TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

Rida Fatima 234023IVSM

SIGNAL PROCESSING, FEATURE DETECTION, AND DATA VALIDATION OF LOW-COST SENSING DRIFTER DATA

Master's Thesis

Supervisor: Laura Piho
PhD

Tallinn 2025

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Rida Fatima 234023IVSM

TRIIVIVA ANDURI SIGNAALANDMETE TÖÖTLEMINE, TUNNUSTE TUVASTAMINE JA VALIDEERIMINE

Magistritöö

Juhendaja: Laura Piho
PhD

Tallinn 2025

Author’s Declaration of Originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Rida Fatima

11.05.2025

Abstract

With the ongoing climate change, glacial meltwater channels have been of significant interest to researchers. These channels are river-like structures formed on the surface of glaciers and carry a large volume of meltwater flowing downstream from the glaciers. They are important to ice sheet mass balance as their development and drainage are connected to changes in ice flow dynamics. Understanding the characteristics of these channels helps estimate the ice sheet melting and predict the ice flow dynamics. To study the characteristics of these channels, researchers at the Centre for Biorobotics, Tallinn University of Technology, developed special drifter sensors. These drifters are low-cost submersible flow sensors that can be deployed on englacial, subglacial, and supraglacial channels in the Arctic region. Englacial channels are penetrated below the surface of a glacier and carry meltwater towards the bed. Subglacial channels are located beneath the ice mass, directing the meltwater parallel to the main ice flow direction. However, supraglacial channels are the streams that form on top of the glaciers that drain the meltwater into the englacial or subglacial channels. Recently, in July 2021, the drifter sensors were deployed in a supraglacial channel in Kongsvegen (Svalbard) to study its characteristics. The collected data is in raw form and requires a substantial effort in cleaning and organising it before conducting any analysis. This hinders the researchers from performing a rapid assessment of the collected data quality and complicates the process of filtering the useful data for further processing. This study primarily focuses on developing *Python* scripts robust enough to filter, preprocess, and analyze the collected dataset. These scripts were customised to investigate possible events occurring in the supraglacial channel, referred to as features, using various statistical techniques. Specifically, the study examined drifter sensor stalls and step-pool formations within the data. The analysis shows that the sensor stall occurred in all the sensors with varying frequency. For some sensors, the stall period was excessively long, up to eight minutes. Similarly, the number of step-pool events varied both among sensors on the same day and across different days. Moreover, the step-pool events were overlaid with GPS data to estimate their spatial locations. This work provides valuable insights for researchers studying supraglacial hydrology by offering a methodological framework to detect and analyze the features and estimate the spatial mapping of dynamic events. These findings contribute to improved monitoring of glacial systems, supporting predictions of meltwater dynamics and related climatic impacts.

Annotatsioon

Triiviva anduri signaalandmete töötlemine, tunnuste tuvastamine ja valideerimine

Seoses kliimamuutuste suurenemisega pakuvad liustikupealsed kanalid teadlastele märkimisväärset huvi. Liustikupealsed kanalid kujutavad endast jõesarnaseid struktuure, mis tekivad liustiku pinnale ning suunavad märkimisväärse osa sulamisveest liustiku sisemusse. Nende morfoloogia ja kuivenduskäitumine mõjutavad jääkihi massibilanssi, kuna on otseselt seotud jäävoolu dünaamikaga. Nende kanalite ruumilise topoloogia mõistmine aitab jälgida jää kiiruse voolumustreid ja võib aidata prognoosida jõgede voolukiirust. Tallinna Tehnikaülikooli biorobootika keskuses on teadlased töötanud välja triivivad andurid nende muutuste uurimiseks. Need seadmed on odavad vooluandurid, mida saab kasutada erinevates liustikukanalites. Hiljutised katsed viidi läbi juulis 2021, kus need seadmed paigutati liustikupealsesse kanalisse Kongsvegeni liustikul (Svalbard), et uurida kanali omadusi. Kogutud andmed on töötlemata kujul, mis raskendab nende kiiret analüüsimist, kuna kasuliku teabe eraldamine ja eeltöötlus nõuavad märkimisväärset töömahtu. Käesolevas töös keskendutakse peamiselt Pythoni skriptide arendamisele, mis võimaldavad kogutud andmeid eraldada, eeltöödelda ja analüüsida. Need skriptid kohandati kanalis esinevate võimalike protsesside (nn tunnuste) uurimiseks, kasutades erinevaid statistilisi meetodeid. Analüüs keskendus eelkõige triivandurite tööseisakute ning jääastmeliste moodustiste tuvastamisele andmestikus. Analüüs näitab, et andurite seisakud esinesid kõigil anduritel erineva sagedusega. Mõne anduri puhul oli seisakuperiood väga pikk, kuni kaheksa minutit. Samamoodi varieerus sammupoolsete sündmuste arv nii sama päeva andurite vahel kui ka eri päevade vahel. Need sündmused pandi üle GPS-andmetega, et hinnata nende ruumilist asukohta. See töö annab väärtuslikke teadmisi liustikupealsete kanalite hüdroloogiat uurivatele teadlastele, esitades metoodilise raamistiku dünaamiliste protsesside tuvastamiseks, analüüsimiseks ja nende ruumilise paiknevuse hindamiseks. Need tulemused aitavad kaasa liustikusüsteemide paremale seirele, toetades sulamisvee dünaamika ja sellega seotud kliimamõjude prognoosimist.

List of Abbreviations and Terms

GPS	Global Positioning System
IMU	Inertial Measurement Unit
GNSS	Global Navigation Satellite System
GLOF	Glacial Lake Outburst Flood
IPCC	Intergovernmental Panel on Climate Change
ACIA	Arctic Climate Impact Assessment
UAV	Unmanned Aerial Vehicles
NASA	National Aeronautics and Space Administration
USGS	United States Geological Survey
InSAR	Interferometric Synthetic Aperture Radar
LiDAR	Light Detection and Ranging
GPR	Ground Penetrating Radar
RDF	Radio Direction Finding
PEARL	Persistent Environmental Awareness Reporting and Location
NaN	Not a Number
ROI	Regions of Interest
IQR	Interquartile Range
hPa	Hectopascal (unit of pressure)
PCC	Pearson Correlation Coefficient
ML	Machine Learning

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Motivation

Climate change is rapidly affecting the northern hemisphere [1]. The northern ice cap, including Svalbard, is significantly losing ice mass [2], which is ultimately contributing to sea level rise. The resulting water from the melting ice runs through the glacial channel system, comprising en-, supra-, and subglacial channels. These channels are important for understanding the topography of glacial networks. Data collection in glacial channels is challenging due to inaccessibility, lack of Global Positioning System (GPS) signals, and harsh conditions. Researchers and environmental engineers have studied the meltwater streams in glaciers to understand their characteristics and reveal their pathways. In previous years, multiple studies were conducted using various sensors to track water channels in glacial environments [3, 4, 5]. A recent experiment conducted by the *Centre for Biorobotics* at Tallinn University of Technology has shown significant advancements in tracking subsurface water flow. These experiments involved low-cost submersible drifter sensors developed in-house and deployed in a supraglacial channel at Kongsvegen (Svalbard) [6]. The drifter sensors were equipped with pressure sensors and an inertial measurement unit (IMU), allowing for precise measurements of linear acceleration (x, y, z) and rotational movement (roll, pitch and yaw) as the sensor moved through the channel. The pressure sensors measured the temperature at the sensor to provide real-time thermal compensation, ensuring high accuracy in pressure readings by correcting for temperature-induced drift and material expansion effects. Additionally, the drifter sensors are equipped with GPS to record the deployment and recovery location with extended logging time. The geometrical configurations and the components of the drifter sensor are illustrated in Fig. 1.

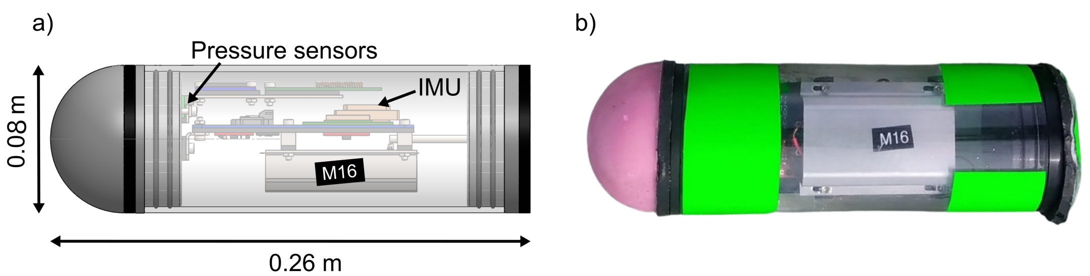

Figure 1. Low cost drifting sensor developed at *Centre for Biorobotics*, Tallinn University of Technology. a) Computer-aided design (CAD) model. b) An image of the sensor used for measurements (adapted from [6]).

The submersible drifter sensor is essentially a cylindrical tube, sealed at both ends to protect its internal components. The diameter of the tube is around 0.08 m with a total length of 0.26 m. One end of the sensor is affixed with a silicone cap, which protects the sensor against large drop impacts during deployment. The components equipped in the drifter sensor and their technical specifications are provided in Tab. 1.

Table 1. Technical specifications of drifter sensor components.

Component	**Model**	**Manufacturer**	**Board/Carrier**
IMU (accel, gyro, mag)	BMX160	Bosch	SEN0373 (DFRobot)
High-g accelerometer	H3LIS331DL	STMicro	SEN-14480 (SparkFun)
2× Pressure (30 bar)	MS583730BA01-50	TE Connect.	Custom PCB
Temp. sensor	(From pressure)	–	–
GNSS module	ZOE-M8Q	U-blox	GPS-15193 (SparkFun)
Radio	RC1701HP-MBUS4	Radiocrafts	–
MCU (logging)	Feather M0	Adafruit	#2796
Batteries (2×)	ICR18650-320PCM	Keeppower	–

The pressure sensors and the IMU in the sensing drifters operate at a sampling rate of 50 Hz. This ensures high-resolution data collection for both pressure, acceleration, orientation, angular velocity, and magnetic field measurements. The data obtained from the sensors is stored in a numeric format to a 16 GB micro SD card. The obtained data was later analysed by reconstructing the water flow path between the start and end coordinates obtained through a global navigation satellite system (GNSS) sensor, providing an insight into subsurface water dynamics. An advantage of using the low-cost sensor is that data quality can be improved by deploying multiple units. However, a major issue with the sensor is the fast validation of data and ensuring its accuracy.

The data obtained from the supra-glacial channel varies from the sub-glacial or englacial channels in terms of the channel characteristics. Sub-glacial channels usually encompass deeper and narrower streams relative to the supra-glacial channel. Another significant difference is the flow hydrology between the two. The supra-glacial channels exhibit faster flow velocity as compared to the sub-glacial channels. Therefore, analysing the characteristics of the supraglacial channel through the signal processing methods is deemed to be challenging. However, in contrast to subglacial environments, the features inferred from supraglacial channel data can often be validated using auxiliary sources such as GPS measurements or video recordings of surface flow events. Although submersible drifters are less accurate than GPS devices and unmanned aerial vehicles (UAVs), which are more advantageous as they can be safely deployed in supraglacial channels with less risk of loss. Moreover, by training these drifters to collect data in the supraglacial channel, we can gather knowledge that helps us understand the subglacial and englacial channels. Also, unlike GPS and drones, drifters can be deployed under the ice, where other technologies cannot reach. In this study, the data collected in Kongsvegen-Svalbard from a supraglacial channel are cleaned, preprocessed, and analysed to investigate the characteristics or features of the glacial stream. This was achieved by composing *Python* scripts to first segregate the useful data, which is later used in detecting the significant features. The *Python* scripts compiled are robust, which assist the researchers in identifying and pre-processing.

1.2 Problem statement

In the context of low-cost sensing drifters, rapid and reliable sensor data validation is essential to ensure the quality and usability of the measurements collected in dynamic and in situ environments such as glacial channels. Submersible drifters, which record data at high frequencies, i.e. 50 Hz, result in gathering extensive and noisy datasets. Effective data validation becomes a critical step before any meaningful analysis can take place, as inaccurate data could lead to incorrect conclusions about subsurface flow patterns. Therefore, a robust and effective method is required to screen out the useful data and detect the features within the preprocessed data.

1.3 Research questions

This study aims to resolve the aforementioned issues by providing a reliable solution for analysing low-cost drifter sensor data. While processing the data, two main research questions are investigated, both of which are central to developing robust and reliable methods for analysing such datasets. These questions aim to explore the effectiveness of signal processing techniques and validation strategies that can be applied to glacial monitoring datasets:

RQ1 Which signal processing method offers the most accurate and efficient solution to assess data quality?

RQ2 What is the most effective method to rapidly validate data collected from low-cost sensing drifters?

Addressing these questions is essential for researchers, as it enables early identification of data quality issues, reducing the time and effort on analysing unreliable datasets and informing whether further measurements are necessary. The outcomes of this research will contribute to developing standardised approaches for processing and validating drifter data, ultimately enabling more reliable monitoring of glacial channels and supporting timely decision-making in environmental management.

2. Background

Climate change predictions show that sea levels will continue to rise [7], but the rate of rise is unknown. The Greenland ice sheet contains sufficient water to contribute to a 7-meter rise in global sea levels [8]. Various models predict changes in ice masses, but they differ in how complex they are and what results they produce. Meanwhile, ocean temperatures around Svalbard are rising [9]. On glaciers, water running on the surface flows into cracks and holes, escaping into the glacial system. To model future scenarios of glacier dynamics, it is important to understand their topography [10]. As glaciers shrink due to climate change, the landscape under them becomes exposed. Getting insights into these dynamics allows scientists to create accurate models capable of predicting the sea-level rise and other environmental impacts.

2.1 Ice flow dynamics

The meltwater flows through the glacier and plays an important role in ice dynamics and its properties. This meltwater can soften the base, aiding in sliding and accelerating the ice flow. The hydrology of glaciers can be classified into four categories: supra-glacial, sub-glacial, englacial and pro-glacial systems [11]. The supraglacial channel forms when surface melting occurs in the firn, an in-between state between snow and ice. The saturation of firn at the surface creates a swamp-like zone, forming the pools of standing water. With the onset of the melting season, the surface drains quickly and exposes the ice beneath [12]. Consequently, the firn zone becomes filled with water. Surrounding ice carries water, mixed with sediment along the glacier's surface, resembling a typical river system. The current study examines the characteristics of a supraglacial channel and explores the features of this channel. As an example, Fig. 2 shows a supraglacial channel at Kongsvegen, Svalbard in which these experiments were conducted. In contrast, subglacial is a passage beneath a glacier or ice sheet through which meltwater flows. These subglacial channels vary in size and their formation depends on several factors [13] such as water temperature, meltwater volume, ice thickness, and surface mass balance. Subglacial processes are essential for understanding glacier movement. Basal meltwater flows [11, 14] through large subglacial networks, which accelerate glacial erosion and ice velocity. The presence of water flows beneath the glaciers and ice sheets influences their response to stress and strain [15]. The englacial channels are formed within the glacier by the tension in the ice that allows water to flow through, such as a moulins. Despite the pressure within the ice sheet, they remain open and are maintained by the continuous melting of water. Also, many hidden englacial channels exist within the glaciers and ice sheets where melted water remains trapped for some time [16]. Water enters the subglacial and englacial drainage systems through features such as moulins, deep cracks, and cut-and-closure systems [17].

