index fiber is often created by doping high-purity fused silica glass (SiO2) with various amounts of minerals such as titanium, germanium, or boron [10].

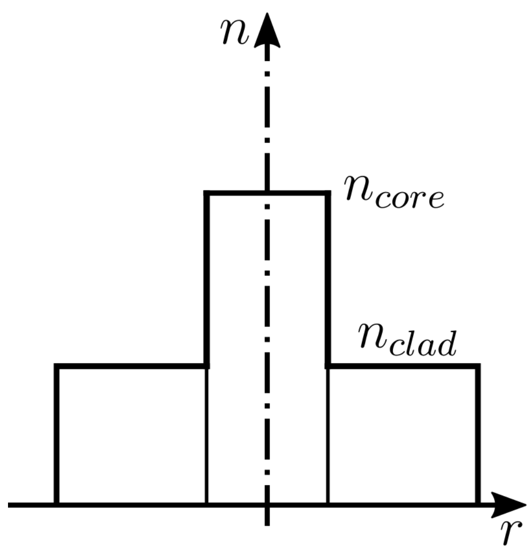


Figure 1.2: Refractive-index Profile for Step-index Fiber. (Image from [10])

1.4 Optical Fiber Sensing

The results of this thesis were obtained using a sensor system made of optical fibers. Therefore, it is critical to provide a brief overview of the current state of technology in this subject, including alternatives to the one used in my research.

Since laser ligth was established in 1960, researchers have been interested in studying the possibilities of optical fiber communication systems for sensing, data communications, and a variety of other applications. As a result, fiber optic communication systems have emerged as the preferred method for gigabit and beyond gigabit data transmission. This technology has found a way not only on the telecommunication domain but also in civil engineering, petroleum industry, medicine, environmental monitoring, and much more. This evolution led to the development of optical fiber sensors, which utilize light to detect various physical and chemical parameters, such as temperature, pressure, vibrations, and chemical concentrations. These sensors offer key advantages, including immunity to electromagnetic interference, compact size, remote sensing capabilities, and suitability for extreme environments.

Optical fiber-based sensors can be broadly classified into three categories: discrete, distributed, and quasi-distributed sensing. Discrete sensing involves the localized sensors that measure parameters at specific points along the fiber. A widely used example is Fiber Bragg Gratings (FBGs), which consist of periodic refractive index modulations in the fiber core. FBGs reflect specific wavelengths of light, and any strain or temperature variations cause shifts in the reflected wavelength, which enables precise measurements.

Unlike discrete sensing, distributed sensing uses scattering phenomena—Rayleigh, Brillouin, and Raman scattering—to provide continuous measurements along the entire fiber length. These techniques allow us to determine the spatially distributed parameters, making them highly effective for structure health monitoring, environmental sensing, and a wide of industrial applications. Raman sensing technique relies on Raman scattering, where incident light interacts with molecular vibrations in the fiber, which causes a frequency shift. It is used for distributed temperature sensing (DTS), making it suitable for application like fire detection and pipeline monitoring, as well as geothermal studies. Brillouin sensing depends on Brillouin scattering, a method that measures frequency shifts induced by temperature and strain changes. It is widely used in distributed strain and temperature sensing (DTSS) for applications such as structural health. Rayleigh backscattering is used to detect acoustic vibrations along an optical fiber. It enables real-time monitoring of seismic activity, pipeline leaks, perimeter security, and even traffic movements by capturing minute disturbances along the fiber.

SOP, the focus of this research, is considered as a quasi-distributed sensing technique. This technique leverages the use of the whole optical fiber as a sensor. However, and even though it is highly sensitive to environmental changes, SOP is unable to localize the exact site of an event along the fiber. This is the major advantage of distributed sensing over quasi-distributed sensing. Raman sensing typically achieves a spatial resolution of 1 meter, and Brillouin can reach sub-meter resolutions, while the DAS system surpasses both with a resolution up to 0.1 meter [12].

1.5 Polarization

To understand polarization, let us start by introducing two fundamental concepts: electric fields and magnetic fields of an electromagnetic wave. We can write the electric field and magnetic fields for a plane wave propagating over time t along the z-axis, as follows:

$$\mathbf{E}(z,t) = \hat{x}E_x \cos(kz - \omega t + \phi_x) + \hat{y}E_y \cos(kz - \omega t + \phi_y) \tag{1.2}$$

$$\mathbf{B}(z,t) = \hat{z} \times \mathbf{E}(z,t) = \hat{y}B_y \cos(kz - \omega t + \phi_x) - \hat{x}B_x \cos(kz - \omega t + \phi_y) \tag{1.3}$$

Where:

- z: Position along the propagation direction (typically the fiber axis).
- t: Time, representing how the electric field changes over time.
- E_x: Amplitude of the electric field in x-direction.

- E_y: Amplitude of the electric field in y-direction.
- ϕ_x: Phase of x-component.
- ϕ_y: Phase of y-component.
- $k = \frac{2\pi}{\lambda}$: Wave number.
- $\omega = 2\pi f$: Angular frequency.

Now, let us focus on the expression $\mathbf{E}(z, t)$, because the polarization is defined by the behavior of the electric field vector. The two components of the electric field, E_x and E_y, can be expressed as:

$$\begin{cases} E_x = A_x \cos(kz - \omega t + \phi_x) \\ E_y = A_y \cos(kz - \omega t + \phi_y) \end{cases} \tag{1.4}$$

where A_x and A_y are the magnitudes of the electric field components E_x and E_y, respectively. Let the phase difference between the components be:

$$\delta = \phi_y - \phi_x \tag{1.5}$$

Then Equation 1.4 becomes:

$$E_x = A_x \cos(kz - \omega t) \tag{1.6}$$

$$E_y = A_y \cos(kz - \omega t + \delta) \tag{1.7}$$

We can expand Eq. 1.7 by using trigonometric identities:

$$E_y = A_y \left[\cos(kz - \omega t)\cos(\delta) - \sin(kz - \omega t)\sin(\delta)\right] \tag{1.8}$$

By substituting (1.6) in (1.8), it is possible to find a more general formula (1.9):

$$\left(\frac{E_x}{A_x}\right)^2 + \left(\frac{E_y}{A_y}\right)^2 - 2\left(\frac{E_x E_y}{A_x A_y}\right)\cos(\delta) = \sin^2(\delta) \tag{1.9}$$

Eq. 1.9 describes an ellipse on E_x, E_y plane, whose shape depends on A_x, A_y, and δ. It also stands as the base to describe the types of polarization that can be obtained:

1.5.1 Linear Polarization

When the electric field oscillates only along a single axis, this is called Linear Polarization. This type of polarization is generated when $\delta = 0$ or $\delta = \pi$. In other words, the x- and y-components are either in phase or completely out of phase. Substituting δ in (1.9), we find:

$$E_y = \left(\frac{A_y}{A_x}\right) E_x \tag{1.10}$$

Eq. 1.10 shows that the trajectory of the electric field vector is a straight line. The orientation of the line depend on the ratio $\frac{A_y}{A_x}$.

1.5.2 Circular Polarization

Circular polarization happens if the electric field vector instead traces a circular rotation. This occurs when the magnitudes of the electric field components, A_x and A_y, are equal and the phase difference $\delta = \pm\frac{\pi}{2}$. Substituting them in Eq. 1.10, we get:

$$\begin{cases} E_x = A\cos(kz - \omega t) \\ E_y = A\sin(kz - \omega t) \end{cases} \tag{1.11}$$

and thus:

$$(E_x)^2 + (E_y)^2 = A^2 \tag{1.12}$$

Eq. 1.12 illustrates the electric field vector, which traces out a circle in the x-y plane.

1.5.3 Elliptical Polarization

So far, we have seen special cases of polarization. Elliptical Polarization is the most general case. It happens when the magnitudes A_x and A_y in (1.6) and (1.7), respectively, are not equal and the phase difference is different from 0 or $\frac{\pi}{2}$. This is indeed the same equation found in (1.9), the ellipse equation:

$$\left(\frac{E_x}{A_x}\right)^2 + \left(\frac{E_y}{A_y}\right)^2 - 2\left(\frac{E_x E_y}{A_x A_y}\right)\cos(\delta) = \sin^2(\delta)$$

The length of the major ellipse's axis is proportional to the larger of A_x and A_y, while the length of the minor axis is proportional to the smaller one. The phase difference δ accounts for the orientation and tilt of the ellipse. Figure 1.3 shows the three mentioned types of polarization.

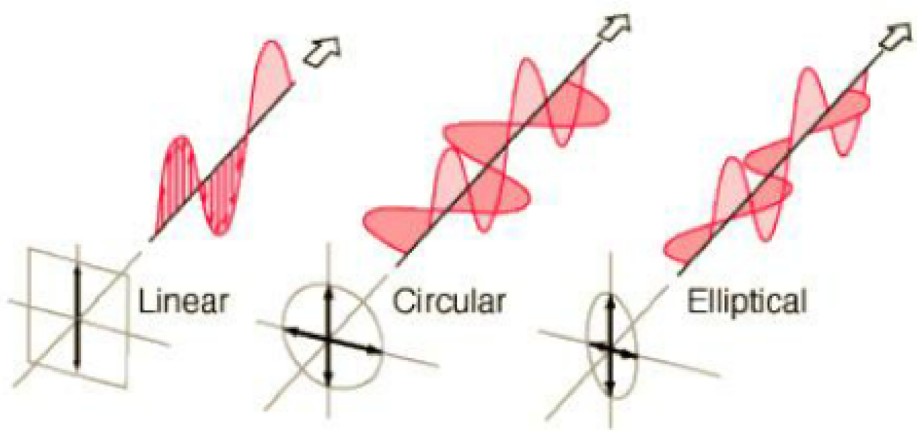


Figure 1.3: Type of Polarization. (Image from [8])

1.6 Stokes Parameters and Poincaré Sphere

Conventionally, to describe the light state of polarization, some useful parameters are employed, called **Stokes Parameters**. They provide a thorough mathematical representation of the trajectory of the electric field. Let us recall the expression of the electric field components in (1.6) and (1.7):

$$
\begin{cases}
E_x = A_x \cos(kz - \omega t) \\
E_y = A_y \cos(kz - \omega t + \delta)
\end{cases}
\tag{1.13}
$$

And the ellipse equation in (1.9):

$$
\left(\frac{E_x}{A_x}\right)^2 + \left(\frac{E_y}{A_y}\right)^2 - 2\left(\frac{E_x E_y}{A_x A_y}\right)\cos(\delta) = \sin^2(\delta)
$$

We define the Stokes parameters S_0, S_1, S_2, and S_3 to characterize the polarization, as follows:

$$
\begin{cases}
S_0 = A_x^2 + A_y^2 \\
S_1 = A_x^2 - A_y^2 \\
S_2 = 2A_x A_y \cos(\delta) \\
S_3 = 2A_x A_y \sin(\delta)
\end{cases}
\tag{1.14}
$$

S_0 represents the total intensity of light. S_1 is the difference in intensity between the horizontal and vertical components. S_2 and S_3 quantify the phase-dependent interference between the components. These parameters can fully describe the state of polarization. However, to account for how much an electromagnetic wave is polarized or not, we need to define another term called the **degree of polarization**- acronymed DOP but let's call it p. It is determined as:

$$
p = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0}
\tag{1.15}
$$

Linear Polarization occurs when $S_3 = 0$. For circular polarization to appear, $S_1 = S_2 = 0$ and $S_3 \neq 0$, meaning the phase difference δ is $\pm\frac{\pi}{2}$. The sign of S_3 defines the direction of polarization, either right-hand or left-hand circularly polarized. Additionally, by varying p, we can describe the three cases of light polarization:

- $p = 1$: Completely polarized light.

- $0 < p < 1$: Partially polarized light.
- $p = 0$: Unpolarized light.

To visualize the state of polarization geometrically, we use the **Poincaré Sphere**. To do so, let us first express the Stokes parameters in spherical coordinates:

$$\begin{cases} S_0 = I \\ S_1 = Ip\cos 2\psi \cos 2\chi \\ S_2 = Ip\sin 2\psi \cos 2\chi \\ S_3 = Ip\sin 2\chi \end{cases} \tag{1.16}$$

where:

- I: Total intensity of the light.
- ψ: Orientation angle.
- χ: Ellipticity angle.

The orientation angle, ψ, represents the orientation of the ellipse formed by the electric field vector in the xy-plane, while the ellipticity angle, χ, describes the shape of the polarization ellipse. These two variables can be defined in terms of the parameters of the polarization ellipse as follows:

$$\tan(2\psi) = \frac{2A_xA_y}{A_x^2 - A_y^2}\cos(\delta) \quad , \quad 0 \leq \psi \leq \pi \tag{1.17}$$

$$\sin(2\chi) = \frac{2A_xA_y}{A_x^2 + A_y^2}\sin(\delta) \quad , \quad -\pi/4 < \chi \leq \pi/4 \tag{1.18}$$

These parameters, along with the Stokes parameters, can be used to create a geometric representation of the SOP of an electromagnetic wave using the Poincaré sphere. The latter, as shown in Figure 1.4, provides a powerful visualization tool to understand the SOP and its changes. Each point on the Poincaré sphere corresponds to a unique SOP, and the Stokes parameters S_1, S_2, and S_3 correspond to the position of this point:

- **Linear Polarization**: Represented by the equator of the sphere; the angles vary from 0° to 180°.
- **Circular Polarization**: Determined by the north and south poles, where the north pole defines right-handed polarization and the south pole defines left-handed polarization.
- **Elliptical Polarization**: Represented by points along the upper and lower hemispheres. The degree of ellipticity increases as the point moves from the equator to the poles.

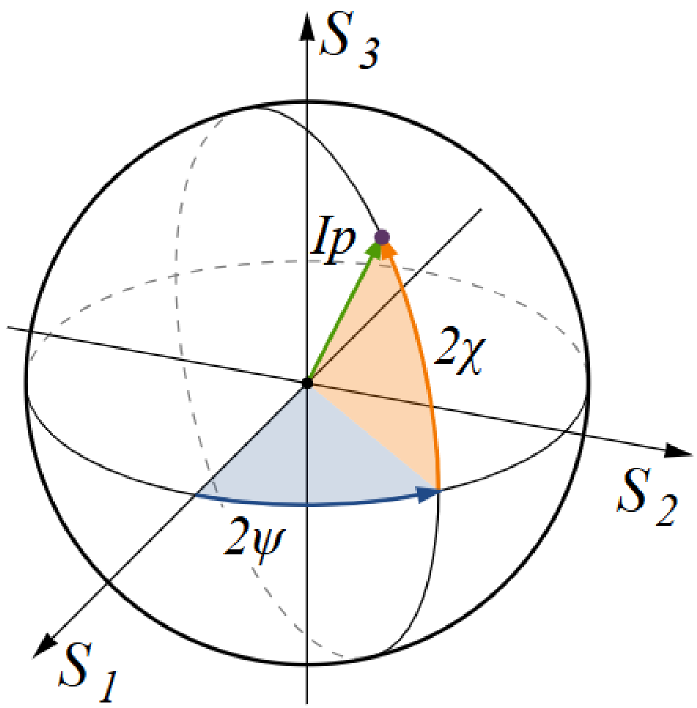


Figure 1.4: The Poincaré Sphere. (Image from [9])

1.7 Birefringence

Another important term that should be mentioned when dealing with polarization is birefringence. In a nutshell, it is the fact that, as different polarization components travel through the fiber, they experience different refractive indices. To better understand this phenomenon, we should first introduce the concept of **effective refractive index** (n_{eff}). This is only for the case of single-mode fibers (SMF), which is the focus of my research.

The effective refractive index seen by the mode propagating in the fiber is defined as:

$$n_{\text{eff}} = \frac{\beta}{k_0} \tag{1.19}$$

where:

- β: propagation constant.
- $k_0 = \frac{2\pi}{\lambda}$: free-space wavenumber.

Ideally, all the polarization components would have the same n_{eff}, but a polarization-dependent variation is introduced. This happens due to imperfections in the optical fiber, such as the geometry of an elliptical core, residual stresses during fiber fabrication, and external perturbations (e.g., bends, pressure, temperature changes). This is what introduces birefringence, which is evaluated as:

$$\Delta n = n_x - n_y \tag{1.20}$$

where n_x and n_y are the effective refractive indices experienced by the two polarization components. This birefringence causes the SOP of light to change as it propagates and results in a phase delay between its polarization components. For instance, in Linear Polarization, where otherwise the polarized input remains the same, differential phase accumulation between the orthogonal modes can induce a phase shift of $\pm\frac{\pi}{2}$ and may cause the polarization to rotate.

1.8 Thesis Structure

This thesis has four Chapters, excluding the introduction and conclusion, each building on the preceding one to provide a thorough knowledge of the research problem, methods, findings, results, and discussion. A full overview of each chapter and its contents is provided below.

Chapter II: Methodology and Experimental Setup

This chapter describes the experimental setup and the algorithm used for event detection. It also details the tools developed for data acquisition and analysis. The chapter is divided into five sections:

- **Section 2.1**: Outlines the general architecture of an optical fiber sensor.
- **Section 2.2**: Describes the polarimeter, the core component of the entire setup.
- **Section 2.3**: Explains the event detection algorithm and the digital signal processing (DSP) scheme used in the detection process.
- **Section 2.4**: Details the experimental setup deployed at the gully site.
- **Section 2.5**: Introduces the Polaralp system, the interface used for monitoring the gully setup.

Chapter III: Rockfall Event Simulation

This chapter describes the experiments conducted to simulate rockfall events. It outlines the tests performed across five different tranches, each corresponding to a separate optical fiber circuit. These experiments are collectively referred to as the *event conditions*.

- **Section 3.1**: Provides an introduction to the experiments and the materials used.
- **Section 3.2**: Explains the application of the moving average filter to the data, discussing its advantages and limitations.

- **Section 3.3**: Describes and analyzes Tests 01–05, conducted on Circuit-4.
- **Section 3.4**: Describes and analyzes Tests 06–13, conducted on Circuit-3.
- **Section 3.5**: Describes and analyzes Tests 14–21, conducted on Circuit-2.
- **Section 3.6**: Describes and analyzes Tests 22–27, conducted on Circuit-1.
- **Section 3.7**: Describes and analyzes Tests 28–31, conducted on Circuit-4.
- **Section 3.8**: Presents additional analysis using power spectrum and wavelet transform techniques.

Chapter VI: Quiet Conditions

This chapter examines the second state of the SOPAS system, referred to as the *quiet conditions* state.

- **Section 4.1**: Provides an introduction to the chapter.
- **Section 4.2**: Discusses system bugs and outlines an algorithm developed to remove them.
- **Section 4.3**: Describes additional filtering techniques used to smooth SOPAS data.
- **Section 4.4**: Analyzes the quiet conditions through various comparisons, considering both the presence and absence of anomalies.

Chapter V: SOPAS Data States and Threshold Determination

This chapter compares the two SOPAS states using various parameters. It also introduces the concept of thresholds and their role in distinguishing actual rockfall events from anomalies. The relationship between the SOPAS threshold and the filtering window size is explored to optimize event detection.

- **Section 5.1**: Defines the SOPAS threshold, ω_{th}.
- **Section 5.2**: Investigates the relationship between the SOPAS threshold and the filtering window size, W.
- **Section 5.3**: Analyzes samples exceeding the threshold ω_{th} and compares both SOPAS states.
- **Section 5.4**: Introduces the time-based threshold, d_{th}, and examines events that surpass both thresholds.
- **Section 5.5**: Explores the use of the Complementary Cumulative Density Function (CCDF).
- **Section 5.6**: Summarizes key statistical metrics used for comparing the two SOPAS states.
- **Section 5.7**: Introduces the final results: the plots of exceedances vs. thresholds.
- **Section 5.8**: Discusses the final results.

Chapter VI: Conclusion

This Chapter summarizes the key findings of the thesis, highlighting the methodology used for SOPAS data analysis, the distinction between quiet and event conditions, and the definition of detection thresholds. It also outlines potential future improvements.

Chapter 2

Methodology & Setup

This Chapter describes the experimental setup and the algorithm used for event detection. It also lists the tools designed to acquire and analyse the data. Before delving into the setup used in this research thesis, and since we are dealing with fiber sensing, it is important to first describe the general architecture of optical fiber sensing. In this way, we will have a broad picture of all elements used to establish the experiment.

