Deliverable 1.6: DAS Processing Workflow

DigiMon

Digital monitoring of CO_2 storage projects

Prepared by:
Antony Butcher, Wen Zhou & Alan Baird (University of Bristol)
Boris Boullenger, Bob Paap & Vincent Vandeweijer (TNO)
Tom Hudson & J-Michael Kendall (University of Oxford)
Anna Stork (Silixa Ltd)

DigiMon Deliverable D1.6 Version 1,
September 2021

Revision

Version	Date	Change	Page
1.0	15.09.2021	First version	All

Document distribution

ACT Coordinator

- Research Council of Norway

ACT national funding agencies

- Forschungszentrum Jülich GmbH, Projektträger Jülich, (FZJ/PtJ), Germany.
- Geniki Grammatia Erevnas kai Technologias/The General Secretariat for Research and Technology (GSRT), Greece.
- Ministry of Economic Affairs and Climate/Rijksdienst voor Ondernemend Nederland (RVO), the Netherlands.
- The Research Council of Norway (RCN), Norway.
- Gassnova, Norway.
- Development and Executive Agency for Higher Education, Research, Development and Innovation Funding (UEFISCDI), Romania.
- Department for Business, Energy and Industrial Strategy (BEIS), UK.
- Department of Energy (DoE), USA.

DigiMon partners

- NORCE Norwegian Research Centre AS
- MonViro AS (Ex OCTIO Environmental Monitoring AS)
- NTNU Norwegian University of Science and Technology
- University of Bristol
- University of Oxford
- CRES Centre for Renewable Energy Sources and Saving
- Helmholtz–Centre for Environmental Research
- Sedona Development SRL
- TNO Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek
- Geotomographie GmbH
- LLC Lawrence Livermore National Security
- SILIXA LTD
- EQUINOR ASA
- REPSOL –NORGE AS

Table of contents

1 Introduction 5

2 Microseismics 6

2.1 Event Detection 7
2.1.1 QuakeMigrate 7
2.1.2 RadDetect 10

2.2 Transfer Function 13
2.2.1 Method 13
2.2.2 Application and Results 15

2.3 Source Mechanism Inversion 18

2.4 Shear-Wave Splitting Inversion 22

3 Ambient Noise Interferometry 27

3.1 Characteristics of Ambient Noise Recorded on the Rutford DAS Array 28

3.2 Choice of Virtual Source: DAS versus Geophone 30
3.2.1 Selective Stacking of Transient Noise 31

3.3 Dispersion and 1D velocity structure ***Feil! Bokmerke er ikke definert.***
3.3.1 Dispersion Analysis 32

3.4 Velocity Inversion 33

4 References 36

1 Introduction

This report addresses deliverable D1.6 of the DigiMon project, which covers the processing workflow for datasets acquired by Distributed Acoustic Systems (DAS), and follows on from the DAS Preprocessing workflow report (DigiMon Deliverable 1.4), which captured the key stages required to prepare the raw seismic data for the main processing stages. The workflows are specifically for microseismic and ambient noise interferometry methods, which are both passive seismic methods that seek to image CO_2 movement within a storage reservoir and potential breaches of the reservoir. Active seismic methods, such as seismic reflection and refraction, are o

The objective of the DigiMon project is to develop an early-warning system for Carbon Capture and Storage (CCS) which utilises a broad range of sensor technologies including DAS. While the system is primarily focused on the CCS projects located in shallow offshore environment of the North Sea, it is also intended to be adaptable to onshore settings. Some of the key areas that the systems will monitor include the movement of the plume within the reservoir, well integrity and CO2 leakage into the overburden. A combination of both active and passive seismic methods will be deployed to track the movement of CO2, for example seismic reflection to image seismic velocity changes or microseismics to capture fault activation. Acquiring seismic surveys using DAS is highly novel and offers cost-effective approach which can significantly increase the spatial resolution of the survey data.

Processing workflows presented in this report have been developed during the project using datasets acquired as part of the DigiMon project, with the real-world data described in DigiMon reports D1.1 and D1.2, and synthetic datasets detailed in the DigiMon deliverable D1.3 report. Alongside this report, processing algorithms relating to the workflows are detailed in DigiMon deliverable D1.5 report.

2 Microseismics

Seismic activity in and around CCS reservoirs can provide valuable information on the movement of CO_2 and the characteristics of a storage reservoir. Any subsurface activity that alters the state of stress in the ground is capable of triggering seismic activity on pre-existing faults, and in the case of CCS, this can be caused by the movement of CO_2 within the subsurface. Microseismic monitoring has become an established method of imaging the subsurface for industrial applications, such as enhanced geothermal systems (e.g. Kwiatek et al., 2019), hydraulic stimulation of "tight" hydrocarbon reservoirs (e.g. Kettlety et al., 2020), and has been successfully deployed in several CCS projects (e.g. Stork et al., 2015). The technology provides an important tool for monitoring CO_2 plumes and potential breaches from the containing reservoir.

The primary goal of microseismic monitoring is to detect, locate and characterise very small earthquakes (Figure 1), which occur at or below the micro-scale of seismicity (M<2). These typically relate to fault ruptures which are 10s of meters in length that produce displacement amplitudes at the detectability limits of seismometers (Bohnhoff et al., 2009). The position of these events within a storage reservoir provides valuable information on the extent of the CO2 plume, while the waveforms can be used to characterise the fracture properties and stress regime within the reservoir. The resolution of these datasets can be significantly improved by lowering the detectability limits of the microseismic array. This limit is not only a function of the instrument but is also controlled by the levels of ambient seismic noise. One approach to improving the signal to noise ratio (SNR) is through applying array processing methods (e.g. Verdon & Budge, 2018) which require relatively highly spatially sampled datasets.

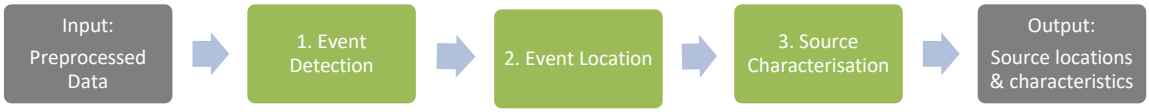

Figure 1: *Microseismic processing workflow.*

The adoption of DAS for microseismic monitoring is at a relatively early stage. Although the instrumentation is well developed and increasingly deployed for commercial projects, there does exist several technical challenges still to overcome. Potentially DAS can significantly improve the resolution of microseismic datasets due to high temporal and dense spatial sampling. However, the single component nature of the measurement and challenges in correcting the instrument response has several limitations that still need to be resolved. The single-component nature of the measurement not only results in sited and orientation effects due to the limited broadside sensitivity of DAS along straight fibres (Baird et al., 2020) but also prevents the use of standard seismic anisotropy methods for fracture analysis (Teanby et al., 2004).

2.1 Event Detection and Location

In this section we use the naturally occurring microseismicity recorded within the Rutford Ice Stream, Antarctica, to assess different methods for locating events recorded on DAS surface arrays. This dataset is described in more detail in DigiMon Deliverable 1.1 (DAS field dataset to compare technologies and deployment scenarios – Antarctica Dataset) and Hudson et al. (2021). Amplitude-based methods, such as a Short-Term-Average to Long-Term-Average amplitude ratio (STA/LTA) algorithm, are commonly used to identify phase arrivals but they required high SNR to be effective. This is not always the case with DAS measurements and instead we are focussing on two different processing approaches which exploit the spatial resolution of the DAS recording: a waveform-based migration method called QuakeMigrate and an approach based on slant-stack transforms named RadDetect. Both approaches identify and pick microseismic events and then locate these events using the non-linear earthquake relocation software, NonLinLoc (Lomax & Virieux, 2000).

2.1.1 QuakeMigrate

QuakeMigrate approximates the seismic energy associated with a seismic phase arrival at a particular receiver as an onset function, which in our case is defined as a continuous STA/LTA function through time. Onset functions for each receiver are stacked and back-migrated through time and space, in order to search for a coalescence of seismic energy corresponding to an event. One particular strength of this method is that although incoherent noise is back-migrated, it will not coalesce, therefore reducing the possibility of false detections. Furthermore, the initial back migration does not require phase picks to be made, which is often challenging given the SNR of DAS data. Phase-picking is only attempted for data where a peak in coalescence is detected. Another advantage of the method is that data from multiple instrument types can be combined once the waveform observations are approximated by onset functions, allowing for DAS and geophone time series data to potentially be used in combination in the detection and location algorithm. A further advantage of the QuakeMigrate algorithm is that it is open source, therefore allowing others to apply the methods demonstrated in this paper for other studies involving DAS, and hopefully improve upon our methods in the future.

Once the events have been detected and initially located, we refine the event locations using the non-linear earthquake relocation software, NonLinLoc. We refine event locations using NonLinLoc because it provides quantification of the statistical spatial and temporal uncertainty of the icequakes, allowing us to quantify the performance of DAS only versus geophone only versus combined network detection and location.

An important question for the applicability of DAS for studying natural microseismicity is whether one can detect and locate seismicity using DAS alone. To address this question, we compare results of icequakes detected using DAS and geophones separately.

Unlike the three-component geophone records, P-wave arrivals are not visible in our DAS data because only the horizonal component of strain is recorded along the fibre. A near-surface firn layer, of substantially lower seismic velocity, refracts P-waves towards vertical incidence. Surface DAS recordings

in areas without a firn layer would yield P-wave observations. For icequake detection and location, we assume an isotropic fabric, as QuakeMigrate can only pick a single S-wave arrival, typically the fast arrival. Figure 2 and Figure 3 show the icequake detection results. Only the hypocentral distance and azimuth are resolved adequately for our DAS geometries. The depths of the DAS-only results are therefore artificially constrained prior to detection and location to between 1700 and 2100 m bsl, based on previous Rutford icequake observations. The geophone-only depths are not artificially constrained. The lateral spatial clustering observed in Figure 2 for all network configurations, interpreted to be sticky patches of the bed, is expected from icequake datasets. It shows that both the DAS line and triangle geometries are able to resolve spatial clustering in the data, if depths are independently constrained. The addition of a vertical section of fibre could also provide depth constraint in the vertical plane, in the same way that the triangle array breaks the symmetry, therefore providing better epicentral constraint in the horizontal plane.

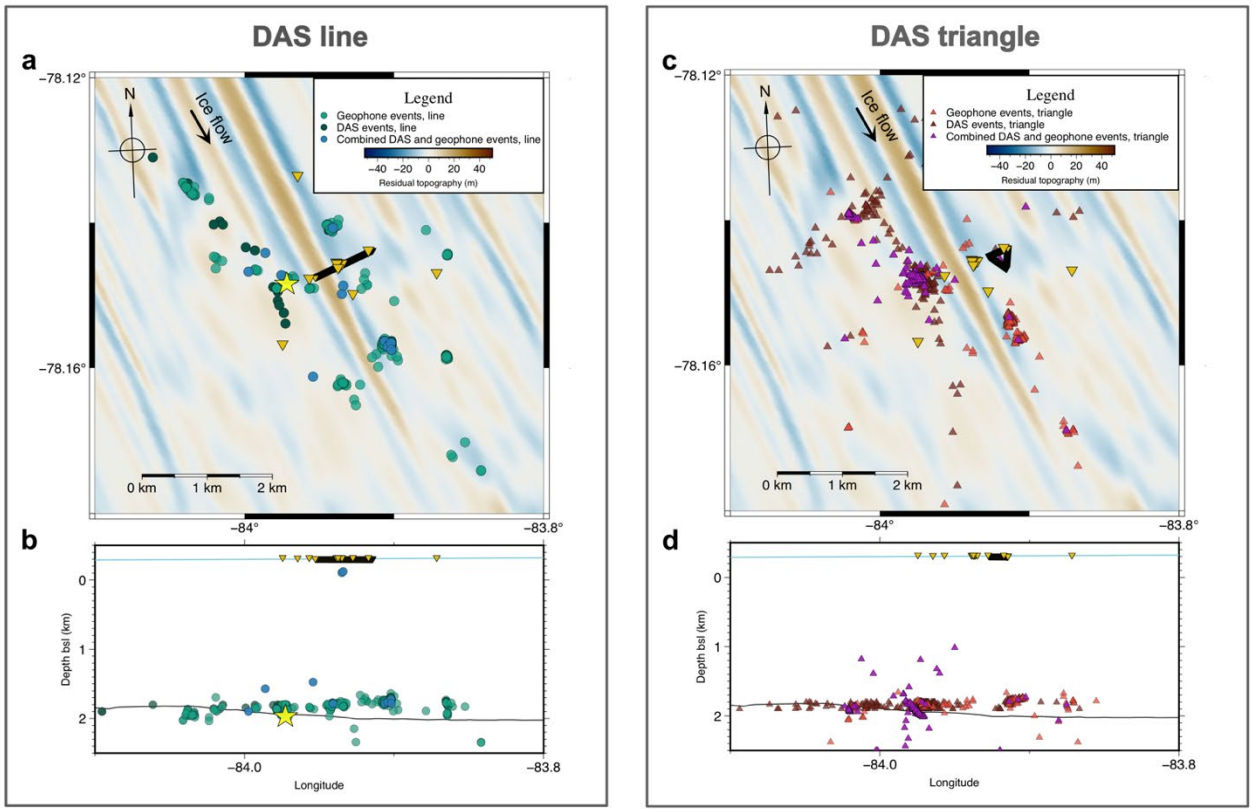

Figure 2. *Detected icequake locations using DAS and geophones independently and together. A) Icequake locations during the DAS line deployment, colored by instruments used for the detection (geophones or DAS). Inverted triangles indicate receivers, with the DAS fiber shown by the SSW to NEE line and triangle near the center of the figure. The yellow star corresponds to the example event shown in Figure 6. B) Same as (a), but with the events plotted with depth vs. longitude. C) Same as (a), but for the DAS triangle deployment. D) Same as (b) but for the DAS triangle deployment. The residual bed topography is calculated based on the difference between the short-wavelength bed topography and the long-wavelength bed topography from King et al. (2016).*

A second observation is that significantly fewer events are detected in data from the DAS-only configuration compared to the geophone-only configuration. Only 499 events are detected using the DAS triangle, compared to 1321 geophone event detections, and only 139 events are detected using the DAS

line, compared to 1270 geophone event detections (see Figure 3a,b,c,d). The line is therefore a less effective configuration than the triangle for event detection, likely partly due to the two-dimensional nature of the triangle configuration resulting in a higher peak coalescence of energy at a single location, rather than being split by the geometric ambiguity of the line, but also because it may record more S-wave energy due to its effectively quasi two-component measurement. The addition of a vertical section of fibre could also provide depth constraint in the vertical plane, in the same way that the triangle array breaks the symmetry, therefore providing better epicentral constraint in the horizontal plane. Such a vertical section would also have reduced surface wave noise compared to a horizontal section. A spatial detection bias is also apparent in the data, with icequakes from clusters beyond the SW end of the linear DAS configuration detected, but icequakes at closer offsets but perpendicular to the cable not detected. This sensitivity is likely a result of the single component nature of the DAS making it insensitive to certain S-phase polarizations, for the basal icequake radiation pattern orientations observed at Rutford Ice Stream.

