POLYPUBLIE
Polytechnique Montréal

POLYTECHNIQUE MONTRÉAL
UNIVERSITÉ D'INGÉNIERIE

Titre: Title:	Advanced Seismic Processing Supported by Deep Learning for Subsea Permafrost Characterization
Auteur: Author:	Jefferson Bustamante Restrepo
Date:	2023
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Bustamante Restrepo, J. (2023). Advanced Seismic Processing Supported by Deep Learning for Subsea Permafrost Characterization [Thèse de doctorat, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/56995/

Document en libre accès dans PolyPublie
Open Access document in PolyPublie

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/56995/
Directeurs de recherche: Advisors:	Gabriel Fabien-Ouellet, & Mathieu J. Duchesne
Programme: Program:	Génie minéral

Ce fichier a été téléchargé à partir de PolyPublie, le dépôt institutionnel de Polytechnique Montréal
This file has been downloaded from PolyPublie, the institutional repository of Polytechnique Montréal

https://publications.polymtl.ca

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Advanced seismic processing supported by deep learning for subsea permafrost characterization

JEFFERSON BUSTAMANTE RESTREPO
Département des génies civil, géologique et des mines

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*
Génie minéral

Décembre 2023

© Jefferson Bustamante Restrepo, 2023.

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Cette thèse intitulée :

Advanced seismic processing supported by deep learning for subsea permafrost characterization

présentée par **Jefferson BUSTAMANTE RESTREPO**
en vue de l'obtention du diplôme de *Philosophiæ Doctor*
a été dûment acceptée par le jury d'examen constitué de :

Philippe PASQUIER, président
Gabriel FABIEN-OUELLET, membre et directeur de recherche
Mathieu J. DUCHESNE, membre et codirecteur de recherche
Gilles BELLEFLEUR, membre
Eileen MARTIN, membre externe

ACKNOWLEDGEMENTS

I wish to begin by expressing my gratitude to my supervisor, Professor Gabriel Fabien-Ouellet, for providing me the opportunity to pursue my PhD studies. This journey has been remarkable, and I owe its success to your constant support and vast knowledge. Our weekly meetings and engaging discussions on the project and related subjects have been not only enlightening but also deeply enriching. Thank you for taking me as your first PhD student.

I extend my gratitude to Dr. Mathieu J. Duchesne, my co-supervisor, for his immense contribution to the development of this thesis. Your profound expertise in permafrost and seismic processing has been invaluable to me. I thoroughly enjoyed our meetings and could not be more thankful for your wise counsel and unwavering support throughout this entire period.

My gratitude also goes to the Geological Survey of Canada (GSC), more specifically the Environmental Geoscience Program of the GSC for its support, and the Korea Polar Research Institute (KOPRI) for granting me access to their datasets. Additionally, I would like to extend my appreciation to the KEGS Foundation and the Fondation J.A de Sève for providing financial support in the form of scholarships.

I would like to acknowledge Dr. Amr Ibrahim for sharing his extensive knowledge with me over the past years. Additionally, I would like to extend my gratitude to my friends, Anne and Adrien, for accompanying me on this journey and to my colleagues at Polytechnique, including Karine, Dany, Camilla, Jeremy, and Lahiru, for their incredible conversations and for dedicating their time to Polytechnique.

Lastly, I would like to express my deepest thanks to my wife, Lina, for all your support during this process. No podría haber terminado este doctorado sin tu apoyo.

RÉSUMÉ

Une caractéristique distinctive de la région arctique est la présence étendue du pergélisol, défini comme un matériau terrestre sous-surface gelé depuis au moins deux ans. La dégradation du pergélisol due à l'augmentation des températures est préoccupante car elle peut libérer des réserves de carbone, augmentant ainsi la quantité de carbone dans l'atmosphère et contribuant au réchauffement climatique mondial. Il est donc crucial de surveiller la distribution et la dégradation du pergélisol pour prédire son impact sur le réchauffement climatique. Alors que les techniques géophysiques évaluent efficacement les conditions du pergélisol sur terre, leur capacité à caractériser avec précision les zones de pergélisol sous-marin reste limitée. Le traitement sismique permet de faire la distinction entre les sédiments sous-marins contenant de la glace et ceux sans glace en exploitant la vitesse sismique de la glace dans le pergélisol sous-marin. Cependant, le traitement conventionnel est confronté à des défis en raison de l'absence de refractions reconnaissables à la base de la couche de pergélisol et de l'interférence entre différents types d'arrivées sismiques. En particulier, les techniques conventionnelles ne sont pas assez précises pour déterminer la distribution verticale du pergélisol, ce qui entrave notre compréhension de la distribution du pergélisol à l'échelle régionale. En revanche, les approches d'apprentissage profond contournent la nécessité de séparer les différentes arrivées en utilisant les données enregistrées pour établir une corrélation entre les images sismiques et les structures sous-surface. L'apprentissage profond peut potentiellement exploiter des arrivées traitées de manière non conventionnelle pour mieux résoudre le sous-surface.

Cette thèse présente une méthodologie d'apprentissage profond pour imager les variations de la vitesse sismique causées par le pergélisol sous-marin afin de permettre une caractérisation et une cartographie efficaces à grande échelle. Pour atteindre cet objectif, nous avons formulé une approche en trois étapes. Tout d'abord, une analyse de sensibilité a été menée pour déterminer la pertinence de l'utilisation des relevés sismiques pour détecter les changements dans les propriétés du pergélisol. Nos résultats suggèrent qu'une approche combinée intégrant les données sismiques après l'application de multiples transformations de domaine peut améliorer l'approche d'inversion pour discerner les changements dans la distribution du pergélisol. Deuxièmement, une méthodologie d'apprentissage profond a été élaborée pour inverser les propriétés sismiques à partir des données enregistrées collectées dans les zones de pergélisol. En s'appuyant sur les résultats de l'analyse de sensibilité, nous avons élaboré un réseau neuronal à entrées multiples et à sorties multiples qui prend en entrée des données sismiques de domaines distincts et fournit en sortie les vitesses sismiques inversées et les valeurs d'atténuation. Les résultats démontrent la capacité du réseau à imager avec précision

le modèle de vitesse et d'atténuation du sous-surface dans les zones de pergélisol. Enfin, nous avons évalué l'approche d'apprentissage profond proposée en utilisant un jeu de données sismiques à l'échelle régionale acquis dans la mer de Beaufort canadienne pour cartographier le pergélisol sous-marin dans la zone d'étude.

En résumé, les résultats de cette thèse montrent comment les données sismiques combinées à des techniques d'apprentissage profond peuvent améliorer la cartographie et la caractérisation du pergélisol sous-marin. Ces résultats constituent la première étape vers un objectif plus vaste de surveillance de la dégradation du pergélisol sous-marin à l'échelle régionale et de l'élaboration d'un outil pour prédire les impacts de sa dégradation sur le réchauffement climatique. Des recherches ultérieures pourraient être menées afin d'améliorer les capacités du réseau neuronal ou pour réaliser d'autres études dans la mer de Beaufort afin d'approfondir notre compréhension de la distribution du pergélisol et des processus associés. De plus, notre méthodologie proposée peut être étendue à d'autres conditions souterraines où différentes interférences sismiques peuvent entraver le traitement sismique et l'interprétation.

ABSTRACT

A distinctive feature of the Arctic region is the extensive presence of permafrost, defined as subsurface earth material frozen for at least two years. Permafrost degradation owing to rising global temperatures is concerning because it can release carbon stores, increasing atmospheric carbon and contributing to global warming. Therefore, it is crucial to monitor the distribution and degradation of permafrost to predict its impact on global warming. While geophysical techniques effectively assess onshore permafrost conditions, their ability to accurately characterize subsea permafrost zones remains limited. Seismic processing can distinguish between ice-bearing and ice-free subsea sediments by the velocity increase induced by the ice content. However, conventional processing faces challenges due to the lack of recognizable refractions at the base of the permafrost layer and interference between different types of seismic arrivals. In particular, conventional techniques are not precise enough in determining the vertical distribution of permafrost, which hampers our understanding of permafrost distribution on a regional scale. In contrast, deep learning approaches bypass the need for segregating different arrivals by utilizing recorded data to establish a direct mapping between seismic images and subsurface structures. Deep learning can potentially leverage unconventionally processed arrivals to better resolve the subsurface.

This thesis introduces a deep learning methodology for the effective large-scale characterization and mapping of seismic velocity and attenuation variations caused by different subsea permafrost conditions. To achieve our objective, we have formulated a three-step approach. Firstly, a sensitivity analysis has been conducted to determine the suitability of employing seismic surveys for detecting changes in permafrost properties. Our findings suggest that a combined approach integrating seismic data after applying multiple domain transformations can improve the inversion approach in discerning changes in permafrost distribution. Secondly, a deep-learning methodology has been devised to invert seismic properties from recorded data collected within permafrost zones. Building upon results from the sensitivity analysis, we designed a multi-input, multi-output neural network that inputs seismic data from distinct domains and outputs inverted seismic velocities and attenuation values. The results demonstrate the capability of the network to accurately image the subsurface velocity and attenuation model in permafrost zones. Lastly, we evaluated the proposed deep learning approach, using two regional-scale seismic surveys conducted in the Canadian Beaufort Sea to identify subsea permafrost in the study area. To effectively map permafrost distribution, we propose adopting predetermined thresholds on inverted parameters. This enables classification into three categories: ice-free, ice-bearing, and ice-bonded permafrost.

In summary, the results from this thesis show how seismic data combined with deep learning techniques can enhance mapping and characterization of subsea permafrost. These findings are the first step towards a larger goal of monitoring subsea permafrost degradation at a regional scale and providing a tool for predicting the impacts of its degradation on global warming. Further research could be conducted to enhance the capabilities of the neural network or to conduct additional studies in the Beaufort Sea to increase our understanding of permafrost distribution and associated processes. Moreover, our proposed methodology can be extended to other subsurface conditions where different seismic wave interference can hinder seismic processing and interpretation.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii

RÉSUMÉ iv

ABSTRACT vi

TABLE OF CONTENTS viii

LIST OF TABLES xi

LIST OF FIGURES xii

LIST OF SYMBOLS AND ACRONYMS xx

CHAPTER 1 INTRODUCTION 1
1.1 Measuring permafrost distribution 2
1.2 Problem identification 4
1.3 Research Objectives 6
1.4 Dissertation structure 7

CHAPTER 2 LITERATURE REVIEW 9
2.1 Permafrost 9
2.1.1 Global permafrost distribution 9
2.1.2 Permafrost formation and the associated thermal regime 11
2.1.3 Structures associated with permafrost 12
2.1.4 Subsea Permafrost 20
2.2 Seismic mapping of permafrost 24
2.2.1 Seismic properties of permafrost 24
2.2.2 Fundamentals of seismic imaging 28
2.2.3 Seismic methods and permafrost 32
2.2.4 Challenges in subsea permafrost imaging 34
2.2.5 Seismic datasets 36
2.3 Deep learning for seismic inversion 37
2.3.1 Optimization 39
2.3.2 Neural networks architectures and components 42

2.3.3 Deep learning for seismic inversion 52
2.3.4 Challenges of using deep learning 54

CHAPTER 3 ARTICLE 1: UNDERSTANDING SEISMIC VELOCITY VARIATIONS OF SUBSEA PERMAFROST: A SENSITIVITY STUDY 58
3.1 Abstract 58
3.2 Introduction 58
3.3 The seismic signature of subsea permafrost 61
3.4 Seismic Velocity of subsea permafrost 65
3.4.1 A simplified conceptual model 67
3.5 Sensitivity analysis 67
3.5.1 Qualitative analysis 67
3.5.2 Quantitative sensitivity analysis 72
3.6 Reproducing seismic data with simple models 75
3.6.1 An integrated velocity model 77
3.7 Conclusion 79
3.8 Acknowledgments 82
3.9 Addendum A: Sensitivity analysis in Radon domains 82

CHAPTER 4 ARTICLE 2: DEEP-LEARNING VISCOELASTIC SEISMIC INVERSION FOR MAPPING SUBSEA PERMAFROST 88
4.1 Abstract 88
4.2 Introduction 88
4.3 Methodology 91
4.3.1 Dataset 92
4.3.2 Neural Network architecture 99
4.3.3 Training 102
4.4 Results 104
4.4.1 Testing dataset 104
4.4.2 Synthetic 2D model 107
4.4.3 Field data 110
4.4.4 1D vs 2D inversion 112
4.5 Conclusion 116
4.6 Acknowledgements 117
4.7 Addendum A: Cross talk analysis between the parameters evaluated 117
4.8 Addendum B: Using single input NNs 119

CHAPTER 5 ARTICLE 3: MAPPING SUBSEA PERMAFROST DISTRIBUTION IN THE BEAUFORT SEA WITH MARINE SEISMIC AND DEEP LEARNING . . 122
5.1 Abstract . . . 122
5.2 Introduction . . . 123
5.3 Survey area and seismic dataset . . . 126
5.3.1 Subsea permafrost in the Beaufort Sea . . . 126
5.3.2 Survey lines . . . 128
5.4 Processing seismic data with a Neural Network . . . 129
5.5 Results . . . 133
5.5.1 Seismic inversion . . . 133
5.5.2 Line intersections . . . 139
5.5.3 Well log comparisons . . . 140
5.5.4 Permafrost distribution in the Beaufort Sea . . . 142
5.6 Discussion . . . 144
5.6.1 Comparisons with previous studies . . . 144
5.6.2 Schematic permafrost distribution in the Canadian Beaufort Sea. . . . 146
5.7 Acknowledgments . . . 149
5.8 Addendum A: Limitations of the NN . . . 149

CHAPTER 6 CONCLUSION AND FUTURE WORK . . . 153
6.1 Future work . . . 155
6.1.1 Enhancing the capabilities of the NN . . . 155
6.1.2 Additional studies and surveys in the Beaufort Sea . . . 156

REFERENCES . . . 158

LIST OF TABLES

Table 3.1 Lithology and seismic parameters found in the Beufort Sea. We refer to the same Q for P- and S-waves. 66

Table 4.1 Lithology and seismic parameters used to create the training dataset. In the table, F., PF., UF., refers to frozen, partially frozen, and unfrozen sediments. The last column represents the probability of the layer being skipped from the model. 94

Table 4.2 Encoder's architectures for the $t-x$, $\tau-p$, $f-c$, and $f-p$. Information in the cells describes the resulting size after passing the layer. The filter kernel in the convolutional 2D layers is presented in parentheses. . . . 99

Table 4.3 2D NN encoder's architectures for the $t-x$, $\tau-p$, $f-c$, and $f-p$. Information in the cells describes the resulting size after passing the layer. The filter kernel in the convolutional 2D layers is presented in parentheses. 101

Table 4.4 Calculated L1 and L2 errors and the SSIM index for the 1D testing dataset and the 2D wedge model. 110

Table 4.5 L1 and L2 norm error for the single-input and the multi-input NN evaluated in the 1D dataset and the 2D wedge model. Green cells display lower errors per column while the red cells show higher ones. . 121

Table 5.1 Typical seismic properties found in the Canadian Beaufort Sea (Bellefleur et al., 2007; Buckingham, 1997; Dou et al., 2016; Matsushima et al., 2016). 128

Table 5.2 Description of the hidden layers in the encoders and decoders of the NN as presented by (Bustamante et al., 2023b) and summarized in Fig. 5.2. The cells indicate the output dimension of the layer and the filter size (in parenthesis). 131

LIST OF FIGURES

Figure 2.1 Permafrost in the Northern Hemisphere. Source: GRID-Arendal/ Nunataryuk (`https://www.grida.no/resources/13519`). 10

Figure 2.2 Sketch of the thermal regime on permafrost areas. Adapted from Allard et al. (2012). 11

Figure 2.3 Schematic representation of onshore and subsea permafrost distribution and associated geological structures. Source: Geological Survey of Canada. 13

Figure 2.4 Thermokarst lake formation as presented in Liebner et al. (2020). . . 14

Figure 2.5 Pingos in the Mackenzie Delta. Photo released as public domain by Emma Pike. 15

Figure 2.6 Eroding coastal permafrost blust, Drew Point, Alaska. Photo Benjamin M. Jones, University of Alaska Fairbanks. 16

Figure 2.7 C17 Gas Emission Crater in the Yamal Peninsula as presented in Bogoyavlensky (2015). Photo by Andrey Umnikov. 17

Figure 2.8 Stratigraphic description in the Upper Cretaceous and Cenozoic of the Beaufort-Mackenzie Basin, foraminiferal zonation, approximated latitude of the basin at 10 Myr intervals, climate and oceans and major events as presented in Dixon et al. (2019) (copied with permission) . 18

Figure 2.9 Sketch of thermal regime on subsea permafrost areas. Adapted from Ruppel et al. (2016). 20

Figure 2.10 Main physiographic elements of the Canadia Beaufort Sea. Inspired from Dixon et al. (2019). 21

Figure 2.11 P-wave velocity for permafrost core samples as a function of temperature. Data taken from measurements on the 17 core samples presented in the work of King (1984). 25

Figure 2.12 P-wave velocity of brine-saturated coarse sand presented as a function of the estimated ice saturation as originally presented in Dou et al. (2016) 27

Figure 2.13 a) Schematic description of an offshore seismic survey on permafrost zones. A source (black circle) and an array of receivers (red triangles) are tied to a survey vessel. Seismic waves are emitted from the source. The receivers record the propagated waves, including direct waves, reflections (white lines), refractions (purple lines) and multiples (yellow lines). b) Example of a shot gather recorded on a subsea permafrost zone in the Canadian Beaufort Sea. 31
Figure 2.14 Map of ice occurrences in sediments, as presented in Pullan et al. (1987). 34
Figure 2.15 Stack section of a seismic survey collected in the Canadian Beaufort Sea over subsea permafrost zone. The top of the permafrost (TP) and the free surface multiples (M) are indicated in the Figure 35
Figure 2.16 Track lines from the ARA04C and ARA05C surveys used in this study. 37
Figure 2.17 Sketch of a NN with two hidden layers. σ represents the activation function applied over the weighted summation of the neuron input. . 38
Figure 2.18 Example of a fully convolutional encoder-decoder network as presented in Yang et al. (2016) (Open source) 43
Figure 2.19 Skipt connections used in the ResNet architecture 44
Figure 2.20 U-net architecture as presented in Ronneberger et al. (2015) 45
Figure 2.21 a) A basic recurrent neural network architecture in its compressed and unfolded representation. b) One-unite Long-Short-Term network (original images by F. Deloche under Creative Commons licence). 46
Figure 2.22 Standard GAN architecture as presented in Temizel et al. (2018). . . 46
Figure 2.23 Model architecture of the transformer NN as presented in Vaswani et al. (2017) 47
Figure 2.24 Sketch of convolutional (a) and pooling (b) layers as presented in Maier et al. (2019) 48
Figure 2.25 Output values for different activation functions. 52
Figure 3.1 Survey map of the ARA05C expedition. Line 06 discussed in this study appears in red. Shot point (SP) numbers increase towards the SE. Crystal cable velocity profiles are available in wells Kopanoar I-44 and M-13. 62

Figure 3.2 a) Stack section of line06 ARA05C survey indicating the top of the permafrost (TP) and the free surface multiples (M). Blue lines mark the position of shot gather used to infer velocity models. Example shot gathers are presented for locations SP 7900 (b), SP 8125 (c), SP 8201 (d) and SP 8275 (e). Guided P-waves (GW), reflection multiples (RE-M), refracted waves (RA) and its multiples (RA-M) interfere with primary reflections. Corresponding dispersion plots are presented in f-i. M0 to M5 are the different dispersion modes identified on f-i. . . . 64

Figure 3.3 The velocity model used for the sensitivity analysis. NFS refers to unfrozen sediments, PIBSI to partially ice-bonded silts, and IBSA to ice-bonded sand. 68

Figure 3.4 Effects on varying depth between 25 m and 100 m (a-d), thickness between 50 m and 150 m (e-h), velocity of the ice-bonded permafrost between 3.0 km/s and 4.5 km/s (i-l) and its Q attenuation factor from 7 to 1000 (m-p) on the modeled SP 8275. Guided P-waves (GW), reflection multiples (RE-M), refracted waves (RA), and its multiples (RA-M) are affected by variations in the depth, thickness, velocity of the ice-bonded permafrost and its Q attenuation factor 69

Figure 3.5 Dispersion curves for modeled shot 8275 when varying permafrost depth between 10 m and 150 m (a-d), thickness between 10 m and 150 m (e-h), ice-bonded permafrost velocity between 3 km/s and 4.4 km/s (i-l) and Q factor between 10 and 80 (m-p). Red lines mark the position of the inflection points and the cutoff frequencies at 10 m depth (a-d) at 10 m thickness (e-f), at 3.0 km/s velocity (i-l), and Q factor equal to 10 (m-p). 70

Figure 3.6 Effects of varying depth (a) and thickness (b) from 10 m to 150 m, and Q factor (c) from 10 to 80 for modeled shot 8275. 73

Figure 3.7 Normalized square sensitivities for depth (a), thickness (b), permafrost velocity (c) and Q factor (d) for the shot gather (blue) and the dispersion plot (orange) 74

Figure 3.8 Recorded, synthetic shot and corresponding velocity model for SP 7900 (a) located in a low ice content zone, models SP 7950 (b) and SP 8125 (c) located in a discontinuous ice-bonded zone, and models SP 8201 (d), and SP 8275 (d) located over continuous ice-bonded permafrost. In general, the modeled shots and dispersion plots reproduce the main features found in the observed data. 76

Figure 3.9 a) Resulting V_P model after coupling velocity analysis on shots gathers and dispersion plots on SP 7900, SP 7950, SP 8125, SP 8185, SP 8201, SP 8225, SP 8250 and SP 8275, and linear interpolation in intermediate shot position. b) Resulting velocity model after refraction travel time tomography. The dashed-dotted black line indicates water depths from bathymetry. 78

Figure 3.10 Recorded vs Synthetic shot gathers at intermediate SP extracted from the velocity model presented on Figure 3.9a. 80

Figure 3.11 Shot gathers in the $\tau - p$ domain for modeled shot 8275 (see Chapter 3) when varying permafrost depth between 10 m and 150 m (a-d), thickness between 10 m and 150 m (e-h), ice-bonded permafrost velocity between 3 km/s and 4.4 km/s (i-l) and Q factor between 10 and 80 (m-p). Red arrows mark the upper boundary of the conical shape of the arrivals, and red dashed lines show the cutoff velocity at 10 m depth (a-d), 10 thickness (e-h), at 3.0 km/s velocity (i-l) and Q factor equal to 10 (m-p). 83

Figure 3.12 Effects on varying depth between 10 m and 70 m (a-d), thickness between 10 m and 70 m (e-h), ice-bonded permafrost velocity between 3 km/s and 4.4 km/s (i-l) and Q factor between 10 and 80 (m-p) on the $f - p$ domain. Magenta arrows mark the cutoff velocity of the modes observed at 10 m depth (a-d), 10 m thickness (e-h), 3.0 km/s velocity (i-l), and Q factor equal to 10 (m-p) 85

Figure 3.13 Normalized square sensitivities for depth (a), thickness (b), permafrost velocity (c) and Q factor (d) for the shot gather (blue), the dispersion plot (orange), Radon (green) and Frequency Radon (red) images. . . 86

Figure 4.1 (a) Recorded data, and (e) synthetic data in the time-offset ($t - x$) domain. (b, f) Recorded and synthetic shots in the intercept-slope ($\tau - p$) domain. (c, g) Recorded and synthetic shots in the frequency-phase velocity ($f - c$) domain. (d, h) Recorded and synthetic shots in the frequency-velocity ($f - p$) domain. 93

Figure 4.2 P- (a) and S- (b) wave velocity and inverse quality factor ($1/Q$) distributions (c) of the samples in the 1D training dataset. 96

Figure 4.3 (a) Overview of the neural network (NN) architecture. (b) 1D encoder and decoder for single-input ($t - x$) data. (c) 2D encoder and decoder for single-input ($t - x$) data. "Conv" represents the combination of 2D convolutional and max-pooling layers, "TC" refers to the transpose convolution layer, and "enc out" denotes the output vector of the encoder. 100

Figure 4.4 (a) Total loss after training 64 similar 1D NNs. (b-d) Losses specifically evaluated for V_P, V_S, and $1/Q$ across the 64 trained NNs. An epoch represents one complete pass through the training data. 103

Figure 4.5 Total losses for the 1D and 2D datasets after training 64 similar NN. An epoch represents one complete pass through the training data. . . 104

Figure 4.6 Results of the 64 1D neural networks in the 1D testing dataset (grey lines). a), e), and i) present the results from the sample corresponding to the median error on the thicker water layer subset, b), f), and j) on the 100 m thickness subset, c), g) and k) on the 75 m subset, and d), h) and l) on the 50 m subset. Red lines denote the labels, whereas broken blue lines present the average value of the 64 predictions. The average L2 error is presented for each subset and true model. 105

Figure 4.7 L2 V_P (a), V_S (b), and $1/Q$ (c) error density distribution for the four subsets in the testing dataset. 107

Figure 4.8 True model and average prediction for the ensemble of 64 2D neural networks evaluated in the 2D testing dataset. The columns present the results for V_P, V_S and $1/Q$. The rows show, from top to down, random examples of the ticker layer subset (a), the 100 m subset (b), the 75 m subset (c) and the 50 m subset (d). The average structural similarity index (SSIM) and L2 error are presented for each subset and true model. 108

Figure 4.9 The evaluation of the NN on a synthetic dataset. a), d) and g) present the true V_P and V_S velocity and $1/Q$ models. b), e) and h) present the results after evaluating a NN with a single time-offset domain input. c), f) and i) present the results from the multi-input NN. The curves on the right compare the labels and the inverted (Inv.) results from the single-input ($t - x$) and multi-input NNs evaluated at 1 km, 4.5 km and 9 km. 109

Figure 4.10 Evaluating NN on seismic line ARA05C-6 recorded in the Beaufort Sea. Black dashed lines represent the location of shallow velocity measurements collected in the wells Kopanoar I44 and M13, located close to the seismic line. The plots on the right compare the well logs and the inverted V_P velocities. 111
Figure 4.11 Inverted V_P velocities for seismic lines ARA05C-06 (first column) and ARA05C-07 (second column) using the 1D (a) and the TL approach (b). The white dashed lines mark the intercept between the lines. The third column shows the inverted seismic parameters at the intercept. 113
Figure 4.12 Comparison between results from travel time first break tomography and the multi-input NN evaluated over the last 20 km of line ARA05C-6. The red, cyan and white arrows mark the distance at which changes in the trend velocities are observed. The dashed black line shows the interpreted base of the permafrost. 114
Figure 4.13 Correlation matrix in single domains for modeled shot 8275 with a permafrost layer at 57 m depth, of 150 m thickness, a velocity of 3.7 km/s and a Q factor of 60. 118
Figure 4.14 Correlation matrix using multiple domains for modeled shot 8275 with a permafrost layer at 57 m depth, of 150 m thickness, a velocity of 3.7 km/s and a Q factor of 60. 119
Figure 4.15 Evaluating single input NNs on a synthetic dataset from a wedge subsea permafrost structure. The columns display the label and inverted models using NNs training on the $t-x$, $f-c$, $\tau-p$ and $f-c$. The rows show the inverted V_P, V_S and $1/Q$ values. The last column represents the average models of all the inversions. 120
Figure 5.1 ARA05C survey lines and well locations. Red dots indicate the wells for which crystal cable survey is available 129
Figure 5.2 a) Summary of the multi-encoder multi-decoder approach as presented in Bustamante et al. (2023b). b) Architecture of the encoder that inputs multiple images in the CMP gather domain. c) Decoder architecture. 130
Figure 5.3 (a) Progression of the loss function during the training sessions for the 1D and 2D datasets. Processed and ground true (label) V_P, V_S and $1/Q$ on random samples of the 1D (b) and 2D testing datasets. . . . 133
Figure 5.4 Fence diagrams displaying the inverted V_P (a), V_S (b) and $1/Q$ (c) values using the NN for all seismic lines. Line locations are presented in (d) 134

Figure 5.5 Inverted V_P (first columm), V_S (second column), and $1/Q$ (third column) for the seismic lines ARA05C-11 (a), ARA04C-04 (b), and ARA05C-07 (c). White dashed lines show the position of the intersections of the line with other seismic lines. Comparisons between inverted parameters at line intersections are presented in Fig. 5.8. Black dotted lines show the closest position (245 m) to well K-59 in (a), where a velocity log is available. 135

Figure 5.6 Inverted V_P (first columm), V_S (second column), and $1/Q$ (third column) for the seismic lines ARA05C-17 (a), ARA04C-10 (b), and ARA04C-02 (c). White dashed lines show the position of the intersections of the line with other seismic lines. Comparisons between inverted parameters at line intersections are presented in Fig. 5.8. Black dotted lines show the closest position (580 m and 975 m) from wells B-35 in (a) and M-13 in (b), where velocity logs are available. 136

Figure 5.7 Inverted V_P for the seismic lines ARA04C-11 (a), ARA05C-08 (b), ARA05C-06 (c), ARA05C-16 (d), ARA05C-05 (e), ARA04C-09 (f), ARA05C-01 (g), ARA04C-01 (h), and ARA05C-03 (i). White dashed lines show the position of the intersections of the line with other seismic lines. Comparisons between inverted parameters at line intersections are presented in Fig. 5.8. Black dotted lines show the closest position (360 m, 75 m, 445 m and 440 m) to the wells K-59 (b), M-13 (c) and B-35 in (h, i), where velocity logs are available. 137

Figure 5.8 Inverted V_P (km s^{-1}) at the intersections of the seismic lines moving perpendicular (red) and parallel (green) to the shelf for the ARA04C and ARA05C surveys. Empty space in the figure means there is no intersection between lines. 140

Figure 5.9 Comparison of well logs K-59, M-13, and B-35 (represented by black dotted lines in Figures 5.5-5.7) with the corresponding V_P velocity models processed along the 500 m section closest to the wells. Seismic lines ARA05C-11 (a), ARA05C-08 (b), ARA05C-06 (c), ARA04C-10 (d), ARA05C-17 (e), ARA05C-03 (f), and ARA04C-01 (g) are used for comparisons. Note that the color scale of the sections matches the minimum and maximum values of the velocity scale of the well logs. . 141

Figure 5.10 Interpretation of subsea permafrost distribution based on the defined thresholds from the inverted V_P, V_S and $1/Q$ values. The inset shows lines locations. 144

Figure 5.11 Comparison between the permafrost conditions inferred by using NN versus ice occurrence in sediments as interpreted in Pullan et al. (1987) 145

Figure 5.12 Schematic of the general distribution of subsea permafrost in the Canadian Beaufort Sea as observed from the interpreted model. The left side of the figure is inspired by the conventional wedge-shaped permafrost distribution presented in Osterkamp (2001). 146

Figure 5.13 Stacked image (a) and Inverted V_P (b), V_S (c), and $1/Q$ (d) on 30 km of line ARA05C-16 recorded on the NW section of the Beaufort continental slope. 151

LIST OF SYMBOLS AND ACRONYMS

K	Hydraulic conductivity
GECs	Gas Emission Craters
PLFs	Pingo Like Features
RC	Reflection coefficient
TC	Transmission coefficient
Z	Seismic impedance
ρ	Density
$\mathscr{L}$	Loss function
y_p	Modeled data
y_l	Observed data
η_k	Step size or the learning rate in optimization
$\nabla\mathscr{L}$	Gradient matrix of the loss function
γ	Momentum parameter
σ	Activation function. Referred to the sigmoid activation function in the activation function description
$tanh$	Tanh activation function
ReLU	Rectified activation function
LReLU	Leaky ReLU activation function
ELU	Exponential linear unit
$S()$	Softmax activation function
δ	Dirac delta
ω_{ab}	Filter of the convolutional layer
NN	Neural Network
GAN	Generative adversal networks
RMS	Root mean square
FWI	Full Waveform Inversion
KOPRI	Korea Polar Research Institute
GSC	Geological Survey of Canada
SP	Shot position
CMP	Common Middle Point
TP	Top of the permafrost
M	Free surface multiples
RE-M	Reflection multiples

RA	Refracted waves
RA-M	Refracted multiples
RD	Amplitude decay with the offset of the refracted waves
GW	Guided waves
t,T	Time
x,X	Offset
f,F	Frequency
c,V	Phase velocity
M0,...,M4	Dispersion modes
V_P	P-wave velocity
V_S	S-wave velocity
Q	Attenuation quality factor
NFS	Unfrozen sediments
PIBSI	Partially Ice-Bonded Silts
IBSI	Ice-Bonded Silts
PIBSA	Partially Ice-Bonded Sands
IBSA	Ice-Bonded Sands
NFSH	Unfrozen shale
IBSH	Ice-bonded shale
d,t	Ice-bonded layer depth and thickness
C	Normalization constant in the sensitivity equation
TTFBT	Travel Time First Break Tomography
NSERC	Natural Science and Engineering Research Council of Canada
VMB	Velocity Model Building
TL	Transfer learning
τ	Time of intercept in the Radon transform
p	slope-velocity in the Radon transform
d_{100}^{1D}	Subset of the 1D training dataset with 100 m layer thickness
d_{75}^{1D}	Subset of the 1D training dataset with 75 m layer thickness
d_{50}^{1D}	Subset of the 1D training dataset with 50 m layer thickness
d_{deep}^{1D}	Subset of the 1D training dataset with a thickness water layer
d_{100}^{2D}	Subset of the 2D training dataset with 100 m layer thickness
d_{75}^{2D}	Subset of the 2D training dataset with 75 m layer thickness
d_{50}^{2D}	Subset of the 2D training dataset with 50 m layer thickness
d_{deep}^{2D}	Subset of the 2D training dataset with a thickness water layer
$t-x$	Shot gather (1D dataset) or CMP gather (2D dataset) domain

$\tau - p$	Radon domain
$f - c$	Dispersion plot domain
$f - p$	Image after applying a Fourier transform to the time dimension of the Radon domain
Conv	Convolution
TC	Transpose convolution
Ecn	Encoder
SSIM	Structural Similarity Index

CHAPTER 1 INTRODUCTION

Global warming have deep implications for the physical environment, ecosystems, and human societies. Over the past and current centuries, the Earth's temperatures have experienced a continuous rise (Pachauri et al., 2014). Among the regions most profoundly impacted by this warming trend, the Arctic stands out, as it has encountered a pronounced surge in surface air temperature, surpassing the global average by more than twice (Meredith et al., 2019; Notz and Stroeve, 2016; Signs, 2022). This increase in surface air temperature is accompanied by a corresponding rise in Arctic Ocean water temperature (Lind et al., 2018; Meredith et al., 2019; Polyakov et al., 2017; Signs, 2022). Extensive research has linked this temperature surge and anthropogenic activities, particularly the emission of greenhouse gases (Fyfe et al., 2013; Meredith et al., 2019; Najafi et al., 2015). Consequently, there is a high confidence level in projecting further warming in the Arctic region.

One of the prominent characteristics of the Arctic regions lies in the extensive distribution of permafrost (Zhang et al., 1999, 2000). Defined as any subsurface earth material that remains frozen for at least two years (Harris et al., 1988), permafrost holds a critical role in shaping hydrological and geological processes in the Arctic. The presence of ice exerts a dominant influence on the properties of permafrost soil and associated geological phenomena. Notably, when ice within permafrost reaches bonding conditions, it significantly reduces the hydraulic conductivity of the soil, thereby regulating the storage and movement of groundwater and giving rise to distinct hydrologic systems (Bring et al., 2016; Walvoord and Kurylyk, 2016). Additionally, the ice present in the pore space enhances the strength of permafrost soil (Fish et al., 1997) and thawing permafrost can create settling and destabilize built infrastructure.

Fluctuations in permafrost conditions, particularly changes in ice content and water flow patterns, lead to the formation of diverse permafrost-related features. These features arise from variations in volume due to ice formation or melting, soil instabilities caused by reduced strength following ice degradation, and the creation of gas and fluid pathways resulting from increased permeability after ice melting.

Moreover, permafrost, though primarily developing on land, extends over a considerable area beneath the seafloor of continental shelves in Arctic regions (Miner et al., 2022; Sayedi et al., 2020). This subsea permafrost originated at the surface during glacial periods when the continental shelves were subaerially exposed. As the glaciers underwent retreat and the sea level experienced a rise, these permafrost regions became submerged (Angelopoulos et al., 2020; Taylor et al., 2013). The marine transgression led to alterations in the thermal

conditions of the submerged permafrost, triggering a process of warming and degradation that persists to the present day (Sayedi et al., 2020).

Increasing temperatures in the Arctic accelerate permafrost degradation both on land and offshore. A significant concern associated with permafrost degradation is the release of carbon content from within and beneath the frozen layer. When onshore permafrost thaws, the organic material contained within becomes exposed, setting off a decomposition process facilitated by soil microbes (Schuur et al., 2015). Consequently, methane and carbon dioxide are emitted into the atmosphere. Conversely, subsea permafrost acts as a barrier, preventing the release of greenhouse gases from beneath (e.g., from gas hydrates) and within the permafrost.

Thus, permafrost degradation plays a significant role in the feedback loop of global warming. On one side, the escalating temperatures accelerate permafrost degradation and promote ice thawing within the pore space. This increases hydraulic conductivity, creating flow channels that enable water and gases to migrate to the surface. Paradoxically, these newly established flow paths can accelerate the release of a substantial amount of greenhouse gases trapped into and below the permafrost layer (Schuur et al., 2015), asseverating the effects of global warming. Hence, it is important to monitor the distribution and degradation rate of permafrost in both onshore and offshore regions, given its significant implications for global warming.

1.1 Measuring permafrost distribution

A range of direct and indirect methods are available for estimating and characterizing permafrost distribution. Notably, borehole logging presents a direct approach, wherein rock properties are directly measured to determine the vertical extent of permafrost. For instance, Hu et al. (2013) employed well seismic surveys and geophysical well logs to determine the base of permafrost in the Mackenzie Basin, while Ruppel et al. (2016) used a similar technique to characterize permafrost in the Canadian Beaufort Sea. While borehole logging remains a precise means to determine the vertical extent of permafrost, its capacity to resolve lateral details is limited. Moreover, extensive mapping is costly and impractical due to high uncertainties between borehole locations. In contrast, indirect geophysical methods estimate physical properties influenced by ice, such as electrical resistivity or seismic impedance. Those methods offer a more cost-effective approach to estimate the lateral extent of permafrost and characterize its condition.

Onshore, electrical and electromagnetic methods (e.g., those based on the solution of Maxwell's equations in different regimes) such as Electrical Resistivity Imaging (ERI), Electromagnetic

Methods (EM), Ground Penetrating Radar (GPR) have proven to achieve higher resolution for inverting permafrost distribution. ERI and EM techniques rely on electrical resistivity/conductivity measurements of the rock. In permafrost regions, the presence of ice leads to an elevated soil resistivity. This phenomenon is attributed to the exclusion of free ions in the water during ice formation. The electrical response primarily occurs in the unfrozen portion of the pore fluid (i.e., water). As the increase in ice content reduces the pore space fraction occupied by the pore fluid, the resulting resistivity displays increased values.

Recent examples of ERI techniques applied to map onshore permafrost include the works of Tomaškovičová and Ingeman-Nielsen (2023), who quantified the freeze-thaw hysteresis of unfrozen water in West Greenland, Wu et al. (2023), who monitored permafrost distribution and degradation for the safety of infrastructures near the China-Russia Oil Pipeline, and Tourei et al. (2023), who investigated the spatial variability of permafrost and its degradation in undisturbed tundra in Utqiagvik, Alaska.

While ERI acquires subsurface information by introducing a direct current into the earth and analyzing the resultant potential distribution, EM methods employ a time-varying magnetic field as a source, inducing electrical currents within the ground through electrical induction. Numerous publications have also proposed using EM methods for mapping onshore permafrost distribution. Nikitenko et al. (2023) developed a mathematical model for monitoring the state of permafrost using EM methods, while Thomas et al. (2023) employed electromagnetic induction measurements to map permafrost degradation in Stordalen mire, Sweden. Additionally, Buddo et al. (2022) analyzed the applicability of transient electromagnetic (TEM) surveys to identify Northern Siberia permafrost structures on continuous, mostly continuous and discontinuous permafrost zones. Sherman (2017) employed a controlled source electromagnetic method in offshore areas to collect data on the depth, thickness, and horizontal extent of shallow-water marine permafrost near Prudhoe Bay, Alaska.

Ground Penetrating Radar (GPR) is another electromagnetic technique extensively employed for assessing permafrost distribution on land. It involves the transmission of high-frequency electromagnetic pulses into the subsurface. As these pulses encounter variations in the electromagnetic properties of the subsurface, they give rise to reflections, refractions, and diffractions, which are subsequently captured by a sensor at the Earth's surface (Heagy et al., 2016). Permafrost regions exhibit a substantial contrast in dielectric permittivity between frozen and unfrozen sediments, leading to distinct reflections facilitating the interpretation of permafrost distribution. Recent examples of GPR surveys applied to onshore permafrost investigations are showcased by Forte et al. (2022), who used this method to survey permafrost-associated orthogonal polygons in northern Victoria Land (Antarctica), Qi et al. (2023), who studied

permafrost distribution and embankment damages along the Qinghai-Tibet Highway, and Greer et al. (2022) who monitored thermokarst sinkholes in Fairbanks, Alaska.

Kneisel et al. (2008) provides a comprehensive overview of the strengths and limitations of electrical and electromagnetic methods for investigating permafrost. One key factor influencing the depth of exploration is the instrument's geometry in the case of ERI techniques, and for EM methods, it is determined by factors such as instrument geometry, frequency, resistivity of the upper layer, and loop size. The effectiveness of GPR investigations, on the other hand, hinges on soil attenuation and survey frequency. However, it is important to note that GPR techniques are generally less suitable for extensive permafrost mapping due to their limited penetration depth, which typically remains below 20 meters. Similarly, ERI methods come with certain constraints related to electrode placement, as a good electrode-terrain contact is crucial and can influence the accuracy of results. Electric and electromagnetic methods have demonstrated their efficiency in characterizing and ascertaining permafrost distribution onshore. However, their resolution for inverting the resistivity of subsea permafrost zones remains limited. The principal drawback of employing electromagnetic methods to determine the extent of subsea permafrost lies in their low resolution at depth in offshore environments. This limitation is attributed to the limitations of the inversion process and the substantial electromagnetic attenuation caused by the presence of highly conductive water column.

1.2 Problem identification

Seismic surveys have proven to be a reliable tool for measuring the lateral extent of subsea permafrost. By exploiting the higher seismic velocity of ice in comparison to seawater, ice-bearing subsea permafrost can be differentiated from ice-free sediments. In fact, laboratory measurements have demonstrated that seismic velocities increase with higher levels of ice saturation within the same type of rock (Dou et al., 2016; King, 1984).

Seismic methods analyse different types of arrivals to identify the structures under the seabed. Among these methods, seismic refraction examines the refracted arrivals and can be used to determine the lateral distribution of subsea permafrost structures. Notably, the most extensive investigation of permafrost distribution in the Beaufort Sea to date takes advantage of the refracted arrivals to generate a map of permafrost distribution, which has been presented in two independent publications (Hunter et al., 1976; Pullan et al., 1987). Seismic refraction has demonstrated its efficiency in assessing the lateral distribution of subsea permafrost and in identifying ice-bonded permafrost layers that are separated by unfrozen layers. However, given the absence of critical refractions at the base of the permafrost, this technique cannot provide vertical information regarding the subsea permafrost distribution.

Similarly, conventional seismic reflection methods rely on imaging reflected arrivals. Seismic reflection, in theory, provides an advantage over seismic refraction as it can get information about the thickness of the permafrost layer from the reflection at the permafrost base. However, in permafrost regions, highly energetic guided waves and free surface multiples lead to interference with reflections from deeper structures and, more critically, from the base of the permafrost (Duchesne et al., 2022). This substantial interference between guided waves, free surface multiples, and reflections poses a significant challenge to applying conventional seismic processing techniques to identify subsea permafrost distribution accurately (Kang et al., 2021).

State-of-the-art seismic processing techniques, such as Full Waveform Inversion (FWI), use all recorded arrivals in an optimization methodology to identify subsea structures. FWI compares synthetic seismic data from a proposed model of subsea structures with the recorded data, progressively minimizing a cost function until reaching the best approximation or until no further improvements in the model can be achieved (Virieux and Operto, 2009). However, due to the pronounced heterogeneities observed in permafrost zones, establishing an initial model sufficiently close to the true model for achieving the best approximation becomes challenging.

Despite the limited capacities of different seismic processing techniques in characterizing and fully determining permafrost distribution, the recorded seismic waves contain valuable information about the vertical and lateral subsea permafrost distribution (Duchesne et al., 2022). For instance, the slope of refracted waves can tell us about the velocity of the permafrost layer, while the point where it intercepts with the zero offset can provide information regarding the depth of the permafrost layer. Similarly, the hyperbolic shape of reflections at the top and bottom of the permafrost layer and their intercepts with the zero offset can give us information about the velocity, depth and thickness of the permafrost layer. Additionally, the characteristic modal distribution of guided waves can provide information about the velocity of the permafrost layers and the depths of the top and base of the permafrost. However, the existing processing methods fail to use that information.

In recent years, deep learning has emerged as an alternative tool for processing seismic data (Araya-Polo et al., 2018). Deep learning could potentially use non-conventionally processed arrivals such as guided waves and multiples, along with reflections and refractions to better resolve the subsurface. One key advantage of deep learning over other seismic processing techniques is that it avoids the need for separating different arrivals or defining a precise initial model as required by Full Waveform Inversion. Instead, deep learning leverages the recorded data to generate a direct mapping between the seismic data and subsurface structures. Once

a deep neural network is trained, it can quickly and easily process large seismic surveys, on the alignment of survey features with those used during the training phase. This makes deep learning seismic inversion a promising technique for imaging permafrost distribution on a large scale and potentially resolving permafrost features and structures we could not previously see. However, despite its potential, no attempts have been made to use deep learning to map subsea permafrost distribution.

1.3 Research Objectives

The general objective of this thesis is to develop a methodology based on deep learning to image seismic velocity variations caused by subsea permafrost for effective large-scale characterisation and mapping. To achieve this, we have structured our investigation around three specific objectives.

The first objective is to evaluate the potential of seismic surveys to image the velocity distribution of subsea permafrost. More precisely, we seek to determine if the seismic data are sensitive to variations in permafrost depth, thickness, and changes in seismic properties such as P- and S-wave velocities and attenuation. Achieving this objective will help us to evaluate the feasibility of using deep learning to image seismic velocity variations in subsea permafrost regions. This topic is discussed in detail in Chapter 3 and the results have been published in the journal *Geophysics*.

- Bustamante, J., Fabien-Ouellet, G., Duchesne, M.J. and Ibrahim, A., 2023. Understanding seismic velocity variations of subsea permafrost: A sensitivity study. Geophysics, 88(5), R655-R669.

The second objective of our research is to design a deep-learning methodology for estimating the distribution of P- and S-wave velocity and attenuation from raw seismic data. Current deep-learning methods for inverting P- and S-wave velocities from seismic data are not suitable for permafrost zones due to high interference between different arrivals commonly observed in the recorded seismic data. To address this issue, we propose a novel multi-input, multi-output neural network that assimilates seismic image data from four distinct domains and yields the inverted seismic values at depth. Fulfilling this objective will enable us to image seismic velocity variations in subsea permafrost regions using deep-learning, which is a crucial step towards achieving the main objective of our thesis. This topic is discussed in detail in Chapter 4. This work has been submitted in the journal *Geophysics*, and is currently under revision:

- Bustamante, J., Fabien-Ouellet, G., Duchesne, M.J. and Ibrahim, A., 2023. Deep-learning viscoelastic seismic inversion for mapping subsea permafrost. Geophysics, *Submitted*

The third objective of this study is to assess the distribution of permafrost in the Canadian Beaufort Sea. This is the subject of Chapter 5. To achieve this objective, our deep learning approach is applied to a series of seismic lines acquired during two expeditions in 2013 and 2014. The details of this investigation were submitted to the Journal of Geophysical Research: *Solid Earth* as:

- Bustamante, J., Fabien-Ouellet, G., Duchesne, M.J. and Ibrahim, A., 2023. Mapping subsea permafrost distribution in the Beaufort Sea with marine seismic and deep learning. Solid Earth, *Submitted*

1.4 Dissertation structure

This thesis dissertation is divided as follows:

- Chapter 1: Introduces the research project by delineating the research problem and outlining the dissertation's objectives.
- Chapter 2: This section presents a literature review of the fundamental concepts investigated within the dissertation. This review aims to offer readers a better understanding about the nature of permafrost and its associated features, the application of seismic techniques in mapping subsea permafrost, and a general overview of the principles underlying deep learning and neural networks. The content of this chapter mirrors the present state of understanding and ongoing research directions in these subject areas.
- Chapter 3: Presents the first article of this thesis, entitled *"Understanding seismic velocity variations of subsea permafrost: A sensitivity study."* This article justifies using a multi-input approach for inverting seismic parameters in permafrost zones. This is achieved by conducting a sensitivity analysis on the capacity of seismic data to retrieve information about permafrost geometry and its seismic parameters.
- Chapter 4: Introduces the second article of this thesis entitled *"Deep-learning viscoelastic seismic inversion for mapping subsea permafrost."* This article presents the data-based inversion approach proposed for mapping seismic properties in permafrost regions.

- Chapter 5: Presents the third article of this thesis, entitled *"Mapping subsea permafrost distribution in the Beaufort Sea with marine seismic and deep learning."* In this article, we assess permafrost distribution from two surveys conducted in the Canadian Beaufort Sea with the proposed data-based inversion approach. The results are used to propose a map of subsea permafrost distribution in the region.

- Chapter 6: Presents the general conclusion of this thesis and succinctly summarizes its contributions to geophysics and permafrost research. It also discusses areas for further research stemming from the findings presented in this thesis.