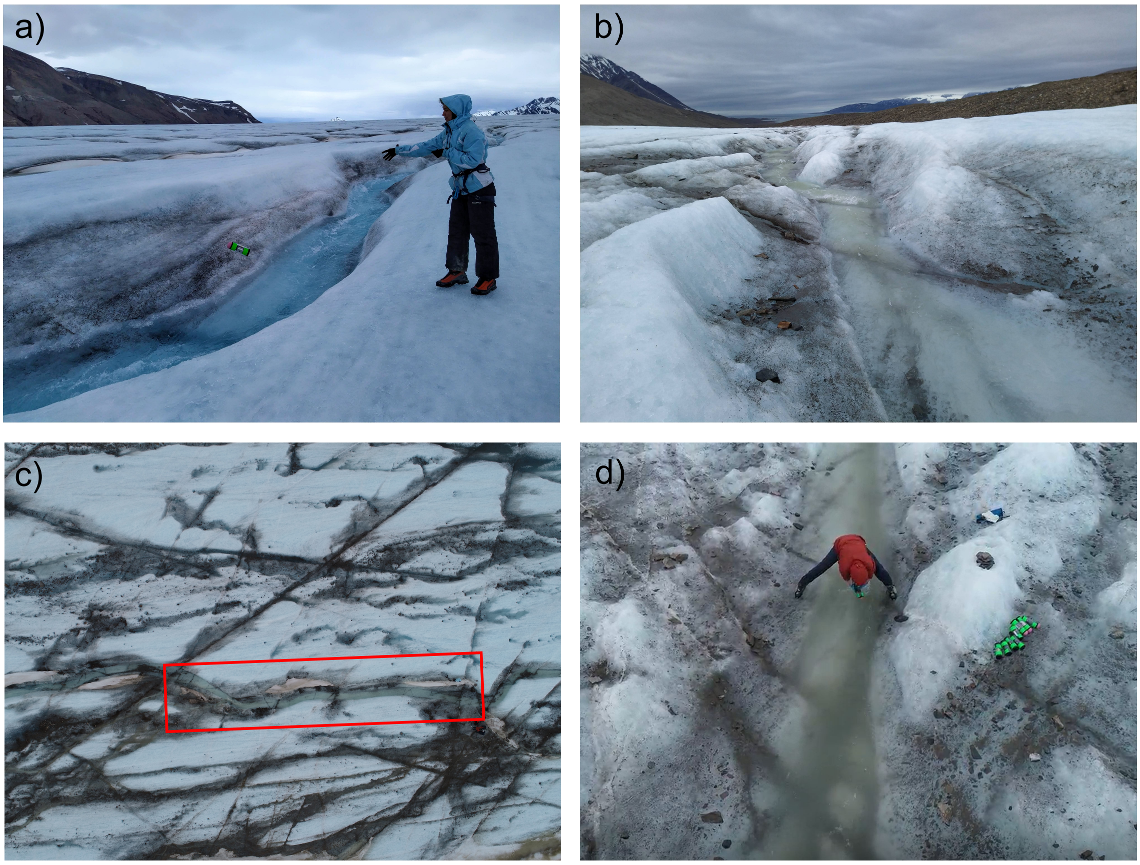

Figure 2. Supraglacial channel on Kongsvegen, Svalbard, documented during field experiment. (a) Sensor deployment in the channel. (b) Aerial view highlighting the channel's morphology. (c) Characteristic features of a supraglacial channel. (d) Sensor recovery after data collection.

Proglacial channels in temperate glaciers are identified by the presence of overflowing meltwater throughout the year, as the ice remains near its melting point, with liquid water present both on the surface and internally within the glacier. Such hydrology can affect glacier behavior by promoting the detachment of icebergs, which accelerate glacier mass loss. Thinning of glaciers has resulted in the formation of moraine-dammed lakes that block the flow of meltwater [18]. The rock debris carried by glaciers builds these moraines. However, when moraine dams fail, a large amount of water is released suddenly, which triggers glacial lake outburst floods (GLOFs) [19]. Such floods [20] are hazardous as they have severe impacts on downriver populations and damage infrastructure. On this matter, research on the Greenland ice sheet states that the presence of meltwater at the bed via basal ice melt is caused by geothermal heat and sliding friction, which affects ice flow and overall glacier dynamics [21]. Understanding these dynamics can lead to a broader study of glacier behavior in response to changing climate conditions.

2.2 Remote sensing in glaciology

Mountain glaciers are a good indicator for understanding ongoing climate change [22] as their properties, size, length, and volume can be measured over time. Glaciological research primarily focuses on the polar regions, and here the impact of climate change is anticipated to be most significant [23], because the loss of ice mass influences the global water resources, and it is a driver of sea level rise [24]. The forecast reported by the Intergovernmental Panel on Climate Change (IPCC) and Arctic Climate Impact Assessment (ACIA) indicates that climate change is extending the summer season in polar regions. This prolonged period of warmth accelerates ice melting, and with more ice melting during summers, the volume of water flowing from glaciers into rivers, lakes, and oceans continues to increase. Also, surface mass balance shows decreased trends as glaciers melt away more ice than they gain each year. Over the past 50 years, these changes have been observed in the Arctic, revealing a consistent trend of glaciers shrinking due to warming climates [25]. These findings highlight the potential of remote sensing to provide valuable information.

Moreover, accurate measurements of ice flow are essential to develop models that can predict the future behavior of glaciers [26]. Inaccurate measurement of ice thickness can lead to uncertainties regarding the timing of future changes in glaciers and their impact on sea level rise [27]. Therefore, the need to observe and detect physical characteristics of glaciers, such as their velocity, ice thickness, ice mass, length, and area, is required as it helps to understand the ongoing transformations. The changes can be monitored through remote sensing, a technique that uses data from sensors mounted on platforms like unmanned aerial vehicles (UAV) [28], commonly known as drones or satellites. Additionally, based on ground *in situ* techniques [29] are available for measuring glacier characteristics, but they often involve intensive labour measurements to be conducted directly on the ice surface during the field experiments.

In the past century, specifically in the 1960s and 1970s, glaciological research utilized satellite imagery from the U.S. Corona program to study glacier changes [22, 30]. The researchers used these corona images to analyze changes in glacier size, mass balance, retreat, and other dynamics over time [31]. In 1972, a program, Earth Resources Technology Satellite (ERTS), later renamed as the Landsat [32], was launched as a partnership between National Aeronautics and Space Administration (NASA) and the United States Geological Survey (USGS), becoming the longest-running Earth observation satellite program for monitoring glaciers. Recently, [33], on September 27, 2021, the Earth observation communities worldwide celebrated the successful launch of LANDSAT 9. These satellite captures high-resolution multispectral images of Earth's surface, which allows optical remote sensing of glaciers, and data can be used in many cases, such as to improve field-based glacial hazard assessments [34].

Revolutionary advancement in glaciological methods was made by Goldstein et al. [35] while observing ice motion by using interferometric synthetic aperture radar (InSAR) on an Antarctic ice stream. The first measurement of ice flow velocity over the outlets of Greenland and Antarctica provided insights into the ice-sheet flow processes [36]. A recent study on Collier Glacier, Oregon, introduced a photon-counting detector by utilizing light detection and ranging (LiDAR) technology to address high losses linked to diffuse optics [37]. These measurements provided valuable information about the glacier ice's physical structure and composition. Ground penetrating radar (GPR) is another widely used technique for remote sensing in my areas, such as earthquake disaster monitoring [38], archaeology [39], soil water dynamics [40], etc. Glacier monitoring is also achieved by remote sensing using GPR. A recent research [41] utilized a set of three multi-temporal Landsat scenes and analysed the status of some Alpine glacier groups (Adamello, Ortles-Cevedale and Bernina) located in northern Italy, commonly known as the "water tower" of Europe. The objective of the study was to investigate the internal structure of the glacier by transmitting radar pulses into the ice and analysing the reflected signals, which gives information about ice thickness and subglacial features.

These technological advancements in remote sensing techniques are very important as they enhance the understanding of glaciology, as they give valuable data for research and environmental monitoring in spatial analysis. However, hydrological variables are hard to observe. Supra-glacial channels, which are numerous meters wide, can be observed via satellite imagery, but measuring their depths and water flow remains challenging[42]. Since optical satellite sensors rely on capturing visible light and other electromagnetic waves, their effectiveness is significantly reduced due to the presence of continuous clouds, which can obstruct signals. [43] Radar sensors can enter clouds and detect changes such as ice thickness, which is critical for observing surface changes and glacier movements. However, they also face a challenge when signals cannot reach certain areas due to steep topography, causing a lack of data in those regions. Additionally, in areas where glaciers are changing surfaces, the radar signal may not match with previous data or features over time [44]. Field-based remote sensing techniques address these problems and offer detailed spatiotemporal resolution, tailored to specific needs.

2.3 Flow sensing drifters

2.3.1 Low-cost drifting sensors in glaciology

Earth's many hydrological systems are located in remote and inaccessible regions. Several methods, such as gas tracing, dye tracing, borehole drilling, and sensor-based direct measurements, have been used to identify the flow path characteristics of water beneath glaciers [45]. Subglacial water pressure has been recorded using tethered borehole sensors on small lowland glaciers [46]. However, it is very challenging to obtain measurements of ice sheets due to their thickness (exceeding 1 km). Also, techniques offered invaluable observations into glacial channels, and helped to validate the existing glacier dynamics [47, 48]. Another approach is to deploy Wireless sensors; for example, the Glacsweb program deployed several sensor 'nodes' to monitor glacier break-up related to climate change on Briksdalsbreen glacier in Norway and Iceland, which transmitted data to a base station on the ice surface [49]. Data gathered through these methods is confined to specific locations.

Lagrangian measurements are based on tracking individual objects or particles as they move through a fluid, and provide a dynamic view of how fluid motion behaves [50]. Using the trajectories of individual particles or tracers, Lagrangian measurements provide information on the position and velocity of fluid elements over time, supporting the study of flow patterns [51]. Surface tracking devices or surface drifters are a type of Lagrangian instrument. They are designed to remain slightly buoyant on the water's surface and move passively with the flow of water. Typically, drifters are tracked using the Argos satellite system, which provides location updates multiple times per day with a positional error ranging from 150 to 1000 meters [52]. Recent models are designed for nearshore applications, and they can be tracked via GPS and cellular phones. They offer improved accuracy of 100 meters with location updates every 10 minutes. This allows researchers to track drifters' movement, such as speed and position over time [53].

In the past, the development of sensing drifters in glaciology has been recorded, with one notable example being the Moulin Explorer [54], which was lost during its initial deployment. A new, successful development of low-cost electronic tracer (E-tracer), more or less the size of a table tennis ball, was introduced [55] for exploring the sub-surface hydrological system of Leverett Glacier, Greenland. This device can travel through the sub-glacial channels and measure pressure directly beneath the ice as it flows. It is equipped with a radio direction-finding (RDF) transmitter, which enables the sensor to determine its location once it reaches beneath the ice sheet. A later version of this E-tracer included pressure sensors, as Lagrangian instruments can be equipped with multiple sensors to gather data along the flow path. To continue advancements in drifters and inspired by [55], a researcher [6] presented low-cost sensing drifters equipped with various sensors to measure water pressure, linear acceleration, magnetic field intensity, and rotational speed within a supra-glacial channel, ensuring consistent and repeatable measurements. The sensor data was used to investigate time series features in a 450-meter-long supra-glacial channel. In correspondence, a recent study [56] conducted in 2019 on Austre Brøggerbreen, Svalbard, utilized low-cost sensing drifters containing IMU and pressure sensors to collect *in situ* data and developed a model for reconstructing the 2D water path flow and pressure distribution of an englacial channel.

2.3.2 Sensing drifters in aquatic environments

Low-cost sensing drifters have been utilized in studies of many hydrological systems. A recent experiment [57] was carried out on the Pirita River, a 105 km-long river in northern Estonia that flows into the Baltic Sea. The low-cost (<150 EUR) drifter equipped with various sensors was employed as a tool for river characterization, to identify large-scale river flow patterns, and to understand river dynamics. Moreover, sensing drifters have also been utilized in coastal monitoring to collect data. The low-cost drifting sensors were designed, and several units were deployed to measure Lagrangian currents. Each drifter was equipped [58] with a visual sensor (a camera facing the ocean floor), an IMU, and GPS, which improved the understanding of coastal water dynamics through data collection. Recently, with advancements in the application of sensing drifters, a study [59] discusses drifters for remotely tracking the trajectory of oil spills in oceans with recent improvements allowing for more continuous and accurate location tracking and data transmission. Furthermore, a new low-cost drifter, the Persistent Environmental Awareness Reporting and Location (PEARL), presented in a study [60], records oceanographic data and is widely used for environmental monitoring. It collects and processes the data quickly on the spot using advanced edge analytics.

Low-cost sensing drifters enable large-scale deployments feasible in limited-resource settings and can operate in difficult-to-access areas like glacial channels, underground rivers, and other subsurface environments. Advancements in sensing drifters are essential for collecting data from glaciological and other environments, improving environmental monitoring, and supporting informed decision-making to tackle global challenges. However, challenges include data processing, such as noise removal, limited battery life, and difficulties in signal processing to obtain accurate information.

2.4 Signal classification methods

Researchers obtain data from sensing devices and store it. Later, this sensing (signal) data from devices like submersible drifters is analyzed, and after results are used in different applications. As stated in a book [61], a signal represents information about the behavior of some physical phenomenon and varies with one or more independent variables, such as time, space, or frequency. Signals take place naturally and can be synthesized, such as an electromagnetic wave used to transform image information. Signals are classified based on various properties, such as time-continuous signals and time-discrete signals. An analog or continuous-time signal $x(t)$ is defined for all values of the time variable t, where t can take any real value. A discrete-time signal is defined at specific, distinct time intervals. It is represented as $x[n]$, where n is an integer (e.g., values like 0, 1, 2, 3, $\ldots$) that indicates the index of a particular time sample.