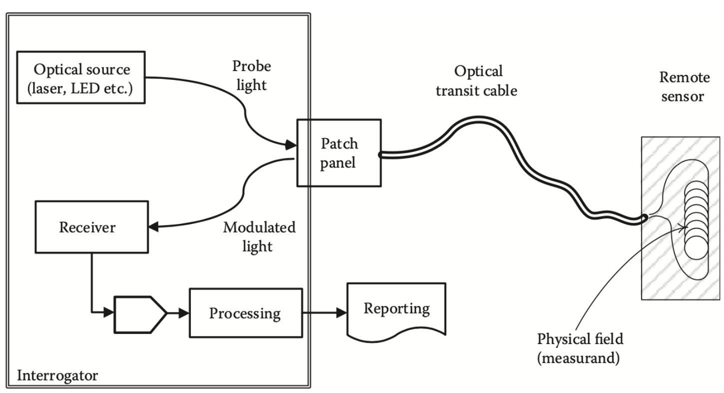


Figure 2.1: The Architecture of Optical Fiber Sensor. (Image from [13])

2.1 Architecture of Optical Fiber Sensor

Optical fiber sensing relies on the interaction between light and the surrounding environment to detect and measure physical phenomena. The system consists of several key components, as illustrated in Figure 2.1, each playing an important role in data acquisition and signal processing. In general, it consists of three main units:

- **Interrogator Unit**

  This is the central element of the system, responsible for generating, transmitting, receiving, and processing optical signals. It consists of an **optical source**, which is a laser used to generate a stable probe light to be injected into the fiber. The laser used in the experiment is shown in Figure 2.2. It is a Novoptel laser instrument that emits light at 1550 nm. One of its key features is its high output power, reaching 17 dBm. The reason for this high power is to compensate for the losses caused by switches, connections, and splices in the fiber patch. It also features an ultra-narrow linewidth, which is fundamental for phase measurements; however, polarization measurements do not require a narrow linewidth and could employ a larger linewidth with much cheaper lasers.

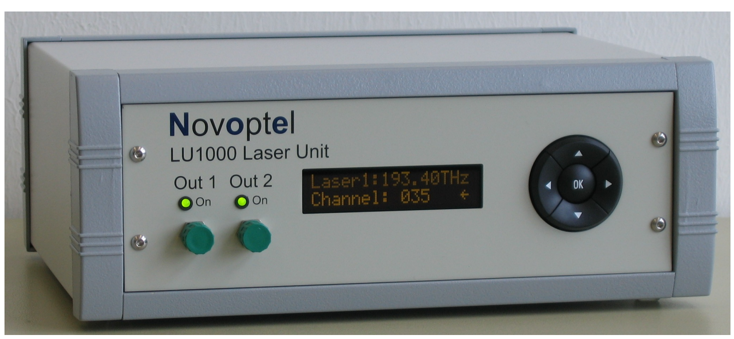


Figure 2.2: Laser Used During the Experiments. (Image from [14])

  The unit also contains a **receiver**, responsible for detecting the modulated light that returns to the interrogator after propagating through the sensing fiber and interacting with the physical field. For further analysis, it converts the optical signal into an electrical signal. To receive the SOP measurement, a polarimeter is used—see the next subsection.

Once the signal is received, a **processing** block extracts relevant information about the measured quantity. Then, the processed data is formatted and stored in a unit called **reporting** for further interpretation and visualization.

To interact with the sensing fiber, a **patch panel** block is present. It routes the probe light into the optical transit cable and maintains proper signal distribution.

- **Optical Transit Cable**

  This is where the probe light propagates. The optical transit cable functions as the medium for transmitting light between the interrogator and the remote sensor. In this research, and as mentioned earlier, the optical transit cable used is a SMF.

- **Remote Sensor**

  This is where the interaction with the physical field occurs. At the sensing location, external forces such as vibrations and acoustic waves act as inputs to the remote sensor. The output is the measured and modulated signal, which is sent back to the interrogator unit for further processing. In this thesis, the terms optical transit cable and remote sensor refer to the same entity. In the experimental field (see Subsection 3.4), fiber optic circuits are installed and utilized as a sensor to extract polarization measurements.

2.2 Polarimeter

The polarimeter is the core of the whole setup. The one used in this research, reported in Figure 2.3 and 2.4, is the PM1000 model by Novoptel. It is a high-speed instrument that is capable of performing the four Stokes parameters measurement with a sampling rate up to 100 MHz. It has three different normalization modes that the user can select depending on the measurement needs. The mode relevant to this thesis is the standard normalization mode. In this mode, the Stokes vector is normalized to unit length. By doing so, its tip will always appear on the surface of the Poincaré sphere, as illustrated Figure 2.5. The Figure shows also the user-interface that appears when performing tests— it is only a template, and the measurements are not taken from the experiments.

Figure 2.3: front side of the polarimeter. (Image from [15])

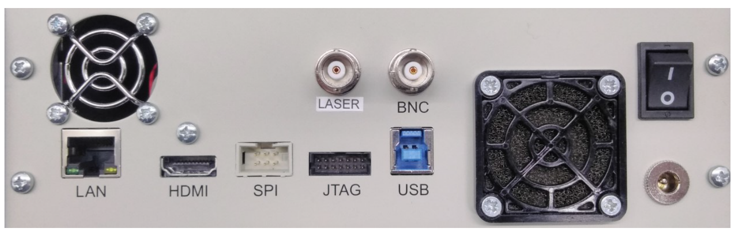


Figure 2.4: rear side of the polarimeter. (Image from [15])

Another hardware parameter that is relevant to the study is the average time exponent (ATE). The ATE parameter controls the internal averaging performed by the polarimeter after the 100 MS/s analog-to-digital (AD) conversion. The number of samples averaged is given by 2^{ATE}, where ATE can range from 0 to 20. Consequently, the sampling frequency can be computed from Eq. 2.1:

$$f_s = \frac{100 \text{ MS/s}}{2^{ATE}} \tag{2.1}$$

The maximum achievable sampling frequency occurs when no averaging is applied ($ATE = 0$), resulting in $f_s = 100$ MS/s, while the minimum sampling frequency is 95.4 S/s, found by substituting Eq. 2.1 with $ATE = 20$. It is worth noting that the software connected to the polarimeter, responsible for processing the acquired data, applies a downsampling operation, reducing the data rate to 100 Hz. As a result, the final dataset consists of samples spaced at 10 ms intervals.

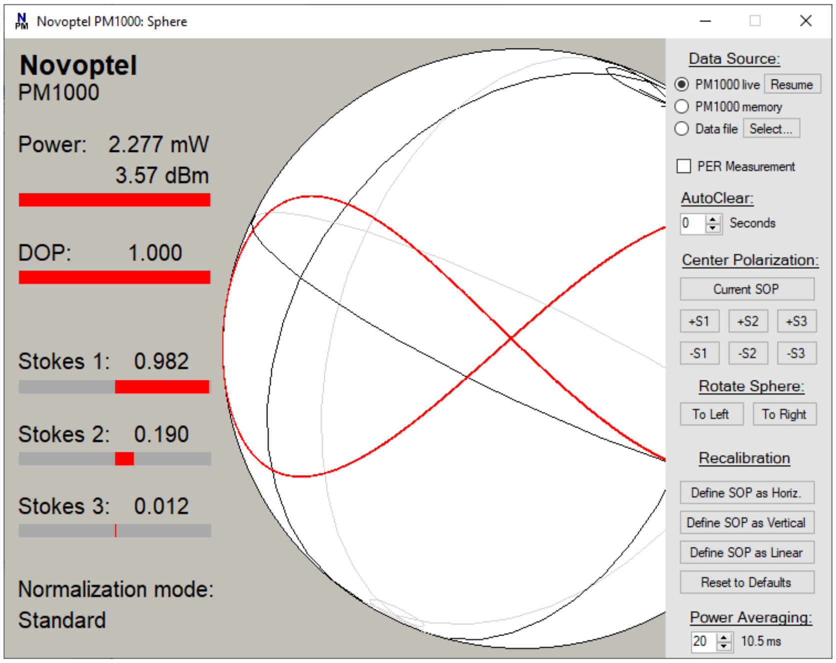


Figure 2.5: Poincaré Sphere on the User Interface. (Image from [15])

2.3 Detection Algorithm & DSP Scheme

As discussed early in Chapter 1, Section 1.6, the Stokes parameters are used to describe the state of polarization (SOP) of light. However, detecting polarization changes over time using these parameters is a challenge, as the variations occur in a three-dimensional space.

Additionally, when an event occurs, the changes in the Stokes parameters are often subtle, making direct detection difficult. To overcome these challenges, instead of analyzing the Stokes parameters directly, we compute the variation over time of the angle θ between two consecutive Stokes vector samples over the Poincaré sphere (see Figure 2.6). This is referred to as the angular speed (AS). This approach is simple since the computation has just one sample memory. Moreover, it reduces the dimensionality problem by evaluating a single dimension metric, while also enhancing sensitivity compared to direct Stokes parameter analysis.

For the computation of the angular speed, The Stokes vector $\vec{S}$ is consid-

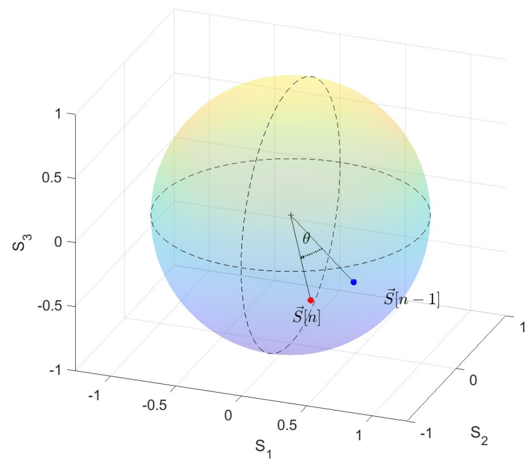


Figure 2.6: Poincaré sphere showing two consecutive Stocks vectors $\vec{S}[n]$ (in red), $\vec{S}[n-1]$ (in blue), and θ, the angle between them.

ered, defined as:

$$\vec{S} = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} \tag{2.2}$$

where the first parameter S_0 is used as a normalization factor and is not angle dependent. We denote the state of polarization angular speed (SOPAS) signal as $\Omega[n]$. Given the Stokes vector $\vec{S}[n]$, where n represents the discrete time index, the SOPAS signal can be represented by Eq 2.3:

$$\Omega[n] = f_s \cdot \arccos\left(\frac{\vec{S}[n-1]^\top \cdot \vec{S}[n]}{\|\vec{S}[n-1]\| \cdot \|\vec{S}[n]\|}\right) \tag{2.3}$$

Equation 2.3 is the fundamental computation used in all subsequent analyses. It represents the SOPAS signal evaluated on a sample-by-sample basis, serving as the foundation for all further tests. Note that the SOPAS is bounded by the utilized sampling frequency f_s. This is because in case the arc length drawn over the Poincaré sphere surface exceeds π, there will be undersampling [16]. Therefore, the maximum measurable SOPAS is:

$$\Omega_{\max} = f_s \cdot \pi \quad [\mathrm{rad/s}] \tag{2.4}$$

This upper limit, however, is usually sufficient if mechanical vibrations and stresses are considered [17].

After we have defined the SOPAS, we can introduce the DSP scheme of this research. This scheme, depicted in Figure 2.7, outlines the signal processing pipeline for analyzing the SOPAS data. First, the discrete time evolution of the Stokes vector $\vec{S}[n]$ are acquired by the polarimeter. Then, the dot product of the normalized Stokes vectors $\vec{S}[n]$ and $\vec{S}[n-1]$, which represente the SOP at consecutive time steps, is computed. The dot product is then passed through an arccosine function and multiplied by the sampling frequency to calculate the angular difference between the two SOP states, resulting in the angular speed.

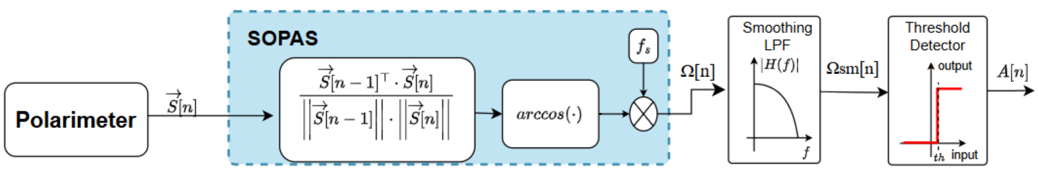


Figure 2.7: DSP Scheme for the Event Detection Algorithm.

After computing the angular speed, the signal undergoes a smoothing filter, primarily implemented as a moving average (more details will be given in the next Chapters). The moving average filter is used to attenuate the SOPAS fluctuations caused by environmental noise of the SOP. The smoothed signal, $\Omega sm[n]$, is then compared against a predefined threshold to detect events. If the angular speed exceeds the threshold for a certain duration, an event is detected, and the system generates an alarm signal, $A[n]$. Otherwise, if the angular speed does not exceed the threshold, no event is detected and alarm is not generated.

2.4 The Experimental Setup in the Gully

Now that we have the components which make up the experiment, it is time to describe the site where the experiment took place and where the data was collected. We will use Figure 2.8 as a reference to understand the layout of this setup.

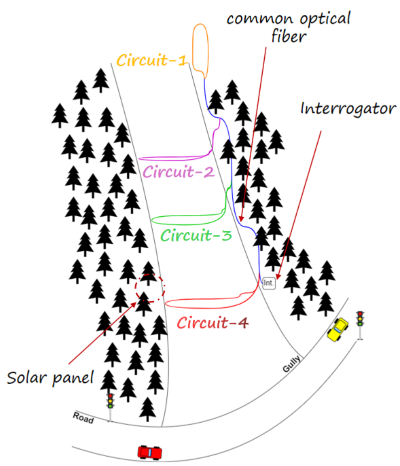
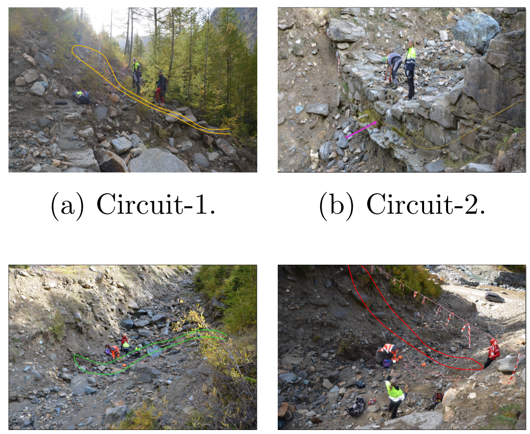


Figure 2.8: Experimental Setup in the Gully.



(a) Circuit-1. (b) Circuit-2.

(c) Circuit-3. (d) Circuit-4.

Figure 2.9: Overview of the fiber optic circuits layouts in the gully.

The chosen location is a gully situated above a roadway just outside a town in the Valle D'Aosta region of Italy, called Cogne. It is a natural waterway where rocks, debris, and water can slide due to gravity. This makes it an ideal site for investigating real-world environmental disturbances that may pose risks to the communities nearby. Regarding this, during the summer of 2024, specifically on Saturday, June 29, the Cogne (AO) area, particularly Valnontey, was severely impacted by a major flood event. This event triggered multiple debris flows along various gullies and caused the overflow of streams [18], including the gully where our setup is deployed.

The experimental setup consists of four fiber-optic circuits, labeled Circuit-1 to Circuit-4, which are placed separately along the gully:

- **Circuit-1** (orange): positioned vertically at the top of the gully on its right side, taking the road as a reference. Image in 2.9(a) shows the layout of circuit-1.

- **Circuit-2** (purple): placed horizontally further down the gully on a check-dam, as shown in image in 2.9(b).
- **Circuit-3** (green): shown in image in 2.9(c), is also horizontally positioned, located just below Circuit-2.
- **Circuit-4** (red): is the lowest circuit, situated closest to the road (approximately 30 meters) at the end of the gully. Image in 2.9(d) shows its layout.

These four circuits are linked by an additional optical fiber (blue), which connects them to the interrogator unit. The interrogator is installed on the right side of the gully, adjacent to Circuit-4. On the left side of the gully, between Circuit-3 and Circuit-4, a solar panels are placed providing energy to power on the interrogator. Figure 2.10 depicts the interrogator. Mounted above the interrogator is a video camera, which monitors the area to provide supplementary observational data.

2.5 The PolarAlp system

The PolarAlp system is a user-friendly interface used to monitor the environmental and structural changes in the gully setup, by detecting changes in SOPAS. The system is accessed through a web application. As illustrated in Figure 2.12, it is designed to display SOPAS data across the four fiber circuits ('CIRCUITO'). However, for now, only one circuit can be monitored at a time. Users can select a date range ('data inizio' stands for starting date, and 'data fino' for the ending date) and the number of samples to display (Misurazioni), with a hard threshold of 1000 samples. The web application also provides status indicators for the laser and polarimeter, including DOP, polarimeter's input power (Input power polarimetro), laser temperature (Temparatura laser), and laser optical power. These indicators are important to ensure the system operates within optimal parameters. Figure 2.13 shows the interface of these indicators. In the same Figure and in Figure 2.12, one can notice jumps and discontinuities in the data. The jumps are due to the first debugging phase that was made on the system, while the discontinuities are due to the burst mode. In this mode, only a segment of data are taken in a period of time. In this case, only a few seconds were recorded every 10 minutes.

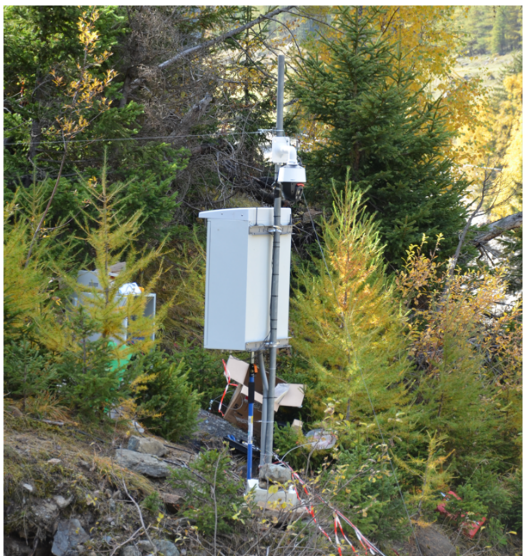


Figure 2.10: The Interrogator.

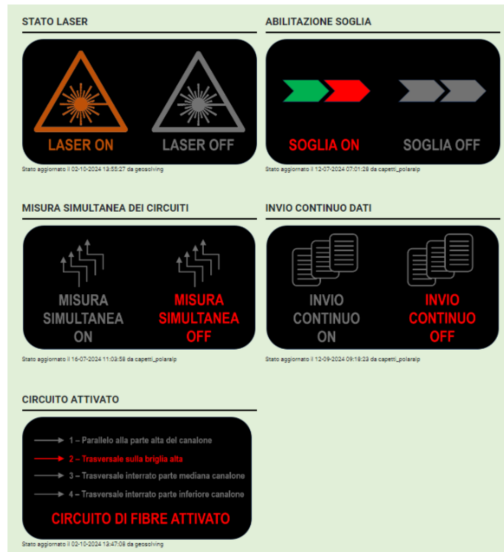


Figure 2.11: System Status Indicators.

2.5.1 System Status and Configuration

The PolarAlp system includes several status indicators and configuration options, as shown in Figure 2.11, to ensure accurate and reliable monitoring.

- **Laser Status (STATO LASER)**: The system allows users to turn the laser on or off, controlling the output optical power.

- **Threshold Comparison (ABILITAZIONE SOGLIA)**: Whenever SOPAS exceeds the threshold, two things happen:
  - The SOPAS data will be recorded,
  - An SMS and email will be sent to the monitor.
  In the analysis phase, it is turned off to avoid having SMS sent all the time.

- **Simultaneous Measurement (MISURA SIMULTANEA DEI CIRCUITI)**: Although the system currently does not support parallel monitoring of all four circuits, this feature is planned for future implementation.
- **Continuous Data Transmission (INVIO CONTINUO DATI)**: Users can choose between continuous data saving and burst mode, where data is recorded in intervals.
- **Activated Circuit (CIRCUITO ATTIVATO)**: This allows users to choose one of the circuits to be monitored.

The system also allows for parameter tweaking, such as setting threshold values, adjusting the moving average filter, and configuring the sampling frequency.

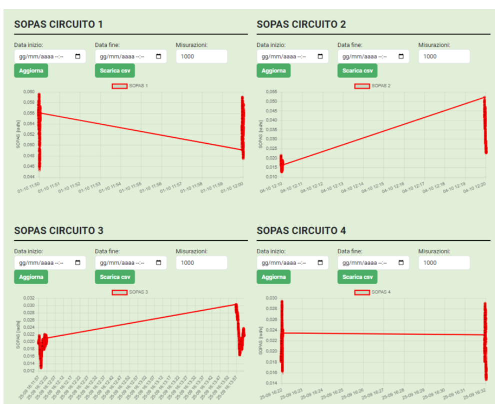


Figure 2.12: SOPAS Display.

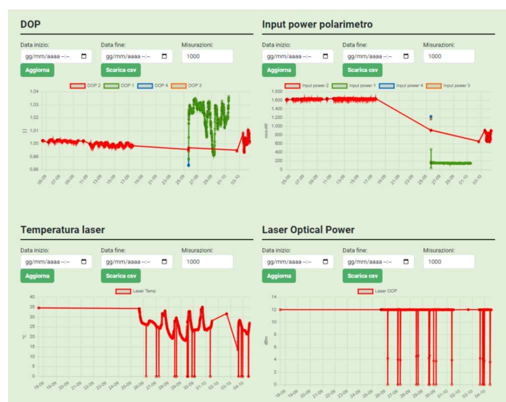


Figure 2.13: Status Display.