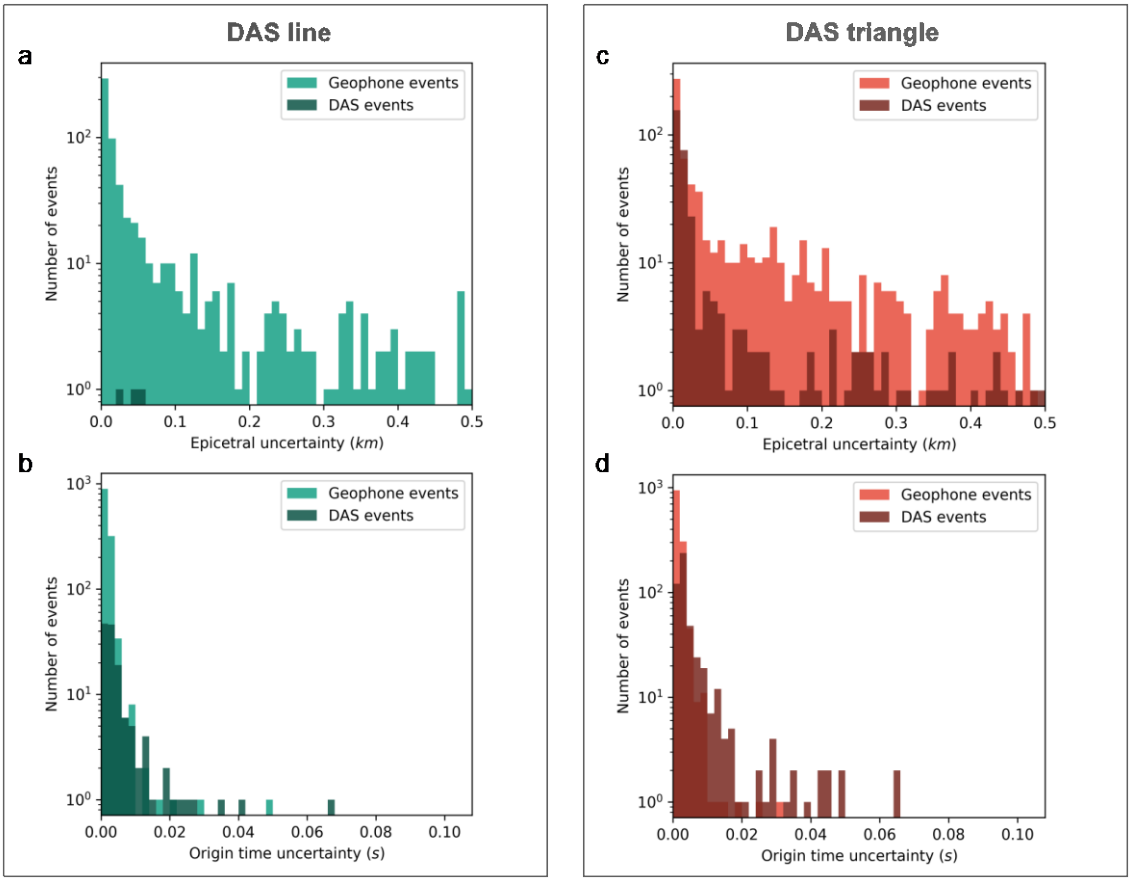

Figure 3. *Uncertainties associated with the icequakes detections presented in Figure 2. A) Histogram of epicentral uncertainty associated with the icequakes, for the linear DAS fiber deployment. B) Histogram of origin time uncertainty associated with the icequakes, for the linear DAS fiber deployment. C,d) Same as (a), (b), respectively, except for the triangle fiber deployment. Note that epicentral uncertainties are clipped at an upper limit of 0.5 km for clarity. There are events with uncertainties greater than this but these are not plotted.*

We also compare the spatial and temporal uncertainty of detected events. Figure 3a and Figure 3c show the epicentral uncertainty of events detected for the line and the triangle, respectively. We compare

epicentral uncertainty, rather than hypocentral uncertainty, since the depths of events for the DAS-only detection are artificially constrained. While for some events the line and the triangle DAS-only epicentral uncertainties are similar to geophone-only detection uncertainties, there are a number of DAS-only events that have spatial and temporal uncertainties greater than the plotted range (i.e. epicentral uncertainty > 0.5 km and origin-time uncertainty > 0.1 s). In summary, the geophones provide better spatial constraint of the icequake epicentres and can also constrain icequake depth, while both the DAS and geophones appear to provide similar constraint on event origin times. There are insufficient events detected by the line to quantify whether the triangle or line has better spatial constraint. Figure 3b and Figure 3d show the origin time uncertainty of the events. For both DAS geometries, the geophone-only and DAS-only data suggest similar constraint on the origin time uncertainty. Although the epicentral uncertainties using DAS presented here are large, DAS could still be particularly useful for monitoring applications where rates of seismicity, rather than accurate locations, are of primary importance.

These results suggest that applying QuakeMigrate to DAS data has limitations for microseismic detection. Firstly, for this experiment arrangement with sources ~2 km below surface, geophones are significantly better than DAS for detecting microseismicity. This is because: (1) the spatial extent of the geophone network is much greater than the DAS (see Figure 1), a limitation specific to our DAS deployment that could be overcome by deploying more fibre; (2) the geophones measure three components of ground motion, and so are sensitive to P-, SV- and SH-phase arrivals; and (3) the geophones have a much higher SNR than single DAS channels. Although the firn velocity structure causes the DAS to be sensitive to only S-phases in this study, horizontally deployed DAS on ice without firn would be sensitive to both P- and S-phases, providing better event depth constraint. A second limitation is the complexity of combining DAS and geophone data together for detection and location. Theoretically, DAS and geophone data could be combined to reduce spatial and temporal uncertainty. However, there could be a trade-off between the gain of additional observations and detrimental additional noise. Weighting the combined data to mitigate for this is complex, and likely site- and network- geometry specific. We therefore do not investigate weighting methods in this study, instead suggesting this as an area for future work. Ideally one would also convert the time-series for each type of instrument into mutually consistent units, so as to avoid any arrival-time uncertainty due to possible phase shifts caused by the difference between a strain-rate and velocity measurement. Overall, the poor performance of our deployment for detection and location is likely dominated by the spatial extent of the DAS compared to the geophones. We suggest that if the horizontal spatial extent of the DAS is comparable to or greater the spatial extent than the depth of seismicity, then better performance may be achieved, such as in Walter et al. (2020).

2.1.2 RadDetect

Seismic arrays deployed using DAS generally suffer from a poorer SNR than those using conventional seismometers or geophones, however their high spatial resolution provides opportunities to supress noise and enhance coherent signals. We seek to boost the detection limits of the Rutford DAS dataset though employing a radon transform-based detection method, which can provide a robust and automatic approach to detected microseismic events (Sabbione et al., 2015). Radon transforms are a slant-stack method, which are regularly used to increase the SNR in seismic shot-gathers during reflection processing.

The transform converts the data into the t-*p* domain, where if the arrivals are well sampled, they can be easily identified based on their apparent slowness.

Seismic data is initially pre-processed using a 2-120Hz bandpass filter followed by an f-k filter, which removes signals outside the range of interest and improves the coherence of the arrivals. We then apply a linear radon transform to a 200m-wide sliding window that progresses along the array in 50m steps. Coherent arrivals are then automatically picked in the t-*p* domain, which provides both a pick time and the apparent slowness of the arrival (Figure 4). These parameters are then used in a coincidence filter to cluster common arrivals and remove spurious picks.

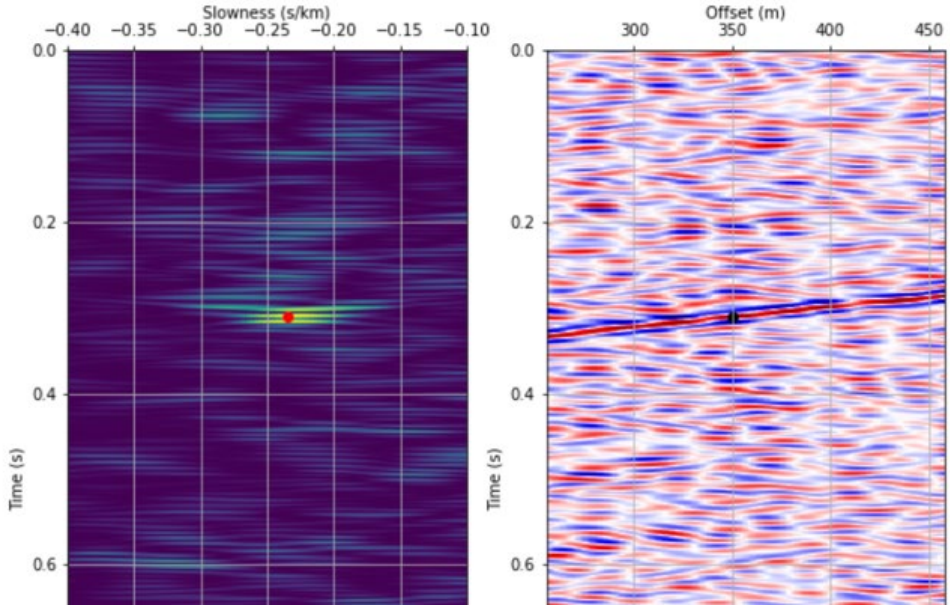

Figure 4. *Event detection in the t-p domain which providing an arrival time and apparent slowness.*

For the linear array this approach increases the detection rate to 2192 events for a 6hr period, which is significantly more than both the geophone array (1270 events) and QuakeMigrate (139 events). The triangular array detects 1234 events, which is slightly lower than the geophone network (1321 events) but still an improvement on QuakeMigrate (499 events). Example arrivals recorded on the linear and triangular array are shown in Figure 5. Only S-wave arrivals appear to have been recorded by the Rutford DAS array, which is due to the presence of a low velocity firn layer in the near-surface. This layer acts to refract the P-waves towards near vertical incidence making the fibre, which is only sensitive to the horizontal component of strain.

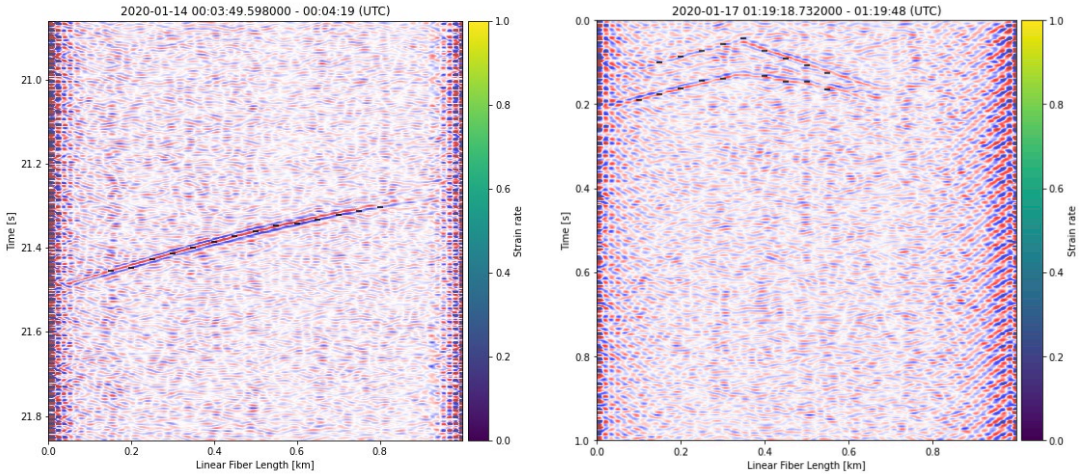

Figure 5. *Example arrivals for the linear (left) and triangular (right) arrays detected using a radon transform-based detection method. Picks indicated by black horizontal lines.*

For each S-wave arrival we invert for event locations using an Eikonal solver (Lomax et al., 2012) constraining their locations to the base of the glacier (Figure 6). For both the linear and triangular array this produces clusters of seismicity which are broadly comparable to the event locations identified by Hudson et al. (2021). Due to the one-dimensional nature of the linear array, there are clear location ambiguities which produces semi-circles of events off the ends of the array. In addition, two 'mirrored' clusters of events are observed at the western end of the fibre, which are likely due to the inherent 180° ambiguity resulting from a linear array.

The triangular array has detected fewer events compared to the linear array, but the multi-dimensional array configuration has removed much of the positional ambiguity. An example of this is to the west of the array, where a diffuse cluster of events located by the linear array has become more focused in the triangular array dataset. There is a, however, a much broader scatter of events around the triangular array compared to the linear array, and this is likely a result of the smaller spatial extent of the array.