CHAPTER 2 LITERATURE REVIEW

Developing a methodology based on deep learning that can image seismic velocity variations caused by subsea permafrost is a challenging task that requires knowledge from various fields of study. These fields include geology for understanding the processes associated with permafrost distribution, seismic processing for comprehending the effects of permafrost presence on seismic surveys, and deep learning for designing a suitable methodology that can tackle this task effectively. This chapter gives an overview of these fields. We begin by explaining what is permafrost, focusing on both its traditional definition and the geological features linked with onshore and subsea permafrost. Then, we review how seismic data can be used to map subsea permafrost. We describe the seismic properties of permafrost, explain the basics of wave propagation in permafrost, and discuss the conventional methods to process seismic data in permafrost regions, along with their associated challenges. We present the current state of knowledge on the distribution of subsea permafrost in the Beaufort Sea. We stress the importance of a quick approach to handle seismic data for analyzing permafrost distribution on a large scale. Finally, we introduce the main concepts of deep learning and its application to seismic inversion. We begin by explaining the fundamental mechanisms of deep learning. Next, we outline the optimization methods used in deep learning training. Then, we provide an overview of various neural network architectures and components. To conclude, we review the applications of deep learning to seismic data analysis and highlight the main challenges.

2.1 Permafrost

2.1.1 Global permafrost distribution

Permafrost, defined as subsurface Earth material that remains frozen for a minimum duration of two years at temperatures below 0 ^{o}C (Harris et al., 1988), exhibits a widespread distribution across the Arctic and boreal regions of the Northern Hemisphere (Figure 2.1). Permafrost covers approximately 30 million km^2 of the Earth's surface across various regions. The Arctic region alone is estimated to contain 13–18 million km^2 of permafrost, with an additional 1.06 million km^2 on the Tibetan Plateau and a significant portion of 16–21 million km^2 submerged under the sea (Miner et al., 2022). In the Arctic regions, permafrost covers approximately 23% of the exposed land surface (Zhang et al., 1999, 2000). Beyond these regions, permafrost can also be found in high mountain areas, subantarctic islands, and the

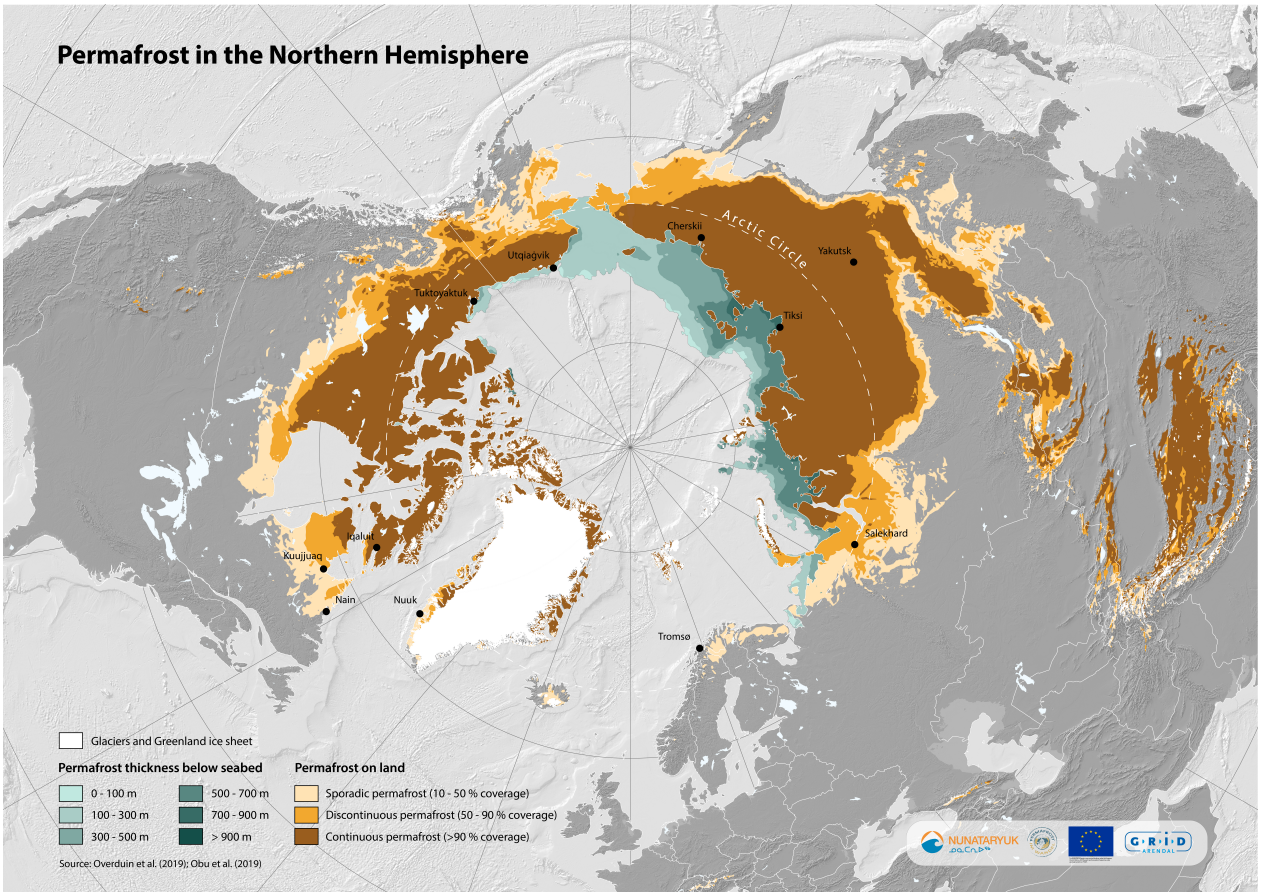


Figure 2.1 Permafrost in the Northern Hemisphere. Source: GRID-Arendal/ Nunataryuk (`https://www.grida.no/resources/13519`).

Antarctic continent (Schuur et al., 2008).

Global warming is causing permafrost degradation to accelerate, especially in the Arctic (Meredith et al., 2019). The temperature increase contributes to the thawing of ice within the pore spaces, resulting in the alteration of various soil properties, such as volume, cohesion, rock strength, and permeability. These changes in permafrost condition can have significant implications for land infrastructures, including the occurrence of landslides (Patton et al., 2021, 2019), land subsidence, and the formation of thermokarst bogs (Heffernan et al., 2020).

Furthermore, the reduced permeability of ice-rich permafrost leads to the development of intricate hydrologic systems by influencing the movement and fluxes of water (Bring et al., 2016; Patton et al., 2019; Walvoord and Kurylyk, 2016). Consequently, permafrost degradation has a profound impact on the hydraulic permeability of permafrost, resulting in modifications to hydrologic systems. Additionally, the increased permeability permits the migration of greenhouse gases stored within and beneath the permafrost into the atmosphere (Schuur

et al., 2015).

2.1.2 Permafrost formation and the associated thermal regime

In regions where the average yearly air temperature is below 0°C, some of the soil that freezes during winter remains frozen throughout the summer. Consequently, a permafrost layer develops and extends deeper each year from the initially frozen ground. This expansion of the permafrost layer occurs incrementally with each winter, depending on the equilibrium between underground heat transfer and outward heat flow into the air. This balance is influenced by the average annual air temperature and the geothermal gradient (Péwé, 2015).

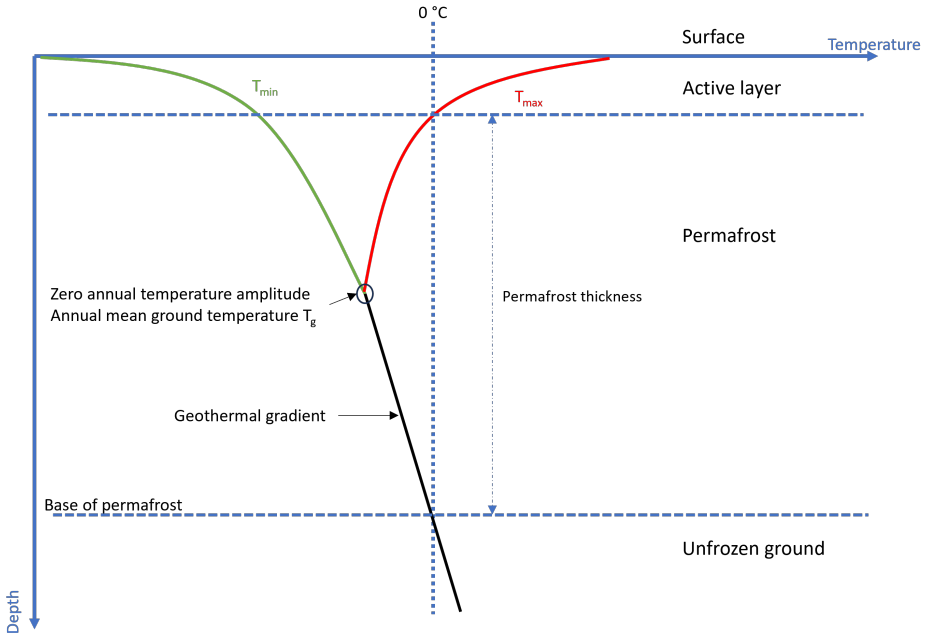


Figure 2.2 Sketch of the thermal regime on permafrost areas. Adapted from Allard et al. (2012).

Figure 2.2 illustrates temperature variations with depth in permafrost zones. Near the surface, temperatures undergo seasonal fluctuations, with warmer temperatures in summer and colder temperatures in winter. These variations result in distinct temperature profiles at different depths. Within the figure, the green line depicts a standard temperature profile representing the minimum observed surface temperature, while the red line signifies the temperature profile associated with the maximum observed surface temperature.

The difference in temperature between these profiles diminishes gradually as depth increases until zero annual temperature amplitude is reached. At this junction, the discrepancy between the two profiles becomes negligible. The temperature at this specific depth corresponds

to the annual average ground temperature of the permafrost (T_g). Beyond this depth, the thermal profiles merge following the geothermal gradient. Subsequently, the temperature gradually rises until reaching 0^o C at the base of the permafrost.

2.1.3 Structures associated with permafrost

Onshore and subsea permafrost can be defined in terms of thickness and geographical extent. The thickness of permafrost depends on its upper and lower boundaries (Fig. 2.2), which are the permafrost table and the base of the permafrost, respectively. These boundaries mark the transition from temperatures below or equal to 0^oC to positive values. Moreover, the permafrost limit indicates the geographical area where permafrost exists. Various factors have the potential to influence the distribution of permafrost. For example, more intense winters can result in the expansion of permafrost thickness, whereas rising global temperatures can impact its spatial distribution. Consequently, we use the terms "permafrost aggradation" for the increase in thickness or area of permafrost and "permafrost degradation" for the decrease in thickness or area of permafrost (Harris et al., 1988).

Continuous permafrost is the term used to describe the occurrence of permafrost throughout an entire geographic region beneath the exposed surface. On the other hand, when permafrost is present in certain areas beneath the land surface while other areas remain free of permafrost, we categorize it as a discontinuous permafrost region (Van Everdingen, 2005). The discontinuous permafrost zone is divided into three subzones (Burn, 2011), widespread discontinuous (50-90% of permafrost), sporadic discontinuous (10-50% of permafrost), and isolated patches of permafrost (less than 10%). The boundaries between zones are not sharp but gradual transitions.

Permafrost occurrence is influenced by three main factors. First, the annual ground temperature is the primary determinant of the presence or absence of permafrost and is shaped by climate-related aspects such as the annual air temperature and the extent of snow cover. Second, the latitude as it impacts the strength of the solar radiation that reaches the ground surface. And third, continentality as it is related to the increasing aridity as we move away from the ocean.

In permafrost zones, the alterations in permafrost conditions, specifically varying ice content and water flow patterns, give rise to different permafrost-related features within permafrost areas. These features emerge due to volume fluctuations caused by ice formation or melting, soil instabilities resulting from diminished strength following ice degradation, and gas and fluid conduits formed due to enhanced permeability after the ice melts. Figure 2.3 presents a schematic diagram illustrating the features associated with both onshore and subsea per-

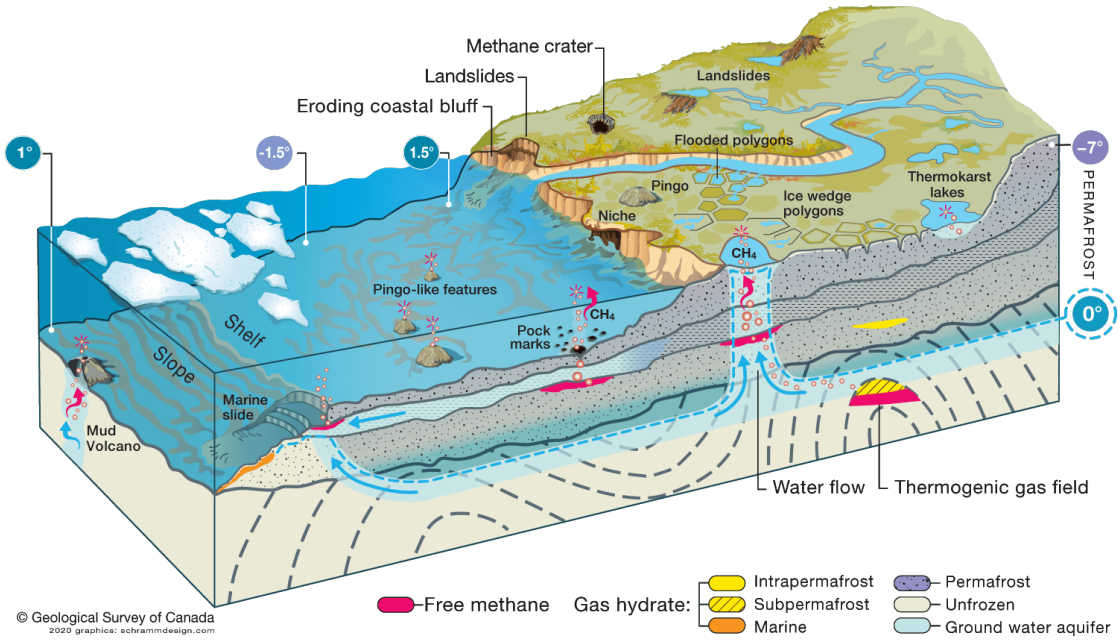


Figure 2.3 Schematic representation of onshore and subsea permafrost distribution and associated geological structures. Source: Geological Survey of Canada.

mafrost.

Taliks

Discontinuous permafrost regions exhibit a significant characteristic known as taliks, which refer to areas of unfrozen ground within the permafrost zones (Harris et al., 1988). Taliks can manifest in different forms depending on their relationship with the surrounding frozen ground. They may exist as open taliks fully penetrating the permafrost, isolated taliks enclosed by perpetually frozen terrain, or closed taliks occupying depressions beneath lakes or rivers in the permafrost table.

Active layer

Onshore, above the permafrost, a layer of soil undergoes seasonal freezing and thawing, known as the active layer (Figure 2.2). This layer is dominant in the geological and hydrological processes occurring above the permafrost. The thickness of the active layer is influenced by the variability of climatic conditions and the geological properties of the lithosphere (Dobiński, 2020). Its thickness can range from centimeters to $\sim$ 25 meters. Acting as a shallow aquifer, the active layer influences the runoff patterns and surface water response to precip-

itation and snowmelt (Bring et al., 2016; Patton et al., 2019; Yamazaki et al., 2006). As temperatures rise, the active layer thickness expands, leading to a larger portion of the subsurface experiencing seasonal groundwater flow, significantly impacting the local hydrologic system (Walvoord and Kurylyk, 2016). Permafrost degradation triggers short-term surges in groundwater saturation and local streamflow (Patton et al., 2019), resulting in enduring effects on groundwater dynamics.

Thermokarst lakes, ice wedge polygons and flooded polygons

Thermokarst refers to any geological process that gives rise to distinct landforms due to the thawing of ice-rich permafrost (Harris et al., 1988). One notable manifestation of thermokarst is the formation of thermokarst lakes illustrated at the right in Figure 2.3. Thermokarst lakes emerge as enclosed depressions resulting from ground settlement caused by the melting of ground ice. The frozen surroundings of these lakes impede water infiltration and exchange, rendering the hydrological balance significantly reliant on factors such as snowmelt, evaporation, precipitation, and water inflow from local permafrost soils (Vonk et al., 2015).

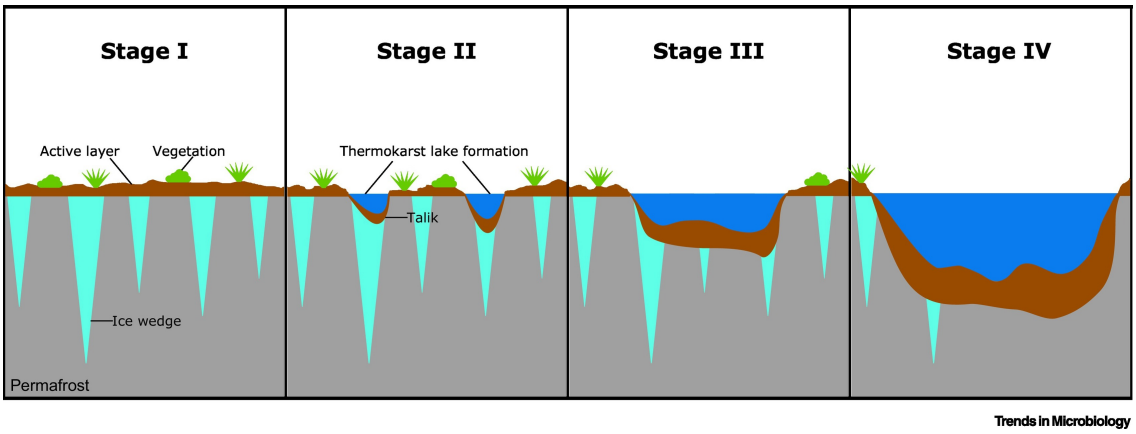


Figure 2.4 Thermokarst lake formation as presented in Liebner et al. (2020).

In areas characterized by continuous permafrost, the initiation of thermokarst lakes (Figure 2.4) typically begins with merging ice-wedge polygons to create pools above melting ice-wedge networks (Bouchard et al., 2016). An ice wedge (located to the left of thermokarst lakes in Figure 2.3) is a fissure in the ground resulting from a slender or thin ice formation, typically measuring up to 3 meters at the surface and extending downward for several meters (Harris et al., 1988). In regions characterized by discontinuous permafrost, ice lens melting induces surface subsidence and the formation of water bodies within topographic depressions (Bouchard et al., 2016). If the permafrost thaws completely, a mature thermokarst pond or

lake, enclosed by a peripheral ridge, can achieve stability if there is a layer of impermeable silts and clays beneath it.

Terrestrial pingos

Figure 2.5 Pingos in the Mackenzie Delta. Photo released as public domain by Emma Pike.

Pingos (Figure 2.5 and located close to the coast in figure 2.3) consist of permanent mounds of frozen soil characterized by a central core composed of massive ice (Harris et al., 1988). The formation process of pingos involves water injection followed by subsequent coverage with soil and vegetation (Mackay, 1998; Paull et al., 2022). The characteristics of pingos can vary depending on the mechanism that pressurizes the groundwater, resulting in two distinct types: open systems (hydraulic) or closed systems (hydrostatic) (Yoshikawa, 2013). Hydraulic system pingos develop at locations where intra-permafrost or sub-permafrost groundwater reaches the surface (French, 2017). In contrast, hydrostatic system pingos occur due to pore water expulsion instigated by permafrost aggradation in water-saturated sediments. When the ice core within a pingo melts, the collapse of the structure gives rise to a shallow-rimmed depression.

Eroding permafrost coast and erosional niche formation

Permafrost degradation also plays a crucial role in intensifying coastal erosion (Figures 2.6, 2.3). In the Arctic region, this phenomenon is primarily attributed to the decreased ice content, which directly reduces the cohesion of the surrounding sediments (Lantuit et al., 2013). While permafrost containing abundant ice exhibits some resistance against mechanical wave forces, its erodibility experiences a significant increase upon thawing (Lim et al., 2020).

Figure 2.6 Eroding coastal permafrost blust, Drew Point, Alaska. Photo Benjamin M. Jones, University of Alaska Fairbanks.

The destabilization of the permafrost coastline leads to the substantial release of particulate organic carbon into the Arctic coastal regions (Jong et al., 2020), affecting the sediment and organic carbon mobilization and carbon pathways.

Permafrost landslides

Likewise, permafrost degradation bears considerable implications for the cohesion, friction, and groundwater dynamics of near-surface materials, thereby exerting a substantial influence on landslide processes (Patton et al., 2019). Among the permafrost-related landslides (Figure 2.3), active-layer detachment slides are particularly interesting in the Arctic. These landslides are shallow-seated translational movements that transpire within the seasonally thawed layer above permafrost (Lewkowicz and Harris, 2005; Patton et al., 2021). As the ice melts at the base of the active layer, the loss in ice and subsequent rise in pore water pressure become evident. The reduced cohesion, heightened water saturation, and diminished friction enable landslides to occur on low-angle hillslopes.

Gas emission crater (GECs)

A novel phenomenon associated with permafrost features is the emergence of gas emission craters (GECs) (Figures 2.7, 2.3). These intriguing structures were first documented in 2014 on the Yamal Peninsula (Bogoyavlensky, 2015). Since then, numerous GECs have been reported in the Yamal and Gydan peninsulas (Chuvilin et al., 2020; Dvornikov et al., 2019).

The origin of GECs is linked to explosive gas emission events, but the precise source of the gas remains a subject of ongoing discussion. Chuvilin et al. (2020) propose a compelling

Figure 2.7 C17 Gas Emission Crater in the Yamal Peninsula as presented in Bogoyavlensky (2015). Photo by Andrey Umnikov.

hypothesis suggesting that GECs develop from thermokarst lakes with underlying taliks. As the lake gradually dries up, the talik becomes frozen, accumulating confined pressure. Within the frozen talik, gas, water, and soil undergo stratification during the freezing process until the cap above the talik is breached due to the buildup of pressurized gas. This sudden collapse results in an active eruption of gas, water, and soil, initiating the formation of the crater. Notably, towards the culmination of this process, the resulting crater may eventually be refilled with water, giving rise to the formation of a new lake.

Gas hydrates

Gas hydrates are crystalline solids formed through the association of water molecules and gas molecules. The gas molecules, commonly called "guests," are effectively confined within water cavities, known as "hosts," which consist of hydrogen-bonded water molecules. Notably, the guest gases in gas hydrates are typically composed of methane, ethane, propane, and carbon dioxide, (Sloan Jr and Koh, 2007). Gas hydrates are prevalent worldwide in settings that bring gas and water together in areas with suppressed subsurface temperatures (Boswell et al., 2020).

Gas hydrates related to permafrost can form within the pressure-temperature range in the lower permafrost (intra-permafrost) and the upper sedimentary layer beneath the ice-bearing permafrost base (Dallimore and Collett, 1995; Ruppel, 2015). These gas hydrates remain stable in a zone extending from the base of the permafrost to a considerable depth of several hundred meters below it. The degradation of permafrost, which modifies the prevailing pressure and temperature conditions, can induce gas release by destabilizing gas hydrates

situated within and below the permafrost layers.

Stratigraphy of the Beaufort Shelf

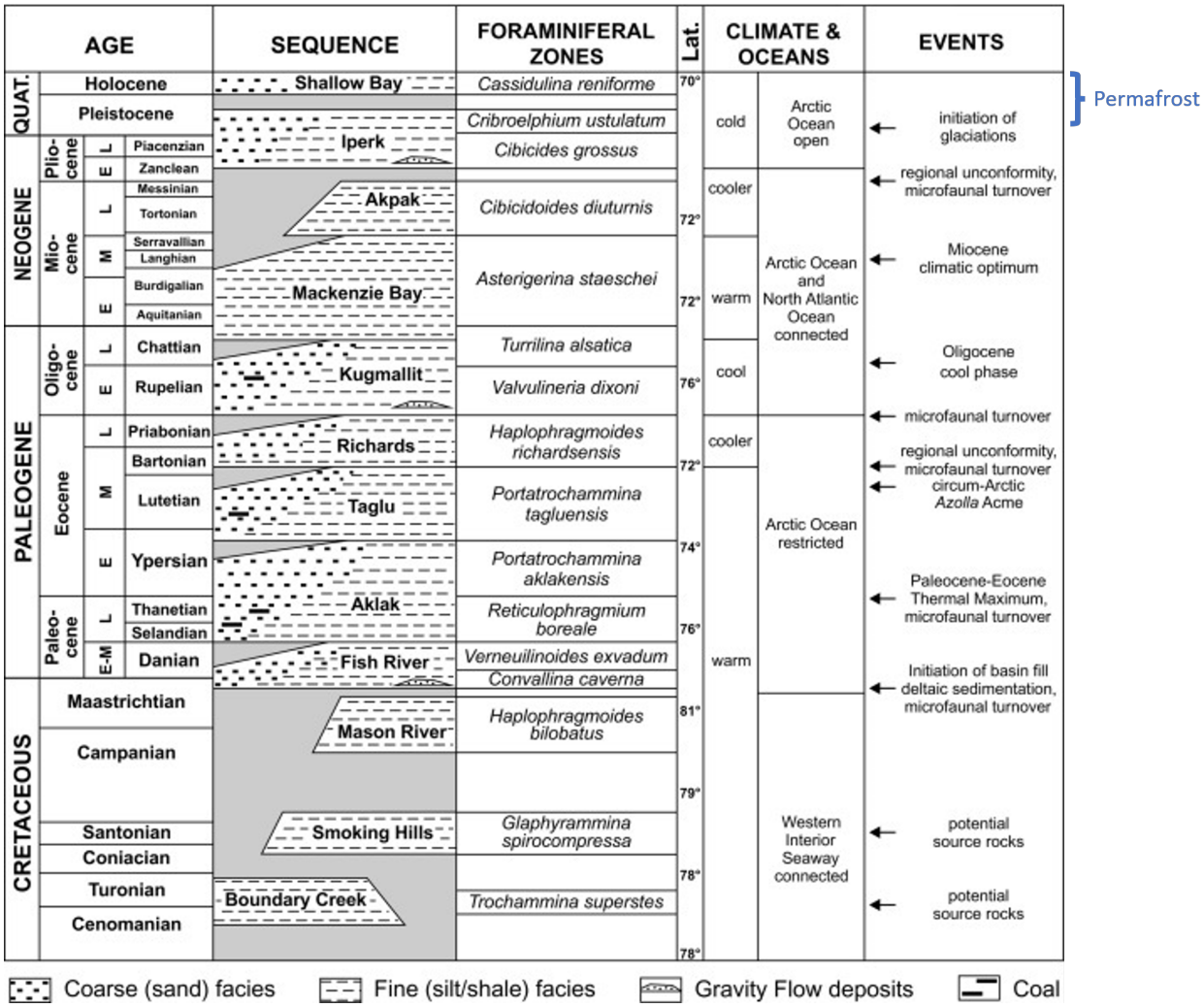


Figure 2.8 Stratigraphic description in the Upper Cretaceous and Cenozoic of the Beaufort-Mackenzie Basin, foraminiferal zonation, approximated latitude of the basin at 10 Myr intervals, climate and oceans and major events as presented in Dixon et al. (2019) (copied with permission)

The stratigraphy units found on the Beaufort shelf from the upper Cretaceous consist of various sequences (Figure 2.8). These include the early Pleocene Fish River sequence, which contains submarine fan deposits, sandstones, conglomerates, mud deposits and fluvio-deltaic strata. Followed by the late Paleocene to early Eocene Aklak sequence and the medium to late Eocene Taglu and Richards sequences, which encompass interbedded sandstone, siltstone, and shale sections with some coal beds. The Oligocene Kugmallit sequences, which contain the

Kopanoar Sequence submarine fan deposits, overlain by a sequence of progradational deposits. The Miocene Mackenzie Bay sequence and the Akpak sequences which are composed by shale with subordinate beds of siltstone and sandstone. The late Pliocene and early Pleistocene Iperk sequences formed by a prograding wedge of fluvio-deltaic sediments containing weakly consolidated to unconsolidated sandstone and conglomerate with larger contents of shale basin-ward. And finally, the uppermost Quaternary Shallow Bay sequences (Dietrich et al., 1985; Dixon et al., 2019; Sippel et al., 2013; Young et al., 1976).

Dixon et al. (2019) describe the stratigraphic history as follows: From the late Maastrichtian to the Pliocene period, sedimentation in the Beaufort-Mackenzie Basin was primarily dominated by the deposition of large delta complexes. These delta complexes gave rise to significant transgressive-regressive sequences (Fish River, Aklak, Taglu, Richards, Kugmallit, Mackenzie Bay, and Akpak sequences). The Fish River was the first major deltaic complex to form in the basin, composed of submarine fan deposits, sandstones, conglomerates, mud deposits and fluvio-deltaic strata. Subsequently, the deposition process can be divided into three distinct phases.

The first phase concluded during the middle Eocene when an unconformity developed between the Taglu and Richards sequences. This phase is characterized by abundant and well-developed coals in the delta-plain deposits of the Aklak sequence.

The second phase lasted until the late Miocene era, during which the Richards, Kugmallit, Mackenzie Bay, and Akpak sequences were deposited. In the late Eocene period, a significant drop in base level caused the erosion of the Richards sequence and exposed the Eocene shelf to erosion from the land. Following this, there was a rise in base level and progradation, leading to the formation of the Kugmallit sequence, which is approximately 4000 m thick and covers a substantial area beneath the central Beaufort Shelf. During the Miocene sedimentation (Mackenzie Bay and Akpak sequences), sediment volume was notably decreased, resulting in less extensive delta complexes forming.

Towards the end of the Miocene epoch, a significant erosional unconformity emerged, marking the peak of another tectonic event. The Pliocene period witnessed the development of a thick succession of prograding sediments, shaping the modern shelf geometry known as the Iperk sequences. This phase lacked deformation and saw a considerable shift in the locus of deposition towards the interior of the basin. The late Miocene unconformity also indicates a major climate shift between the Miocene and Pliocene, with a substantial cooling of the climate during this time.

The quaternary deposits forming the upper portion of the continental shelf of the Canadian Beaufort Sea comprise a sedimentary wedge >800 m-thick.These sediment layers consist of

abundant clastic material, including thick, fine sand with fresh pore water sourced from glaciofluvial origins (Blasco et al., 2012).

2.1.4 Subsea Permafrost

Subsea permafrost is a prevalent feature found on the continental shelves of polar regions. Subsea permafrost initially formed at the surface during glaciations when the shelves were exposed. It became submerged as the glaciers retreated and the sea level rose (Angelopoulos et al., 2020; Taylor et al., 2013).

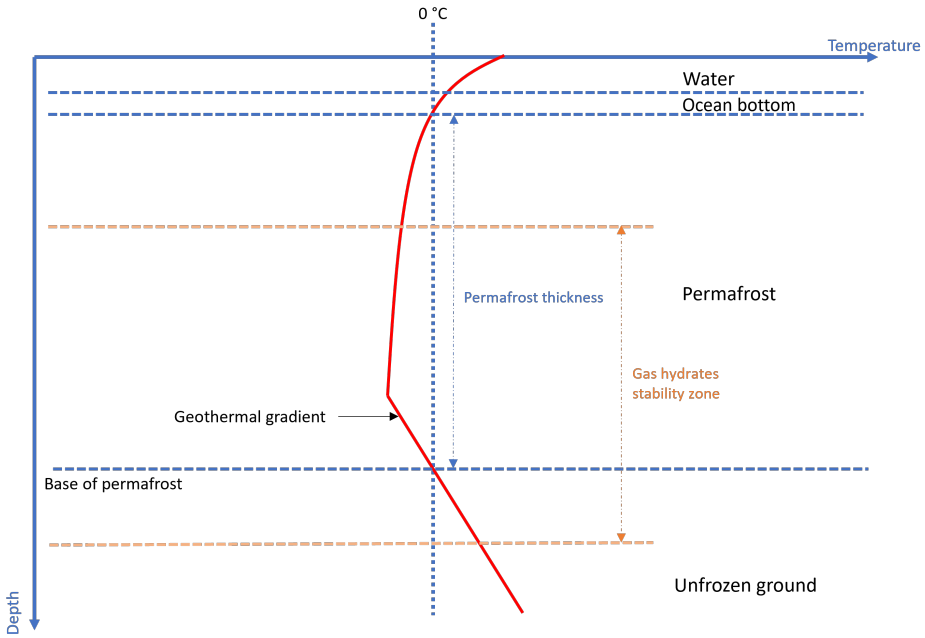


Figure 2.9 Sketch of thermal regime on subsea permafrost areas. Adapted from Ruppel et al. (2016).

Although subsea permafrost remains unaffected by seasonal thawing and freezing (hence lacking an active layer), it gradually degrades. This degradation is attributed from top to bottom to the heat transfer between the warmer overlying sea layer and the frozen sediments (Romanovskii et al., 2005) and from the bottom to the top following the temperature changes associated with the geothermal gradient (Figure 2.9). However, as the ocean temperatures in the Arctic are slightly below the freezing point (Riedel et al., 2015), the degrading rate of subsea permafrost is slow (Paull et al., 2007; Taylor et al., 2013).

Subsea permafrost in the Beaufort Sea

The Beaufort Sea is a marginal sea located north of Canada and Alaska (Figure 2.10). Its continental shelf is narrow, especially to the east (10 km), and widens to the north of the Mackenzie delta, reaching up to 120 km from the shores. The mean annual water temperatures in the Beaufort Sea vary from -1.6°C at ∼30-200 m water depth to 0.5°C at ∼ 350 m water depth (McLaughlin et al., 2004). In fact, Taylor et al. (2013) calculates that for the last 125,000 years, the seabed temperature in the Beaufort Sea has not surpassed 5°C and most of the time kept under 0°C. Those low temperatures indicate a process of slow warming over the underlying subsea permafrost.

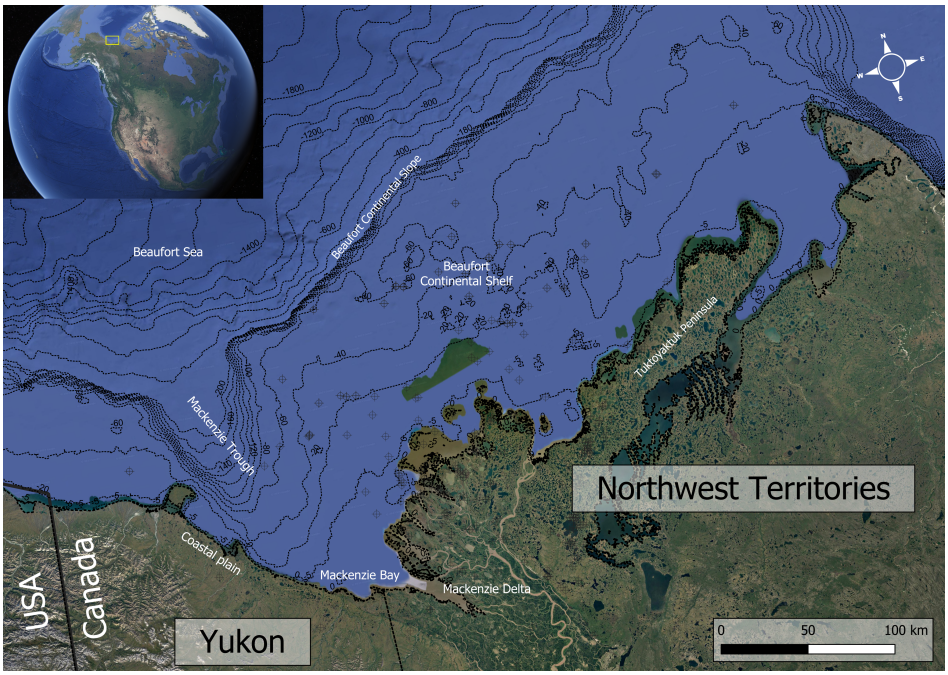


Figure 2.10 Main physiographic elements of the Canadia Beaufort Sea. Inspired from Dixon et al. (2019).

The Mackenzie River plays a significant role in the oceanography and sedimentology of the Beaufort Sea and the Beaufort-Mackenzie Basin. About 300 km^3 of freshwater and 85 million tonnes of sediment annually pass from the Mackenzie Delta to the Canadian Beaufort Shelf (Dunbar et al., 2015).

The Beaufort-Mackenzie Basin is a structural depression occurring near the continental margin of the Beaufort Sea (Sippel et al., 2013, 2015). It is part of the Canadian Beaufort Sea on the Southeast margin and extends for short distances into the Yukon and the Northwest Territories (Dixon et al., 2019). Sedimentation in the basin began in the Late Triassic and

continues up to date. Different depositional and faulting processes have affected the basin, including marine water floodings, epicontinental sedimentation, development of fluviodeltaic systems, and strike-slip faulting. Those elements and the movement of the Alaska block generated a set of complex structures known as the "tectonokinetic diapir field" (Pinet et al., 2023; Young et al., 1976).

The Mackenzie Trough, situated west of the Mackenzie Delta, splits the shelf into two distinct regions. This trough reaches depths of up to 400 m before joining the continental slope. Towards its southern end, the 100 m isobath comes within 10 km of the coastline (Dixon et al., 2019; Young et al., 1976). The Mackenzie trough is a distinctive geological characteristic in the Canadian Beaufort Sea, linking the Mackenzie River Delta to the continental slope. This relatively recent paleovalley took shape due to glacial erosion during the late Wisconsinan period (approximately 25,000 to 10,000 years ago) and is currently filled with around 400 meters of unconsolidated sediments (Batchelor et al., 2014; Duchesne et al., 2020).

Subsea permafrost in the Beaufort Sea formed at the surface when the shelf was subaerially exposed during the last glaciation. Borehole stratigraphy indicates that the inner shelf comprises a thick section of eight regressive/transgressive sediment cycles consisting of sea-level low-stand deposits of thick clastics and sea-level highstand deposits of thin marine muds (Blasco et al., 2012; Ruffell et al., 1990).

Blasco et al. (2012) explains the formation and submergence of subsea permafrost as follows: During glacial maxima and periods of low sea levels, the Beaufort Shelf was exposed, and due to ice sheets blocking the Mackenzie River drainage, it experienced a lack of sediment supply. The initial deglaciation phase was characterized by significant meltwater events that inundated the shelf with thick glaciofluvial clastic outwash sediments. This deposition of meltwater sediments on the exposed shelf occurred rapidly. During interglacial periods, rising sea levels led to the deposition of thin marine mud, mainly because the Mackenzie River sediment discharge bypassed the inner shelf.

Throughout the low sea levels associated with glaciation, the shelf experienced exposure to subzero temperatures, preserving the underlying sediments that contained frozen sediments. The rapid flow of sediment-laden meltwater across the shelf resulted in the quick subaerial deposition of clastic sediments and the simultaneous accumulation of ice-bearing permafrost. As sea levels rose during deglaciation, the shelf flooded, and marine mud was deposited. During interglacial periods, seawater slightly warmed the subsurface up to -1.5^{o}C, but it was insufficient to melt the ice within the sediments, which remained frozen.

Subsequent glaciation and the resulting drop in sea level once again exposed the shelf, causing the marine muds now exposed to freeze and lowering the temperature of the still-frozen sed-

iments beneath -1.5^{o}C. With deglaciation, another layer of glaciofluvial outwash was rapidly deposited, which promptly froze in the cold subaerial environment. This process repeated in eight cycles, causing the accumulation of ice-bearing sediments on the shelf as the Tertiary age basin continued to subside.

Groundwater dynamics

The Arctic's subsea permafrost is a dynamic system that undergoes constant substantial environmental shifts. Many of these changes can be attributed to the Arctic hydrologic cycle (Charkin et al., 2017). Freshwater delivery is of particular significance, which plays a crucial role in shaping the distribution of salinity and the formation of sea ice in the region. The combined effect of alterations in the freshwater flow to the Arctic Ocean could have a notable influence on the global ocean circulation by impacting the volume of North Atlantic deep-water formation (Stocker et al., 2013).

Apart from the freshwater input coming from rivers, the Arctic and subarctic watersheds are experiencing the effects of climate warming, permafrost thawing, and thermokarst formation. These changes are causing significant alterations in how surface water and groundwater interact (Charkin et al., 2017). For instance, approximately 25% of the water discharge from the Yukon River to the Bering Sea consists of groundwater. This groundwater is then carried into the Arctic Ocean through the Bering Strait (Walvoord and Striegl, 2007). However, up to date, there has been limited research on the Arctic Ocean shelf groundwater system, and the submarine groundwater discharges remain inadequately understood (Charkin et al., 2017; Gwiazda et al., 2018).

Heat advection can accelerate permafrost thaw, modify the hydrate stability zone, and facilitate the diffusion of methane gas to the surface (Angelopoulos et al., 2020; Shakhova et al., 2017b). Nonetheless, submarine groundwater discharge has the potential to refresh marine sediment porewater, elevate the freezing point, and thus protect subsea permafrost from degradation, while increasing gas hydrate stability (Frederick and Buffett, 2015). Methane present in Arctic water may also have its origins in terrestrial sources, using submarine groundwater discharge conduits as a means of transportation (Lecher et al., 2016).

In the Canadian Beaufort Sea, the observation of freshwater seepage into the shelf-edge and slope has been attributed to several factors, including the infiltration of mixed sea and river water, subsea permafrost degradation, and submarine groundwater discharge (Gwiazda et al., 2018). Another contributing factor to subsea permafrost thawing may involve offshore water flow within the thawed permafrost layer, pushed by convective processes (Semenov et al., 2020), and potentially contributing to gas hydrate destabilization (Angelopoulos et al., 2020).

Pockmarks and pingo-like features

Seafloor features associated with subsea permafrost are present in different shapes, including pockmarks and pingo-like features (Angelopoulos et al., 2020) (Figure 2.3). Pockmarks manifest as expansive, circular depressions on the seabed (Harrington, 1985). The existence of pockmarks is often linked to subsurface fluid flows triggered by permafrost thaw (Hovland et al., 2002; King and Maclean, 1970; Sullivan et al., 2021).

The formation mechanism of pockmarks involves upward fluid flows, incorporating processes like hydrodynamic blowouts, submarine hydrate dissolution, catastrophic subterranean hydrate dissolution, and explosive cryovolcanism events (Brothers et al., 2012b; Judd and Hovland, 2009; King and Maclean, 1970). Fluid expulsion, bottom currents, and sediment deposition can modify the morphology of pockmarks (Sullivan et al., 2021).

Pingo-like features (PLFs) are structures associated with subsea permafrost. These structures are circular mounds that emerge from the seafloor and resemble pingos on land. Several PLFs have been observed scattered across the Canadian Beaufort Sea but are mostly found on the outer shelf and upper slope (Paull et al., 2007; Shearer et al., 1971). Recently, based on four bathymetry surveys conducted between 2010 and 2019, Paull et al. (2022) revealed that some PLFs experienced a volume loss that could be attributed to permafrost thawing.

The formation of PLFs remains an area of ongoing study, with two main theories attempting to explain their origin. One theory suggests that PLFs may form due to the accumulation of methane, similar to the process observed GECs on the surface (Bogoyavlensky, 2015; Serov et al., 2015). On the other hand, an alternative theory proposes that these structures could rise as a result of low salinity water flowing from the surface, beneath the permafrost, and reaching the permafrost limit at the shelf-slope (Paull et al., 2022; Serov et al., 2015). As sediment temperatures in the outer shelf are below the freezing point (Riedel et al., 2015), low-salinity water could freeze and create intra-sediment ice, exceeding the sediment pore space. This freezing-induced volume expansion would cause the seafloor to heave, as observed in onshore pingos (Paull et al., 2022).

2.2 Seismic mapping of permafrost

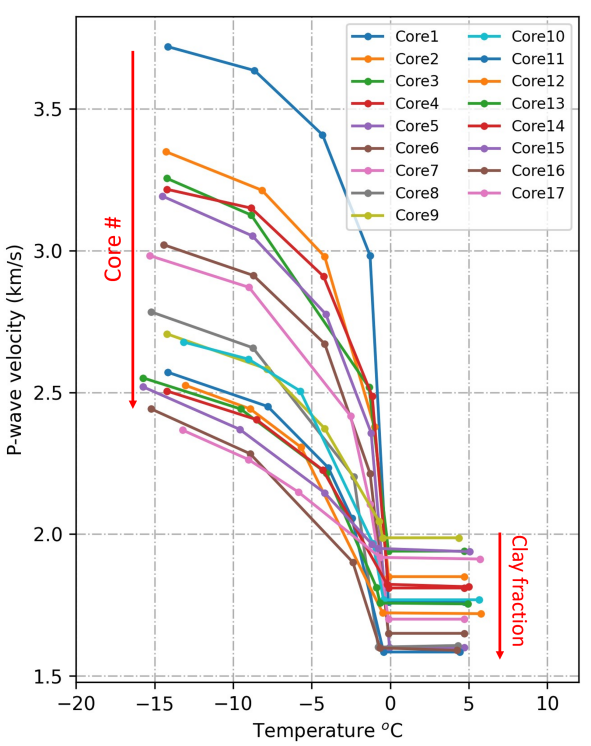
2.2.1 Seismic properties of permafrost

Permafrost can be categorized according to its ice content or water saturation. For example, dry permafrost is permafrost that neither contains water nor ice. Ice-bearing permafrost describes permafrost containing ice, while ice-bonded permafrost refers to ice-bearing per-

mafrost in which the soil particles are cemented together by ice (Van Everdingen, 2005).

The presence of ice has a dominant effect on the properties of the permafrost soil and associated processes. For instance, ice-bonded permafrost has a low hydraulic conductivity (K). This reduced conductivity plays a crucial role in regulating the storage and movement of groundwater, giving rise to intricate hydrologic systems (Bring et al., 2016; Walvoord and Kurylyk, 2016).

On the other side, the strength of the ice, characterized by its internal angle of friction, cohesion, shear strength, and tensile strength, typically exhibits higher values at lower temperatures and gradually diminishes as the temperature approaches the bulk-melting point (Fish et al., 1997). Nonetheless, it is important to acknowledge that ice strength is influenced by various parameters. These parameters include hydrostatic pressure, formation conditions, crystal size, presence of impurities, and the inclusion of fine material (Gruber and Haeberli, 2007). Consequently, the response of ice-filled soils, rocks, and discontinuities to warming displays a complex behaviour owing to these effects.



Core	Porosity	% Particle size			Total clay Fraction
		Sand	Silt	Clay	
1	0.37	26	30	14	0.05
2	0.36	37	31	32	0.05
3	0.35	32	45	23	0.05
4	0.37	18	56	26	0.05
5	0.41	48	25	27	0.1
6	0.43	26	34	40	0.1
7	0.39	23	42	35	0.1
8	0.45	0	45	55	0.25
9	0.31	2	58	40	0.1
10	0.39	3	50	47	0.2
11	0.43	1	38	61	0.3
12	0.39	3	47	50	0.2
13	0.38	7	49	44	0.2
14	0.34	1	44	55	0.2
15	0.32	5	49	46	0.2
16	0.42	6	37	57	0.25
17	0.33	3	47	50	0.2

Figure 2.11 P-wave velocity for permafrost core samples as a function of temperature. Data taken from measurements on the 17 core samples presented in the work of King (1984).

Due to the higher seismic velocity of ice with $V_P \sim 3.8$ km/s and $V_S \sim 1.9$ km/s, ice-bearing subsea permafrost presents higher velocities than ice-free sediments. These contrasting veloc-

ities have been extensively documented in laboratory studies (Dou et al., 2016; King, 1984; King et al., 1982; Matsushima et al., 2008). However, the seismic velocities of ice-bearing permafrost can vary significantly due to several influencing factors (Figure 2.11), including grain size (King et al., 1982), clay fraction (King, 1984), cohesion (Winters et al., 2004, 2007), pore-water salinity, and ice saturation (Dou et al., 2016).

King et al. (1988) proposed that in the presence of clay, two interfaces are contained in the porous medium: a silicate-water-ice interface and a silicate-water-silicate interface where a continuous layer of unfrozen water surrounds each silicate grain. The thickness of this layer decreases with temperature. Under these conditions, it is assumed that ice forms faster in coarse material as the surface forces are lower on it. Comparisons of seismic velocities measured on coarse and fine-grained rocks agree with this theory (Dou, 2015; King et al., 1982). Further studies by King (1984) also suggest that a reduction in the compressional velocity could arise due to an increase in the clay fraction. They also found a linear relationship between the water-filled porosity and the compressional velocity regardless of the original porosity, clay content or temperature.

Winters et al. (2004, 2007) proposed that ice and gas hydrates can serve either as pore fillers, frame builders, or cement. Laboratory analysis validates the idea that hydrates can be integral components of the sediment frame rather than solely cementing the grains together. In the study by Winters et al. (2007), no evidence was found of ice cementing the sediments in the core samples. However, acoustic models based on measured V_P did suggest the possibility of a cementing effect.

An analysis of the pore-water salinity effects on ice-bearing permafrost is presented in Dou (2015). She measured seismic velocities with decreasing temperatures and observed that an increase in salinity affects P-wave velocity. As their findings suggest, a reduction in the P-wave velocity with salinity within the interval between the freezing and the eutectic points eludes to a supercooling state of water.

In the study by Dou (2015), it was also observed that there was a sudden decrease in the measured amplitudes and centroid frequency below the freezing point. The researchers identified the minimum amplitude and centroid frequency values between the freezing and eutectic points, which was dependent on the pore fluid salinity. As the eutectic point was approached, there was a rapid increase in both amplitude and centroid frequency. Interestingly, the researchers also noted that salinity seemed to have less influence on the measured P-wave velocities, amplitudes, and centroid frequency above the freezing point and below the eutectic point.

Dou et al. (2016) used temperature and salinity measurements to calculate ice saturation.