In hydrological systems, signal classification is essential to interpret sensor data and to gain useful information about physical processes. Identifying and classifying signals helps to understand the flow of glacial channels, hence contributing to the study of glacial hydrology and ice flow dynamics. Classification is sometimes hard and challenging due to environmental factors causing noise in sensor data, and also different physical phenomenon produces distinct signals, which leads to overlapping of classification. A study [62], mapping glacial parameters from space, highlights that remote sensing signals from satellites often encounter noise and inconsistencies due to atmospheric conditions and sensor limitations. Additionally, glaciers have complex surface and subsurface dynamics, making it challenging to classify signals that represent different glacier states or processes. The author also mentions that, at times, it is challenging to distinguish between rock glaciers and debris-covered glaciers in medium-resolution satellite imagery.

In a recent study [63], the author demonstrated the importance of processing seismic signals for identifying subsurface features, successfully detecting a previously unknown subglacial lake on an Alaskan glacier. The study utilized seismometers, which are commonly used to detect earthquakes, to measure vibrations caused by flowing water beneath the ice. By mapping areas of subsurface pressure, the study demonstrated the importance of signal classification in distinguishing between different subsurface materials, such as ice and sediments, based on their seismic output. Seismic signals are analyzed as time-series data, and classification is focused on identifying patterns associated with specific subglacial features. However, reflected signals often contain noise, leading to incorrect classification if signal processing methods are not appropriately tailored.

3. Research Method

This chapter provides an overview of the drifters' data and methods used in analysing these datasets. The research method adopted in this study focuses on the drifters' data collected in July 2021 at Kongsvegen glacier (Svalbard). In these experiments, drifter sensors with unique IDs ranging from M01 to M24 were deployed. On each of the six deployment days, eleven sensors were released. Before each deployment, the flow velocity in the supraglacial channel was measured using a propeller velocimeter. The purpose of measuring the flow velocity through a velocimeter was to check the water velocity inside the channel over a week. The collected data was grouped based on the day of deployment and the sensor ID. The structure of the collected data files is illustrated in Fig. 3. Fortunately, none of the drifter sensors were lost or damaged during deployment. In total, one hundred and seventy measurements were recorded over the six days.

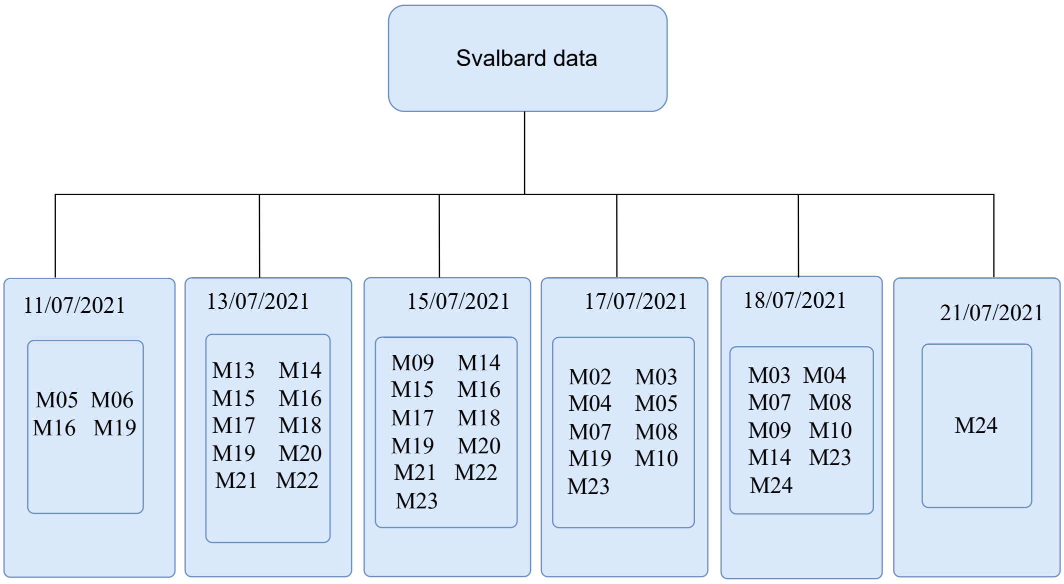

Figure 3. Structure of the metadata directory with sensor data.

Data collected from the drifter sensors is in raw form, and extracting meaningful insights from such data requires an open-source platform. In this study, the Python environment was chosen due to its extensive library support and flexibility in handling complex data preprocessing and analysis tasks. In the initial step, the data was manually classified as either good or bad. The data is considered good or usable if it does not contain data quality errors [64], such as missing values or frozen signals, and includes measurements with timestamps spanning at least five to twenty minutes. The data which did not comply with these criteria was classified as bad or unusable. In the next step, the useful data was filtered and pre-processed before being used in further analyses. This was necessary to ensure the reliability of the collected data beforehand. The final analysis included feature detection through the signal processing methods and validating these features using the auxiliary data. The research approach employed in this thesis was based on a standard data wrangling approach [65], which involves various stages, i.e. data cleaning, transformation, and validation. This approach ensures consistency, reduces errors, and improves data quality. It also enhances reproducibility and efficiency across analyses. Fig. 4 provides a detailed overview of each step involved in the data processing workflow.

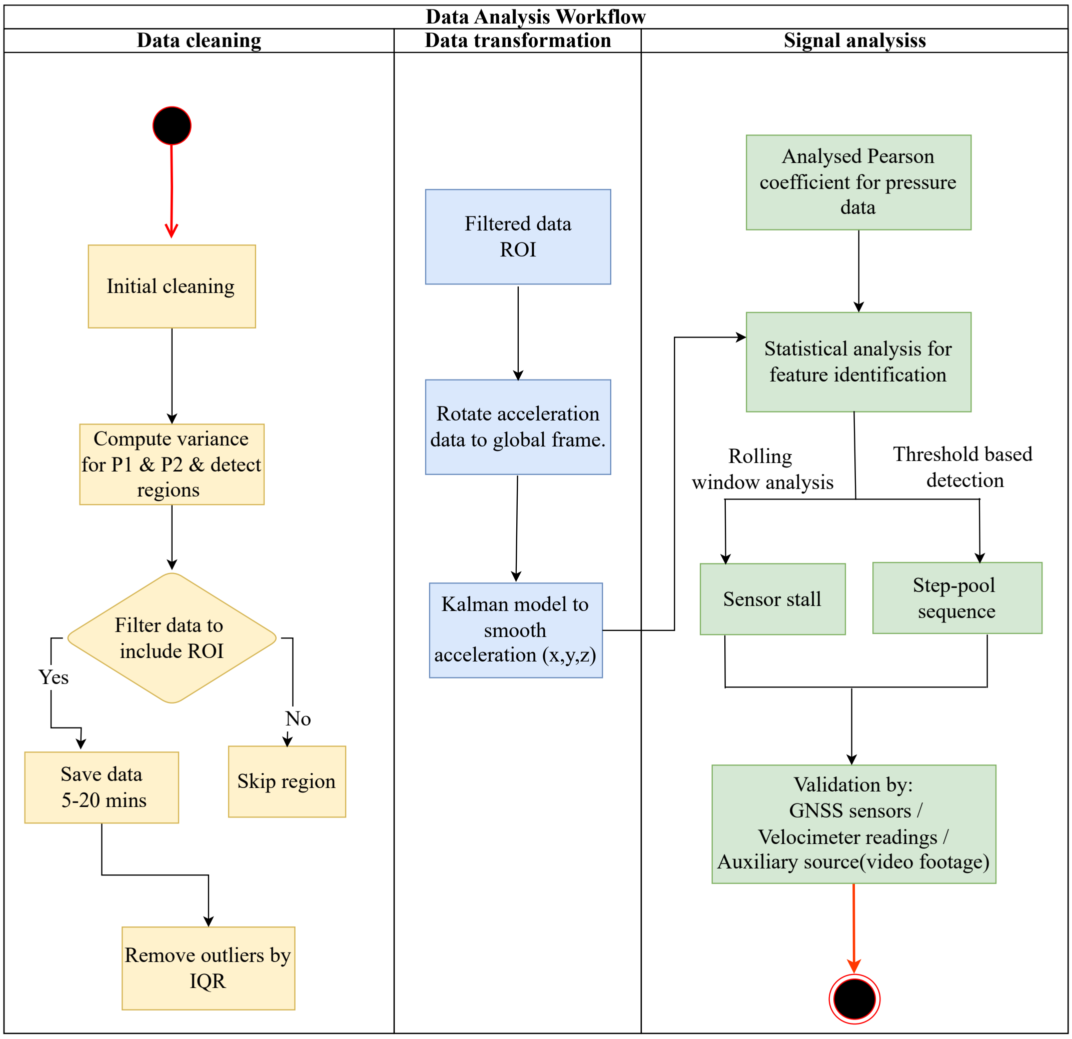

Figure 4. Data analysis workflow illustrating the process from raw sensor data to feature detection. The steps include raw data cleaning, transformation, and signal analysis for identifying supraglacial features.

3.1 Data cleaning

In the cleaning stage, the data was cleaned and classified as good or bad data. The numeric format of the recorded data consisted of frequency, pressure, acceleration, gyroscope, magnetometer, and other unordered GPS data. Initially, the first seventeen columns were taken into consideration, ignoring the redundant columns, and invalid entries were replaced with a Not a Number (*NaN*) expression, later converted into numeric types. The headers of the first seventeen columns in each file were as follows:

`time, pressure1, temp1, pressure2, temp2, accx, accy, accz, gyx, gyy, gyz, magx, magy, magz, hgax, hgay, hgaz`

The first column represents the time stamps or frequency. The `pressure1` and `pressure2` columns record the pressure measurements from two pressure sensors, along with their corresponding temperatures `temp1` and `temp2`. The Inertial Measurement Unit (IMU) records three-axis acceleration (`accx`, `accy`, `accz`), angular velocity (`gyx`, `gyy`, `gyz`), and magnetic field strength (`magx`, `magy`, `magz`). Moreover, a high-g accelerometer captures high-range acceleration data along the x, y, and z axes (`hgax`, `hgay`, `hgaz`). These variables are sufficient to understand the flow characteristics in a supraglacial channel.

The data classification was based on the data from the pressure sensors. For that, the frequency was converted into time over which the pressure data was plotted for each sensor measurement. The total time of each sensor measurement varied depending on when the sensor was switched on before deployment and switched off after the retrieval. Before classification, the sensor's active deployment period, defined as the region of interest (ROI), was identified, corresponding to the time between when the sensor was released into the stream and when it was retrieved. Following this, a signal classification criterion was established: the pressure data within the region of interest had to span at least 5 minutes and no more than 20 minutes to be considered good. Data outside this range was labelled as bad data. Among the recorded data files, some contained only a few entries or lasted less than five minutes. Besides that, some data files had no measurements at all or contained extended frozen signals or non-varying data throughout the file. These data files were not considered for further analysis.

The identification of the region of interest (ROI) within the pressure signal was carried out using a statistical approach (see Fig. 5). A rolling window analysis was used in which the variance of both pressures, i.e. pressure 1 and pressure 2, was calculated within a defined window size. A minimum threshold in rolling variance was defined to identify regions of high variability within the signal. Segments exceeding this threshold were designated as regions of interest corresponding to the movement of the sensor within the channel. The variance (σ^2) of the pressure signal was calculated using the following expression:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \tag{3.1}$$

where μ is the local mean, x_i represents the individual pressure readings within the window, and N is the window size.

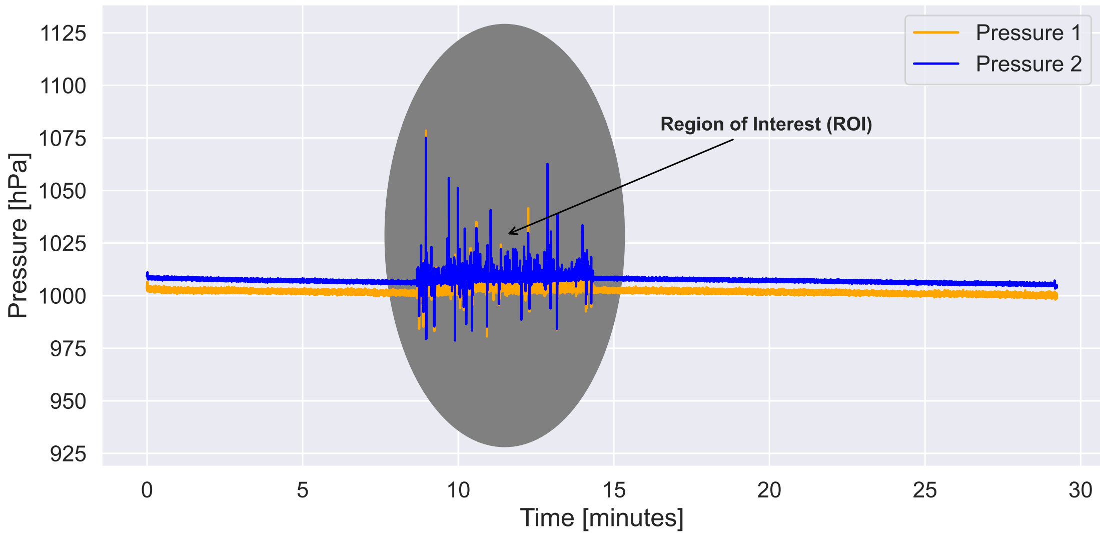

Figure 5. Pressure 1 and Pressure 2 plotted over the complete sensor deployment time. The gray region shows the region of interest later screened.

The regions of interest from each sensor were saved as separate *.csv* files, later used for preprocessing. To eliminate the outliers in the pressure signal caused by anomalies or noise, an Interquartile Range (IQR) method was applied to both Pressure 1 and Pressure 2. This outlier removal step ensured cleaner data for further identification of features from the signal.

3.2 Data preprocessing

Following successful data cleaning, the dataset was prepared for the next phase: preprocessing. This stage involved structured steps to refine the data further, ensuring its suitability for subsequent feature detection analysis. The following sections briefly explain the steps involved in the preprocessing of the data.

3.2.1 Data transformation

The Inertial Measurement Unit (IMU) used in the sensing drifter records acceleration data along three local axes: x, y, and z as shown in Fig. 6. In the local drifter's frame, the axes represent the orientation of the drifter, such that the x-axis points lateral, the y-axis points forward, and the z-axis points downwards to the drifter's body. However, the orientation of the IMU within the drifter can vary and might not always stay in the same position due to external forces and water flow dynamics during the deployment. Therefore, the raw acceleration values do not directly correspond to a global frame of reference, thus making physical interpretation and feature extraction difficult. To address this, acceleration vectors were transformed into the fixed global coordinate system.