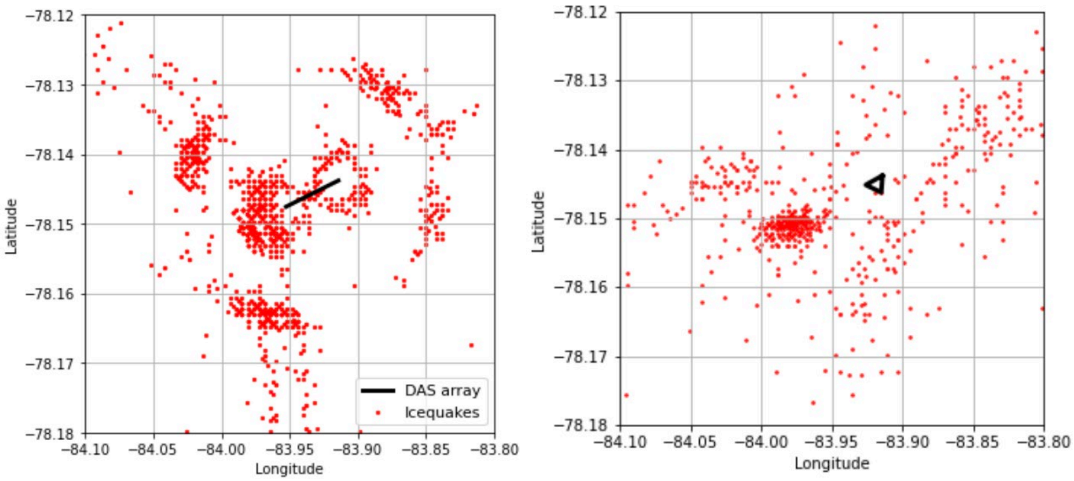

Figure 6. Events located using a 1km long linear (left) and 350m sided triangular (right) array.

2.2 Source Characterisation

2.2.1 Transfer Function

DAS measurements are measurements of optical phase changes which are then straightforwardly translated into the dynamical mechanical strain (or strain rate) of a fiber gauge length in the axial direction of the fiber (Lindsey et al., 2020). The resulting strain or strain rate measurements are typically expressed in nm/m or nm/m per second, respectively. However, commonly used algorithms for estimating an earthquakes magnitude and source parameters require ground displacement or velocity measurements (expressed in m and m/s respectively) as input data. There is, therefore, a requirement to determine a transfer function that converts strain-based DAS measurements into ground motion units so that standard seismological methods, such as magnitude scales, can be applied to DAS datasets.

In this section, we report a method to achieve the conversion between the corresponding physical quantities and demonstrate it using collocated seismic DAS and geophone records from the FORGE site.

2.2.1.1 Method

The measured strain is related to velocity through the apparent phase velocity C_x (Aki and Richards, 2002) which, when expressed as the ratio of frequency and wavenumber ($C_x = \frac{\omega}{k}$), leads in the frequency domain to :

$V_{DAS}(\omega, k) = -\frac{\omega}{k} E_{xx}(\omega, k)$,

where E_{xx} represents the DAS strain measurements in the axial direction of the fiber and V_{DAS} represents the derived ground particle velocity as indirectly measured by the fiber along its axial direction.

In practice, the time-domain velocity $v_{DAS}\,(t, x)$ can be obtained from $E_{xx}(\omega, k)$ by performing the following steps:

1) transforming the time-domain strain measurements into the frequency-wavenumber (F-K) domain by applying twice a Fourier transform.
2) rescaling the values by the factor $-\frac{\omega+n}{k+n}$ where n is an additional term to limit undesired numerical instabilities when k tends to and reaches 0.
3) applying the inverse Fourier transform twice.

This F-K rescaling procedure which allows to obtain DAS-derived ground velocity measurements has been first proposed and described by Lindsey et al. (2020). In this practical approach, the parameter n is determined relatively to the peak ground velocity value as measured by a collocated conventional seismometer.

The following two remarks should be taken into account about the intended result (v_{DAS}) of this proposed F-K rescaling method:

- It corresponds to a ground velocity as measured by the fiber which is unlikely to be perfectly equal to the ground velocity as recorded by a collocated geophone device, due to the difference of nature and ground coupling between the two instruments.
- It includes the inner DAS instrument response which may not be neglectable depending on the considered seismic frequency bandwidth.
-

As shown in Lindsey et al. (2020), the DAS instrument response can be quantified by deconvolving DAS-based velocity records with ground velocity records from collocated geophones (records after geophone instrument response removal). This may offer the possibility to obtain even more accurate ground velocity records from DAS measurements which in turn will result in more accurate micro-seismic characterization.

2.2.1.2 DAS data

The method is applied on seismic records from the FORGE site in Utah, USA. FORGE is a field laboratory for developing and testing technologies for enhanced geothermal systems. As induced seismicity is one of the key research themes, seismic monitoring data has been recorded by various installations on the site, including a monitoring well equipped with both a DAS fiber and a geophone string for performing micro-seismic analysis during stimulation campaigns.

This monitoring well with identification code 78-32 is nearly 1000 meters long. The geophone string includes 12 three-components sensors spaced by 30.5 meters and is lowered towards the bottom part of the well whereas the engineered optical fiber is cemented along the full depth range of the well. The DAS interrogator is based on the Carina system developed by the firm Silixa (as for the fiber itself) and is set to produce measurements with a channel spacing of 1 m using a gauge length of 10 meters. The publicly available data are expressed as strain rate measurements in nm/m per second which means that these are already converted from the raw optical-phase change measurements.

To demonstrate the method, we focused on the records of seismic waves generated by a perforation shot. As the source location of such controlled seismic event is known, the analysis and interpretation of the recorded signals are associated with less uncertainties than in the case of microseismic events. In this case, the source is located within a stimulation well at a horizontal distance of approximately 370 meters and at a depth of 1200 meters with respect to the bottom of the monitoring well 78-32. Therefore, the generated seismic waves are recorded at the deepest DAS channels and geophones with an incidence angle of approximately 20 degrees.

The DAS strain rate and geophone velocity records of the perforation shot are shown in Figure 7 in the depth range beyond 600 m where all geophones lie. The processing steps to obtain the plotted DAS strain rate records include: applying a bandpass filter in the range 10-250 Hz, detrending (linear trend and mean removal) and removing of the zero wavenumber noise by tapering in the frequency wavenumber domain. The geophone velocity records are obtained from the original acceleration records using the following processing sequence: detrending (linear trend and mean removal), applying a bandpass filter in the range 10-250 Hz and integrating.

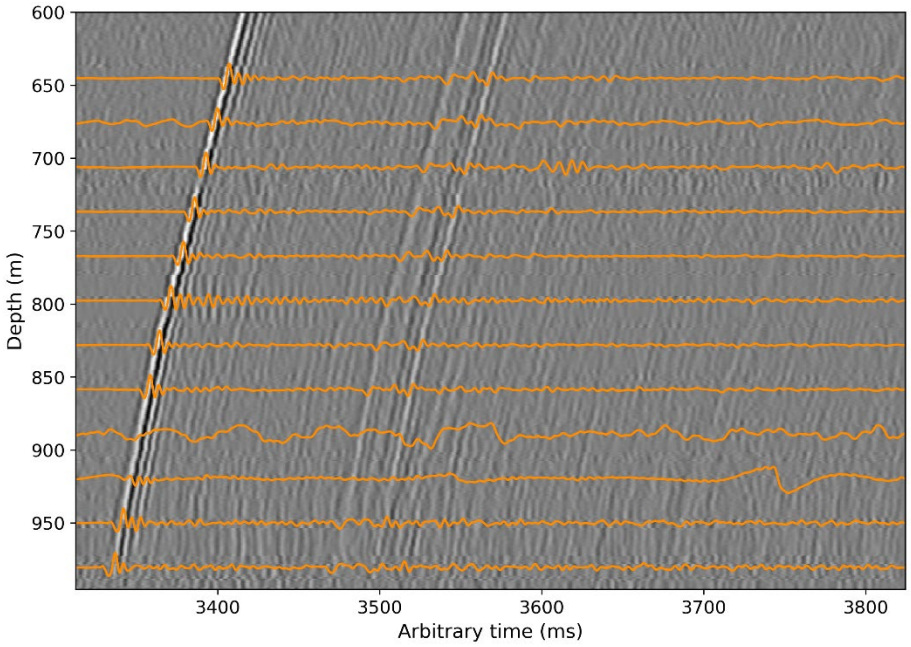

Figure 7. *DAS and geophone perforation shot records: pre-processed DAS strain rate measurements in the depth range beyond 600 meters (grayscale background mesh), with superimposed pre-processed vertical component of the particle velocity records from the 12 downhole geophones (in orange). For display purposes, the individual DAS and geophone traces are normalized by their root-mean-square and maximum values, respectively.*

2.2.1.3 Application and Results

Prior to implementing the conversion method, the traces from a group of 151 adjacent channels are selected to be transformed into the F-K domain and converted from strain rate to strain measurements unit with integration. The group is centered around the channel number 842 which is collocated with the 8th geophone (starting from top) at a depth of approximately 859 m.

Figure 8 shows the result of transforming these 151 strain records in the F-K domain and the effect of the rescaling method in that domain. The stabilization parameter is set as the peak ground velocity observed on the geophone trace multiplied by a factor of 1000. To further prevent numerical noise after back transformation in the time domain, the rescaled data is also tapered at and around the zero wavenumber ($k = 0$).

The transformation of the rescaled and tapered data (bottom row in Figure 8) back to the time-domain results in 151 retrieved DAS-based velocity records. Figure 9 shows the retrieved DAS-based velocity record of the central channel (in black) as well as the collocated geophone-based velocity record (in orange) and the original DAS strain record (in violet) for comparisons. Note that the plots in Figure 9 do not involve any normalization of the traces. We observe that the retrieved DAS velocity record presents a very good match with the geophone ground velocity, thereby demonstrating a significant accuracy of the proposed method and its effectiveness in controlling the level of noise. This example also shows that

relevant features such as the peak ground velocity present some differences. If we assume that the geophone record is the most trustworthy, this means that the DAS velocity record could be further improved by quantifying more precisely the DAS instrument response. This latter would then allow through deconvolution to access the true ground velocity corresponding to any seismic event using only DAS measurements.

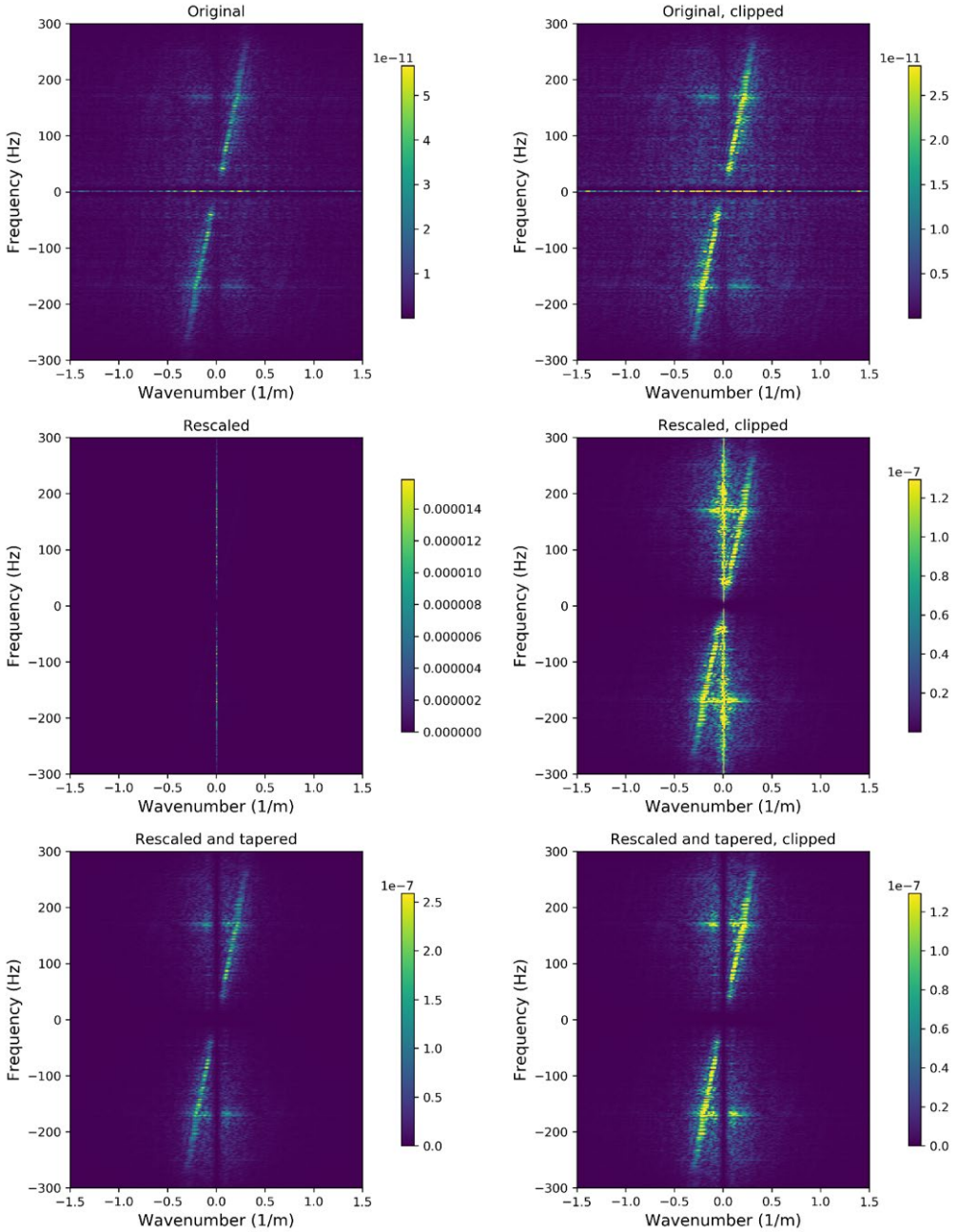

Figure 8. *DAS data rescaling in the F-K domain: amplitude spectrum in the F-K domain of the selected DAS strain records (top raw), after applying the rescaling formula (middle row) and after additional tapering around k=0 (bottom row). The right plot is a repetition of the left plot where values above 50% of the maximum value are*

clipped; for both rescaled results (middle right and bottom right plots), the maximum value is taken from the result after tapering (bottom left plot).