2.5.2 Data Extraction and Analysis

PolarAlp system features the ability to connect to the database, which contains the SOPAS data to be extracted and analyzed. The database acquires the data from the interrogator and by then the user can establish an SQL connection to the database using tools like DBeaver to read the data. This way is better than using the webapp interface on the PC to connect to the interrogator, which is slow. Once connected, users can execute SQL commands to extract data based on the number of samples, circuit number, and date range. The data can be exported in various formats, with CSV being the preferred option for further analysis. In the thesis, we used the webapp to monitor the system status and also to change the parameters, while we utilized the DBeaver database to download massive data.

Chapter 3

Rockfall Events Simulation

3.1 Introduction

On Tuesday, October 15th, experiments were performed to simulate and analyze the impact of falling rocks on fiber optic sensors deployed in a mountain gully. The experiment involved releasing blocks and rocks from different launch positions to investigate their interaction with the terrain and the fiber circuits. To replicate rockfall events, specifically designed blocks were employed. As illustrated in Figure 3.1, these blocks were crafted to closely mimic the rolling and falling behavior of natural rockfalls on a slope. Four such blocks, weighing 13 kg each, were utilized. Additionally, naturally occurring rocks of various sizes, collected from the gully, were tested to compare detection capabilities. To further assess sensitivity, finer materials such as gravel and sand were also introduced.

The experiments were executed in multiple phases, with the blocks and rocks being released from different positions along the gully. The selection of release positions aimed to realistically represent rockfall scenarios in relation to the fiber optic circuits. A total of five tranches (launch sequences) were conducted, each comprising several trials. Each trial was recorded through short video clips, which began a few seconds before the release and continued a few seconds after, ensuring that they could later be cross-referenced with the acquired data. From this point forward, we will use the following notations:

Figure 3.1: The Blocks Used in the Experiments.

- **Events**: The results of the rockfall experiments that we aim to detect.
- **Anomalies**: Any other SOPAS variations that are not relevant to our detection goals.

3.2 Data Smoothing

Before beginning the analysis, the data must be smoothed to mitigate the background noise and smooth fluctuations to define the events peaks better. Background noise refers to the oscillations that do not correspond to actual rockfall events but can still be detected by the fiber, even in quiet conditions where no events occur. Filtering techniques such as low-pass filters are commonly used to reduce its impact while preserving meaningful signals. However, these techniques introduce a delay in the processed signal. Figure 3.2 illustrates the noise present in the analyzed tranches (in blue), showing that its intensity is, for these experiments, generally very small and rarely exceeds 0.2 rad/sec. A deep analysis of the background noise will be presented in the next Chapter.

3.2.1 Moving Average Smoothing

The first approach to addressing this issue is smoothing the data using MATLAB's `movmean` function. The moving average computes the average of a set number of neighboring points and replaces each data point with this average.

For a given SOPAS signal $\Omega[n]$, the smoothed value at index n is computed as:

$$\Omega_{\text{sm}}[n] = \frac{1}{W} \sum_{i=n}^{n+W-1} \Omega[i] \tag{3.1}$$

where W represents the window size, i.e., the number of samples used for averaging. The window size is derived from:

$$W = T_w \times f_s \tag{3.2}$$

where T_w is the window size in seconds, and f_s is the sampling frequency. In this study, the sampling frequency is fixed at 100 Hz (samples per second). For instance, if we set the window size to $W = 50$ samples, then the moving average smooths the signal over:

$$T_w = \frac{50}{100} = 0.5 \text{ seconds} \tag{3.3}$$

3.2.2 Smoothing Effect

The results of applying the moving average on the noise are shown in Figure 3.2 (in red), where a considerable amount of fluctuation is removed. Additionally, Figure 3.3, which presents a zoomed-in view of an anomalous period, shows the smoothing effect on a segment with a detected event. It can be observed that the variation in event values is reduced, allowing a better peak definition. This peak enhances the ability to classify the events types more effectively.

However, this is not always the case. The moving average filter tends to flatten peaks as the window size increases.
Consider, for example, the event in Figure 3.4.

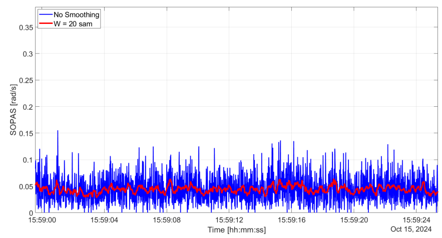


Figure 3.2: Noise (in blue) and its Smoothed Version (in red).

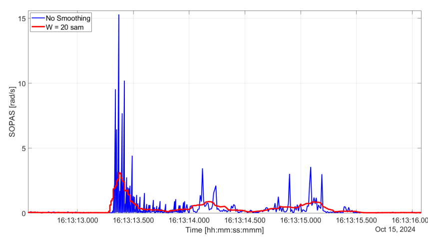


Figure 3.3: Smoothed Event (taken from Tranche 4).

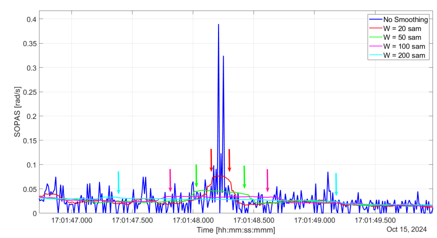


Figure 3.4: Effect of Smoothing with Different Window Sizes.

- Smoothing with $W = 20$ samples preserves the peak but may obscure details about the number of underlying spikes.
- Smoothing with $W = 50$ samples further reduces visibility, making the event less distinct.
- Smoothing with $W = 100$ samples nearly hides the event.
- Smoothing with $W = 200$ samples completely eliminates the event.

This is because the moving avergae filter acts like a low pass filter by replacing the current sample value with the average of W previous sample values, to reduce the impact of sudden spikes or rapid oscillations. This, however, raises concerns about the filter's ability to preserve important features. The excessive flattening observed with larger window sizes ($W = 50, 100, 200$ samples) underscores the need for careful selection of smoothing parameters to avoid significant loss of detail.

3.3 Tranche 1: Circuit-4 Tests

3.3.1 Tests Description

The first tranche was conducted upstream of Circuit-4, the closest circuit to the road (see Figure 2.9(d)). Table 3.1 provides details of the performed tests. Each test employed four blocks, with three individuals launching them simultaneously. Two participants held a single block each, while the third handled two.

Tranche	Test Code	Start Time	End Time	Material
1	01-04A	13:26:26	13:26:30	4 Blocks
1	02-04A	13:30:31	13:30:36	4 Blocks
1	03-04A	13:35:04	13:35:09	4 Blocks
1	04-04A	13:39:04	13:39:08	4 Blocks
1	05-04A	13:41:54	13:41:58	4 Blocks

Table 3.1: Summary of Tranche 1 Rockfall Tests.

In test 01-04A, the blocks were initially placed near Circuit-4 and then rolled onto the fiber using hand movements. The video analysis revealed that two blocks came to rest almost simultaneously, while the remaining two stopped milliseconds later. In test 02-04A, a similar procedure was repeated, but the blocks stopped shortly after their release. The participants then manually rolled them further by approximately one meter. Conversely, in test 05-04A, the blocks were released in parallel over the circuit, traveling from the left to the right side of the gully. They continued rolling until reaching the gully's center, where they came to a stop. In Figure **??**, the recorded SOPAS values corresponding to these tests can be seen. In the plot of this tranche, as well as in all other tranches, we use arrows of different colors to indicate events and anomalies:

- **Green arrows** point to SOPAS variations that have a high probability of being related to one of the tests (an event).
- **Orange arrows** indicate anomalies that are unlikely to correspond to any test and are instead attributed to unintentional contact with the fiber during the experiments.

- **Red arrows** highlight other spikes or noise present in the data.

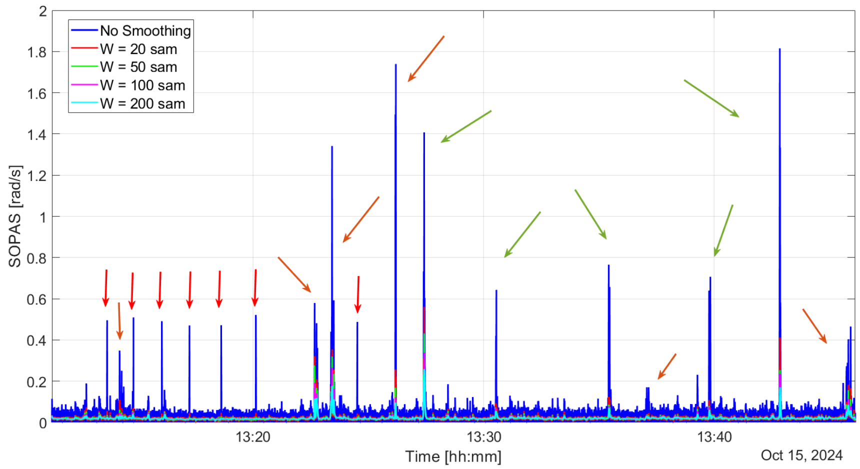


Figure 3.5: SOPAS Values of Tests in Tranche 1.

3.3.2 Analysis of Events

Based on observations from the experiments, we know that events typically last only a few seconds, beginning when the blocks or rocks are released and ending when they come to a stop. This information is crucial for defining and distinguishing between consecutive events.

Referring to the tables above, we note that the maximum duration of an event does not exceed six seconds. Furthermore, the recorded timestamps help determine the separation between two events. However, in some cases, a discrepancy was found between the recorded times in the videos and the actual timestamps of the extracted SOPAS values. This makes it difficult to definitively associate a given event with a specific test. Fortunately, the different materials used in the experiments—blocks, rocks, and gravel—provide additional clues for distinguishing events. Additionally, we assume that the number of blocks and rocks used in each test helps to identify the corresponding events.

As mentioned, all tests were conducted using four blocks. This suggests that the events may contain four distinct peaks. To detect and define these events, we will use Table 3.1 as a reference.

Test 01-04A started at 13:26:26. Looking at Figure 3.5, the closest detected event occurs at 13:26:10, which might correspond to this test. However, according to the table, we expect the second test's event to appear three minutes later, which is not the case, as the next detected event is at 13:26:45—only about 35 seconds apart. Instead, the following event at 13:30:32 appears exactly three minutes after the first, suggesting that this is the true event for test 01-04A. Consequently, the event at 13:30:32 corresponds to test 02-04A.

Following this pattern, we identify:

- The event for test 03-04A at 13:35:27.
- The event for test 04-04A at 13:39:47.
- The event for test 05-04A at 13:42:52.

When comparing these detected events to the recorded test times in the table, we observe a small time shift. However, this shift is consistent across all tests, strongly suggesting that these events correspond to the actual test events.

Figure 3.6 illustrates the event for test 01-04A. When smoothing the data, we observe two clear peaks and a third, slightly smaller peak, while some additional peaks are also present. Notice the effect of applying a large window size: the peaks are almost entirely removed.

3.3.3 Unexpected Anomalies Before the Tests

We also observe two additional anomalies before the first test:

1. The first anomaly starts at 13:22:40 and lasts for approximately 10 seconds (see Figure 3.7). When cross-checking with the table, we find no test matching this duration, suggesting that it may not correspond to any planned test. Instead, it could be attributed to people stepping on the fiber before the experiments began.

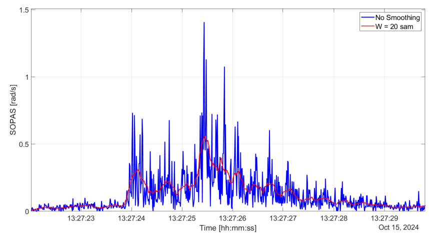


Figure 3.6: 01-04A Test's Event.

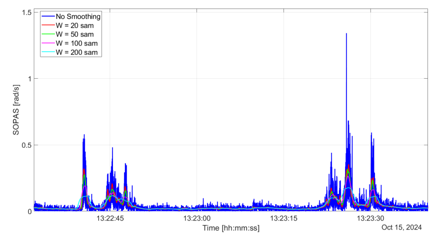


Figure 3.7: Unintentional Anomalies.

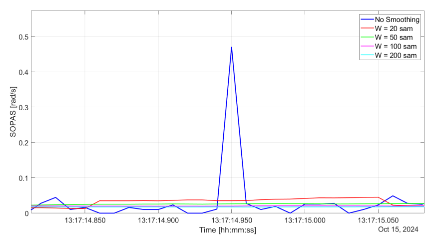


Figure 3.8: System Bug.

2. The second anomaly appears at 13:23:22 and also lasts for around 10 seconds. It exhibits three strong peaks (see Figure 3.7), which remain even when using the largest smoothing window (200 samples). This suggests that the anomaly may have been caused by three individuals stepping on the fiber simultaneously.

3. A third small spike is observed at 13:24:31, approximately one minute after the second anomaly. It is a single-sample peak, resembling a system-induced spike rather than a true event. It is entirely smoothed out with a window size of $W = 50$ samples. Similar spikes appear before the first recorded anomaly (indicated by red arrows in Figure 3.8).

These small spikes are classified as **system bugs**, likely caused by system failures. More details on system bugs and their treatment are provided in the next chapter.

3.4 Tranche 2: Circuit-3 Tests

3.4.1 Tests Description

The second tranche was performed upstream of Circuit-3 (see Figure 2.9(c)). This phase introduced both the designed blocks and naturally occurring rocks. A total of eight tests were executed, with Table 3.2 summarizing the test times. Video recordings were unavailable for tests 11-03B, 12-03B, and 13-03B.

Tranche	Test Code	Start Time	End Time	Material
2	06-03B	14:44:07	14:44:09	4 Blocks
2	07-03B	14:46:19	14:46:24	4 Blocks
2	08-03B	14:50:00	14:50:05	4 Blocks
2	09-03B	14:53:17	14:53:23	4 Blocks
2	10-03B	14:59:00	14:59:01	2 Rocks
2	11-03B	15:01:10	15:01:14	2 Rocks
2	12-03B	15:03:31	15:03:34	2 Rocks
2	13-03B	15:05:45	15:05:49	2 Rocks

Table 3.2: Summary of Tranche 2 Rockfall Tests.

In test 06-03B, the blocks were gently placed on the terrain and then left to roll over Circuit-3. Three blocks came to rest directly over the fiber circuit, while one continued rolling, eventually stopping approximately three meters away. A similar pattern emerged in tests 07-03B and 08-03B. However, in test 09-03B, two blocks were launched in a parallel manner to the fiber circuit, while the other two were released from positions similar to previous trials. One block veered away from the circuit due to encountering a large, smooth rock.

Test 10-03B introduced a different approach: two flat, regular rocks were thrown from an elevated position near the participants' heads. Due to their rectangular shape, they came to an abrupt stop over the fiber upon impact. The recorded SOPAS values for these tests are illustrated in Figure 3.9.

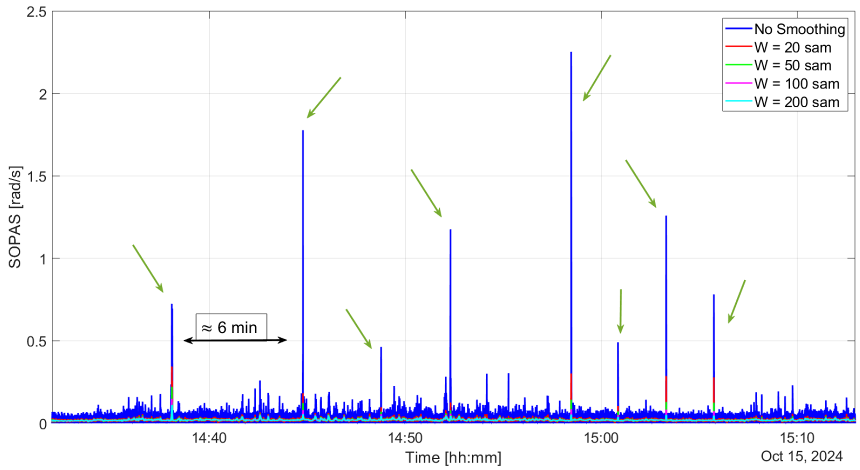


Figure 3.9: SOPAS Values of Tests in Tranche 2.

3.4.2 Analysis of Events

At first glance, Figure 3.9 appears to show eight distinct events, which matches the number of tests conducted. If the events followed the exact order of the tests, this would suggest that the first event corresponds to the first test in this tranche, 06-03B, the second event to the second test, and so on.

Examining the first event, illustrated in Figure 3.10, we observe that it has four distinct peaks, with the first being a small peak occurring 2 seconds before the second peak. This matches the number of blocks used in test 06-03B. However, this event occurs at 14:38:03, whereas Table 3.2 records the test at 14:44:07, indicating a time shift of approximately 6 minutes. To confirm whether this anomaly corresponds to test 06-03B, we need to check the time differences between subsequent events.

Unexpectedly, the time difference between the first two events—measured from the end of the first event to the beginning of the second—appears to be around 6 minutes and 45 seconds (see Figure 3.9), while Table 3.2 reports a difference of only 2 minutes and 10 seconds. Repeating this process for the second and third events, we find that the time differences align almost exactly. According to Table 3.2, the difference between tests 07-03B and

08-03B is 3 minutes and 36 seconds, which is also observed in the plot. Similarly, the time difference between the third and fourth events matches the interval between tests 08-03B and 09-03B. This pattern holds for all subsequent events, confirming that they correspond to their respective tests. The only discrepancy occurs with the first test.

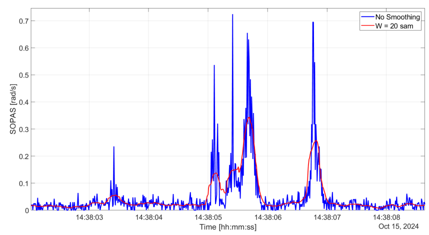


Figure 3.10: 06-03B Test's Event.

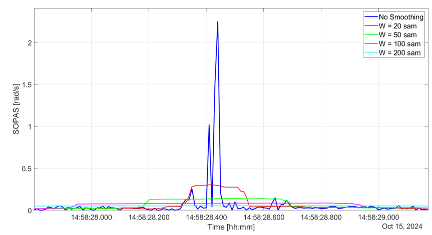


Figure 3.11: 10-03B Test's Event.

3.4.3 Highest SOPAS Intensity Event

The event with the highest SOPAS intensity in this tranche corresponds to test 10-03B, which involved two regular flat rocks. As described earlier, these rocks were thrown from an elevated position near the participants' heads, giving them greater momentum to strike above the fiber. This resulted in two distinct peaks, as shown in Figure 3.11. However, when applying a moving average filter, the event is smoothed to the extent that it appears as a single peak. Add to that it is short in time thus a filter will completely kill it.

3.5 Tranche 3: Check Dam Tests

3.5.1 Tests Description

The third tranche was conducted near the check dam (see Figure 2.9(b)). This series of tests included the four early-mentioned blocks, gravel, and sand, as listed in Table 3.3. Tests 18-02C, 20-02C, and 21-02C were not recorded. Figure 3.12 shows the corresponding SOPAS values.

In test 14-02C, two regular rocks, similar to those used in test 10-03B, were launched. They initially struck the upper edge of the check dam, just

Tranche	Test Code	Start Time	End Time	Material
3	14-02C	15:26:52	15:26:55	2 Rocks
3	15-02C	15:28:10	15:28:14	2 Rocks
3	16-02C	15:29:33	15:29:37	14 Small Rocks
3	17-02C	15:31:07	15:31:11	2 Rocks
3	18-02C	15:32:19	15:32:21	Gravel and Sand
3	19-02C	15:33:29	15:33:30	Gravel and Sand
3	20-02C	15:34:53	15:34:55	Gravel and Sand
3	21-02C	15:35:56	15:36:11	Hand

Table 3.3: Summary of Tranche 3 Rockfall Tests.

centimeters above the fiber circuit, before falling into the gully. In test 15-02C, the rocks rebounded twice—once on the upper edge above the circuit and again on the lower edge. Test 16-02C involved smaller rocks, which were first positioned on the check dam and then pushed onto the fiber circuit. In test 17-02C, two larger rocks were used: one impacted only the upper edge, while the other struck both edges, moving additional small rocks in the process.

Gravel and sand were thrown over the circuit in test 19-02C, while test 21-02C featured direct manual tapping (Hand) on the fiber to evaluate detection response. Notably, certain launches, particularly those involving finer materials (tests 16-02C and 18-02C), resulted in no detectable signals.

3.5.2 Analysis of Tests with Gavels and Sand

We now define the tests and focus on those conducted using gavels and sand, specifically tests 18-02C, 19-02C, and 20-02C. The last test, 21-02C, involved shaking the fiber manually to test the actual response. Indeed, if we examine the final detected event—ignoring the smaller peaks that follow for the moment—we can observe that its shape differs from all previous events (see Figure 3.13). This event lasted from 15:35:56 to 15:36:12, which exactly matches the recorded time in Table 3.3. This strongly suggests that this event corresponds to the last test of this tranche, 21-02C.