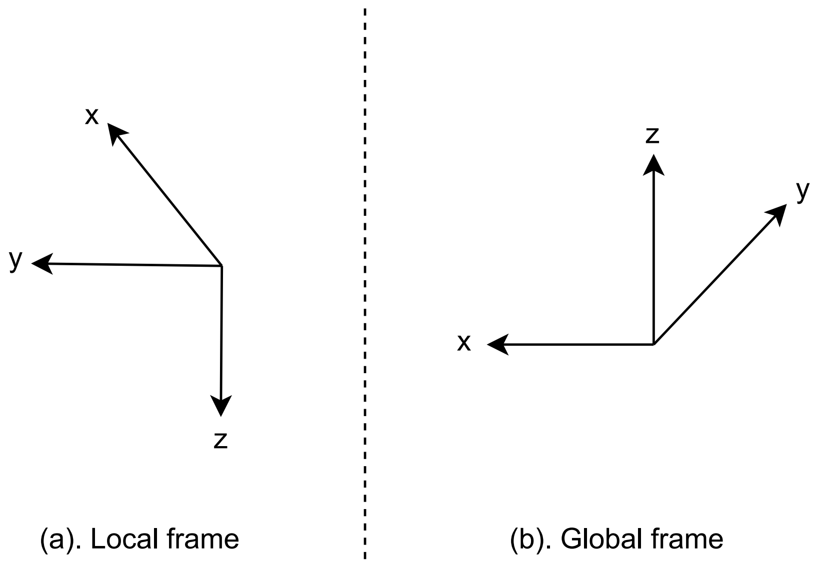

Figure 6. (a) Orientation of axes for drifter's local frame. (b) Orientation of axes for global frame

A rotation matrix was applied to convert the body-frame acceleration vector into the global frame. The transformation used is a 90-degree rotation around the z-axis combined with a flip in the z-direction to match the orientation of the sensor when deployed. The rotation matrix transforms the original acceleration vector into the global frame by matrix multiplication and is defined as follows:

$$R = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (3.2) \qquad a_{local} = \begin{bmatrix} accx \\ accy \\ accz \end{bmatrix} \quad (3.3)$$

$$a_{global} = R \, . \, a_{local} \tag{3.4}$$

The transformation results in the following global axis interpretations:

X_{forward}: aligned with the original lateral direction ($accy$).
Y_{lateral}: aligned with the negative of the original forward direction ($-accx$).
Z_{upward}: aligned with the negative of the original upward direction ($-accz$).

Acceleration transformation was essential to align the data with the direction of gravity and horizontal flow, as it corrects sensor tilt and orientation changes during movement, thus allowing a meaningful comparison and analysis, such as pattern recognition and feature detection across different deployments and periods.

3.2.2 Kalman filter

Following the coordinate transformation of acceleration data, the data exhibited significant environmental noise, which complicated the interpretation of the features. Therefore, to enhance signal quality and extract significant features, the acceleration data was smoothed using a Kalman filter algorithm [66]. It is a mathematical technique that is commonly used in signal processing to help remove noise or errors from data. It is quite efficient, especially when there is missing information in the data; it can give the best possible estimate of how a system is changing over time. This filtering approach allowed for effective noise reduction while preserving the dynamic characteristics essential for analysis. This method provides an optimal solution for tracking and predicting data in various applications. Fig. 7 demonstrates a potential use case for the Kalman filter.

Ideally, there are known inputs and the Kalman filter uses them to predict system changes over time [67]. The drifter sensors move passively through the supra-glacial channel, carried by the natural flow of water while encountering slopes, obstacles, and varying hydraulic conditions along their path. The IMU units inside the drifter sensors record the motion of the drifter, i.e, acceleration and orientation, which are subject to environmental noise that comes from ice water dynamics. Fig. 7 illustrates the Kalman filtering process employed in this study.

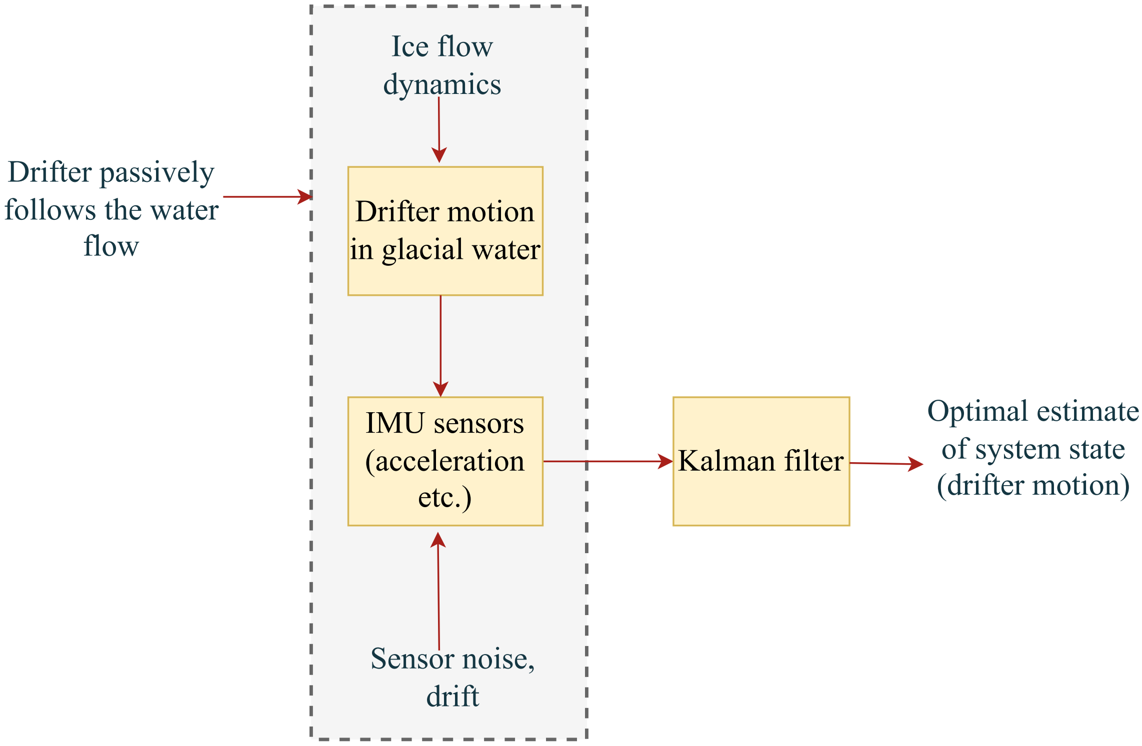

Figure 7. Kalman filtering process employed in current study [adapted from [68]].

The mathematical used for Kalman filtering is expressed as follows:

$$\hat{x}_k = \hat{x}_{k|k-1} + K_k(z_k - H\hat{x}_{k|k-1}) \tag{3.5}$$

where:

$\hat{x}_k$ is the updated state estimate,
$\hat{x}_{k|k-1}$ is the prior state estimate,
z_k is the measurement at time k,
H is the observation model (matrix),
K_k is the Kalman gain, calculated as:

$$K_k = \frac{P_{k|k-1}H^T}{HP_{k|k-1}H^T + R} \tag{3.6}$$

Where $P_{k|k-1}$ is the predicted error covariance, Q and R are the process noise covariance and measurement noise covariance, respectively. The Kalman filter was applied to each axis of IMU recorded data. i.e. in acceleration data $accx$, $accy$, and $accz$. The filter was adjusted by using a process variance Q of $1 \times E^{-4}$, reflecting the assumption that the drifter's acceleration is smooth and does not fluctuate suddenly, unless external forces act on it. Also, a measurement covariance R of 0.05 was configured, which represents the expected noise from the IMU sensor readings. The Kalman filter combines both predictions using a weighted average, where the weights depend on the relative certainty (inverse of covariance) of each source. This balances the expected smoothness of the true signal against the uncertainty/errors in the sensor reading. This process is recursive, efficient, and allows real-time state estimation. The resulting filtered signals provide a reliable estimate of the drifter's true acceleration and are used for further analysis.

3.3 Data analysis

The final step in this study is to analyse the preprocessed data of the drifter sensor to identify the features or characteristics of the supraglacial channel. The drifter sensor transits through the stream, encountering key events referring to the characteristics of the channel. These events were identified as features such as a step pool or sensor stall, etc. The data was qualitatively assessed using statistical techniques. The identification of such features helps us understand the characteristics of these streams as well as possible sensor anomalies that occur in such remote environments.

3.3.1 Pressure correlation

The pre-processed data includes two sets of pressure readings, i.e, pressure 1 and pressure 2, recorded simultaneously by two sensors. The pressure readings are measured in hectopascals (hPa) and capture atmospheric pressure variations over time. It was reasonable to investigate whether the two pressures, mounted on the cylindrical anterior part of the drifters, are correlated. To assess consistency between the two pressure sensors across multiple measurements, a statistical correlation analysis was performed on the entire dataset.

To compare two pressure data readings to see how similar they are, a *Python* script was written to recursively search all subdirectories containing filtered pressure data files and compute the Pearson correlation coefficient (PCC). The method assesses the strength of the linear relationship between the two pressure signals across the entire dataset. The Pearson correlation coefficient r is calculated using the formula:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{3.7}$$

where x_i and y_i represent the individual data points of pressure 1 and pressure 2, respectively, and $\bar{x}$ and $\bar{y}$ are their corresponding means.

The analysis extracts metadata such as date and sensor identifiers from the folder structure and assigns a unique label to each field measurement. The results are compiled, and plots are generated to visualize the distribution of correlation values across the sensors. This approach provides insight into sensor agreement trends over time across drifters and highlights datasets with notably strong, moderate, or weak inter-sensor correlation for all data samples.

3.3.2 Feature detection in supraglacial channel

Supraglacial channels are formed on the surface of a glacier due to melting ice. The identification of key features and events occurring within these channels requires signal processing methods. In this study, two different key events, i.e. drifter sensor stall and step pool, were investigated using distinct analytical methods. The identified events were later overlaid on the GPS collected data to estimate their positioning along the track. Moreover, video recordings of the experiments were used to fine-tune the analytical parameters for feature identification and to validate the results.

3.3.3 Drifter sensor stall

In this study, specific segments of the drifter trajectory called stall regions were identified, where the sensor temporarily stopped moving or became lodged in place. These regions are of particular hydrodynamic significance, as they often correspond to distinct stream features. For example, the sensor occasionally remained stationary behind submerged ice boulders or became trapped on ice meanders. In the latter case, manual assistance was required to push the drifter into the main flow. Fig. 8 shows the drifter sensor holding a station behind an ice boulder.

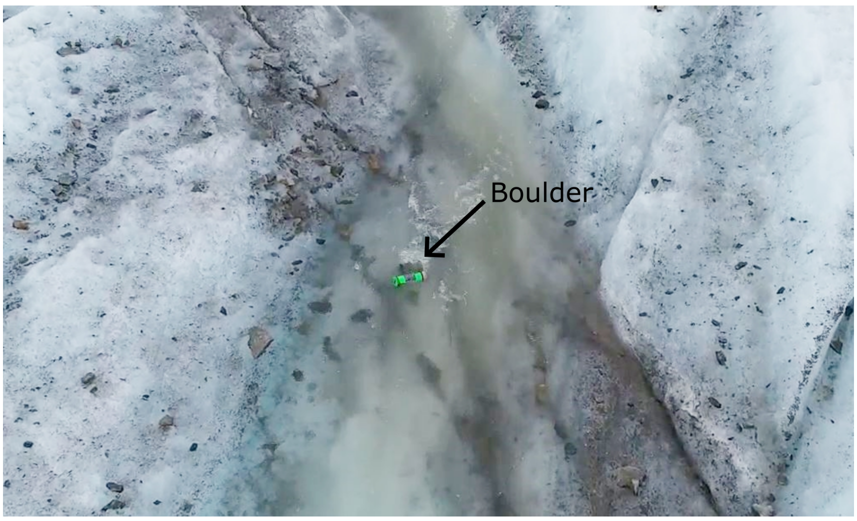

Figure 8. Drifter sensor holding a station behind the boulder during sensor stall.

The identification of these stall regions was primarily based on two key variables: acceleration and pressure. During sensor stall, acceleration in the x, y, and z directions remained constant, indicating no significant movement or change in the sensor's position. Meanwhile, the mean pressure values remained stable in these regions, showing a low variance in the pressure measurement, again suggesting minimal sensor activity. Thus, the following criterion was established to identify instances of drifter sensor stall within the signal:

$$\left.\begin{array}{l}\text{- Constant acceleration in the } x, y, \text{ and } z \text{ directions}\\ \text{- Minimal variance in Pressure 1 and Pressure 2}\end{array}\right\} Sensor\ stall$$

Sensor stall regions were identified using rolling window analysis applied to the time series data. A detailed description of this analysis method, including the parameters used, is provided in the following section.

Rolling window analysis

Rolling window analysis is a common technique in signal processing, used to analyze local patterns in the signal by applying statistical measures over a fixed window size. Rather than computing statistics on the entire dataset, rolling analysis slides the window forward one data point at a time, allowing dynamic observation of how metrics evolve over time.

In this study, the variance of the filtered acceleration data in all three directions ($accx$, $accy$ and $accz$) was calculated in a rolling window of 100 samples. Regions where the variance dropped below 10% and remained consistently low for at least 3 seconds were classified as stall regions. The rolling variance σ_t was computed by evaluating, at each time step t, the variance of data points within a predefined window of size n, as given by the following expression:

$$\sigma_t = \frac{1}{n} \sum_{i=0}^{n-1} (x_{t-i} - \overline{x}_t)^2 \tag{3.8}$$

This method enables the detection of subtle changes in signal dynamics. The mean acceleration values of the filtered data in both $accx$ and $accy$ were close to zero. However, the mean acceleration in the upward direction $accz$ remained $9.8\ \mathrm{m/s^2}$, due to the gravitational acceleration while traversing supra-glacial channels. In regions where the sensor experienced a stall, the acceleration values remained closer to their mean values with low or zero variance.

3.3.4 Step-pool events

Supraglacial channels exhibit significant geomorphological features such as steps and pools, which form due to differential melting of the glacier surface. These features influence the hydraulic behaviour of the meltwater flow, creating zones of high acceleration at steps and relatively low acceleration regions in pools, which can contribute to further ice erosion and alter the channel morphology over time, Fig. 9. A comprehensive description of step-pool features can be found in the work of [6], which serves as a key reference for this study.