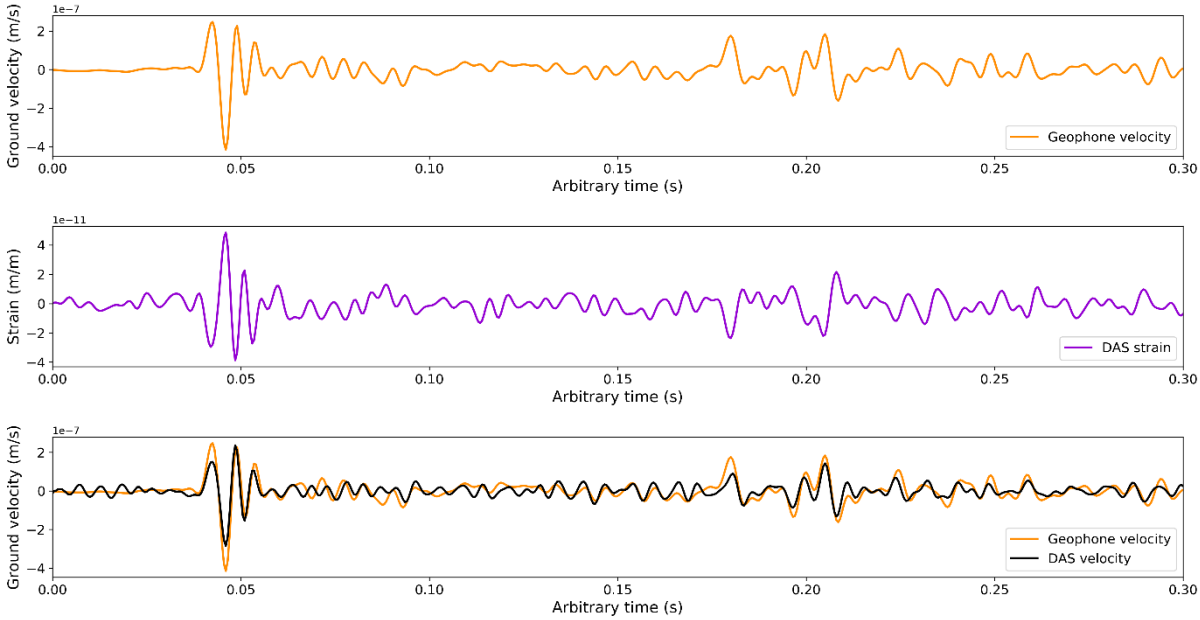

Figure 9. *Application results of the F-K rescaling method: the geophone-based ground velocity record (in orange) is a collocated reference trace, the original DAS strain record (in violet) is obtained after integration of the pre-processed strain rate measurements and converted through the proposed method to the retrieved DAS-based ground velocity (in black).*

2.2.2 Source Mechanisms

An earthquake source mechanism inversion involves using the observed radiation pattern of an earthquake to constrain the physical elastic failure mechanism that generates the earthquake. The earthquake source mechanism inversion presented here is based upon the method described in Hudson et al. (2021) and available as the open-source package SeisSrcInv (Hudson, 2020a). We apply this method to the Antarctica dataset, which is described in more detail in DigiMon Deliverable 1.1 (DAS field dataset to compare technologies and deployment scenarios – Antarctica Dataset).

The source inversion method is a full-waveform Bayesian source mechanism inversion, randomly sampling the model space millions of times in order to obtain an estimate of the posterior probability distribution. We constrain the source model to be a Double-Couple (DC) model, which is appropriate for the predominantly stick-slip seismicity observed at Rutford Ice Stream. The DAS source inversion workflow is as follows:

1. **Spatially downsample the DAS data:** Due to the computational expense associated with calculating Green's functions and performing the source inversion, we only use every 10th channel, approximately equal to the gauge length (10 m), along the DAS cable to approximately sample the wavefield. This is found to be a sufficient resolution of spatial sampling for our case. Alternatively one could improve the SNR by stacking all these channels together, if they could be correctly aligned.
2. **Filter the DAS data:** We first apply a rectangular fk-filter passing wavenumbers, k, of -0.04 to $0.04\ m^{-1}$ and frequencies between -150 to $150\ Hz$. To remove surface wave noise from a generator, we also apply notch filters centered at 33 Hz and 66 Hz, with a bandwidth of 2.5 Hz. Alternatively, a more sophisticated fk filter informed by apparent surface wave velocities would likely be capable of removing the generator noise.
3. **Generate synthetic modelled waveforms:** We generate synthetic modelled waveforms, for comparison to the observed data, for each moment tensor component ($m_{xx}, m_{yy}, m_{zz}, m_{xy}, m_{xz}, m_{yz}$), for each DAS channel, for an event. These are calculated using the program fk, which uses the Thompson-Haskell propagator matrix method (Zhu & Rivera, 2002). They are modelled for an isotropic, homogeneous ice medium overlaid with a 100 m firn layer (Smith, 1997) of decreasing velocity towards the surface.
4. **Account for anisotropy:** Rutford Ice Stream has a strong elastically anisotropic fabric (Harland et al., 2013; Smith et al., 2017), causing Shear Wave Splitting (SWS) that has to be accounted for. This can be done either by applying a linearization and time shift correction to the observed data, or by simulating the effect of SWS on the isotropic modelled S wave phases to produce anisotropic synthetic waveforms. Correcting for such effects is valid as long as the anisotropy is a path effect, and the source region can be assumed to be isotropic. We implement the latter method, approximating the effect of SWS on the synthetic DAS data by applying an average anisotropic splitting angle and a fast-slow S-wave delay time. We assume that the S-waves arrive at approximately normal incidence (vertical) to the surface, due to the firn velocity structure. We can

then approximate the fast and slow S-wave arrivals in the North and East axes from the LQT coordinate system of the synthetics using the equations,

$$N_{fast} = -Q_{model}\cos(\theta)\cos(\phi) \quad ,$$
$$N_{slow} = Q_{model}\cos(\theta)\sin(\phi) \quad ,$$
$$E_{fast} = T_{model}\cos(\theta)\cos(\phi) \quad ,$$
$$E_{slow} = -T_{model}\cos(\theta)\sin(\phi) \quad ,$$

where θ is the azimuthal angle from the source to the receiver and ϕ is the average anisotropic splitting angle. These can then be combined into single N and E traces using,

$$N_{model,aniso}(t) = N_{fast}(t) + N_{slow}(t+\delta t) \quad ,$$
$$E_{model,aniso}(t) = E_{fast}(t) + E_{slow}(t+\delta t) \quad ,$$

where δt is the fast-slow S-wave delay time.

5. **Rotate modelled data into DAS axis:** Once the simulated anisotropy has been applied to the synthetic modelled data, the North and East model components can then be rotated into the DAS axis, $DAS_{model}(t)$, given by,

$$DAS_{model}(t) = N_{model,aniso}\cos(\gamma) + E_{model,aniso}\sin(\gamma) \quad ,$$

where γ is the angle of the DAS fiber clockwise from North. It is important that this angle corresponds to the positive strain-rate direction along the fiber, as defined by the DAS interrogator.

6. **Convert data into same units:** The observed DAS data is in units of strain-rate, and the modelled DAS data is in units of particle velocity. In order to compare the modelled data to the observations, one therefore has to convert all the data either into strain-rates or particle velocities. We opt for converting the synthetic modelled data into strain-rate. The axial strain-rate, $\dot{\varepsilon}_x$, the native measurement of DAS, by differentiating spatially, as given by,

$$\dot{\varepsilon}_{xx} = \frac{\partial v_x}{\partial x} = \frac{\partial DAS_{model}(t)}{\partial x} \quad ,$$

where v_x is the velocity and x is the distance in the direction parallel to the cable. To make the modelled strain-rate consistent with the DAS, this differentiation should be applied over a length scale equal to the gauge length.

7. **Perform source mechanism inversion**: Now that the observed and modelled data are in the same coordinate system and both in units of axial strain-rate, they can be compared to one another in a source mechanism inversion. We do this for a DC-constrained inversion with 1×10^6 samples.

The source mechanism inversion results for the icequake highlighted by the yellow star in Figure 2 are shown in Figure 10. This icequake occurred when the DAS was configured in a linear arrangement. Figure 10a shows the event arrival at every channel along the DAS cable. The fast shear wave is dominant at smaller epicentral distances, with the slow shear wave prevalent at greater distances. The north component of a collocated geophone at the SW end of the linear cable is shown in black, at its approximate location. Figure 10b shows the DAS source mechanism inversion result. It should be noted that although we use S-waves in the inversion, we plot the P-wave radiation pattern, for consistency with other studies. The double-couple (DC) upper hemisphere focal mechanism is as one might expect (Kufner et al., 2021), suggesting horizontal slip with the slip vector aligned approximately with the ice flow direction. Uncertainty in the slip vector, shown by the red dashed lines, is of the order of $\pm 20°$. It should be noted

that a P-wave radiation pattern is plotted for consistency with other studies, although the fibre, and hence the inversion, is actually only sensitive to S-wave arrivals. Due to the $45°$ rotation of the S-wave radiation pattern relative to the P radiation pattern, the DAS cable lies near an S-wave radiation pattern maximum. The fit between the model and the observations is quantified by a variance reduction value of 0.65. The 2D observation, model, and difference fields show that the model provides a good fit, with negligible discernible coherent energy shown in the difference field for at least the majority of the slow S-wave. The scatter in the best fitting model result is due to the automatic alignment algorithm not correctly aligning every individual channel to the data. However, the general trend and the small proportion of scatter relative to fitted signal provides us with confidence that the inversion is successful. To our knowledge this is the first microseismic source mechanism inversion performed using DAS observations.

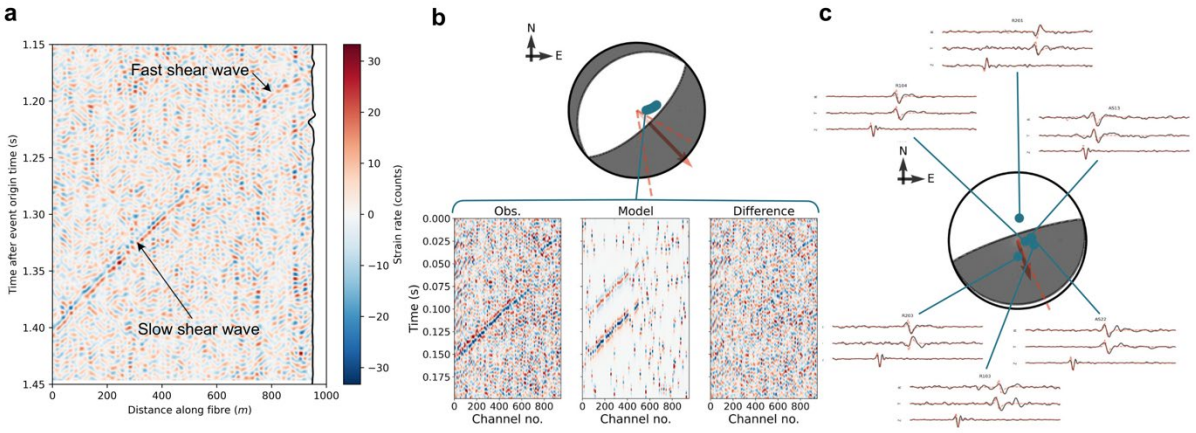

Figure 10. *Example of icequake source mechanism inversion results using DAS compared to geophone observations. The icequake occurred at 04:01:43 on the 14th January 2020, with the hypocentre shown by the yellow star in Figure 2a. a) The event S phase arrival at the DAS fibre, colored by strain-rate amplitude. The north component of a geophone collocated at one end of the fiber is shown by the black line at ~950 m. Key features of the event arrival are labelled. b) Double-Couple (DC) constrained full-waveform source mechanism inversion result using the DAS fiber observations, showing the observed waveforms used in the inversion, the modelled result, and the difference, colored by normalized strain-rate. The blue scatter points on the focal mechanism plot indicate the fiber and the red solid arrow and dashed lines indicate the slip vector and its associated uncertainty, respectively. c) DC constrained full-waveform source mechanism inversion result using geophone observations. The Z components include the P phase arrival only, and the R and T components include the S phase arrival only. Note that the Z and RT components are not plotted temporally aligned with one another, so as to clearly see P- and S- phase arrivals on the Z and RT components, respectively. Both focal mechanism radiation patterns are for an upper hemisphere projection for P-wave radiation, for consistency with other similar studies. Black quadrants correspond to compressional first arrivals. Both inversions used 10^6 samples of the moment tensor space.*

For comparison, we also show DC-constrained moment tensor source inversion results using geophones instead of DAS. All P-phase polarities are correctly fitted by the geophones, as are the majority of S-phase polarities, although there is more uncertainty for the S-wave matches due to some of the shear wave splitting not being entirely compensated for. Unfortunately, geophone coverage of the focal sphere for the network configuration is not as well configured for source mechanism analysis as previous studies (Kufner et al., 2021), leaving the most likely DC source poorly constrained compared to previous

observations. While the slip vector is approximately the direction of ice flow, and therefore in agreement with the DAS source inversion, the uncertainty, denoted by the dashed lines, indicates $\sim 270°$ azimuthal uncertainty. This is significantly greater than the $\sim 45°$ azimuthal uncertainty for the DAS source inversion. These results show that DAS provides better constraint of the source mechanism than the geophone network configuration, at least for this icequake.