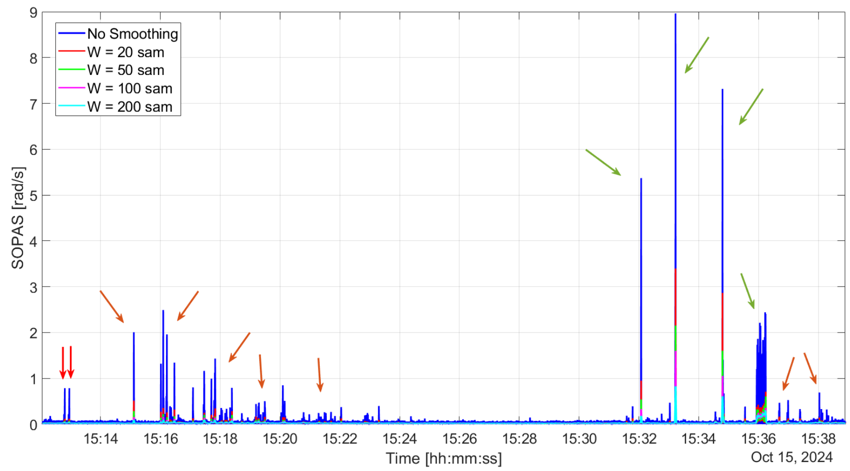


Figure 3.12: SOPAS Values of Tests in Tranche 3.

Taking this event as a reference, we can determine the timing of the other tests. According to Table 3.3, the time difference between tests 21-02C and 20-02C is 1 minute and 1 second. In the plot, we observe a very small peak 23 seconds away from the reference event, followed by a larger anomaly ending at 15:34:50. This is most likely the event corresponding to test 20-02C. Similarly, the time difference between tests 20-02C and 19-02C is 1 minute and 23 seconds—almost the same value observed in the plot. Additionally, the time difference between tests 19-02C and 18-02C is 1 minute and 10 seconds, both in the table and in the plot, suggesting this event corresponds to test 18-02C.

Regarding the other tests, it appears that the fiber did not detect them, as no events were found at the timestamps recorded in Table 3.3. This is probably because the rocks either:

- Didn't touch the fiber,
- Induced too weak vibration on the fiber by hitting the surrounding terrain.

3.5.3 Unexpected Anomalies and Possible Causes

At the beginning of this tranche's plot, we observe two identical anomalies approximately five minutes apart (see Figure 3.14). These do not match any of the planned tests and may have been caused by people moving around the fiber. The repeated anomaly could also be due to a system failure.

Additionally, we notice a series of anomalies occurring between 15:16 and 15:24. These could potentially be matched to certain tests, as they are spaced several minutes apart. However, it is also possible that these anomalies were caused by the movement of people unintentionally stepping on the fiber.

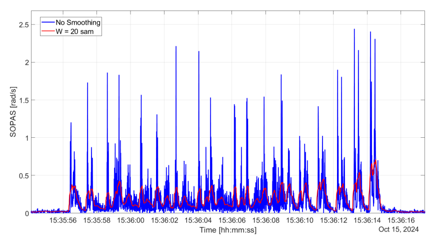


Figure 3.13: 21-02C Test's Event.

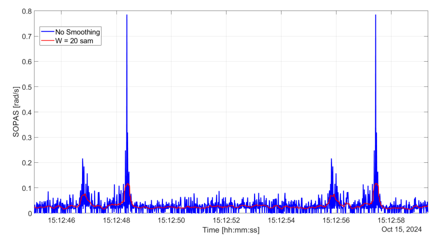


Figure 3.14: Duplicate Event.

3.6 Tranche 4: Circuit-1 Tests

3.6.1 Tests Description

This tranche was carried out downstream of Circuit-1 (see Figure 2.9(a)). Table 3.4 outlines the details.

Tranche	Test Code	Start Time	End Time	Material
4	22-01D	15:56:19	15:56:24	4 Blocks
4	23-01D	15:58:59	15:59:01	4 Blocks
4	24-01D	16:07:37	16:07:43	2 Small Rocks
4	25-01D	16:08:44	16:08:50	2 Small Rocks
4	26-01D	16:10:37	16:10:38	3 Small Rocks
4	27-01D	16:12:11	16:12:12	1 Small Rock

Table 3.4: Summary of Tranche 4 Rockfall Tests.

In test 22-01D, the four blocks were once again deployed. Three blocks continued rolling down the gully, while one stopped approximately three meters from its release point. Test 23-01D had a similar pattern. Instead, test 24-01D and 25-01D featured the release of two small rocks from a slightly higher position than in the previous test. These rocks rolled downward until halting. However, a fiber-cut alarm was triggered in test 25-01D, and a loss of data for about 30 seconds was observed. In test 26-01D, three small rocks were thrown; they traveled no more than three meters. In test 27-01D, a single small rock was launched, ultimately stopping over the fiber. Figure 3.15 displays the SOPAS values for this tranche.

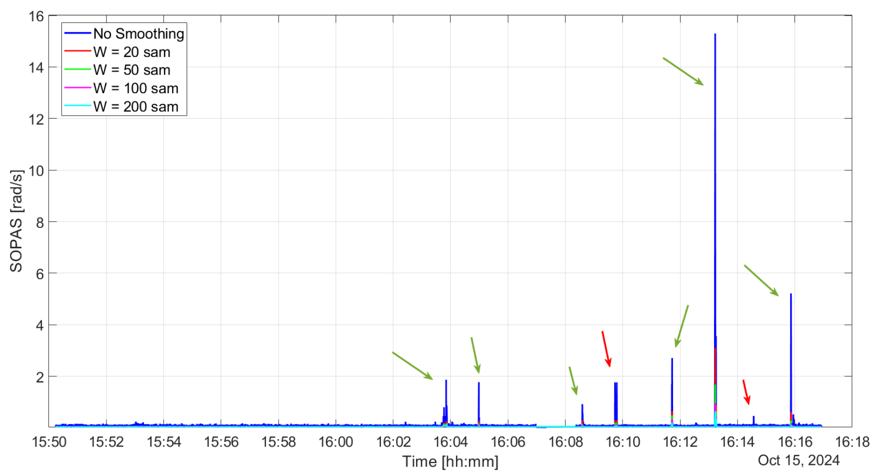
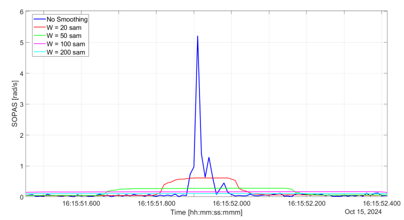


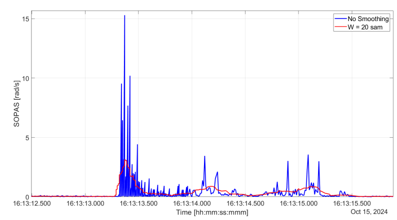
Figure 3.15: SOPAS Values of Tests in Tranche 4.

3.6.2 Analysis of Events

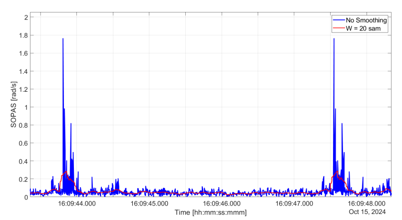
Figure 3.15 shows two distinct positions where events occur, identified as the largest and second largest in SOPAS intensity among all other anomalies. A zoomed-in view of the first event (see Figure 3.16b) reveals three distinct peaks, with the first one being the highest. In contrast, a zoomed-in view of the second event (see Figure 3.16a) displays only a single peak. The time difference between these two events is approximately 160 seconds. Comparing this to Table 3.4, we can infer that these events correspond to tests 26-01D and 27-01D, respectively. This conclusion is based on the number of rocks used in each test—three rocks for 26-01D and one rock for 27-01D—which aligns with the detected signal patterns.



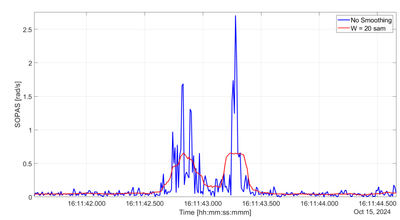
(a) 27-01D Test's Event.



(b) 26-01D Test's Event.



(c) Duplicate Anomaly



(d) 25-01D Test's Event.

Figure 3.16: Events and Anomalies from 01D Tests .

Now, let us again examine Figure 3.15, which reveals five distinct positions, before tests 26-01D and 27-01D, where events are present. In the first two tests, four blocks were used, which suggests that the corresponding events might contain four peaks. A zoomed-in view of the first event, as seen in shows two separate sets of peaks, occurring approximately two seconds apart. Each set contains multiple peaks. When blocks are thrown, we expect one peak per block, but in this case, the blocks bounced, resulting in additional detected peaks. This event may correspond to the first test in this tranche, 22-01D.

The neighboring event occurs approximately 70 seconds after the end of the first event. A zoomed-in view revealed four peaks. Despite a time discrepancy of more than two minutes between this event and the expected test time (see Table 3.4), it is the only event with four peaks, strongly suggesting that it corresponds to test 23-01D. In the next two tests, two rocks were used per test. The third detected event consists of two peaks, as expected. This event likely corresponds to test 24-01D.

3.6.3 Duplicate Anomaly

A problem arises in identifying the fourth test, 25-01D, because data loss was observed during this test. Indeed, the recorded data confirms this issue, as the software appears to have duplicated samples upon restarting. This results in an unusual pattern of two nearly identical peaks (see Figure 3.16c). Since data was lost, it is possible that the timestamps were also affected, making it difficult to confirm whether this anomaly belongs to 25-01D or not.

Finally, a zoomed-in view, shown in Figure 3.16d, reveals a two-peak event. However, the time difference compared to the recorded time in Table 3.4 is quite large. Given the circumstances, the most reasonable conclusion is that this is the event which corresponds to 25-01D. Due to the system failure, a time shift likely occurred, causing the event to appear later than expected.

3.7 Tranche 5: Additional Circuit-4 Tests

3.7.1 Tests Description

In the final tranche, additional tests were performed at Circuit-4. Four trials were conducted, in which blocks were released in a parallel manner over the fiber circuit, originating from the left side of the gully. Table 3.5 shows the tests.

Tranche	Test Code	Start Time	End Time	Material
5	28-04A	17:02:10	17:02:14	4 Blocks
5	29-04A	17:04:05	15:04:11	2 Blocks
5	30-04A	17:06:08	17:06:13	4 Blocks
5	31-04A	17:08:39	17:08:44	4 Blocks

Table 3.5: Summary of Tranche 5 Rockfall Tests.

Test 29-01D utilized only two blocks: one stopped in the middle of the gully over the circuit, while the second traveled only half the distance of the first. In test 30-01D, four blocks were released, all coming to rest near the center of the gully, a pattern that repeated in test 31-01D. Figure 3.17 illustrates the SOPAS values for the remaining tests in this tranche.

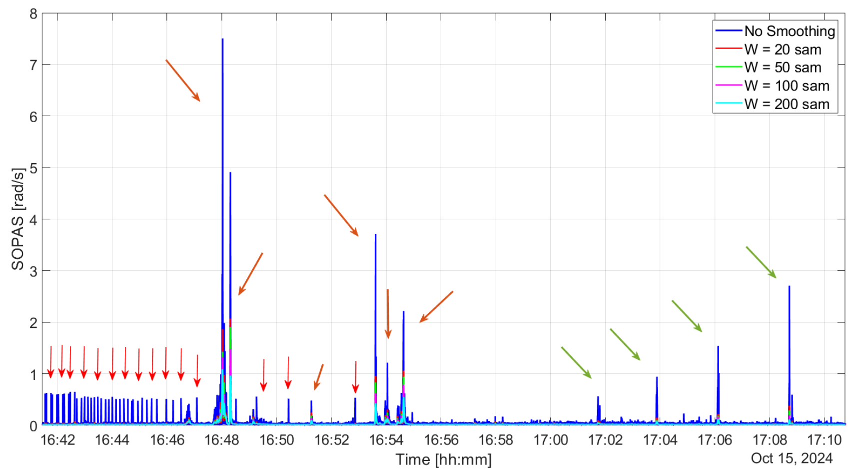


Figure 3.17: SOPAS Values of Tests in Tranche 5.

3.7.2 Analysis of Events

In the first two tests of this tranche, only two blocks were used, whereas in the last two tests, four blocks were used. The time difference between the events is nearly constant, approximately 2 minutes. Four events appear as expected based on the description. Interestingly, SOPAS intensity increases progressively, as illustrated in Figure 3.18. There is, however, a slight time shift. The first event, which likely corresponds to test 28-04A, starts at 17:01:44, whereas its recorded time in Table 3.5 is 17:00:10. Despite this shift, the difference is relatively small.

Notably, a series of spikes appears at the beginning of the time axis. These are system bugs and are entirely smoothed out when applying the moving average filter. After the system bug artifacts, five significant anomalies can be observed. The first, centered around 16:48:00, exhibits the highest SOPAS values among all events and anomalies encountered across all tranches (see Figure 3.19). The second anomaly follows 18 seconds later as seen in Figure 3.19, while the third, fourth, and fifth occur at 16:53:37, 16:54:00, and 16:54:38, respectively.

Since this tranche contains only four tests, these additional anomalies are likely due to human interference, with a high probability that individuals unintentionally stepped on the fiber.

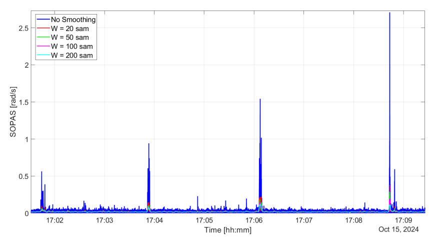


Figure 3.18: Event of the Last Four Tests (28/29/30/31).

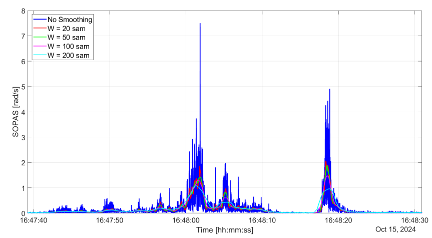


Figure 3.19: Highest SOPAS Intensity Anomaly (+ the one next to it).

Discussion

Throughout the analysis, we confirmed that the system is indeed capable of detecting rockfall events across all circuits. Despite the complexities involved, the SOPAS-based sensing approach successfully identified events, demonstrating the sensitivity of the system to rockfall-induced disturbances. This validation is a crucial step, as it confirms that the technology is functional and capable of capturing meaningful data. The challenges we encountered primarily concern system-level reliability rather than the fundamental ability to sense events, that can be addressed with further improvements.

The challenge lies in categorizing events based on SOPAS data alone. While the system detects variations effectively, it does not inherently provide information about the specific cause behind each event. Further data acquisition and analysis are necessary to develop a more precise classification approach. Additionally, the application of smoothing filters, such as the moving average, does not always yield ideal results. Even though smoothing helps to reduce noise, it can also obscure significant details of the anomalies. Specifically, it tends to broaden peaks, making them less distinct. This introduces uncertainty in determining whether a peak represents an actual event or a residual effect of the smoothing process. Future improvements in filtering techniques was made to enhance the accuracy of event detection, and the results will be discussed in the next Chapter.

Despite these challenges, the analysis has provided valuable insights into the effects of different rockfall events on the fiber circuit. By cross-referencing experimental observations, test conditions, and SOPAS data, we have been able to assess and interpret the detected events, paving the way for further refinements in the system.

3.8 Time and Frequency Analysis

To further analyse of the SOPAS data, we initially relied on the power spectrum to investigate frequency components. Later, we introduced the wavelet transform to refine our understanding of time-frequency variations. The transition between these two methods was driven by the need for better localization of transient events, which are essential for detecting events in the gully experiments.

3.8.1 Power Spectrum Analysis

The power spectrum, typically obtained through the Fourier Transform, helps identify dominant frequency components in a signal. This method was our initial choice for analyzing SOPAS data. The main aim of using the power spectrum was to identify specific shapes or patterns associated with events, aiding in the classification of different event types. As previously discussed, the material used in the experiments may influence the shape of the SOPAS signal, and the manner in which these materials were released may help determine the frequency components of each event. Since the SOPAS signal consists of both noise fluctuations and events, the power spectrum can allows us to examine the dominant frequency components. We could observe variations in spectral energy distribution by applying the power spectrum to different test cases which can offer us insights into the behavior of the SOPAS signal under different conditions.

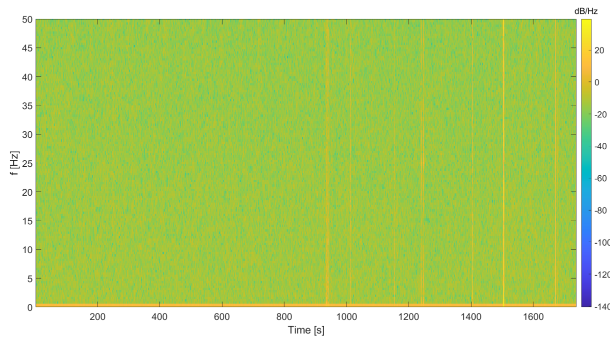


Figure 3.20: PSD of Circuit-1 Tests.

For this purpose, we used the `stft()` function in MATLAB (Short-Time Fourier Transform). Figure 3.20 shows the power spectrum for Circuit-4 tests using raw SOPAS data. Note that the frequency range inside which these spectrograms are represented goes from zero to half fs because of the Nyquist theorem.

We can now observe the frequency components of the events. As established earlier, seven distinct events are present—six corresponding to the six tests of Circuit-4 and one caused by system failure. The power spectrum shows the energy exhibited by both the background noise and the events. The background noise has values around 1-Hz which is reasonable as the natural fluctuations in SOPAS data occur mainly at this frequency. Some events, like 26-01D, exibits higher power than the rest, which may indicate that the thrown blocks had a bigger impact on the fiber. However, the power spectrum does not provide much information about the shape of this event. The frequency components of the 26-01D test (see Figure 3.21) are not clearly distinguishable.

The power spectrum can offer a global frequency representation by indicating which frequencies contain significant energy, but it does not provide temporal localization of transient events. This limitation became evident when analyzing rockfall events, as different phases of the rockfall could generate varying frequency content over time. It suffers from the fixed FFT window size trade-off, where:

- Narrow windows provide good time resolution but poor frequency resolution.
- Wide windows provide good frequency resolution but poor time resolution.

These challenges led us to explore an alternative approach—the wavelet transform.

3.8.2 Wavelet Transform Analysis

To overcome the time-frequency trade-off inherent in Fourier-based methods, we employed the wavelet transform. The wavelet transform decomposes a signal into wavelets—localized oscillatory functions that vary in both scale (frequency) and position (time). This decomposition is achieved using a *mother wavelet*, a prototype function that is scaled to analyze the signal across different time-frequency resolutions. The transformed signal provides information about the time and the frequency. Therefore, wavelet-transformation contains information similar to the STFT, but with additional special properties of the wavelets, which show up at the resolution in time at higher analysis frequencies of the basis function.

Mathematically, the *Continuous Wavelet Transform* (CWT) of a signal $\omega(t)$ is defined in Eq 3.4:

$$W(a,b) = \int_{-\infty}^{\infty} \omega(t) \cdot \frac{1}{\sqrt{a}} \psi^* \left(\frac{t-b}{a} \right) dt. \tag{3.4}$$

where a is the *scale parameter* (inversely related to frequency), b is the *translation parameter* (related to time), and $\psi(t)$ is the *mother wavelet*. The wavelet coefficients $W(a,b)$ provide a joint time-frequency representation of the signal, allowing for the identification of transient features and their evolution over time.

Unlike power spectrum analysis, wavelets transform enabled a precise localization of sudden changes in the signal, which offers a better and deeper insight into the underlying physical processes. We used the `cwt()` function in MATLAB for this analysis. The wavelet transform revealed frequency bursts at specific moments, which were not clearly visible in the power spectrum. Figure 3.22 illustrates the wavelet transform of the 26-01D test.

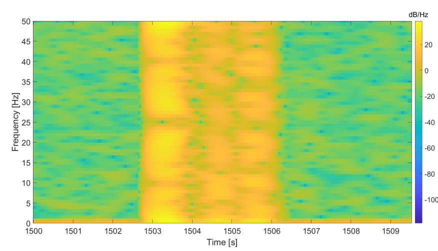


Figure 3.21: PSD of 26-01D Test.

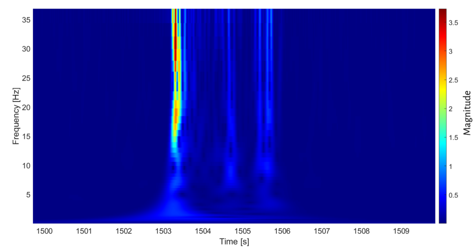


Figure 3.22: WT of 26-01D Test.

This test featured three peaks, with the first peak exhibiting higher intensity than the other two. While the power spectrum in Figure 3.21 does not clearly highlight this difference, the wavelet transform allows us to identify both low- and high-frequency components through a colormap representation. Additionally, it enables precise determination of the frequency values at the exact moments when the rocks impacted the fiber optic.

This capability plays a crucial role in event detection. The wavelet transform allows us to classify different types of events by analyzing their frequency characteristics over time. For instance, the weight of the rocks could be inferred from the variations in the SOPAS signal. Heavier rocks exert greater pressure on the fiber, leading to more pronounced frequency variations. This information could be useful to refine event classification and enhance the reliability of event detection.