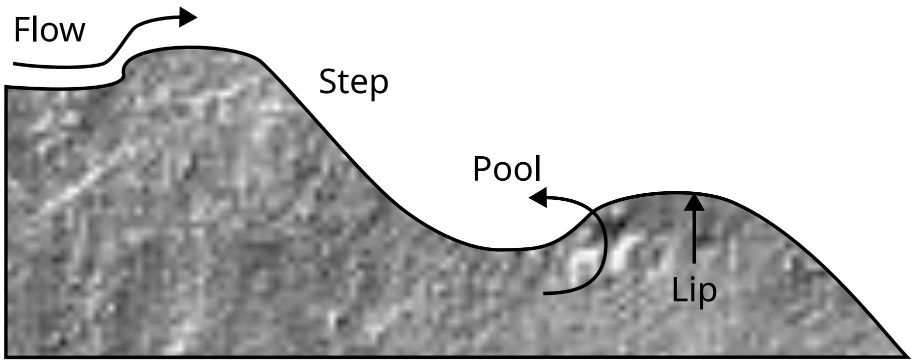

Figure 9. Illustration of the step-pool feature [adapted from [69]]

In this study, steps and pools were investigated using a multi-variable analysis. Step-pool features were identified by applying a predefined criterion to the acceleration and pressure data, based on distinct signal characteristics outlined below.

$$\left.\begin{array}{l}\text{- Higher acceleration in the forward direction}\\ \text{- Low pressure variance}\end{array}\right\} Steps$$

$$\left.\begin{array}{l}\text{- Decrease in gravitational acceleration}\\ \text{- High variance in pressure}\end{array}\right\} Pools$$

These areas were initially identified as step-pool regions based on the time series data and were subsequently cross-checked across all drifter datasets collected each day. Given the high probability of false positives in such detections, only those regions that consistently appeared across multiple datasets were ultimately labelled as step-pool regions.

Threshold-based event detection

As the step-pool event is explicitly linked to multi-variable dependency, a relatively simple approach is applied to detect it within the signal. The threshold-based detection is quite simple as it examines point by point the successive data values in the trace. If the value is above or below the defined thresholds, the event is marked as started, and when the value crosses the threshold in the opposite direction, the event ends. As in our case, the threshold values were set for multiple variables, i.e., acceleration and pressure, and the event detection was split into two steps. In the first step, the mean gravitational acceleration $\overline{accz}$ was deducted from the filtered acceleration $\widehat{accz}$ and was divided by its standard deviation to obtain the normalised gravitational acceleration. The normalised acceleration accounts for the sensor sensitivity and biases and is therefore preferred. The normalised acceleration was later multiplied by a constant to set the threshold for acceleration. Similarly, in the second stage, the median pressure ($\widetilde{p}$) from both pressure sensors was added to the product of their standard deviation and a constant. The value of the constant was input to the system to tune the sensitivity of the algorithm in capturing the pressure peaks. Finally, the regions where the steps and pools occurred repeatedly in sequence were called the step-pool events. The current study was tuned to identify the step pool events where the pool depth was a minimum of 6 cm.

Table 2. Thresholds specified for the step-pool events.

Steps		**Pools**	
gravitational acceleration g [m/s]	min step duration [samples]	pressure threshold [hPa]	rolling window size [samples]
$< \frac{\widehat{accz}-\overline{accz}}{\sigma_{accz}} \times c$	5	$\widetilde{p} + (\sigma_p \times c)$	2

3.3.5 Data validation through GPS tracking

Once the key features were identified, they were overlaid onto the GPS data to estimate their positions along the drifter's path. The drifter sensors were capable of recording Global Positioning System (GPS) data, which could be used to identify the positioning of the sensor, but unfortunately, among all the drifters deployed, only some could record it in sections with little accuracy. However, the surface drifter data from the Global Navigation Satellite System (GNSS) were accurate and provided an estimate of the path, which was later overlaid with the drifter sensor data. As mentioned, the drifters' data were in sections; therefore, the closest time steps to the available data were selected to represent the estimated location of the events. The total track length of the GPS data was around 590 m with a slope of 19 m (see Fig. 10).

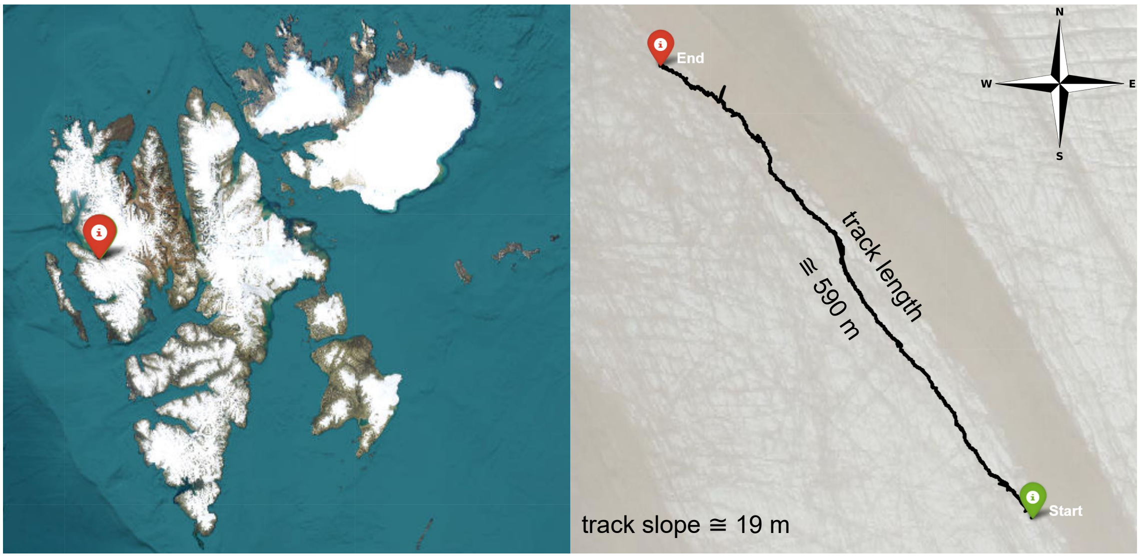

Figure 10. Drifter's path on supraglacial channel in the polar region.

4. Results and Discussions

This chapter presents the results of analyses conducted on the acquired drifter sensor datasets, including data cleaning, preprocessing, and feature detection. It also highlights the outcomes of the statistical approaches used, their effectiveness in identifying key features, and the advantages they offer in signal interpretation.

4.1 Data cleaning and preprocessing

The raw data, once cleaned using a rolling window analysis, yielded a section defined as the region of interest (ROI). As discussed in the previous chapter, the ROI corresponded to data segments between 5 and 20 minutes in duration within each dataset. Following the cleaning step, the data underwent outlier removal using the Interquartile Range (IQR) method. The resulting pressure data, for example, is illustrated in Fig. 11.

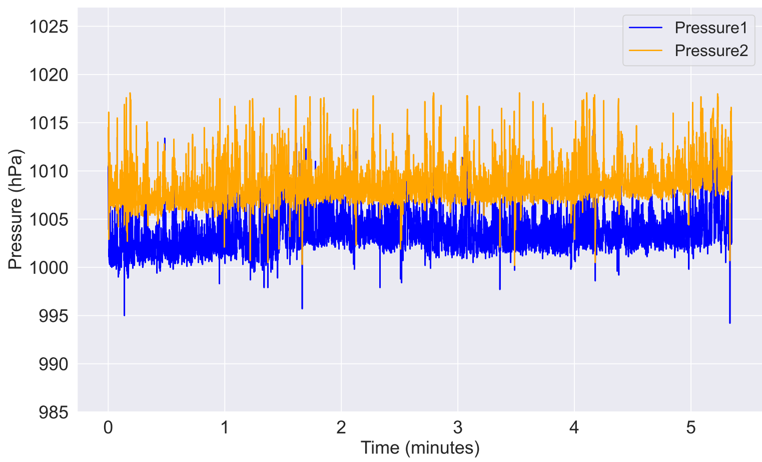

Figure 11. Region of interest ROI extracted with outlier removed from the Pressure 1 and Pressure 2 data.

Both the rolling window analysis and the IQR method were quite effective in extracting relevant signal sections while minimising the impact of noise and anomalies, thereby ensuring cleaner input for the subsequent feature identification analysis. These methods not only enhanced the reliability of the data but also helped retain meaningful variations crucial for detecting events such as sensor stalls and step-pool interactions. Their combined use provided a robust framework for preprocessing, balancing sensitivity to signal fluctuations with resistance to outliers.

4.2 Pressure correlation

After the successful data cleaning and preprocessing, the pressure sensor data was further analysed to identify the correlation between pressure 1 and pressure 2. At this stage, it was essential to quantify the correlation between the two pressure sensors before proceeding with further analysis. The Pressure 1 and Pressure 2 are pressure sensor data collected by the two pressure sensors located around the spherical head of the drifter sensor. The Pearson correlation coefficient r was calculated for each sensor measurement, followed by a box plot (see Fig. 12).

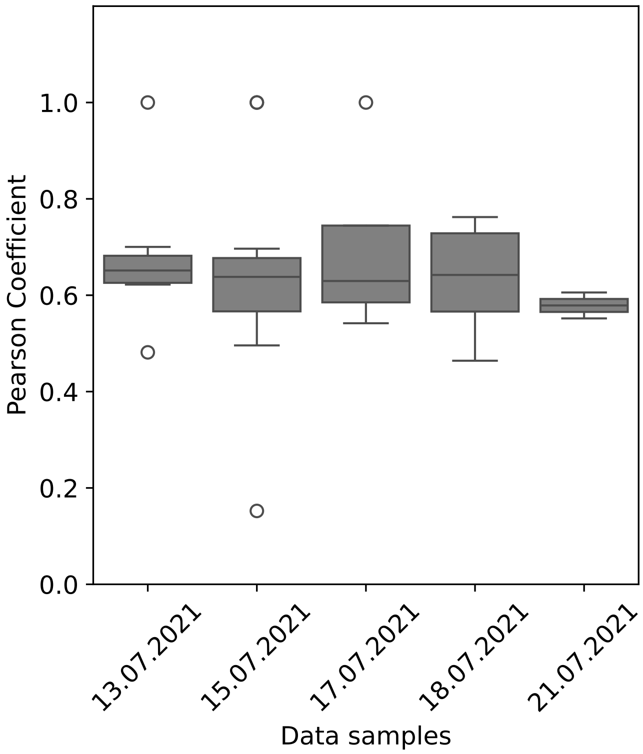

Figure 12. Correlation between the measured Pressure 1 and Pressure 2, based on the Pearson coefficient of the sampled data.

From the data analysed, it was observed that the Pearson correlation coefficient remained within a moderate range, between 0.5 to 0.6, suggesting that both sensors generally recorded similar pressure readings and responded equally to any changes while the sensor moved through the channel. Additionally, a few datasets showed a strong correlation, with values exceeding 0.7, indicating that the sensors captured consistent trends in pressure variations. An outlier with weak similarity was observed in the M16 sensor readings from the field experiment conducted on 15-07-2021. This dataset showed a low Pearson correlation of 0.152, indicating that M16 recorded distinctly different pressure patterns compared to the other sensor. An ideal case of perfect correlation was also observed, with a Pearson correlation value of 1.0 between pressure 1 and pressure 2. However, this was later considered an outlier, as both pressure readings contained identical values, resulting in the perfect correlation. Specifically, this occurred in the M17 sensor measurements taken on 13.07.2021 and 15.07.2021, as well as in the M05 sensor readings recorded on 17.07.2021. Overall, correlation results showed consistency in the pressure readings and reflect reliable sensor performance across different sensors embedded in the drifting sensors.

4.3 Feature detection

The preprocessed dataset from the drifter sensor was further analysed to identify key features encountered within the supraglacial channel. The subsequent sections describe the results of the two potential features, i.e., drifter sensor stall and step-pool event, identified in the measured data using different statistical techniques. These findings show the estimated number of times these events happened along the drifter's transit path and possible reasons causing these events.

4.3.1 Sensor stall

The drifter sensor moving through the supra-glacial channel stalled multiple times during its transit. Data analysis revealed two distinct stall mechanisms: (1) brief interruptions (lasting seconds) caused by ice boulders obstructing the sensor, and (2) prolonged immobilisation when high-velocity flow displaced the sensor onto ice bars. These stall periods were systematically identified in each sensor's dataset to evaluate individual performance metrics, including total run time versus stall duration. Fig. 13 shows a heat map of the sensor stall frequency on each measurement day. A comprehensive overview of the data shows that the average stall frequency of the measurements was around 5, with an average stall time of approximately 3.8 mins.

The M04 sensor on 18.07.2021 exhibited the highest stall frequency of 22 events, with a total stall duration of 8 minutes. Table 3 shows the mean flow velocities measured using the propeller velocimeter along the channel during the field experiments. The elevated number of stall events recorded on 18.07.2021 is truly supported by the velocity data in the table, as it corresponds to the lowest average flow velocity observed during the study period, measured at 1.13 m/s.

Table 3. Mean flow rate observed using anemometer on various days.

Measurement day	13 Jul	15 Jul	17 Jul	18 Jul
Velocity [m/s]	1.40	1.77	1.82	1.13

The inverse relationship between flow velocity and stall frequency suggests that reduced water movement increased the likelihood of stall conditions, as slower velocities limited the drifter's ability to overcome local flow resistance and pressure fluctuations, therefore, an increased count of stalls is observed on 18.07.2021. Moreover, it can also be observed that on 17.07.2021, the average stall frequency count was 2, the lowest observed during the day, which reflects the highest recorded flow velocity of 1.82 m/s on that day.

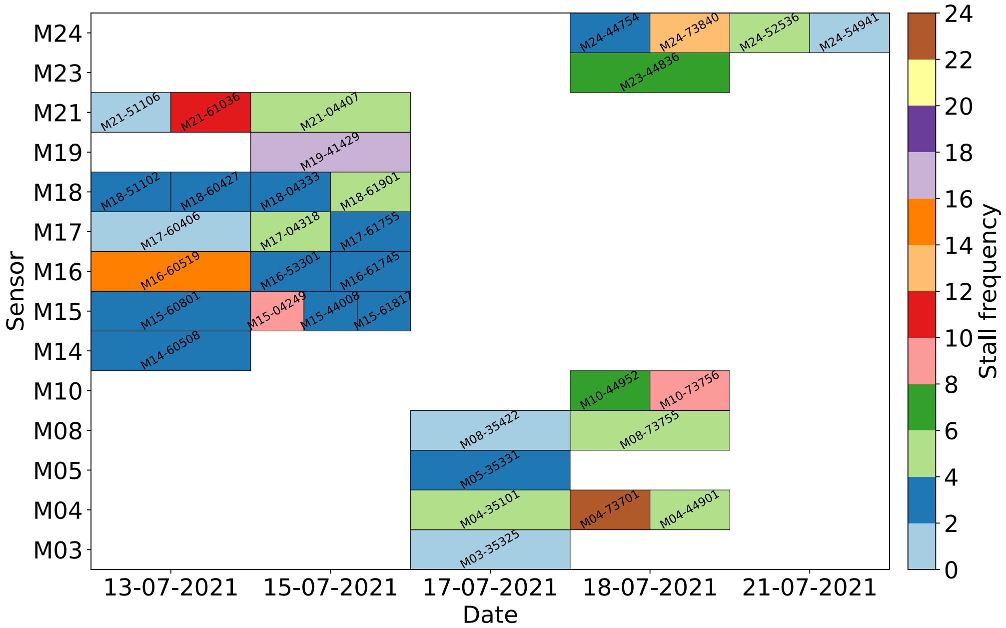

Figure 13. Heat map illustrating the stall frequency of the deployed sensors on different days.