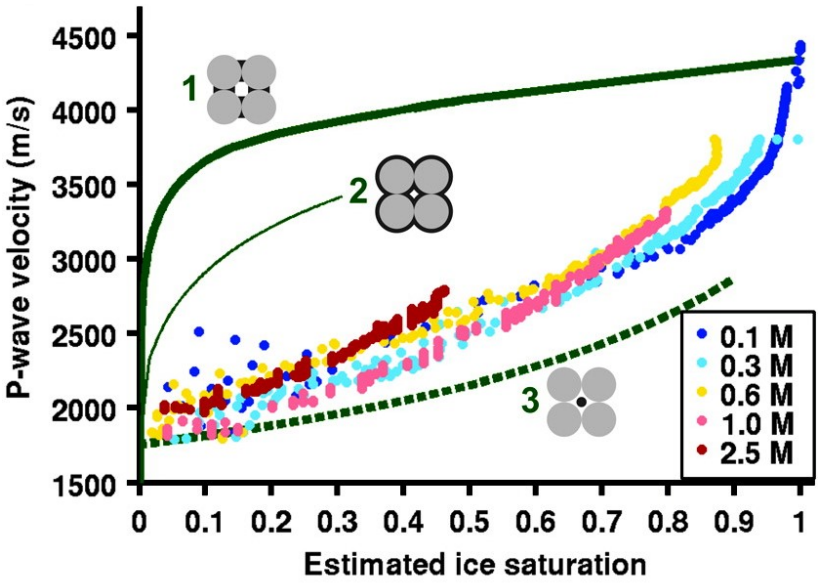


Figure 2.12 P-wave velocity of brine-saturated coarse sand presented as a function of the estimated ice saturation as originally presented in Dou et al. (2016)

They discovered that the P-wave velocity increases with ice saturation, regardless of the pore-fluid salinity (Figure 2.12). Interestingly, the observed trend exhibited an intermediate curve between the theoretical curves representing ice cementing the pore grains and ice acting as pore fill. The study also revealed that amplitude and centroid frequency demonstrated an apparent minimum at approximately 70% ice saturation.

The literature has extensively documented seismic attenuation in ice-bearing samples (Dou, 2015; Matsushima et al., 2008, 2016; Wu et al., 2017). Generally, seismic attenuation has a low impact when temperatures are above the freezing point. However, as temperatures fall between the freezing and eutectic points, the arrival time decreases, and the waves are attenuated, leading to a decrease in high-frequency content during the supercooling stage. Subsequently, below the eutectic point, attenuation decreases again, and the high-frequency values reappear on the frequency spectrum.

In a recent study by Ji et al. (2023), a comprehensive comparison of seven seismic models was conducted to predict the seismic wave velocity of permafrost, along with six unfrozen water content models. The researchers used laboratory and in-situ data from thirteen publications to achieve this. The seven seismic models analyzed in this study assumed that permafrost soil consists of a soil skeleton, water, and ice. Regardless of variations in porosity, grain size, and temperature, the researchers discovered that all soil types with permafrost displayed a consistent linear correlation between P-wave and S-wave velocity characterized by the

equation $V_S = 0.6116V_P - 210.4$. Additionally, the study found that the performance of the seismic velocity models exhibited differences depending on the specific type of permafrost soil being considered.

2.2.2 Fundamentals of seismic imaging

The following summary is based on the description of the fundamentals of seismic imaging presented in Dix (1952), Sheriff (2002), Heagy et al. (2016) and Mavko et al. (2020). Seismic surveys employ waves to construct images of the subsurface by analyzing vibrations produced by these waves. Seismic methods involve generating elastic energy from a source point that moves through the ground. Geophones capture ground movements that reflect from points where the properties of the subsurface materials change. At these points, two fundamental physical effects occur: reflection and refraction.

In reflection, a portion of the energy from an incoming seismic wave interacting with an interface gets sent back, or reflected, toward the surface. Similarly, as energy propagates, some of it bends or is refracted at the interface. This bending might cause a refracted wave to travel along the interface between two subsurface materials, eventually returning to the surface. These two phenomena form the basis for the two primary techniques in exploration seismology: reflection and refraction surveying.

Seismic surveys are responsive to the velocity of seismic waves within the subsurface. This velocity is influenced solely by the elastic properties and density of the material through which the waves travel. The properties that influence the amount of energy reflected or transmitted at a boundary are known as the reflection and the transmission coefficients (RC and TC, respectively):

$$\begin{aligned} RC &= \frac{Z_2 - Z_1}{Z_2 + Z_1} \\ TC &= \frac{2Z_1}{Z_2 + Z_1} \\ Z_i &= V_i \rho_i \end{aligned} \tag{2.1}$$

In this equation, Z is the seismic impedance, ρ and V symbolize density and seismic wave velocity, respectively. Subscript 2 relates to the material situated beneath the reflecting boundary, while subscript 1 pertains to the material above the boundary (Dix, 1952).

When elastic materials are present, such as those found beneath the surface, a P-wave (compressional wave) that approaches an interface at an angle other than perpendicular will

generate both reflected and transmitted S-waves (shear waves) and P-waves. Analogous conversions occur when an S-wave encounters a flat boundary (Heagy et al., 2016). These transformations in wave types can make interpretation more complex. Nonetheless, S-waves travel slower than P-waves, and consequently, initial wave arrivals are always P-waves. Moreover, while S-waves are consistently transmitted through elastic materials, they cannot be propagated through liquids like a water layer due to the absence of shear stress resistance.

While the law of reflection states that the angle of reflection matches the angle of incidence, the angle of refraction follows Snell's law, given by:

$$\frac{\sin\theta_1}{V_1} = \frac{\sin\theta_2}{V_2} \tag{2.2}$$

When a wave moves from a medium with a lower velocity to one with a higher velocity, it bends away from the normal upon refraction. Conversely, if the wave travels from a higher velocity medium to a lower velocity one, it bends toward the normal. The critical angle of refraction occurs when the angle of incidence leads to a corresponding angle of refraction of 90^o. At this point, the refracted ray moves horizontally along the boundary, and this critical angle is defined by $\theta_c = V_1/V_2$.

In cases where the wave within the second medium experiences critical refraction, it moves parallel to the boundary at a velocity of V_2. While it travels, it radiates energy into the upper medium along its path, creating an angle θ_c. Note that the energy will be refracted towards the surface only when the velocity of the lower layer is larger than the velocity of the upper layer. If this is not the case, the ray will curve towards the vertical direction.

When seismic waves propagate through an isotropic and homogeneous medium, they travel at a constant velocity. This implies that the time it takes for a seismic wave to travel from the source to the receiver (known as the direct arrival) within a uniform layer can be calculated as $t_{direct} = x/V$, where x represents the distance covered in the layer. Similarly, the travel times for the reflected (t_{rfl}) and refracted (t_{rfr}) arrivals can be expressed based on the distance x as follows:

$$\begin{aligned} t_{rfl} &= \sqrt{t_0^2 + \frac{x^2}{V_{RMS}^2}} \\ t_{rfr} &= \frac{x}{V_i} + t_{i-1} \\ V_{RMS} &= \frac{\sum_{i=1}^{N} V_i^2 t_i}{\sum_{i=1}^{N} t_i} \end{aligned} \tag{2.3}$$

In these equations, t_0 represents the two-way travel time of the reflected wave at zero offset,

V_{RMS} corresponds to the root mean squared velocity encompassing all the layers traversed, t_i signifies the intercept time at which the refraction line intersects the y-axis, and V_i stands for the velocity of the i-th layer. Note that while the formulas for direct and refracted arrivals display a linear relationship with respect to x, the arrival time for reflected waves follows a hyperbolic pattern.

Attenuation is an important phenomenon that occurs in wave propagation. It refers to the decrease in energy that waves experience when they travel through a medium (Robinson and Treitel, 2008). Attenuation is related to the quality factor Q, which helps us determine the amplitude of vibration as a function of time (t) and the natural frequency of the system (ω_0). The following equation can be used to calculate the amplitude at any given time:

$$A(t) = A_0 e^{-\omega_0 t/2Q} \tag{2.4}$$

where A_0 represents the amplitude at the source. It is important to note that as $1/Q$ increases, the attenuation also increases.

Marine seismic imaging and subsea permafrost

In offshore seismic surveys conducted in permafrost zones, a typical setup involves a boat equipped with a source and an array of receivers (Figure 2.13a) that traverse the study area. The primary objective is to examine the propagation of seismic waves through the subsurface beneath the seawater and its interaction with underlying geological layers. Figure 2.13 shows an example of a shotgather collected on subsea permafrost zones.

Seismic waves are emitted from the source and travel through the seawater. The receivers capture the direct arrivals - waves travelling from the source to the receivers through the water layer. At the ocean bottom, the significant impedance contrast between the seawater layer and the unfrozen sediment generates a reflection. These reflections then travel upward through the water column, appearing as hyperbolic events on the shot gathers. Simultaneously, due to the velocity of the sediments is higher than the velocity of the water, a portion of the downward waves undergo critical refraction, following the interface at the velocity of the unfrozen sediments. Refractions are displayed as straight lines in the seismograms.

The seismic waves continue propagating through the unfrozen sediments until they encounter the ice-bearing permafrost layer. This interface yields another impedance contrast, passing from lower velocity (sediments) to higher velocity (the permafrost). The wave hits the interface, resulting in additional reflected and refracted waves, which travel upwards and are recorded by the receivers. The downward waves propagate through the ice-bearing per-

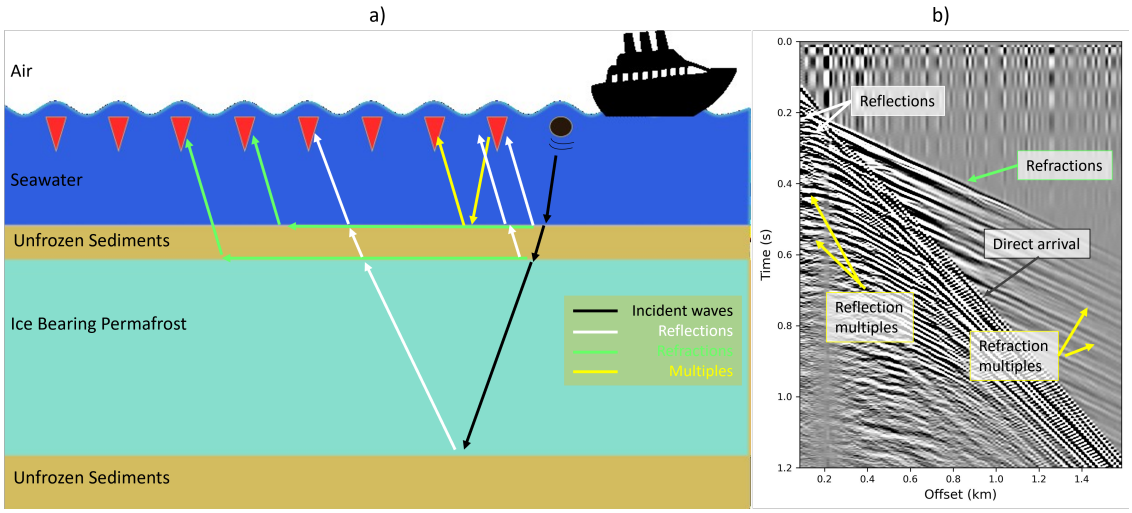


Figure 2.13 a) Schematic description of an offshore seismic survey on permafrost zones. A source (black circle) and an array of receivers (red triangles) are tied to a survey vessel. Seismic waves are emitted from the source. The receivers record the propagated waves, including direct waves, reflections (white lines), refractions (purple lines) and multiples (yellow lines). b) Example of a shot gather recorded on a subsea permafrost zone in the Canadian Beaufort Sea.

mafrost until they reach the underlying unfrozen sediments at the base of the permafrost. Notably, at this interface, there is a decrease in the velocity, preventing critical refractions. Still, reflections are generated at this interface.

Upon reaching the surface, the direct, reflected, and refracted waves undergo another round of reflection due to the high reflection coefficient between the top of the water layer and the air, called the free surface. These subsequently reflected waves propagate downwards, giving rise to a series of arrivals called free surface multiples. Additionally, multiples can also arise at various interfaces and are known as intralayer multiples.

The amplitudes of the reflected and refracted waves and their multiples depend on the reflection and transmission coefficients at the respective interfaces. Notably, higher reflection coefficients lead to increased energy for the reflected and critical refracted waves while diminishing the energy of the waves continuing to propagate downwards. Given the considerable impedance contrast between the air and the water layer, between the seabed and unfrozen sediments, and between the unfrozen sediments and the ice-bearing permafrost, a leaky waveguide can emerge. This waveguide is confined upwards by the interface at the sea level and downwards by the interface between the unfrozen sediments, effectively trapping a substantial amount of the emitted energy from the source.

2.2.3 Seismic methods and permafrost

Due to the increase in seismic velocity caused by the occurrence of ice in the pore space of the sediments, seismic surveys have been proposed to map the distribution of ice-bearing permafrost at a regional scale. For example, Johansen et al. (2003) conducted a seismic survey along a 5 km long line in Advetdelta, Spitsbergen, Norway, to analyse near-surface sediment thickness and structure and to investigate changes in reflectivity properties as a function of variations in pore ice concentration. Using numerical modelling, they discovered that the P- and S-wave velocities show a significant increase of 80% - 90% when the fully water-saturated sediment transitions from unfrozen to a frozen state. Furthermore, their research highlighted that the seismic reflectivity in fully water-saturated sediments experiences a substantial reduction when they become frozen.

In the Swiss Alps, Hilbich (2010) employed time-lapse refraction seismic tomography to observe alterations in ice content and permafrost degradation. Their study confirmed that this method effectively captures temporal changes in alpine permafrost and can identify ground ice degradation. However, they noted that a significant limitation lies in the shallow penetration depth caused by strong velocity contrasts between the active layer and the permafrost table. To limit the impact of this issue on the recordings, they recommended conducting measurements in the late summer.

In the Mackenzie Delta (Canada), Ramachandran et al. (2011) conducted a high-resolution 3D traveltime tomography to create detailed maps of the permafrost velocity structure. Their findings revealed that the permafrost structures in the shallow layers are not continuous, and regions exhibiting strong amplitude blanking and frequency attenuation coincided with deep lakes containing unfrozen water.

Overduin et al. (2015) employed ambient vibration seismic wavefield spectral ratios and S-wave velocity data obtained from dispersion curves of surface waves to assess the thickness of sediment layers above ice-bearing permafrost in Muostakh Island in the Laptev Sea, Siberia. The researchers observed that the significant contrast in shear wave velocity between permafrost and the unfrozen sediments made the analysis of ambient noise spectral ratios an appropriate method. Furthermore, they successfully demonstrated that it was feasible to transform the spectral ratio peaks into layer thickness estimates using cross-correlations of ambient noise data.

In central Alaska, James et al. (2017) used seismic refraction interferometry with moving window cross-spectral analysis on a two-year-long recording of ambient noise data to explore the feasibility of using seismic noise to quantify relative velocity changes resulting from

seasonal active-layer dynamics. While the conventional stationary reference approach proved inadequate in capturing the decrease in seismic velocity, the researchers successfully achieved continuous velocity variation measurements in the permafrost environment by adopting a moving-reference approach.

Lastly, in a study by Cheng et al. (2022), researchers explored the practicality of monitoring permafrost using permanently installed fibre-optic seismic networks. During a controlled thawing experiment in Fairbanks, Alaska, they conducted a 2-month seismic monitoring campaign. The findings from the experiment demonstrated a 15% decrease in S-wave velocity, indicating a reduction in the soil's rigidity as it warmed. Additionally, the controlled thawing effects and permafrost degradation resulted in a deepening of the observed permafrost table by 2 meters.

Subsea permafrost mapping in the Beaufort Sea

Various seismic methods have been used in the Beaufort Sea to investigate the lateral extent of subsea permafrost. One of the early attempts to map the distribution of subsea permafrost in the Canadian Beaufort Sea was documented by Hunter et al. (1976). They analysed seismic refraction using data recorded with a 24-channel refraction array and by the oil industry. The results indicate that subsea permafrost exists over much of the Beaufort Sea shelf. The subsea permafrost zone was mostly continuous, encompassed by a surrounding discontinuous permafrost zone. Interestingly, the study also noted that the seabed topography in areas with ice-bonded permafrost was notably rougher than in regions without permafrost. Furthermore, the investigation revealed no evidence of ice-bonded permafrost in water depths exceeding 90 m.

In a subsequent investigation, Pullan et al. (1987) built upon the research conducted by Hunter et al. (1976) by analyzing additional seismic lines obtained from the oil industry using seismic refraction. Their findings align with those of Hunter et al. (1976). Until now, these two publications constitute the most comprehensive map of permafrost distribution in the Beaufort Sea region (Figure 2.14) characterizing a large area of the Canadian Beaufort Sea between low ice-content sediments, little or no ice-bonded sediments, discontinuous ice-bonded sediments and continuous ice-bonded sediments. However, no raw data from the measurements used in these studies has been preserved.

Brothers et al. (2012a) used refraction analysis on 5,000 km of multichannel seismic data collected between 1977 and 1992 to produce the first regional map displaying subsea ice-bearing permafrost on the U.S. Beaufort Sea continental shelf. The findings indicated that the offshore extent of subsea permafrost in that area did not exceed the 20 m isobath.

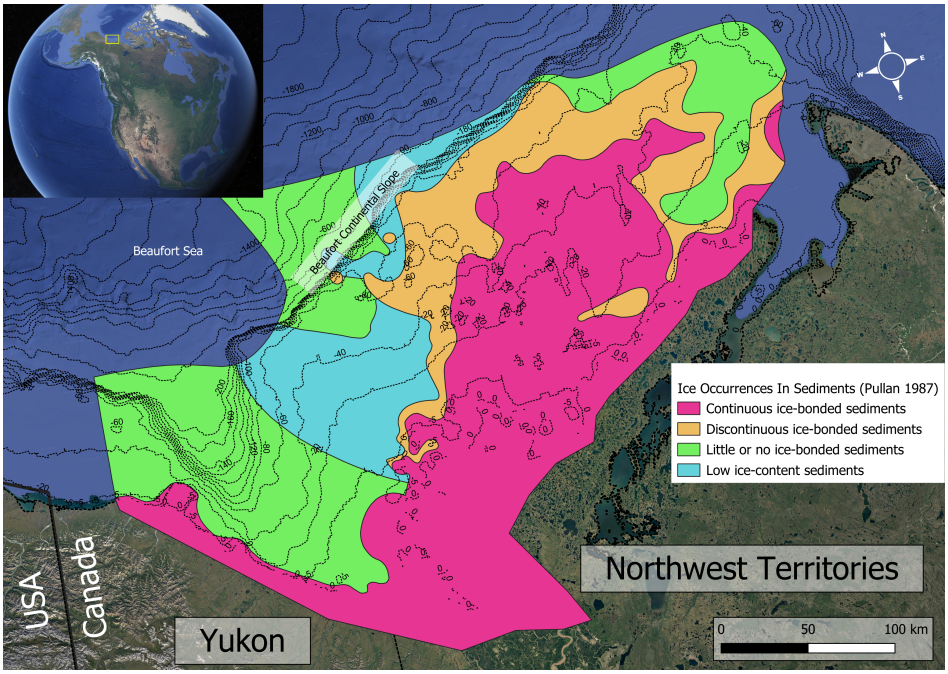


Figure 2.14 Map of ice occurrences in sediments, as presented in Pullan et al. (1987).

Likewise, Brothers et al. (2016) employed stacking velocity analysis of multichannel seismic reflection data gathered by industry on the U.S. Beaufort shelf. They aimed to identify continuous ice-bearing permafrost distribution. The study revealed that ice-bearing permafrost is present near the coastline, covering the entire U.S. Beaufort margin. However, as they moved seaward, the permafrost showed a rapid decrease in thickness.

In a recent study, Kang et al. (2021) employed full-waveform acoustic inversion to create P-wave velocity models on the continental shelf of the Canadian Beaufort Sea. The outcomes of their research provided valuable insights into the upper and lower boundaries of subsea permafrost. Nevertheless, their analysis of the permafrost base contradicted previous findings regarding the distribution of subsea permafrost in that region.

2.2.4 Challenges in subsea permafrost imaging

Seismic methods analyse various types of arrivals to unravel subsurface structures beneath the seabed. Among these methods, seismic refraction examines refracted arrivals and can be used to deduce the lateral distribution of subsea permafrost structures. Seismic refraction methods have proven effective in assessing the lateral distribution of permafrost and identifying ice-bonded permafrost layers separated by unfrozen layers (Hunter et al., 1976; Pullan et al., 1987). However, due to the absence of critical refractions at the base of the permafrost,

seismic refraction cannot provide vertical information regarding the distribution of subsea permafrost.

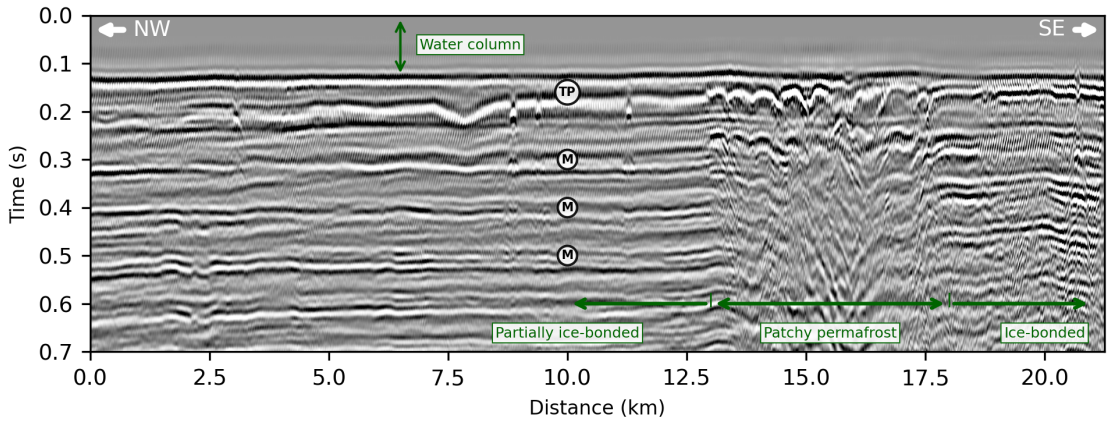


Figure 2.15 Stack section of a seismic survey collected in the Canadian Beaufort Sea over subsea permafrost zone. The top of the permafrost (TP) and the free surface multiples (M) are indicated in the Figure

Similarly, conventional seismic reflection methods image reflected arrivals. Nonetheless, in permafrost regions, the presence of highly energetic guided waves and multiple reflections from the free surface can disrupt the detection of reflections from deeper structures, particularly from the base of the permafrost (Duchesne et al., 2022). This interference between guided waves, free surface multiples, and reflections poses a challenge when applying conventional seismic processing techniques to identify subsea permafrost distribution (Kang et al., 2021). An example of this can be seen in Figure 2.15, which shows the stacking section of a seismic survey taken in a region where permafrost is expected in the Canadian Beaufort Sea. We can observe that while the reflection from the top of the permafrost is clearly visible, the presence of other arrivals (i.e., free surface multiples) interferes with the identification of other reflections from the base of the permafrost and deeper structures.

State-of-the-art seismic processing techniques, such as Full Waveform Inversion (FWI), leverage all recorded arrivals within an optimization framework to identify subsurface structures. FWI involves comparing synthetic seismic data generated from a proposed model of subsurface structures with the actual recorded data. This process iteratively minimises a cost function until it achieves the best possible approximation or until further improvements in the model are unattainable (Virieux and Operto, 2009). However, due to the pronounced heterogeneities observed in permafrost zones, establishing an initial model sufficiently close to the true model to achieve the best approximation presents a challenge.

Despite the limited capacities of different seismic approaches for characterizing and fully determining permafrost distribution, the recorded seismic waves still contain valuable information about the vertical and lateral subsea permafrost distribution (Duchesne et al., 2022). However, the existing processing methods fail to use that information. Moreover, abundant seismic data from the public domain and industry sources are available in the Arctic. While these surveys were not originally acquired for permafrost imaging purposes, they can be reprocessed to identify and describe subsea permafrost distribution.

2.2.5 Seismic datasets

Throughout this dissertation, we assess various seismic processing techniques and evaluate their suitability on recorded offshore seismic data. These data were obtained during the ARA04C and ARA05C expeditions from a region of the Canadian Beaufort Sea that is known to contain subsea permafrost. The ARA04C and ARA05C expeditions represent two collaborative research initiatives involving the Korea Polar Institute (KOPRI), the Geological Survey of Canada (GSC), the Department of Fisheries and Ocean (DFO), Monterrey Bay Aquarium Research Institute (MBARI), and the Alfred Wegener Institute (AWI). ARA04C occurred between September 6 and September 24, 2013, while ARA05C occurred from August 30 to September 15, 2014. Both expeditions conducted various research experiments investigating geological processes linked to permafrost degradation, fluid movement, degassing, seismostratigraphy, and associated geohazards in the Beaufort shelf and upper slope region.

Although both expeditions explored different areas of the Beaufort Sea using different methods, the present study concentrates on the multichannel seismic data gathered from the Canadian Beaufort Sea shelf and upper slope. In total, the ARA04C survey encompassed 14 lines, covering approximately 435 line kilometres, with 4,500 recorded shot gathers. Similarly, the ARA05C survey included 20 seismic lines, spanning around 1,000 line-kilometers and involving about 20,000 shot gathers. For our analysis, we used 6 lines collected during the ARA04C survey and 15 seismic lines from the ARA05C survey (see Figure 2.16).

Both surveys were conducted aboard the icebreaker research vessel Araon and employed similar tools and techniques. The energy source comprised a collection of eight airguns that represented a total volume of 1,200 cubic inches. The streamer used was a solid-state streamer (Sercel Sentinel) measuring 1.5 kilometres in length, which was equipped with 120 channels, each spaced by 12.5 meters. The time recording window was set at 10.0 seconds, and the information was captured using a sampling rate of 1 millisecond. Shots were fired at intervals of 93.75 meters for most of the ARA04C investigation and at 50 meters for the ARA05C study. Throughout both studies, the energy sources and receivers were maintained

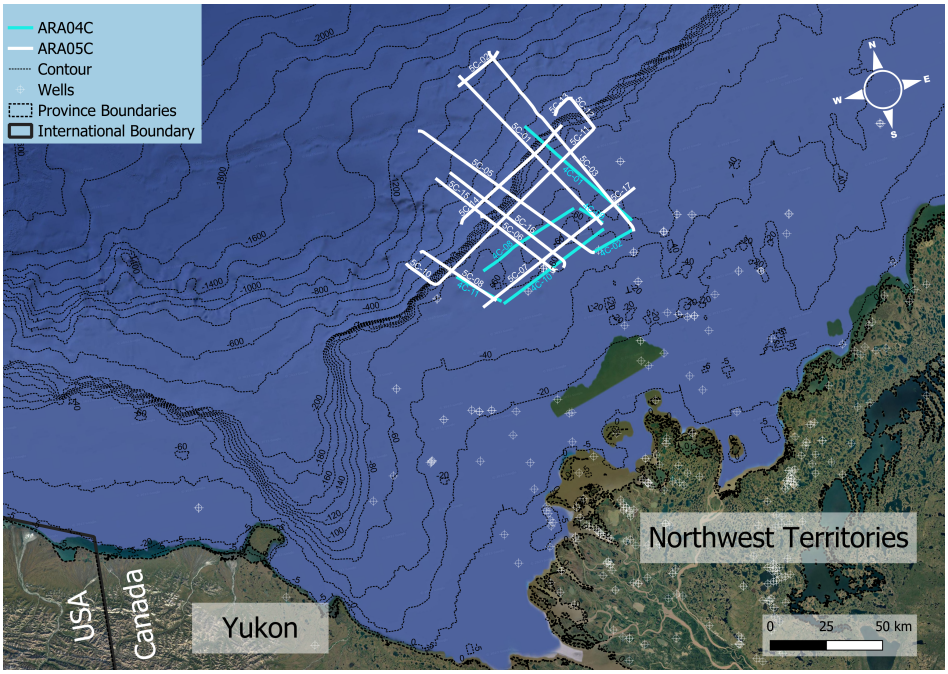


Figure 2.16 Track lines from the ARA04C and ARA05C surveys used in this study.

at a constant depth of 6 meters. To ensure the appropriate depth of the streamer, a total of eight cable levellers (birds) were fixed to the streamer at 300-meter intervals (Jin and Dallimore, 2016; Jin et al., 2015).

2.3 Deep learning for seismic inversion

The purpose of this thesis is to develop a methodology based on deep learning that can be used to image seismic velocity variations caused by subsea permafrost. To achieve this, we have focused on deep learning as the main approach. In this section, we will discuss the key concepts associated with deep learning, which includes a description of the commonly used optimization algorithms, neural network architectures, and their components. We will also discuss the challenges associated with applying deep learning to seismic inversion.

Deep learning is a subset of machine learning employed for performing tasks such as recognizing speech, classifying text, detecting fraudulent activities, enabling computer vision, processing natural language, and refining data, among other applications. This approach involves using neural networks (NN), an interconnected group of artificial neurons that process information based on connections. In deep learning, NNs are arranged in layers to handle data and acquire the ability to identify and categorize specific elements present in the data. Each layer in these networks builds upon the knowledge gained from the preceding one,

progressively enhancing the precision of object classification and predictive capabilities.

An NN consists of multiple neurons, which receive one or more inputs, perform a weighted summation, apply a non-linear activation function, and generate an output (Bishop, 1994). To achieve general approximation capabilities in an NN, having more than two layers is essential. Each neuron in one layer passes its output as inputs to the next layer, as shown in Figure 2.17.

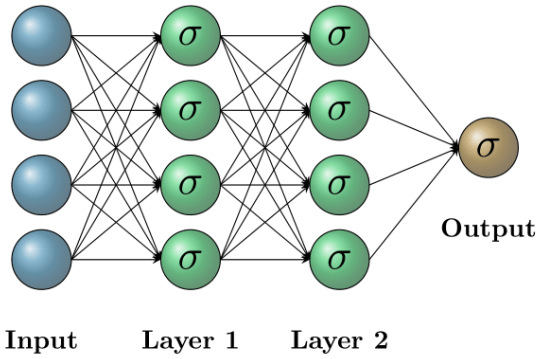


Figure 2.17 Sketch of a NN with two hidden layers. σ represents the activation function applied over the weighted summation of the neuron input.

In the context of using deep learning for seismic inversion, the seismic data acts as the input to the NN. This input is conceptualized as an image, where the attributes (such as time and offset in a shot gather) correspond to the image's width and length. The amplitude registered at a specific position can be likened to a pixel value within the image. The image pixels are weighted and summed, and the result is passed as input to neurons in the subsequent layer after applying a non-linear activation function to the output. This process continues through multiple layers until the output layer is reached, providing an estimated velocity profile in depth. Consequently, the objective of training an NN for seismic velocity inversion is to determine the weights and bias parameters for each neuron in all the layers, resulting in improved reproduction of the output velocity models.

In supervised learning, the training dataset includes samples with input data (e.g., seismic image) and corresponding labels (e.g., velocity profile). An optimization algorithm minimizes the error between the output and label images (LeCun et al., 2015) during training. A loss function (e.g., the least square error) defines this error. To appropriately adjust the weight vector, the learning algorithm computes a gradient vector that indicates the direction and magnitude by which the error would increase or decrease if the weight were slightly adjusted.

The weight vector is then updated in the opposite direction of the gradient vector to reduce the error during training.

In a typical deep-learning setup, optimizing a vast number of weights, often in the thousands or millions, is essential. Additionally, deep neural networks rely on numerous labelled examples, usually in the thousands or millions, to effectively learn and replicate underlying features during training. As a result, using efficient algorithms becomes imperative, especially for computing gradients.

2.3.1 Optimization

In supervised learning, an NN creates a map linking the input data (e.g., seismic images) and the corresponding labels (e.g., the velocity profile). Throughout the training process, the NN compares its predictions against the actual labels. An optimisation algorithm then gradually adjusts the weights and biases within its hidden layers. This tweaking aims to minimize the difference between the NN's outputs and the true labels.

During optimization, the aim is to find a minimizer (ω^*) that reduces the difference between the output of the neural network and the labels to the noise level (Menke, 2018; Nocedal and Wright, 2006). However, it is not always possible to find the exact value of the minimizer. To evaluate the difference, different loss functions can be defined. For instance, the least square equation ($\mathscr{L}$) can be used as a loss function, which is defined as follows:

$$\mathscr{L}(x) = \frac{1}{2}\sum_{j=1}^{m}\left(y_p(\omega) - y_l\right)^2 \tag{2.5}$$

Here, $y_p(\omega)$ represents the modeled data and y_l denotes the labels. To find the model parameters ω^* within a specified precision error, different optimization methods can be employed. Typically, optimization algorithms iterate the values of ω until either a solution point has been approximated or no more progress in the approximation can be made.

Gradient descent (GD) is the fundamental method for optimizing deep learning. It is based on the following equation:

$$\begin{aligned} \omega_{k+1} &= \omega_k - \eta_k \nabla \mathscr{L}(\omega_k) \\ \mathscr{L}(\omega_k) &\geq \mathscr{L}(\omega_{k+1}) \end{aligned} \tag{2.6}$$

In this equation, ω_k represents the model parameters (such as the weights and biases of the hidden layers) used in iteration k, $\nabla\mathscr{L}(\omega_k)$ is the gradient of the loss function for the k-th iteration, and η_k is the step size or the learning rate. For gradient descent, η_k can change at

each iteration. Certain choices of η_k, such as using the Wolfe conditions (Wolfe, 1969) or the Barzilai-Borwein method (Barzilai and Borwein, 1988), can guarantee convergence.

A variant of the GD is the Stochastic Gradient Descent (SGD) algorithm (Robbins and Monro, 1951). It is based on the assumption that the loss function can be expressed as the sum of training losses for a mini-batch of training samples. In other words, it computes the average of the losses for a subset of the data rather than the entire dataset. The SGD algorithm updates the model parameters using the following formula:

$$\begin{aligned} \mathscr{L}(\omega_k) &= \frac{1}{n}\sum_{i=1}^{n}\mathscr{L}_i(\omega_k) \\ \omega_{k+1} &= \omega_k - \frac{\eta_k}{n}\sum_{i=1}^{n}\nabla\mathscr{L}_i(\omega_k) \end{aligned} \tag{2.7}$$

By replacing the actual gradient calculated from the entire dataset with an estimated gradient calculated from a randomly selected subset of the data, SGD reduces the computational burden in high-dimensional optimization problems. This method achieves faster iterations in exchange for a lower convergence rate (Bottou and Bousquet, 2007). SGD is commonly used in training multiple models in machine learning, and when combined with the backpropagation algorithm, it represents the standard algorithm for training neural networks.

A further development was acquired with the momentum algorithm. Momentum (Rumelhart et al., 1986) is a technique that can help speed up SGD in the right direction and smooth out oscillations by incorporating a fraction γ of the past update vector to the current update vector (Ruder, 2016). The method can be summarized as follows:

$$\begin{aligned} v_k &= \gamma v_{k-1} + \eta_k\nabla\mathscr{L}(\omega_k) \\ \omega_{k+1} &= \omega_k - v_k \\ 0 &\leq \gamma \leq 1 \end{aligned} \tag{2.8}$$

Here, the momentum parameter γ controls the influence of the past update vector on the current update vector. By using momentum, the update process can benefit from the accumulated gradients of the past steps, which can help the algorithm to escape local minima and to converge faster. Similarly, the Nesterov accelerated gradient (Nesterov, 1983) is a momentum-based optimizer that adds momentum to the stochastic gradient descent (SGD) algorithm. This is done by approximating the next position of the parameters at $\omega - \gamma V_{t-1}$.

The Nesterov accelerated gradient can be calculated as follows:

$$\begin{aligned} v_k &= \gamma v_{k-1} + \eta_k \nabla \mathscr{L}(\omega_k - \gamma v_{k-1}) \\ \omega_{k+1} &= \omega_k - v_k \end{aligned} \tag{2.9}$$

Adagrad (Duchi et al., 2011) is an SGD method that adjusts the learning rate for each parameter at every iteration. The learning rate is different for each parameter ω_i at each iteration k. Adagrad can be described as follows:

$$\begin{aligned} \omega_{k+1,i} &= \omega_{k,i} - \frac{\eta}{\sqrt{G_{k,ii} + \epsilon}} g_{k,i} \\ g_{k,i} &= \nabla \mathscr{L}(\omega_{k,i}) \end{aligned} \tag{2.10}$$

Here, G_k is the diagonal matrix of the sum of the square of the gradients at iteration k, and ϵ is a smoothing term that prevents division by zero. Adagrad eliminates the need for tuning a learning rate, which is its main advantage. However, it suffers from accumulating summation in the denominator, making the learning rate vanish after several iterations.

Another improvement to the SGD was presented with the Adadelta algorithm, proposed by Zeiler (2012). It addresses the problem of vanishing learning rates by restricting the window of accumulated past gradients to a fixed size. The denominator of the update equation is defined as a decaying average of all past square gradients. The update equation is as follows:

$$\begin{aligned} \omega_{k+1} &= \omega_k + \Delta\omega_k \\ \Delta\omega_k &= -\frac{\eta}{\sqrt{E[g^2]_k + \epsilon}} g_k \\ E[g^2]_k &= \beta E[g^2]_{k-1} + (1-\beta) g_k^2 \end{aligned} \tag{2.11}$$

Finally, the Adaptive Moment Estimation (ADAM) introduced by Kingma and Ba (2014) also computes adaptive learning rates for each parameter. In addition to storing an exponentially decaying average, it also keeps an exponentially decaying average of past gradients. The

algorithm updates the parameters as follows:

$$\begin{aligned}
\omega_{k+1} &= \omega_k - \frac{\eta}{\sqrt{\hat{v}_k} + \epsilon}\hat{m}_k \\
\hat{m}_k &= \frac{m_k}{1 - \beta_1^k} \\
\hat{v}_k &= \frac{v_k}{1 - \beta_2^k} \\
m_k &= \beta_1 m_{k-1} + (1 - \beta_1) g_k \\
v_k &= \beta_2 v_{k-1} + (1 - \beta_2) g_k^2
\end{aligned} \tag{2.12}$$

The ability of the ADAM algorithm to adapt momentum and learning rates contributes to its fast convergence, making it highly effective for dealing with sparse and noisy datasets.

As can be deducted from the previous section, the computation of the gradients is crucial for the commonly used algorithms in supervised learning. One of the most widely utilized algorithms for gradient computation in deep learning (and the algorithm we will use in this dissertation) is reverse mode automatic differentiation, often called backpropagation. This method is rooted in the concept that numerical computations can be expressed as compositions of a finite set of elementary operations with known derivatives. Combining these elementary operations' derivatives using the chain rule, the overall composition's derivative can be obtained (Baydin et al., 2018). The algorithm records the sequence of operations during the forward pass to perform automatic differentiation. During the subsequent backward pass, it traverses this list of operations in reverse order to compute gradients. While this approach consumes more memory in one pass, as it accumulates derivatives for every variable with respect to each quantity in the computation, it enables the calculation of derivatives for multiple inputs in a single reverse pass (Murray, 2019).

2.3.2 Neural networks architectures and components

Neural networks are used in deep learning to process information through connections. However, due to the availability of multiple NN architectures, choosing the right network configuration can be challenging. This section aims to provide a brief overview of the most commonly available NN architectures, highlighting their applicability and the input data they require. We will also discuss convolutional and transpose convolutional networks in more detail, as they are used in the proposed architecture for this thesis. Additionally, we will give a brief description of the evolution of CNNs and explain the activation functions that give the NN its non-linearity capabilities.

While the fundamental elements of deep learning were initially introduced from the 1940s to the 1980s (Fukushima, 1975; Kelley, 1962; McCulloch and Pitts, 1943; Rumelhart et al., 1986), their practical utility was considerably limited due to the computational constraints of that era. It was not until the 2010s that deep learning experienced a remarkable surge in popularity. This resurgence can be attributed to the notable advancements in GPU processing speed, which enabled the effective application of deep learning techniques.

Since the resurgence of deep learning, numerous architectures have been consistently proposed and updated. While NNs were initially limited to basic classification tasks, their capabilities have progressed to encompass a wide array of problems across various domains. In practical application, deep learning has transitioned from modest artificial NNs, capable of processing only limited datasets, to intricate architectures comprising millions of parameters trained on extensive volumes of data.

The most basic NN architecture is the perceptron or neuron. It takes multiple inputs, multiplies them by a weight matrix, sums the output and applies a non-linear function to generate an output. The Feed-forward Network (as the NN illustrated in Figure 2.17) is constructed by arranging multiple perceptrons in a sequential manner and organizing them into different layers (Sazli, 2006). This type of NN is called a multilayer network, and it operates by forwarding information in one direction. Initially, data enters through the input layer. It then travels through a sequence of hidden layers before reaching the output layer.

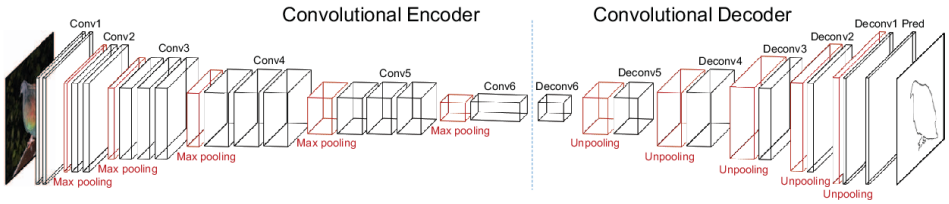


Figure 2.18 Example of a fully convolutional encoder-decoder network as presented in Yang et al. (2016) (Open source)

Convolutional neural networks (CNNs) are a type of neural network designed to analyze data with a grid-like structure, like images. They were initially introduced by LeCun et al. (1998). CNNs consist primarily of convolutional layers, which extract features from the input image using filters. These filters, essentially weight arrays, capture specific portions of the image. Additionally, CNNs often include pooling layers that shrink the image dimensions, reducing computational requirements.

Deconvolutional networks are a variant of CNNs that operate in a reversed fashion. While convolutional and pooling layers are commonly employed in CNNs to decrease the dimen-

sions of an image, deconvolutional networks work to restore the original image size. This is achieved through the use of upsampling and transpose convolutional layers. Unlike many other layers in neural networks, upsampling does not possess trainable parameters. Its role primarily involves duplicating the rows and columns of the image data according to specified dimensions.

Using a CNN to simultaneously extract image features, alongside a Deconvolution NN for subsequent feature processing and precise spatial localization (Figure 2.18), exemplifies the application of an Encoder-Decoder NN (Schmidhuber, 2015). Encoder-decoder NNs represent a class of deep learning models with paired neural networks. Typically, these two networks exhibit identical architectures, differing only in their operational directions, one working in a standard feedforward manner while the other operates in reverse.

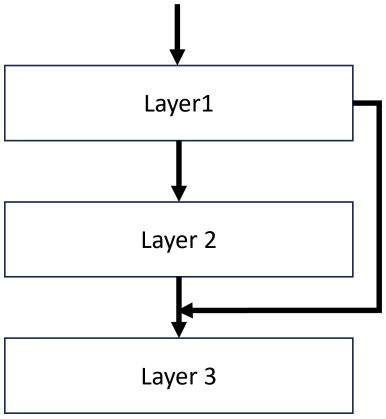


Figure 2.19 Skipt connections used in the ResNet architecture

ResNet represents a deep learning structure designed to facilitate the training of deep neural networks (He et al., 2016). Within this architectural pattern, information originating from prior layers is transmitted to the present layer via skip connections (Figure 2.19). The fundamental objective of this architecture is to mitigate the issue of vanishing gradients, which is achieved by incorporating the outputs from preceding layers into the ongoing layer's computations.

U-Net (Figure 2.20) is a fully convolutional neural network that uses a combination of a CNN encoder-decoder architecture along with skip connections (Ronneberger et al., 2015). The architecture of U-Net consists of two main parts: a contracting path (encoder) that uses convolutional and pooling layers, and an expansive path (decoder) that uses upsampling and transpose convolution layers. These two parts give U-Net its characteristic U-shaped architecture. Additionally, similar to ResNet, the decoder of U-Net leverages skip connections to combine features and spatial information from both the encoder and the expanding path.

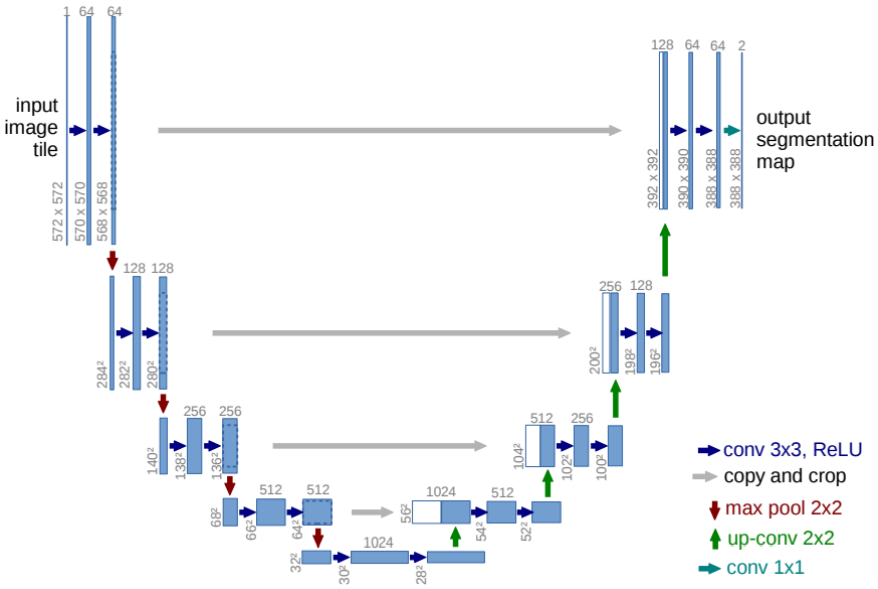


Figure 2.20 U-net architecture as presented in Ronneberger et al. (2015)

Recurrent Neural Networks (RNNs) are constructed to retain previously acquired knowledge and employ it for forthcoming predictions (Jordan, 1997). This structure is fed with sequential data, and its hidden internal states are revised each time it processes the subsequent sequence from the input data. An alteration of the basic RNN is the Long Short-Term Memory network (LSTM), first proposed by Hochreiter and Schmidhuber (1997). The LSTM incorporates gating mechanisms to store information across different time points effectively.

Generative Adversarial Networks (GANs) are a type of NN built upon unsupervised learning principles. They produce synthetic data by identifying patterns present within the input dataset (Goodfellow et al., 2014). Comprising two core elements, GANs consist of a generator responsible for fabricating synthetic data, and a discriminator assigned the task of categorizing input data as either real or artificial (Figure 2.22). The fundamental concept involves a competitive interplay between these networks, wherein the generator continually refines its output until it produces examples of such high quality that the discriminator can no longer reliably differentiate between synthetic and real data.

The Transformer NN architecture (Figure 2.23) was introduced to address memory limitations encountered by RNN and LSTM architectures (Vaswani et al., 2017). It relies on attention mechanisms to establish connections between an encoder and a decoder. The attention mechanism leverages all the hidden states from the encoder within the decoder. The underlying concept is rooted in the notion that valuable information could be present within each data point of the dataset. Consequently, at each sequential step, the attention mech-

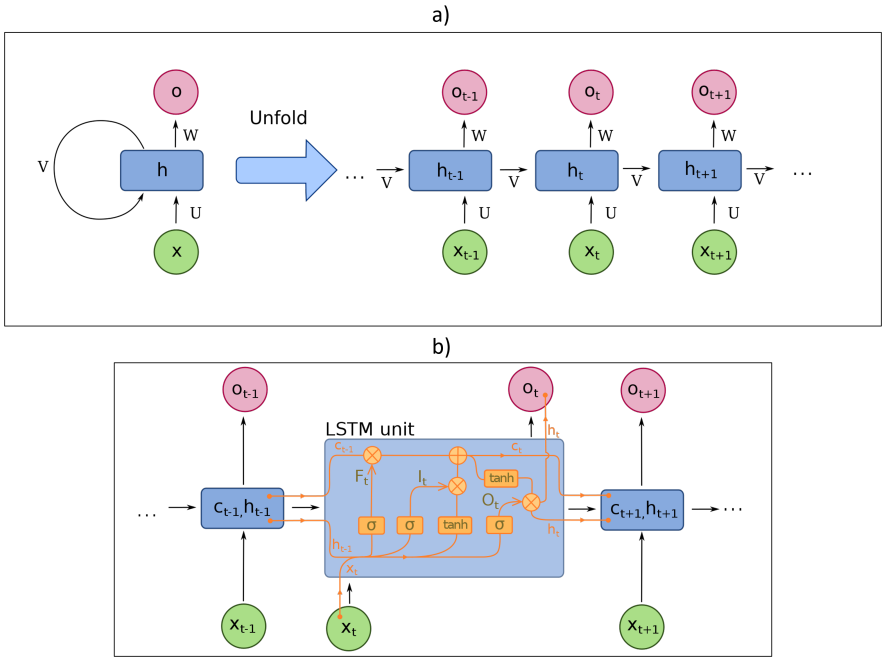


Figure 2.21 a) A basic recurrent neural network architecture in its compressed and unfolded representation. b) One-unite Long-Short-Term network (original images by F. Deloche under Creative Commons licence).

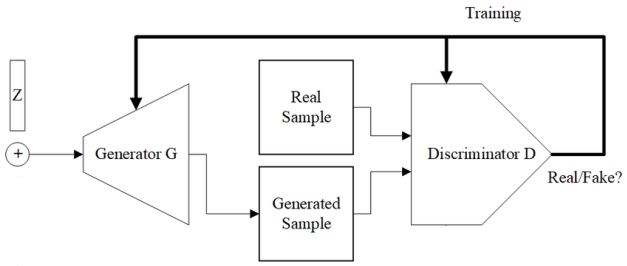


Figure 2.22 Standard GAN architecture as presented in Temizel et al. (2018).

anism learns the significance of different portions of the data, distinguishing between what holds importance and what is of lesser relevance.