Chapter 4

Quiet Conditions

4.1 Introduction

The experiments explained in Chapter IV provide a foundation for simulating a rockfall event scenario that may occur. A major advantage of these experiments is that the SOPAS intensity is significantly lower than that of real events in the gully. This is because we used small blocks and rocks, whereas, in a real scenario, the rocks could be much larger and more severe. This helps us establish a threshold that can detect such events, maximizing correct detections while minimizing false detections, which could lead to false alarms.

Before defining the threshold and its working mechanism, we need to examine another state of our data, referred to as "quiet conditions." This state represents the SOPAS values recorded by the polarimeter when no real events or experiments are occurring. We can also refer to it as the background noise, which was briefly discussed in the previous chapter. Analyzing this noise is crucial because we want our detector to identify only dangerous events without being triggered by mere background noise. Additionally, it serves as a reference for comparison and is essential for establishing a baseline in event detection.

4.2 System Bugs

Analyzing noise can sometimes be a challenging task. The SOPAS data, collected from fiber optic circuits, consists of time-indexed angular speed values. While the majority of these values represent real physical phenomena, some errors may arise due to:

- Sudden spikes or glitches,
- External disturbances affecting polarization,
- Missing or corrupted data,
- Noise interference from environmental conditions.

What we encountered in our data is a sudden, isolated spikes that deviate significantly from neighboring SOPAS values, which we refer to by **system bugs**. These bugs are neither part of the natural signal variations nor events caused by rockfall experiments. Figure 4.1 illustrates a sample taken during what is considered a quiet conditions, where this large number of events is not expected. A zoomed-in view (see Figure 4.2) shows the nature of these bugs.

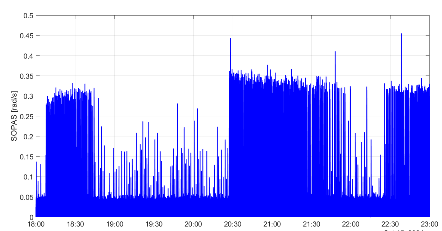


Figure 4.1: SOPAS Values of 15-10-2024, From 18:00 to 23:00; Circuit-2.

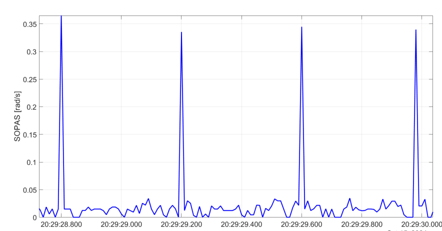


Figure 4.2: One-sample Spikes (bugs) - Zoom in on Figure 4.1

Notably, a bug is a peak or spike lasting for only one sample. It is usually abrupt, short-lived, and uncorrelated with surrounding data points. These bugs originate from the system hardware and may be caused by a lack of calibration in the polarimeter. Due to temperature variations outdoors, the polarimeter experiences sudden temperature fluctuations. if it is kept in a lab where the temperature remains nearly constant or changes gradually, it is less likely to undergo decalibration.

4.2.1 Moving Average Smoothing

The first approach to addressing this issue is smoothing the data using the `movmean` function in MATLAB, which we have seen the previous Chapter. Figure 4.3 shows the results of applying the moving average filter to noisy data. The limitation of this technique is that it is not dynamic. As indicated by the arrows in Figure 4.3, it uniformly smooths but fails to account for fluctuations where bugs values are high in certain periods and low in others.

A closer inspection reveals that while the moving average reduces noise in the SOPAS signal, it struggles with high-intensity system bugs. This is true even for larger windows (W = 50, 100, and 200 samples). Instead of fully suppressing these bugs, it merely smooths them, leaving behind residual peaks, as shown in Figure 4.4. Although these peaks have lower intensity compared to the original ones, they can still resemble events, which may lead to false detections.

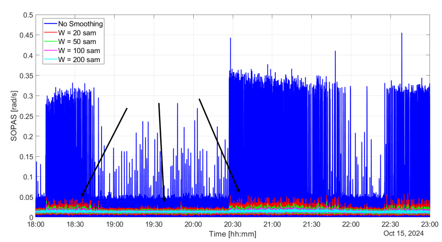


Figure 4.3: Raw SOPAS values vs. Filtered SOPAS using *movmean*.

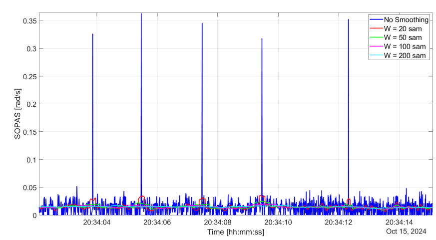


Figure 4.4: Zoom in on Figure 4.3 reveals the smoothed bugs.

4.2.2 Dynamic Thresholding

To address these issues, a dynamic thresholding method based on standard deviation within a sliding window was employed. This method is based on observations from experimental trials, which revealed that events follow a distinct pattern involving multiple samples, unlike single-sample system bugs (see Figure 3.16b above). When a possible event is detected, the variation between its SOPAS values tends to be small, while in case of the bugs it is a one-sample abrupt change.

The goal of the dynamic thresholding was to detect and correct this values that significantly deviated from their neighboring points while accounting for the intensity of the bugs in particular periods. The approach follows these steps:

Sliding Window Approach

Instead of evaluating each point in isolation, the algorithm processes the SOPAS data using a sliding window technique. This localized approach ensures adaptive thresholding for identifying events.

First, we **define the window**:

- A variable-sized window is used to analyze local statistics. This ensures that the threshold adapts to local fluctuations.
- The window extends symmetrically around each data point, bounded by the dataset's length.

Second, we **extract local data**:

- Within each window, a subset of the data centered around the current point is extracted.
- This subset represents the "local neighborhood" of the point under evaluation.

Statistical Calculation & Dynamic Thresholding

Within the sliding window, the algorithm computes the local mean and standard deviation. The mean provides a central reference for detecting deviations, while the standard deviation quantifies data variability. Mathematically, the standard deviation of a SOPAS dataset with n samples is given by:

$$\sigma_{\text{local}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(SOPAS_i - \mu_{\text{local}})^2} \tag{4.1}$$

Here, σ_{local} refers to the standard deviation taken within the sliding window. Likewise, μ_{local} is the mean of the values in the window.

Now, we apply a dynamic threshold, defined as:

$$T_{\text{dynamic}} = k \cdot \sigma_{\text{local}} \tag{4.2}$$

where k is a user-defined multiplier (`k_factor`) controlling sensitivity to outliers.

By setting the dynamic threshold using Eq 4.2, only significant deviations from local patterns are identified as events. A higher `k_factor` allows more variation, reducing false positives, while a lower `k_factor` makes detection stricter.

Anomaly Detection & Correction

For each SOPAS data point, $SOPAS(j)$, the algorithm compares it to its two immediate neighbors, $SOPAS(j-1)$ and $SOPAS(j+1)$. The bug detection rule is:

$$\left| SOPAS(j) - \frac{SOPAS(j-1) + SOPAS(j+1)}{2} \right| > T_{\text{dynamic}} \tag{4.3}$$

If the absolute difference exceeds the dynamic threshold, the value is deemed a bug. To correct bugs, $SOPAS(j)$ is replaced with the mean computed within the window (μ_{local}).

This approach effectively removes system bugs while accounting for their quantity. Figure 4.5 illustrates cases before and after removing the bugs. Furthermore, this approach account successfully for the presence of events without removing them. Figure 4.6 presents results from tranche 5 tests. On the left side, the presence of system bugs is evident, but they are effectively removed by the algorithm. On the right side, the algorithm is applied to an actual event. Here, we can clearly observe that while the event is smoothed, its overall structure and characteristics remain intact.
Additionally, we have tried alternative correction methods to handling the system bugs such as:

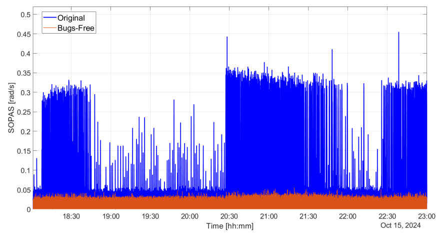


Figure 4.5: Raw SOPAS vs. Filtered SOPAS using the Dynamic Thresholding.

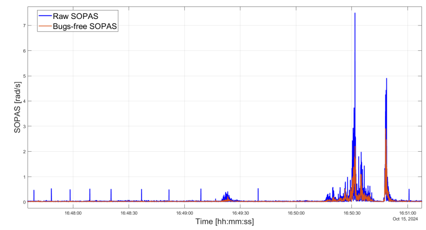


Figure 4.6: Raw SOPAS vs. Filtered SOPAS using the Dynamic Thresholding (Tranche 5).

- Replacing the bug with the **median** of the window instead of the mean.
- Replacing the bug with the **smaller** of its two neighboring values.

However, trials for different datasets suggest that the best method is the local mean replacement.

4.3 Filtering the SOPAS Data

The moving average filter performs well in reducing noise, but it tends to flatten peaks, making event detection more difficult. To address this limitation, we applied various filters to the data to determine the most effective one. In this section, we present, along with the moving average filter, the two other filters used in our analysis: the Savitzky-Golay (Sgolay) filter and the Median Filter.

4.3.1 Savitzky-Golay Filter

The Savitzky-Golay filter operates by fitting a low-degree polynomial to a moving window of data points using the least-squares method. This approach allows for smoothing without destroying essential characteristics of the signal [19]. The Sgolay filter performs a local polynomial regression; this enables it to reduce noise and maintain the structure of the original signal, unlike traditional moving average filters, which may distort sharp features.

It has two parameters: a sliding a window of fixed size over the data, and a polynomial fit to the points within this window. The value of the polynomial at the central point of the window is then taken as the smoothed value. This process is repeated for each point in the data set, resulting in a smoothed signal. The mathematical formulation is as follows:

- **Polynomial Fitting:** For a data point y_i, a window of size $2m + 1$ (centered at y_i) is selected. A polynomial of degree n is fitted to the data points in the window:

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

  where x represents the index of the data points within the window, and $a_0, a_1, \ldots, a_n$ are the polynomial coefficients.

- **Matrix Representation:** The polynomial fitting can be expressed in matrix form as:

$$\mathbf{y} = \mathbf{X}\mathbf{a}$$

  where:

  - $\mathbf{y}$ is the vector of data points in the window:

$$\mathbf{y} = [y_{i-m}, y_{i-m+1}, \ldots, y_{i+m}]^T.$$

  - $\mathbf{X}$ is the design matrix, with rows representing powers of the indices x:

$$\mathbf{X} = \begin{bmatrix} 1 & (-m) & (-m)^2 & \ldots & (-m)^n \\ 1 & (-m+1) & (-m+1)^2 & \ldots & (-m+1)^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & m & m^2 & \ldots & m^n \end{bmatrix}.$$

  - $\mathbf{a}$ is the vector of polynomial coefficients:

$$\mathbf{a} = [a_0, a_1, \ldots, a_n]^T.$$

- **Least Squares Solution:** The polynomial coefficients $\mathbf{a}$ are obtained by solving the least squares problem:

$$\mathbf{a} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}.$$

  The smoothed value $\hat{y}_i$ at the center of the window is given by:

$$\hat{y}_i = p(0) = a_0.$$

- **Convolution with Filter Coefficients:** The Sgolay filter can also be implemented as a convolution with precomputed filter coefficients $\mathbf{c}$, derived from the first row of $(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$:

  $$\mathbf{c} = [c_{-m}, c_{-m+1}, \ldots, c_m].$$

  The smoothed value $\hat{y}_i$ is then computed as:

  $$\hat{y}_i = \sum_{k=-m}^{m} c_k y_{i+k}.$$

Implementation in MATLAB

In this thesis, the Sgolay filter is applied to SOPAS data to smooth fluctuations using MATLAB's built-in `sgolayfilt` function, with parameters tuned to optimize the balance between noise reduction and event retention. Specifically, a polynomial order of p and a window length of W were chosen based on empirical evaluation of the dataset. An example implementation in MATLAB is:

```
% Define filter parameters
p = 3; % Order of polynomial
W = 11; % Filter's window (must be an odd number)

% Apply the Savitzky-Golay filter
filtered_SOPAS = sgolayfilt(SOPAS_data, p, w);
```

4.3.2 Median Filter

The median filter is a non-linear filtering technique used to reduce noise while preserving edges in a signal. It replaces each data point with the median value of its neighboring points within a defined window. This makes it effective for removing impulsive noise, such as sudden spikes in SOPAS data.

Implementation in MATLAB

MATLAB provides a built-in function called `medfilt1`, used for applying a one-dimensional median filter. The syntax for using this function is:

```
filtered_data = medfilt1(raw_data, window_size);
```

where:

- `raw_data` is the original time-series signal (e.g., SOPAS data).
- `window_size` specifies the number of neighboring points used for computing the median.

4.3.3 Performance of Filters on SOPAS Data

Figure 4.7 presents the raw SOPAS data (in blue) alongside three different filtering techniques: Sgolay filter (red), moving average filter (green), and median filter (black). The choice of window size is crucial: a small window preserves finer details but may not remove all noise, whereas a larger window provides stronger filtering but risks suppressing valid rapid fluctuations. Therefore, selecting an optimal window size depends on the nature of the noise and the characteristics of the data. For this specific burst of data, the filters were applied with the same window size ($W = 20$, except for the Sgolay filter, which requires an odd window size, $W = 21$) to allow a fair comparison.

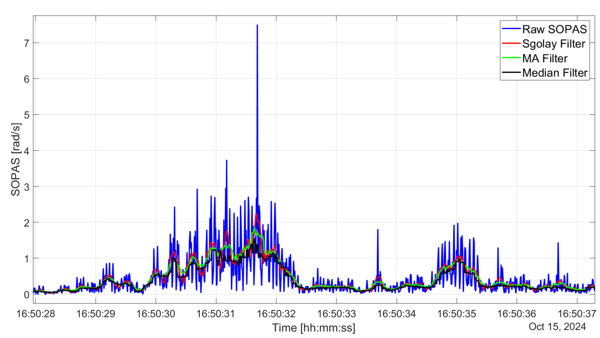


Figure 4.7: Raw SOPAS vs. Filtered SOPAS .

Each filter can be assessed based on how well it smooths the data while retaining important features such as peaks and events. They are applied to the unfiltered SOPAS data, which exhibits high-frequency noise with sharp spikes and fluctuations, making it difficult to identify significant events or trends.

The moving average filter reduces random noise and produces a smooth trend. This makes it useful for capturing general patterns in SOPAS data while suppressing small fluctuations. However, it blurs sharp peaks, making it difficult to detect sudden events. The median filter, on the other hand, is effective at reducing noise, making it ideal for reducing extreme spikes and system bugs. However, it lacks flexibility, as it can slightly distort the signal by flattening rapid oscillations.

The Sgolay filter smooths the signal while better preserving peaks and transitions compared to both the moving average and median filters. It retains sharp variations, making it useful for detecting fast changes in SOPAS data. Sgolay filter might sometimes result in negative values, which can be an issue in computaion, when using logarithmic scale, for example. It smooths the data by fitting a local polynomial to a sliding window of points. If the original data has values close to zero and the polynomial fit oscillates slightly, it can produce negative values in regions where the true signal should remain non-negative, as shown in Figure 4.8. Also, instead of just smoothing the signal, it can sometimes amplify small variations by overfitting the local polynomial. This can cause sharp oscillations, which might lead to negative values. Figure 4.9 shows an example of this.

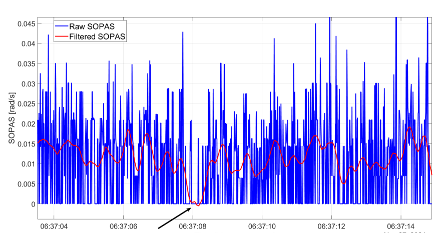


Figure 4.8: Drawbacks of Sgolay Filter (Example 1).

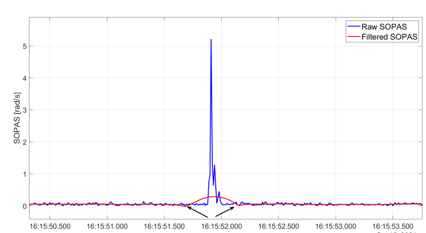


Figure 4.9: Drawbacks of Sgolay Filter (Example 2).

4.3.4 Histogram Analysis of Filtered Data

To further analyze the distribution of SOPAS values, histograms were plotted to assess how effective is each filter in altering the data distribution. Figure 4.10 shows the overlaid histograms of the three filters applied to SOPAS data from Tranche 5.

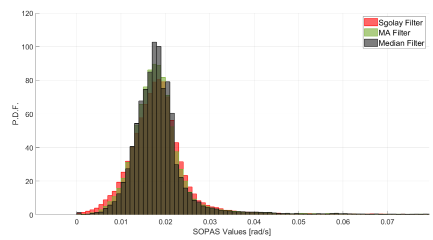


Figure 4.10: Histograms of the Three Filters

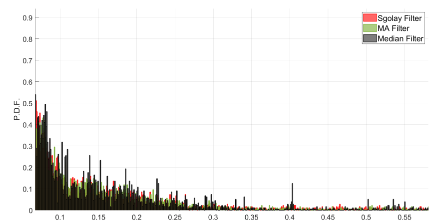


Figure 4.11: Tail of Histograms in Figure 4.10

We notice that the median filter produces a more uniform spread with a higher peak compared to the other filters. But it still retains more outliers compared to the other two filters which seem to smooth these outliers better, as seen in Figure 4.11. Though, they are significantly reduced compared to the raw SOPAS data. On the other hand, The moving average filter results in a distribution that is more compressed around the mean, with fewer extreme values. Its lower variance indicates strong smoothing, and it exhibits a higher peak near the center as fluctuations are averaged out. While it removes noise, it may over-smooth sharp transitions. The Sgolay filter has to a broader distribution at the left and right, which may indicate retaining more details than the other filters. Among the three filters, it exhibits the lowest variance while retaining sharp transitions. This suggests that the Sgolay filter is particularly effective in preserving peaks.

A similar pattern is observed in Figure 4.12, which corresponds to the histograms of the filters applied to another SOPAS data burst. In this case, the number of outliers was lower than in the first dataset. However, the median filter exhibited higher variance, while the moving average and Sgolay filters maintained similar distributions as before.

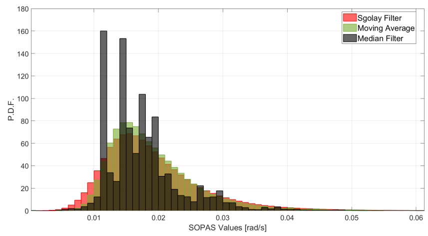


Figure 4.12: Histograms of the Three Filters for SOPAS Data in 16-10-2025 from 09:00 to 14:00

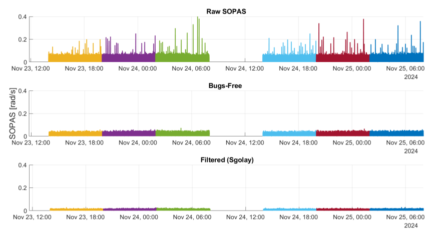


Figure 4.13: Effect of Removing Bugs and Applying Sgolar Filter on Six Circuit-2 Batches.

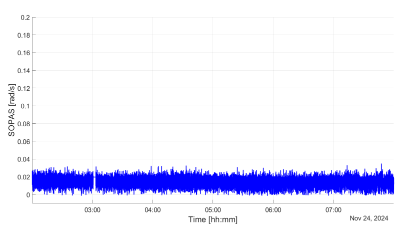
The analysis suggests that the Sgolay filter performs best among the three filters, because it achieves efficient data smoothing while preserving features like peaks and transitions. Therefore, it is the most suitable choice for SOPAS data processing when both noise reduction and event detection are required. Figure 4.13 shows the results of removing the system bugs and smoothing the data for six different data batches acquired from Circuit-2. Notably, the algorithm had effectively removed all the system bugs, and the Sgolay filter smooths the data to a great extent.

4.4 Analysis of Quiet Conditions

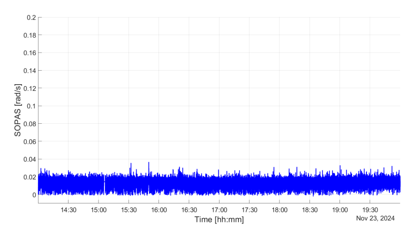
We have acquired extensive SOPAS data from different circuits and various days. These data are just background noise (considered as quiet conditions) that the fiber circuits acquire continuously. By analyzing the distribution of these data, we can compare it with the SOPAS data we collected during the experiments (considered as event conditions) and then set a threshold to determine when an alarm should be triggered.

The SOPAS data that will be presented from now on will be after removing the system bugs and applying the Sgolay filter for smoothing. Unless otherwise mentioned, the filter's window size used is 50. We will first perform a comparison between the data taken from different circuits. To ensure a fair comparison, we first analyze SOPAS data recorded from the same circuit at different times of the day and then compare them across different days.

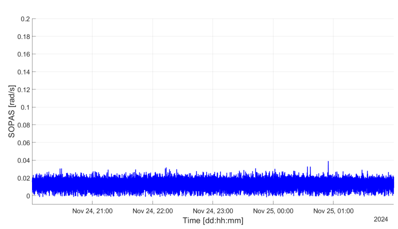
The first comparison is based on three distinct time periods: morning, evening, and night. This approach, even though these terms don't always match the exact time period with the collected data, covers the entire day, allowing us to investigate whether SOPAS variations are time-dependent. If variations exist, we can adapt the threshold accordingly. Furthermore, comparing different circuits on different days helps assess the consistency of data acquisition across them. Unfortunately, the circuits cannot be activated simultaneously. This prevents direct verification of whether an event detected in one circuit appears in another during the same time period. Additionally, Circuit-1 was not operational during the data collection period, and thus, no data is available for comparison.