The individual variable, such as acceleration (x, y and z), post Kalman filter application of the M04 sensor on 18.7.2021, also shows the signal stall durations. The sections of the stall signals can be viewed in Fig. 15, highlighted with grey hatched regions. A cross comparison shows that during these signal stalls, the Pressure 1 and Pressure 2 exhibited minimum variance (see Fig 14).

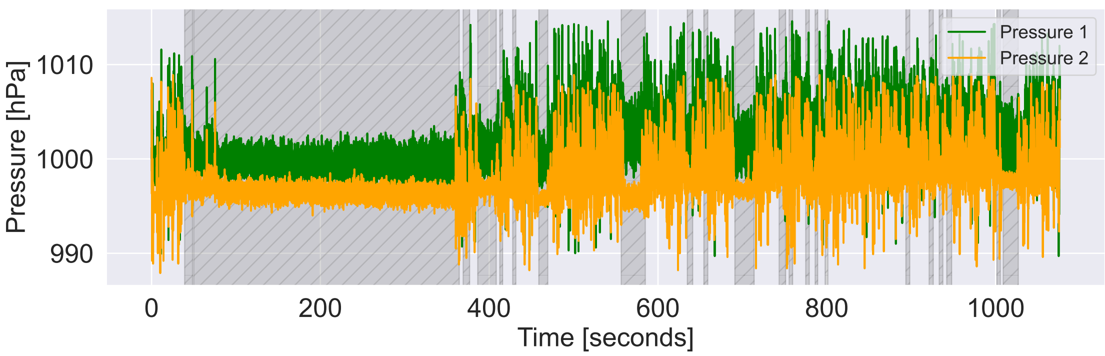

Figure 14. Pressure 1 and Pressure 2 plotted over time from screened data.

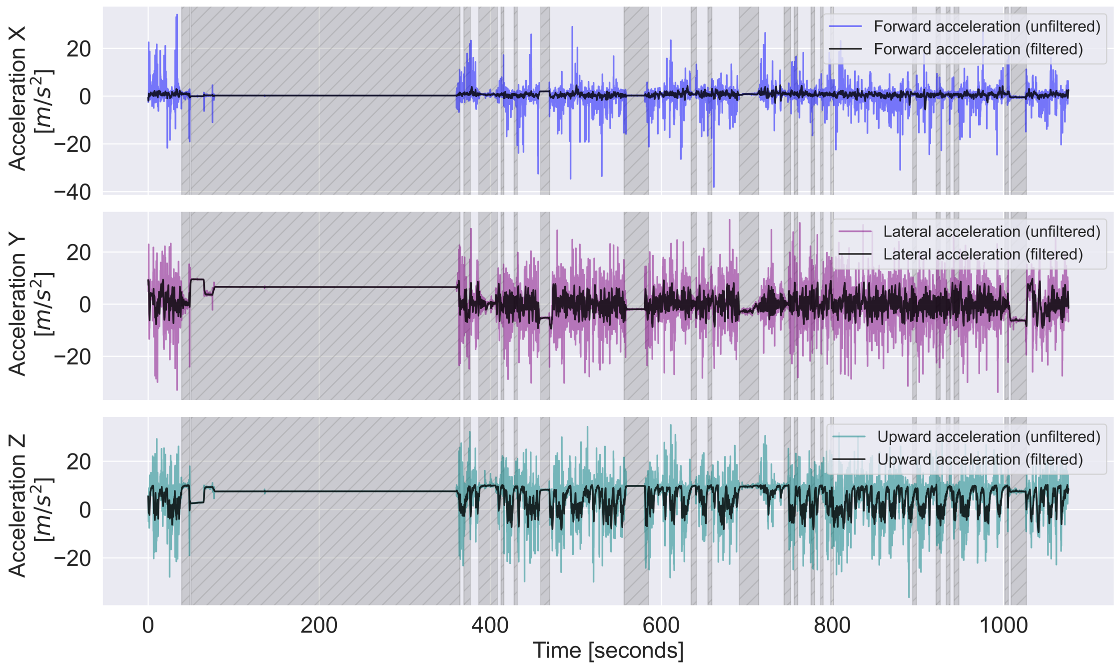

Figure 15. Time series acceleration data of all three axis overlaid with the filtered acceleration obtained through Kalman filter (black line). The gray hatched regions represent drifter stall regions.

4.3.2 Step-pool event

The step-pool events described in the method section were identified by a two-step identification method. The initial step involved the search for regions where the upward/gravitational acceleration falls below zero. These regions were marked as the step regions. In the second step, the pressure peaks were checked throughout the signal in both Pressure 1 and Pressure 2, marking these regions as pools. Finally, Fig. 16, those time stamps where both the step and pools were occurring in sequence, were identified as the step-pool events.

For example, the M04 sensor from July 18, 2021, shows the step-pool event occurring at approx 27.5 sec (see Fig. 17). During the step-pool event, it is observed that there was a deceleration in the forward acceleration, whereas the lateral acceleration exhibits subtle fluctuations. The gravitational acceleration starts to descend at 22.5 sec where a smaller peak was observed in the pressure as well. Later at 27.5 secs when the gravitational acceleration touches its minimum value of zero, the pressure also exhibits a higher value up to 1025 hPa, which represents the hydrostatic pressure on the drifter at that specific depth in the pool. The depth of the pool was estimated empirically to be 25 cm at this location.

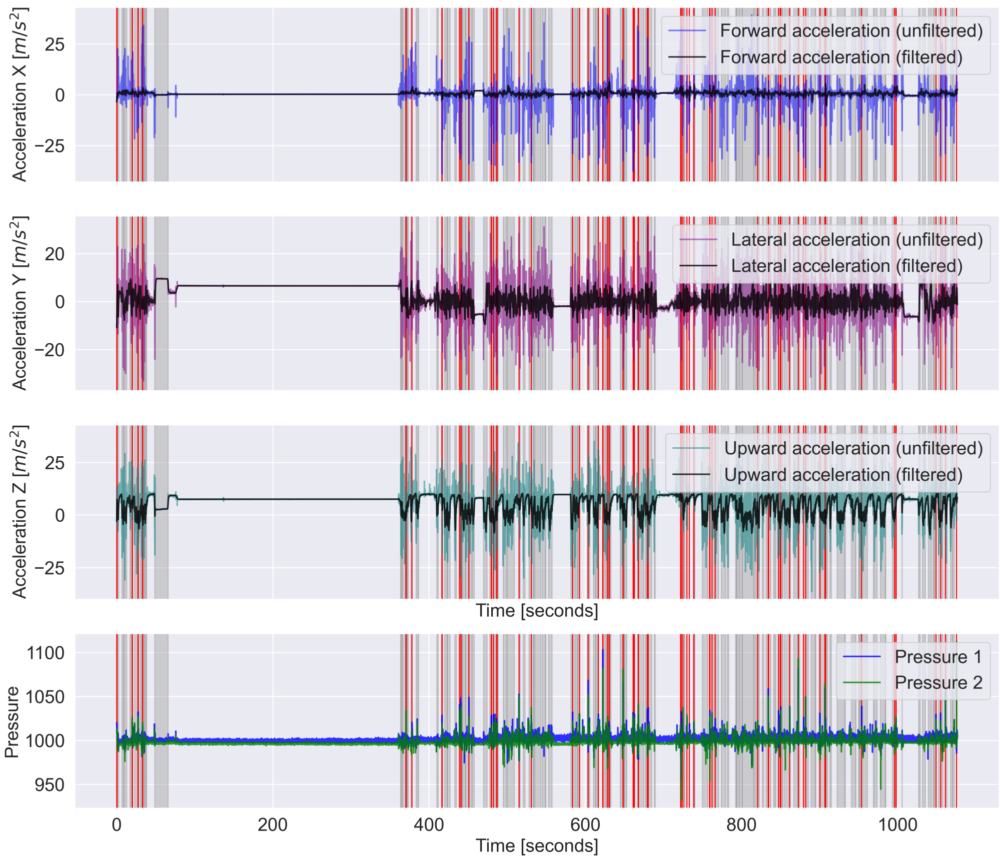

Figure 16. Step-pool sequence in M04 sensor from July 18, 2021

For each sensor and deployment day, the frequency of step-pool sequences was quantified (Fig. 18). The maximum occurrence was recorded in M05 sensor data on 17 July 2021, while no step-pool sequences were detected in M21 sensor data on 13 July. The average number of step-pools increased between July 13 and July 21, 2021, starting from 7.88 on July 13 and peaking at 14 on July 2021. The highest count was detected on July 21. A noticeable increase was also observed on July 17 with 12-step-pool events as compared to lower values observed on July 13 and 18. Different sensors detected different numbers of features. Previous studies [69] documented a seasonal increase in step-pool frequency in supra-glacial channels from mid-to-late July across northern Arctic glaciers.

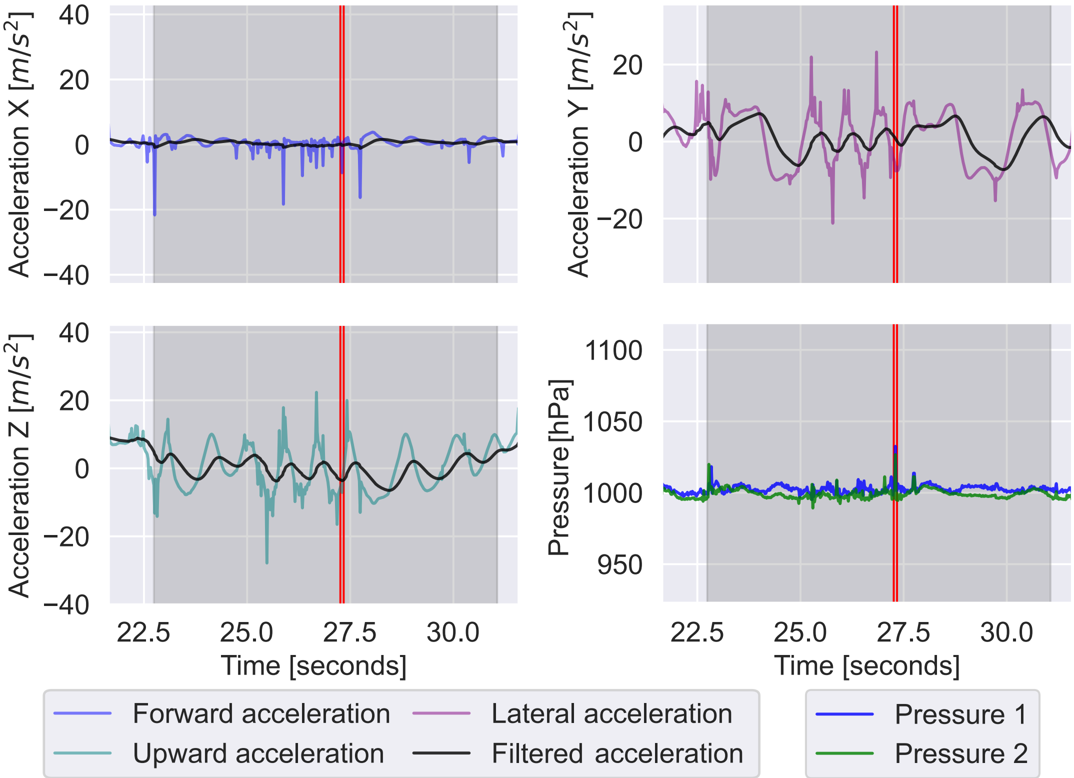

Figure 17. Step-pool sequence identified in acceleration and pressure data (M04 sensor, 18 July 2021). Grey hatched regions denote steps, while red vertical lines mark pools.

4.4 GPS tracking of events

The drifter's step-pool events identified in the earlier sections were overlaid on the GPS data obtained from GNSS to determine the estimated positioning of step-pool features along the drifter's trajectory. The analysis was limited to drifters equipped with GPS loggers (M08, M23, and M24), as only these sensors provided usable data after processing steps. On July 18, 2021, the estimated location of the step pool events from all these sensors was plotted overlaid with the GNSS trajectory (see Fig. 19). The M08 drifter detected 23 step-pool events, distributed along the trajectory, with a noticeable concentration observed in the mid-section of the drifter's path. While M23 and M24 sensors each recorded two step-pool events, both estimated the location either just after the deployment or approaching the retrieval point. This distribution suggests that the water flow conditions varied laterally across the channel. M08 recorded more step pool events, possibly because it always followed the section of the channel with the highest probability of a step pool feature.