In summary, each NN architecture possesses its own set of strengths and weaknesses, making their effectiveness depend on the nature of the input data they encounter. For instance, standard feedforward NNs are primarily applied in resolving classification and regression challenges tied to straightforward, structured data. On the other hand, NN models like RNNs, LSTMs, and Transformers excel at retaining information over extended periods, particularly when dealing with sequential data. Conversely, CNNs and deconvolutional NN are more adapted for handling grid-like structures. Since seismic images can typically be conceptual-

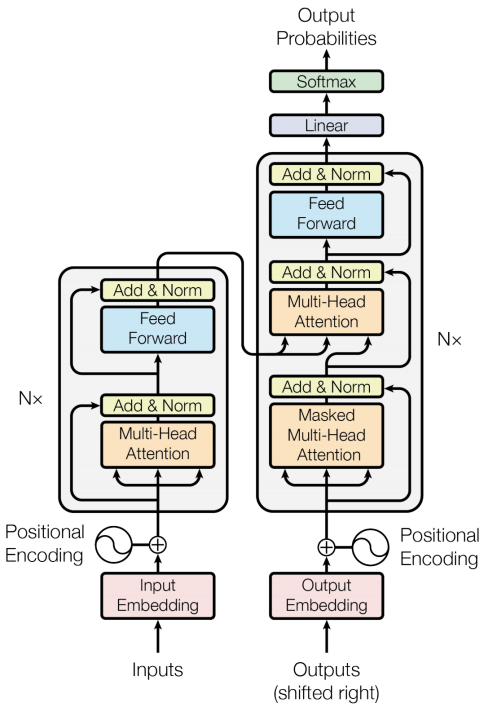


Figure 2.23 Model architecture of the transformer NN as presented in Vaswani et al. (2017)

ized as grid structures, employing a combination of CNNs and deconvolutional NN emerges as a suitable approach for effectively processing such datasets.

Convolutional neural networks

This thesis aims to develop a methodology based on deep learning to image seismic velocity variations caused by subsea permafrost. At the core of this approach lies a NN architecture based on CNN encoders and deconvolutional decoders. In the subsequent portion of this section, our focus will be on providing a more detailed explanation of the fundamental components of CNNs and deconvolutional NNs. We will commence with a comprehensive description of CNN encompassing convolutional and transposed convolutional layers, followed by elaborating on the role of the activation functions. To wrap up, we will provide a description of the evolutionary progression of CNNs over time.

A convolutional neural network (CNN) is a NN characterized by multiple hidden layers, including convolutional and pooling layers. These layers work to extract and process features from the input data (Valueva et al., 2020). Eventually, the CNN culminates with a fully connected layer, which further analyzes the extracted features to make final predictions or classifications.

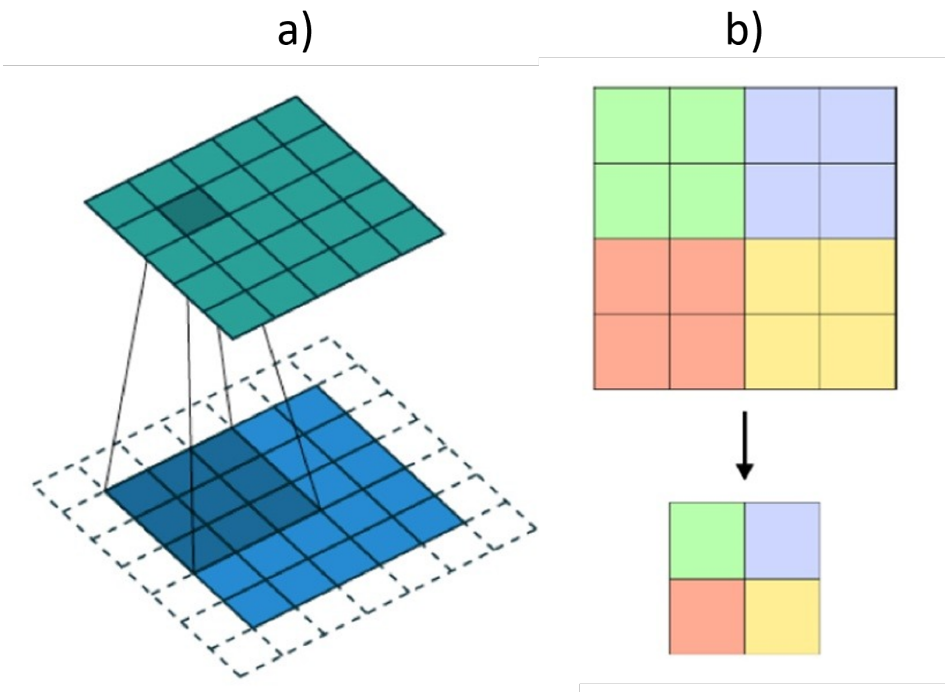


Figure 2.24 Sketch of convolutional (a) and pooling (b) layers as presented in Maier et al. (2019)

The convolutional layer (Figure 2.24) serves as the fundamental building block in a CNN. In the forward pass, the layer convolves a filter, an n-dimensional array of weights, over the input image's width and height to generate a 2D activation map for that filter (Goodfellow et al., 2016). Mathematically, the convolutional operation can be represented as follows:

$$x_{ijk}^{l} = \sum_{b,c,d} y_{b,j+c-1,k+d-1}^{l-1} \omega_{i,b,c,d} \tag{2.13}$$

In this equation, x_{ijk}^{l} denotes the output of the multichannel convolutional layer at position (i, j, k) before passing through the activation function, ω represents a 4D kernel tensor, and y_{ijk}^{l-1} corresponds to the input of the layer. Note that i, j refers to the pixel coordinates with respect to the height and width of the image, while k refers to the channel's position. Throughout the training process, the CNN learns to optimize the weights on the filters to activate when detecting specific types of features at certain spatial positions in the input. This learning is accomplished by updating the filters based on their ability to detect these features.

Common CNN architectures have multiple CNN layers followed by a pooling layer. The main purpose of the pooling layer is to perform a non-linear downsampling process (Géron,

2022). While various functions have been developed for pooling, the most widely used one is `max pooling`. This method involves dividing the input image into several rectangles and determining the maximum value within each sub-region. The mathematical expression for max pooling is as follows:

$$x_{ij}^{l} = \max_{m,n} y_{is_x+m,js_y+n} \tag{2.14}$$

Here, s_x and s_y represent the stride values in the horizontal and vertical directions, respectively, and the pooling window is defined at the output index (i, j).

The pooling layer plays a crucial role in progressively reducing the spatial dimensions of the representation. This reduction helps to decrease the number of parameters, memory usage, and computational load within the neural network. As a common practice, a pooling layer is often inserted between successive convolutional layers in a CNN architecture.

Transpose convolutional layers

The need for incorporating transposed convolutional layers arises from the intention to apply a transformation that operates in the reverse direction of a standard convolution operation (Dumoulin and Visin, 2016). This type of transformation is useful, for example, in serving as the decoding layer within a convolutional decoder, facilitating the projection of feature maps into a higher-dimensional space.

Transposed convolutions work by interchanging the forward and backward passes of a typical convolutional process. To elaborate, while a convolutional filter ω conventionally defines the convolution operation (Equation 2.13), the computations for transposed convolutions involve multiplying this filter with the transpose of $(y_{(i+a)(j+b)}^{l-1})$. The process of increasing spatial dimensions is executed through the specification of distinct strides and the use of padding, ensuring that the resulting output dimension is $\texttt{Output}_{size} = \texttt{input}_{size} \times \texttt{stride}$.

CNN evolution

The Neocognitron, as presented by (Fukushima, 1980), stands as an early precursor to CNNs. In his publication, Fukushima (1980) introduces fundamental concepts, including feature extraction, pooling layers, and the integration of convolutions within a NN. The terminology "convolutional neural networks" was later accepted following the efforts of LeCun et al. (1998), who unveiled the LeNet-5 architecture tailored for handwritten character recognition. This architecture design comprises four convolutional layers and two fully connected layers that complete the network's structure.

The benefits of employing CNNs for image classification became evident through their participation in the ImageNet Large Scale Visual Recognition Challenge. This annual competition took place from 2010 to 2017 and aimed to assess various algorithms for object detection and image classification on a large scale. Among the successful entries, noteworthy outcomes emerged from the work presented by Alexnet (Krizhevsky et al., 2012). Their approach used five convolutional layers and three fully connected layers, culminating in a final softmax operation to classify an extensive dataset containing 1.2 million images across 1000 distinct classes.

ZFNet (Zeiler and Fergus, 2014) adopted a similar architecture to AlexNet but with a meticulous tuning of the hyperparameters. VGGNet (Simonyan and Zisserman, 2014), on the other hand, explored the impact of convolutional network depth on accuracy in large-scale image recognition. This investigation led them to secure the first and second positions in the competition. Inception (Szegedy et al., 2015) employed a CNN with an impressive arrangement of 22 layers, devoid of fully connected layers. Additionally, ResNet (He et al., 2016) introduced a novel concept of residual blocks, which are interconnected through skip connections, further advancing the state-of-the-art of the field.

After the results presented during the Imagenet challenges, DenseNet (Huang et al., 2017) expanded the idea of residual connectors by connecting the outputs of each layer with every other layer in a feed-forward fashion. Thus, for each layer, the feature maps of all preceding layers are used as inputs, and its own feature maps are used as inputs into all subsequent layers. This approach alleviates the vanishing-gradient problem, strengthens feature propagation, encourages feature reuse and reduces the number of parameters. However, the feature maps must be the same size for this method to be applied.

The impact of scaling CNNs was investigated by Tan and Le (2019). This research Tan and Le (2019) leads the authors to a novel CNN architecture named EfficientNet. This architecture demonstrated notable accuracy while maintaining a reduced size and enhanced speed compared to previously suggested CNNs.

In a subsequent advancement, Xie et al. (2020) introduced a semi-supervised technique named "Noisy Student Training." This method involved a pair of CNNs, a teacher and a student. The teacher network underwent training on a subset of the dataset to formulate pseudo-labels, which were then employed to train the student network alongside the originally labelled images. By alternating the roles of teacher and student, Xie et al. (2020) achieved a considerable enhancement in accuracy on the ImageNet dataset, surpassing the existing state-of-the-art performance.

Later, Pham et al. (2021) modified the Noisy Student training approach by enabling the

refinement of teacher-generated pseudo-labels based on the feedback obtained from the student's performance. This adaptability empowered the teacher network to generate improved pseudo-labels for effectively instructing the student network. The results are presented in a new architecture named "Meta pseudo labels".

Activation functions

Evaluating activation functions at the output of each neuron accounts for the non-linear capabilities of an NN. Without an activation function, an NN acts as a linear regression model, limiting its performance and effectiveness (Sharma et al., 2017). In the field of machine learning, various non-linear activation functions are commonly used (Szandała, 2021). The following list presents some examples of commonly used activation functions

- The Binary Step Function, the simplest activation function, which yields 1 for positive neuron outputs and 0 for negative outputs. It has a derivative of 0 throughout its interval. Another type is the linear activation function, which produces a value directly proportional to the input and maintains a constant gradient.
- The sigmoid activation function transforms input values to a range between 0 and 1, represented by the function $\sigma(x) = 1/(1+e^{-x})$. It is continuously differentiable, and its derivative is given by $\sigma'(x) = \sigma(x)(1-\sigma(x))$. Similarly, the tanh function maps input values to a range between -1 and 1 following the function $tanh(x) = (e^{z}-e^{-z})/(e^{z}+e^{-z})$, with its derivative being $tanh'(x) = 1 - tanh(x)^2$.
- The rectified linear unit (ReLU) activation function (Agarap, 2018), defined as $ReLU(x) = max(0, x)$, is more efficient than other functions since it keeps some neurons inactive.
- A modified version of ReLU is the leaky ReLU function (Maas et al., 2013), expressed mathematically as $LReLU = x$ for $x > 0$, and $LReLU = 0.01x$ otherwise. Its derivative is 1 for $x > 0$ and 0.01 otherwise.
- Another variation of ReLU is the exponential linear unit (ELU) defined as $ELU(x) = x$ for $x > 0$ and $ELU(x) = a(e^{x} - 1)$ otherwise, where a is a positive constant (Clevert et al., 2015). The derivative of ELU is 1 for $x > 0$ and $ELU'(x) = f(x) + a$ otherwise.
- The softmax activation function (Bridle, 1989) comprises multiple sigmoid functions, providing a vector of values ranging from 0 to 1. Mathematically, the softmax activation function can be represented as $S(x)_i = e^{z_i} / \sum_{j=1}^{N} e^{z_j}$ and its derivative is $S'(x)_i = S_i(\delta_{ij} - S_j)$, where δ is the Dirac delta function.

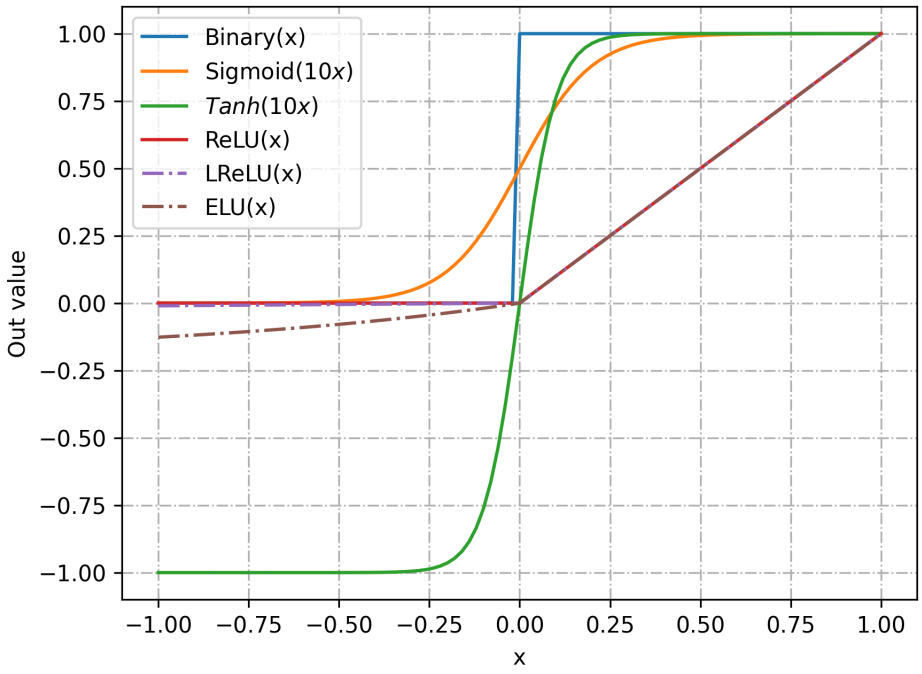


Figure 2.25 Output values for different activation functions.

Figure 2.25 summarizes the results of evaluating various activation functions from -1 to 1. Like the neural network architecture, each activation function possesses its own advantages and disadvantages, making them better suited for specific tasks over others. For instance, when dealing with binary classification tasks, the binomial, sigmoid, and tanh functions find more frequent use. Conversely, the ReLU, LReLU, and ELU functions are more commonly employed in predictive modelling scenarios. Furthermore, the softmax activation function demonstrates greater suitability for tasks involving multivariable classification.

2.3.3 Deep learning for seismic inversion

Employing deep learning techniques for seismic inversion offers numerous advantages compared to traditional and advanced seismic processing methods. One key benefit is that this approach learns directly from the dataset, eliminating the requirement to identify and isolate different arrivals, a task inherent in conventional seismic processing methods. Unlike state-of-the-art Full Waveform Inversion (FWI), which demands a meticulously defined initial model, deep learning employs an optimization algorithm to reduce errors without the need for such precise initialization. Furthermore, despite the potential computational demands during the training of an NN, the subsequent dataset evaluation becomes a straightforward task. This evaluation process can be easily extended to encompass large-scale datasets after the model

has been trained.

In recent years, deep learning has emerged as a promising and viable solution for handling seismic data and seismic inversion. The pioneer work by Araya-Polo et al. (2018) demonstrated the application of NN in velocity model building (VMB) using 3D semblance images. The outcomes of this study indicated that deep learning could effectively identify subsurface structures while being cost-efficient.

Expanding the advancement described above, Alfarraj and AlRegib (2019) developed a recurrent neural network (RNN) for elastic impedance inversion, incorporating multi-angle seismic data within a semi-supervised modelling framework. In this workflow, well-log data was used to guide the inversion process. The outcomes demonstrated exceptional results, achieving an average correlation of 98% for the benchmark Marmousi 2 model (Martin et al., 2006) by implementing the semi-supervised approach. Despite these remarkable achievements in semi-supervised learning, it is essential to acknowledge that access to well-log information is generally limited.

Fabien-Ouellet et al. (2020) discussed the simplified task of 1D velocity estimation using neural networks for general velocity inversion. Inspired by semblance analysis, they designed an NN for estimating 1D layered velocity models. The performance of this NN was evaluated using synthetic data and real 2D marine data. The results demonstrated that the NN can accurately estimate velocity from seismic data. Interestingly, even when trained on synthetic data, the NN performed well on real data, highlighting its practical applicability.

In another investigation, Mustafa et al. (2021) presented a bidirectional transfer learning technique to retrieve impedance values. This approach allows for utilizing mutually beneficial information from multiple datasets to enhance individual estimations across all datasets. The authors claim this method leads to more resilient estimations, particularly when dealing with noisy seismic data and limited well-log information.

Similarly, Simon et al. (2023) used a transfer learning approach to train a neural network. Initially, the NN was trained on horizontally layered velocity models, and subsequently, an augmented network was trained to estimate 2D dipping layered models. Their results demonstrate that applying transfer learning techniques can facilitate the use of deep learning for inverting 2D and 3D surveys by lowering the data requirements for training.

GANs (Generative Adversarial Networks) have proven to be a successful approach in seismic data analysis for estimating velocity and impedance models. In this method, two neural networks, namely the generator and the discriminator, engage in a competitive process to improve their predictions. The main objective of the generator is to produce synthetic out-

puts, while the discriminator aims to distinguish between the artificially generated outputs and the real ones it receives.

An illustrative example of GAN's effectiveness in seismic data analysis is presented in the work by Zhang and Lin (2020). They used GAN to reconstruct subsurface velocities and designed the generator with an encoder-decoder architecture that takes shot gathers as input and outputs velocity models. The discriminator, in this case, employed a CNN architecture. The study's findings demonstrated that GAN outperformed state-of-the-art methods like Full-Waveform Inversion (FWI) in producing inverted models.

Wang et al. (2022) introduced a GAN method that simultaneously addresses seismic inversion and forward calculation. The inversion generator used a U-net architecture with an encoder-decoder structure to produce synthetic acoustic impedance models based on seismic data. Additionally, the forward GAN employed the U-net approach to map seismic data to acoustic impedance models. The system also included two discriminators to distinguish between synthetic and real seismic data as well as acoustic models. Their study demonstrated that this combined inversion and forward calculation approach significantly improved the lateral continuity of the generated models.

Marano et al. (2023) offers a critical assessment of the current literature concerning using GANs in the seismic domain. Furthermore, Adler et al. (2021) conducts an in-depth review of the emerging neural network techniques used for seismic inversion. Additionally, Mousavi and Beroza (2022) presents a comprehensive overview, including the trends, obstacles and potential advancements associated with the application of deep learning in seismology. This extensive analysis is founded on the meticulous examination of 637 papers published between 1988 and 2022.

2.3.4 Challenges of using deep learning

Although deep learning has demonstrated effectiveness in seismic inversion, it also introduces its own challenges. Some of those challenges include the absence of labels, data representativity, the large size of the datasets, the varying size of the seismic data, generating a training set and the uncertainty of predictions.

Training a deep neural network for seismic inversion necessitates a substantial dataset that accurately mirrors the underlying data distribution. In the case of supervised learning, it is essential to have precise labels for every sample. In the context of seismic inversion, the training dataset comprises pairs of seismic data (input) and models for P-wave and S-wave velocity, as well as its attenuation (labels/outputs) (Araya-Polo et al., 2018). While abundant

seismic data is available, obtaining accurate labels is challenging since the true Earth velocity is rarely known. Consequently, many deep learning approaches in seismic inversion rely on synthetic data, where velocity and seismic response can be computed precisely (Mousavi and Beroza, 2022). However, for a synthetic-trained NN to be applied to a real target dataset, the distribution of the training set must closely resemble the unknown distribution of the actual subsurface.

The representativeness of data (Clemmensen and Kjærsgaard, 2022) is, therefore, a concern related to deep learning for seismic inversion. In statistical terms, representative sampling is a method that allows us to make conclusions about a larger group based on a smaller selection (D'Exelle, 2014). In the domain of deep learning, this means that the dataset used for training should accurately mirror the characteristics of the overall target data. This alignment is crucial so that the learned patterns and features during training can be effectively applied to the entire target dataset. For instance, in the case of using deep learning for estimating seismic parameters, research by Fabien-Ouellet et al. (2020) shows that even when the NN is trained on synthetic data, it performs well on real data, showing its practical usability. However, it is important to note that to ensure the representativeness of the samples, the synthetic data must be generated to mimic the geological structures expected in the target dataset. For instance, Liu et al. (2021) proposed a method to construct realistic structural models for deep learning networks, including dense-layers, faults and salts within the models, while Wu et al. (2020) proposed a method to build diverse structure models with realistic folding and faulting features.

Creating a training dataset that accurately reflects the recorded data is a significant challenge when using deep learning for seismic inversion (Fabien-Ouellet et al., 2020; Liu et al., 2021; Wu et al., 2020). To overcome this challenge, it is necessary to use all available information Zhang et al. (2021). For instance, stratigraphy information can be used as a starting point for generating synthetic data models. Moreover, incorporating additional data, such as information from well logs and ocean bottom tomography, can further improve the representativeness of the training dataset. Additionally, including information about the hydrological and geological processes that occur in the zone can help create a more accurate dataset.

Furthermore, the survey parameters play a crucial role in generating synthetic data that is representative of the recorded data. For instance, it is imperative to analyze the frequency content of the synthetic data carefully to ensure that it accurately represents the recorded data Yuan et al. (2022); Zheng et al. (2019). By doing so, we can create synthetic data that is more representative of the recorded data and achieve better results in seismic inversion with deep learning.

Dealing with the size of the dataset is also a major challenge that need to be overcome for inverting seismic data using NN. Seismic surveys typically collect a significant amount of data. For example, the ARA04C and ARA05C surveys collected approximately 25,000 shot gathers (Jin and Dallimore, 2016; Jin et al., 2015). Each shot gather in these surveys includes traces from 120 receivers that recorded samples every 0.0001 seconds for 10 seconds. To reduce the size of the dataset, we can downsample the time domain to 0.002 seconds, which accounts for a Nyquist frequency of 250 Hz, covering the frequency band of interest. Additionally, since we expect permafrost to be located in the upper 700 meters depth, we can analyze only the first part of the recorded data (i.e., 2 seconds).

Moreover, to train an NN, we require a training dataset that contains information about the seismic data and the corresponding label (Araya-Polo et al., 2018). The dataset must have the same features as the recorded data, including survey geometry and sample rates. Although we can reduce the size of the dataset using the same approach as the recorded data, the resulting size of the dataset is still significant. For instance, while developing this thesis, we created two seismic datasets, a 1D dataset and a 2D dataset, with sizes of 175 GB and 650 GB, respectively. Although stochastic gradient descent optimization can help to deal with such a large dataset, we still need large computational resources to train the NN efficiently.

Inverting seismic data using deep learning poses an additional challenge due to the varying sizes of the data. For example, during the 1970s and 1980s, a large number of seismic surveys were collected in the Beaufort Sea as oil and gas exploration was at its peak. However, since multiple companies conducted the surveys at different times, there are differences in the survey geometry of the recorded data. As the training dataset needs to have the same features as the recorded data, these differences in acquisition geometry can lead to errors in the inverted models. One way to address this issue is by transforming the shot gather to domains that are not sensitive to the survey geometry, such as the frequency-velocity domain (Ibrahim and Fabien-Ouellet, 2021; Ibrahim et al., 2018). However, in this thesis, we propose using seismic data in multiple domains for seismic inversion, and limiting the domain transforms could restrict our approach. Another solution could be to apply seismic data interpolation or regularization to transfer samples from irregularly recorded locations to locations on a regular grid, which is the standard seismic interpolation techniques (Spitz, 1991). Moreover, Wei et al. (2023) proposed using deep learning through denoising diffusion probabilistic networks to interpolate seismic data.

Finally, a less evident challenge associated with seismic inversion using a deep learning approach is related to the difficulty of estimating the uncertainty of the predictions. The reason

behind this challenge is the insufficient information about the true seismic velocities. FWI handles this situation by comparing synthetic data generated using the inverted model with the recorded data. However, this comparison may not be advantageous for deep learning as the NN will not necessarily be trained to reproduce the synthetic data. Instead, to estimate the uncertainty we can compare the inverted models with information from well logs data. However, this estimate will be local and may not represent the general uncertainty of the predictions.

CHAPTER 3 ARTICLE 1: UNDERSTANDING SEISMIC VELOCITY VARIATIONS OF SUBSEA PERMAFROST: A SENSITIVITY STUDY

Bustamante, J., Fabien-Ouellet, G., Duchesne, M. J., & Ibrahim, A. (2023). Understanding seismic velocity variations of subsea permafrost: A sensitivity study. Geophysics, 88(5), R655-R669.

Initial submission on September 7, 2022. Published online on July 27, 2023. Printed version on September 2023.

3.1 Abstract

Monitoring the change in permafrost conditions and distribution is crucial for forecasting global warming. As seismic velocities increase with ice content, marine seismic surveys can map the top of ice-bearing subsea permafrost on a large scale. However, conventional seismic methods cannot fully map internal velocity variations linked to changes in ice content because the removal of guided waves and multiples is challenging for reflection processing. Nevertheless, these arrivals carry information about velocity variations with depth. We investigate if a joint analysis of various wave arrivals could provide information about velocity variations within permafrost and accurately estimate the permafrost thickness. Through a sensitivity analysis, we estimate the feasibility of using different wave arrivals, such as reflections, refractions, multiples, and guided waves, to better characterize permafrost conditions. We show that guided waves are sensitive to velocity variations within permafrost and can detect the depth of the permafrost's base under certain geological conditions. Then, we combine the analysis of all seismic arrivals to derive a subsurface permafrost model for a seismic line collected in the Beaufort Sea. The joint analysis reveals the transitions in depth between ice-bonded and partially ice-bonded permafrost and offers a crude estimate of permafrost thickness.

3.2 Introduction

Subsea permafrost is defined as any frozen ground under the sea level that remains frozen for more than two years (Osterkamp, 2001). It is widespread on the continental shelves of the polar regions. Subsea permafrost was initially formed at the surface when the shelves were exposed during glaciations. As the glaciers retreated, the sea level rose and flooded part of that relic permafrost. Subsea permafrost slowly degrades as the overlying water body warms

it (Romanovskii et al., 2005). The rate of degradation has recently increased in response to global warming (Shakhova et al., 2017b).

Subsea permafrost acts as a methane reservoir and keeps gas hydrates stable under and inside the permafrost (Dallimore and Collett, 1995; Dallimore et al., 2015). Therefore, subsea permafrost degradation is not only a consequence of global warming but can also accelerate global warming through the emission of greenhouse gases through a feedback effect (Anthony et al., 2012; Schuur et al., 2008). Since temperatures have increased by more than twice the global average in the Arctic over the last two decades (Meredith et al., 2019), monitoring the state of subsea permafrost and its lateral and vertical distribution is crucial to forecast global warming.

Various direct and indirect methods exist to estimate permafrost's lateral and vertical extents. For instance, borehole logging performs direct measurements of rock properties to determine the vertical extent of subsea permafrost. For example, Hu et al. (2013) successfully used well seismic surveys and geophysical well logs to determine the permafrost base in the Mackenzie Basin, and Ruppel et al. (2016) used a similar technique to characterize permafrost in the Canadian Beaufort Sea.

Although direct methods such as borehole logging can accurately estimate the vertical extent of subsea permafrost, it provides a limited lateral extent estimation. Moreover, it is prohibitively expensive for mapping large areas and introduces distribution uncertainty between borehole locations. On the other hand, indirect geophysical techniques can estimate the lateral extent of permafrost with acceptable cost. Such methods are based on measuring physical properties affected by ice, such as electrical resistivity or seismic impedance.

Electromagnetic methods can determine the extent of subsea permafrost from the response of the ground to the propagation of an electromagnetic field used to estimate the subsurface electrical conductivity. In permafrost zones, a sudden increase in the electrical resistivity values is expected because of the free ions expelled from the ice structure during the freezing process. An example of electromagnetic surveys used to map subsea permafrost is presented in Sherman et al. (2017) who used a controlled source electromagnetic method on the Beaufort shelf in Alaska. The main disadvantage of electromagnetic methods for determining the extent of subsea permafrost is the low resolution of the method at depth. This is caused by the limitations of the inversion and the high electromagnetic attenuation of the high conductivity of water.

Alternatively, seismic methods detect the presence of subsea permafrost since the ice in the sediment pore space increases the velocity of seismic waves (Dou, 2015). Seismic surveys are a reliable tool to measure the lateral extent of subsea permafrost (Hunter et al., 1976; Kang

et al., 2021). Brothers et al. (2012a) used seismic refraction to map permafrost distribution in the US Beaufort Sea. Brothers et al. (2016) utilized root mean squared (RMS) velocities derived from semblance analysis of seismic reflections to map the subsea permafrost on the continental shelf off Alaska. More recently, Kang et al. (2021) employed Full Waveform Inversion (FWI) of seismic waves in the Laplace domain for inferring seismic velocities of permafrost bodies in the Beaufort Sea. Other examples of seismic analysis of permafrost regions are presented by Rekant et al. (2005) (Yugorsky Peninsula, Russia), Johansen et al. (2003) (Svalbard Archipelago, Norway), Hilbich (2010) (Swiss Alps), and Ramachandran et al. (2011) (Mackenzie Delta, Canada).

Although most of the studies mentioned above were able to map the lateral extent of the permafrost accurately, none fully described its vertical extent. Seismic refraction is inherently unable to estimate the permafrost thickness because it cannot handle the velocity inversion found at the transition from ice-bearing to water-bearing sediments (Geldart and Sheriff, 2004). On the other hand, seismic reflection, in theory, can detect the base of the permafrost, at least when a sharp transition from ice-bearing to unfrozen sediments occurs. However, at typical frequencies used for seismic surveying, the strong seismic impedance contrast between frozen sediments and the water column creates a waveguide that traps energy in the overlying water column. Guided waves and multiples dominate seismic recording in such conditions. Early reflections, such as the reflection from the permafrost base, are hidden under these highly energetic arrivals and are hardly retrieved even with the most sophisticated denoising approaches.

Nevertheless, other arrivals can carry useful depth information. For instance, guided P-waves and Scholte waves interact with the seafloor, and their dispersion carries depth information. As the seismic waves at different frequencies have different depths of penetration, the guided wave modes propagate at different wave velocities, which can be used to estimate with depth-dependent seismic properties (Klein et al., 2005; Wang et al., 2016). For instance, Kugler et al. (2007) and Li et al. (2018) use Scholte waves and acoustic guided waves in a full-wavefield inversion of P-wave and S-wave velocities of the shallow subsurface. Boiero et al. (2013) use the secular function describing modal curves to estimate the seismic velocity on shallow marine data. Although it has been proposed that guided P-waves and free surface multiples can be used to estimate the depth and thickness of subsea permafrost (Duchesne et al., 2022; Fabien-Ouellet et al., 2020), to our knowledge, there is no published attempt on this purpose.

In this paper, we study the feasibility of using seismic data to estimate the vertical extent of subsea permafrost. We first show common features found in seismic data collected in a

marine environment where the occurrence of subsea permafrost is known from a 2D dataset collected in the Canadian Beaufort Sea. Then we review how ice content impacts seismic velocities. Afterward, we perform a sensitivity analysis from a simplified subsea permafrost model to study how and when depth to the base of permafrost can be retrieved from different seismic arrivals. Finally, based on the sensitivity analysis, we combine different wave arrivals to build a velocity model for a seismic line and we compare the results with conventional travel-time first break tomography.

3.3 The seismic signature of subsea permafrost

The present study focuses on seismic line ARA05C-06, collected in 2014 in a joint effort between the Korea Polar Research Institute (KOPRI) and the Geological Survey of Canada (GSC). The objective of the seismic survey was to investigate geological conditions, subsea permafrost, and gas hydrates on the outer continental shelf and upper slope of the Canadian Beaufort Sea (Jin and Dallimore, 2016). Figure 3.1 presents the map of the ARA05C seismic survey. Line ARA05C-06 is perpendicular to the shelf and intersected by wells I-44 and M-13. The seismic source consisted of a 1200 in^3 G-gun array fired every 50 m. The seismic streamer contained 120 channels with a 12.5 m spacing, with the streamer and airgun arrays kept at a constant depth of 6 m.

The Beaufort Sea shelf extends over 120 km offshore from the Mackenzie delta in the Beaufort Sea and narrows to 10 km in the west. Subsea permafrost extends offshore from thick permafrost ($\leq$700 m) located close to the shore to thin or no permafrost at the shelf break limit ($\sim$100 m isobath) (Taylor et al., 2013). Ice distribution can thus be highly variable along the shelf and exhibit sudden lateral changes in zones of discontinuous and degrading permafrost. However, the knowledge of the true ice distribution in the Beaufort Sea is still an open question due to the limited capability of geophysical methods to accurately detect such changes.

The Beaufort shelf stratigraphic units (Dixon et al., 2019) include the Kugmallit formation associated with the most recent basin extension, the thick Mackenzie Bay formation (atop the Miocene unconformity), the Akpak formation, and the Pliocene-age Iperk formation that has been affected by slope failures at the outer shelf. The glaciations have incised the most recent units by up to 300 m in some places, and Quaternary deposits have filled depressions such as the Mackenzie trough, a unique geological feature of the Canadian Beaufort Sea, connecting the Mackenzie River Delta to the continental slope. The Mackenzie trough is considered a young paleo-valley formed by glacial erosion during the Late Wisconsinan (25 to 10 kyr) filled by $\sim$400 m of unconsolidated sediments (Batchelor et al., 2014; Duchesne

et al., 2020). In the Canadian Beaufort Sea, relict permafrost can be found in sedimentary rocks or Quaternary sediments just under the seafloor.

The line is representative of most marine data acquired over subsea permafrost, both in terms of seismic acquisition and permafrost conditions. Line ARA05C-06 shows continuous and discontinuous permafrost, as well as zones with ice-bounded and ice-bearing permafrost (Figure 3.2a).

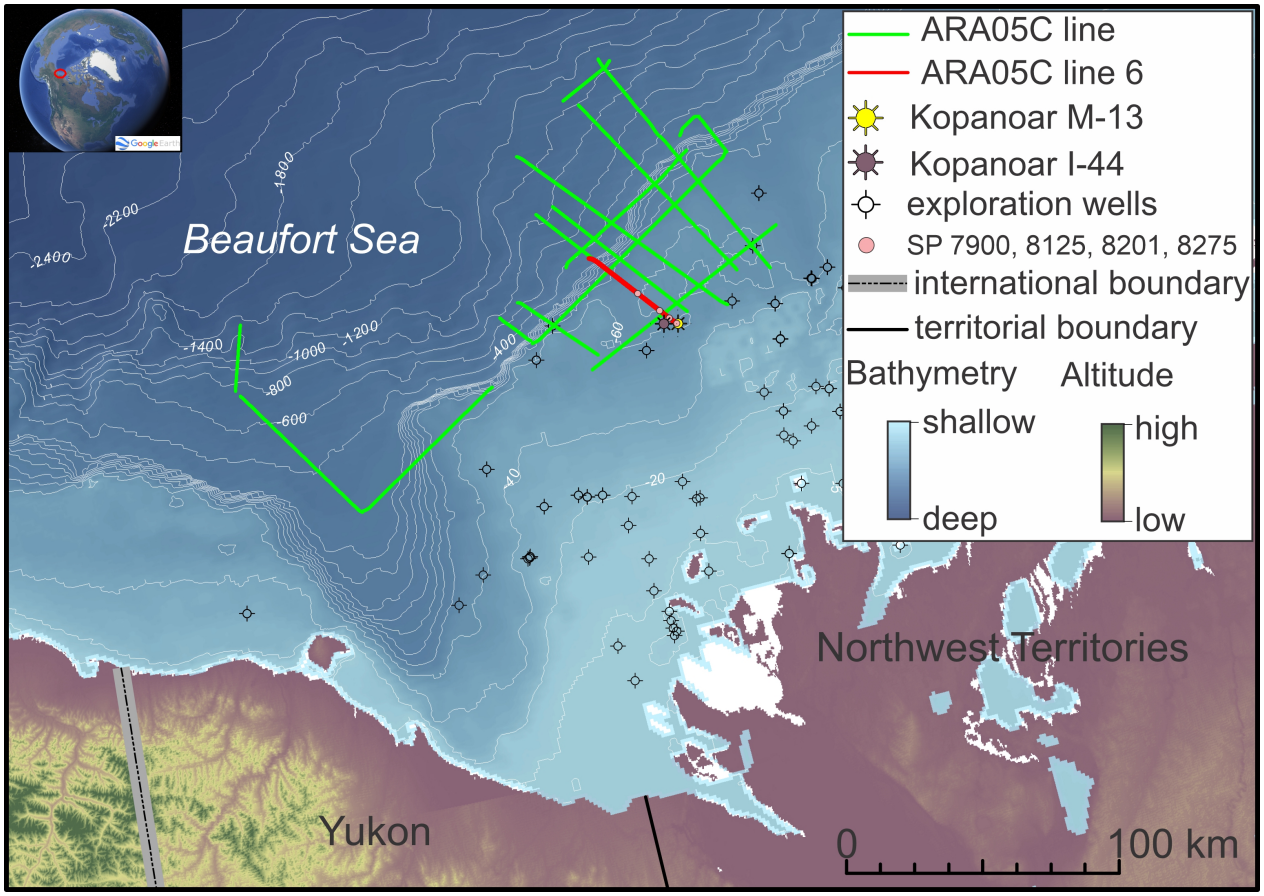


Figure 3.1 Survey map of the ARA05C expedition. Line 06 discussed in this study appears in red. Shot point (SP) numbers increase towards the SE. Crystal cable velocity profiles are available in wells Kopanoar I-44 and M-13.

Figure 3.2a shows a portion of the brute stack seismic section over 22 km of line 06. The top of permafrost (TP) can be identified on most of this portion of the line by a continuous, strong event. Previous studies in this zone identified sediments with low ice content on the NW end of the stack section (Pullan et al., 1987). A continuous ice-bonded sediment region is present on the SE end (over 18 km in 3.2a). Note that there is a transition zone between the low-ice-content and the ice-bonded regions (i.e., between 13 and 18 km). That transition zone

is characterized by undulated events associated with patchy and discontinuous permafrost distributions. The limits of permafrost patches act as diffractors, where a phase change of the pore space occurs (Bustamante et al., 2021).

Even though seismic methods detect the top of the permafrost, it fails to provide information on ice content or its depth extent. In fact, numerous free-surface multiples (M) hide deeper reflections and, most importantly, the reflection from the base of the permafrost. These highly energetic multiples are created by the large impedance contrast between water, frozen and unfrozen sediments that happen over short vertical distances. Given their high energy and interference with other arrivals, removing these multiples during processing frequently fails, even with the most advanced multiple removal workflows.

Refraction analysis is another robust method to detect the presence of permafrost (Brothers et al., 2012a; Hunter et al., 1976). The seismic refraction method provides valuable information about the presence of permafrost, can estimate its depth below the seafloor, and can be used to infer ice content through velocity variations. However, because seismic refraction cannot detect velocity inversions, the method cannot locate the permafrost base. Figures 3.2b-e display the observed shot gathers for SP 7900, 8125, 8201, and 8275, respectively. The analysis of refractions provides velocities of 1770 m/s, 1820 m/s, 2220 m/s, and 3530 m/s for SP 7900, 8125, 8201, and 8275, respectively. The velocity increases towards the coast, most likely caused by increased ice content and the resulting bounding effect.

By further examining the shot gathers of Figures 3.2b-e, several wave arrivals can be identified in addition to reflections and refractions. Strong reflection multiples (RE-M) interfere with primary reflections for near offsets. In fact, primary reflections are difficult to discern from these shots gathers. Note that below the refracted waves (RA), a train of refracted multiples (RA-M) interferes with each other and potentially hides deeper refractions. The refractions show varying amplitude decay with offset (RD) along the seismic line. We will later show that this decay can be associated with changes in the ice content and the thickness of the frozen layer, as hinted by (Fabien-Ouellet et al., 2020).

In the far field, the interference of multiples and refractions can be described as guided P-waves, which have a modal and dispersive nature. In fact, the near-surface permafrost acts as a leaky waveguide (Jensen et al., 2011), created by the high impedance contrasts between the water and permafrost. The dispersive and modal nature of guided P-waves is best seen in the F-V domain through dispersion plots (Wang et al., 2016). Dispersion plots for SPs 7900, 8125, 8185, and 8275 are presented in Figures 3.2f-i. We observe the characteristic modal distribution for all the shot points. Higher phase velocities are linked to higher refraction velocities in the T-X domain. Those modes are discontinuous and, in some cases, have very

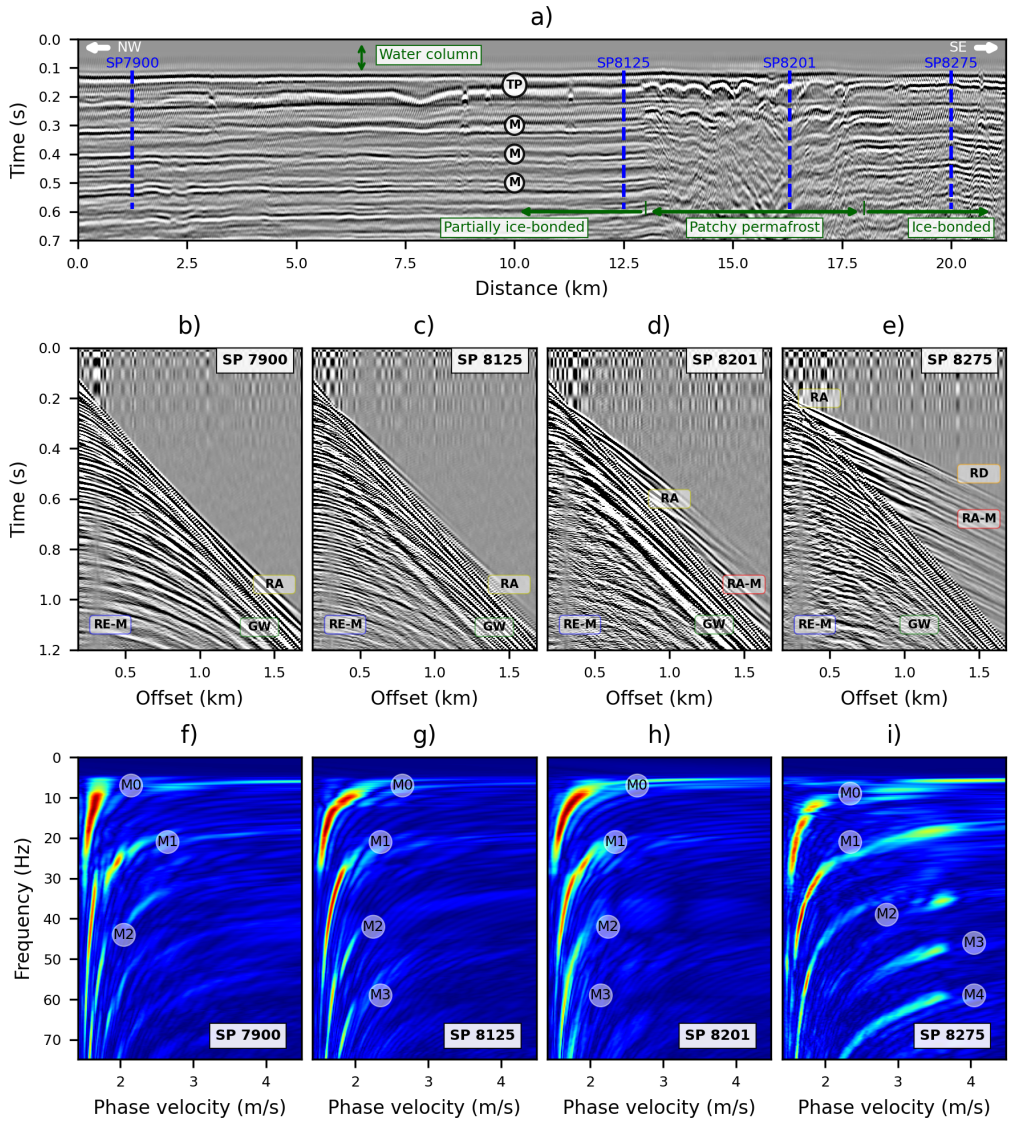


Figure 3.2 a) Stack section of line06 ARA05C survey indicating the top of the permafrost (TP) and the free surface multiples (M). Blue lines mark the position of shot gather used to infer velocity models. Example shot gathers are presented for locations SP 7900 (b), SP 8125 (c), SP 8201 (d) and SP 8275 (e). Guided P-waves (GW), reflection multiples (RE-M), refracted waves (RA) and its multiples (RA-M) interfere with primary reflections. Corresponding dispersion plots are presented in f-i. M0 to M5 are the different dispersion modes identified on f-i.

low energy, especially at high phase velocity, which makes a traditional modal phase velocity inversion challenging.

All the aforementioned wave arrivals or wave attributes can potentially carry information on lateral and vertical velocity variation of permafrost. The question whether they are sensitive enough to map velocity variations within the permafrost or to locate the base of the permafrost is the subject of the rest of the paper.

3.4 Seismic Velocity of subsea permafrost

The unique seismic signature of subsea permafrost is created by the significant velocity changes that occur when ice is present in the pore space. In fact, water can be frozen, partially frozen, or even completely unfrozen under 0 ^{o}C. Since ice has a higher seismic velocity ($V_P \sim 3.96$ km/s) than water ($V_P \sim 1.43$ km/s), higher ice content leads to an increase of V_p. Hence, the velocity of permafrost is highly dependent on ice content and bonding characteristics between the pore space and the matrix, which in turn depends on several factors (e.g., water salinity, multi-phase pore filling, and clay content, among others) (Dou, 2015).

Another important factor impacting permafrost velocity is grain size. King et al. (1988) proposed that in the presence of clay, two interfaces are contained in a porous medium, a silicate-water-ice interface and a silicate-water-silicate interface where there is a continuous layer of unfrozen water surrounding each silicate grain. The thickness of this water layer decreases with temperature. They also suggested that ice forms faster in coarse-grained materials, given the lower surface forces. Comparisons of seismic velocities measured on coarse and fine-grained samples agree with this theory (Dou, 2015; King et al., 1982). Further studies (King, 1984) also suggest that an increase in the clay fraction can induce a reduction in P-wave seismic velocities. King (1984) also found that for unconsolidated permafrost, there is a correlation between P-wave seismic velocities and water-filled porosity, independent of original porosity, clay content, or temperature.

An analysis of the pore-water salinity effects of seismic velocities on permafrost samples is presented in Dou et al. (2016). They observed that an increase in salinity has two main effects on P-wave velocity. First, it reduces the freezing point, characterized by increased seismic velocities as ice forms. Second, a reduction in the P-wave velocity with salinity is observed in the interval between the freezing point, where the ice starts forming, and the eutectic point, where all water has solidified. Due to salinity, ice content increases gradually with decreasing temperatures between the freezing point and the eutectic point. They note

that measured P-wave velocities, amplitudes, and centroid frequency are not highly affected by salinity over the freezing point and under the eutectic point.

From the analysis presented above, it is clear that temperature is not the only factor explaining ice content variations in permafrost. In fact, the ice content and permafrost consolidation can be highly variable for the same temperature due to porosity, grain size, and salinity variations. This is evidenced by well loggings, where extreme velocity variations can be observed over very short intervals within ice-bearing permafrost. Velocity variations are thus a good indicator of ice content and permafrost conditions. In fact, seismic methods can detect permafrost varying from ice-bearing to ice-bonding conditions (Vosoughi et al., 2021).

The presence of ice in the pore space also affects the attenuation (Q) of seismic waves. For example, in seismic laboratory measurements, Wu et al. (2017) found a peak of the inverted seismic attenuation at temperatures immediately below the freezing point for saline permafrost soils. Similar results were presented by Matsushima et al. (2008, 2016) who saw an increase in inverted seismic attenuation of the P- and S-waves between 0^oC and -3^oC. They also reported a correlation between the ultrasonic waves' attenuation and frozen brine's presence in the pore space. As attenuation plays a crucial role in seismic wave propagation, it must be considered when inferring velocity models in areas where permafrost exists.

Table 3.1 summarizes laboratory measurements of seismic velocities and attenuation factors (for P- and S-waves) measured on samples representative of the permafrost conditions found in the Beaufort Sea found in the literature (Bellefleur et al., 2007; Buckingham, 1997; Dou et al., 2016; Matsushima et al., 2016). As discussed above, in the presence of ice, coarse-grained materials (e.g., PIBSA and IBSA) have higher seismic velocity values than fine-grained materials (silts). The lowest Q values are found in partially-ice bonded material at intermediary ice content.

Table 3.1 Lithology and seismic parameters found in the Beufort Sea. We refer to the same Q for P- and S-waves.