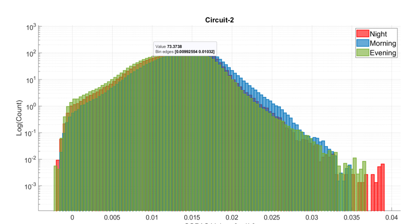
(a) Batch 1: 24-11-2024 from 02:00 to 08:00.



(b) Batch 2: 23-11-2024 from 14:00 to 20:00.



(c) Batch 3: from 24-11-2024, 20:00 to 25-11-2024, 02:00.



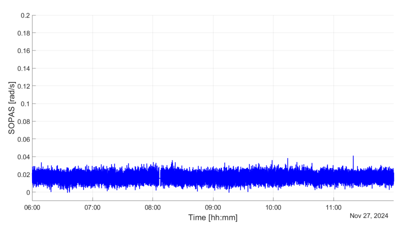
(d) Overlaid Histograms of the Three Batches

Figure 4.14: SOPAS Data & Overlaid Histograms Corresponding to Circuit-2.

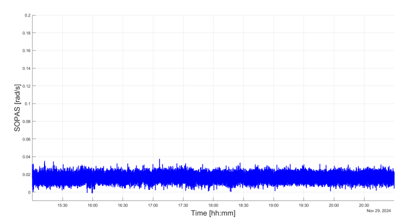
For all the data batches that will be shown in plots, each data batch consist of bursts of 60 seconds every 70 seconds of data, resulting in a 10-second gap between bursts. Some of the batches have a larger gap of missing data due to the failure of the system. However this won't affect our analysis.

4.4.1 Comparison Based on Different Times of the Day

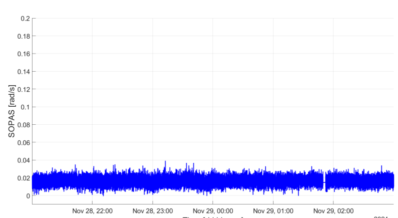
For Circuit-2, three data batches were collected, each containing six hours of recorded SOPAS data. The recorded times are specified in the titles of the corresponding SOPAS plots. Figures 4.14a, 4.14b, and 4.14c show the SOPAS value plots for the morning, evening, and night, respectively. Additionally, Figure 4.14d overlays the histograms of the three batches. Similarly, Circuit-3 and Circuit-4 (Figures 4.15 and 4.16, respectively), each have three SOPAS value plots and one histogram displaying the overlaid batches. For Circuit-4, the batches are not always 6-hours long.



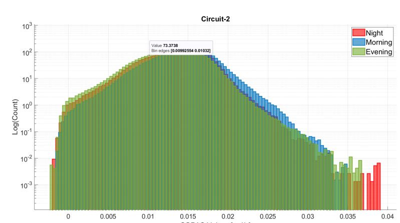
(a) Batch 1: 27-11-2024 from 06:00 to 12:00.



(b) Batch 2: 29-11-2024 from 15:00 to 21:00.



(c) Batch 3: 28-11-2024 from 21:00 to 03:00.

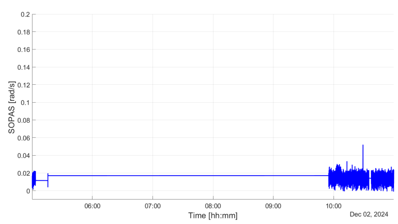


(d) Overlaid Histograms of the Three Batches

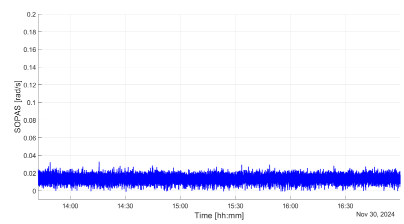
Figure 4.15: SOPAS Data & Overlaid Histograms Corresponding to Circuit-3.

All batches are considered to be in quiet conditions with no or very few values that exceeds the average values. The summarized statistical metrics for this Section are presented in Table 4.1. This table includes the mean, standard deviation, and 99th percentile of SOPAS values for each dataset. The 99th percentile is particularly useful for defining an event threshold, as it identifies extreme but still naturally occurring values in quiet conditions.

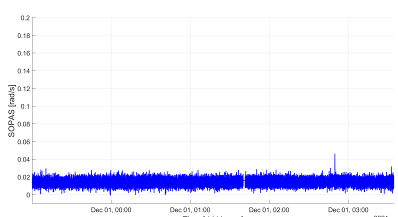
The SOPAS values range from approximately -0.01 (due to the Sgolay filter) to 0.045 rad/s, indicating that SOPAS variations are relatively small. This suggests that the system is generally stable, with minor fluctuations. The histograms show a distribution that is not perfectly symmetric, with a slight skewness to the left. This indicates that lower SOPAS values occur more frequently, which is expected since the data primarily consists of background noise rather than significant events.



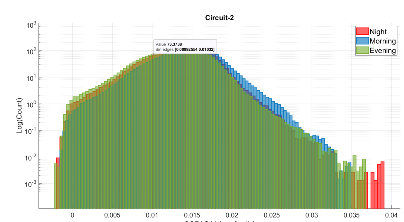
(a) Batch 1: 02-12-2024 from 05:00 to 11:00.



(b) Batch 2: 30-11-2024 from 13:42 to 17:00.



(c) Batch 3: 30-11-2024 from 23:00 to 03:34.



(d) Overlaid Histograms of the Three Batches

Figure 4.16: SOPAS Data & Overlaid Histograms Corresponding to Circuit-4.

The peak of the histograms is likely around mid-range values (0.012 to 0.018 rad/s). Compared to the mean values of each batch (see Table 4.1), the data is concentrated around the mean, indicating a stable baseline with occasional deviations. The presence of values at the higher end of the range (0.03 rad/s and above) suggests occasional anomalies, particularly for Circuit-4.

However, these cannot correspond to actual rockfall events since the values are significantly smaller than those observed in Chapter IV.

Circuit	Data Batch	Mean	Std	99th P
3*Circuit-2	23/11 14:00 - 23/11 20:00	0.013	0.004	0.020
	24/11 02:00 - 24/11 08:00	0.014	0.003	0.022
	24/11 20:00 - 25/11 02:00	0.014	0.004	0.021
3*Circuit-3	27/11 06:00 - 27/11 12:00	0.016	0.003	0.025
	28/11 21:00 - 29/11 03:00	0.016	0.003	0.024
	29/11 15:00 - 29/11 21:00	0.016	0.003	0.024
3*Circuit-4	02/12 05:00 - 02/12 11:00	0.010	0.004	0.022
	30/11 13:42 - 30/11 17:00	0.013	0.003	0.021
	30/11 23:00 - 01/12 03:34	0.014	0.003	0.021

Table 4.1: Filtered SOPAS Statistics for Different Circuits

4.4.2 Comparison Based on Different Days

Figures 4.17, 4.18, and 4.19 show the three overlaid histograms from Circuit-2, Circuit-3, and Ciruit-4 batches, respectively. Similar to the previous comparison, the batches are considered to be in quiet conditions with no or very few values that exceeds the average values. These batches were taken on different days to ensure that the values range found on the previous comparison is not much different. Indeed, we can observe that the SOPAS values range is almost the same as found when we made comparisons based on different times of the day, with a slight increase in the tail, reaching this time a value of 0.5 rad/sec. This suggests that the background noise is consistent throughout different circuits and on different days. In addition, most SOPAS values are in the range of 0.012 to 0.018 rad/s, same as previously found.

The observations highlight the importance of adjusting event detection thresholds based on the variation in SOPAS values. The threshold definition might not be very complex since the variations are quite small compared to event conditions. The statistics in Table 4.1 may be useful for defining

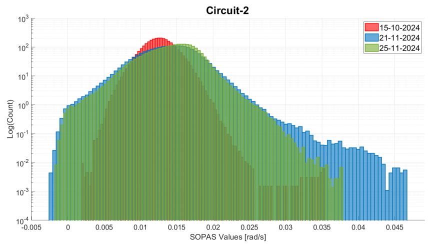


Figure 4.17: Overlaid Histograms of Circut-2 (Different Days).

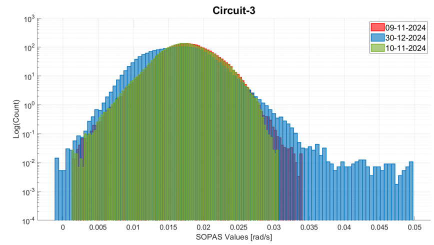


Figure 4.18: Overlaid Histograms of Circut-3 (Different Days).

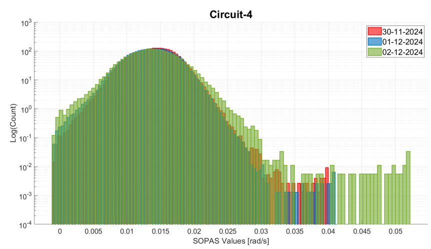


Figure 4.19: Overlaid Histograms of Circut-4 (Different Days).

a primary threshold and testing it against different data types, where not only quiet conditions but also some anomalies are present. By taking into account these insights, we can refine the detection criteria and ensure a good and reliable event identification while minimizing false alarms.

4.4.3 Comparison Based on Different days — Anomalies Present

Now that we have examined the case where the SOPAS data is in its quietest conditions, let us explore instances where some anomalies appear within these quiet conditions. To do so, we plotted the histograms of different data batches from Circuit-2, where at least one anomaly was observed. These anomalies were neither caused by actual events, as seen in the experiments, nor by system bugs affecting the data.

Figure 4.22 presents ten histogram plots for batches of five- and six-hour data intervals. The histogram in purple exhibits the longest tail among all the histograms, reaching a value of 0.24 rad/sec. Its SOPAS plot is displayed in Figure 4.20. In contrast, the yellow histogram (Figure 4.21 shows its SOPAS values) has the smallest tail, not exceeding 0.05 rad/sec. However, the peak values or all batches remain within the same range as observed in the previous comparison, between 0.01 and 0.02 rad/sec. This suggests that the overall data distribution is consistent across different days.

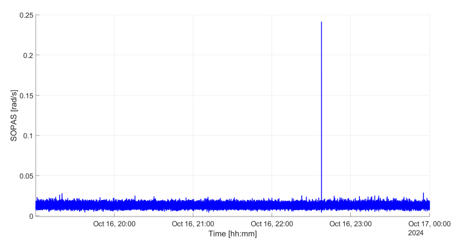


Figure 4.20: Batch 1: from 16-10-2024, 19:00 to 17-10-2024, 00:00.

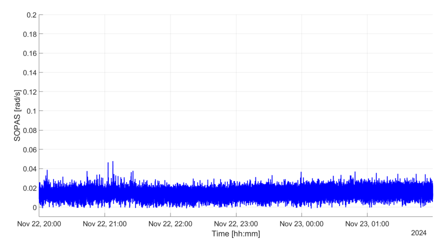


Figure 4.21: Batch 2: from 22-11-2024, 20:00 to 23-11-2024, 02:00.

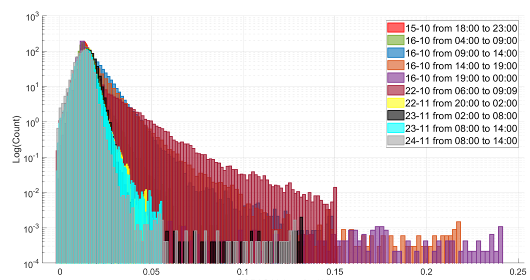


Figure 4.22: Histogram Plots of Ten Different Batches From Circuit-2.

This comparison highlights the variations in SOPAS data and reinforce the general stability of its distribution. The presence of longer tails in certain histograms suggests the presence of occasional disturbances that do not align with typical experimental events or system bugs. These anomalies might be attributed to environmental factors (perhaps animals or birds stepping over the fiber) or minor external influences.

Despite the observed differences in tail lengths, the peak values remain within a narrow range. This indicates that the baseline SOPAS conditions are stable over time. This consistency is essential for threshold-based detection, as it helps distinguish real events from normal fluctuations. Future analysis could involve correlating these anomalies with external factors, such as animal movements, to better understand their origins.

Chapter 5

SOPAS States & Threshold Determination

As we have seen earlier, categorizing events based solely on SOPAS values remains challenging. In this chapter, we aim to analyze both quiet conditions and event conditions (rockfall event simulations). By comparing these states under different scenarios, we can establish criteria to determine whether a given event corresponds to an actual rockfall event or merely an environmental disturbance that should be disregarded. To do so, we want to establish a threshold that can be efficient in detecting anomalies and events.

5.1 SOPAS-based Threshold

The threshold is one of the core parameters in this thesis. By setting a good threshold, we can control the alarm to be triggered in case of an event occurs. The chosen threshold is designed to balance sensitivity to events while minimizing false detections. The first approach to define the threshold is relying on the statistics of the SOPAS data, namely the mean, standard deviation, and percentile. Using these parameters will help to rebust against normal fluctuations and to adapt to the data's actual distribution. The primary goal of setting the threshold is to reduces false positives, that is avoiding flagging minor fluctuations as anomalies.

In this work, two types of thresholds were defined. The first type is the SOPAS threshold, denoted as $\boldsymbol{\omega_{th}}$. This threshold is applied to the $\Omega_{sm}[n]$

signal, where any significant impact on the fiber results in SOPAS values exceeding this threshold. This exceedance allows us to classify the event as either an anomaly or a valid event.

Initially, ω_{th} is defined as:

$$\omega_{th}^{\text{initial}} = \text{mean}(\Omega_{\text{sm}}[n]) + 2 \times \text{std}(\Omega_{\text{sm}}[n]) \tag{5.1}$$

To refine this threshold further, we incorporate the 99th percentile (P_{99}), leading to the final definition:

$$\omega_{th} = \max\left(2 \times \omega_{th}^{\text{initial}}, 2 \times P_{99}\right) \tag{5.2}$$

These statistics can tell us a lot about our data. The **mean** represents the central tendency of the SOPAS values. It provides a measure of the typical SOPAS intensity that is observed in the data. The **standard deviation** (std) quantifies the spread or variability of SOPAS values around the mean: a higher std indicates larger fluctuations in SOPAS values, while a lower one suggests more stable and consistent conditions. **The percentile** represents the threshold below which percent of the SOPAS data points lie. It helps detect extreme anomalies by highlighting values that are significantly higher than normal variations. If SOPAS values exceed this threshold, it may indicate an unusual event, such as a rockfall. The 99th percentile of the smoothed signal, $P_{99}(\Omega_{sm}[n])$, is useful for setting an alert threshold, which allows for detecting significant changes while minimizing false alarms.

5.2 Relationship Between ω_{th} and W

To analyze the effect of applying the threshold, we applied it to some data batches and compared them over a specific time period on a selected date, in the afternoon from 14:00 to 17:00. In addition, we applied different filter's window sizes ($W = 5, 50, 100$ samples) to observe the relationship between smoothing and thresholding. To ensure a reliable comparison, we selected three data batches from the three circuits and clipped them to be 3-hours each (from 14:00 to 17:00).

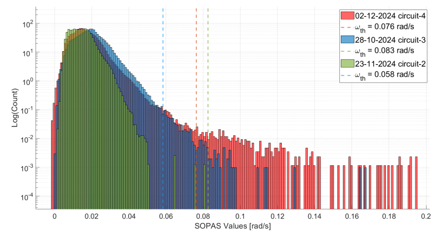


Figure 5.1: Histograms of Circuit-2, -3, and -4 ($W = 5$ samples).

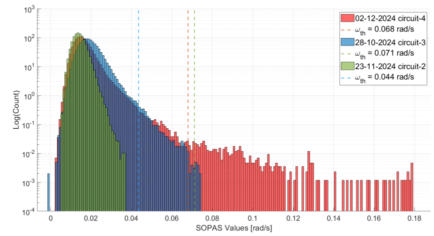


Figure 5.2: Histograms of Circuit-2, -3, and -4 ($W = 50$ samples).

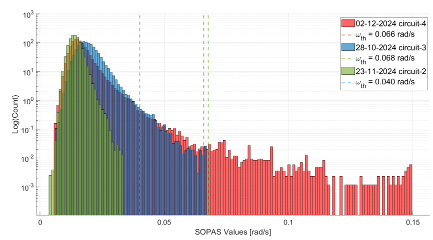


Figure 5.3: Histograms of Circuit-2, -3, and -4 ($W = 100$ samples).

Observing the plots, we note that the SOPAS data From Circuit-4 (histogram in red color) exhibits more noise than the other circuits (Circuit-2 in green and Circuit-3 in blue), where its tail extends for a longer period, suggesting the presence of more anomalies. A similar pattern is observed Circuit-3, though with lower intensity. To compare the distributions, we overlaid the histograms of the three batches. The results are displayed in Figure 5.1 for the circuits' batches with $W = 5$ samples, followed by Figure 5.2 with $W = 50$ samples, and finally Figure 5.3 with $W = 100$ samples, all for the same data batches.

For $W = 5$ samples, the tail of the SOPAS data from Circuit-4 is the largest, reaching a value of 0.2 rad/sec. In contrast, Circuit-3's tail reaches 0.1 rad/sec, which is half the value observed in Circuit-4. Meanwhile, Circuit-2, which does not exhibit any apparent anomalies, has a small tail with only a few outliers. The calculated thresholds, however, did not follow the expected trend. The highest threshold was found in the Circuit-3 batch ($Th = 0.083$ rad/sec), despite Circuit-4 ($Th = 0.076$ rad/sec) exhibiting more outliers, whereas Circuit-2 had the lowest threshold ($Th = 0.058$ rad/sec). The same pattern persisted when increasing the filter window size ($W = 50$), but the threshold values for all batches decreased (0.071 rad/sec for Circuit-3, 0.068 rad/sec for Circuit-4, and 0.044 rad/sec for Circuit-2). Increasing the window size further ($W = 100$) led to an even greater reduction in thresholds. As expected, the number of outliers also decreased as the smoothing window increased.

Discussion

From these results, we can draw several conclusions. Firstly, the filter window size plays a crucial role in determining the threshold values. A smaller window preserves short-term variations, making anomalies more pronounced, but it also leads to higher thresholds. On the other hand, larger window sizes suppress noise and smooth out fluctuations, reducing the threshold values but potentially masking significant anomalies.

The unexpected behavior in threshold values (where Circuit-3 had the highest threshold despite Circuit-4 showing more outliers) suggests that the thresholding method is sensitive not only to extreme values but also to the distribution of SOPAS data. This indicates that while thresholding is an effective way to detect anomalies, its accuracy depends on both the filtering process and the statistical properties of the data.

Therefore, choosing an appropriate window size is critical. A very small window may overestimate the threshold, leading to unnecessary alarms, while an excessively large window may underestimate the threshold, potentially missing real anomalies. A balanced approach is necessary, possibly by adapting the window size dynamically based on the characteristics of the data.

5.3 Samples Above Threshold (SAT)

Now that we have defined a threshold, we can further analyze the SOPAS data by identifying samples exceeding the threshold. This is the goal: we want to examine the type of SOPAS values that surpass ω_{th} and categorize them by distinguishing between actual rockfall events and those caused by environmental disturbances, such as animals stepping on the fiber. To achieve this, a code was written to count and plot samples that exceed the set threshold for each data batch. What we aim to observe is how the number of samples above the threshold (SAT) changes with different thresholds and filter window sizes. For this purpose, we are currently showing only two batches, as this should be sufficient to illustrate the point. Table 5.1 summarizes the results.

Batch	**W (sam)**	**Mean**	**Std**	P_{99}	ω_{th}	**SAT**
28-10-24	5	0.019	0.006	0.036	0.073	362
	20	0.019	0.005	0.036	0.072	300
	50	0.019	0.005	0.035	0.070	268
	100	0.019	0.005	0.034	0.067	172
02-12-24	5	0.016	0.006	0.040	0.070	956
	20	0.016	0.006	0.036	0.069	884
	50	0.016	0.006	0.034	0.067	826
	100	0.016	0.005	0.033	0.066	735

Table 5.1: Statistical Parameters and Threshold/SAT Values for Different Batches and Window Sizes

Before starting the analysis, the expected results are: increasing the ω_{th} will decrease the number of SAT, and vice versa. Additionally, using a larger W should also decrease the SAT, as applying a smoothing filter removes some fluctuations, reducing the number of samples exceeding the threshold. It is also important to remember that ω_{th} depends on statistical parameters (mean, std, and P_{99}), which themselves change when adjusting W, since the data being analyzed consists of smoothed SOPAS values.

In the first batch (28-10-24), the number of SAT decreases as and W increases and ω_{th} decreases. Similar trend us observed in the second batch (02-12-24), where SAT decreasing by increasing W. These results suggest that the window size has more impact, on these particular SOPAS data, than the SOPAS threshold ω_{th}, because the number of SAT always decreases even though ω_{th} decreases. Nevertheless, there is a trade-off between W and ω_{th}, which should be taken into account in the analysis.