In contrast, M23 and M24 detected fewer events, which indicates that the drifter sensor might have transited through the sections where the probability of the step pool was quite

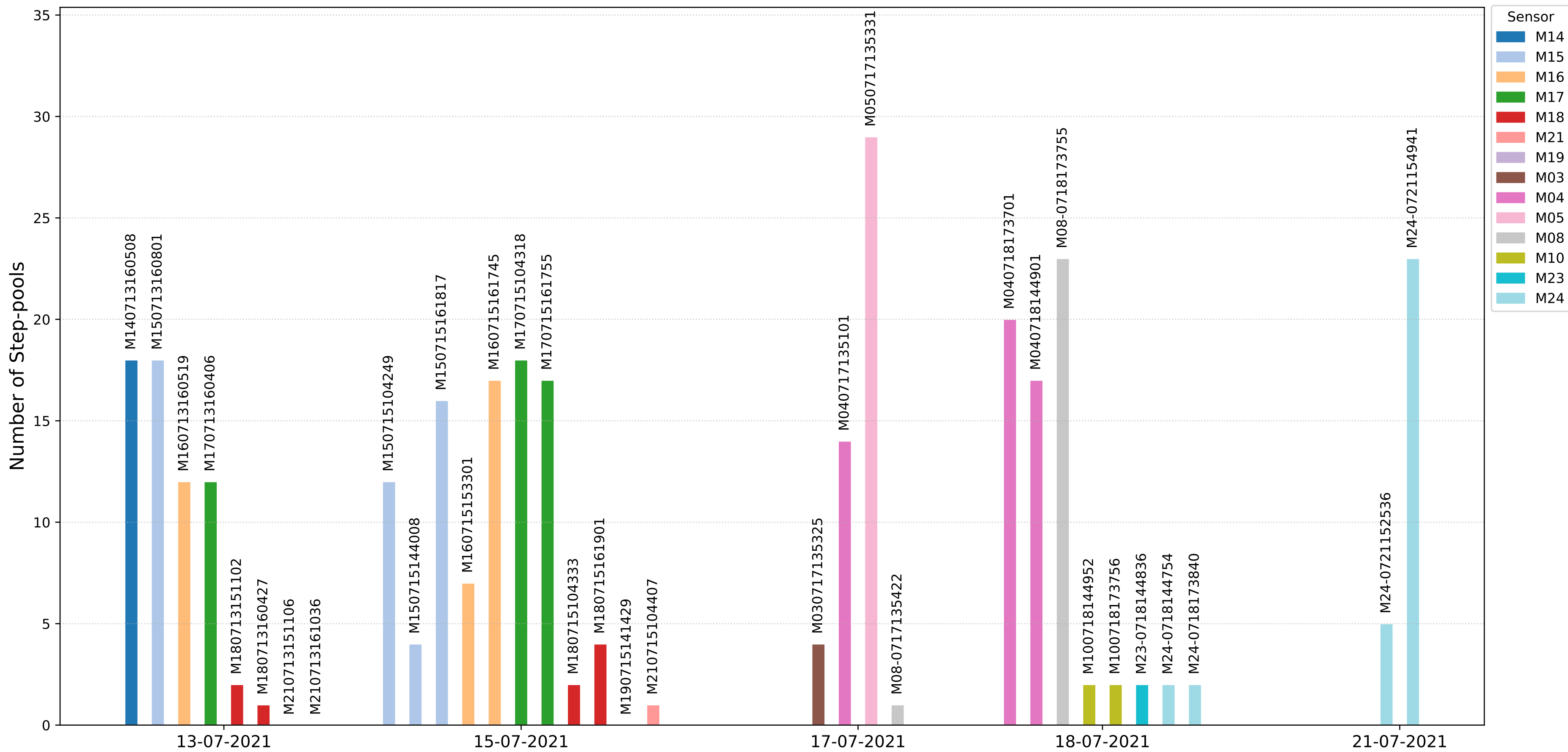

Figure 18. Number of step-pool events observed through each sensor daily.

low. Although the actual physical presence of the step pool feature can not be explicitly validated but the regions within which all drifter sensors identified the step pool event on the same day strongly support its physical presence.

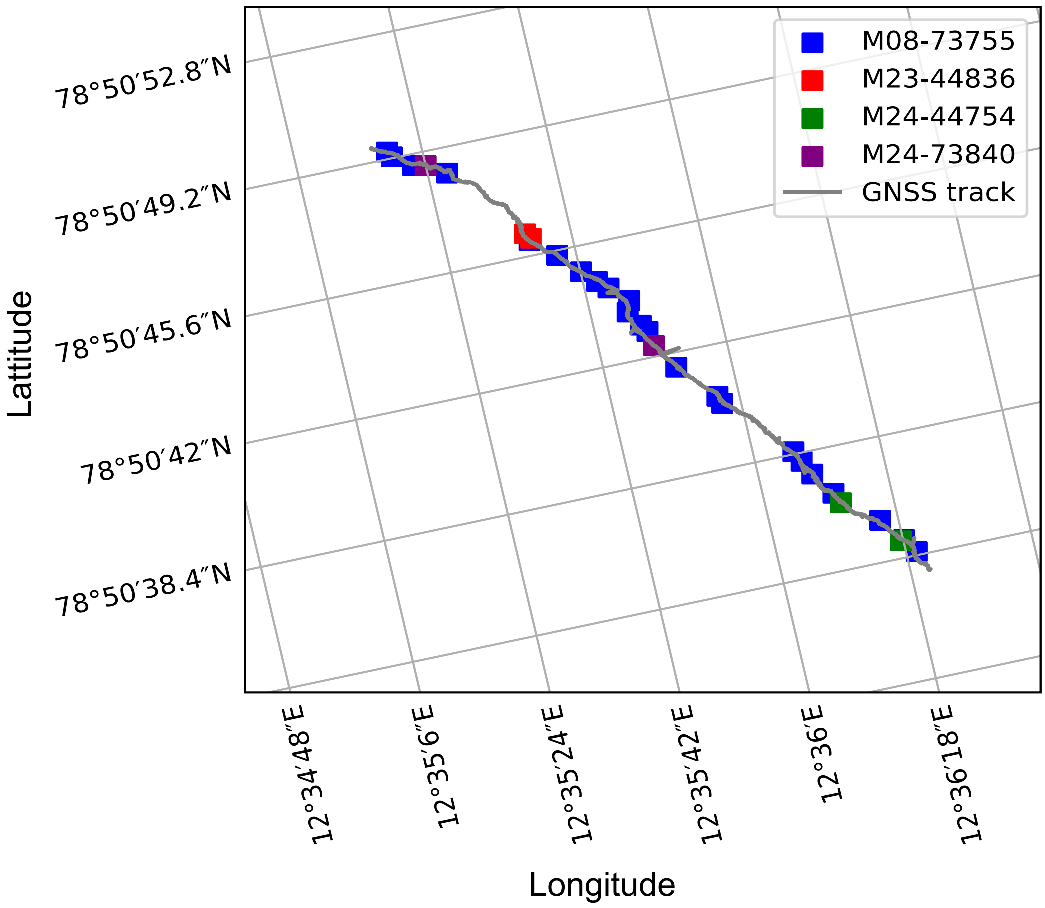

Figure 19. Step-pool events occurring along the trajectory of the drifter. The data in the plot represents the data from M08, M23, and M24 from July 18, 2021.

5. Conclusion and Outlook

In this work, the data obtained from a low-cost sensing drifter were cleaned, preprocessed, and analysed. The data under investigation is from the latest experiments carried out in a supraglacial channel at Kongsvegen (Svalbard). The data was in numeric format, which was first cleaned to exclude the unnecessary data, and later preprocessed using various techniques, resulting in the removal of outliers. All this was achieved using simple statistical techniques, which provided promising results. The preprocessed data was then analysed to identify key features, i.e., drifter sensor stall and step-pools, using the statistical and analytical approaches. This study aimed to investigate two research questions focusing on the methods for assessing the data quality of the drifter sensor and rapidly validating it. The following are the outcomes of this study in answering the anticipated research questions.

RQ1 Which signal processing method offers the most accurate and efficient solution to assess data quality?
This work utilises a common statistical approach, i.e., rolling window analysis, to detect firstly the region of interest (ROI) in the time series data. Besides this, it was also used to identify the features, i.e., sensor stall and step-pool events, by analysing the acceleration and pressure dataset. In both applications, the rolling window analysis provided reasonably good results. This approach was simple and robust as it did not require fine-tuning the parameters; instead, certain thresholds were defined to segregate the frozen or steady signal from the region of interest. Moreover, the rolling window method allows controlling the window size, which can be adjusted based on the temporal resolution of the event. Additionally, the rolling window method Python code was not much complex to understand, highlighting the efficiency and applicability.

RQ2 What is the most effective method to rapidly validate data collected from low-cost sensing drifters?
Ideally, the validation of low-cost sensing drifters could be done using the GPS data gathered through the same IMU. Unfortunately, the GPS logger's data was not available for all drifter sensors, and thus, this study utilised the data from only four sensors to overlay them with the GNSS data. This approach yielded estimated event locations from each sensor, which were validated by the fact that multiple sensors independently detected the event at the same position on the same day, confirming

the accuracy of event detection. Additionally, these events could be physically validated through any auxiliary source, such as video footage. This was a quick approach to validate the data from a low-cost sensing drifter.

Using the statistical approach, it was easier to detect sensor stalls than step pools because the signal pattern associated with stalls was more distinct and less affected by environmental variability. Stall events were identified as flat regions or nearly constant sensor readings over a period, clearly indicating that the sensor was either not fully submerged or positioned in an area with minimal water movement.

In contrast, step-pools were relatively harder to detect because of multivariable dependency, and the changes in the signal representing the features were quite subtle due to the relatively smaller vertical drops in supraglacial channels, which could get camouflaged by environmental factors such as sensor fall orientation, turbulence, and varying hydraulic conditions, etc. The morphology of the supraglacial channel is characterised by a sequence of relatively small step-pools than the englacial or subglacial channels, with vertical drops typically ranging from 20 to 60 cm [70].

The rolling window analysis proved to be quite effective in detecting sensor stalls by identifying periods of minimal change in acceleration, with small pressure variations. It is a simple method for capturing moments of inactivity without the implication of complex machine learning algorithms. However, it relies on the assumption of a completely flat signal, which may not hold in dynamic flow conditions such as those when the drifter holds a position behind a boulder. Therefore, the rolling window analysis was used with certain thresholds, which certainly filter out smaller stalls, for example, less than 3 s. Alternatively, a mixed approach utilising the rolling window analysis and the threshold-based event detection employed to detect the step pool events effectively captures transient changes in acceleration and pressure. It allows flexibility in adjusting window size and thresholds for varying flow conditions. However, it struggles with step-pools of minimal elevation, where subtle changes are easily masked by noise or turbulence and requires manual evaluations. Additionally, careful calibration is needed to avoid missing smaller features or falsely detecting unrelated fluctuations. In summary, the rolling window analysis is effective for detecting sensor stalls, while the mixed rolling window and threshold approach for step-pool detection was more susceptible to challenges with subtle or noisy signals. Overall, both methods enable a rapid feature identification using low-cost drifter sensor data, and can be improved in future for greater accuracy in such dynamic conditions.

5.1 Limitations of the study

During the data processing, the study encountered many limitations. Among the collected data from July 2021, some sensors were missing data or had no data at all. The more data is available from the experiment, the better the results can be deduced. As the study was based on the supraglacial channel data, the feature detection was quite challenging due to the high levels of noise in the data. Moreover, for most of the field measurements, the GPS data was missing, which led to a reduction in the effectiveness of the spatial estimation of features. Although there was video footage of the channel, a drone video of the complete channel from the sensor deployment till retrieval could provide additional assistance in finding the true locations of the events on each day.

5.2 Future outlook

In future research, many improvements can be made to effectively analyse the data from a low-cost sensing drifter. Specifically, to identify the features of a supraglacial channel, advanced machine learning (ML) techniques capable of detecting small patterns in the data can be used. These methods can significantly improve the accuracy of feature identification, especially in challenging scenarios with high signal noise. Also, if possible, the next generation of sensors can be equipped with height sensors which read the sensor height along the track. The height data would enable the validation of specific feature events, such as steps or pools. Moreover, analysing the gyroscope data in conjunction with pressure and acceleration measurements would provide a better understanding of the sensor's motion and position. Implementing these advancements can significantly enhance future research, enabling it to achieve more accurate and efficient feature identification.

Code availability

The scripts used for analysis are publicly available in the following GitHub repository:

Github link: `https://github.com/ridafatimakhan/Kongsvegen_data-analysis_master-thesis`

References

[1] Peter U. Clark, Richard B. Alley, and David Pollard. "Northern Hemisphere Ice-Sheet Influences on Global Climate Change". In: *Science* 286.5442 (1999), pp. 1104–1111. DOI: 10.1126/science.286.5442.1104.

[2] Barbara Barzycka et al. "Changes of glacier facies on Hornsund glaciers (Svalbard) during the decade 2007–2017". In: *Remote Sensing of Environment* 251 (2020), p. 112060. ISSN: 0034-4257. DOI: 10.1016/j.rse.2020.112060.

[3] Florent Gimbert et al. "Subseasonal changes observed in subglacial channel pressure, size, and sediment transport". In: *Geophysical Research Letters* 43.8 (2016), pp. 3786–3794. DOI: doi.org/10.1002/2016GL068337.

[4] Shin Sugiyama and G. Hilmar Gudmundsson. "Short-term variations in glacier flow controlled by subglacial water pressure at Lauteraargletscher, Bernese Alps, Switzerland". In: *Journal of Glaciology* 50.170 (2004), pp. 353–362. DOI: 10.3189/172756504781829846.

[5] Timothy C. Bartholomaus, Robert S. Anderson, and Suzanne P. Anderson. "Response of glacier basal motion to transient water storage". In: *Nature Geoscience* 1.1 (2008), pp. 33–37. DOI: 10.1038/ngeo.2007.52.

[6] A. Alexander et al. "Pressure and inertia sensing drifters for glacial hydrology flow path measurements". In: *The Cryosphere* 14 (2020), pp. 1009–1023. DOI: 10.5194/tc-14-1009-2020.

[7] Nicholas R. Golledge. "Long-term projections of sea-level rise from ice sheets". In: *WIREs Climate Change* 11.2 (2020), p. 634. DOI: 10.1002/wcc.634.

[8] I. Bartholomew, P. Nienow, D. Mair, et al. "Seasonal evolution of subglacial drainage and acceleration in a Greenland outlet glacier". In: *Nature Geoscience* 3 (2010), pp. 408–411. DOI: 10.1038/ngeo863.

[9] C. Nuth et al. "Decadal changes from a multi-temporal glacier inventory of Svalbard". In: *The Cryosphere* 7.5 (2013), pp. 1603–1621. DOI: 10.5194/tc-7-1603-2013.

[10] K. Lindbäck et al. "Subglacial topography, ice thickness, and bathymetry of Kongsfjorden, northwestern Svalbard". In: *Earth System Science Data* 10 (2018), pp. 1769–1781. DOI: 10.5194/essd-10-1769-2018.

[11] Kevin R. Gamache et al. "Chapter 12 - The Impact of Glacial Geomorphology on Critical Zone Processes". In: *Developments in Earth Surface Processes*. Ed. by John R. Giardino and Chris Houser. Vol. 19. Section 12.10.6: Glacier Hydrology. Elsevier, 2015, pp. 363–395. DOI: 10.1016/B978-0-444-63369-9.00012-4.

[12] Richard Alley. "The physics of glaciers. Fourth edition. Amsterdam, etc." In: *Journal of Glaciology* 57.202 (2011), pp. 383–384. DOI: 10.3189/002214311796405906.

[13] Boronina Alina, Popov Sergey, and Qiao Gang. "On the factors and the degree of their effect on subglacial melt and changes in the state of Antarctic subglacial lakes". In: *Arctic, Antarctic, and Alpine Research* 56.1 (2024), p. 2406622. DOI: 10.1080/15230430.2024.2406622.

[14] Frédérique Rémy and Benoit Legrésy. "Subglacial hydrological networks in Antarctica and their impact on ice flow". In: *Annals of Glaciology* 39 (2004), pp. 67–72. DOI: 10.3189/172756404781814401.