Although the DAS source mechanism inversion outperforms the geophone inversion here, there are a number of challenges and limitations of the inversion. Indeed, some other DAS-constrained source mechanism inversions of other icequakes in this dataset have large slip-vector uncertainties compared to the icequake in Figure 10. The first limitation is that due to the low velocity firn layer, body wave phases arrive approximately vertically. This means that DAS is not sensitive to P-wave arrivals that have particle motions perpendicular to the fibre, unlike studies with no firn layer, and so the source mechanism can only be constrained using S-waves. The SNR of the DAS strain-rate data also limits the performance of our source inversion method. Comparing DAS strain-rate observations to model outputs can also be challenging. While some models output strain-rate, the results from the wave propagation code used here, fk (Zhu & Rivera, 2002), have to be converted from particle velocities. Even if the model used did output strain-rate directly, one still needs to simulate gauge-length effects. Another challenge is accounting for any anisotropy at the study site. Our results suggest that we have adequately accounted for anisotropy, but our method only holds for anisotropy path effects that can be approximated as homogenous on the spatial scale of the cable. If anisotropy were present at the source, or varied and was unknown across the length of the cable, then our method would be invalid. However, one could instead view the clear anisotropy observations in Figure 10a as an advantage of DAS, as an inversion for ice fabric anisotropy is theoretically possible. This does not always align with the observed phase arrival but instead on noise at a different time. This behaviour is likely to be primarily due to the low SNR of individual DAS channels, but could also be partly caused by inhomogeneous attenuation of the medium and/or assumptions related to the frequency content of the source-time function. Given these limitations, we therefore present these source mechanism inversion results as a foundation for future analysis of higher SNR data associated with more earthquakes, possibly in combination with more advanced inversion methods yet to be developed.

Here, we have treated the DAS and geophone data separately for the source mechanism inversions to provide a clear comparison between DAS and geophones. However, it is theoretically possible to combine DAS and geophone observations in a joint inversion. While this would not provide substantial gains for the event discussed here due to little increase in spatial constraint over the focal sphere, it might in other situations. One could imagine constraining the inversion better by deploying DAS within the centre of a network and geophones located more sparsely and at greater distances, for example. Performing such a joint inversion would provide its own challenges, such as how to weight DAS and geophone observations, since SNR and the single-component vs. three-components may introduce bias into the solution.

2.2.3 Shear-Wave Splitting

The strongly anisotropic nature of Rutford Ice Stream, as observed by shear-wave splitting, prompts the question of whether DAS can be used to observe seismic anisotropy. Shear-wave splitting provides a measure of the anisotropy along the ray path between the source and the receiver, which can be characterised by a delay time, *dt*, between the fast and slow S-wave arrivals, and the polarization, ϕ_f, of the fast S-wave. Shear wave splitting is typically estimated using three-component particle motion analysis on geophones. However, such an approach is not possible with linear DAS data because of its single-component nature and because it measures strain-rate rather than particle motion. Shear-wave splitting has previously been observed and analysed using DAS in strongly anisotropic shales, where S-waves arrive as two distinct arrivals allowing dt to be measured. In that example the anisotropy was known to have a relatively simple Vertical Transverse Isotropy (VTI) symmetry fabric such that determining the polarization was straightforward. However, at Rutford Ice Stream, the anisotropy has a more complicated orthorhombic symmetry (Smith et al., 2017), making determination of the polarization using DAS challenging.

Here we describe a methodology for estimating shear-wave splitting using the triangular DAS array. The motivation for this is to provide a proof of concept demonstrating that a 2D DAS geometry can be effectively used as a multi-component sensor capable of measuring shear-wave splitting. Using the triangular array partially alleviates for the inherent single component nature of DAS fiber because it records strain in a 2D plane rather than a 1D line, albeit with measurements at different orientations not at precisely the same location. However, if we assume that at the scale of the array the S-waves can be approximated as plane waves, we can approximate the triangular array as a point sensor by correcting for the spatial distribution, similar to the approach of Innanen et al. (2019). We then need to determine how any amplitude changes or polarity reversals recorded on the three sides of the array relate to the polarization of the S waves. The strain sensitivity pattern of an S-wave depends on both the orientation of the ray slowness vector (i.e. wavefront propagation direction) and of the polarization vector (Baird et al., 2020). Thus, if we can estimate the orientation of the slowness vector, then we can invert for the polarization that best fits the observed data.

For the splitting analysis we use a simplified velocity model with a constant firn layer S velocity of 1456 m s^{-1}, rather than a velocity gradient. We also assume an isotropic firn layer, such that the majority of the observed splitting is accrued as the wave passes through the ice column, as in Smith et al (2017).

The anisotropy inversion processing steps to find the fast-slow S-wave delay-time and polarization are as follows:

1. The triangular array is divided into the three linear segments, before applying a slant-stacking processing technique to stack data over various possible linear velocities represented by apparent slowness values. These are normalised by the number of channels in the stack to preserve amplitudes. This single processing step: (1) provides an estimate of one of the components of the slowness vector, which is required to estimate the propagation direction in order to forward model the strain sensitivity; (2) reduces the varying travel times over the length of the array to a

single point measurement at its midpoint; and (3) aids the identification and picking of the two S phases by separating them by their apparent slowness.

2. The fast and slow S-wave arrivals are then picked in the slowness-time domain, with the peak amplitude of the phase immediately after the pick taken as the recorded strain-rate amplitude of that arrival. The travel time difference between the fast and slow S-waves provides an estimate of the shear-wave splitting delay time *dt*. This analysis is then repeated for each side of the triangle, with the estimated *dt* at the midpoint of the array taken as the average of all those recorded on each side.
3. We then determine orientation of the slowness vectors (i.e. the wavefront propagation directions) of the two S waves. Since we have measured the apparent slowness of each S wave on the three sides of the array, we can apply a least squares inversion to solve for the best fitting horizontal slowness components. One should note that this slowness method provides an independent estimate for the azimuth of the source epicentre relative to the array from the hypocentral locations derived earlier in this study. This independence from our detection and location method is important because our original locations could be biased by the assumption of no anisotropy. In addition, we need to estimate the inclination of the propagation through the firn layer (to estimate emergence angle) and through the ice column (to define the path over which most of the splitting was accrued). We estimate this using Snell's law with the estimated horizontal slowness and a simplified two-layer velocity model.
4. Once we have determined the S-wave slowness vectors, we can then perform an inversion to obtain the fast and slow S-wave polarizations. The strain tensor, $\mathbf{e_S}$, associated with an S-wave propagating in the x_1 direction and polarized in the x_3 direction has the form,

$$\mathbf{e_S} = \begin{bmatrix} 0 & 0 & e_S \\ 0 & 0 & 0 \\ e_S & 0 & 0 \end{bmatrix} \quad ,$$

where the scalar factor e_S determines the overall amplitude of the strain. Thus, to simulate the DAS response to an S-wave with arbitrary propagation and polarization directions, we rotate this tensor into the appropriate orientation, and project the resulting tensor onto the fiber geometry. To eliminate the need to solve for e_S, we instead model the ratio of strain on different sides of the array, relative to a reference side (for example, $A_{side\ 1}/A_{side\ 2}$ and $A_{side\ 3}/A_{side\ 2}$ if side 2 is the reference). Since we have already estimated the azimuth and emergence inclination in step 3, this then leaves only the polarization to be determined, which is described by a rotation about the propagation axis. We then apply a least squares inversion to solve for the S polarization that minimises the misfit between the modelled and observed amplitude ratios. We independently invert for the fast and slow S-waves.

In order to compare our splitting measurements with previously published splitting results at Rutford from Smith et al. (2017), we convert *dt* to a percentage difference in fast and slow S velocities dv_S, by using the formula from (Wuestefeld et al., 2010),

$$dv_S = 100\frac{v_{Smean}dt}{r} \quad ,$$

where *r* is the source-receiver distance and v_{Smean} is the mean S-wave velocity along the full path.

Figure 11 shows how the anisotropy-inversion methodology is applied to measure the slowness of the icequake S-wave arrivals. The icequake clearly shows the differing moveout observed on each side of the array (see Figure 11a). The slant-stacking result is shown in the time-slowness domain in Figure 11c. The fast and slow S-wave arrivals can then be identified in this domain, as well as the peak strain-rate amplitude. To find the fast and slow S-wave polarizations, we choose side two of the array as our reference (see Figure 11b) and normalize the recorded strain amplitudes on each side by the amplitude on side two. These amplitude ratios allow us to determine the polarization, and are plotted in Figure 12a,c for the fast and slow S-waves, respectively.

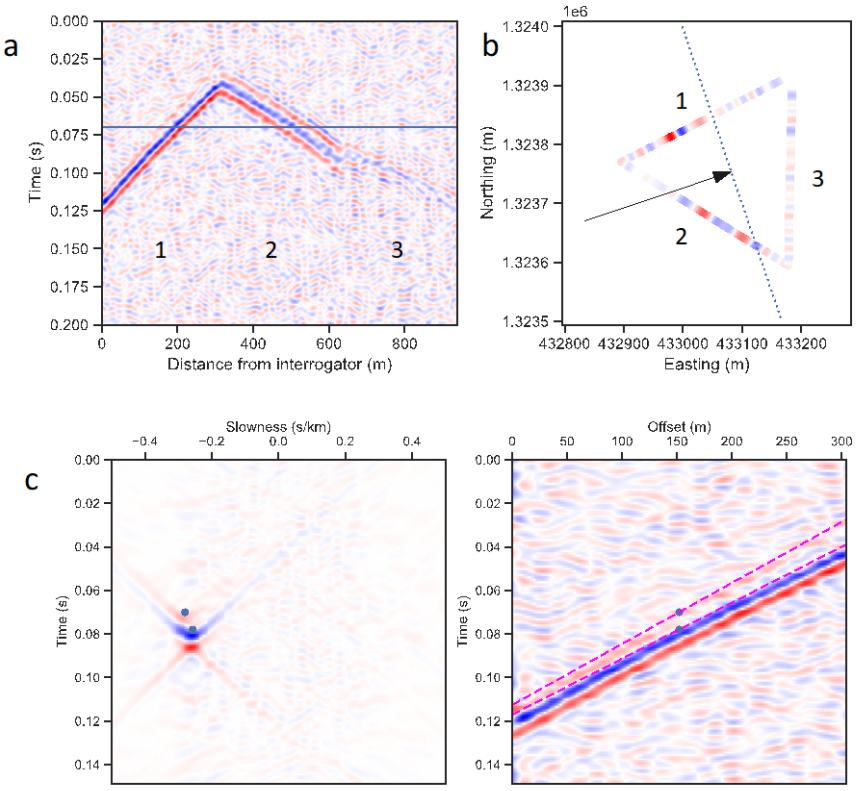

Figure 11. *(a) An example recording of an icequake on the triangular DAS array showing the S arrivals for an event at 04:18:07 on 17th January 2020. Clear linear moveouts can be observed on the three linear segments of the array. (b) Plot of triangular array geometry with a snapshot of the strain at the time indicated by the blue line in (a). Arrow and dotted line indicates the horizontal projection of the propagation vector and wavefront orientation, respectively, estimated from the slowness analysis. (c) Slowness analysis of the data from side 1 of the triangular array shown in (a). Left panel shows a slant stack of the DAS data shown on the right. Two distinct S waves can be observed and picked in the slant stack with travel times corresponding to the midpoint of the linear segment (indicated by blue points). The dashed magenta lines overlaid on the DAS data correspond to the linear moveout indicated by the picked slownesses in the slant stack.*

The S-wave polarization inversion results for the fast and slow S-wave for this example icequake are shown in Figure 12, along with the modelled strain sensitivity pattern of the best fitting inversion results. The best fit solution for the fast S-wave polarization, ϕ_f, is -74˚, where ϕ_f is defined in the ray frame according to the convention of Wuestefeld et al. (2010) (i.e. ϕ_f is 0˚ for SV and ±90˚ for SH, where positive angles are

measured clockwise looking along the ray towards the source). The best fit solution for the slow S-wave polarization is 15˚. The observation that the fast and slow S-wave polarizations, ϕ_f, are approximately perpendicular gives us confidence that this method is correctly measuring shear wave splitting.

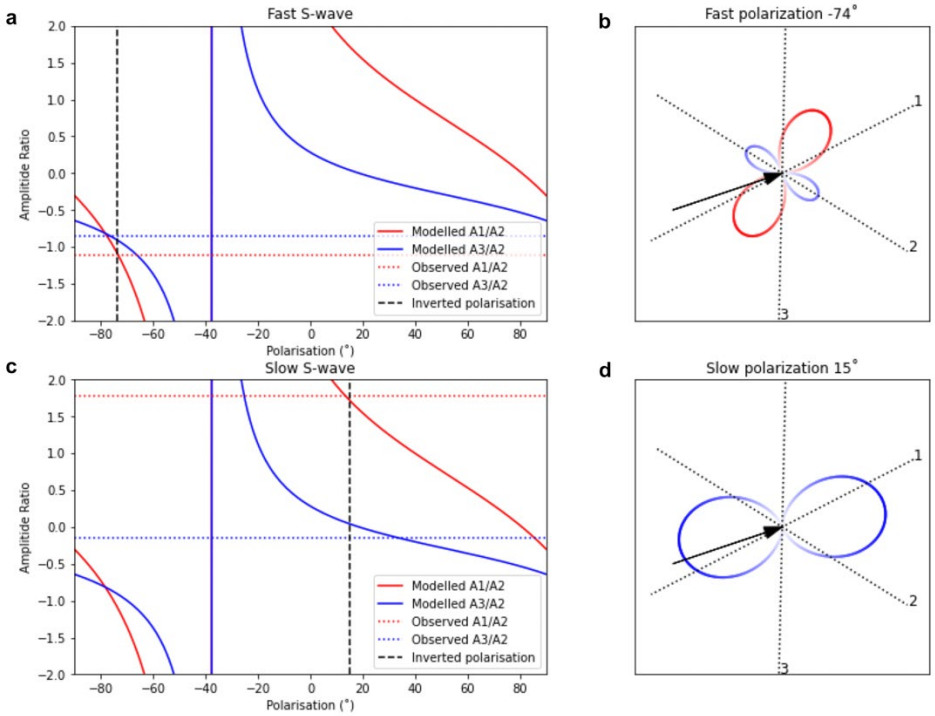

Figure 12. *a) Results of the polarization inversion of the fast S-wave for the same event as in Figure 7. Solid lines indicate modelled amplitude ratios as a function of polarization angle, dotted horizontal lines indicated the measured amplitude ratios and vertical dashed black line indicates the inverted polarization (-74˚). (b) Predicted strain sensitivity pattern for the inverted fast S arrival with polarization of 74˚. Dotted lines indicate the orientation of the three sides of the array, and black arrow indicates the horizontal projection of the propagation direction. (c) and (d) same as (a) and (b) but for the slow S-wave with inverted polarization of 15˚.*

Overall, this example event is found to have a shear-wave splitting delay time of 0.0093 s, based on averaging all three vertices of the triangular array to estimate the splitting at the centre of the array. We find a best fitting propagation azimuth of 72˚, as indicated by the arrow in Figure 11b, with horizontal slownesses of 0.28 s/km and 0.26 s/km for the fast and slow waves, respectively. Furthermore, we estimate a ray emergence inclination of ~22–25˚ based on an estimate of a mean firn layer S-wave velocity of 1456 m s^{-1}. However, we expect that the true inclination may be steeper than this, due to a non-linear gradient in the firn layer velocity structure. This is evidenced by the lack of P wave sensitivity. We find that reducing the inclination modulates the magnitude of the strain recorded but has only a minor effect on the relative strain sensitivity patterns observed, so we do not believe that uncertainty in emergence angle is a significant source of error.