	Vp (m/s)	**Vs (m/s)**	**Q**	ρ **(kg/m^3)**
Water	1430	0	1000	1000
Unfrozen Sediments (NFS)	1500-1900	300-600	30-90	1900
Partially Ice-Bonded Silts (PIBSI)	2000-2900	720-1044	10-30	1900
Ice-Bonded Silts (IBSI)	3000-3600	1600-2000	40-70	1900
Partially Ice-Bonded Sands (PIBSA)	3000-3500	1050-1250	10-15	1900
Ice-Bonded Sand (IBSA)	3500-3900	1350-1850	30-90	1900
Unfrozen Shale (NFSH)	2100-2400	700-1350	50-120	1900
Ice-Bonded Shale (IBSH)	2650-3350	1450-1850	70-130	2300

3.4.1 A simplified conceptual model

We adopt a simplified velocity model of subsurface observed in the Beaufort Sea for the sensitivity analysis. In general, ice content within permafrost varies with depth as follows. Near the seafloor, permafrost usually has a low ice content and is partially ice-bonded due to the heat transfer between seawater and the frozen sediments. Ice content increases with depth and, in some cases, reaches ice-bonding conditions. The ice-bonded permafrost maintains its stability until the temperature increase (due to the geothermal gradient) reduces the ice content enough that grains are not bonded anymore (Dobinski, 2011). From that depth, ice content decreases until it reaches zero. In what follows, we use a simplified model containing 6 layers: a water layer on top, followed by the unfrozen seafloor (NFS), 3 layers of permafrost with varying ice content, followed finally by a layer of unfrozen sediments (Figure 3.3). The seismic properties of the permafrost layers are set to partially ice-bonded silts (PIBSI), ice-bonded sands (IBSA), and partially ice-bonded silts (PIBSI), in this order (Table 3.1). Combining those 3 permafrost layers allow us to emulate the ice content variation as a function of depth that is representative of the Canadian Beaufort Sea.

3.5 Sensitivity analysis

The sensitivity analysis is based on the simplified velocity model presented in Figure 3.3. We vary four parameters of the ice-bonded sands layer: depth (d), thickness (t), velocity (v), and Q factor. Note that we analyze a single value of Q for P- and S-waves, given the limited information in the literature, and to keep the degree of complexity low. Synthetic shots are modeled by a viscoelastic code implemented with the time-domain finite difference method (Fabien-Ouellet et al., 2017). First, we present a qualitative analysis to determine seismic features affected by the variation of permafrost conditions. The data is analyzed in two domains: the T-X domain (shot gathers), which contains all seismic arrivals, and the F-V velocity domain (dispersion plots), which is most sensitive to guided waves. In the second part, a quantitative analysis is presented to better estimate the sensitivity of these features.

3.5.1 Qualitative analysis

Figure 3.4 summarizes the effects of changing the depth of the permafrost (Figure 3.4a-d) from 25 m to 100 m, the thickness of the permafrost (Figure 3.4e-h) from 50 m to 150 m, the velocity of the ice-bonded layer (Figure 3.4i-l) from 3 km/s to 4.5 km/s, and the Q factor of the ice-bonded permafrost from 7-1000 (Figure 3.4m-p). On these figures, we changed one parameter at a time and kept the other parameters constant to a value of 57 m for the

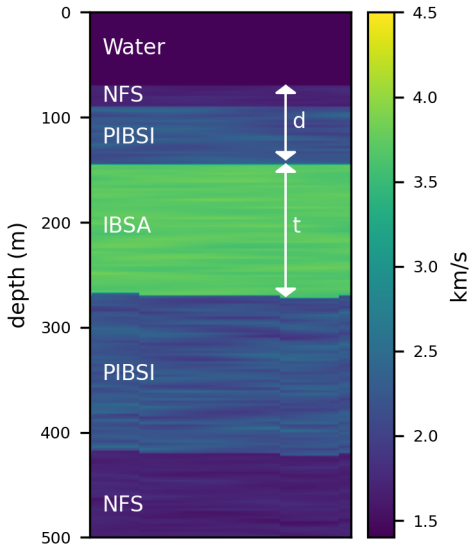


Figure 3.3 The velocity model used for the sensitivity analysis. NFS refers to unfrozen sediments, PIBSI to partially ice-bonded silts, and IBSA to ice-bonded sand.

depth, 150 m for the thickness, 3.7 km/s for the ice-bonded permafrost velocity, which are representative of permafrost conditions at SP 8275, and 60 for its Q factor. For all the cases, we observe that the interference between reflection multiples (RE-M), refraction arrivals, and its multiples with primary reflections hinder the identification of deeper reflections, including the base of the permafrost.

Similarly, Figure 3.5 summarizes the effects on the dispersion plots of changing the depth of the permafrost (Figure 3.5a-d) from 10 m to 150 m, the thickness of the permafrost (Figure 3.5e-h) from 10 m to 150 m, the velocity of the ice-bonded layer (Figure 3.5i-l) from 3 km/s to 4.4 km/s and the Q factor of the ice-bonded permafrost (Figure 3.5m-p) from 10 to 80. For all the evaluated cases, we can observe the characteristic modal distribution associated with guided waves.

Variation of depth

As expected, changes in the permafrost depth affect the crossover distance between the direct and refracted arrivals (Figure 3.4a-d). Furthermore, the increasing depth decreases the number of refracted arrivals associated with the free surface multiples (RA-M). Note also that the amplitude decay with offsets of the refracted waves remains constant for all evaluated depths. Changes in the permafrost depth also control the amplitude decay, the frequency content, and the arrival times of the reflection multiples (RE-M).

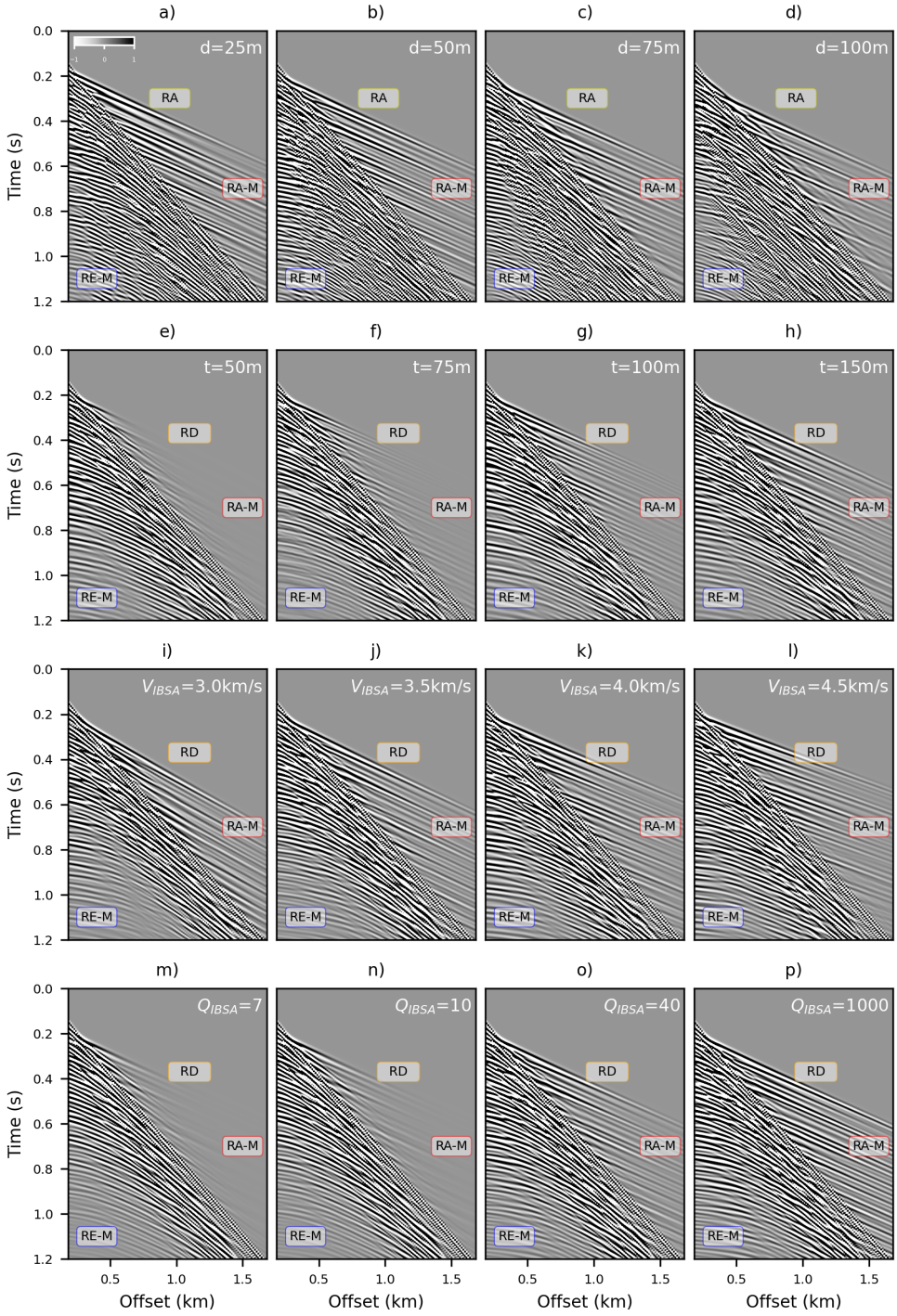


Figure 3.4 Effects on varying depth between 25 m and 100 m (a-d), thickness between 50 m and 150 m (e-h), velocity of the ice-bonded permafrost between 3.0 km/s and 4.5 km/s (i-l) and its Q attenuation factor from 7 to 1000 (m-p) on the modeled SP 8275. Guided P-waves (GW), reflection multiples (RE-M), refracted waves (RA), and its multiples (RA-M) are affected by variations in the depth, thickness, velocity of the ice-bonded permafrost and its Q attenuation factor

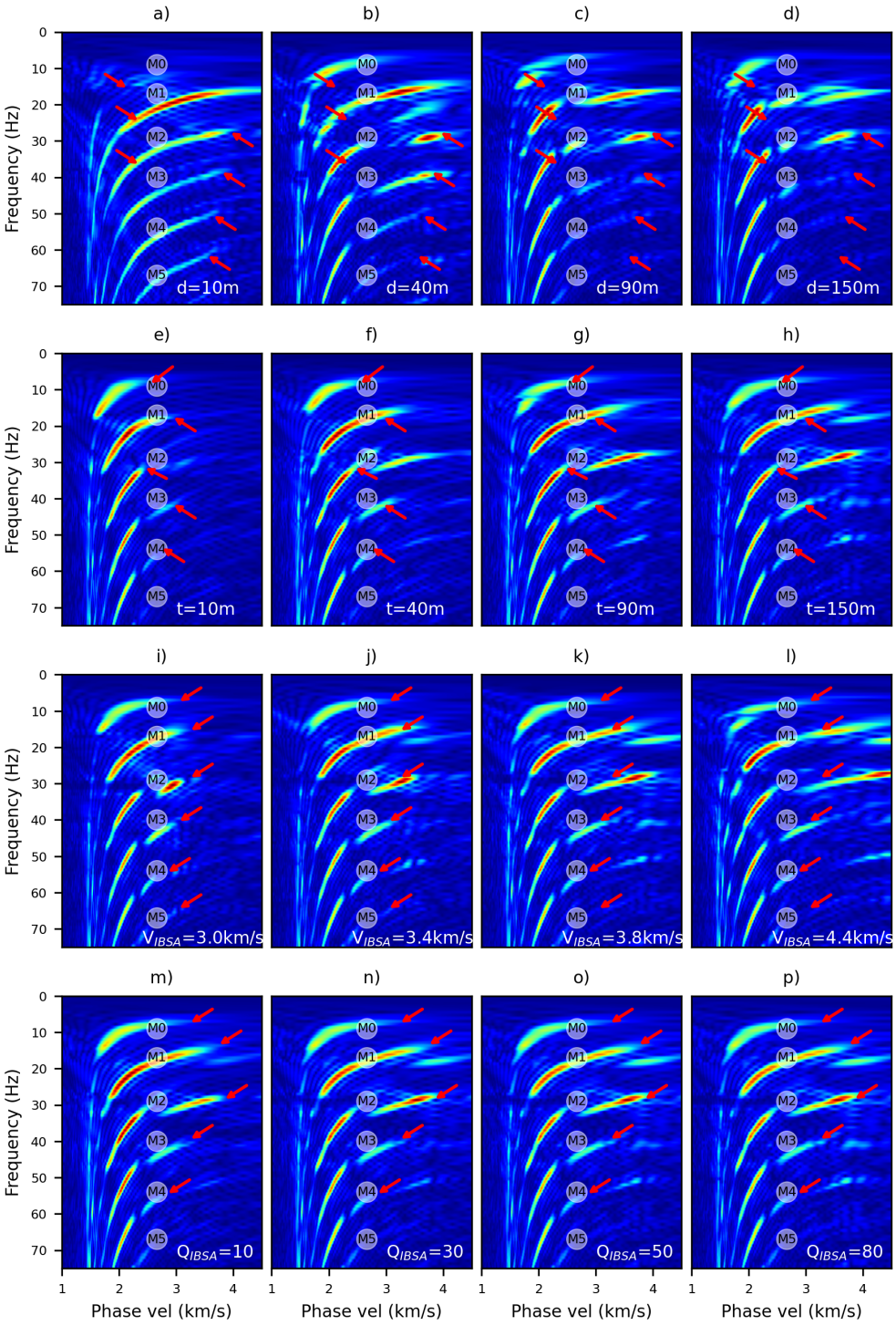


Figure 3.5 Dispersion curves for modeled shot 8275 when varying permafrost depth between 10 m and 150 m (a-d), thickness between 10 m and 150 m (e-h), ice-bonded permafrost velocity between 3 km/s and 4.4 km/s (i-l) and Q factor between 10 and 80 (m-p). Red lines mark the position of the inflection points and the cutoff frequencies at 10 m depth (a-d) at 10 m thickness (e-f), at 3.0 km/s velocity (i-l), and Q factor equal to 10 (m-p).

Red arrows in Figure 3.5a mark the inflection points for modes M0, M1, and M2 and the cutoff frequencies of modes M2, M3, M4, and M5 at 10 m depth. For comparison, we kept the red arrows at the same location in the remaining subplots (Figure 3.5b-d). As depth increases, the inflection points move towards lower frequencies. The changes are larger at smaller depths and gently diminish at medium to large depths. The cutoff frequency of higher modes decreases with increasing depth. Note also that modes become discontinuous and hard to identify from a depth of 40 m onward.

Variation of thickness

Thickness variations (Figure 3.4e-h) significantly impact the amplitude decay with offsets of the refracted waves (RD). At small thicknesses (50 m), the refracted waves are quickly attenuated. The amplitude at large offset gradually increases with increasing thickness. The permafrost thickness also affects both the arrival times and the amplitude of the reflection multiples. In all cases, it is impossible to discern the primary reflections under the train of multiple.

In Figure 3.5e there is a gradual expansion of the modes to higher phase velocities for increasing thickness. All modes are affected, but the expansion is larger for modes M1 and M2. Also, note that mode expansion is faster for smaller thicknesses (10 m - 40 m) and becomes more subdued for larger values (90 m - 150 m).

Variation of Q factor

Almost identical to varying thickness, varying Q impacts the amplitude decay with offset of the refracted waves (RD) and the reflection multiples (Figure 3.4m-p). The amplitude decay of the refracted waves decreases with increasing Q. Note, however, that the arrival times of the reflection and refraction multiples are not affected by Q variations, contrary to thickness.

In Figure 3.5m, guided P-wave modes are mostly unaffected by Q. A discontinuous fragment of what we identify as M1 does appear at larger Q, but all other modes remain close to constant.

Variation of velocity

As expected, variations in the ice-bonded permafrost velocity affect the slope of the refracted arrivals (RA), their multiples (RA-M), and the crossover distance (Figure 3.4i-l). On the other hand, the amplitude decay with offset (RD) is mostly unaffected by the IBSA velocity.

Velocity variations also affect the arrival times and amplitudes of the reflection multiples (RE-M).

In Figure 3.5i, we note a gradual expansion of the modes towards higher phase velocities for increasing velocities, without any changes on the lower velocity parts of the modes. Velocity affects mostly higher modes, and M0 remains mostly unaffected in this case.

Summary

Figure 3.6 summarizes the effects of changing depth, thickness, velocity, and Q factor. The black and red contours compare the smallest and largest parameter value presented in columns one and four of Figure 3.5, allowing us to visualize the total change better. With increasing depth, the inflection point of the lower modes (M0, M1, and M2) moves towards lower frequencies as higher modes (M3, M4, and M5) gradually shift to lower velocities. On the other hand, increasing thickness and permafrost velocity expands the recorded modes to higher phase velocities. Their effects on dispersion plots are very similar and cannot be easily distinguished.

However, as seen in the previous section, variations of velocity and thickness have distinct effects in the T-X domain. Similarly, changes in Q do not significantly affect the dispersion plots, but effects are notable in the T-X domain. Moreover, the interference between arrivals has a different behavior in each domain. For example, in the T-X domain, the interference between arrivals generates complicated patterns associated with the constructive/destructive effect that hinders their identification. Conversely, it could generate discontinuities and modes overlapping in the dispersion plots. To provide an in-depth understanding of the effects of varying permafrost's depth, thickness, velocity, and attenuation on the seismic response, we quantitatively combine the information of both domains in the next section.

3.5.2 Quantitative sensitivity analysis

To assess to what extent shot gathers and dispersion plots can be used to estimate depth, thickness, velocity, and Q factor, we perform a quantitative sensitivity analysis (Razavi and Gupta, 2015). We measure the normalized sensitivity according to

$$S_j = \frac{1}{C} \sum_i \left| \frac{\partial u_i}{\partial m_j} \right| \tag{3.1}$$

where u_i is the seismic data (either the shot gather of the dispersion image) and m_j are the perturbed parameters, either d, t, v or Q. The normalization constant C is chosen so that

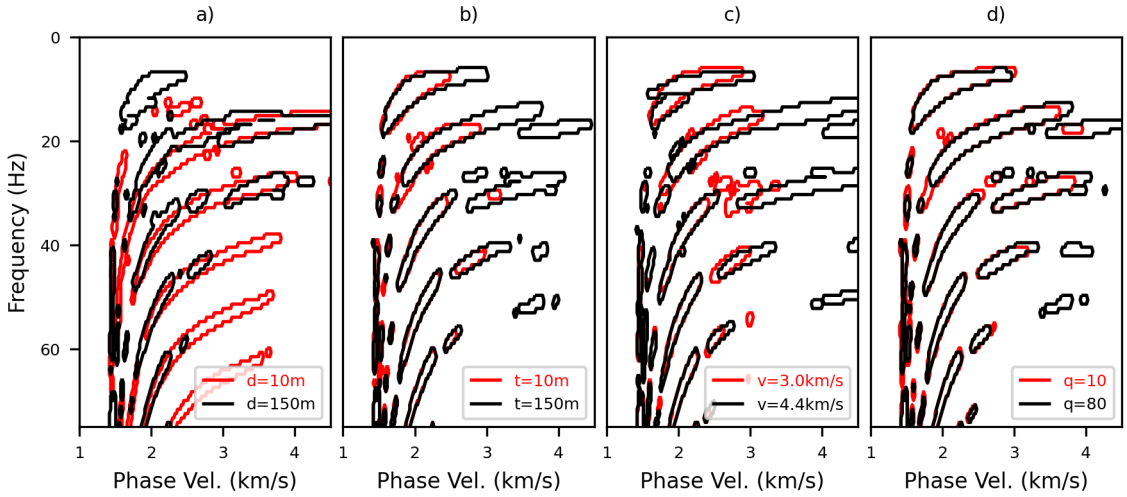


Figure 3.6 Effects of varying depth (a) and thickness (b) from 10 m to 150 m, and Q factor (c) from 10 to 80 for modeled shot 8275.

sensitivity has a maximum of 1 for each parameter. The sensitivity can be viewed as the sum of the rows of the Jacobian matrix that would be used for inversion. We use centered finite-difference to compute the partial derivatives.

Figure 3.7 presents the resulting normalized sensitivities of the shot gathers and dispersion plots for the four parameters. We observe that the shot gather sensitivity is not monotonic regarding depth, thickness, and velocity. In fact, several sensitivity peaks can be observed, which can be linked to constructive/destructive interference of multiples. An inversion methodology may find multiple local minimums and hinder an accurate retrieval of the evaluated parameters.

On the other hand, the sensitivity of the dispersion plots is monotonically decreasing for all tested parameters, aside from minor local variability. The sensitivity exponentially decreases for depth and thickness, owing to the evanescent nature of guided waves. The decrease is more or less linear with respect to velocity (Figure 3.7c). Usually, in an inversion procedure, one defines an acceptable error (Tarantola, 2005). Assuming that the error is negligible under 5% of the maximum gradient, we could say that the dispersion plot can retrieve permafrost depths up to 168 m and permafrost thickness up to 452 m for the evaluated model and Q factors up to 27. For permafrost velocities, they are retrieved for the whole evaluated interval (3.0 km/s - 4.5 km/s).

Interestingly, shot gathers can retrieve a higher range of Q factors (up to 89). However, the curve displays an additional local minimum at a Q value of 10. The peak of the normalized

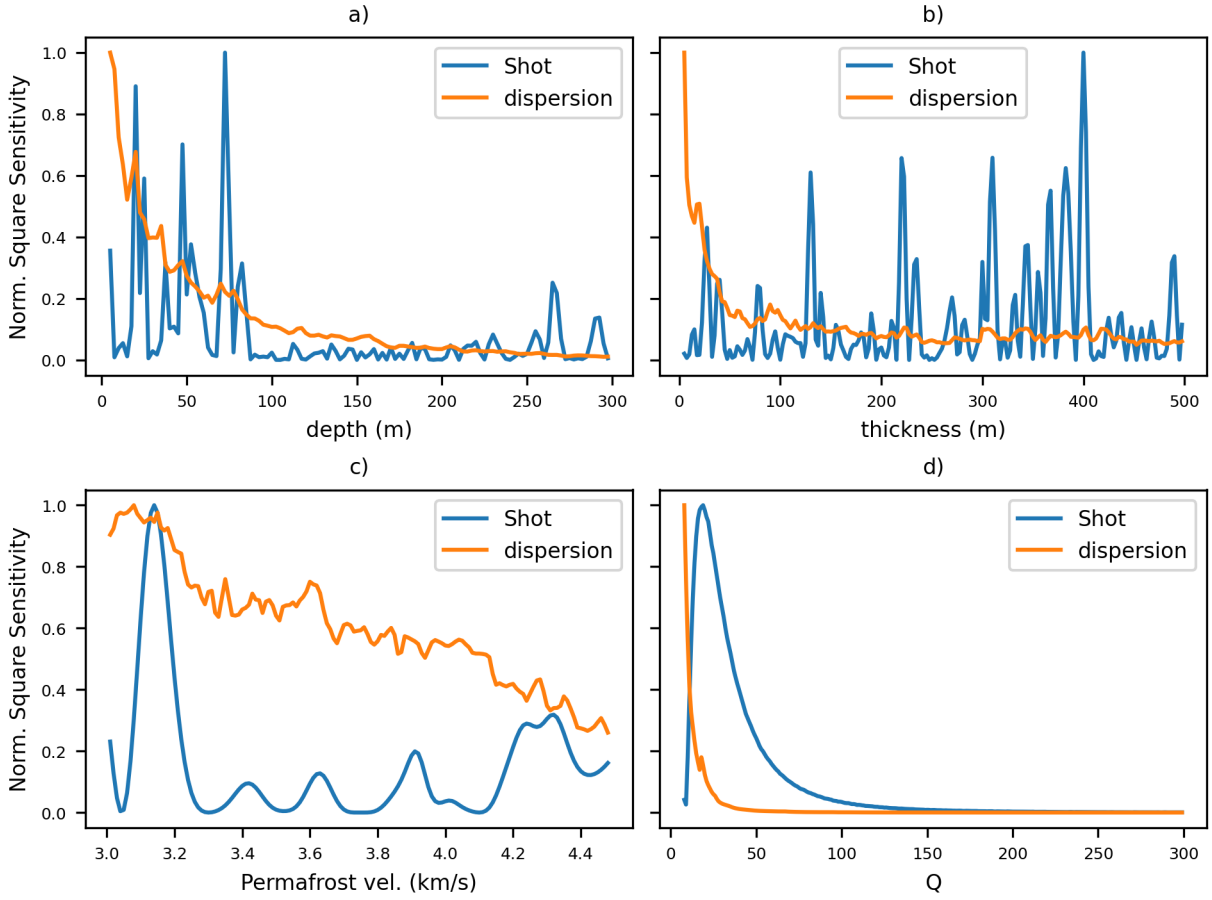


Figure 3.7 Normalized square sensitivities for depth (a), thickness (b), permafrost velocity (c) and Q factor (d) for the shot gather (blue) and the dispersion plot (orange)

sensitivity for the shot gather is located at a Q factor of 18.

The sensitivity analysis confirms that the seismic data, and more specifically guided P-wave, are sensitive to depth variation of permafrost. Furthermore, both the T-X and F-V domains are complementary, as both contain attributes that can be used to build a velocity model. Therefore, we propose a coupled approach to build a velocity model where both shot gathers and dispersion plots are used jointly.

3.6 Reproducing seismic data with simple models

The previous analysis is based on simplistic permafrost models. However, as we will show here, simplified model schemes are justified when seismic data are acquired over permafrost areas that can be reproduced to a large extent by tabular models.

In what follows, we build a tabular model of permafrost by a manual trial-and-error methodology. Each layer's thickness, velocity, and Q factor is varied until a good visual similarity is attained between the observed and modeled shots and dispersion plot. The fitting is guided by the features identified in the previous section (mode shape, arrival times, amplitude decay, and interference between the coherent arrivals).

We applied this procedure at various locations along the line (see blue lines in Figure 3.2a for SP locations). Figure 3.8 shows the recorded data, the synthetic shots, and the corresponding Vp velocity models for SPs 7900, 8125, 8201, and 8275. In general, the modeled shots reproduce the main features found in the observed data, such as guided waves, reflected arrivals, and multiples. The observed and modeled shots also show similar amplitude decay with offset and refraction velocities. However, a perfect fit was not sought in this experiment, and several differences can be seen between modeled and observed data. For instance, the arrival time of the reflection multiples are not reproduced exactly, as their complex interference pattern is not linearly correlated with layer thickness and velocity.

Figure 3.8a-d presents the recorded and synthetic shots and dispersion plots for different permafrost conditions: Figure 3.8a is located in a zone of low ice content, Figure 3.8b-c are located over discontinuous ice-bonded permafrost and Figure 3.8d is located over a zone of continuous ice-bonded permafrost (see Figure 3.2 for shot locations along the line). The dispersion plots show increasing energy for high phase velocities towards the SE of the line, i.e., where permafrost conditions are changing from low ice content sediments to discontinuous ice-bonded sediments. For instance, at SP 7900 (Figure 3.8a), the energy of the mode hardly goes above 2000 m/s, while for SPs 8125 and 8201 (Figures 3.8b-c), there is significant energy within 2000 m/s and 3000 m/s and for SP 8275, we observe a significant amount of

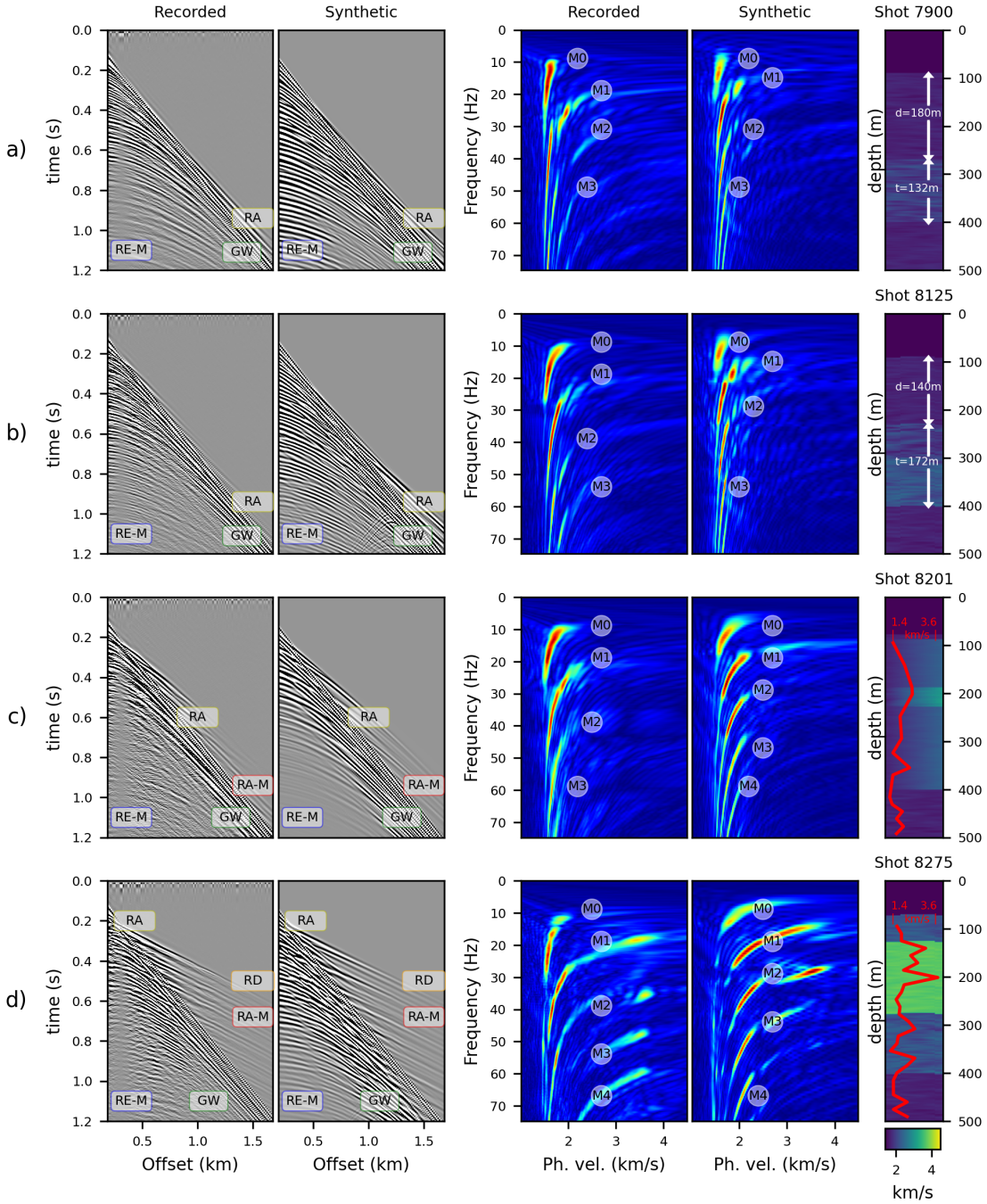


Figure 3.8 Recorded, synthetic shot and corresponding velocity model for SP 7900 (a) located in a low ice content zone, models SP 7950 (b) and SP 8125 (c) located in a discontinuous ice-bonded zone, and models SP 8201 (d), and SP 8275 (d) located over continuous ice-bonded permafrost. In general, the modeled shots and dispersion plots reproduce the main features found in the observed data.

high-velocity energy for the first modes (i.e., 3500-4500 m/s). As shot point locations move towards the SE end of the line, the spacing between modes increases, and phase velocities become larger. As presented in the previous section, those displacements can be associated with changes in permafrost depth.

The red line in Figures 3.8c-d represents the interval velocity recorded in boreholes I-44 and M-13, which are close to the position of SPs 8201 and 8275, respectively. We can observe that, in general, our inferred velocity model agrees with the measured velocity in the well. In particular, there is a good agreement for the velocities at the top of the ice-bonded permafrost in the transition from ice-bonded to partially ice-bonded permafrost ($\sim$ 220 m for SP 8201 and $\sim$ 270 m for SP 8275) and the transition from partially ice-bonded to unfrozen sediments ($\sim$ 400 m).

3.6.1 An integrated velocity model

The trial and error methodology cannot be viewed as a robust or general method to map permafrost on a large scale. It is obviously too subjective and time consuming. However, we can still get a general idea of the lateral variations of permafrost over the line ARA05C-06. To do so, we repeated the procedure described in the previous sections to estimate velocity models at SPs 7950, 8185, 8225, and 8250. A velocity model is obtained for the entire seismic line. To avoid extensive processing, we applied a linear interpolation to all analyzed shots assuming lateral continuity of the layers. Although this is a simplified approach, it is sufficient for evaluating the capability of seismic data to retrieve permafrost information with depth. Figure 3.9a presents the resulting V_P velocity section obtained from modeling. The white lines in the figure mark the SPs used to build the velocity model.

On the left of the model until SP 8185 ($\sim$16 km), we can observe partially ice-bonded permafrost (1700 m/s $< V_P <$ 2900 m/s) from a depth of around 270 m. Seismic velocities are slowly increasing to the SE in this zone. Between SPs 7900 ($\sim$1.7 km) and 8185 ($\sim$16 km), there is a sudden rise in the permafrost table, and the velocity increases from 2000 m/s to 3200 m/s. The velocity increase is interpreted as higher ice content. This sudden change between SPs 7900 and 8185 is associated with the undulated seismic event observed between 13 km and 17 km in Figure 2a.

Ice-bonded permafrost ($V_P >$ 3500 m/s) is observed from SP 8185 ($\sim$16 km) to the SE end of the line. This ice-bonded permafrost is delimited at its top and base by layers of low ice content sediments. The velocity model suggests that at SP 8185, ice-bonded permafrost is approximately 200 m. At that position, the permafrost table (top of the low ice content layer) is approximately 150 m deep, and its depth rapidly decreases until it approaches the

seafloor at SP 8275 (∼20 km). The thickness of the ice-bonded permafrost increases while the depth of the top of the permafrost decreases to the SE. From SP 8185 (∼16 km) to SP 8250 (∼19 km), there is a gradual increase in the observed V_P velocities, supporting a higher ice content. The top of the ice-bonded permafrost remains at a constant depth from SP 8250 until the end of the line while its thickness increases towards the SE.

We compare our results with a velocity model obtained by applying conventional travel time first break tomography (TTFBT) (Li et al., 2013; Taillandier et al., 2009). Figure 3.9b presents the velocity model obtained using the open-source software pyGIMLi (Rücker, C. and Günther, T. and Wagner, F. M., 2017). We collected first-break picks every 10 shots and at intermediate locations where the trend of the adjacent picks did not match. Because the maximum offset is positioned 1572.5 m behind the source, it restricts the water layer and the shallow subsurface layers from being accurately resolved.

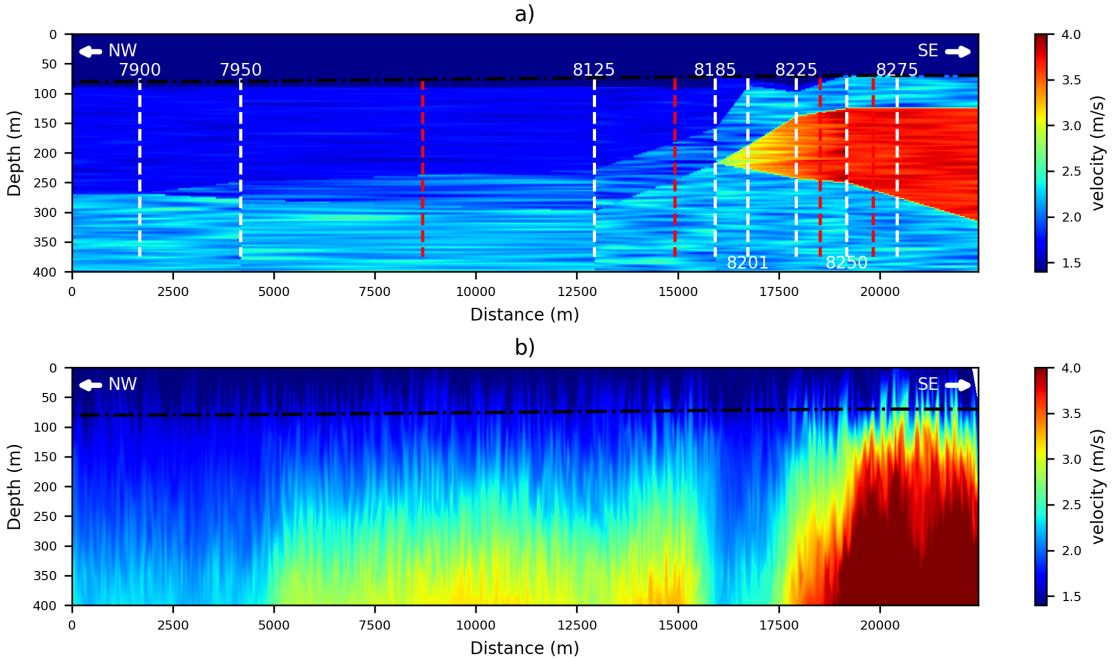


Figure 3.9 a) Resulting V_P model after coupling velocity analysis on shots gathers and dispersion plots on SP 7900, SP 7950, SP 8125, SP 8185, SP 8201, SP 8225, SP 8250 and SP 8275, and linear interpolation in intermediate shot position. b) Resulting velocity model after refraction travel time tomography. The dashed-dotted black line indicates water depths from bathymetry.

In Figure 3.9, we compare the results of the TTFBT and the trial and error approach. We notice that the two models agree on the permafrost table's velocity increase trend. The permafrost table varies in both models from 270 m at the NW end to 150 m depths at 16 km

(SP 8185). From 17 km (SP 8201) to the SE end, the depth of the permafrost table continues reducing until reaching the ocean bottom in both models. The most significant discrepancies between the models occur for the velocities below the permafrost table between 5 km and 15 km. However, the TTFBT depends on the starting model at larger depths, and most of the velocity increases between 5 km and 15 km are explained by the linear trends imposed on the starting model.

In both models, high velocities are observed from 16 km (SP 8185) to the SE end of the seismic line. Those high velocities support the occurrence of ice-bonded sediments. In this zone, the method that visually matches synthetic data and observed data not only detects the top of the ice-bonded permafrost but can also delimit the lower transition between the ice-bonded and partially ice-bonded permafrost. This is not possible with TTFBT. Still, the transition between partially ice-bonded permafrost (∼2500 m/s) and unfrozen sediments (∼1700 m/s) is difficult to estimate by visually matching synthetic data with observed data. The sensitivity analysis indicates that it should be possible, but a more rigorous processing approach (e.g., full-waveform inversion or deep learning) is required to do so.

To further assess the reliability of the model obtained by visual agreement, synthetic shots are evaluated at intermediate SP positions (red lines in Figure 3.9a). Figure 3.10 presents the recorded and synthetic shot gathers built from the velocity model in Figure 3.9a at SP 8040, 8190, 8237, and 8263. Once again, synthetic shots are in good agreement with the recorded data. All arrivals included in the real data are present in the synthetic shot gathers. The crossover distance between the direct arrival and the refracted waves is accurately reproduced for all SPs. The apparent velocity on the refracted waves also seems to follow the velocity observed in the recorded shots.

3.7 Conclusion

In this paper, we assessed the potential of seismic data to estimate the lateral and vertical distribution of subsea permafrost. Free-surface multiples and guided waves hinder the identification of the permafrost base in the data of line ARA05C-6 on the outer continental shelf of the Beaufort Sea. In addition, the presence of different seismic arrivals and the constructive/destructive interference between them makes it challenging to identify reflected waves.

We presented a sensitivity analysis showing that variations of the permafrost depth, thickness, velocity, and attenuation impact the seismic arrival times, amplitude, and frequency content. The sensitivity of the data in the T-X domain is non-linear and subject to con-

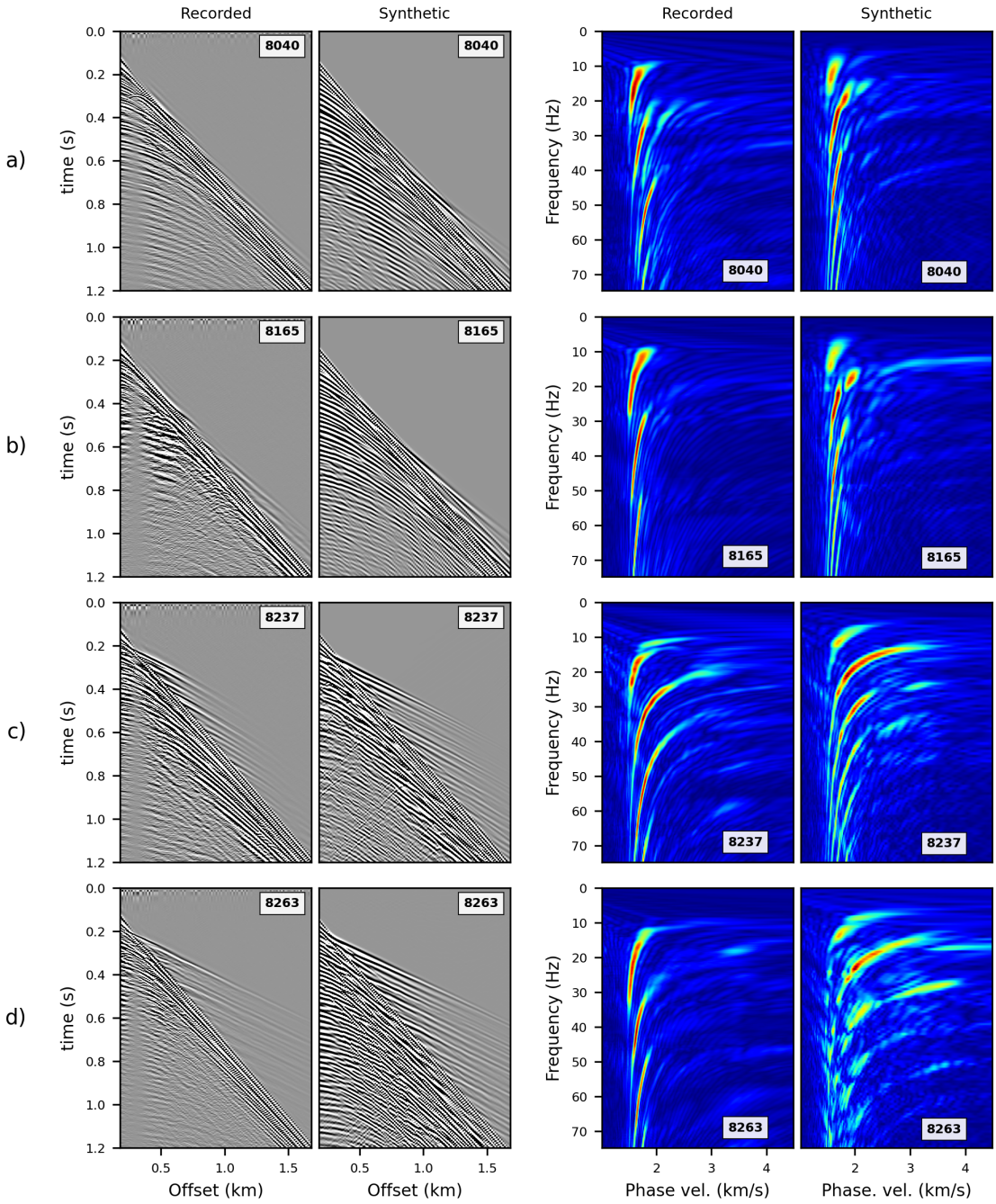


Figure 3.10 Recorded vs Synthetic shot gathers at intermediate SP extracted from the velocity model presented on Figure 3.9a.

structive/destructive interference of multiples. Using the full-waveform for inversion would thus be especially challenging. We used dispersion plots to analyze highly energetic guided P-waves. We showed that guided P-waves display a predictable behavior linked to the depth of the permafrost table and its thickness. Given its evanescent nature, guided P-waves lose their sensitivity as permafrost deepens and becomes thicker. We estimated a penetration depth of nearly 600 m, representing approximately the maximum expected depth of subsea permafrost in the Beaufort Sea.

We suggested a combined approach to retrieve 1D layered velocity models based on both shot-domain and dispersion-domain features. We derived 1D velocity models by selecting the model parameters that provide the best visual fit between synthetic and recorded shots in the T-X and F-V domains. We showed that most features found in seismic data over subsea permafrost can be explained by these simple layered models.

We built a 2D velocity model for the whole seismic line ARA05C-6 by the linear interpolation of selected shots along the line. The resulting model shows vertical and lateral seismic velocity variations linked to varying ice content of subsea permafrost. A comparison with results from TTFBT demonstrated that a combined analysis of shot gathers and dispersion plots could provide a more accurate vertical velocity distribution. In particular, a joint approach can detect the transition between ice-bonded and partially ice-bonded permafrost. In contrast, the TTFBT cannot detect any interface below ice-bonded sediments, given the limitations of the first break picking.

This paper aimed to show that seismic data contains information that can constrain the extent of subsea permafrost at depth. By visually matching features found in synthetic and real data, we demonstrate that information about subsea permafrost (e.g., depth, thickness, velocity, and attenuation) can be retrieved from seismic data. We believe this visual agreement approach provides a good first-order approximation to identify the presence of ice-bearing and ice-bonded permafrost. Our findings, however, support the possibility of designing a more robust inversion approach. Indeed, future work will focus on the application of deep-learning processing to map lateral and vertical variations of subsea permafrost so our approach is easier to use on larger datasets. Our results also support the implementation of an FWI methodology that would minimize discrepancies jointly in the T-X and F-V domains. Finally, additional strategies could use the visual agreement approach to generate initial models for an FWI workflow.

3.8 Acknowledgments

This work benefited from the support of the Environmental Geoscience Program of the Geological Survey of Canada (GSC). We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), RGPIN-2018-06163.

3.9 Addendum A: Sensitivity analysis in Radon domains

The sensitivity analysis detailed in Chapter 3 only focuses on the time-offset ($t - x$) and frequency-phase velocity ($f - c$) domains. In the subsequent chapter, we use two additional domains: the hyperbolic radon transform in the time domain and frequency domain.

The $\tau - p$ or zero offset time-velocity domain can be obtained by applying the hyperbolic Radon transform on the $t-x$ domain. The hyperbolic Radon transform provides information about reflection and multiples by integrating the data along a hyperbolic travel-time curve. This transform can be expressed using Equation 3.2 (Ibrahim et al., 2018):

$$\tilde{u}(\tau, p) = \sum_{k=1}^{N} d(t = \sqrt{\tau^2 + x_k^2/p^2}, x_k), \tag{3.2}$$

where $\tilde{u}$ is the hyperbolic Radon model, d is the recorded data, x is the offset, τ is the zero-offset time, N is the number of offsets, and p is the stacking velocity. Besides, we can also perform a sensitivity analysis in the $f - p$ domain, which is obtained by applying a Fourier transform to the time dimension on the hyperbolic Radon model $\tilde{u}(\tau, p)$.

Figure 3.11 illustrates the influence of varying parameters, such as the depth, thickness, velocity, and Q factor of the ice-bonded permafrost layer, for the $\tau - p$ domain. When transforming the different shots into the $\tau - p$ domain, arrivals form a characteristic conical shape. Modifications to the depth, thickness, velocity and Q of the ice-bonded layer influence the shape of this cone as follows:

- Increasing the depth of the ice-bonded permafrost layer from 10 m to 150 m (Figures 3.11a-d) leads to a gradual time delay in the appearance of the first and subsequent arrivals within the cone. Furthermore, the continuity of the arrivals is also impacted by these depth variations. However, the cutoff velocity (e.g., the velocity for which the arrival amplitude is attenuated) demonstrates a consistent trend regardless of the depth changes.
- Varying the thickness of the permafrost layer from 10 m to 150 m (Figure 3.11e-h) impacts the cutoff velocity, revealing a direct correlation between cutoff frequency and

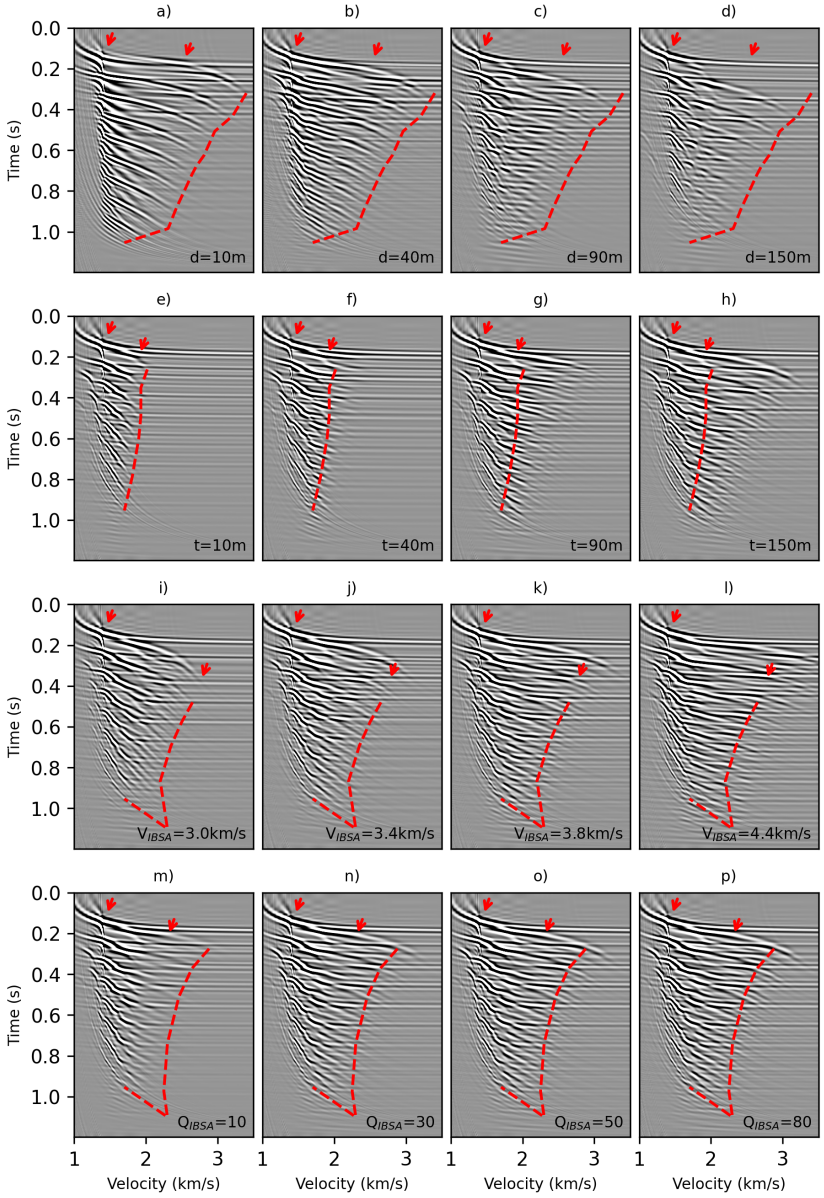


Figure 3.11 Shot gathers in the $\tau - p$ domain for modeled shot 8275 (see Chapter 3) when varying permafrost depth between 10 m and 150 m (a-d), thickness between 10 m and 150 m (e-h), ice-bonded permafrost velocity between 3 km/s and 4.4 km/s (i-l) and Q factor between 10 and 80 (m-p). Red arrows mark the upper boundary of the conical shape of the arrivals, and red dashed lines show the cutoff velocity at 10 m depth (a-d), 10 thickness (e-h), at 3.0 km/s velocity (i-l) and Q factor equal to 10 (m-p).

thickness. Nevertheless, at lower velocities, the arrivals are unaffected by changes in thickness. This behaviour is expected as the shot gathers at low velocities are dominated by multiples in the water column, unaffected by the thickness of the permafrost layer.