5.4 Events Above Threshold (EAT)

Taking the results from the previous Section into account, we employed an alternative approach for analyzing by counting for anomalies and events consisting of samples that surpass the threshold ω_{th}. We refer to those anomalies or events by **exceedances**. The start of an exceedance is marked by the first sample that exceeds ω_{th}, while the end is determined by the last sample that falls below ω_{th}, as shown Figure 5.4. The circular markers correspond to the SOPAS values that comprise these exceedances.

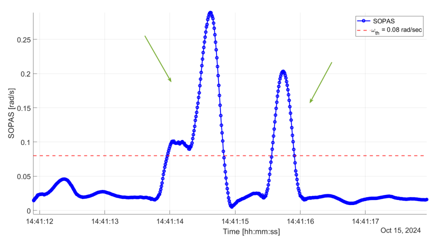


Figure 5.4: Definition of an Exceedance (From 03B Tests).

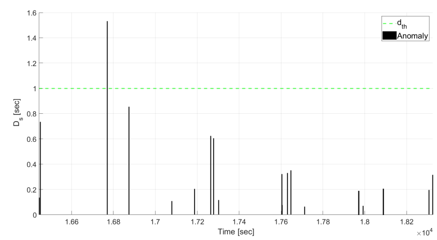


Figure 5.5: EAT of Batch present in Figure **??**.

We further defined the second type of threshold, a time-based threshold, denoted as $\boldsymbol{d_{th}}$. What we denote EAT are the exceedances that exceed ω_{th} for a specific duration, D_s. If this duration exceeds a time-threshold, d_{th}, then the exceedance will be classified as an event. This additional criterion helps distinguish between different events and anomalies. For example, certain events, such as debris flows, tend to last for a longer period of time, while environmental disturbances may only last for a few seconds. Figure 5.5 shows an example for detecting EAT, taken from the data batch in Circuit-4 where only one EAT is found. As we observed in Chapter 3, the rockfall event simulations typically last for a few seconds (on average, around 5 seconds). Although real events have completely different dynamics and duration (tens of seconds to even minutes), considering the generated rockfall as events to be detected, can be a useful guide to set the thresholds.

Let us first analyze the experiments presented in Chapter 3. The experiments consist of a total of 5 tranches, each containing a number of exceedances. Table 5.2 illustrates the results. Each tranche is tested against three parameters: threshold, SAT, and EAT. Also, for each parameter, 5 different filter's window sizes are computed. The extreme values of the window sizes were used only to examine the relationship between the parameters, but they are not employed in practices. The last column represents the EAT that exceed both thresholds for $D_s = 1$ **second**.

Tranche	**ω_{th} (rad/sec)**	**SAT**	**EAT**
	W= 5 — 50 — 100 — 200 — 500	W = 5 — 50 — 100 — 200 — 500	W = 5 — 50 — 100 — 200 — 500
04A (01 - 05)	0.126 — 0.126 — 0.129 — 0.134 — 0.145	855 — 873 — 866 — 739 — 444	1 — 2 — 2 — 3 — 2
03B (06 - 13)	0.082 — 0.080 — 0.079 — 0.077 — 0.074	164 — 203 — 234 — 132 — 184	0 — 0 — 0 — 1 — 1
02C (14 - 21)	0.297 — 0.310 — 0.303 — 0.309 — 0.324	653 — 504 — 398 — 381 — 531	0 — 1 — 1 — 2 — 2
01D (22 - 27)	0.228 — 0.217 — 0.201 — 0.181 — 0.164	286 — 319 — 327 — 426 — 408	0 — 0 — 0 — 2 — 1
04A (28 - 31)	0.359 — 0.355 — 0.359 — 0.344 — 0.392	738 — 737 — 704 — 868 — 844	3 — 3 — 3 — 5 — 3

Table 5.2: SAT and EAT statistics for different test cases

The maximum number of EAT found was 5, corresponding to tranche 04A (28 to 31 tests) using window W = 200 samples, while the minimum number occurred three times in tranche 03B and 01D and one time in 02C. On the other hand, the highest number of SAT was observed in tranche 04A (1 to 5 tests) using $W = 50$ samples, whereas the lowest count was found in tranche 03B using W = 200 samples.

We observe a clear pattern in the number of EAT across all tranches. The number of EAT increases up to the fourth window size row (W = 200 samples), after which it starts to decrease. This demonstrates the impact of using a larger W. Regarding the threshold ω_{th}, there is no obvious pattern of increasing or decreasing. Instead, the results appear to be more random. This randomness comes from the way the threshold is defined. As the W increases, it influences the statistical parameters that define ω_{th}. With larger W, there may be instances where $\omega_{th}^{\text{initial}}$ (see Eq 5.1) is smaller than the P_{99}, causing the final threshold to be set at twice the 99th percentile instead.

Notably, there is no clear relationship between SAT and EAT, because the events differ significantly in how they are defined. Some exceedances exhibit an inverse V-shape, where their SOPAS values increase and then decrease once (second curve in Figure 5.4), while others have a random shape (first curve in Figure 5.4), where samples count increase and decrease randomly. This variability in shape leads to different numbers of SAT for different values of threshold, and explains why the number of EAT is random.

Discussion

The analysis of SAT and EAT, as well as the two threshold types, is valuable in distinguishing between exceedances. The ability to define EAT based on time allows to have more refined detection of rockfall events, which typically last longer than environmental disturbances. But while this approach aids to narrow down potential events, some uncertainty remains in the precise duration of real-world events. Further research and more extensive data are needed to refine the threshold for different event types, especially as the timescale of environmental disturbances may vary.

The relationship between ω_{th}, SAT, and EAT is complex, as demonstrated by the observed variability across the tranches. Although larger W impacts the number of EAT detected, their effect on the threshold definition introduces some randomness. This suggests that more sophisticated methods may be needed to account for the nature of these parameters, especially when applying this analysis to real-world data. Ultimately, this approach can serve as a useful tool for events validation.

5.5 Complementary Cumulative Density Function (CCDF) Analysis

To further analyze the distribution of the SOPAS data, particularly its tail behavior, we employed the Complementary Cumulative Density Function (CCDF). It is defined as the probability that a random variable X exceeds a given value x:

$$CCDF(x) = P(X > x) = 1 - CDF(x) \tag{5.3}$$

which is the inverse of the Cumulative Distribution Function ($CDF(x)$), representing the probability that X is less than or equal to x. Figure 5.6 shows the CCDF for Circuit-3. The filter's window for this batch and all others in this Section is set to $W = 50$ samples.

Before proceeding, we need to address an issue that arose while plotting the CCDF. Notice the flatness in the tail of the CCDF (indicated by black arrows), which appears as stair-like steps. Upon analyzing the histogram of this batch and checking its SOPAS values, we found duplicate values at the tail, as shown in Figure 5.7. These duplicates occur occasionally due to system failures, as we have seen in Chapter 3. This frequent duplication could bias the data distribution, since they exceed the threshold and thus increasing the number of SAT. Though, this might not be a big issue in our dataset since it happens rarely compared to the overall amount of data.

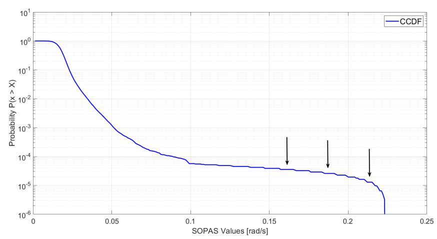


Figure 5.6: CCDf of Circuit-3 Batch (09-1-24).

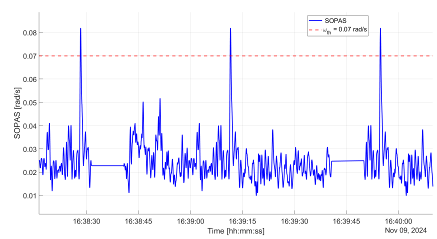


Figure 5.7: Duplicate Exceedance From Batch in Figure 5.6.

Looking at tranche 03B from the rockfall experiments, which represent the event conditions (Figure 5.8), we observe that only 10% of the values exceed 0.026 rad/sec, where a steep drop begins. Additionally, applying the threshold to this set of tests reveals that only 0.1% of the values exceed the threshold ω_{th}, which in this case is 0.080 rad/sec. A similar pattern is observed across the other tranches. We overlaid all tranches on the same plot (Figure 5.9), along with their respective thresholds. In each case, only 10% of the values exceed the range of 0.025 to 0.03 rad/sec, except for tranche 01D, where the corresponding value is 0.051 rad/sec. The thresholds crossings occur between 0.1% and 0.4%, indicating the presence of only a few extreme values, which is the case as the tests periods are not long compared to the total acquired data.

To further investigate, we analyzed SOPAS data during quiet conditions by dividing the batches into two categories:

- Batches with no exceedances (BT - Below Threshold).
- Batches with at least one exceedance (AT - Above Threshold).

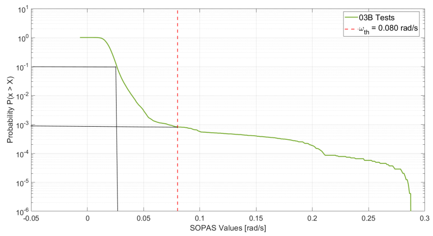


Figure 5.8: CCDF of Tranche 03B.

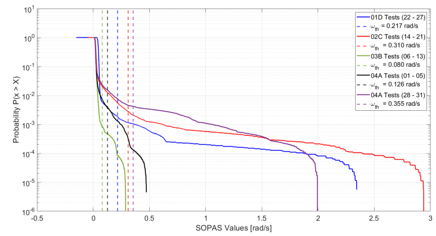


Figure 5.9: CCDFs of All Tranches.

This classification allows us to examine the distribution in the presence and absence of anomalies and how the tail behavior differs in each case. Additionally, we compare these distributions with the event conditions previously analyzed.

For Circuit-2, we identified eight batches where all SOPAS values remained below the threshold (BT), shown in Figure 5.10, and another eight batches where at least one exceedance exists (AT), presented in Figure 5.11. In the BT batches, only 10% of SOPAS values exceed the values ranging between 0.015 and 0.018 rad/sec, depends on each batch. Additionally, the tail end consists mainly of values ranging between 0.03 and 0.04 rad/sec. Since no values exceed ω_{th} in these batches, there are no threshold crossings with probability values. Conversely, in the AT batches, only 10% of the SOPAS values exceeded values that range between 0.015 and 0.030 rad/sec, except for one batch (blue curve), where this value reaches approximately 0.045 rad/sec. This batch exhibits the longest tail among all, with a maximum value of 0.745 rad/sec. The threshold crossings with probability values range between 0.01% and 0.001%.

For Circuit-3, we analyzed twelve BT batches (Figure 5.12) and six AT batches (Figure 5.13). Here, 10% of the SOPAS values exceed 0.019 to 0.022 rad/sec in BT batches, whereas this range increases slightly to 0.020 to 0.024 rad/sec for AT batches, showing only a minor difference. The threshold crossings with probability values are similar to those in Circuit-2, occurring between 0.01 and 0.001 %.

Circuit-4, with four BT batches and three AT batches, follows a similar trend. However, for this circuit it is even less: only 1% of SOPAS values

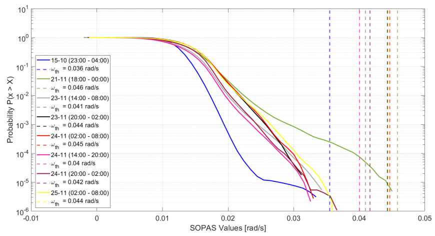


Figure 5.10: CCDFs of BT Circuit-2 Batches.

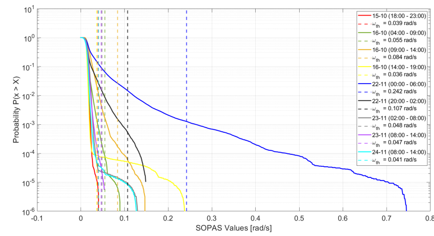


Figure 5.11: CCDFs of AT Circuit-2 Batches.

exceed the 0.021 to 0.029 rad/sec range in the BT batches, whereas in the AT batches, this range is slightly extended. Figure 5.14 and Figure 5.15 illustrate the CCDF plots of the BT and AT batches, respectively.

In addition, we computed the average ω_{th} values for each case (see Table 5.3). In the table, C denotes Circuit, while RFE represents the average ω_{th} values for the rockfall experiments. The BT case for Circuit-2 and Circuit-4 have the same average values, while Circuit-3 has a slightly higher value. For the AT case, three different average values are observed, with Circuit-2 showing the highest value. The rockfall experiments, on the other hand, exhibit the highest average threshold among all cases.

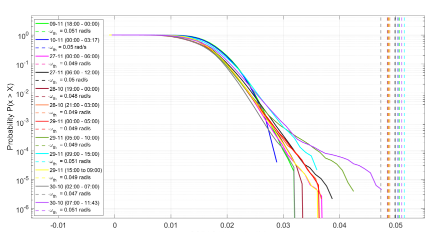


Figure 5.12: CCDFs of BT Circuit-3 Batches.

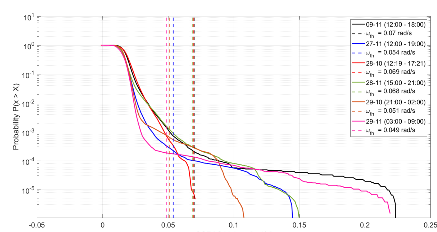


Figure 5.13: CCDFs of AT Circuit-3 Batches.

Circuits(AT/BT)	**BT_C2**	**BT_C3**	**BT_C4**	**AT_C2**	**AT_C3**	**AT_C4**	**RFE**
mean(ω_{th}) [rad/sec]	0.042	0.049	0.042	0.078	0.060	0.049	0.218

Table 5.3: Average Threshold Values.

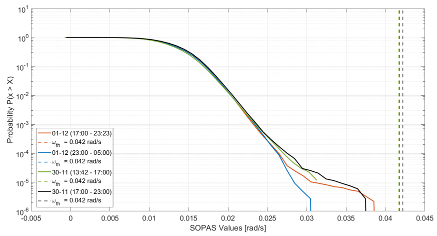


Figure 5.14: CCDFs of BT Circuit-4 Batches.

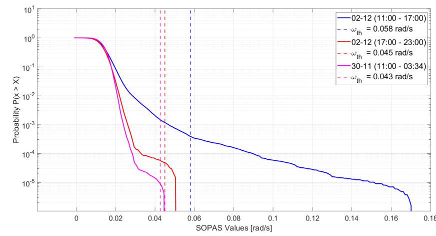


Figure 5.15: CCDFs of AT Circuit-4 Batches.

Discussion

The CCDF analysis provides valuable insights into the tail behavior of SOPAS data under different conditions. When comparing quiet conditions (BT batches) and event conditions, we confirmed that extreme values are rare, which indicates that major disturbances occur infrequently relative to the overall dataset. One key observation is that the distribution of SOPAS values remains relatively stable across different batches and circuits. The 10% cutoff values for BT and AT batches are close to each other, particularly in Circuit-3 and Circuit-4, where the difference is minimal.

Although the probability of threshold crossings remains low, even small percentage values can translate into actual samples, given that each batch spans several hours of data at a 100 Hz sampling rate. This implies that, over time, false alarms will inevitably occur every few hours, which is expected.

Evidently, ω_{th} remains a useful parameter for distinguishing between event and quiet conditions. Furthermore, the similarity in the average value of ω_{th} between Circuits in the BT case suggests comparable background noise levels, while in the AT case the higher ω_{th} in Circuit-2 suggests existence of anomalies but with low SOPAS intensity compared the rockfall experiments where the average ω_{th} value is three times (0.218 rad/sec) the highest average value in this case (0.078 rad/sec). This again reinforces the importance of ω_{th} selection in differentiating between environmental disturbances and actual rockfall events.

5.6 Statistical Metrics

In this section, we present statistical analyses applied to SOPAS data under both quiet and event conditions. These statistics provide insights into setting key parameters such as the SOPAS threshold ω_{th} and the window size W.

Table 5.4 summarizes the SOPAS data used in this study, particularly the data acquired under quiet conditions. It includes the number of data batches for each circuit, the total batch duration, and the SOPAS values duration (since data batches consist of bursts depending on the acquisition rate). Additionally, the total number of samples recorded for each circuit is reported.

Circuit	**# Batches**	**$\sum$ Batches Duration**	**$\sum$ SOPAS Duration**	**$\sum$ # Samples**
Circuit-2	19	105 h 09 m	90 h 08 m	32,448,000
Circuit-3	18	98 h 00 m	83 h 58 m	30,228,000
Circuit-4	07	37 h 51 m	32 h 26 m	13,626,000

Table 5.4: Summary of Batches and Data durations for each Circuit.

Table 5.5 provides a summary of the rockfall experiments. It includes the statistical metrics used to define the SOPAS threshold ω_{th}, the computed ω_{th} values, and the SAT and EAT (lasting more than $D_s = 1$ second) for each tranche. Moreover, these parameters are evaluated for different filter window sizes, specifically $W = [50, 100, 200]$ samples.

2*Metric	**01D (Tests: 22-27)**			**02C (Tests: 14-21)**			**03B (Tests: 6-13)**			**04A (Tests: 1-5)**			**04A (Tests: 28-31)**		
	W=50	W=100	W=200	W=50	W=100	W=200	W=50	W=100	W=200	W=50	W=100	W=200	W=50	W=100	W=200
Mean	0.046	0.046	0.046	0.027	0.027	0.027	0.020	0.020	0.020	0.020	0.020	0.020	0.025	0.025	0.025
Std	0.031	0.027	0.022	0.054	0.049	0.044	0.007	0.006	0.005	0.015	0.014	0.013	0.063	0.062	0.060
P_{99}	0.062	0.060	0.068	0.155	0.151	0.154	0.040	0.039	0.039	0.063	0.065	0.067	0.177	0.180	0.172
ω_{th}	0.217	0.201	0.181	0.310	0.303	0.309	0.080	0.079	0.077	0.126	0.129	0.134	0.355	0.359	0.344
SAT	319	327	426	504	398	381	203	234	132	873	866	739	737	704	868
EAT	0	0	2	1	1	2	0	0	1	2	2	3	3	3	5

Table 5.5: Summary of Statistical Metrics for Rockfall Experiments.

To gain further statistical insights, we compute the minimum and maximum values of the mean, standard deviation, and 99th percentile of the smoothed SOPAS data. Similarly, the minimum and maximum values of ω_{th} are calculated, all for different window sizes W (50, 100, and 200 samples). These results are summarized in Table 5.6. The table includes also the number of SAT and EAT for each W.

Metric	W = 50 samples	W = 100 samples	W = 200 samples
min(Mean)	0.020	0.020	0.020
max(Mean)	0.046	0.046	0.046
min(Std)	0.007	0.006	0.005
max(Std)	0.063	0.062	0.060
min(P_{99})	0.040	0.039	0.039
max(P_{99})	0.177	0.180	0.172
min(ω_{th})	0.080	0.079	0.077
max(ω_{th})	0.355	0.359	0.344
SAT	2636	2529	2546
EAT	6	6	13

Table 5.6: Summary of Statistical Metrics for All Tranches

A similar analysis is conducted for quiet conditions. For each circuit, the same window sizes ($W = 50, 100, 200$ samples) are used to ensure a fair comparison with the event conditions. The same statistical metrics are computed, and the number of SAT and EAT occurrences is counted and reported for each W value. These statistics are summarized in Table 5.7.

2*Metric	Circuit-2			Circuit-3			Circuit-4		
	W=50	W=100	W=200	W=50	W=100	W=200	W=50	W=100	W=200
min(Mean)	0.013	0.013	0.013	0.015	0.015	0.015	0.013	0.013	0.013
max(Mean)	0.025	0.025	0.025	0.019	0.019	0.019	0.015	0.015	0.015
min(Std)	0.002	0.002	0.001	0.003	0.003	0.002	0.003	0.002	0.002
max(Std)	0.024	0.023	0.023	0.005	0.005	0.005	0.005	0.004	0.004
min(P_{99})	0.018	0.017	0.016	0.024	0.023	0.021	0.021	0.018	0.018
max(P_{99})	0.121	0.120	0.119	0.035	0.034	0.033	0.029	0.027	0.027
min(ω_{th})	0.036	0.033	0.031	0.047	0.044	0.043	0.042	0.036	0.036
max(ω_{th})	0.242	0.240	0.238	0.070	0.068	0.067	0.058	0.055	0.055
SAT	4139	4152	3943	2960	3348	3771	790	746	746
EAT	4	14	12	5	10	10	1	3	3

Table 5.7: Summary of Statistical Metrics for Circuits.

Discussion

The dataset consists of approximately 200 hours of acquisition. This provides a solid groundwork for analysis and increases the reliability of the results. The extended duration ensures that the findings are representative of long-term trends rather than short-term fluctuations. Additionally, the use of data from multiple circuits enhances the analysis by allowing a broader assessment of the system's performance across different locations.