[15] J. Weertman. "General theory of water flow at the base of a glacier or ice sheet". In: *Reviews of Geophysics* 10.1 (1972), pp. 287–333. DOI: 10.1029/RG010i001p00287.

[16] Andrew G. Fountain and Joseph S. Walder. "Water flow through temperate glaciers". In: *Reviews of Geophysics* 36.3 (1998), pp. 299–328. DOI: 10.1029/97RG03579.

[17] J.D. Gulley et al. "A cut-and-closure origin for englacial conduits in uncrevassed regions of polythermal glaciers". In: *Journal of Glaciology* 55.189 (2009), pp. 66–80. DOI: 10.3189/002214309788608930.

[18] Shaun D Richardson and John M Reynolds. "An overview of glacial hazards in the Himalayas". In: *Quaternary International* 65-66 (2000), pp. 31–47. DOI: 10.1016/S1040-6182(99)00035-X.

[19] S. Harrison et al. "Climate change and the global pattern of moraine-dammed glacial lake outburst floods". In: *The Cryosphere* 12.4 (2018), pp. 1195–1209. DOI: 10.5194/tc-12-1195-2018.

[20] John J. Clague and Jim E. O'Connor. "Chapter 14 - Glacier-related outburst floods". In: *Snow and Ice-Related Hazards, Risks, and Disasters (Second Edition)*. Ed. by Wilfried Haeberli and Colin Whiteman. Second Edition. Hazards and Disasters Series. Elsevier, 2021, pp. 467–499. DOI: 10.1016/B978-0-12-817129-5.00019-6.

[21] J. Harper et al. "Generation and fate of basal meltwater during winter, western Greenland Ice Sheet". In: *The Cryosphere* 15.12 (2021), pp. 5409–5421. DOI: 10.5194/tc-15-5409-2021.

[22] Bruce H. Raup et al. “Remote sensing of glaciers”. In: *Remote Sensing of the Cryosphere*. John Wiley & Sons, Ltd, 2015. Chap. 7, pp. 123–156. ISBN: 9781118368909. DOI: 10.1002/9781118368909.ch7.

[23] VM Kattsov et al. “Arctic climate impact assessment”. In: *Arctic Climate Impact Assessment* (2005), pp. 99–150.

[24] Richard B. Alley et al. “Ice-Sheet and Sea-Level Changes”. In: *Science* 310.5747 (2005), pp. 456–460. DOI: 10.1126/science.1114613.

[25] Tristram D. L. Irvine-Fynn et al. “Polythermal glacier hydrology: A review”. In: *Reviews of Geophysics* 49.4 (2011). DOI: 10.1029/2010RG000350.

[26] Andreas Aschwanden, Mark Fahnestock, and Martin Truffer. “Complex Greenland outlet glacier flow captured”. In: *Nature Communications* 7 (2016), p. 10524. DOI: 10.1038/ncomms10524.

[27] A. Rabatel et al. “Estimation of Glacier Thickness From Surface Mass Balance and Ice Flow Velocities: A Case Study on Argentière Glacier, France”. In: *Frontiers in Earth Science* 6 (2018), p. 112. DOI: 10.3389/feart.2018.00112.

[28] G. Jouvet et al. “In situ measurements of the ice flow motion at Eqip Sermia Glacier using a remotely controlled unmanned aerial vehicle (UAV)”. In: *Geoscientific Instrumentation, Methods and Data Systems* 9.1 (2020), pp. 1–10. DOI: 10.5194/gi-9-1-2020.

[29] Jonathan L. Bamber and Andres Rivera. “A review of remote sensing methods for glacier mass balance determination”. In: *Global and Planetary Change* 59.1 (2007). Mass Balance of Andean Glaciers, pp. 138–148. ISSN: 0921-8181. DOI: 10.1016/j.gloplacha.2006.11.031.

[30] Tobias Bolch et al. “Planimetric and volumetric glacier changes in the Khumbu Himal, Nepal, since 1962 using Corona, Landsat TM and ASTER data”. In: *Journal of Glaciology* 54.187 (2008), pp. 592–600. DOI: 10.3189/002214308786570782.

[31] Chiyuki Narama et al. “Spatial variability of recent glacier area changes in the Tien Shan Mountains, Central Asia, using Corona (1970), Landsat (2000), and ALOS (2007) satellite data”. In: *Global and Planetary Change* 71.1 (2010), pp. 42–54. ISSN: 0921-8181. DOI: 10.1016/j.gloplacha.2009.08.002.

[32] Donald T Lauer, Stanley A Morain, and Vincent V Salomonson. “The Landsat program: Its origins, evolution, and impacts”. In: *Photogrammetric Engineering and Remote Sensing* 63.7 (1997), pp. 831–838.

[33] Kamlesh Lulla et al. “Mission to earth: LANDSAT 9 will continue to view the world”. In: *Geocarto International* 36 (Oct. 2021), pp. 1–3. DOI: 10.1080/10106049.2021.1991634.

[34] D. J. Quincey et al. “Optical remote sensing techniques in high-mountain environments: application to glacial hazards”. In: *Progress in Physical Geography: Earth and Environment* 29.4 (2005), pp. 475–505. DOI: 10.1191/0309133305pp456ra.

[35] Richard M. Goldstein, Howard A. Zebker, and Charles L. Werner. “Satellite radar interferometry: Two-dimensional phase unwrapping”. In: *Radio Science* 23.4 (1988), pp. 713–720. DOI: 10.1029/RS023i004p00713.

[36] Ian Joughin, Ben E. Smith, and Waleed Abdalati. “Glaciological advances made with interferometric synthetic aperture radar”. In: *Journal of Glaciology* 56.200 (2010), pp. 1026–1042. DOI: 10.3189/002214311796406158.

[37] Markus Allgaier et al. “Direct measurement of optical properties of glacier ice using a photon-counting diffuse LiDAR”. In: *Journal of Glaciology* 68.272 (2022), pp. 1210–1220. DOI: 10.1017/jog.2022.34.

[38] Motoyuki Sato. “Disaster Monitoring by SAR, Gb-SAR and GPR”. In: *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*. 2019, pp. 4752–4755. DOI: 10.1109/IGARSS.2019.8897911.

[39] Maurizio Forte and Stefano Campana, eds. *Digital Methods and Remote Sensing in Archaeology. Archaeology in the Age of Sensing*. Quantitative Methods in the Humanities and Social Sciences. Springer Cham, 2017, pp. XIX, 496. ISBN: 978-3-319-40656-5. DOI: 10.1007/978-3-319-40658-9.

[40] G. Serbin and D. Or. “GPR measurement of crop canopies and soil water dynamics-implications for radar remote sensing”. In: *Proceedings of the Tenth International Conference on Grounds Penetrating Radar, 2004. GPR 2004.* 2004, pp. 497–500.

[41] Eva Savina Malinverni, C. Croci, and F. Sgroi. “GLACIER MONITORING BY REMOTE SENSING AND GIS TECHNIQUES IN OPEN SOURCE ENVIRONMENT”. In: *EARSEL EProceedings* 7.2 (2008), pp. 120–132.

[42] G. E. Flowers. “Hydrology and the future of the Greenland Ice Sheet”. In: *Nature Communications* 9 (2018), p. 2729. DOI: 10.1038/s41467-018-05002-0.

[43] Liam S Taylor et al. “Remote sensing of the mountain cryosphere: Current capabilities and future opportunities for research”. In: *Progress in Physical Geography: Earth and Environment* 45.6 (2021), pp. 931–964. DOI: 10.1177/03091333211023690.

[44] Ian Joughin, Ben E Smith, and Waleed Abdalati. “Glaciological advances made with interferometric synthetic aperture radar”. In: *Journal of Glaciology* 56.200 (2010), pp. 1026–1042. DOI: 10.3189/002214311796406158.

[45] Elizabeth Bagshaw et al. "Novel wireless sensors for in situ measurement of sub-ice hydrologic systems". In: *Annals of Glaciology* 55 (July 2014). DOI: 10.3189/2014AoG65A007.

[46] B. P. Hubbard et al. "Borehole water-level variations and the structure of the subglacial hydrological system of Haut Glacier d'Arolla, Valais, Switzerland". In: *Journal of Glaciology* 41.139 (1995), pp. 572–583. DOI: 10.3189/S0022143000034894.

[47] Wilson Blake et al. "Instruments and Methods: Direct measurement of sliding at the glacier bed". In: *Journal of Glaciology* 40 (1994), pp. 595–599. URL: https://api.semanticscholar.org/CorpusID:140575234.

[48] Engelhardt Hermann and Kamb Barclay. "Basal sliding of Ice Stream B, West Antarctica". In: *Journal of Glaciology* 44.147 (1998), pp. 223–230. DOI: 10.3189/S0022143000002562.

[49] J. Hart et al. "Assessing the catastrophic break-up of Briksdalsbreen, Norway, associated with rapid climate change". In: *Journal of the Geological Society* 168 (Apr. 2011), pp. 673–688. DOI: 10.1144/0016-76492010-024.

[50] Christian Küchler et al. "Lagrangian particle tracking at large Reynolds numbers". In: *Rev. Sci. Instrum.* 95.10 (2024), p. 105110. DOI: 10.1063/5.0211508.

[51] Kyle C. Landon et al. "Bathymetry Estimation Using Drifter-Based Velocity Measurements on the Kootenai River, Idaho". In: *Journal of Atmospheric and Oceanic Technology* 31.2 (2014), pp. 503–514. DOI: 10.1175/JTECH-D-13-00123.1.

[52] J. C. Ohlmann. "GPS-cellular drifter technology for coastal ocean observing systems". In: *Journal of Atmospheric and Oceanic Technology* 22.9 (2005), pp. 1381–1388. DOI: 10.1175/JTECH1786.1.

[53] Annalisa Griffa et al. *Lagrangian analysis and prediction of coastal and ocean dynamics*. Cambridge University Press, 2007.

[54] Alberto Behar et al. "The moulin explorer: A novel instrument to study Greenland ice sheet melt-water flow". In: *IOP Conference Series: Earth and Environmental Science* 6.1 (2009), p. 012020. DOI: 10.1088/1755-1307/6/1/012020.

[55] Elizabeth Bagshaw et al. "E-tracers: Development of a low cost wireless technique for exploring sub-surface hydrological systems". In: *Hydrological Processes* 26 (Sept. 2012), pp. 3157–3160. DOI: 10.1002/hyp.9451.

[56] L. Piho, A. Alexander, and M. Kruusmaa. "Topology and spatial-pressure-distribution reconstruction of an englacial channel". In: *The Cryosphere* 16.9 (2022), pp. 3669–3683. DOI: 10.5194/tc-16-3669-2022.

[57] Juan Francisco Fuentes-Pérez, Francisco Javier Sanz-Ronda, and Jeffrey A. Tuhtan. "An Open Surface Drifter for River Flow Field Characterization". In: *Sensors* 22.24 (2022). DOI: 10.3390/s22249918.

[58] Daniel Boydstun et al. "Drifter Sensor Network for Environmental Monitoring". In: *2015 12th Conference on Computer and Robot Vision*. 2015, pp. 16–22. DOI: 10.1109/CRV.2015.10.

[59] Guillaume Novelli, Cédric M Guigand, and Tamay M Özgökmen. "Technological advances in drifters for oil transport studies". In: *Marine Technology Society Journal* 52.6 (2018), pp. 53–61.

[60] Eric Cocker et al. "Low-Cost, Intelligent Drifter Fleet for Large-Scale, Distributed Ocean Observation". In: *OCEANS 2022, Hampton Roads*. 2022, pp. 1–8. DOI: 10.1109/OCEANS47191.2022.9977209.

[61] Roland Priemer. *Introductory signal processing*. Vol. 6. World scientific, 1991.

[62] Adina E. Racoviteanu et al. "Challenges and recommendations in mapping of glacier parameters from space: results of the 2008 Global Land Ice Measurements from Space (GLIMS) workshop, Boulder, Colorado, USA". In: *Annals of Glaciology* 50.53 (2009), pp. 53–69. DOI: 10.3189/172756410790595804.

[63] Celeste R. Labedz et al. "Seismic Mapping of Subglacial Hydrology Reveals Previously Undetected Pressurization Event". In: *Journal of Geophysical Research: Earth Surface* 127.3 (2022), e2021JF006406. DOI: 10.1029/2021JF006406.

[64] Hock Yew Teh, Anna W Kempa-Liehr, and Kuan-Ing Wang. "Sensor data quality: a systematic review". In: *Journal of Big Data* 7.1 (2020), p. 11. DOI: 10.1186/s40537-020-0285-1.

[65] Sean Kandel et al. "Wrangler: interactive visual specification of data transformation scripts". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. Vancouver, BC, Canada: Association for Computing Machinery, 2011, pp. 3363–3372. ISBN: 9781450302289. DOI: 10.1145/1978942.1979444.

[66] L. Bailey. "The Kalman Filter". In: 2010. URL: https://api.semanticscholar.org/CorpusID:13372487.

[67] Tony Lacey. "Tutorial: the Kalman Filter 11.1 Introduction 11.2 Mean Squared Error". In: *The Kalman Filter*. 1998. URL: https://api.semanticscholar.org/CorpusID:48574040.

[68] Peter S. Maybeck. *Stochastic models, estimation, and control*. Vol. 141. Mathematics in Science and Engineering. 1979.

[69] S. L. ST. Germain and B. J. Moorman. “The development of a pulsating supraglacial stream”. In: *Annals of Glaciology* 57.72 (2016), pp. 31–38. DOI: 10.1017/aog.2016.16.

[70] G. Vatne and T. D. L. Irvine-Fynn. “Morphological dynamics of an englacial channel”. In: *Hydrology and Earth System Sciences* 20.7 (2016), pp. 2947–2964. DOI: 10.5194/hess-20-2947-2016.

Appendix 1 – Non-Exclusive License for Reproduction and Publication of a Graduation Thesis[1]

I Rida Fatima

1. Grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis "Signal processing, feature detection, and data validation of low-cost sensing drifter data", supervised by Laura Piho
   1.1. to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;
   1.2. to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.
2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.
3. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

11.05.2025

[1]The non-exclusive licence is not valid during the validity of access restriction indicated in the student's application for restriction on access to the graduation thesis that has been signed by the school's dean, except in case of the university's right to reproduce the thesis for preservation purposes only. If a graduation thesis is based on the joint creative activity of two or more persons and the co-author(s) has/have not granted, by the set deadline, the student defending his/her graduation thesis consent to reproduce and publish the graduation thesis in compliance with clauses 1.1 and 1.2 of the non-exclusive licence, the non-exclusive license shall not be valid for the period.