- Varying the velocity of the ice-bonded layer from 3 km/s to 4.4 km/s (Figure 3.11i-l) impacts both the arrival times and the cutoff velocity of the cone. As anticipated, increasing the velocity of the ice-bonded layer leads to the earlier appearance of the first and subsequent arrivals within the cone. Moreover, a clear positive correlation exists between the ice-bonded layer's velocity and the cutoff velocity. Also, the continuity of the arrivals remains unaffected, although changes in the slope are observed.
- Finally, varying the Q factor from 10 to 80 (Figure 3.11m-p) directly affects the cutoff velocity, showing an increase with increasing Q factor. However, similar to varying the thickness of the permafrost, varying the Q factor does not affect the arrival times or their continuity.

Figure 3.12 presents a similar analysis in the $f-p$ or frequency-velocity domain. The results reveal a distinct modal distribution of arrivals following the domain transformation. Notably, alterations in the assessed parameters have an impact on both the shape of the modes and their corresponding cutoff velocities as follows:

- Varying the depth of the permafrost layer impacts the cutoff frequency of the modes, resulting in an observable rise in its values as the depth increases. Additionally, the depth variations influence the shape of the modes, leading to an augmentation in their thickness and a gradual mode bifurcation.
- For changes in the thickness, velocity, and Q factor of the permafrost layer, the cutoff velocity increases as the value of the evaluated parameter increases and the effect of bifurcations has more influence on the mode shape.
- Changes in the velocity of the permafrost layer exhibit more pronounced variations, as expected, significantly influencing the cutoff velocity.
- When the Q factor is varied from 10 to 30, there is a rapid alteration in the modes, but further increasing the Q factor to 80 yields minimal variation.
- Gradually increasing the thickness affects the cutoff velocity and mode shape through bifurcation. However, the variation rate is not as pronounced as that observed for changes in velocity.

Note that even though it may be more difficult to observe the effects of different parameters in the frequency-velocity domain, it is still possible to track changes in the evaluated parameters in this domain, particularly for permafrost depth and thickness. This is crucial to gain a

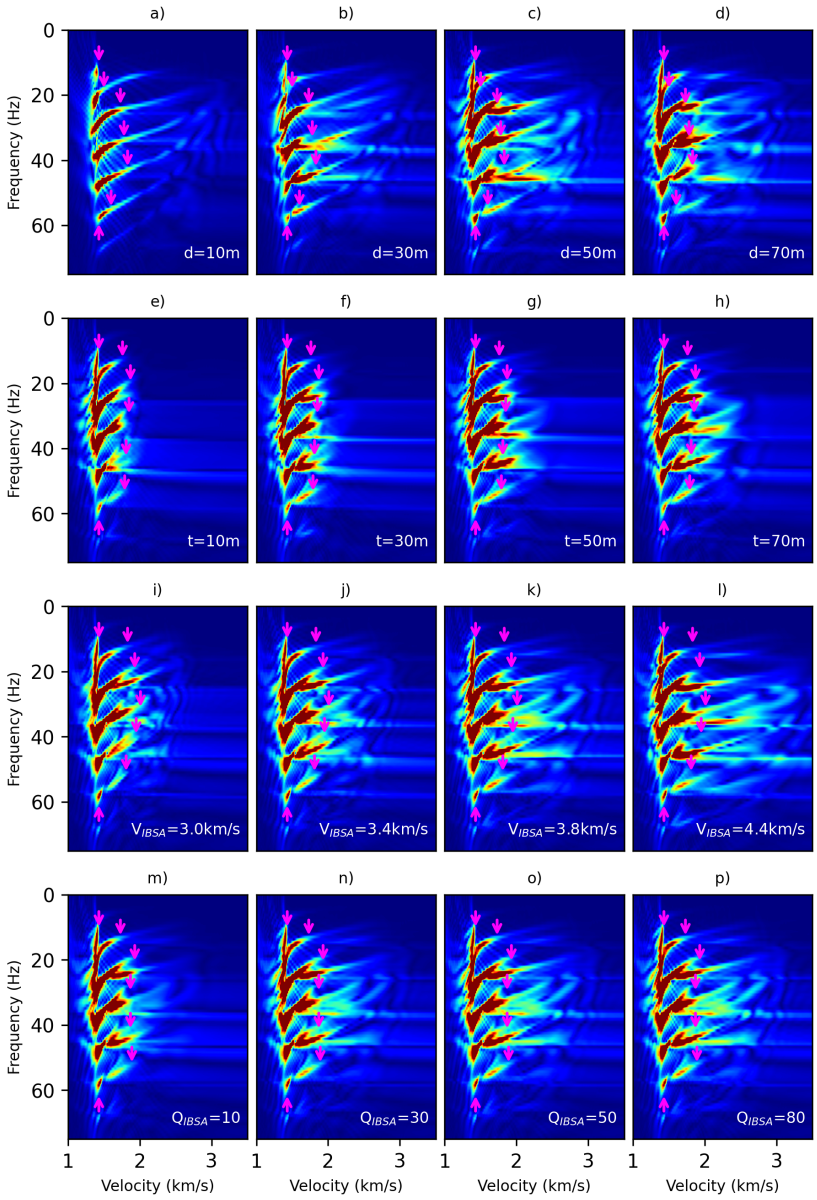


Figure 3.12 Effects on varying depth between 10 m and 70 m (a-d), thickness between 10 m and 70 m (e-h), ice-bonded permafrost velocity between 3 km/s and 4.4 km/s (i-l) and Q factor between 10 and 80 (m-p) on the $f - p$ domain. Magenta arrows mark the cutoff velocity of the modes observed at 10 m depth (a-d), 10 m thickness (e-h), 3.0 km/s velocity (i-l), and Q factor equal to 10 (m-p)

thorough understanding of subsea permafrost distribution since the depth and thickness of the permafrost are key parameters.

Following up, the normalized sensitivity of the time-velocity and frequency-velocity domains with respect to the depth thickness and attenuation are calculated following Equation 3.1.

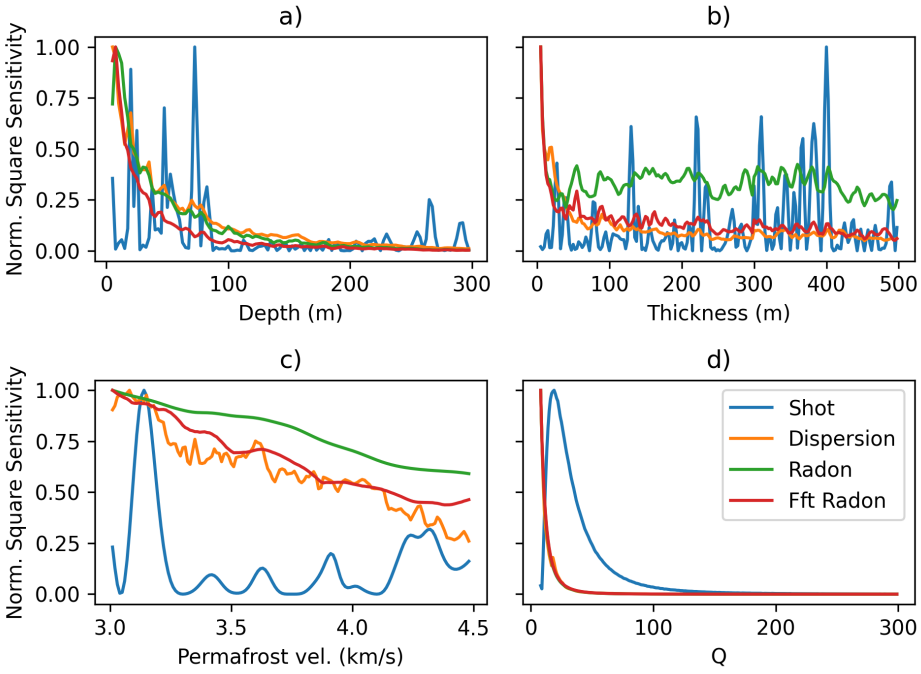


Figure 3.13 Normalized square sensitivities for depth (a), thickness (b), permafrost velocity (c) and Q factor (d) for the shot gather (blue), the dispersion plot (orange), Radon (green) and Frequency Radon (red) images.

Figure 3.13 shows the normalized sensitivity of the shot gathers in the four domains that are input to the NN. Similar to the sensitivity observed in $t-x$ and in the $f-c$ domain, both the $\tau-p$ (Radon) and $f-p$ (frequency Radon) domains demonstrate a consistent monotonic decrease across all tested parameters. Nevertheless, slight variations are discernible in the sensitivity curve of the $\tau-p$ domain with respect to the thickness, forming a stable curve for the evaluated depths. This distinctive behaviour highlights the Radon domain's superiority in providing crucial insights into permafrost thickness compared to other evaluated domains. Furthermore, the sensitivity of the Radon domain exhibits higher values within the range of ice-bonded permafrost velocities, enhancing its capacity to discern relevant information.

Interestingly, the sensitivity trends of the time-velocity and frequency-velocity domains with respect to the Q factor mirror those observed in the frequency-phase velocity (dispersion) domain. Lastly, the time-offset (shot gather) domain still exhibits a superior range of reliability when it comes to the inversion of the Q factor, with values approaching zero up to 100.

In conclusion, by extending the sensitivity analysis to multiple domains, we verify that different domains can provide more precise information about specific parameters. For example,

while the sensitivity curves in the $t-x$ domain reveal multiple local minimums when inverting the depth, thickness, and velocity of the permafrost layer, the $f-c$, $\tau-p$, and $f-p$ domains show a decreasing sensitivity towards the same parameters. Additionally, the sensitivity of the $\tau-p$ domains towards the thickness shows a stable curve compared to the other domains. These results support the notion that analyzing multiple domains simultaneously is necessary for accurately mapping and characterizing subsea permafrost. In fact, we investigate deeper into this approach in the subsequent chapters of this thesis.

CHAPTER 4 ARTICLE 2: DEEP-LEARNING VISCOELASTIC SEISMIC INVERSION FOR MAPPING SUBSEA PERMAFROST

Bustamante, J., Fabien-Ouellet, G., Duchesne, M.J. and Ibrahim, A., 2023. Deep-learning viscoelastic seismic inversion for mapping subsea permafrost. Geophysics, *Submitted*

Initial submission on December 22, 2022. Last submission November 27, 2023.

4.1 Abstract

Marine seismic surveys can be used to map ice-bearing subsea permafrost on a large scale. However, present seismic processing technologies have limited capacity to image permafrost distribution at depth, mainly due to the low sensitivity of primary reflections and refractions to the velocity inversion found at the base of permafrost. Guided waves and multiples are more sensitive to the velocity variations below the top of permafrost, but they remain challenging to use in physics-based inversion approaches. A deep-learning-based seismic inversion might improve seismic imaging below the top of permafrost by automatically extracting information from all wave modes. We present a multi-input neural network to estimate seismic velocities from marine seismic data. The network is trained on synthetic data generated from representative distributions of the seismic properties of subsea permafrost. We show that our network can image large velocity contrasts and reversals in depth, typical of subsea permafrost. We use our network to estimate P- and S-wave velocity and attenuation models from a seismic line in the Beaufort Sea. The neural network indicates highly discontinuous subsea permafrost with variable thickness in the area. The predictions correlate well with travel time tomography results and velocity logs recorded in wells along the line. Our work shows that deep-learning-based seismic inversion could become a cost-effective technology to map on a large scale the lateral and vertical distribution of subsea permafrost and, more generally, high-velocity rocks in shallow water.

4.2 Introduction

Subsea permafrost is any frozen ground under the sea level that remains frozen for more than two years. Subsea permafrost acts as a methane reservoir that stabilizes gas hydrates under and inside the permafrost. It also acts as a seal for additional underground methane reservoirs (Dallimore and Collett, 1995; Dallimore et al., 2015). Due to climate change, the overlying water body is getting warmer (Nielsen-Englyst et al., 2023), and permafrost is

degrading faster. This permafrost degradation releases additional greenhouse gases into the atmosphere and, in turn, accelerates climate change (Anthony et al., 2012; Schuur et al., 2008). Hence, mapping and monitoring the distribution of subsea permafrost is crucial for forecasting global warming.

Seismic surveys have proven to be a reliable tool for determining the lateral extent of subsea permafrost (Hunter et al., 1976; Kang et al., 2021). Ice in the sediments' pore space increases the velocity of seismic waves (Dou, 2015). Therefore, seismic methods can detect the presence of ice-bearing or ice-bonded subsea permafrost by measuring contrasts in seismic velocities (Dou, 2015; King, 1984). For example, Brothers et al. (2012a) mapped the lateral distribution of subsea permafrost in the US Beaufort Sea using seismic refraction. In the continental shelf of Alaska, Brothers et al. (2016) derived root mean squared (RMS) velocities from semblance analysis of seismic reflections to map subsea permafrost. In the Canadian Beaufort Sea, Kang et al. (2021) inferred seismic velocities of permafrost by employing Full Waveform Inversion (FWI) of seismic waves in the Laplace domain. Additional examples of seismic analysis of permafrost regions are found in the literature, including Rekant et al. (2005) in the Yugorsky Peninsula (Russia), Johansen et al. (2003) in the Svalbard Archipelago (Norway), Hilbich (2010) in the Swiss Alps, and Ramachandran et al. (2011) in the Mackenzie Delta (Canada).

Seismic methods can detect ice-bearing permafrost and map its lateral distribution. However, resolving the vertical extent of subsea permafrost is challenging. Seismic methods often fail to detect the base of the permafrost, characterized by the transition from ice-bearing to unfrozen sediments. On one side, no critical refractions are produced from the velocity inversion from ice-bearing (high velocity) to unfrozen sediments (lower velocities), preventing the use of the refraction method. On the other side, seismic reflection can, in theory, detect the base of the permafrost. However, the high seismic impedance contrast between air, the water column, and the frozen sediments acts as a leaky waveguide that traps seismic waves in the water column (Bustamante et al., 2021). As a result, early reflections with weaker amplitudes are hidden by the interference with the high amplitude guided waves and multiples traveling in the waveguide (Duchesne et al., 2022). The severe interference between reflections, refractions, guided waves, and multiples hinders their separation prior to the use of classical inversion (e.g., pre-stack and post-stack inversion) for the estimation of permafrost properties. Moreover, seismic data over permafrost zones are affected by high attenuation given the presence of ice in the pore space (Dou et al., 2016; Matsushima et al., 2016).

Still, guided waves and multiples contain valuable information about the velocity in deeper layers Bustamante et al. (2023a). Additionally, seismic surveys conducted in permafrost

areas can provide insights into shear wave velocity due to the large angle of illumination of the permafrost layer. However, there is no straightforward physic-based method to interpret these different arrivals simultaneously, except for viscoelastic FWI (Fabien-Ouellet et al., 2017). Nonetheless, FWI poses significant challenges in subsea permafrost regions due to the high-order multiples contained in the data. Instead, data-driven methods like deep learning could be a simpler alternative for performing a viscoelastic inversion.

In recent years, deep learning methods have become an efficient tool for inverting seismic velocities (Marques et al., 2022). An advantage of deep learning approaches is that the pre-processing step required to separate the different seismic arrivals is not needed. Instead, during the training phase, deep learning methods learn how to emphasize the features that best correlate for solving a given task (Waldeland et al., 2018). Since the pioneering work of Araya-Polo et al. (2018), who used neural networks (NN) to produce velocity models from semblance cubes, research on velocity model building (VMB) using NN has exploded (Adler et al., 2021). For example, generative adversarial networks (GAN) have been used for retrieving acoustic velocity and impedance models from seismic data (Meng et al., 2021; Wu et al., 2021; Zhang and Lin, 2020). Alfarraj and AlRegib (2019) used recurrent NN for elastic impedance inversion from multi-angle seismic data. Mustafa et al. (2021) proposed a bidirectional transfer learning strategy for retrieving impedance. Similarly, Simon et al. (2021) utilized transfer learning to build acoustic 2D velocity models in time and depth. Fabien-Ouellet and Sarkar (2020) use a recurrent neural network to estimate seismic velocity and show that a network trained on synthetic data can generalize to real seismic data. Zhang and Alkhalifah (2019) combines full waveform inversion (FWI) with fully connected NN to invert elastic velocity models.

Conventionally, seismic processing techniques take advantage of various domains to identify and isolate the desired signal from the noise. For example, semblance analysis utilizes the coherence of reflections in the time-RMS velocity domain to estimate stacking velocities. The frequency-domain slant-stack or linear radon transform transforms the seismic data to the frequency-phase velocity domain to create a dispersion image which can be used to isolate guided wave modes and surface waves (Park et al., 1998). The frequency-wavenumber domain is commonly used to separate surface waves. In addition, the hyperbolic Radon transform uses the zero-offset time-velocity domain to separate seismic reflections and multiples. Similarly, deep learning networks can be trained on different domains to identify specific features or to avoid restrictions associated with the recorded time-offset domain. For example, Abbas et al. (2022) demonstrates that a NN trained on the frequency velocity domain reduces the impact of the array geometry for inverting the near-surface shear-wave velocities model. Furthermore, Kaur et al. (2020) proposed that a deep learning method can improve the

convergence of the Radon transform, reducing lateral smearing and providing faster solutions. Similarly, Feng et al. (2022) proposed a high-resolution Radon transform inversion scheme based on NN, aiming to avoid aliasing problems.

Deep learning can take advantage of the various signatures associated with different domains by using multiple inputs. For example, Lowney et al. (2021) used the time-space, frequency-space, frequency-wavenumber, dip, and greyscale domains to classify diffractions, reflections, and noise enhancing and physical constraining the classification. Yang et al. (2021) presented a similar approach for reservoir classification, performing a clustering analysis to find composite seismic attributes, which are the input of a NN for oil and gas reservoir characterization. Although using multiple domains as input for a NN can provide additional information that could improve the inversion, to our knowledge, there is no published attempt to perform seismic inversion using a multi-input NN.

In this paper, we evaluate the capability of deep learning to estimate velocity changes for surveys collected over subsea permafrost. First, we describe how the synthetic training and validation datasets are created. Then, we introduce the NN architecture and the different strategies to improve the prediction accuracy and stability. We evaluate the performance of our NN to predict synthetic 1D and 2D models. Finally, we invert seismic data from the Canadian Beaufort Sea and compare the NN results with travel time tomography.

4.3 Methodology

An NN comprises multiple neurons that receive one or multiple inputs, apply a weighted summation, passes the results through a non-linear activation function and produces the output (Bishop, 1994). To achieve general approximation properties, a neural network must contain more than 2 layers, where each neuron of one layer passes its output to the inputs of the next layer. In the context of this work, the input of the NN corresponds to the seismic data, and the output is an estimated velocity (or attenuation) profile in depth.

We adopt a supervised learning approach (Bartlett and Maass, 2003). In supervised learning, the parameters of the NN are optimized by using a large quantity of paired inputs and outputs (labels). The training consists of minimizing the prediction error of the NN over the entire training set. Moreover, we adopt the transfer learning (TL) methodology proposed by Simon et al. (2023), in which a series of sub-datasets of increasing complexity is used to train the network. The principle of TL involves applying previously acquired knowledge from a particular task to enhance learning in a related task (Torrey and Shavlik, 2010). Here, we utilize the prior knowledge gained during training with 1D layered models to train a

comparable NN that takes in neighboring CMP gathers centered around a specific location to predict 2D velocity models. The 1D layered dataset can be created in large quantity at low cost. The 2D dataset, which is the target dataset, is more expansive to compute. Pretraining with the 1D models reduces the number of 2D models required to achieve good performance. This approach has been shown to reduce the cost of training and leads to more robust estimations (Siahkoohi et al., 2019).

The performance of a NN greatly depends on the suitability of a training set, the architecture of the network and the training methodology. In the following, we will discuss each of those topics in the context of velocity inversion.

4.3.1 Dataset

Training a deep neural network requires a large dataset that statistically represents the distribution of the target data. For supervised learning, accurate labeling of each sample must be available. In the context of VMB, each sample in a training set consists of a pair of seismic data (input) and P- and S-wave velocity and Q-factor models (labels/outputs). Extensive databases of seismic data are available, but accurate labels (true velocity models) are scarce, as the true Earth velocity is rarely known. Most deep learning approaches for VMB thus rely on synthetic data, for which both velocity and the seismic response can be accurately computed. However, to apply a synthetic-trained NN to a real target dataset, the training set distribution must be realistic and as close as possible to the unknown distribution of the true subsurface. Building such representative synthetic sets is not trivial. To do so, we use as much a priori information as possible and try to reproduce most features found in the target seismic dataset.

Figure 4.1a presents a shot gather from a survey recorded in the Beaufort Sea in 2014. This shot gather is part of the target dataset for which we wish to predict velocity. It is representative of data acquired over continuous, ice-bonded permafrost. The survey consisted of a single streamer with one hundred twenty receivers spaced every 12.5 m with a minimum offset of 85 m. The source was a 1200 in^3 G-gun array fired every 50 m. From the shot gather, we observe that several wave modes interfere with reflections, notably guided waves, refractions, and multiples that hide any reflections below the top of the permafrost. Figure 4.1e shows a simulated shot from a 1D velocity model optimized by trial-and-error by Fabien-Ouellet et al. (2020) to reproduce the real shot gather in Figure 4.1a. A simple layered model reproduces nearly all wave modes found in the true shot gather, showing that 1D models are good first approximations of the local geology. Furthermore, to obtain such a fit, viscoelastic effects must be considered, and P- and S- wave velocity and quality factors (Q) must be

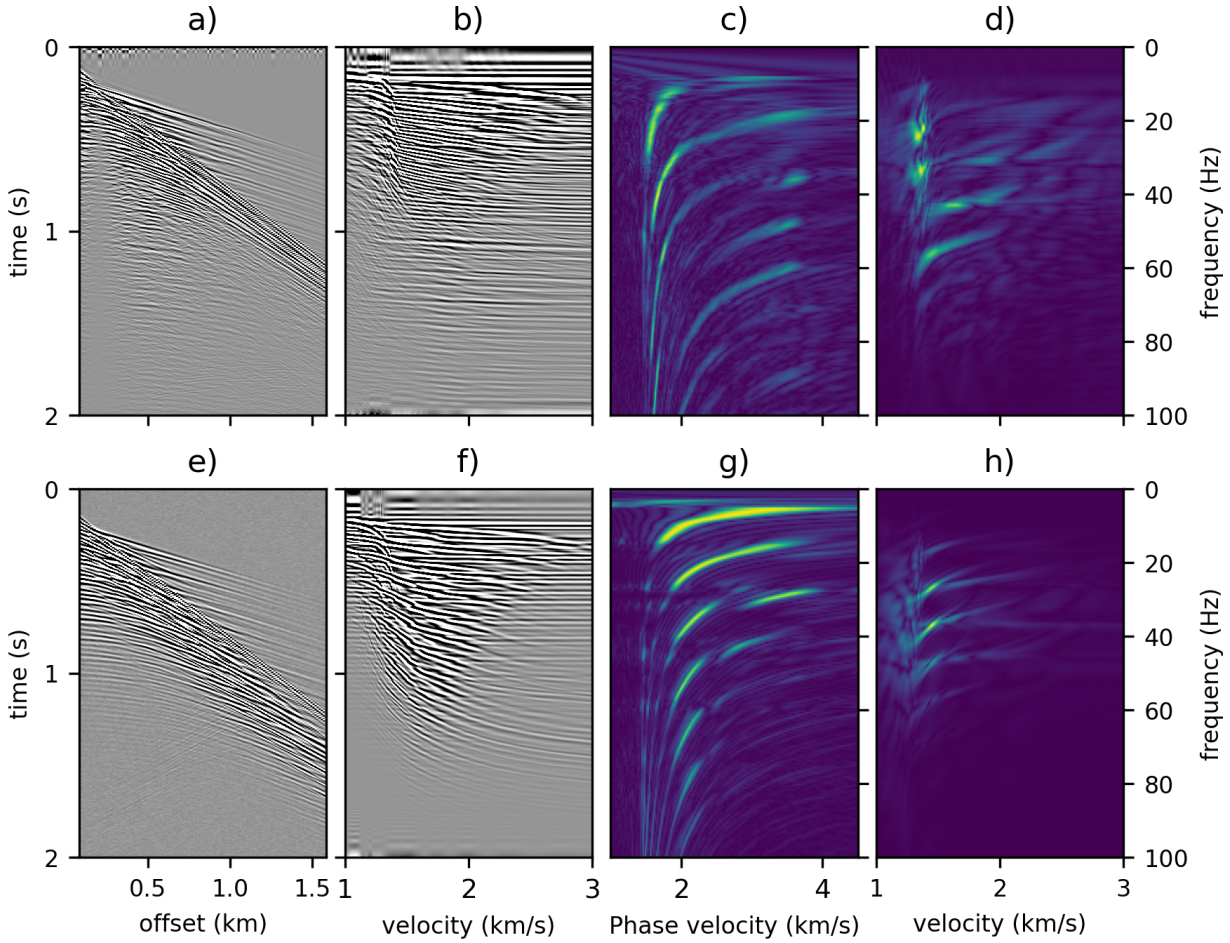


Figure 4.1 (a) Recorded data, and (e) synthetic data in the time-offset $(t-x)$ domain. (b, f) Recorded and synthetic shots in the intercept-slope $(\tau-p)$ domain. (c, g) Recorded and synthetic shots in the frequency-phase velocity $(f-c)$ domain. (d, h) Recorded and synthetic shots in the frequency-velocity $(f-p)$ domain.

carefully defined.

Permafrost model building

The first a priori information that is concluded in the generation of the training set is the expected depth distribution of ice saturation. Ice content in the pore space is highly variable and depends on several variables, including temperature, porosity, grain size (King, 1984; King et al., 1982), and pore fluid salinity (Dou et al., 2016). However, ice content variation with depth in subsea permafrost regions can be generalized as follows: near the seafloor, the ice content of the sediments is usually low or zero due to the heat transfer between the seawater and sediments. Here, ice can appear floating in the pore space or partially

ice-bonding the grains. Ice content increases with depth and, in some cases, reaches fully ice-bonded conditions. Permafrost remains stable under these conditions unless the temperature increases (e.g., due to the geothermal gradient). Then, the ice content decreases gradually until the grains are no longer bonded. Below this depth, the ice content continues to decline until it reaches ice-free conditions.

Sequence	Layer	Vp (km/s)	Vs (km/s)	Q	Thickness (m)	Skip Pr.
Water	Water	1.43	-	-	50-125	0
UF Sediments	UF Sediments	1.50 - 2.10	0.30 - 0.70	45 - 55	20 - 400	0.33
	UF Silts	1.95 - 2.45	0.55 - 1.00	22 - 28		
Permafrost	PF Sands	3.00 - 3.70	0.90 - 1.60	9 - 11	50 - 600	0
	F Shale	2.60 - 3.60	1.10 - 2.00	90 - 110		
	F Sands	3.50 - 4.30	1.25 - 2.15	55 - 65		
	PF Silts	2.10 - 2.90	0.65 - 1.25	18 - 22		
Gas Hydrates	UF Sediments	1.50 - 2.10	0.30 - 0.70	45 - 55	100 - 300	0.5
	PF Silts	2.10 - 2.90	0.65 - 1.25	18 - 22		
UF Sediments	UF Sediments	1.50 - 2.10	0.30 - 0.70	45 - 55	50 - 700	0
	UF Shale	2.10 - 2.50	0.90 - 1.40	70 - 100		

Table 4.1 Lithology and seismic parameters used to create the training dataset. In the table, F., PF., UF., refers to frozen, partially frozen, and unfrozen sediments. The last column represents the probability of the layer being skipped from the model.

To build a dataset representative of the subsurface conditions found at locations where subsea permafrost exists in the Beaufort Sea, we use an object-based geological modeling approach similar to Fabien-Ouellet et al. (2020). We generate random layered velocity models with a specified sequence of units, which are, in order: a water layer, unfrozen sediments, permafrost, gas hydrates, and unfrozen sediments. To generate different conditions of permafrost, we allowed the unfrozen sediments and the gas hydrate units to be absent from the models with a probability of 0.33 and 0.5, respectively. Note that in all the cases, the order of the units in the sequence is consistent.

Expected permafrost velocities are a second important a priori information that was included to build the training set. The seismic properties (V_P, V_S, Q_P, Q_S, and ρ) are defined for common sediment types and ice saturation scenarios based on laboratory measurements and well logs in the Beaufort Sea (Bellefleur et al., 2007; Buckingham, 1997; Dou et al., 2016; Matsushima et al., 2016). Within each unit, a layer can be assigned a specific sediment type, for which seismic properties are drawn from a uniform distribution. Table 4.1 summarizes the limits of the random distribution for each sediment type. Very little information exists in the literature on Q_s of permafrost. We thus make the assumption in what follows that $Q_S = Q_P$ and simply denote it as the quality factor Q. Note that changing the Q_S/Q_P ratio has very little effect on the seismic waveform, which make this assumption reasonable.

1D layered models

The hierarchical transfer learning strategy starts with simple 1D layered models. These can be created cheaply and in large quantity. The 1D training set contains 36,000 synthetic velocity and attenuation models. The 1D models are further divided into four subsets with decreasing layer thickness. The first two subsets have 10,000 samples each, with a minimum layer thickness of 100 m and 75 m (d^{1D}_{100}, d^{1D}_{75}). The third subset has 13,000 samples with a minimum layer thickness of 50 m (d^{1D}_{50}). For the first three subsets, the thickness of the water layer varies from 75 m to 125 m. The fourth subset (d^{1D}_{deep}) contains 3,000 samples with a thicker water layer (125 m to 250 m) and a minimum layer thickness of 50 m. This division into different subsets allows for balancing velocity structures found in the training set and studying the effect of resolution on the NN accuracy.

Figure 4.2 summarizes the P- and S- wave velocity and the inverse quality factor ($1/Q$) distribution as a function of depth for the 1D training dataset. The colormap indicates the number of samples with a specific velocity at a depth position. The dashed white line shows the average velocity variation with depth. The seismic velocities are bound by the water velocities (V_P ∼1.43 km/s, V_S ∼0) at the top of the models, and the values of fully frozen sands ($V_P \sim 4.3$ km/s, $V_S \sim 2.0$ km/s) in the permafrost layer. Similarly, the inverse quality factor is bound by low attenuation in the water layer ($1/Q$ ∼0) and the high attenuation on the partially frozen layers ($1/Q$ ∼0.1). Taylor et al. (2013) conclude that no permafrost should appear 700 m beneath the seafloor in the Beaufort Sea, which is the maximum permafrost thickness in the last glacial maximum. We integrated that information into our training set. Hence, below 500 m, most models have a P-wave velocity between 2.1 km/s and 2.5 km/s and S-wave velocities between 0.9 km/s and 1.3 km/s, which corresponds to the velocities of unfrozen sediments and partially frozen sediments.

2D layered models

To generate the 2D dataset, we kept the same stratigraphic constraints to define random layered velocity models. However, we allowed the layers to have random average dips between -15^o (landward dip) and 15^o (seaward dip) and boundaries following random sinusoidal deformation with random amplitude and frequency.

Similar to the 1D model creation, we defined four subsets of 5 km length models to allow for balancing structures found in the training set. Two subsets (d^{2D}_{100}, d^{2D}_{75}) include 100 models each, with a minimum layer thickness of 100 m and 75 m, respectively. The third subset (d^{2D}_{50}) has 150 models with a minimum layer thickness of 50 m. The water depth for the first

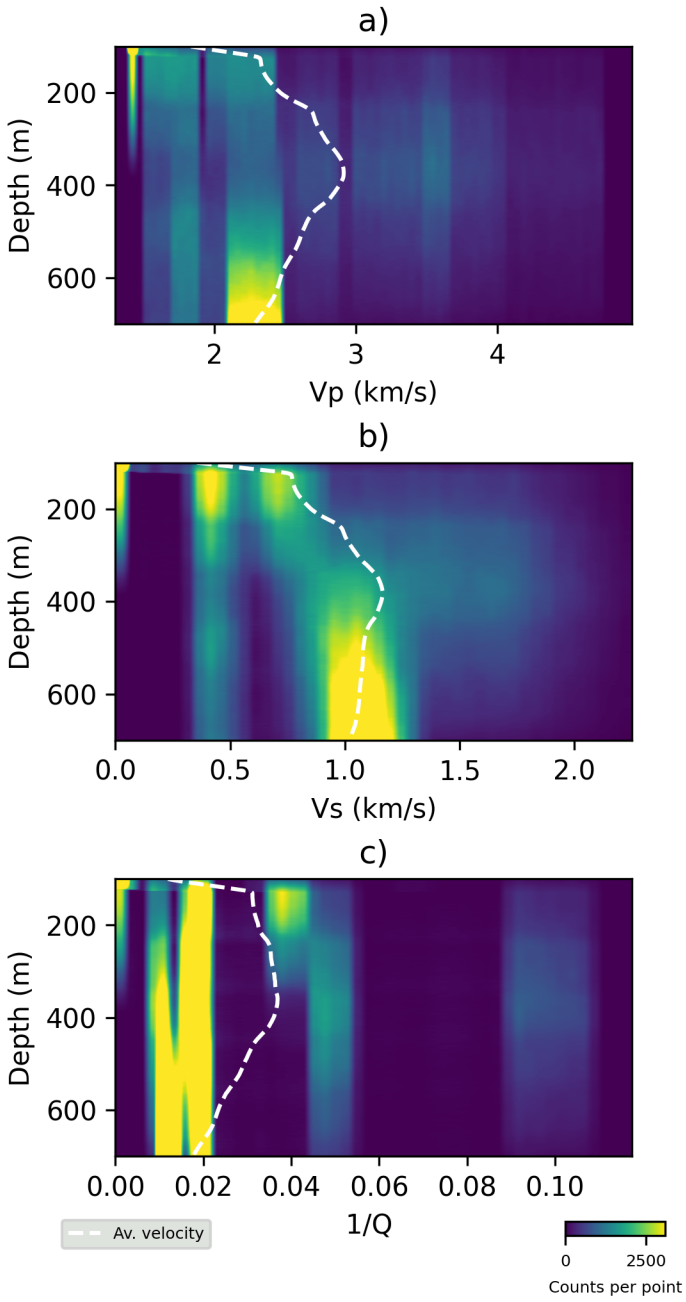


Figure 4.2 P- (a) and S- (b) wave velocity and inverse quality factor ($1/Q$) distributions (c) of the samples in the 1D training dataset.

three models varies from 75 m to 125 m. Finally, the last subset d^{2D}_{deep} includes 50 velocity models with deeper water depths (125 m to 250 m) and a minimum layer thickness of 50 m.

Note that in the 1D mode, the NN outputs the 1D velocity/attenuation profile under a shot/CMP gather. However, in 2D mode, the network does not output the whole 2D model,

but a single 1D profile centered on neighboring CMP gathers. This approach proposed by Kazei et al. (2021) reduces the cost of training and allows generating 2D cross-sections by applying the network to all CMPs of a survey. Hence, each 2D model from the training set can be divided into neighboring CMPs, which leads to a 2D training dataset consisting of 18800 samples in total.

Generation of synthetic seismic data

We generate a synthetic shot gather for each 1D velocity model. For the 2D models, we create 47 groups of synthetic CMP gathers per model. The seismic data is computed with the open-source code `SeisCL` (Fabien-Ouellet et al., 2017). `SeisCL` creates the synthetics by solving the viscoelastic wave equation with finite-difference in the time domain. The source is modeled using a Ricker wavelet with a randomly selected centroid frequency between 30 and 50 Hz. Each shot follows the acquisition geometry of the field dataset recorded in the Beaufort Sea. One hundred twenty receivers are located every 12.5 m with a minimum offset of 85 m. The shot spacing is 50 m. The total length of the synthetic shot is 2 s with a sampling interval of 0.002 s.

Seismic data is usually corrupted by different types of noise that can interfere with the true seismic events. We add two sources of noise: random white noise with a mean amplitude equal to 5% of the maximum amplitude and a random static shift to each trace with a maximum shift 0.002 s. The latter accounts for possible source and receivers position variations during the survey. An example of the resulting time-offset $(t - x)$ gather is presented in Figure 4.1e. Note that the resulting time-offset gather can be interpreted as an image with a size of 1000x120x1.

Seismic events usually overlap in time, hindering the identification of specific phases and limiting our capacity of performing inversion (Duchesne et al., 2022). Transforming the data into a different domain can help amplify and isolate specific seismic events or phases (Bustamante et al., 2023a). In theory, a NN can build automatically a data representation that is optimal for a specified task. In reality, the seismic data contains events with extremely large amplitude and scale variations. Feeding the NN with data in different domains can lead to a better representation that takes into account the features at all scales. Our NN network has four types of inputs, corresponding to the seismic data in different domains.

1- The $(t-x)$ domain is the original domain. If no processing has been applied, the raw data contains all events such as reflections, refractions, multiples and guided waves (Bustamante et al., 2021). Those events provide information on the subsurface structures and can be used to infer velocity models.

2- The zero offset time-velocity or $(\tau - p)$ domain is obtained by applying the hyperbolic Radon transform on the original domain. The hyperbolic Radon transform can be expressed as the integration of the data along a hyperbolic travel-time curve (Ibrahim et al., 2018) following the equation

$$\tilde{u}(\tau, p) = \sum_{k=1}^{N} d(t = \sqrt{\tau^2 + x_k^2/p^2}, x_k), \tag{4.1}$$

where $\tilde{u}$ is the hyperbolic Radon model, d is the recorded data, x is the offset, τ is the zero-offset time, N is the number of offsets, and p is the stacking velocity. Note that $t = \sqrt{\tau^2 + x_k^2/p^2}$ represents the hyperbolic travel times over which the transform integrates. Figures 4.1(b,f) show the $(\tau - p)$ representation of the shot gather in Figures 4.1(a-e). We compute the Radon transform using 200 velocities uniformly spaced between 1.0 km/s and 3.0 km/s generating an image with size 1000x200x1. The $(\tau - p)$ domain should help the NN distinguish between reflections and their multiples and identify the RMS velocity variations in time of the reflections.

3- $(f - c)$ domain. The dispersion of guided waves or surface waves can be visualized in the frequency-phase velocity $(f - c)$ domain (Ibrahim and Fabien-Ouellet, 2021). To transform the data, we apply a linear moveout correction in the frequency-offset domain and sum over the offset axis. The transformation can be expressed as

$$\tilde{u}(f, c) = \sum_{k=1}^{N} d(f, x_k) e^{2\pi j \frac{f x_k}{c}}, \tag{4.2}$$

where $\tilde{u}(f, c)$ is the data in the $(f - c)$ domain, $d(f, x_k)$ is the recorded data in the frequency-offset $(f - x)$ domain, f are the frequencies, and c is the phase velocity. The summation is evaluated over 200 phase velocities uniformly spaced between 1.0 km/s and 4.5 km/s and 200 frequencies uniformly spaced between 0 Hz and 100 Hz. The resulting size of the image is 200x200x1. Figure 4.1(c,g) presents the shot from Figure 4.1(a,e) in the $(f - c)$ velocity image after transforming the information in the $(f - c)$ domain. This domain clearly indicates the modal nature of guided waves, which is sensitive to the P- and S- wave velocity variations with depth.

4- In the frequency-velocity $(f - p)$ domain additional features can also be observed by applying a Fourier transform to the time dimension on the hyperbolic Radon model $\tilde{u}(\tau, p)$. The resulting frequency hyperbolic Radon domain indicates maxima around discrete frequencies and velocities, which depend on the thicknesses of the layers. An example of this domain is presented in Figures 4.1(d,h) after applying a Fourier transformation to Figures 4.1(b,f) along the time axis. The resulting image is evaluated over frequencies varying between 0 Hz

and 100 Hz and has a size of 200x200x1.

4.3.2 Neural Network architecture

The basic architecture of our NN is a multi-input encoder-decoder neural network based on the U-Net architecture (Ronneberger et al., 2015). In this kind of network, the primary function of the encoders is to find features within the inputs. On the other hand, the function of the decoder is to relate those features to the desired output (Badrinarayanan et al., 2017). Figure 4.3a provides an overview of our neural network (NN) architecture. It comprises four encoders designed to extract information from seismic images in separated domains ($t-x$, $f-c$, $\tau-p$, and $f-p$). Additionally, three decoders are employed to accurately estimate the corresponding seismic parameters V_P, V_S, and $1/Q$.

	$(t-x)$	$(\tau-p)$	$(f-c)$	$(f-p)$
Input	1000x120x1	1000x200x1	200x200x1	200x200x1
Conv2D64	1000x120x8 (17x3)	1000x200x8 (17x3)	200x200x8 (17x3)	200x200x8 (17x3)
Conv2D-MaxPooling	200x60x16 (17x3) 40x30x32 (17x3) 20x15x64 (9x3) 10x7x128 (5x3) 1x1x256 (3x3)	250x200x8 (17x3) 125x100x16 (17,3) 62x50x16 (17x3) 31x25x32 (9x3) 15x12x32 (5x3) 7x6x64 (3x3) 3x3x128 (3x3) 1x1x256 (3x3)	100x100x8 (17x3) 50x50x16 (17x3) 25x25x32 (9x3) 12x12x64 (5x3) 6x6x128 (3x3) 3x3x256 (3x3) 1x1x256 (3x3)	100x100x8 (17x3) 50x50x16 (17x3) 25x25x32 (9x3) 12x12x64 (5x3) 6x6x128 (3x3) 3x3x256 (3x3) 1x1x256 (3x3)

Table 4.2 Encoder's architectures for the $t-x$, $\tau-p$, $f-c$, and $f-p$. Information in the cells describes the resulting size after passing the layer. The filter kernel in the convolutional 2D layers is presented in parentheses.

We design four single input encoders based on convolutional 2D layers. For each of the four data domains, the encoder takes the input image and sequentially reduces its size to 1x1x256. Each layer of the encoder is composed of a 2D-convolution, followed by a leaky-RELU activation function (Rectified Linear Unit, Maas et al. (2013)) and a Max-Pooling operation that downsamples the input size. Concurrently, the number of features increases with the number of layers from 16 to 256. The kernel size of the 2D convolution layers sequentially varies from 17x3 applied to the input image to 3x3 applied to the last layer. Figure 4.3b presents the network for the $(t-x)$ input with a size 1000x120x1. Table 4.2 summarizes the encoders' architecture. In total, five 2D Convolutional-MaxPooling layer pairs were employed for the $(t-x)$ domain, eight for the $(\tau-p)$ domain, seven for the $(f-c)$ and $(f-p)$ domains.

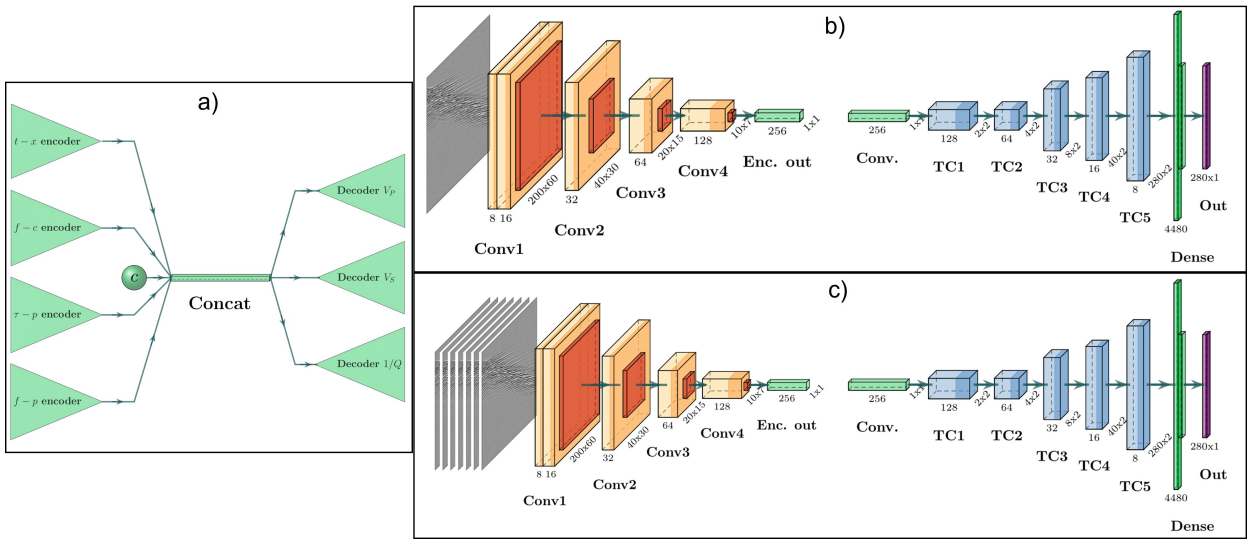


Figure 4.3 (a) Overview of the neural network (NN) architecture. (b) 1D encoder and decoder for single-input $(t-x)$ data. (c) 2D encoder and decoder for single-input $(t-x)$ data. "Conv" represents the combination of 2D convolutional and max-pooling layers, "TC" refers to the transpose convolution layer, and "enc out" denotes the output vector of the encoder.

Multi-input neural network

To build a multi-input network, we concatenate the outputs of the encoders into a single vector of size 1x1x1024 and use this vector as the input of three identical decoders (Figure 4.3a). Each decoder sequentially upsamples the features using multiple transposed 2D Convolutional layers (Figure 4.3b). First, the input vector to the decoder downsamples to 1x1x256 by a convolution with a filter kernel of size 1x1. This convolution is equivalent to applying a dense layer to the input vector but keeps the same number of dimensions. Then, five transposed 2D convolutional layers upsample the features from the previous ones and combine them with the corresponding decoder filter. The upsampling is performed by defining different strides and using padding such as the dimension of the output is $\texttt{output}_{size} = \texttt{input}_{size} \times \texttt{stride}$. The output sizes after each transpose 2D-convolution layers are 2x2x128, 4x2x64, 8x2x32, 40x2x16, and 280x2x8. Finally, a pair of flat and dense layers reduces the size of the features to 280.

During the training, we compare the three vectors returned by the NN to the V_P and V_S velocity and $1/Q$ models created in our dataset section. As the sampling interval for the velocity models is 2.5 m, the 280 samples allow us to train our network to predict velocities up to a subsurface depth of 700 m. This depth interval was selected to match the anticipated maximum depth of permafrost in the Beaufort Sea (Taylor et al., 2013).

2D network

The 2D architecture keeps the same architecture described for the 1D case. However, the inputs to the encoders are now 11 CMP gathers located within a radius of 250 m around the center point. For example, Fig. 4.3c presents such an architecture for CMP gathers in the $(t-x)$ domain. Each CMP gather is transformed into different domains and the input sizes for the 4 encoders are now 1000x120x11 for the $(t-x)$ domain, 1000x200x11 for the $(\tau-p)$ domain, 200x200x11 for the $(f-c)$ domain, and 200x200x11 for the $(f-p)$ domain. Each encoder inputs the groups of images and outputs a single vector of size 1x1x264. Finally, the decoders input the concatenated outputs from the encoders and return the inverted V_P, V_S, and $1/Q$ at the center point. Note that although the output is a 1D model under the center input CMP, we can apply the network to all CMPs in a survey to build a 2D model. Table 4.3 presents the modified description of the encoders' layers. Note that no modification is applied to the decoders.

	$(t-x)$	$(\tau-p)$	$(f-c)$	$(f-p)$
Input	1000x120x11	1000x200x11	200x200x11	200x200x11
Conv2D64	1000x120x8 (17x3)	1000x200x8 (17x3)	200x200x8 (17x3)	200x200x8 (17x3)
Conv2D-MaxPooling	200x60x16 (17x3) 40x30x32 (17x3) 20x15x64 (9x3) 10x7x128 (5x3) 1x1x256 (3x3)	250x200x8 (17x3) 125x100x16 (17,3) 62x50x16 (17x3) 31x25x32 (9x3) 15x12x32 (5x3) 7x6x64 (3x3) 3x3x128 (3x3) 1x1x256 (3x3)	100x100x8 (17x3) 50x50x16 (17x3) 25x25x32 (9x3) 12x12x64 (5x3) 6x6x128 (3x3) 3x3x256 (3x3) 1x1x256 (3x3)	100x100x8 (17x3) 50x50x16 (17x3) 25x25x32 (9x3) 12x12x64 (5x3) 6x6x128 (3x3) 3x3x256 (3x3) 1x1x256 (3x3)

Table 4.3 2D NN encoder's architectures for the $t-x$, $\tau-p$, $f-c$, and $f-p$. Information in the cells describes the resulting size after passing the layer. The filter kernel in the convolutional 2D layers is presented in parentheses.

To apply TL, we use the pretrained weights and bias learned with the 1D dataset. This is possible as the 1D and 2D networks are identical, except for the first layer. To handle the increase in the last component of the image dimension (from 1 to 11), we expanded the shape of the weights in the first convolutional layer from (17,3,1,8) to (17,3,11,8) for the $t-x$ and $\tau-p$ encoders, and from (3,3,1,8) to (3,3,11,8) for the $f-c$ and $f-p$ encoders. To apply TL for the first layer, the initial weights are recalculated as:

$$w^{2D}_{i,j,k,l} = a_k * w^{1D}_{i,j,l}$$
$$a_k = \begin{cases} 1.0 & \text{if } k = 6 \\ 0.1 & \text{otherwise} \end{cases} \tag{4.3}$$

Here, $w^{2D}_{i,j,k,l}$ represents the weights in the first layer of the 2D architecture, $w^{1D}_{i,j,l}$ are the weights in the first layer of the 1D architecture, and $a_k = 1$ if i is the image at the center point, and $a_k = 0.1$ otherwise.