We compare this splitting measurement with that of a previous study at Rutford (Smith et al., 2017) by converting our splitting fast-slow delay time into a S-wave velocity difference, dv_s. From the estimated source location for the event and the approximate path-length between source and receiver, we estimate a dv_s of 0.78% with an average inclination of 29˚. Although Smith et al. (2017) found some splitting as high as 4-5%, they also found that dv_S and ϕ_f varied substantially for different azimuths and inclinations, due to the orthorhombic symmetry of the anisotropy. If we restrict comparison of our results to only splitting parameters from Smith et al. (2017) with ray paths within azimuthal and inclination ranges of ±15˚ of our measurement, they have a mean dv_S and ϕ_f of 0.89% and -75˚, respectively. These are remarkably similar to our result.

The consistency between our result and previously published results gives us confidence in our approach, suggesting that smaller 2D DAS arrays can be effectively used as multicomponent sensors to detect and quantify shear wave splitting. Practically, however, this is unlikely to become the preferred method of characterising shear wave splitting. Deploying a DAS system in a small, azimuthally varying 2D array removes one of the major benefits of DAS systems: their large aperture, instead effectively reducing the array to a point sensor. In our case the source-receiver geometry is such that the ray path sampled a relatively weak splitting axis, producing a splitting measurement not very representative of the broader anisotropic fabric. In contrast, a less dense but broader aperture geophone array could record multiple splitting measurements from a single event, covering a wide range of azimuths and inclinations, which would provide a more comprehensive picture of the anisotropy. Nevertheless, we suggest there could be benefits of including small azimuthally varying 2D segments of DAS arrays as part of a larger aperture deployment to introduce some multi-component sensitivity.

3 Ambient Noise Interferometry

Ambient Noise Interferometry (ANI) is a technique that reconstructs ambient noise, the frequently discarded element of the seismic recording, into impulse signals. Although developed in the past 20 years, the technique itself is now mature and has been widely used in various conditions and scales from onshore to offshore, from applied geophysical exploration to continental imaging. The advantages of ANI are that acquisition costs can be significantly lower, as an active source is not required, and the methodology allows for regular repeat monitoring of the subsurface. A potential disadvantage of ANI is its dependency on noise. Ambient noise is often not isotropic or broad band and might change spatially and temporally, which limits the resolution of ANI images.

The technique obtains velocity measurements from ambient seismic noise by extracting the Green's Function through cross correlating the seismic noise wavefield between two receivers and stacking over time. This produces a response (Green's Function) at the second receiver that would be measured if the first was an impulse source. As a result, this operation finds the travel-time difference between waves recorded on the two receivers which is often termed a virtual seismogram. Creating multiple travel-time measurements between different pairs of receivers enables tomographic inversion of the data, which produces velocity estimates within the confines of the receiver array. Where these arrays are permanent and the receiver geometry is unchanged between surveys, a high degree of repeatability can be achieved which is ideal for monitoring CO_2 storage sites.

Dispersive surface waves are typically used for ANI as these signals are easily extracted from the noise record. When using surface waves, the depth of the measurement is directly related to the period of the signal, with higher frequencies confined to the shallow subsurface and lower frequencies extending to greater depths. Stork et al. (2018) used surface waves to assess to the potential of ANI to monitor the Aquistore CO_2 storage project, Canada, and showed the method was sensitive to depth of 100-400 m for surface wave periods between 0.5-1.4 s when using surface deployed geophones. Borehole arrays provide the opportunity to extend the sensitivity depth and have been shown to be capable of retrieving body waves from noise data which can be used to image both P- and S-wave velocity structure. Using cross correlations from geophones located at ~3km depth, Zhou & Paulssen (2017) retrieved direct P-wave arrivals in the 3-80 Hz band from vertical components and 3-50 Hz S-waves on the horizontal component from noise signal generated on the surface.

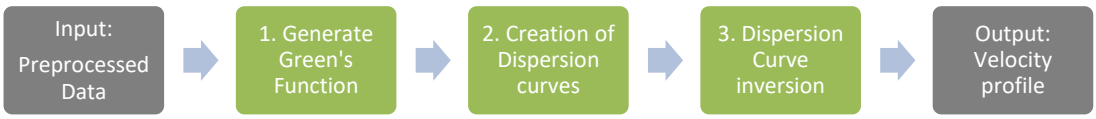

Figure 13. *Ambient noise interferometry processing workflow.*

3.1 Green's Function Generation

3.1.1 Characteristics of Ambient Noise Recorded on the Rutford DAS Array

Ahead of the creation of dispersion curves from the passive dataset, we seek to characterise some of the seismic noise sources recorded on the DAS array. In Figure 14 we present an example of high frequency noise, which begins abruptly and appears to originate close to one end of the array (0 m). We examine the frequency content of this signal by creating a spectrogram at 120m (Figure 15c) along the array using a short time Fourier transform (STFT). The signal starts with a gradually increasing amplitude over 10 to 80 Hz from around 25 s, then settles at a constant frequency around 40 Hz and 50 Hz at ~70 s. From 120 s, the signal frequency content and amplitude both increases until an abrupt stop at 145 s. From the pattern of this signal, we suspect it is likely noises from human activity at the interrogator site.

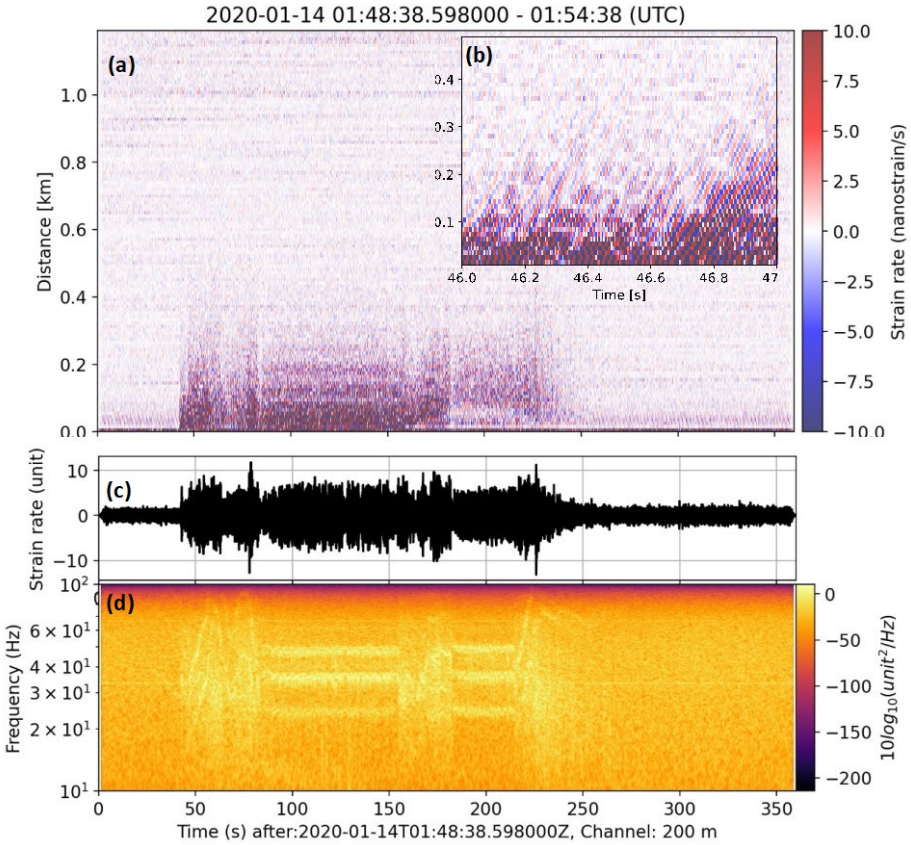

Figure 16. *(a) 6 minutes of DAS recording after median value been subtracted for each time step, and band pass filter 10 to 70 Hz. (b) a zoom in of (a) for 46-47 second at 0 to 500 m. (c) time series for DAS channel at 200 m. (d) spectrogram for (c), with amplitude in log scale.*

An example of a low frequency transient surface wave event is shown in Figure 17. The exact origin and location of this event is not determined with the linear fibre, but it is traveling near parallel to the fibre and from North-East to South-West, with apparent velocity around 1800 m/s. The dispersion feature of the signal is clearly shown in the spectrogram in Figure 18c and appear to originate from the shear margin

of the glacier. The exact sources of these signals is not resolved in the study, but crevasses and ice fracture at and beyond the shear margins are two potential candidates.

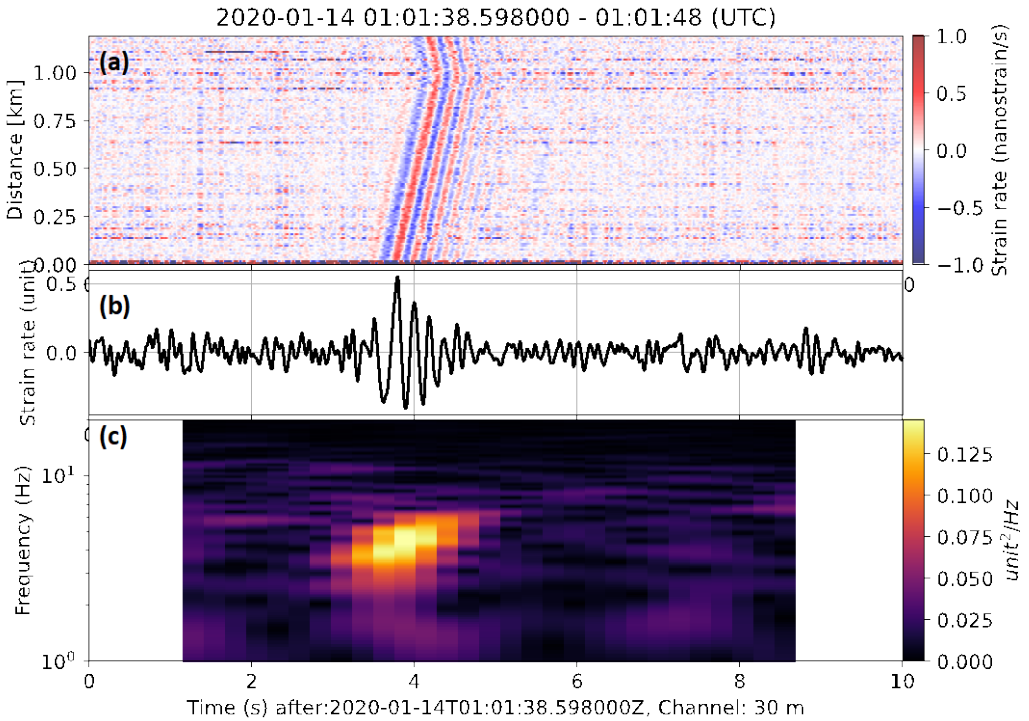

Figure 19. *(a) 10 seconds of DAS strain rate measurement after median removal and band pass filter 1 to 10 Hz. (b) waveform plot for channel at 30 m offset. (c)spectrogram of (b).*

Coherent signals are observed at periods beyond 100 seconds (Figure 20a). These signals are mostly travelling from far offset towards near offset and have propagating speed of the order of a few m/s. Interestingly, at frequency 0.01 to 0.5 Hz, there are three visible oscillating frequencies bands. As shown in Figure 21a and b, the oscillating frequencies increase with the increasing of the low-frequency signal (absolute) amplitude. Compared with its low-frequency counterpart that propagates slowly along the fibre (Figure 22d,e), the oscillating signals are spatially random, as we did not observe clear propagations in space-time plot or f-k domain in Figure 23c & d.

We suspect that this noise relates to variations in atmospheric pressure or temperature variation. Previous studies suggest these play a significant role on the low-frequency part of DAS and is the main source of signal beyond 100s (Ide et al., 2021). In addition, wind may be the source of the oscillations from 0.01 to 0.5 Hz, as it is expected to have a positive correlation with air pressure and temperature. Moreover, in the CCs (Figure 24b), we also see the strain signal of a moving marker-flag driven by the wind, although that signal is with higher frequency (9 Hz). Further investigation on this signal should be of interest for meteorology applications, but it is beyond the scoup of this study.