In contrast, the data acquired from the rockfall experiments covered much shorter periods, with each tranche lasting a maximum of 15 minutes. Despite the shorter duration, the number of SAT occurrences in both states—quiet conditions and event conditions—remains comparable. A similar trend is observed for the number of EAT occurrences. This suggests that actual rockfall events generate a density of samples comparable to what is accumulated over days of acquisition under quiet conditions, which supports the ability of the system to detect significant events.

Regarding the statistical metrics, the minimum and maximum values illustrate the clear distinction between the two states. For example, in Circuit-2 with a window size of $W = 50$ samples, the values obtained during the rockfall experiments are nearly twice as large as those recorded under quiet conditions. This significant increase validates the effectiveness of the rockfall experiments and provides a strong reference for detecting real-world rockfall events.

Furthermore, the EAT values increase across both conditions as the window size increases, except for Circuit-2. This trend suggests that the threshold ω_{th} plays a more critical role in event detection than the choice of window size. However, there remains a trade-off between these two parameters that should be carefully considered to optimize detection performance.

5.7 Exceedances vs. Thresholds

This section presents the final results of this thesis. After analyzing the statistical characteristics of each SOPAS state and identifying the key parameters, we now focus on determining the most effective technique for detecting the presence or absence of events. The objective is to make sure that the detection system does not trigger an alarm for every minor fluctuation above the threshold. Instead, it should only activate for significant events, minimizing false alarms.

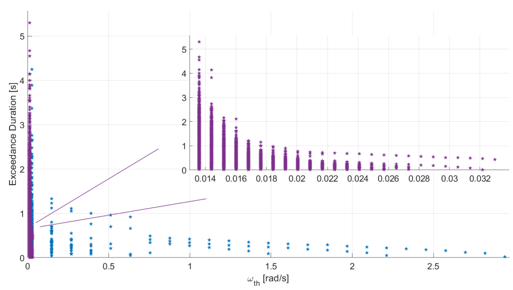


Figure 5.16: ω_{th} vs. Exceedance Duration for Quiet (Purple) and Event (blue) Conditions; Circuit-2.

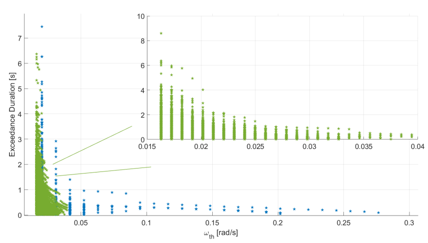


Figure 5.17: ω_{th} vs. Exceedance Duration for Quiet (Purple) and Event (blue) Conditions; Circuit-3.

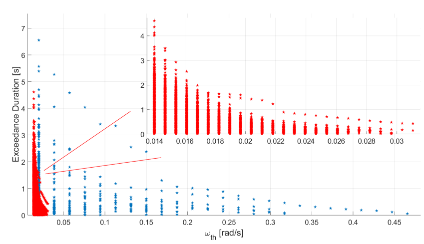


Figure 5.18: ω_{th} vs. Exceedance Duration for Quiet (Purple) and Event (blue) Conditions; Circuit-4.

To achieve this, we employ both thresholds discussed earlier: the SOPAS-based threshold ω_{th} and the time-based threshold d_{th}. The combination of these two criteria ensures a more reliable event detection mechanism while reducing false alarms. To implement this approach, we developed a script to identify exceedances for both quiet and event conditions. Instead of selecting a single fixed threshold, we define a vector of threshold values, ranging from the chosen value of the SOPAS data to its maximum value. This allows us to analyze how different threshold levels affect exceedance duration and count.

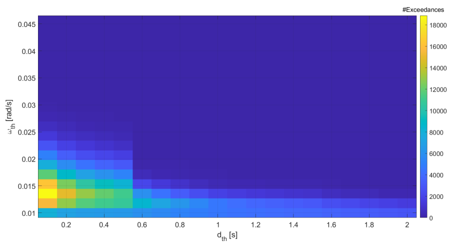


Figure 5.19: Map of Exceedances (Circuit-2 Quite Conditions).

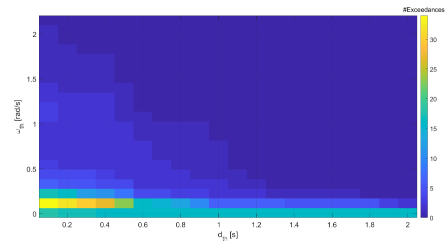


Figure 5.20: Map of Exceedances (Circuit-2 Event Conditions).

Next, we compute the duration of each exceedance and visualize the results. To facilitate comparison between SOPAS states, Figure 5.16, 5.17, and e 5.18 present an overlay of the "exceedance vs. threshold" plots for SOPAS data of Circuit-2, Circuit-3, and Circuit-4, respectively. Each plot shows both the quiet conditions and event conditions, for the purpose of comparison. The blue curve in each plot corresponds to the event conditions, while the other zoomed-in colored curves are for the quiet conditions for each circuit.

Furthermore, to understand how exceedance distributions differ between quiet and event conditions, we generated plots for each case by counting the number of exceedances. Figures 5.19 and 5.20 are for Circuit-2, Figures 5.21 and 5.22 for Circuit-3, while Figures 5.23 and 5.24 are for Circuit-4.

For the first set of figures, each SOPAS state contains 25 different threshold points, spanning a range of values. A key observation is that the highest threshold in the quiet conditions does not even surpass the second threshold value in the event conditions. This suggests that an appropriately chosen threshold can effectively differentiate between normal fluctuations and actual rockfall events. The maximum exceedance duration is observed to be around 7 to 8 seconds (we have observed higher values on some quiet conditions batches). However, this occurs only when setting the threshold to its minimum value, which is impractical for reliable detection. As the threshold

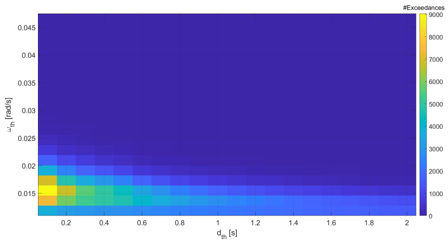


Figure 5.21: Map of Exceedances (Circuit-3 Quite Conditions).

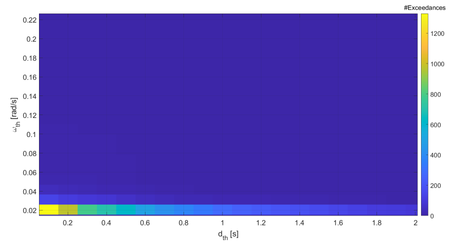


Figure 5.22: Map of Exceedances (Circuit-3 Quite Conditions).

increases, the number of exceedances decreases, which aligns with expectations.

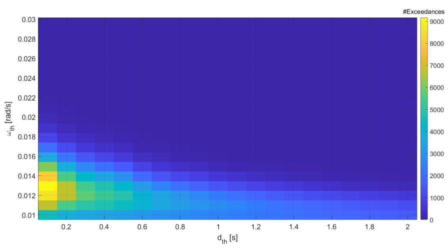


Figure 5.23: Map of Exceedances (Circuit-4 Quite Conditions).

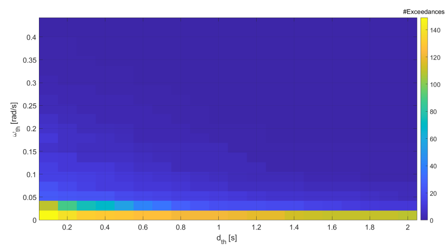


Figure 5.24: Map of Exceedances (Circuit-3 Event Conditions).

Another observation is that there is a significant contrast between the two SOPAS states is in the number of exceedances: the quiet conditions exhibit a much higher number of exceedances compared to the event conditions. This discrepancy arises because the quiet condition dataset spans six hours, whereas the event dataset lasts only around 15 minutes. The difference in acquisition time naturally results in a higher number of exceedances in the quiet conditions, but these exceedances are typically of shorter duration. Since the total count of exceedances is influenced by the duration of the dataset, it cannot be used as a reliable comparison metric. Instead, a more robust approach is to compare exceedance durations. This is precisely why the time-based threshold was introduced. For instance, if we ignore exceedances for very low threshold values, and shorter than 1 second, the remaining long-duration exceedances will be more indicative of actual events.

This filtering process significantly improves event detection by making sure that only meaningful and sustained SOPAS variations are considered.

The combination of SOPAS-based and time-based thresholds provides a balanced trade-off between sensitivity and reliability, which allows for accurate event detection with minimal false alarms. In addition, the maps provide an intuitive way to assess the impact of different threshold selections. By choosing a specific pair of ω_{th} and d_{th} on the map, one can immediately determine the expected number of false alarms. This enables the identification of an optimal threshold pair where no exceedances meet both conditions, effectively eliminating false alarms. It also offers a clear estimate of the number of false alarms within the observed time window, which aids in fine-tuning the detection parameters. The results confirm that defining an optimal threshold range, along with duration-based filtering, is important for distinguishing actual rockfall events from background noise in the SOPAS data.

5.8 Disucssion

We introduced a novel optical fiber sensing system based on State of Polarization Angular Speed (SOPAS) tracking for monitoring and early warning of hazardous mass movements in mountainous regions. The goal is to mitigate the increasing risks associated with growing human activity and infrastructure in these areas. Our system has demonstrated several advantages, including high sensitivity, real-time detection, and potential scalability.

Compared to real rockfall events, our controlled experiments generated weaker SOPAS intensities. However, the system successfully detected even these weaker disturbances, including tests where only small rocks were thrown onto the fiber. This is crucial as it confirms the high sensitivity of the system, ensuring that even minor disturbances, that are potential precursors to larger events, can be identified. Moreover, a detailed analysis of quiet conditions (background noise) versus event conditions showed a clear distinction in SOPAS values. Under quiet conditions, SOPAS remains consistently low, significantly below the values observed during an event. This differentiation is vital for refining the SOPAS-based detection threshold ω_{th}, minimizing false alarms while maintaining reliable event identification.

Our study highlighted W and ω_{th} as critical parameters for event detection. These must be carefully optimized to balance sensitivity and robustness. The results confirmed that the chosen averaging windows ($W = 50, 100, 200$ samples) are effective across a wide range of both weak and intense events. However, the selection of ω_{th} is the most critical factor, as this SOPAS threshold determines when an actual anomalous event should trigger an alarm. Its optimal setting depends on several factors, including the intensity of the anomalous events, the position of the fiber circuit on the gully, and the quiet conditions state that characterizes the background noise.

The Complementary Cumulative Distribution Function (CCDF) analysis revealed that false alarms are expected every few hours, likely due to external factors such as animals or human activity in the gully. Monitoring all four fiber circuits simultaneously could have helped mitigate this issue, but due to system limitations, only one circuit could be analyzed at a time. Future work should explore multi-circuit monitoring to reduce false positives.

In addition, while our focus was rockfall detection, landslides and debris flows are also important hazards that need further investigation. These processes may produce SOPAS values closer to background noise rather than the sharp spikes associated with rockfalls. This may make it challenging to distinguish between these subtle changes and regular noise, which must be addressed by future studies. Furthermore, and while the SOPAS values alone provide detection, they may not be sufficient for event classification. Machine learning techniques could help differentiate between different event types; such as landslides, debris flows, and other anomaly types. Future research should explore classification models trained on labeled datasets to improve event characterization and reduce uncertainty.

Chapter 6

Conclusion

This thesis presents a comprehensive study on the analysis of State of Polarization Angular Speed (SOPAS) data to detect rockfall events. The methodology integrates an optical fiber-based sensing system with an event detection algorithm to distinguish between actual rockfall events and anomalies.

The work began by detailing the experimental setup and the event detection algorithm, outlining the tools and techniques used for data acquisition and analysis. The rockfall experiments were then described, demonstrating how controlled events were conducted to simulate real-world conditions. These experiments provided valuable insights into SOPAS variations during rockfall events. The study also examined the quiet conditions, where SOPAS data was acquired and analyzed in the absence of rockfall events. By comparing the SOPAS states under event and quiet conditions, key statistical parameters were extracted to define detection thresholds. Two critical parameters, the SOPAS threshold ω_{th} and the time-based threshold d_{th}, were introduced and evaluated.

In addition, the issue of system bugs was addressed by identifying the underlying problem and presenting an algorithm designed to eliminate these bugs while preserving the overall SOPAS variations. Moreover, two filtering techniques, the moving average and the Savitzky-Golay (Sgolay) filters, were analyzed in detail, outlining their respective advantages and limitations. Then, a detailed analysis of different filtering window sizes and statistical metrics, which allowed for a robust threshold determination, was made. The results indicated that actual rockfall events generate significant SOPAS variations, distinguishable from background noise. Furthermore, the trade-off between filtering window size and threshold sensitivity was highlighted, emphasizing the need for an optimized detection strategy. The study continued by applying Complementary Cumulative Density Function (CCDF) analysis to further refine event classification.

The findings demonstrate that SOPAS analysis is a promising tool for rockfall detection. The study identified the averaging window W and ω_{th} as key parameters for event detection, requiring careful optimization. While the tested averaging windows performed well, setting ω_{th} is more crucial, as it depends on event intensity, fiber placement, and background noise conditions.

Finally, this research provides a solid foundation for future work in fiber optic-based geohazard monitoring. The findings can be extended by integrating machine learning techniques for automated event classification, optimizing detection thresholds, and deploying the system in broader environmental conditions.

Appendix A

Data Batches & Their Acquisition Periods/Rates

Circuit	Date	Time Range	Duration	Acqui Rate
7*Circuit 4	30-11-2024	13:43 → 17:00	03 h 17 m	60s/70
		17:00 → 23:00	06 h 00 m	60s/70
		23:00 → 03:34	04 h 34 m	60s/70
	01-12-2024	17:00 → 23:00	06 h 00 m	60s/70
		23:00 → 05:00	06 h 00 m	60s/70
	02-12-2024	05:00 → 11:00	06 h 00 m	60s/70
		11:00 → 17:00	06 h 00 m	60s/70

Table A.1: Acquisition periods/rates for Circuit-4.

Circuit	Date	Time Range	Duration	Acqui Rate
18*Circuit 3	28-10-2024	12:20 → 17:20	05 h 00 m	60s/62
		19:00 → 00:00	05 h 00 m	60s/70
	29-10-2024	00:00 → 05:00	05 h 00 m	60s/70
		05:00 → 10:00	05 h 00 m	60s/70
		21:00 → 02:00	05 h 00 m	60s/80
	30-10-2024	02:00 → 07:00	05 h 00 m	60s/80
		07:00 → 11:43	04 h 43 m	60s/80
	09-11-2024	12:00 → 18:00	06 h 00 m	60s/70
		18:00 → 00:00	06 h 00 m	60s/70
	10-11-2024	00:00 → 03:17	03 h 17 m	60s/70
	27-11-2024	00:00 → 06:00	06 h 00 m	60s/70
		06:00 → 12:00	06 h 00 m	60s/70
		12:00 → 19:00	07 h 00 m	60s/70
	28-11-2024	15:00 → 21:00	06 h 00 m	60s/70
		21:00 → 03:00	06 h 00 m	60s/70
	29-11-2024	03:00 → 09:00	06 h 00 m	60s/70
		09:00 → 15:00	06 h 00 m	60s/70
		15:00 → 21:00	06 h 00 m	60s/70

Table A.2: Acquisition periods/rates for Circuit-3.

Circuit	Date	Time Range	Duration	Acqui Rate
19*Circuit 2	15-10-2024	18:00 $\rightarrow$ 23:00	05 h 00 m	1
		23:00 $\rightarrow$ 04:00	05 h 00 m	1
	16-10-2024	04:00 $\rightarrow$ 09:00	05 h 00 m	1
		09:00 $\rightarrow$ 14:00	05 h 00 m	1
		14:00 $\rightarrow$ 19:00	05 h 00 m	1
		19:00 $\rightarrow$ 00:00	05 h 00 m	1
	21-11-2024	18:00 $\rightarrow$ 00:00	06 h 00 m	60s/70
	22-11-2024	00:00 $\rightarrow$ 06:00	06 h 00 m	60s/70
		06:00 $\rightarrow$ 09:09	03 h 09 m	60s/70
		20:00 $\rightarrow$ 02:00	06 h 00 m	60s/70
	23-11-2024	02:00 $\rightarrow$ 08:00	06 h 00 m	60s/70
		08:00 $\rightarrow$ 14:00	06 h 00 m	60s/70
		14:00 $\rightarrow$ 20:00	06 h 00 m	60s/70
		20:00 $\rightarrow$ 02:00	06 h 00 m	60s/70
	24-11-2024	02:00 $\rightarrow$ 08:00	06 h 00 m	60s/70
		08:00 $\rightarrow$ 14:00	06 h 00 m	60s/70
		14:00 $\rightarrow$ 20:00	06 h 00 m	60s/70
	25-11-2024	02:00 $\rightarrow$ 08:00	06 h 00 m	60s/70

Table A.3: Acquisition periods/rates for Circuit-2.

Bibliography

[1] Hilker, N. Badoux, A. Hegg, C.. (2009). The Swiss flood and landslide damage database 1972-2007. Natural Hazards and Earth System Sciences - NAT HAZARDS EARTH SYST SCI. 9. 913-925. 10.5194/nhess-9-913-2009.

[2] U.S. Commercial Service. (n.d.). *Italy: Landslide and flood management and prevention*. Retrieved from https://www.trade.gov/market-intelligence/italy-landslide-flood-management-and-prevention

[3] Joyce, Karen, Belliss, S. E., Samsonov, Sergey, McNeill, Stephen, & Glassey, Phil. (2009). A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters. *Progress in Physical Geography*, 33, 183–207. https://doi.org/10.1177/0309133309339563

[4] Wikipedia Contributors. (n.d.). Optical fiber history. Wikipedia. Retrieved from https://en.wikipedia.org/wiki/Optical_fiber#History

[5] M2 Optics. (n.d.). History of optical fiber. M2 Optics. Retrieved from https://www.m2optics.com/blog/history-of-optical-fiber

[6] National Institute of Information and Communications Technology. (2024, June 26). World record 402 Tb/s transmission speed achieved with fiber optics. NICT. Retrieved from https://www.nict.go.jp/en/press/2024/06/26-1.html#:~:text=World%20Record%20402%20Tb%2Fs,of%20Information%20and%20Communications%20Technology

[7] Bhagyaraj, S., Oluwafemi, O. S., & Krupa, I. (2020). Polymers in optics. In M. A. A. AlMaadeed, D. Ponnamma, & M. A. Carignano

(Eds.), Polymer science and innovative applications (pp. 423-455). Elsevier. https://doi.org/10.1016/B978-0-12-816808-0.00013-5

[8] Garg, A., Sharma, R., and Dhingra, V. (n.d.). Polarization studies in a computer-based laboratory.

[9] Wikipedia contributors. (n.d.). Unpolarized light – Poincaré sphere. Wikipedia, The Free Encyclopedia. Retrieved February 5, 2025, from https://en.wikipedia.org/wiki/Unpolarized_light#Poincar%C3%A9_sphere

[10] Wikipedia contributors. (2024, May 19). Step-index profile. In Wikipedia, The Free Encyclopedia. Retrieved 07:30, February 8, 2025, from https://en.wikipedia.org/w/index.php?title=Step-index_profile&oldid=1224586845

[11] Hausner M. B., Suárez F., Glander K. E., van de Giesen N., Selker J. S., and Tyler S. W.: Calibrating Single-Ended Fiber-Optic Raman Spectra Distributed Temperature Sensing Data. Sensors, vol. 11, pp. 10859 – 10879, 2011.

[12] Brusin, A. M. R., Rizzelli, G., Fasano, M., Morosi, J., Pellegrini, S., Ferrero, V., Bosco, G., Pilori, D., Parolari, P., Madaschi, A., Brunero, M., Boffi, P., & Gaudino, R. (2024). Overview and analysis of optical sensing techniques over deployed telecom networks. In Proceedings of the 24th International Conference on Transparent Optical Networks (ICTON). IEEE. https://doi.org/10.1109/ICTON62926.2024.10647319

[13] Arthur H. Hartog. An Introduction to Distributed Optical Fibre Sensors. Series in Fiber Optic Sensors. CRC Press, 2017.

[14] Novoptel. (n.d.). LU1000 Laser Unit. Novoptel GmbH. Retrieved February 15, 2025, from https://www.novoptel.de/Laser/Laser_LU1000_en.php

[15] Novoptel GmbH, PM1000 User Guide (Novoptel GmbH, 2025), https://www.novoptel.de/Polarimeter/PM1000_UG.pdf.

[16] D. Charlton, S. Clarke, D. Doucet, et al., “Field measurements of SOP transients in OPGW, with time and location correlation to lightning strikes,” Opt. Express 25, 9689–9696 (2017).

[17] Pellegrini, S., Minelli, L., Andrenacci, L., et al. (2025). Overview on the state of polarization sensing: Application scenarios and anomaly detection algorithms. Journal of Optical Communications and Networking, 17(2). https://doi.org/10.1364/JOCN.17.000A196

[18] Geosolving. (2024, July 10). Debris flow detected by EWS GeoAndes. Geosolving. https://www.geo-solving.com/2024/07/10/debris-flow-detected-by-ews-geoandes/

[19] Savitzky, A., & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry, 36(8), 1627-1639.