4.3.3 Training

We trained the 1D network with the Adam optimizer (Kingma and Ba, 2014) with a learning rate of 1e-4 and batches of 10 samples. To help stabilize the optimizer, we normalized the inputs and true models between 0 and 1 for the minimum and maximum expected velocity values. For training the neural network, we define a custom loss function that includes several regularization terms:

$$\begin{aligned} \mathscr{L}(\boldsymbol{y_l}, \boldsymbol{y_p}) = \quad & \iota_1 \left(\frac{1}{n} \sum_{i=0}^{n-1} \frac{|\boldsymbol{y_l}_i - \boldsymbol{y_p}_i|_1^1}{|\boldsymbol{y_l}_i|_1^1} \right) \\ & + \iota_2 \left(\frac{1}{n} \sum_{i=0}^{n-1} \frac{|\boldsymbol{y_l}_i - \boldsymbol{y_p}_i|_2^2}{|\boldsymbol{y_l}_i|_2^2} \right) \\ & + \alpha \left(\frac{1}{n} \sum_{i=0}^{n-1} \frac{|d\boldsymbol{y_l}_i/dz - d\boldsymbol{y_p}_i/dz|_2^2}{|d\boldsymbol{y_l}_i/dz|_2^2} \right) \\ & + \beta \left(\frac{1}{n} \sum_{i=0}^{n-1} \frac{|d\boldsymbol{y_p}_i/dz|_1^1}{|\boldsymbol{y_p}_i|_1^1} \right) \\ & + \theta \left(\frac{1}{n} \sum_{i=0}^{n-1} |\max(\boldsymbol{y_l}_i) - \max(\boldsymbol{y_p}_i)| \right) \end{aligned} \tag{4.4}$$

where $\boldsymbol{y_l}$ and $\boldsymbol{y_p}$ are the labeled and predicted normalized velocities and Q-factors, and n is the batch size. The first and second terms in Equation 4.4 account for the mean absolute and mean squared errors. The third term calculates the mean squared error of the derivative of the velocities with respect to depth. The fourth term aims to promote blocky inversion by minimizing the gradient of the velocities with respect to depth. Finally, the fifth term minimizes the error between the maximum labeled and predicted velocities. The weights of the five terms are set to $\iota_1 = 0.05$, $\iota_2 = 0.63$, $\alpha = 0.02$, $\beta = 0.1$ and $\theta = 0.2$ following a trial and error procedure.

The loss function is independently evaluated for each one of the outputs (V_P and V_S velocities and $1/Q$). The total loss ($\mathscr{L}_{Total}$) is the weighted sum of the resulting losses following:

$$\mathscr{L}_{Total} = 0.49\mathscr{L}_{V_P} + 0.49\mathscr{L}_{V_s} + 0.02\mathscr{L}_{1/Q} \tag{4.5}$$

where $\mathscr{L}_{V_P}$, $\mathscr{L}_{V_s}$ and $\mathscr{L}_{1/Q}$ are the losses evaluated for V_P, V_S and $1/Q$, respectively.

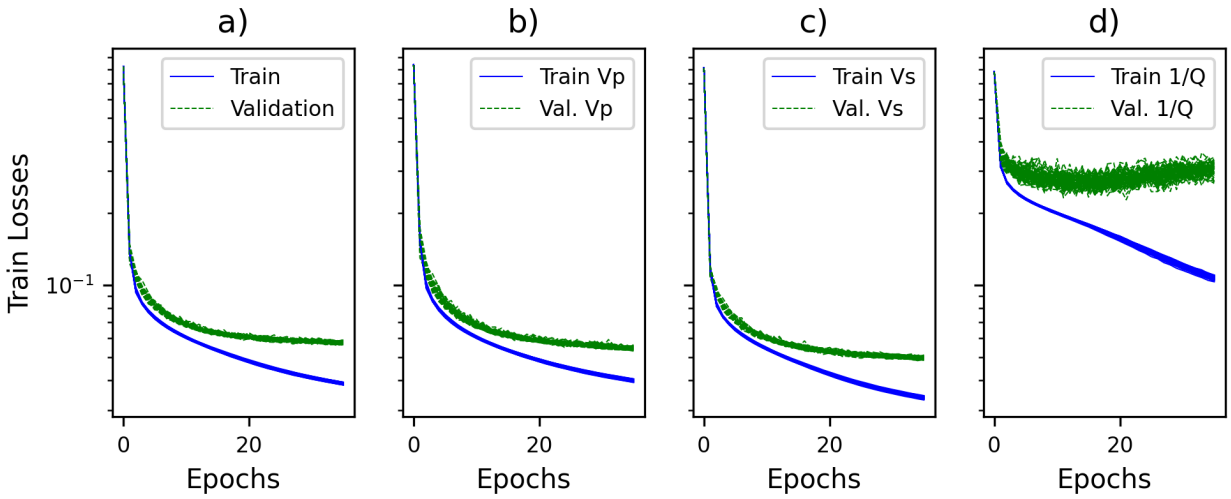


Figure 4.4 (a) Total loss after training 64 similar 1D NNs. (b-d) Losses specifically evaluated for V_P, V_S, and $1/Q$ across the 64 trained NNs. An epoch represents one complete pass through the training data.

We verified the network's performance after using all the training datasets in one cycle of training (each epoch) by evaluating the total loss on the validation dataset. The 1D and 2D validation datasets contains 3600 and 1800 samples, respectively. They were randomly generated following the same procedure as the training set. We stopped the training when the difference between the loss function for the training and validation datasets started increasing. We trained an ensemble of 64 networks with different initialization in parallel, and averaged their output to improve the accuracy in areas with high uncertainties (Opitz and Maclin, 1999; Simon et al., 2023). Figure 4.4 presents the total losses (Figure 4.4a) and individual losses (Figure 4.4b-d) evaluated on the training and validation datasets after each epoch for the 1D training. We observe over-fitting after epoch 35 with an average total loss value of 0.039 for the training dataset and 0.057 for the validation dataset (down from 0.822 and 0.822 before training).

We implemented our TL methodology and retrained the NNs with the 2D dataset, using the weights pretrained in 1D. The retraining lasted for 20 additional epochs until the loss stopped

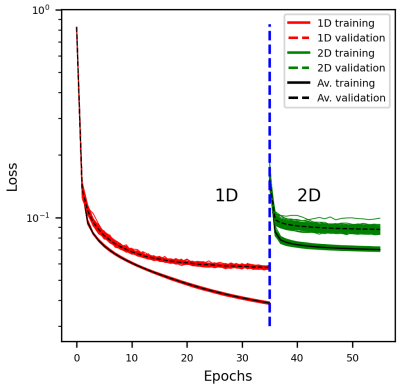


Figure 4.5 Total losses for the 1D and 2D datasets after training 64 similar NN. An epoch represents one complete pass through the training data.

decreasing. Figure 4.5 presents the total losses for the 1D and 2D training and validation datasets. It is worth noting that the loss values increased when transitioning from the 1D to the 2D dataset due to the higher complexity of the latter. This increased complexity also explains the higher standard deviation and final loss values for the 2D dataset after training compared to the resulting loss values for the 1D dataset.

4.4 Results

4.4.1 Testing dataset

To evaluate our 1D neural network performance, 720 new samples were generated using the same methodology as the training set. The test dataset is divided into four subsets with decreasing layer thickness. Two subsets similar to d_{100}^{1D} and d_{75}^{1D} contain 200 samples each. Another subset containing 260 samples is similar to d_{50}^{1D}. Finally, a subset with 60 samples resembling d_{deep}^{1D} accounts for thicker water layers.

Figure 4.6 presents the V_P, V_S and $1/Q$ predictions of the ensemble of 64 1D neural networks for the four subsets of the testing dataset. For each test set, we present the model with the median $L2$ error. In general, we observe that the NN predictions correspond to a smoothed version of the true velocity and attenuation models. The error is lower for V_P and V_S than for the $1/Q$ values, which is consistent with the training curves. Since seismic data are more sensitive to the velocities than the attenuation, the attained accuracy is surprisingly high. As shown in Figure 4.6i-l), attenuation predictions are more susceptible than velocities to contain large errors.

The low resolution of the predictions can be explained by the finite resolution of the seismic

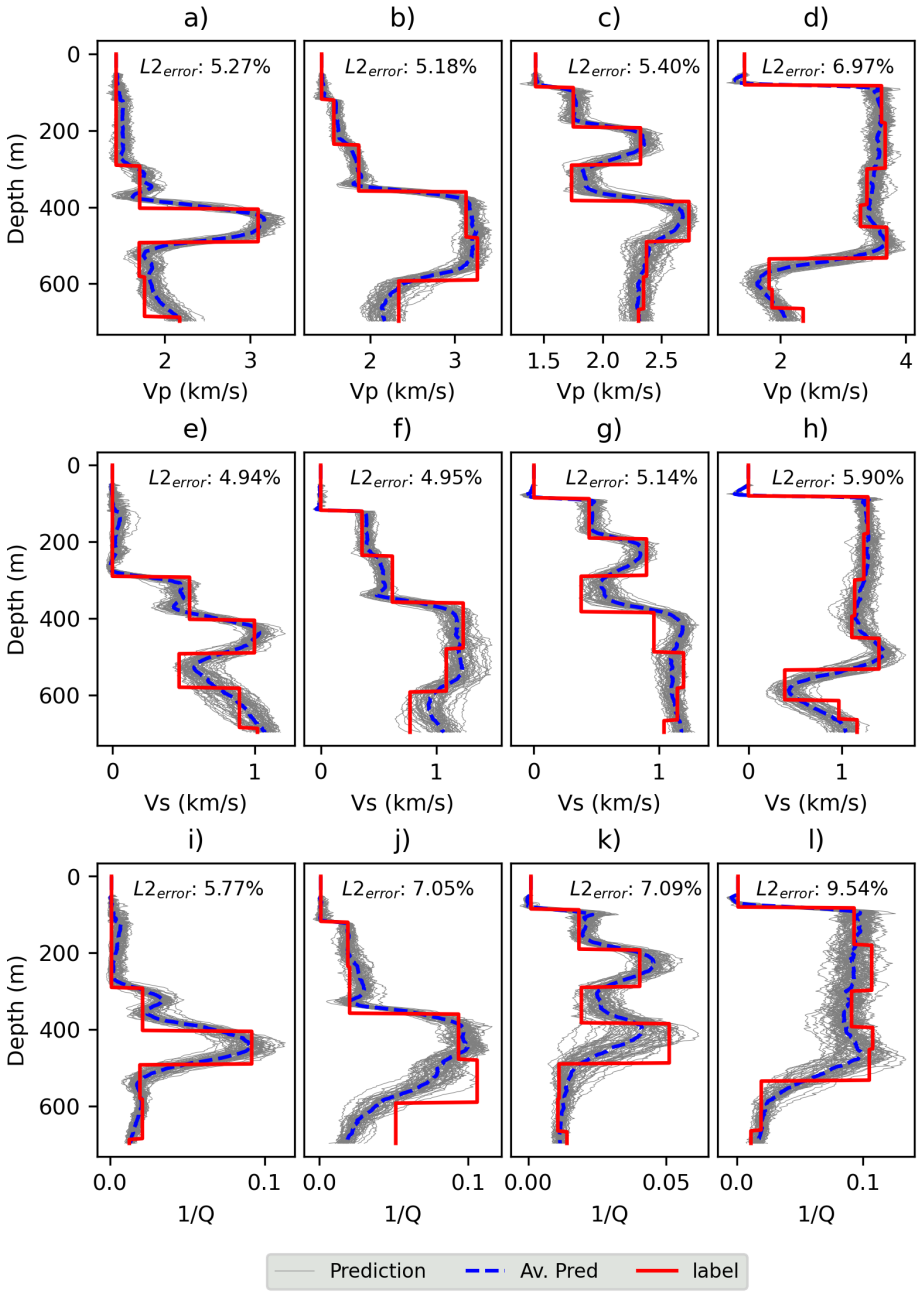


Figure 4.6 Results of the 64 1D neural networks in the 1D testing dataset (grey lines). a), e), and i) present the results from the sample corresponding to the median error on the thicker water layer subset, b), f), and j) on the 100 m thickness subset, c), g) and k) on the 75 m subset, and d), h) and l) on the 50 m subset. Red lines denote the labels, whereas broken blue lines present the average value of the 64 predictions. The average L2 error is presented for each subset and true model.

data. Moreover, it can also be a consequence of the strong influence of the mean square arguments in Equation 4.4, which provide non-sparse solutions but introduce stability to the

inversion (Ellis and Oldenburg, 1994). In addition, we note that the results for the different networks are closer to the mean value at shallow depths. In general, the standard deviation increases below a large velocity reversal and with depth. This behaviour can be explained by multiple arrivals overlapping in different domains, which hinder the identification of velocity changes under the permafrost base. Furthermore, the sensitivity of the seismic data usually drops below a large velocity contrast. We also note that the average $L2$ error values for the four subsets increase with decreasing layer thickness. Thus, the evaluated average $L2$ error in d_{100}^{1D} is 5.10%, 4.91% and 7.02% for V_P, V_S and $1/Q$, while in d_{50}^{1D} are 6.95%, 5.90% and 9.52%. Relatively low errors also occur in the d_{deep}^{1D} subset where the average $L2$ error has values of 5.17%, 4.85% and 5.75% for V_P, V_S and $1/Q$, respectively.

We calculate the point-by-point $L2$ norm of the errors for each one of the samples in the 1D testing dataset. Figure 4.7 presents the distribution of the normalized errors for V_P, V_S and $1/Q$ on the four subsets of samples in the testing dataset. The error generally increases with decreasing minimum thickness of the layers. This error trend is expected as the interference between different arrivals increases with thinner layers which can lead to the misidentification of events. It can also be associated with the finite resolution of the data.

The distribution of the $L2$ errors shows again that the NN predictions are more accurate for velocities than for $1/Q$. Although the mean normalized error is below 5% for each property, the $1/Q$ distribution has a longer tail that reach a maximum error of 28%, whereas the maximum error is 15% and 20% for V_S and V_P, respectively. This means that large errors are much more likely for $1/Q$ than for velocity. This is expected as attenuation is considered more challenging to estimate than velocity.

To assess the performance of the 2D NN, we constructed a 2D testing set comprising 860 samples following the same procedure outlined in the dataset section. We generated 215 new samples for each subset d_{deep}^{2D}, d_{100}^{2D}, d_{75}^{2D}, and d_{50}^{2D}. Figure 4.8 shows the predictions of the ensemble of 64 2D neural networks on the 2D testing dataset. The columns of the figure represent the label and predicted values for V_P, V_S, and $1/Q$, respectively. The rows display random examples from distinct subsets, specifically the 100 m subset, the 75 m subset, and the 50 m subset, from top to down, respectively.

Overall, we find a good agreement between the inverted and true models. The average L2 norm of the error further confirms this observation. Similar to the results in 1D, we note that the values of the L2 norm are higher for 1/Q compared to those for V_P and V_S. This indicates a greater sensitivity of seismic data to velocity variations than attenuation. Additionally, we observe an increasing trend in the average L2 norm with decreasing layer thickness. This behaviour aligns with the patterns observed in the 1D dataset.

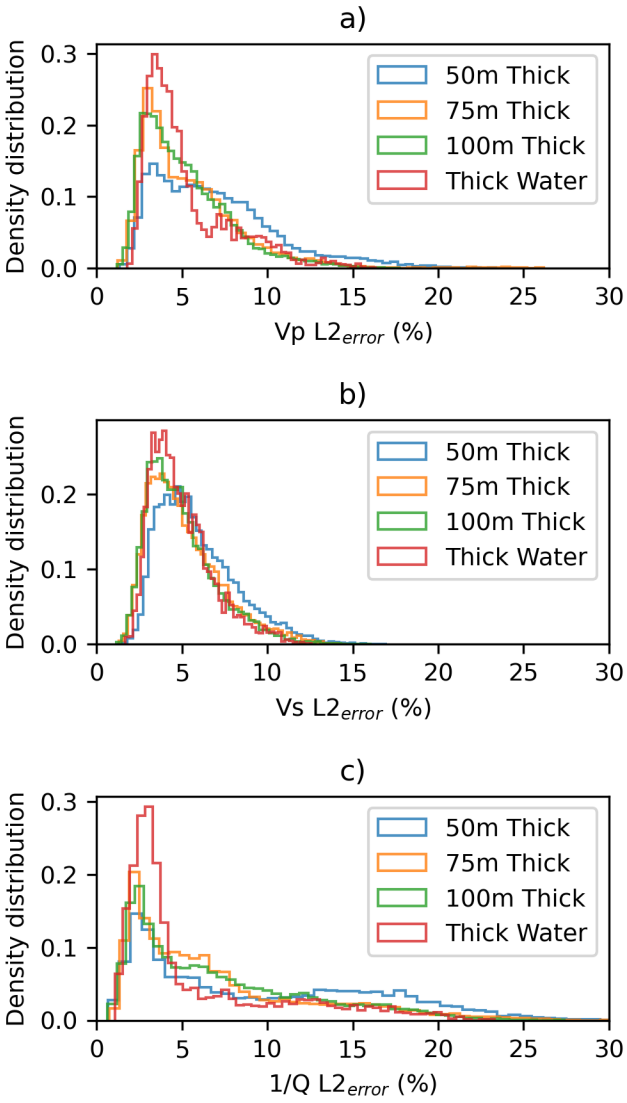


Figure 4.7 L2 V_P (a), V_S (b), and $1/Q$ (c) error density distribution for the four subsets in the testing dataset.

Figure 4.8 also presents the average structural similarity index (SSIM) (Wang et al., 2004) for the assessed subset and model. Similar to the L2 norm of the errors, the SSIM demonstrates lower accuracy (lower SSIM value) for the inverted $1/Q$, ranging from 0.52 to 0.66. Regarding the velocities, the average SSIM ranges from 0.6 to 0.73.

4.4.2 Synthetic 2D model

The previous test cases were created with the same procedures, which does not allow us to judge the generalizability of the network. To evaluate the capacity of the NN to generalize

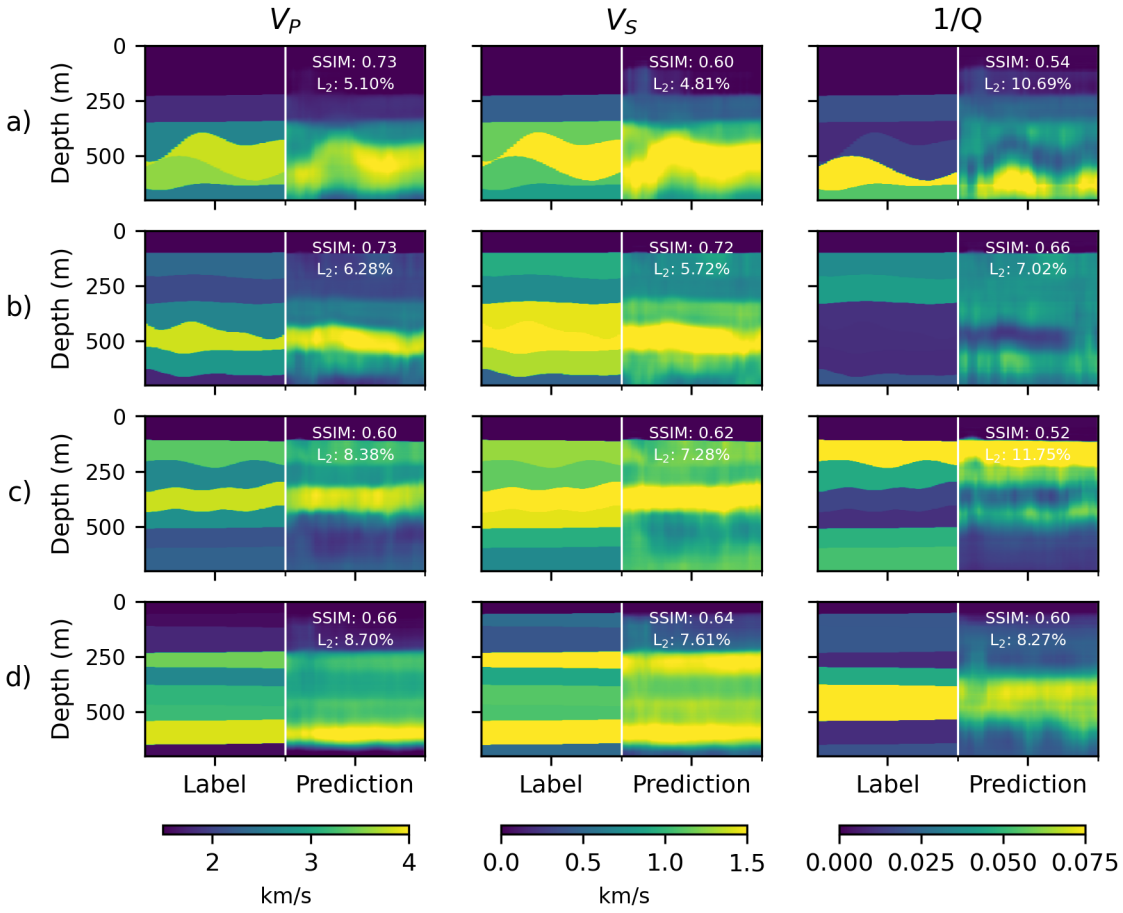


Figure 4.8 True model and average prediction for the ensemble of 64 2D neural networks evaluated in the 2D testing dataset. The columns present the results for V_P, V_S and $1/Q$. The rows show, from top to down, random examples of the ticker layer subset (a), the 100 m subset (b), the 75 m subset (c) and the 50 m subset (d). The average structural similarity index (SSIM) and L2 error are presented for each subset and true model.

to out-of-distribution 2D velocity models, we created a synthetic dataset resembling wedge-shaped permafrost conditions (Figure 4.9). The model contains velocity variations within each layer created with unconditional Gaussian simulations based on the statistics found in well logs (Hu et al., 2013). The stochastic features are not found in the training set and will help assess the capacity of the NN to generalize to models that reassembles more closely to the expected distribution of subsea permafrost.

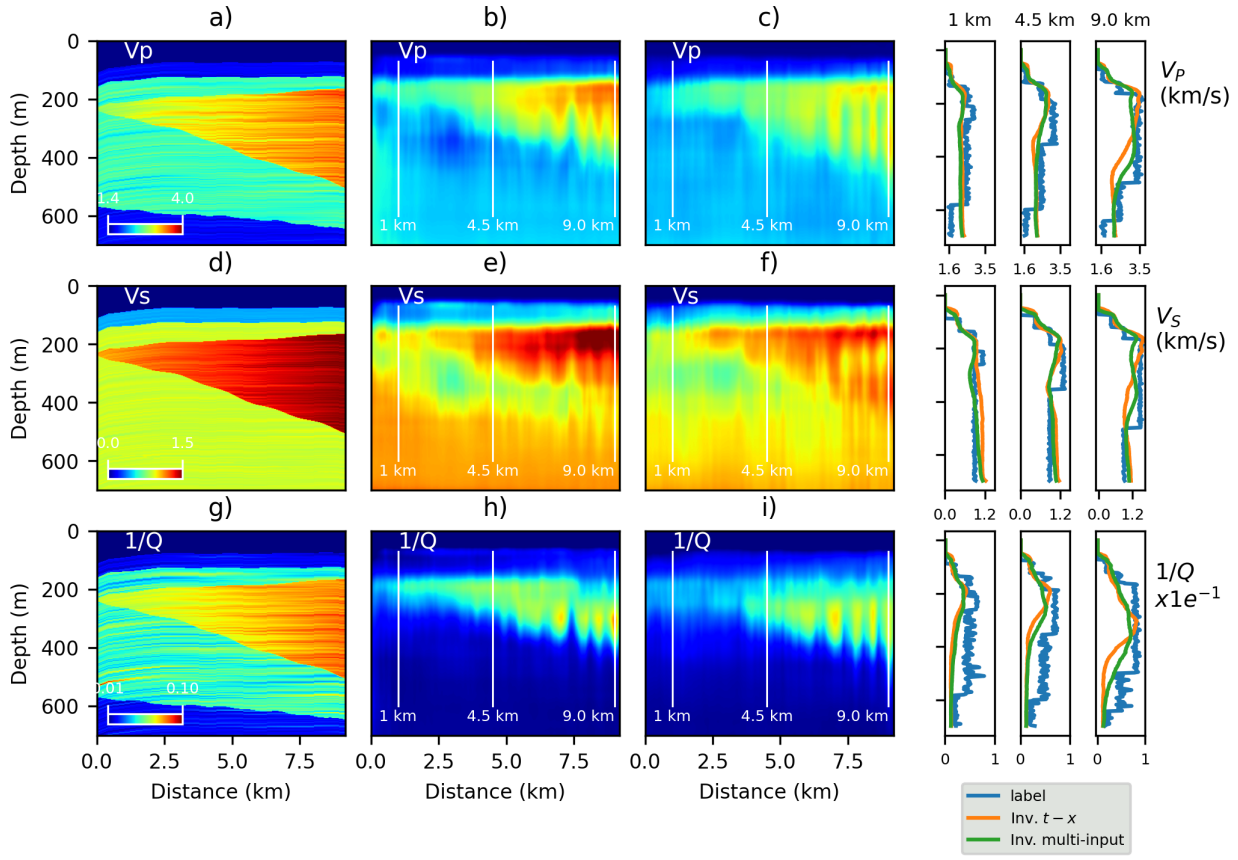


Figure 4.9 The evaluation of the NN on a synthetic dataset. a), d) and g) present the true V_P and V_S velocity and $1/Q$ models. b), e) and h) present the results after evaluating a NN with a single time-offset domain input. c), f) and i) present the results from the multi-input NN. The curves on the right compare the labels and the inverted (Inv.) results from the single-input $(t-x)$ and multi-input NNs evaluated at 1 km, 4.5 km and 9 km.

The outputs of two different NN are compared in Figure 4.9. The first NN (second column) takes a single-input, CMP gathers in the $(t-x)$ domain. The second NN (third column) uses all four inputs, employing CMPs images in four domains. The white lines represent wells at 1 km, 4.5 km and 9 km. The figures located at the right compare labels and results at the positions of the wells from the single-input and multi-input NNs. Although some

inversion artifacts are visible in the resulting velocity models, both NNs recover the wedge structure. In general, both NNs detect the top of the high-velocity layer. The bottom of the wedge is not as clearly defined from the V_P and V_S velocity models but is still visible and can be interpreted. On the other hand, the inverted $1/Q$ model provides a cleaner view of the permafrost base. However, the resulting $1/Q$ model is so smooth that it does not allow the identification of the transition from partially frozen to frozen permafrost and vice-versa.

Note that the NNs cannot reproduce the fine layering due to the limited resolution of the seismic data. We also note that the networks underestimate the velocity in the rightmost part of the model, where permafrost is the thickest and velocity is the largest.

Overall, the multi-input NN performs slightly better than the single-input. In particular, the multi-input NN does a better job of estimating the velocities between 5 to 9.2 km distance and better delimits the base of the wedge, which represents the transition from frozen (V_P >2.5 km/s, V_S >1 km/s) to partially frozen permafrost (2.2 km/s< V_P >3.0 km/s, 0.7 km/s< V_S >1.2 km/s). The profiles at the location of the wells show that the single-input NN underestimates more severely the velocity at the permafrost base for the inverted V_P, V_S and $1/Q$ models. Thus, the multi-input NN performs best for resolving the permafrost base.

	Test Dataset						**Wedge Model**								
	L1 error (%)			**L2 error (%)**			**L1 error (%)**			**L2 error (%)**			**SSIM**		
	Vp	**Vs**	**1/Q**	**Vp**	**Vs**	**1/Q**	**Vp**	**Vs**	**1/Q**	**Vp**	**Vs**	**1/Q**	**Vp**	**Vs**	**1/Q**
t-x	4.28	4.12	5.37	6.25	5.78	8.01	10.02	6.38	12.56	12.38	7.79	15.14	0.459	0.675	0.001
Multi-input	3.97	3.73	5.13	5.84	5.32	7.83	9.57	5.45	11.56	11.34	6.83	13.82	0.470	0.706	0.001

Table 4.4 Calculated L1 and L2 errors and the SSIM index for the 1D testing dataset and the 2D wedge model.

Table 4.4 summarizes the L1 and L2 norms of the errors and the SSIM index calculated for the testing dataset and the wedge model using the single-input and multi-input NNs. The multi-input NN provides the lowest L1 and L2 norm of the error values and the highest SSIM index for the testing and the wedge model. As expected, we observe a better performance of the multi-input approach as it emphasis different features at different scales of the seismic data from which the NN can build a more robust representation during training.

4.4.3 Field data

We evaluate the applicability of our neural network to real dataset recorded in the Beaufort Sea. The data consists of line 06 of the ARA05C survey recorded in 2014 in a joint effort between the Korean Polar Research Institute (KOPRI) and the Geological Survey of Canada (GSC). To align the frequency content of the training dataset and mitigate the influence of

low-frequency guided waves, a bandpass filter ranging from 15 Hz to 65 Hz is applied to the shot gather. This process helps to match the training dataset's frequency characteristics and minimize the discrepancies caused by inaccurately reproduced low-frequency guided waves in synthetic models. Note that the shot gather in Figure 1a is located at position 38.7 km of the seismic line.

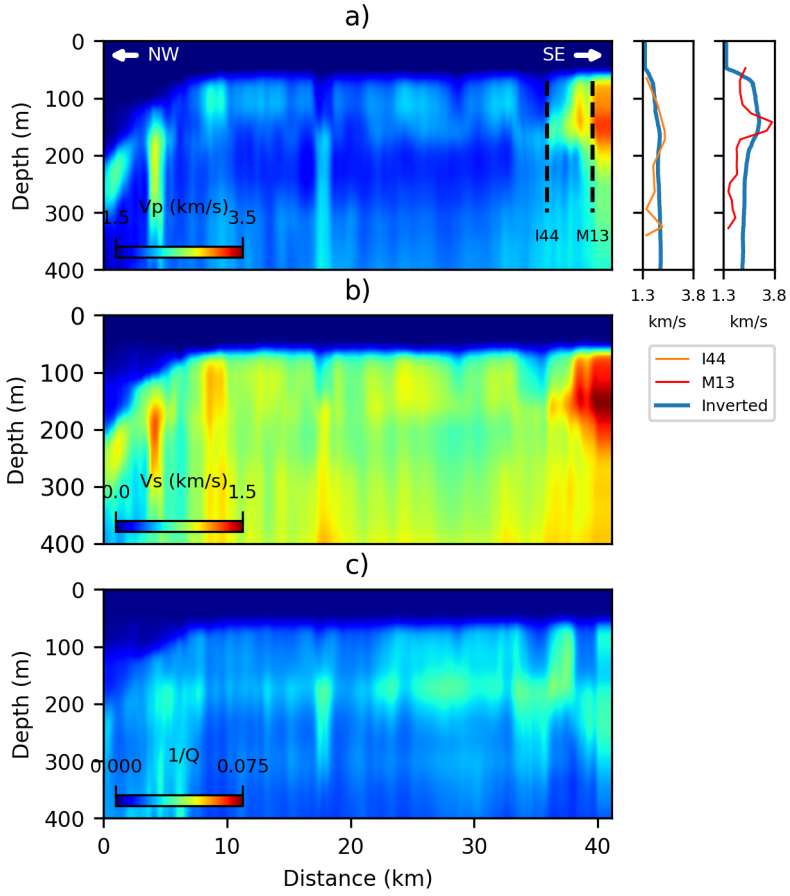


Figure 4.10 Evaluating NN on seismic line ARA05C-6 recorded in the Beaufort Sea. Black dashed lines represent the location of shallow velocity measurements collected in the wells Kopanoar I44 and M13, located close to the seismic line. The plots on the right compare the well logs and the inverted V_P velocities.

Figure 4.10 presents the V_P, V_S and $1/Q$ models estimated by the 2D multi-input NN on seismic line ARA05C-6. Although some inversion artifacts are present, the resulting image shows permafrost non-uniformly distributed along the line (under 300 m depth). Over 10 km towards the SE, the inverted permafrost distribution agrees with the results of Pullan et al. (1987). In their study, Pullan et al. (1987) used refraction analysis to map the transition from low ice-bonded sediments to ice-bonded sediments to the east (from left to right). Our

results locate seismic velocities associated with ice-bonded permafrost over 35 km. In this region, the permafrost table is located a few meters under the ocean bottom and continues until ∼300 m. Under ∼300 m, the seismic velocities decrease to values typically found on partially ice-bonded permafrost (V_P between 2.2 and 3.0 km/s, V_S between 0.7 and 1.2 km/s). Similarly, $1/Q$ reaches its maximum at ∼300 m suggesting the presence of partially frozen sediments. In general, the $1/Q$ model provides a cleaner interpretation of the base of the permafrost. However, this result should be carefully analyzed and linked to the velocity results, given the high uncertainties found during the training and testing approaches. Our results also agree with shallow velocity logs recorded in wells Koponoar I44 and M13, which are close to the SE-end of the seismic line.

Towards the NW, the first 10 km of the survey is located near the shelf edge. Our NN results indicate an increase in seismic velocities towards the continental shelf (from right to left). The maximum velocity values in this region are located around 5 km. The observed increase in velocities corresponds to the presence of "pingo-like" features, which are ice mounds forming within the near-surface or at the seafloor. These features have been documented along the outer-continental shelf of the Canadian Beaufort Sea (Paull et al., 2007).

4.4.4 1D vs 2D inversion

To assess the performance of our TL methodology, we inverted the seismic parameter V_P from lines ARA05C-06 and ARA05C-07 using the 1D and the 2D NNs. Both lines are recorded mainly within the 100 m isobath and intersect in a region previously identified by Pullan et al. (1987) as transitioning from discontinuous ice-bonded to low ice-content permafrost. The inverted seismic parameters are presented in Figs. 4.11a and 4.11b, respectively.

From Fig. 4.11, we can identify velocity values associated with discontinuous ice-bearing permafrost conditions ($V_P \geq 2.2$ km s^{-1}, $V_S \geq 0.7$ km s^{-1}) in the interval 0-30 km of the line ARA05C-06 (first column). Over 30 km, there is a gradual increase in the velocity values and the apparent thickness of the permafrost towards the SE. On the other hand, the inverted seismic parameters from line ARA05C-07 (second column) show velocity values associated with transitions from low ice content to discontinuous ice-bonding permafrost. In particular, there is an increase in seismic velocities in a continuous interval between 15 and 23 km. These transitions from low ice content velocities to discontinuous ice-bonded permafrost in both lines agree with the permafrost distribution map presented in Pullan et al. (1987).

Although the inverted seismic parameters obtained from the 1D and 2D networks shown in Figs. 4.11a and 4.11b exhibit similar velocity distribution and overall characteristics, the 2D NN predictions in Fig 4.11b provide a more precise representation of the geological structures

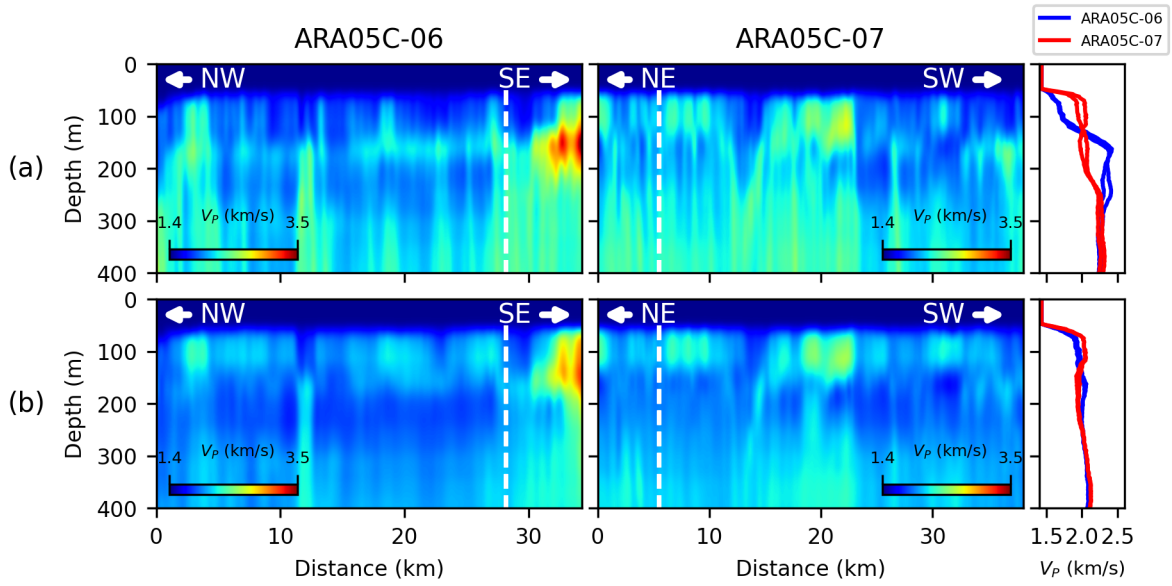


Figure 4.11 Inverted V_P velocities for seismic lines ARA05C-06 (first column) and ARA05C-07 (second column) using the 1D (a) and the TL approach (b). The white dashed lines mark the intercept between the lines. The third column shows the inverted seismic parameters at the intercept.

associated with permafrost. Several factors contribute to the improvement in the inverted parameters. First, the artifacts (vertical lines) resulting from inverting 2D data using the 1D NN have a lesser impact in Fig 4.11b. Second, the TL approach smooths the inverted models due to the inclusion of information from adjacent CMPs. Third, the information about adjacent points allows the NN to better determine the depth and thickness of the geological structures. For instance, the depth and thickness of the discontinuous low ice-content permafrost are better observed in both lines using the TL approach. Additionally, the base of the permafrost can be more accurately identified from the inverted models in Fig. 4.11b.

To further analyze the effect of implementing the TL methodology, we compare the inverted models at the intersection of lines ARA05C-06 (recorded in the NW-SE direction) and ARA05C-07 (recorded in the NE-SW direction). The third column in Fig. 4.11 displays the inverted velocity models at the CMPs around the intercept point using the 1D and TL NNs, respectively. Note that with the 1D NN, despite the similarity in the characteristics of the models, such as the depth of the ocean bottom and the permafrost thicknesses, there are notable variations in the inverted values between seismic lines. Conversely, with the TL methodology, the inverted parameters exhibit consistent values regardless of the azimuth of

the acquisition. This finding confirms that our TL methodology enhances the performance of the 1D NN and performs well in the presence of lateral velocity variations.

In Figure 4.12 we compare our results with those obtained using conventional travel time first break tomography (TTFBT) (Li et al., 2013; Taillandier et al., 2009). For this case, we restrict our analysis to the last 20 km of the seismic line and a maximum depth of 400 m. Figure 4.12a presents the velocity model obtained using the open-source software `py-GIMLI` (Rücker, C. and Günther, T. and Wagner, F. M., 2017). To obtain the model, we picked the first breaks every tenth shot. The initial model used for the inversion process was a triangular gradient model, with a maximum element size of 50 m^2 (corresponding to the square of the shot spacing). The initial velocity distribution ranged from 1.43 km/s at the top to 3.5 km/s at the bottom. The inversion was conducted up to a depth of 400 meters.

Identifying the shallow subsurface layers is difficult for this survey using first-break tomography because the refractions of the frozen layers hide refractions from the unfrozen layers. In addition, the relatively short maximum offset of the survey (1572.5 m) limits the TTFBT resolution. Moreover, TTFBT cannot detect the base of the permafrost as no critical refraction occurs from a velocity inversion.

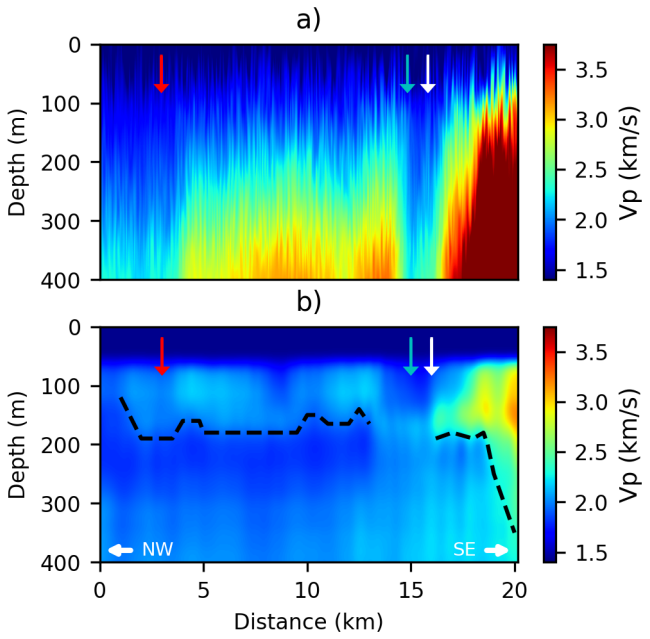


Figure 4.12 Comparison between results from travel time first break tomography and the multi-input NN evaluated over the last 20 km of line ARA05C-6. The red, cyan and white arrows mark the distance at which changes in the trend velocities are observed. The dashed black line shows the interpreted base of the permafrost.

Figure 4.12 shows that both models present a similar lateral distribution of permafrost. From the NW-end of the line to 3 km (red arrow), the TTFBT and the NN both indicate velocities associated with low to no-ice content. Between 3 and 15 km (cyan arrow), the TTFBT method and the NN indicate velocities associated with partially ice-bonded permafrost. However, the TTFBT draws a region with continuous ice saturation, while our NN suggests the presence of discontinuous partially-ice bonded permafrost. In this case, our results from the NN better agree with previous studies in this area of the Beaufort Sea, where permafrost is interpreted from seismic refraction velocities to be discontinuous and partially ice-bonded towards the shelf edge (Pullan et al., 1987). Near the SE-end of the seismic line, both methods retrieve seismic velocities that abruptly change around 15 and 16 km (white arrow). The seismic velocities in both models increase again from 16 km toward the SE-end of the seismic line indicating higher ice saturation in the zone.

The most significant difference between both models is the vertical velocity trend. Although the top of the permafrost is found at approximately the same depths, velocities in the TTFBT model increase with depth, while the NN predictions show a low velocity zone below the high-velocity permafrost layer.

We can explain the discrepancies between the two models with the nature of the TTFBT. For instance, TTFBT fails to retrieve information under the permafrost layer as no refractions exist at the interface from ice-bearing to unfrozen sediments. On the other hand, the NN can use all types of events contained in the data. Note also that the smooth initial model used for TTFBT greatly impacts the vertical trend found in the results, making it susceptible to find a local minimum which could differ from the true model. These variations are thus not fully reliable in the final model.

In general, the multi-input NN performs well for inverting velocity variations. It works relatively well for inverting velocities in the first 300 m. At those depths, the NN predicts velocities associated with a discontinuous layer of permafrost. The ice content conditions of that layer vary from partially ice-bonded permafrost (2.2 km/s$< V_P >$3.0 km/s) to the left to ice-bonded permafrost ($V_P >$2.5 km/s) to the right. The NN can accurately display discontinuities in the permafrost zones and provide an approximated depth of the base of the permafrost (dashed black lines). Less accurate information is obtained beneath the base of the permafrost, where 1D artifacts are mostly noticed. This result is similar to what was observed when statistically evaluating the NN in the 1D testing dataset; i.e. higher standard deviation values appear after the first velocity reversal, which in turn corresponds to the base of the permafrost.

Despite these artifacts, mapping permafrost is therefore possible with the proposed NN.

The NN provides information about the lateral and vertical distribution of permafrost. It can also provide information about the permafrost conditions by comparing velocity values with benchmark measurements. Deep learning also significantly reduces the preprocessing steps required to obtain a reliable velocity model, which can be problematic in conventional seismic processing when subsea permafrost is present. The preprocessing for the neural network only involves a bandpass filter of the shot gathers followed by data transformation to the different domains. In any case, the computing cost of the NN approach is relatively high, mostly due to the training phase. Using the capacities of the Beluga cluster (`https://docs.alliancecan.ca/wiki/B%c3%a9luga/en`) from the Digital Research Alliance of Canada, we trained each of the 64 parallel NN in about 8 hours using one Nvidia V100-SXM2 GPU per NN. However, once trained, evaluating the models is cheap and takes less than 2 minutes per NN.

4.5 Conclusion

In this paper, we evaluated the ability of deep learning to map velocity and attenuation variations of subsea permafrost. We designed a multi-input NN that uses different domains and transfer learning to improve the mapping of the velocity variations. Our approach limits the preprocessing step required to separate the different arrivals before obtaining a final velocity model. Instead, the NN automatically identify seismic signatures from the different domains and relates them to the corresponding seismic velocities and attenuation factors.

We used our multi-input NN on a 2-D seismic dataset collected in the Beaufort Sea. The resulting model not only agrees with horizontal permafrost distribution mapped by previous studies in the region, but also provides a vertical estimate which was impossible using other techniques such as TTFBT or refraction analysis.

Our results show that multi-input NN is a reliable tool for estimating the seismic properties of subsea permafrost from seismic data. Although the computational cost of our approach is relatively high due to the training phase, it still remains much cheaper than state-of-the-art FWI. Furthermore, once trained, the NN's evaluation cost is relatively low. This low computational cost facilitates the evaluation of the NN for different seismic datasets. In fact, extensive seismic surveys are available in the Beaufort Sea area, and our NN opens the door to large-scale mapping of subsea permafrost at an unprecedented resolution and low cost. Further work could also lead to using the inverted models as an initial velocity model for an FWI methodology.

Finally, besides subsea permafrost environments, our method could be applied to many other

contexts where guided waves and high-amplitude multiples overprint near-surface reflections, such as carbonates in shallow water.

4.6 Acknowledgements

The authors would like to thank the Korea Polar Research Institute (KOPRI) and the members of the ARAC05C expedition for providing access to the recorded data. This research was enabled in part by support provided by the Digital Research Alliance of Canada (https://alliancecan.ca/).

4.7 Addendum A: Cross talk analysis between the parameters evaluated

One of the primary concerns of implementing the NN approach is dealing with non-uniqueness and parameter crosstalk. Non-uniqueness occurs when several sets of parameters can lead to the same result, while parameter crosstalk happens when errors in one model parameter are attributed to errors in the estimate of another parameter (Keating and Innanen, 2019). Indeed, the interplay of various factors such as depth, thickness, velocity, and attenuation of the permafrost layer can lead to different combinations of these parameters, resulting in similar responses in the four domains employed in the inversion scheme. This interplay between variables can be analyzed using correlation matrices. The correlation matrix (r_{jk}) reveals the relationships between variables, highlighting non-uniqueness and crosstalk. It can be calculated using the formula (Aster et al., 2018):

$$r_{jk} = \frac{s_{jk}}{\sqrt{s_{jj} s_{kk}}} \tag{4.6}$$

where s_{jk} is the covariance between parameters j and k. Assuming independent and identically distributed data errors, the covariance matrix can be calculated as follows:

$$s(m) = \sigma^2 (J^T J)^{-1} \tag{4.7}$$

here, J is the Jacobian with respect to the evaluated parameters, and σ is the standard deviation.

Figure 4.13 depicts the correlation matrix around the fixed parameters used to model shot 8275. A strong correlation exists between velocity and thickness across the four domains. Likewise, there is a significant correlation between depth and velocity in the $t-x$ (shotgather), $\tau - p$ (radon) and the $f - p$ (frequency radon) domains, between thickness and attenuation within the $t - x$ and $\tau - p$ domains, as well as between velocity and attenuation within the

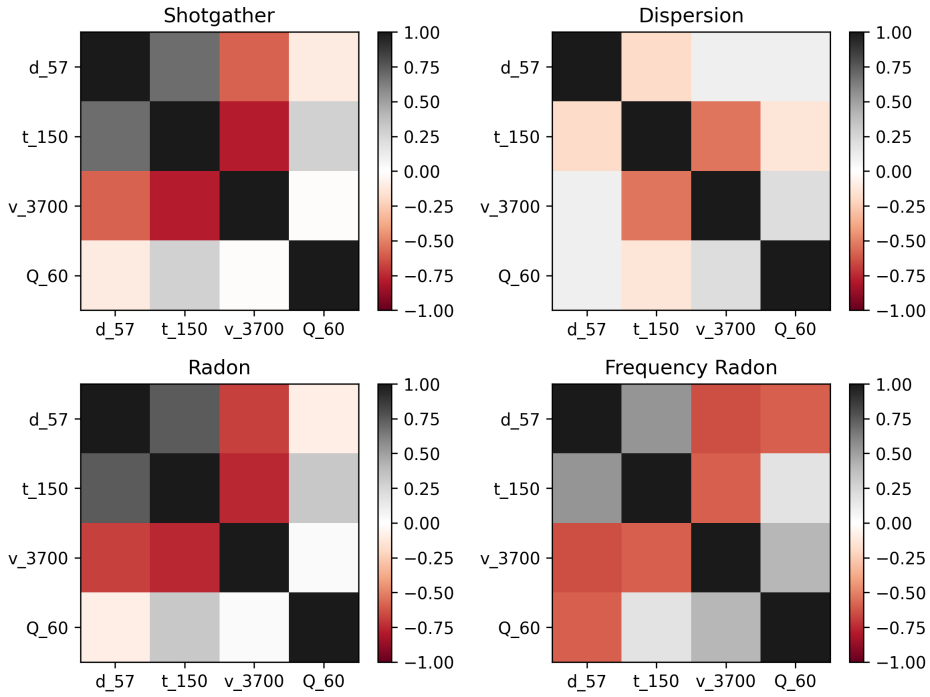


Figure 4.13 Correlation matrix in single domains for modeled shot 8275 with a permafrost layer at 57 m depth, of 150 m thickness, a velocity of 3.7 km/s and a Q factor of 60.