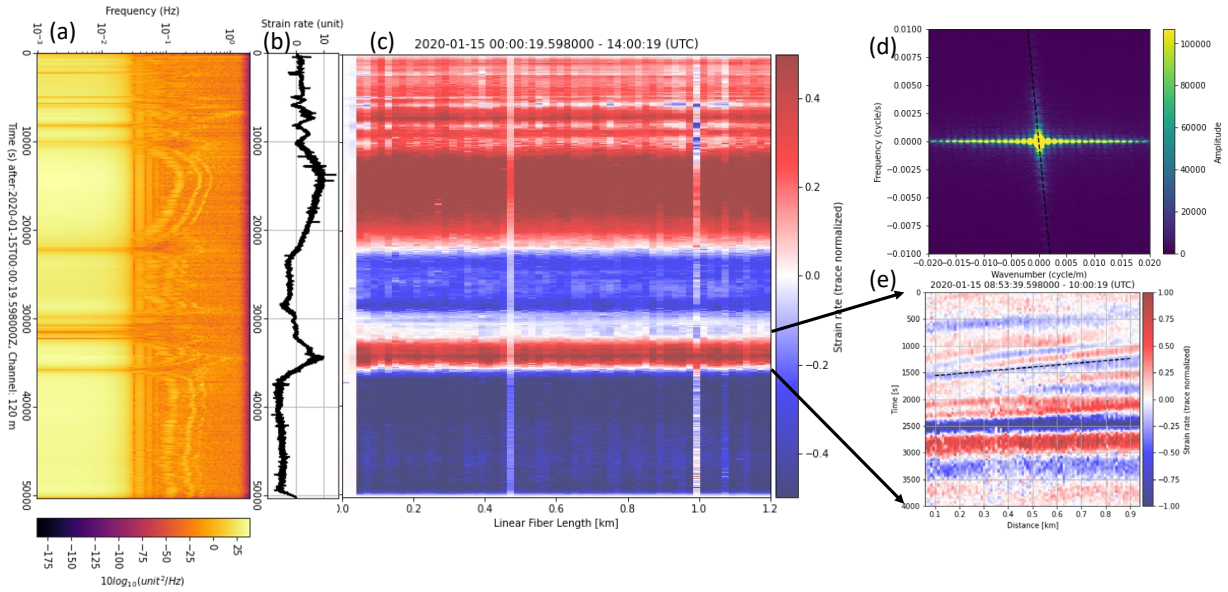

Figure 25. *Low frequency signals recorded by DAS. (a) spectrogram of the signal of (b) for 100-second sliding window and 90% overlapping. (b) is 14 hours of continuous DAS recording, after a low pass filter and a 2 Hz resampling, at one DAS channel at 120 m. (c) Image plot of all DAS channels. (d) f-k transform of the 14 hours DAS data. (e) a 4000 second zoom-in as indicated.*

3.1.2 Choice of Virtual Source: DAS versus Geophone

We initially obtain a DAS interferogram image through linearly stacking 2-minute CCs over the entire recording period of 5 days. We produce two different interferograms, the first uses the 600th DAS channel as a virtual source (Figure 26a), while the second takes the vertical component of a co-located geophone as the virtual source (Figure 27b). The CCs based on the DAS reference channel are noisy and the seismic responses, especially the lower frequency part, are hardly visible. In contrast, the geophone virtual source produces an interferogram with a clear visible seismic signal.

Considering the geophone derived interferogram (Figure 28b), lower frequency seismic signals are observed travelling both forward (from 0 offset to large offset) and backward with an apparent velocity around 1.7 km/s. The higher frequency responses travel primarily in the forward direction and clear dispersion is observed; the higher frequency signals have steeper slope, which indicates a slower velocity. The high frequency oscillating signal presenting at offset 0 to 600 m are due to strong harmonic signal at 33.3 Hz generated by the petrol electricity generator. The strong noise around 550 m of the DAS channel is probably strain produced by wind on a flag which was used to indicate the location of the geophone.

Although the geophone measures particle velocity and DAS measures strain rate, combining the two has counter-intuitively improved the resulting seismic response. We suspect this is due to the instrument noise on DAS channels having a harmful impact on the linear stacking of the cross-correlations. Firstly, we can see clear horizontal bands in Figure 29a which are most dominant at t = 0, and this implies that the

instrument noise on each DAS channel are not independent. This is likely due to the nature of DAS measurement, that senses the entire cable with one interrogator unit. When a geophone is used as virtual source, this breaks the coherency of instrument noises, and thus the linear stacking is unharmful. Secondly, we observe that the dominant noises contributing to the seismic responses are transient nature, therefore a large number of cross-correlations derived using a DAS virtual source contain only instrument noise. The correlation between vertical direction particle velocity (instrument response not corrected) and horizontal strain rate does produce a phase shift on CCs, which could be easily corrected by a second-order cross-correlation. Since, in this study, only apparent velocities (slope of in the time-offset plot) are used, no correction is needed.

3.1.3 Selective Stacking of Transient Noise

To improve the quality of the final interferogram image, previous studies have introduced more sophisticated techniques of stacking CCs, such as phase-weighted stacking (Schimmel et al., 2011) and SNR-weighted stacking (Cheng et al., 2015), or data selection (Dou et al., 2017). Phase-weighted stacking suppresses incoherent noise among two CCs and has several advantages over a standard linear stack (Dou et al., 2017). The approach, however, assumes that noise is continuous and coherent signal on every time span of CCs. With the exception of the harmonic generator noise, this is not the case in our dataset as most of the seismic signals are transient in nature. SNR-weighted stacking is based on the SNR of CCs and has been shown to perform well for anthropogenic seismic noises above 2 Hz, which are often transient and spatially variable(Cheng et al., 2015). For our 2-minute CCs, we find some signals have very low SNR when looking at individual channels which would be smeared (down-weighted by low SNR ratio) applying this method. Therefore, we take another approach, that is, selective stacking, where we only stack CCs based on certain criteria, which seems to be more suitable for the character of DAS data. Previous studies have implemented selective stacking based on SNR (Olivier et al., 2015), or based on signal apparent velocity from beamforming analysis (Vidal et al., 2014).

We choose to select a time period of CCs based on the maximal amplitude on tau-p domain (slant-stack) after a bandpass from 3 to 25 Hz. We select signals with their highest amplitude larger than 0.0014 (CC coefficient), and located at delay time close to zero (0 ± 0.05 s) with an absolute velocity smaller than 2500 m/s. These two criteria pass most of the visually observed signals, while most pure instrument noise CCs are removed. Of the 3068 (5 days with few hours of data loss) 2-minute CCs, 453 time periods met the search criterion and were linearly stacked. The selective-stacked CCs are presented in Figure 30c & d, which show higher SNR of seismic responses compared to the fully stacked CCs (Figure 31a&b). Although the coherent instrument noise is still contained in the selective stacked CCs for a DAS virtual source (Figure 32c), the approach up-weighted seismic signal by eliminating a large chunk of pure instrument noise.

Combining a geophone virtual source and selective stacking (Figure 33d), we achieved the best quality CCs for this dataset.

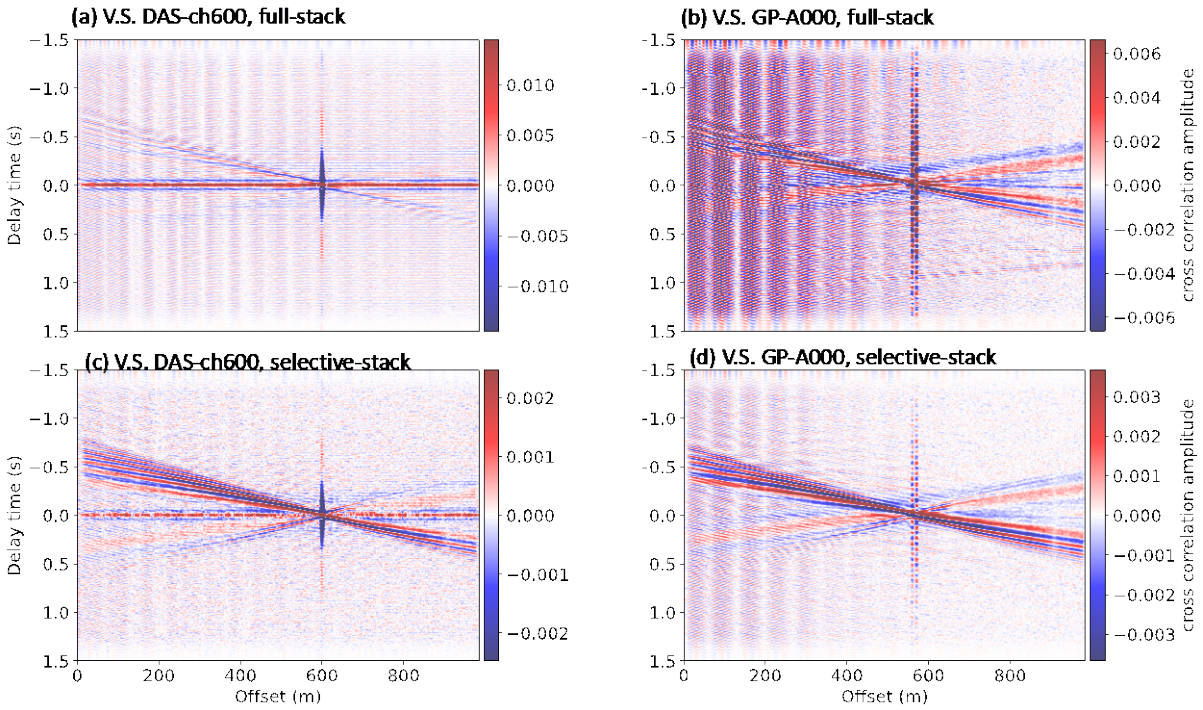

Figure 34. *(a) Stacked cross-correlations with a virtual source at DAS channel at 600 m. (b) Same as (a) but with a virtual source at a geophone (A000) located close to channel 550m. (c) Selective-stacked CCs for the same data as (a). (d) Selective-stacked CCs for the same data as (b)*

3.2 Dispersion Analysis

We create dispersion curves for both the active and passive datasets using a frequency wavenumber (f-k) transform after applying a Hann window to the data. After applying the f-k transform, we stack the positive and negative parts of the K domain to further enhance the signal. These f-k domain plots are then converted to the frequency–velocity domain as shown in Figure 35. Multiple modes of Rayleigh waves are presented in both datasets, but for simplicity, only the fundamental mode dispersion curve is extracted by picking the local maximal amplitude. The passive dataset contains lower frequency content and its dispersion curve is well constrained down to 3 Hz (Figure 36a). At 33.3 Hz, there is an inverse of velocity which is due to near-constant wavenumber of the strong noise observed at this frequency. This strong signal causes spectral leakage in fast Fourier transforms (FFT) even though a Hann window was applied before the FFT. However, since the frequency range is small it does not influence our inversion.

The 21 (2 shots for every 50 m) active survey dataset are processed as follows: First, for each shot gather, DAS channels are split into two segments at the location defined by the active source. Second, a f-k transform is applied to both segments. Third, negative wavenumbers are flipped and stacked with the positive part. Last, stack all shot gathers on f-k domain. From Figure 37b, we see that it contains signals mostly beyond 10 Hz with dispersion most stable between 15 and 50 Hz. This is likely due to the lack of low-frequency energy generated by the hammer and plate source compared to ambient seismic noise. In general, there is strong agreement between the dispersion curves from CCs and active shot gather

between 15 and 50 Hz. This provides us with confidence that the methods adopted when producing the ANI interferogram are appropriate.

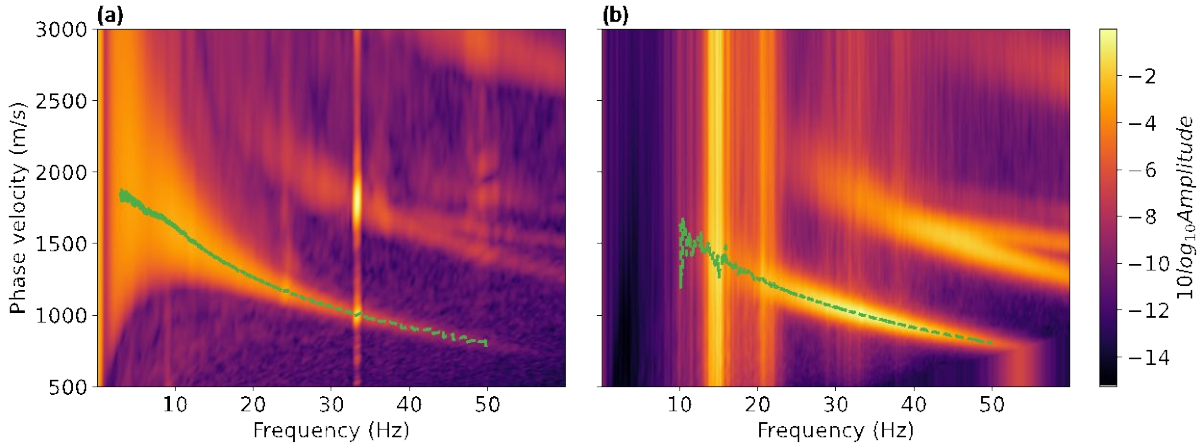

Figure 38. *(a) Extracting dispersion curve from the frequency-velocity domain of stacked CCs. (b) Dispersions obtained from stacked active shot.*

3.3 Velocity Inversion

Most previous surface wave inversion studies treat the subsurface as a layered model with either fixed or unfixed layer thickness, for two-station or multi-station surveys. The glacier firn layer is defined as a layer with continuous metamorphism of snow to ice. The continuous metamorphism results in a smooth increase of P and S velocity as a function of depth, until near-constant beneath ~100 m at RIS, according to a refraction survey. Due to these characteristics of the firn layer, instead of using a layered model with few layers and large thickness, we approach a near-continuous model with 100 layers, with each layer of thickness 1 m, except the bottom layer which represents a half-space. With 100 layers we significantly increase the number of variables and non-uniqueness of the inversion problem.

To simulate the phase velocity dispersion of Rayleigh wave, we use the Python package disba (Luu, 2021), which translated from the well-adapted Fortran program surf96 from Computer Programs in Seismology (Herrmann, 2013). A Gaussian-Newton inversion procedure is applied using the package pyGimli (Rücker et al., 2017), with the regularization lambda to be 20, and a predefined relative error of 10% to prevent overfitting. The large relative error and regularization also mean we are finding a solution that is close to the starting model.

We use a smoothed firn layer elastic model (smoothed P wave from refraction survey, constant Vp/Vs=1.95, gradually increasing density from 400 to 916 kg/m^3 as our starting model. As shown in Figure 39, the starting model has in general consistent phase velocity with the data. But especially at higher frequencies, the starting model has lower phase velocity than the data, which indicates the starting model is underestimating at shallower depth.

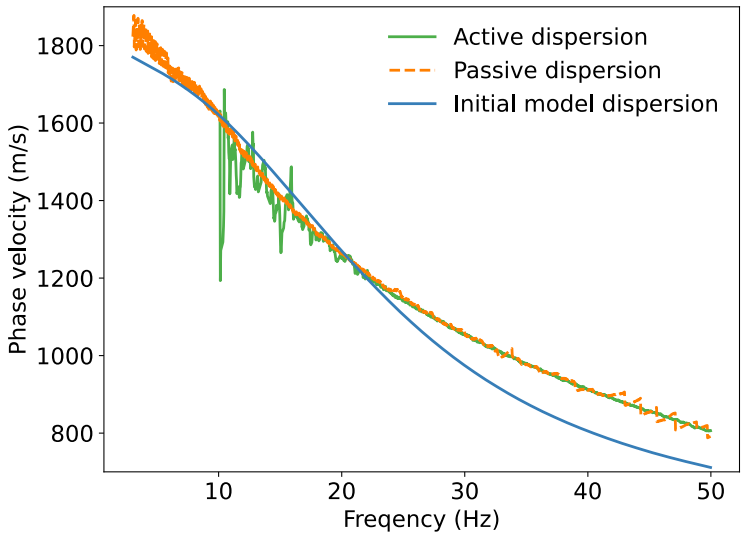

Figure 40. *Observed dispersion curves from selective stacked cross-correlations (CCs), active shot gathers, and modelled dispersion from the initial Vs model.*

To capture the uncertainty in the data, we inverted each individual dispersion curve measured from every 50 selected 2-minute CCs (out of 453 time periods) and 3 virtual sources with the 3 colocated geophones. This produced 27 independent inversions of S-wave velocity, then, a probability density function (PDF) has been calculated over every layer. The amplitudes of PDFs are presented as grey scale in (Figure 41a). The maximal point of PDFs, for each layer represents a Vs model (Vs_1 in Figure 42a). Alternatively, a Vs (Vs_2 in Figure 43a) is directly inverted from fully stacked CCs (all 453 time periods) and 3 virtual sources.

It is shown in Figure 44a that the Vs_1 is slight smoother, while shows slightly divergence with the Vs_2, which further indicates uncertainty of the final models. To verify the inversions, we compare them with an S-wave velocity model produced by taking a P-wave velocity model from a refraction survey by the British Antarctic Survey (BAS) and assuming a constant Vp/Vs ratio 1.95.

In general, our two Vs models agree with the BAS model down to 80 m depth. While for the depth below 80 m, our models suggest a slightly steeper increase, and reaches up to 2100 m/s.
In both our Vs models and the BAS model, we see a clear kink at around 12 m depth, which is similar to the densification observation and modelled by regional atmospheric climate model by van den Broeke, 2008.

A sensitivity analysis is done for discrete frequencies 3, 6, 10, 20, 40 Hz as shown in (Figure 45b) using the BAS firn layer model, which indicates the highest sensitivity of most signals (> 20 Hz) are for 0 to 40 m. Signal below 10 Hz start to have larger sensitivity for the lower part of the model. Signal below 6 Hz are dominantly sensitive to the lowermost layer of our model, which is assumed to be a half space.

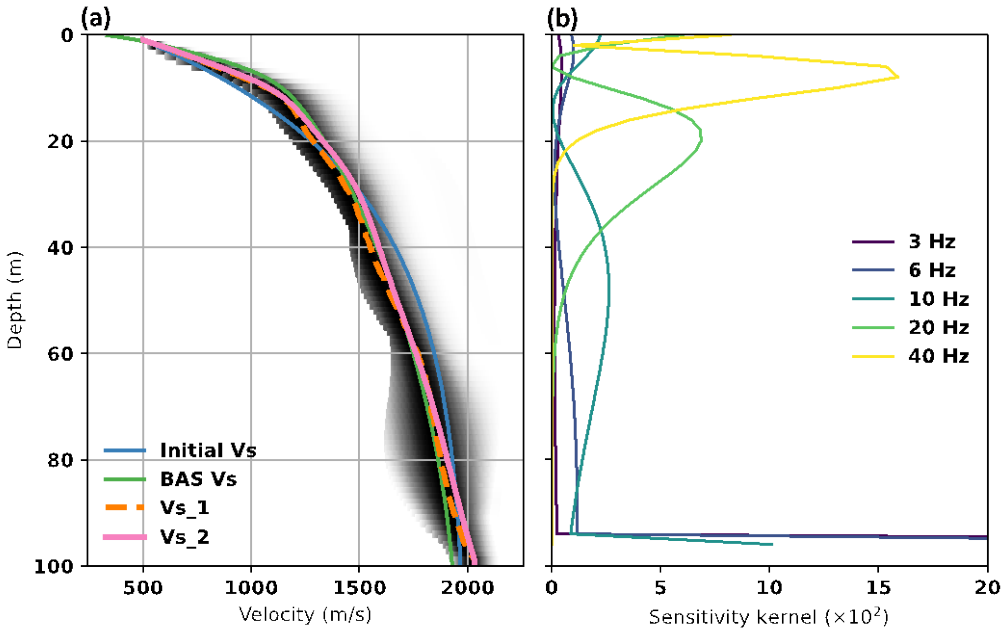

Figure 46. *(a) 2 Inverted Vs models, Vs_1 from maximal PDF (grey scale) and Vs_2 from direct inversion of a fully stacked CCs. (b) Sensitivity kernel calculated from the final model Vs_1.*

4 References

Aki, K., & Richards, P. G. (2002). *Quantitative seismology* (2nd ed.). San Francisco: University Science Books.

Baird, A. F., A. L. Stork, S. A. Horne, G. Naldrett, J. Kendall, J. Wookey, J. P. Verdon, and A. Clarke (2020). Characteristics of microseismic data recorded by distributed acoustic sensing systems in anisotropic media, GEOPHYSICS, 85, no. 4, doi: 10.1190/geo2019-0776.1.

Bohnhoff, M., G. Dresen, W. L. Ellsworth, and H. Ito. (2009). Passive Seismic Monitoring of Natural and Induced Earthquakes: Case Studies, Future Directions and Socio-Economic Relevance, in *New Frontiers in Integrated Solid Earth Sciences*, Springer Netherlands, Dordrecht, 261–285.

Dou, S., Lindsey, N., Wagner, A. M., Daley, T. M., Freifeld, B., Robertson, M., Peterson, J., Ulrich, C., Martin, E. R., & Ajo-Franklin, J. B. (2017). Distributed Acoustic Sensing for Seismic Monitoring of the Near Surface: A Traffic-Noise Interferometry Case Study. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-11986-4

Cheng, F., Xia, J., Xu, Y., Xu, Z., & Pan, Y. (2015). A new passive seismic method based on seismic interferometry and multichannel analysis of surface waves. *Journal of Applied Geophysics*, *117*, 126–135. https://doi.org/10.1016/j.jappgeo.2015.04.005

Herrmann, R. B. (2013). Computer programs in seismology: An evolving tool for instruction and research. *Seismological Research Letters*, *84*(6). https://doi.org/10.1785/0220110096

Hudson, T. (2020). TomSHudson/SeisSrcInv: Initial release for publication. Zenodo. https://doi.org/10.5281/zenodo.3726697

Hudson, T. S., A. F. Baird, A. M. Brisbourne, A. M. Smith, A. Butcher, A. Chalari, and A. Clarke. (2021). Distributed Acoustic Sensing (DAS) for natural microseismicity studies : A case study from Antarctica.

Ide, S., Araki, E., & Matsumoto, H. (2021). Very broadband strain-rate measurements along a submarine fiber-optic cable off Cape Muroto, Nankai subduction zone, Japan. *Earth, Planets and Space*, *73*(1). https://doi.org/10.1186/s40623-021-01385-5

Innanen, K. A., Lawton, D., Hall, K., Bertram, K. L., Bertram, M. B., & Bland, H. C. (2019). Design and deployment of a prototype multicomponent distributed acoustic sensing loop array. In *SEG Technical Program Expanded Abstracts 2019* (pp. 953–957). https://doi.org/10.1190/segam2019-3216304.1

Kwiatek, G. et al., 2019, Controlling fluid-induced seismicity during a 6.1-km-deep geothermal stimulation in Finland, Sci. Adv., 5, no. 5, eaav7224, doi: 10.1126/sciadv.aav7224.

Kettlety, T., J. P. Verdon, A. Butcher, M. Hampson, and L. Craddock. (2020). High-Resolution Imaging of the M L 2 . 9 August 2019 Earthquake in Lancashire, United Kingdom , Induced by Hydraulic Fracturing during Preston New Road PNR-2 Operations, Seismol. Res. Lett., no. August 2019, doi: 10.1785/0220200187.Supplemental.

Kufner, S., Brisbourne, A. M., Smith, A. M., Hudson, T. S., Murray, T., Schlegel, R., et al. (2021). Not all Icequakes are Created Equal: Basal Icequakes Suggest Diverse Bed Deformation Mechanisms at Rutford Ice Stream, West Antarctica. *Journal of Geophysical Research: Earth Surface*, *126*(3). https://doi.org/10.1029/2020JF006001

Lellouch A., Lindsey N. J., Ellsworth W. L., and Biondi B. L., 2020, Comparison between distributed acoustic sensing and geophones: downhole microseismic monitoring of the FORGE geothermal experiment, Seismological Research Letters, 91 (6), 3256–3268. doi: 10.1785/0220200149.

Lindsey N. J., Rademacher H., and Ajo-Franklin J. B., 2020, On the broadband instrument response of fiber-optic DAS arrays, Journal of Geophysical Research: Solid Earth, 125 (2), e2019JB018145, doi: 10.1029/2019JB018145

Lomax, A., C. Satriano, and M. Vassallo, 2012, Automatic Picker Developments and Optimization: FilterPicker--a Robust, Broadband Picker for Real-Time Seismic Monitoring and Earthquake Early Warning, Seismol. Res. Lett., 83, no. 3, 531–540, doi: 10.1785/gssrl.83.3.531.

Olivier, G., Brenguier, F., Campillo, M., Lynch, R., & Roux, P. (2015). Body-wave reconstruction from ambient seismic noise correlations in an underground mine. *GEOPHYSICS*, *80*(3). https://doi.org/10.1190/geo2014-0299.1

Luu, K. (2021). disba: Numba-accelerated computation of surface wave dispersion. https://doi.org/https://doi.org/10.5281/zenodo.3987395

Rücker, C., Günther, T., & Wagner, F. M. (2017). pyGIMLi: An open-source library for modelling and inversion in geophysics. *Computers & Geosciences*, *109*, 106–123. https://doi.org/10.1016/j.cageo.2017.07.011

Sabbione, J. I., Sacchi, M. D., & Velis, D. R. (2015). Radon transform-based microseismic event detection and signal-to-noise ratio enhancement. *Journal of Applied Geophysics*, *113*, 51–63.

Schimmel, M., Stutzmann, E., & Gallart, J. (2011). Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale. *Geophysical Journal International*, *184*(1), 494–506. https://doi.org/10.1111/j.1365-246X.2010.04861.x

Smith, Emma C., Baird, A. F., Kendall, J. M., Martin, C., White, R. S., Brisbourne, A. M., & Smith, A. M. (2017). Ice fabric in an Antarctic ice stream interpreted from seismic anisotropy. *Geophysical Research Letters*, *44*(8), 3710–3718. https://doi.org/10.1002/2016GL072093

Stork, A. L., J. P. Verdon, and J.-M. Kendall, 2015, The microseismic response at the In Salah Carbon Capture and Storage (CCS) site, Int. J. Greenh. Gas Control, 32, 159–171, doi: 10.1016/j.ijggc.2014.11.014.

Stork, A. L., Allmark, C., Curtis, A., Kendall, J. M., & White, D. J. (2018). Assessing the potential to use repeated ambient noise seismic tomography to detect CO2 leaks: Application to the Aquistore storage site. *International Journal of Greenhouse Gas Control*, *71*(February), 20–35. https://doi.org/10.1016/j.ijggc.2018.02.007

van den Broeke, M. (2008). Depth and Density of the Antarctic Firn Layer. *Arctic, Antarctic, and Alpine Research*, *40*(2), 432–438. https://doi.org/10.1657/1523-0430(07-021)

Teanby, N. A., J. Kendall, and M. Van Der Baan, 2004, Automation of Shear-Wave Splitting Measurements using Cluster Analysis, 94, no. 2, 453–463.

Verdon, J. P., and J. Budge, 2018, Examining the capability of statistical models to mitigate induced seismicity during hydraulic fracturing of shale gas reservoirs, Bull. Seismol. Soc. Am., 108, no. 2, 690–701, doi: 10.1785/0120170207.

Walter, F., Gräff, D., Lindner, F., Paitz, P., Köpfli, M., Chmiel, M., & Fichtner, A. (2020). Distributed Acoustic Sensing of Microseismic Sources and Wave Propagation in Glaciated Terrain. *Nature Communications*, *53*(9), 1689–1699. https://doi.org/10.1017/CBO9781107415324.004

Wuestefeld, A., Al-Harrasi, O., Verdon, J. P., Wookey, J., & Kendall, J. M. (2010). A strategy for automated analysis of passive microseismic data to image seismic anisotropy and fracture characteristics. *Geophysical Prospecting*, *58*(5), 755–773. https://doi.org/10.1111/j.1365-2478.2010.00891.x

Zhou, W., & Paulssen, H. (2017). P and S Velocity Structure in the Groningen Gas Reservoir From Noise Interferometry. *Geophysical Research Letters*, *44*(23), 11,785-11,791. https://doi.org/10.1002/2017GL075592

Zhu, L., & Rivera, L. A. (2002). Computation of dynamic and static displacement from a point source in multi-layered media. *Geophysical Journal International*, *148*, 619–627.