$f-c$ (dispersion) and $f-p$ domains. Such pronounced correlations among these parameters indicate a higher likelihood of encountering crosstalk and non-uniqueness phenomena.

However, evaluating the correlation matrix using the combined information from the four domains results in significantly reduced correlation values. Figure 4.14 presents the correlation matrix using the combined information from the four domains. Although the correlation values have reduced significantly, some high correlations could still lead to non-uniqueness and parameter crosstalk. For instance, after combining the four domains, a significant correlation remains between velocity and thickness and velocity and depth. However, a lower correlation is observed between thickness and depth and depth and attenuation.

Non-uniqueness and crosstalk represent a common challenge encountered in various inversion methodologies. One example of this is observed in the context of FWI (Keating and Innanen, 2019; Operto et al., 2013), where a typical strategy involves the sequential inversion of low-frequency data before gradually incorporating higher frequencies. This approach effectively mitigates the impact of nonuniqueness arising because of cycle skipping (e.g., matching no corresponding cycles of the wave arrival).

In our NN approach, we implemented various parameter constraints to address the issue of

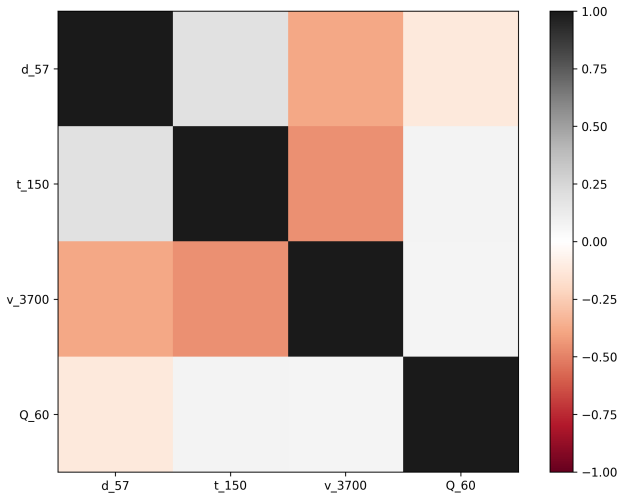


Figure 4.14 Correlation matrix using multiple domains for modeled shot 8275 with a permafrost layer at 57 m depth, of 150 m thickness, a velocity of 3.7 km/s and a Q factor of 60.

non-uniqueness in our solution. Initially, we limited the seismic parameters (V_P, V_S, Q, and density) in the training dataset to values obtained from laboratory measurements on samples that are known to represent permafrost conditions in the Beaufort Sea, as documented in the literature.

Furthermore, we imposed constraints on the depth, thickness, and velocities of the subsurface layers, utilizing data from well logs whenever available. Additionally, to further enhance the precision of our models, we restricted the occurrence of permafrost to the region above the 100 m isobath, ensuring that thicknesses did not exceed 700 m, following the recommendations of Taylor et al. (2013).

Moreover, based on the progressive increase in frequency content commonly used in FWI, we generated distinct subsets with decreasing layer thicknesses in the dataset. This approach aimed to allow the NN to learn features from seismic data with less complexity, involving thicker layers, as well as more complex data with thinner layers, leading to stronger interference between different arrivals.

4.8 Addendum B: Using single input NNs

A question that could arise from chapter 5, is related to the capability of single-input NNs to invert subsea permafrost distribution accurately. To address this question, we developed four distinct NNs, each exclusively utilizing one of the encoders from our proposed multi-input approach. The number of decoders was consistently set to three, designed explicitly

for inverting velocities and attenuation.

Essentially, each NN is tasked with processing a seismic image within a single domain: time-offset, frequency-phase velocity, time-velocity, or frequency velocity. Consequently, the output of each NN are V_P, V_S, and $1/Q$. To ensure robustness and precision, we trained each NN 16 times using the 1D dataset and returned the average prediction. No TL methodology was applied in this test.

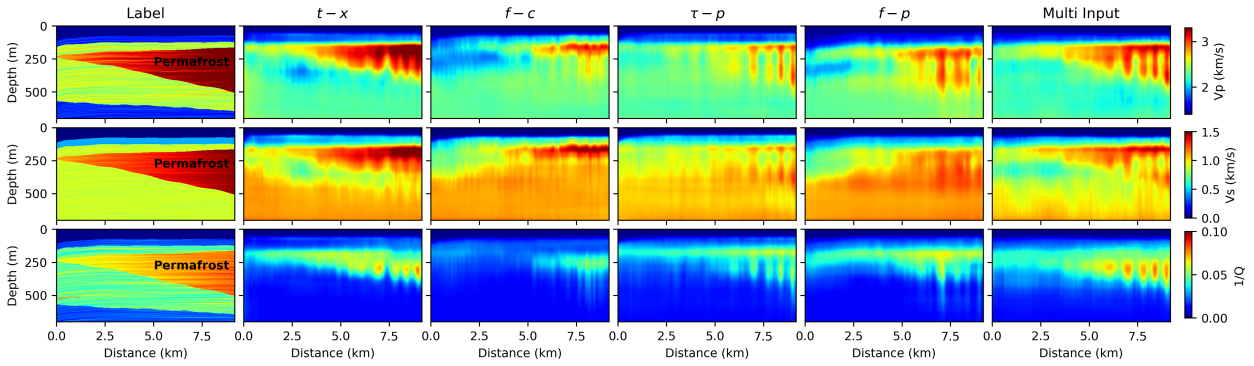


Figure 4.15 Evaluating single input NNs on a synthetic dataset from a wedge subsea permafrost structure. The columns display the label and inverted models using NNs training on the $t-x$, $f-c$, $\tau-p$ and $f-c$. The rows show the inverted V_P, V_S and $1/Q$ values. The last column represents the average models of all the inversions.

We evaluated the NNs on the synthetic wedge model discussed in chapter 4. In Figure 4.15, we present the average predictions of four distinct NNs, each trained on seismic data from four separated domains. The results indicate that training NNs in different domains emphasizes distinct permafrost characteristics.

Notably, employing $t-x$ inputs improves the retrieval of higher velocities. On the other hand, using $f-c$ and $f-p$ inputs offer superior reconstructions of the top of the permafrost and the bathymetry, respectively. Moreover, for the effective representation of the permafrost base, using $\tau-p$ and $f-p$ inputs proves to be more advantageous.

The reason behind the success of each type of input for retrieving distinct permafrost characteristics can be explained by the nature of the different transformations applied to the data. For example, the $t-x$ domain performs better at retrieving high velocities because it contains all the information from the different arrivals. When refracted waves are discernible, using this domain makes it easy to estimate seismic velocities. However, the high interference between arrivals can make it challenging to detect the permafrost base. On the other hand, the $f-c$ and $f-p$ domains display a modal distribution of the arrivals, which is mainly affected by the depth and thickness of the permafrost layer. Therefore, these domains better

provide information about the top of the permafrost and bathymetry. Lastly, the hyperbolic Radon domains ($\tau - p$ and $f - p$) look for information about reflections and multiples following hyperbole defined by distinct velocities. This discretization may favour the separation between arrivals, making it easier to identify the permafrost in this domain compared to other domains.

	Test Dataset						Wedge Model					
	L1 error (%)			L2 error (%)			L1 error (%)			L2 error (%)		
	Vp	Vs	1/Q	Vp	Vs	1/Q	Vp	Vs	1/Q	Vp	Vs	1/Q
t-x	4.28	4.12	5.37	6.25	5.78	8.01	10.02	6.38	12.56	12.38	7.79	15.14
f-c	6.74	5.33	6.17	9.93	7.55	9.13	11.82	7.01	15.11	14.65	7.89	17.90
tau-p	4.19	3.92	5.18	6.11	5.54	7.79	9.99	6.31	13.21	12.31	7.57	15.86
f-p	5.81	4.82	5.75	8.51	6.86	8.62	9.34	6.60	12.67	11.90	7.45	15.33
Average	5.26	4.55	5.62	7.70	6.43	8.39	10.29	6.58	13.39	12.81	7.68	16.06
Multi-input	3.97	3.73	5.13	5.84	5.32	7.83	9.57	5.45	11.56	11.34	6.83	13.82

Table 4.5 L1 and L2 norm error for the single-input and the multi-input NN evaluated in the 1D dataset and the 2D wedge model. Green cells display lower errors per column while the red cells show higher ones.

In general, using single-input neural networks leads to an estimation of the overall distribution of the inverted properties. Nevertheless, the results from the multi-input NN surpass those obtained from the single-input networks. The findings are summarized in Table 4.5, which presents the L1 and L2 norm errors evaluated on both the 1D testing dataset and the 2D wedge model. Notably, the multi-input approach consistently yields the lowest errors among the evaluated models. There are only two exceptions: the L2 norm for the 1/Q inversion from the $\tau - p$ domain on the 1D dataset and the L1 norm for the inverted V_P velocities for the $f - p$ domain in the wedge model. However, even in these cases, the error obtained from the multi-input neural network is remarkably close to the best estimation achieved.

CHAPTER 5 ARTICLE 3: MAPPING SUBSEA PERMAFROST DISTRIBUTION IN THE BEAUFORT SEA WITH MARINE SEISMIC AND DEEP LEARNING

Bustamante, J., Fabien-Ouellet, G., Duchesne, M.J. and Ibrahim, A., 2023. Mapping subsea permafrost distribution in the Beaufort Sea with marine seismic and deep learning. Solid Earth, *Submitted*

Initial submission on October 13, 2023.

5.1 Abstract

Subsea permafrost is widespread across the polar continental shelf, and mapping its distribution is crucial for assessing its impact on global warming. While seismic techniques have been useful in determining the lateral extent of subsea permafrost in the Canadian Beaufort Sea, they fall short when it comes to evaluating its vertical variation and large-scale mapping. However, recent advances in deep learning inversion have enabled the generation of more reliable and accurate large-scale, depth-resolved velocity and attenuation models. In this study, we employ a multi-input, multi-output neural network to evaluate the distribution of subsea permafrost at a regional scale in the Canadian Beaufort Sea. Our methodology involves velocity and attenuation data derived from 15 seismic lines. Our results reveal a gradual reduction in ice content and bonding effects towards the shelf. In particular, we observe discontinuous subsea permafrost volumes with ice-bonding conditions at the centre of the volumes and ice-bearing conditions at the edges. The subsea permafrost volumes become sparser and smaller toward the shelf until no permafrost conditions are discernible. Furthermore, our findings agree with well logs tied with the survey lines and support previous studies on subsea permafrost distribution in the region. We believe our method resolves the entire thickness of the permafrost with greater accuracy than previous studies and provides a practical framework for regional-scale detection of various permafrost conditions. Our work opens the way to a large-scale automated reprocessing of existing Arctic seismic data that could help refine the distribution of subsea permafrost across the continental shelf.

Plain Language Summary

Our study focuses on mapping the distribution of subsea permafrost in the Canadian Beaufort Sea. We used a multi-input, multi-output neural network to evaluate the permafrost con-

ditions at a regional scale. Our findings show that the subsea permafrost volumes become sparser and smaller towards the shelf until no permafrost conditions are discernible. Our method has provided a practical framework for large-scale detection of various permafrost conditions, and we believe that it resolves the entire thickness of the permafrost with greater accuracy than previous studies. Our work could help refine the distribution of subsea permafrost across the continental shelf and open the possibility for automated reprocessing of existing seismic data collected on the Arctic.

5.2 Introduction

Subsea permafrost, defined as frozen soil beneath the seafloor experiencing temperatures below 0^oC for a minimum of two consecutive years (Osterkamp, 2001), is a prevalent feature beneath the continental shelves of polar regions. The significance of subsea permafrost monitoring has gained substantial attention within the scientific community due to its impact on global warming (Cheverda et al., 2022; Creel et al., 2023; Sayedi et al., 2020; Shakhova et al., 2017a). One crucial aspect is the role of subsea permafrost as a methane reservoir, providing conducive conditions for the formation and stability of gas hydrates within and beneath the frozen ground (Dallimore and Collett, 1995; Dallimore et al., 2015). Furthermore, it acts as an effective barrier, sealing off other methane reservoirs at depth. The gradual degradation of subsea permafrost is evident in the Arctic, where ocean temperatures remain slightly above the freezing point (Romanovskii et al., 2005). However, the escalating global temperatures, particularly in the Arctic regions (Meredith et al., 2019), can accelerate permafrost degradation, both on land and at sea. This acceleration can release additional greenhouse gases stored within and below the permafrost into the atmosphere, thereby initiating a feedback loop amplifying global warming (Anthony et al., 2012; Schuur et al., 2008).

In the Beaufort Sea, subsea permafrost initially formed at the surface when the shelf became exposed during glaciations (Taylor et al., 2013). As the last glacial period concluded, this ancient permafrost was gradually submerged due to rising sea levels caused by the melting of glaciers. Currently, subsea permafrost in the Beaufort Sea can be found directly beneath the ocean floor and extend considerably within the shelf. It is limited by the upper continental slope and reaches a maximum depth of 700 meters.

The distribution of subsea permafrost in the Beaufort Sea exhibits considerable variability, which can be attributed to multiple factors, including the complex historical patterns of transgressions and regressions, the discharge and influence of the Mackenzie River, and the impacts of ice cover (Osterkamp, 2001). For instance, unfrozen volumes of sediment, commonly referred to as taliks, can occur near atypical environmental heat sources (Frederick

and Buffett, 2014). Taliks result from localized variations in thermal, hydrological, hydrogeological, or hydrochemical conditions. On the surface, these taliks can form beneath bodies of water that do not completely freeze during the winter or near geological faults where heat flux is elevated (Romanovskii and Hubberten, 2001). While substantial research has been undertaken to explain the implications of talik formation on onshore permafrost dynamics (O'Neill et al., 2020), their occurrence and evolutionary processes in marine settings have remained relatively unexplored. Still, the presence of taliks could help to explain high seawater methane concentration (Kvenvolden et al., 1993) and the presence of gas plumes observed in the Beaufort Sea (Paull et al., 2011) and its possible connection with gas hydrates (Frederick and Buffett, 2014).

In order to investigate the implications of permafrost degradation in the Beaufort Sea, it is crucial to understand the characteristics of permafrost sediments and our ability to detect them. Although several studies have made efforts to map the extent of subsea permafrost in the Beaufort Sea (Brothers et al., 2016; Hu et al., 2013; Pullan et al., 1987), no publication has determined the horizontal and vertical distribution of subsea permafrost in this region.

Given that the presence of ice within the pore space of the sediments significantly increases the velocity of seismic waves (Dou, 2015; King, 1984), seismic methods are the most used tool for determining the lateral extent of subsea permafrost (Hunter et al., 1976; Kang et al., 2021). In particular, various seismic methods have been applied in the Beaufort Sea to assess the lateral extent of subsea permafrost. For example, Brothers et al. (2012a) employed seismic refraction analysis to map the lateral distribution of subsea permafrost in the US Beaufort Sea. Similarly, Brothers et al. (2016) used root mean squared (RMS) velocities obtained by semblance analysis to map subsea permafrost distribution. In the Canadian Beaufort Sea, Hunter et al. (1976) used refraction velocity analysis to estimate ice content distribution, while Kang et al. (2021) used full waveform inversion (FWI) in the Laplace domain to invert seismic velocities and permafrost distribution.

Pullan et al. (1987) provided the most extensive map of permafrost occurrence in the Beaufort Sea. To generate the map, they relied on seismic refraction surveys obtained along a regular grid covering much of the Canadian Beaufort Sea. Their analysis was based on apparent refraction velocity. Extensive public domain, as well as industry-owned seismic data, exist in the polar region. Although their acquisition was not designed for permafrost imaging, they can be reprocessed to detect and characterize permafrost. However, standard reflection and refraction processing are generally not adapted for imaging subsea permafrost.

The first limitation comes from the transition from ice-bearing to ice-free sediments at the base of the permafrost. This represents a velocity reversal that cannot be detected with

seismic refraction analysis as no refraction occurs in the interface. A second limitation comes from the presence of a waveguide formed by the high impedance contrast between the air, the water column, and the frozen sediment (Bustamante et al., 2023a). The propagation of high-amplitude waves along the waveguide interferes with near-surface reflections, which exhibit weaker amplitudes, making their identification and separation challenging (Duchesne et al., 2022). This complex interference among reflections, refractions, guided waves, and multiples poses a significant obstacle for classical inversion techniques (e.g., pre-stack and post-stack inversion) in accurately estimating permafrost properties.

Deep learning has become a promising alternative for processing and inverting seismic data in recent years. The pioneering work of Araya-Polo et al. (2018) used neural networks (NN) for velocity model building (VMB), where they employed 3D semblance images to create the models. Building upon this progress, Alfarraj and AlRegib (2019) designed a recurrent neural network (NN) to conduct elastic impedance inversion using multi-angle seismic data. In a separate study, Mustafa et al. (2021) introduced a bidirectional transfer learning approach to retrieve impedance. Generative adversarial networks (GAN) have also been effectively used for inverting velocity and impedance models in seismic data analysis (Marano et al., 2023; Meng et al., 2021; Wang et al., 2022; Wu et al., 2021; Zhang and Lin, 2020). A more comprehensive review of the emerging NN approaches for seismic inversion is presented in Adler et al. (2021). Mousavi and Beroza (2022) delivers an extensive overview discussing the trends, challenges, and potential advancements associated with the application of deep learning in seismology. This exhaustive analysis is based on a meticulous examination of more than 637 papers published between 1988 and 2022.

Bustamante et al. (2023b) developed a multi-input multi-output neural network with a specific focus on inverting seismic velocities and attenuation for assessing the distribution of subsea permafrost in the Canadian Beaufort Sea. In their study, they build upon three previous studies. Firstly, inspired by the findings of Fabien-Ouellet et al. (2020), who demonstrate that a neural network trained on synthetic data can effectively adapt to real-world scenarios and provide accurate results when the synthetic data closely resembles the target data. Secondly, the work of Bustamante et al. (2023b) was influenced by the results of Simon et al. (2023), who employ transfer learning techniques to train a neural network capable of inverting velocity models in both the time and depth domains. Through transfer learning, they enhance the performance and accuracy of the neural network in inverting velocity models while reducing the cost of training. Lastly, they incorporate the findings of Bustamante et al. (2023a), who emphasize the importance of evaluating shot gathers in different domains to improve inversion results in permafrost zones. The NN proposed by Bustamante et al. (2023b) is trained on synthetic data, incorporates the combined knowledge from the shot

gathers in multiple domains and implements a transfer learning methodology for increasing the capability of the NN to invert seismic velocity and attenuation in permafrost zones.

This study aims to assess the distribution of subsea permafrost in the Beaufort Sea. From the deep-learning methodology of Bustamante et al. (2023b), we extend our analysis beyond a single seismic line and apply it to two regional-scale seismic datasets collected in the Canadian Beaufort Sea. Our investigation focuses on a specific region located offshore the Northwest Territories in Canada, where the subsea permafrost undergoes a transition from ice-bonded to ice-free conditions. First, we present a description of the survey lines and the expected permafrost distribution. Second, we introduce the NN approach. Finally, we present the results from the inversion, discuss their limitations and compare the inverted permafrost distribution to previous studies in the study area.

5.3 Survey area and seismic dataset

5.3.1 Subsea permafrost in the Beaufort Sea

Background Geology

The Canadian Beaufort Sea is a portion of the Arctic Ocean bordering North America that lies off the Mackenzie Delta and Yukon (Fig. 5.1). The eastern part of the Beaufort Sea is characterized by a wide shelf of over 120 km, while in the western part of the Yukon coastline, it is only 10 km wide. A gentle slope (average of $5.7^o \times 10^{-3}$) characterizes the wide shelf, which transitions to a steeper continental slope at water depths ranging from 80 to 120 m beyond the shelf edge (Taylor et al., 2013).

The Beaufort shelf stratigraphy units include the Kugmallit sequences, which are Oligocene in age and related to the latest basin extension, and are the lowest stratigraphic units (Dixon et al., 2019; Taylor et al., 2013). Above it lies the Miocene Akpak and Mackenzie Bay sequences. Finally, the Quaternary/Pliocene Iperk, a sand, silt, and clay sequence, is the uppermost stratigraphic unit.

The deposits from the late Cretaceous to the Cenozoic era consist of organic-rich shale, deep-water submarine fan deposits, shelf shale and sandstone, and deltaic deposits of alternating sandstone, shale, and coal. The deltaic deposits include upper delta-plain, lower delta-plain, and delta-front deposits. Individual progradations can be identified within the transgression and regression, each showing an overall coarsening-upward succession of prodelta shale, gradually covered by delta-front sandstone, followed by delta-plain deposits, usually with erosional contacts on the underlying delta-front beds. This type of succession repeats throughout the

deltaic deposits of the sequences. The shelf/prodelta environment is mainly composed of silty to sandy shales, with occasional thin layers of sandstone (Dixon et al., 2019).

Subsea permafrost in the Beaufort Sea formed during glacial periods when the sea level was lower, and the continental shelf was subaerially exposed. The permafrost interval experienced several cycles of sea level changes characterized by glaciofluvial sand/estuarine-marine mud, respectively bounded by unconformities. Ice-bearing or ice-bonded sands and silts dominate the regressive intervals, and the most recent transgressive marine-estuarine muds were deposited with the Holocen sea level rise. Currently, subsea permafrost extends offshore from thick permafrost ($\leq$700 m) near the shore to thin or no permafrost at the shelf break limit ($\sim$100 m isobath) (Taylor et al., 2013).

Impact of ice on seismic properties

In permafrost sediments, the presence of ice in the pore space significantly impacts seismic velocities and attenuation. Under 0 ^{o}C, ice can be frozen, partially frozen, or completely unfrozen. As ice velocity ($V_P \sim 4$ km s^{-1} and $V_S \sim 1.84$ km s^{-1}) is much higher than the velocity of the unfrozen sediments (typically $V_P \sim 1.5$ - 2.4 km s^{-1} and $V_S \sim 0.3$ - 1.4 km s^{-1}), the velocity of the permafrost layer significantly increases.

Attenuation is a reduction in the energy of a traveling wave as it propagates through a medium (Robinson and Treitel, 2008). It is related to the quality factor Q, which allows us to express the amplitude A of the vibration as a function of time (t) and the natural frequency of the system (ω_0) following $A(t) = A_0 e^{-\omega_0 t/2Q}$ (Ammon et al., 2020). Note that attenuation increases with $1/Q$. Seismic attenuation is high in permafrost, especially when a mixture of water and ice is present in the pores. On the other hand, fully frozen and unfrozen sediments usually have a low seismic attenuation (Matsushima et al., 2016).

Seismic velocities and attenuation are influenced by ice saturation as well as the bonding properties between the pore space and matrix, which are determined by various factors such as pore fluid salinity, grain size, clay content, and multi-phase pore filling, among others (Dou, 2015). Velocity and attenuation changes can thus be used to detect and quantify the ice distribution of permafrost.

Table 5.1 provides a summary of seismic velocity and attenuation factor (for both P- and S-waves) measurements conducted on laboratory samples representative of permafrost conditions in the Beaufort Sea, as reported in previous literature (Bellefleur et al., 2007; Dou et al., 2016; King, 1984; Matsushima et al., 2016). It should be noted that the presence of ice in the material can result in higher seismic velocity values for coarse-grained materials

Table 5.1 Typical seismic properties found in the Canadian Beaufort Sea (Bellefleur et al., 2007; Buckingham, 1997; Dou et al., 2016; Matsushima et al., 2016).

Sediments	**Vp (km s^{-1})**	**Vs (km s^{-1})**	**Q**
Water	1.43	0	1000
Unfrozen Sediments	1.5-1.9	0.3-0.6	30-90
Partially Ice-Bonded Silts	2.0-2.9	0.7-1.1	10-30
Ice-Bonded Silts	3.0-3.6	1.6-2.0	40-70
Partially Ice-Bonded Sands	3.0-3.5	1.0-1.3	10-15
Ice-Bonded Sand	3.5-3.9	1.3-1.9	30-90
Unfrozen Shale	2.1-2.4	0.7-1.4	50-120
Ice-Bonded Shale	2.6-3.4	1.4-1.9	70-130

(e.g., sands), compared to fine-grained materials (such as silts), as suggested in previous studies (Dou, 2015; King et al., 1982). The lowest Q values were found in materials with intermediate ice content that were partially ice-bonded.

5.3.2 Survey lines

Seismic surveys ARA04C and ARA05C were collected in 2013 and 2014 onboard the ice-breaker RV Araon in the Beaufort Sea (Fig. 5.1). The surveys were conducted by a joint effort between the Korea Polar Research Institute (KOPRI) and the Geological Survey of Canada (GSC) and aimed to explore the geological conditions, subsea permafrost, and gas hydrates on the outer continental shelf and upper slope of the region offshore Yukon and Northwest Territories (Jin and Dallimore, 2016).

Our analysis focuses on the sections of the seismic lines beyond the 100 m isobath, representing the farthest extent of subsea permafrost (Taylor et al., 2013). The ARA04C survey collected 6 seismic lines in this region, with lines 2, 8, and 10 running along the shelf and lines 1, 9, and 11 crossing it. A 1200 in^3 G-gun array triggered every 93.75 m was used to conduct this survey. In contrast, the ARA05C survey consists of 9 seismic lines in the same region; lines 2, 7, 11, and 17 are running along the shelf, whereas lines 1, 3, 5, 6, 8, and 16 are perpendicular to it. The seismic source used in the ARA05C survey was the same as for the ARA04C survey, but it was fired every 50 m. Both surveys employed the same seismic streamer with 120 receivers spaced at 12.5 m intervals for a total streamer length of 1500 m. The streamer and airguns were maintained at a constant depth of 6 m. In the ARA04C survey, the minimum offset for lines 1 and 2 was set to 100 m, while for lines 8, 9, 10, and 11 the minimum offset was set to 110 m. For the ARA05C survey, the closest receiver had an 85 m offset for all lines.

Several hydrocarbon exploration wells have been drilled in the Canadian Beaufort Sea during

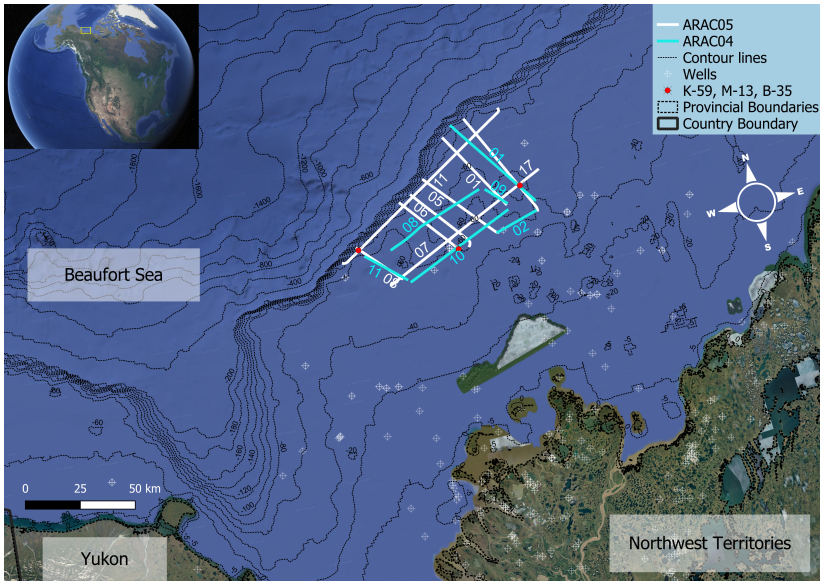


Figure 5.1 ARA05C survey lines and well locations. Red dots indicate the wells for which crystal cable survey is available

the 1970s and 1980s. Some logs recorded in those wells were available for our analysis. However, despite the considerable amount of well logs at our disposal, most are situated outside our research area or over 1 km away from the seismic lines. We restrict our analysis to wells at a 1 km radius from the seismic lines to reduce mismatches associated with the high heterogeneity associated with changes in ice content expected within the studied zone. Furthermore, only a few well logs provide information on the first 700 m depth, as the logging process typically starts below the first casing (i.e. between ∼300 and ∼500 m). We have identified three wells, namely Nektoralik K-59, Kopanoar M-13, and Irkaluk B-35 (red dots in Fig. 5.1), which possess surveys that can be used to validate our inversion results (indicated by red dots in Fig. 5.1). These wells were logged with crystal cables (Walker and Stuart, 1976) to determine the velocity of the near surface above the well casing.

5.4 Processing seismic data with a Neural Network

We use deep learning to process the seismic lines presented above. One notable advantage of employing deep learning techniques is that there is no need for a separate pre-processing step to segregate different seismic arrivals. Instead, these methods learn to accentuate the most relevant features for resolving a specific task during the training phase (Waldeland et al., 2018). Another distinct advantage over state-of-the-art techniques like FWI is that it eliminates the requirement for defining an initial velocity model.

Here, we use the NN described in Bustamante et al. (2023b). This NN is a multi-encoder multi-decoder NN that takes as input seismic images in four distinct domains and that outputs three seismic parameters: P- and S-wave velocities, (V_P, V_S), and the attenuation ($1/Q$). The architecture is inspired by the U-Net model, with two main components: an encoder and a decoder. The encoder is responsible for discerning significant features within the input data, generating a sparse representation known as the latent space. This latent space is optimized for the prediction task at hand. The decoder component uses this latent space to construct predictions and effectively maps the input space to the output space.

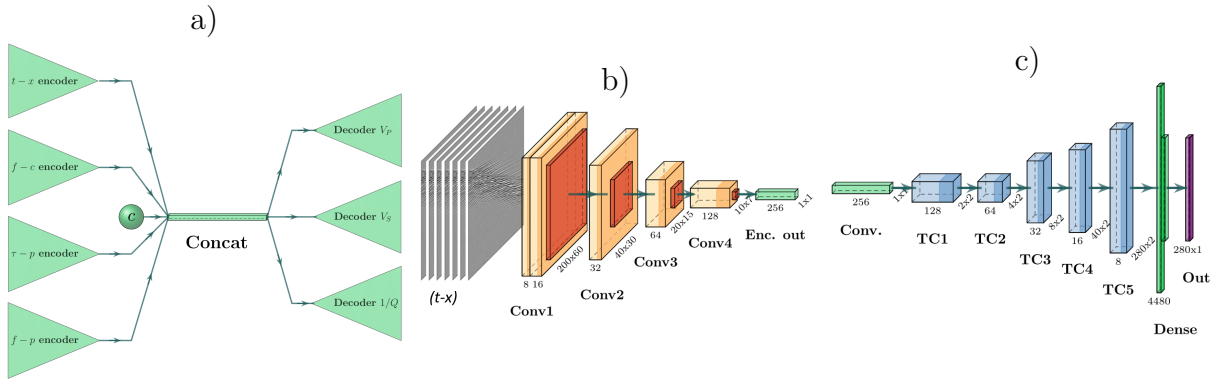


Figure 5.2 a) Summary of the multi-encoder multi-decoder approach as presented in Bustamante et al. (2023b). b) Architecture of the encoder that inputs multiple images in the CMP gather domain. c) Decoder architecture.

The decision to incorporate multiple inputs is grounded in the understanding that diverse seismic events can be effectively separated and enhanced through appropriate transformations into different domains. While NNs possess the potential to build the optimal input representation autonomously, our empirical investigations (Bustamante et al., 2023a) have revealed that the inclusion of carefully selected transforms (such as the time-offset ($t-x$), hyperbolic Radon ($\tau-p$), frequency phase velocity ($f-c$), and frequency velocity ($f-p$) domains) significantly enhances the capacity of the NN to generalize to real data, particularly when trained on synthetic data. Moreover, this approach provides valuable guidance to the network in utilizing distinct seismic events, including reflections, refractions, and guided waves, which is crucial for subsurface imaging below the permafrost table.

Figure 5.2a illustrates the network architecture, which consists of four encoders responsible for analyzing seismic images in four domains: $t-x$, $f-c$, $\tau-p$, and $f-p$, and three decoders that provide information on V_P, V_S, and $1/Q$. The configuration details of the network, including the input, hidden layers, and outputs, as well as the filter size in the encoders and decoders, are presented in Table 5.2. The network is fed by 11 common middle

points (CMP) gathers and their transform, centered on the position of the predicted depth profile. This corresponds to a radius of 250 m around the central point.

Table 5.2 Description of the hidden layers in the encoders and decoders of the NN as presented by (Bustamante et al., 2023b) and summarized in Fig. 5.2. The cells indicate the output dimension of the layer and the filter size (in parenthesis).

	$t-x$	$\tau-p$	$f-c$	$f-p$	
Input	1000x120x11	1000x200x11	200x200x11	200x200x11	
Conv2D 64	1000x120x64 (17x3)	1000x200x64 (17x3)	200x200x64 (17x3)	200x200x64 (17x3)	
Conv2D-MaxPooling	200x60x16 (17x3) 40x30x32 (17x3) 20x15x64 (9x3) 10x7x128 (5x3) 1x1x256 (3x3)	250x200x8 (17x3) 125x100x16 (17,3) 62x50x16 (17x3) 31x25x32 (9x3) 15x12x32 (5x3) 7x6x64 (3x3) 3x3x128 (3x3) 1x1x256 (3x3)	100x100x8 (17x3) 50x50x16 (17x3) 25x25x32 (9x3) 12x12x64 (5x3) 6x6x128 (3x3) 3x3x256 (3x3) 1x1x256 (3x3)	100x100x8 (17x3) 50x50x16 (17x3) 25x25x32 (9x3) 12x12x64 (5x3) 6x6x128 (3x3) 3x3x256 (3x3) 1x1x256 (3x3)	**Encoders**
Concatenate	1x1x1024				
Convol2D 256	1x1x256 (1x1)	1x1x256 (1x1)	1x1x256 (1x1)		**Decoders**
TConv2D	2x1x128 (3x1) 4x1x64 (5x1) 8x1x32 (9x1) 40x1x16 (17x1) 280x1x8 (17x1)	2x1x128 (3x1) 4x1x64 (5x1) 8x1x32 (9x1) 40x1x16 (17x1) 280x1x8 (17x1)	2x1x128 (3x1) 4x1x64 (5x1) 8x1x32 (9x1) 40x1x16 (17x1) 280x1x8 (17x1)		
Flatten-Dense	280	280	280		
	Vp	**Vs**	**1/Q**		

The encoder specifically designed for the $t-x$ domain is depicted in Fig. 5.2b. Each encoder sequentially reduces the size of the image and outputs a single vector. The four output vectors are concatenated, serving as the input of the decoders. Figure 5.2c shows the architecture of the decoders. We use three separate decoders, one for each output, namely V_P, V_S, and $1/Q$. The output represents the velocity and attenuation profile under the CMP location from the seafloor down to 700 m, the maximum depth at which subsea permafrost is expected (Taylor et al., 2013).

To train the NN, we generate synthetic datasets representing subsurface conditions observed in the Beaufort Sea in 1D and 2D. The synthetics were created by modifying the simplified tabular method proposed in Fabien-Ouellet et al. (2020) and Bustamante et al. (2023a). We defined random layered units consisting of a water layer, unfrozen sediments, permafrost, gas hydrates, and unfrozen sediments in that order. To introduce 2D characteristics, the layers were allowed to possess random slopes ranging from -15° to 15° and a random sinusoidal deformation of the layer with random amplitude and frequency. The training dataset consisted of 36,000 samples in the 1D dataset and 18,000 in the 2D dataset.

In order to mitigate the computational burden associated with training the NN, we employ a transfer learning (TL) approach introduced by Simon et al. (2023). This methodology, also known as curriculum learning (Soviany et al., 2022), involves training the network using a

series of sub-datasets with increasing complexity. TL operates on the principle of leveraging prior knowledge obtained from a specific task to enhance learning in a related task (Torrey and Shavlik, 2010). Specifically, we use the knowledge obtained from a pre-trained NN, trained on the 1D layered models, to train a comparable 2D NN. The 2D NN takes neighboring CMP gathers, centered on a particular location, as input and predicts 2D models of velocity and attenuation.

To enhance the accuracy in regions characterized by substantial uncertainties, a training approach employing an ensemble comprising 64 similar NNs was adopted (Opitz and Maclin, 1999; Simon et al., 2023). The NNs were initialized with random weight values and trained using the Adam optimizer (Kingma and Ba, 2014) with a learning rate set to 1e-4. Initially, a 1D dataset and a customized loss function were employed to train the NNs. The loss function was independently evaluated for inverting the seismic parameters V_P, V_S, and $1/Q$. The total loss was then computed as a weighted sum of the individual losses. The training process was concluded after 35 epochs when evidence of overfitting was first observed.

Subsequently, a second training phase was conducted employing the TL methodology to retrain the NNs using the 2D dataset. We maintained the same weights and biases learned in the 1D approach for every layer in the encoders and decoders of the 2D NN. However, the dimensions of the weights in the initial layers were adjusted to accommodate the increased size of the 2D training dataset. This retraining phase lasted for an additional 20 epochs until no further improvement was observed.

Figure 5.3(a) provides an overview of the loss function's progression after each epoch during the training sessions for both the 1D and 2D datasets. Notably, there was an increase in loss values upon transitioning from the 1D to the 2D dataset. This increase in loss values can be attributed to the heightened complexity of the latter dataset. The elevated complexity also explains the greater standard deviation and ultimate loss values for the 2D dataset when compared to those of the 1D dataset. Figures 5.3(b-c) present examples of predictions using 1D and 2D testing datasets. From these examples, we observe that the NN provide a smoother version of the inverted parameters. Although the NN generally provide accurate inversion values, as expected the higher mismatches are observed for the inverted $1/Q$. This is because attenuation depends on multiple factors including geometry divergence, conversion of energy into heat and other factor affecting amplitude, such as transmissivity losses and mode conversion, which are difficult to estimate using seismic data.

To process the recorded data with the trained NN, the ARA04C and ARA05C surveys were sorted into CMP gathers. Each CMP was then transformed into the four input domains. The CMP gathers were arranged into groups of 11 and represented the input of the 64

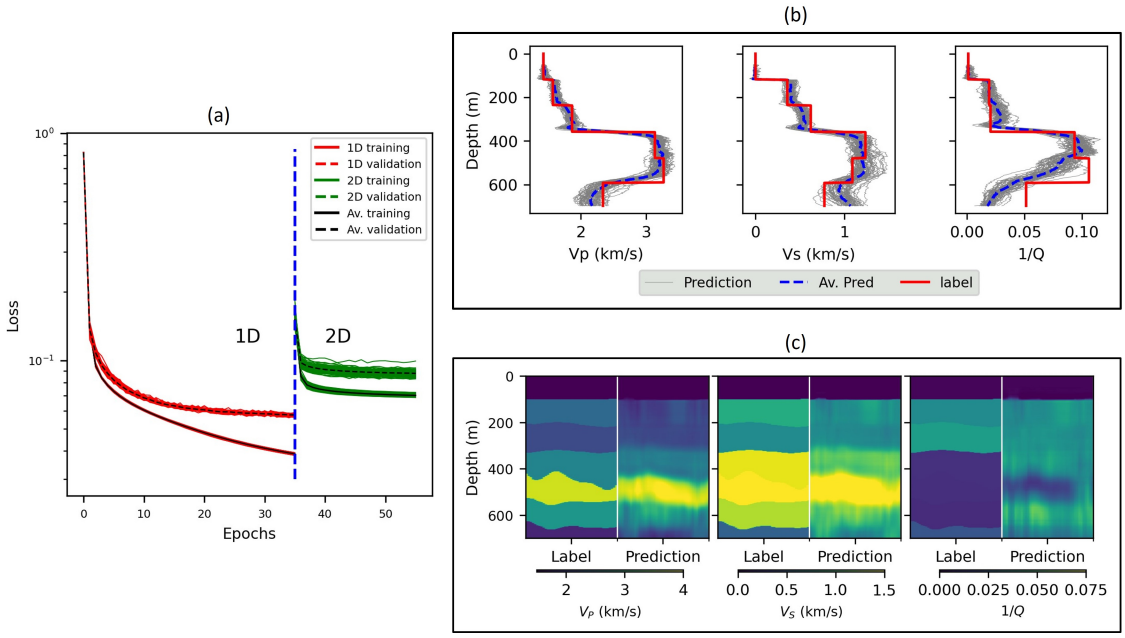


Figure 5.3 (a) Progression of the loss function during the training sessions for the 1D and 2D datasets. Processed and ground true (label) V_P, V_S and $1/Q$ on random samples of the 1D (b) and 2D testing datasets.

NNs contained in the ensemble. The 64 predictions were then averaged to obtain the final estimates of V_P, V_S, and $1/Q$. To produce 2D crosssections, the predictions of adjacent CMPs along the lines were concatenated.

5.5 Results

To evaluate the distribution of subsea permafrost in the area of interest, we processed nine seismic lines from the ARA05C survey and six seismic lines from the ARA04C survey. Some of these lines intersect with each other. All the lines cross areas where subsea permafrost is expected to range from ice-free or low ice content to discontinuous ice-bonded permafrost. In addition, velocity logs are available in wells Nektoralik K-59, Kopanoar M-13, and Irkaluk B-35 in the surveyed area.

5.5.1 Seismic inversion

Figure 5.4 presents a fence diagram of the distribution of inverted velocities and attenuation for all the lines in the ARA04C and ARA05C surveys. The figure shows a consistent agreement among the inverted velocities for all the seismic lines. Below the seafloor, V_P, V_S and $1/Q$ values associated with permafrost are observed in certain locations. Moreover, the V_P

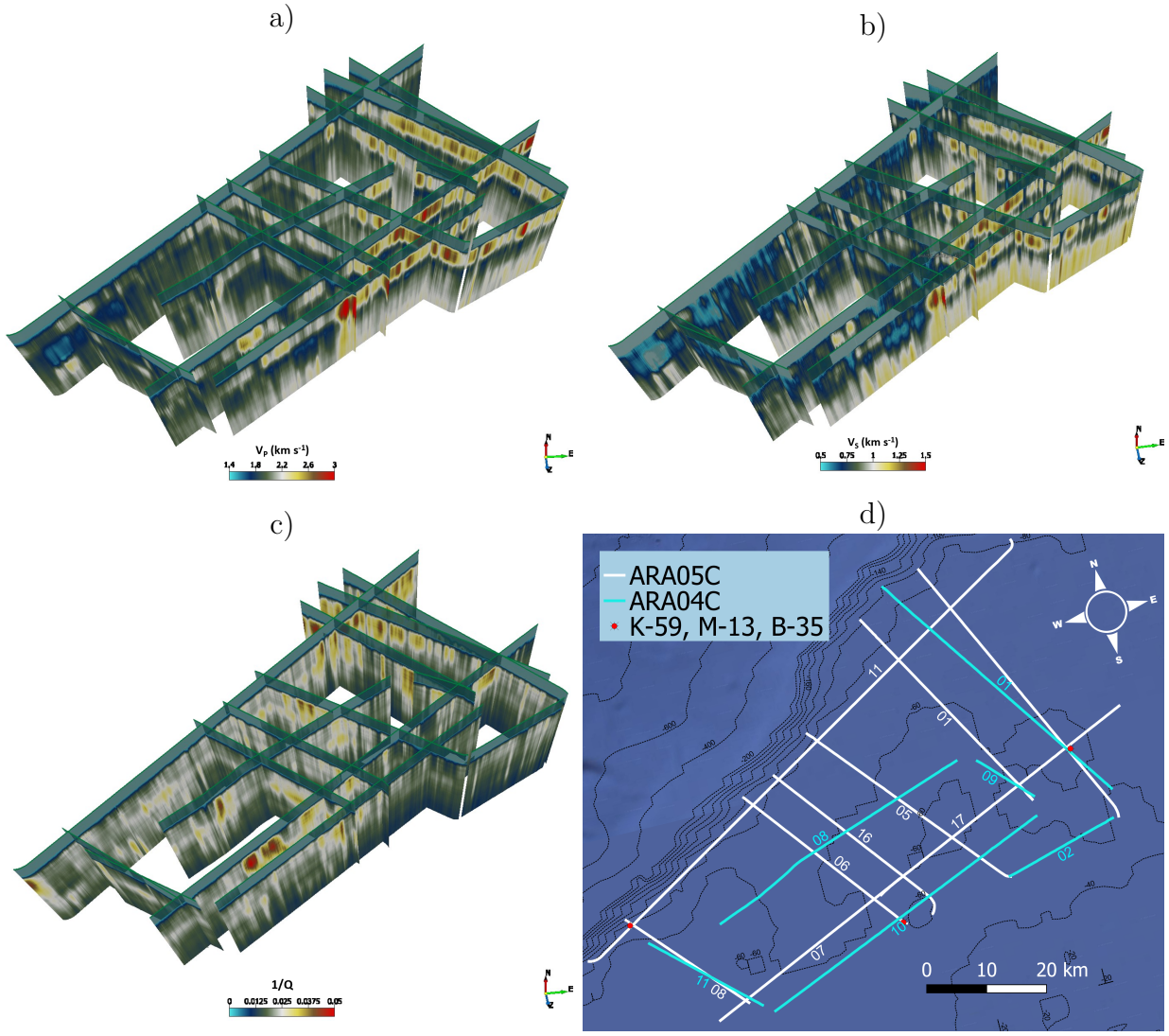


Figure 5.4 Fence diagrams displaying the inverted V_P (a), V_S (b) and $1/Q$ (c) values using the NN for all seismic lines. Line locations are presented in (d)

and V_S values gradually increase towards the coast (NE and SE sections on the map), suggesting increasing ice content toward the coast and corroborating earlier permafrost distribution maps in the region (Hu et al., 2013; Pullan et al., 1987).

We divided our analysis into three groups of seismic lines to better explore the results from the inverted V_P, V_S and $1/Q$ values. The first group comprises seismic lines that are parallel to the shelf and located in the northwest to central portion of the research area. The second group comprises seismic lines parallel to the shelf in the central to southeast portion. Finally, the third group consists of seismic lines orthogonal to the shelf.

Figure 5.5 shows the inverted seismic parameters for lines ARA05C-11, ARA04C-08, and

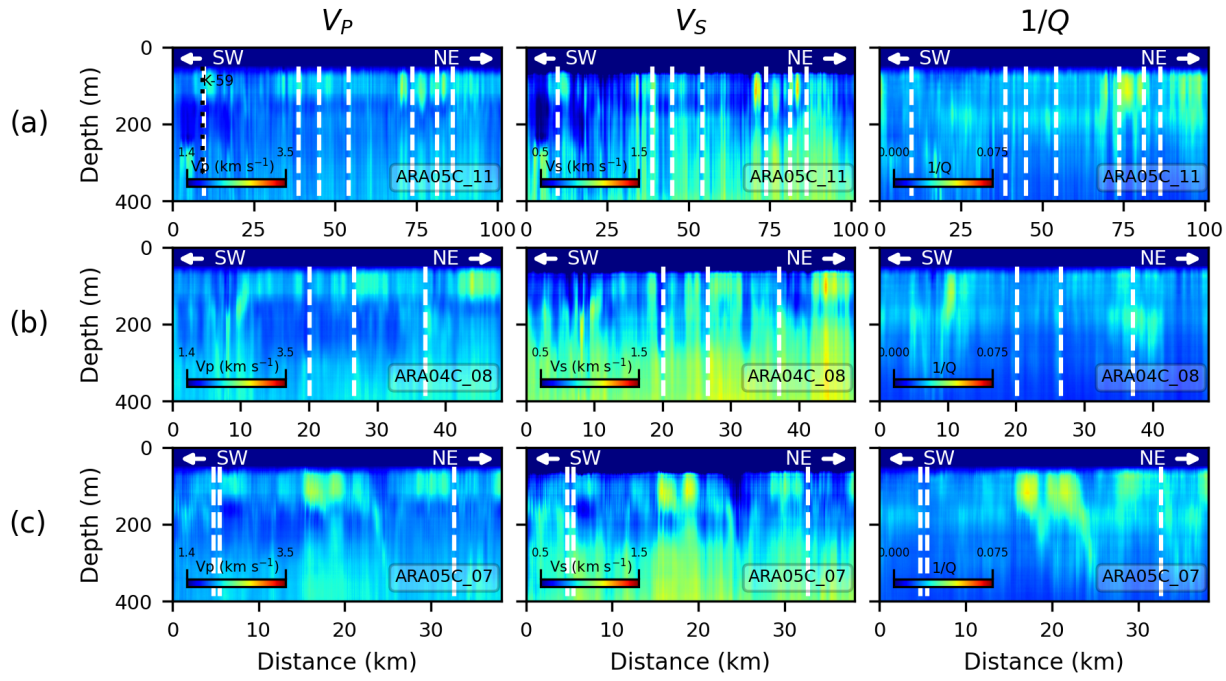


Figure 5.5 Inverted V_P (first columm), V_S (second column), and $1/Q$ (third column) for the seismic lines ARA05C-11 (a), ARA04C-04 (b), and ARA05C-07 (c). White dashed lines show the position of the intersections of the line with other seismic lines. Comparisons between inverted parameters at line intersections are presented in Fig. 5.8. Black dotted lines show the closest position (245 m) to well K-59 in (a), where a velocity log is available.

ARA05C-07, located in the northwest and central sections of the research area, that are parallel to the shelf. The first and second columns represent the inverted seismic velocities (V_P and V_S, respectively), while the third column displays the inverted attenuation parameter ($1/Q$). The seismic profiles in Fig. 5.5 illustrate seismic velocities and attenuation values corresponding to permafrost regions characterized by either low ice content or ice-free sediments ($V_P \leq 2.6$ km s^{-1}, $V_S \leq 1.1$ km s^{-1} and $1/Q \geq 2.52e^{-2}$ at sections). While the inverted parameters indicate low ice saturation, the continuity of ice-rich permafrost becomes more pronounced towards the northeastern segment. For instance, velocity and attenuation gradually increase towards the northeast in line ARA05C-11 at 70 km, line ARA04C-08 at 10 km and line ARA05C-07 at 14 km. While the increase in velocity and attenuation is gradual towards the coast for all the lines, there is a significant rise in inverted values within the 14-19 km range in line ARA05C-07. This observation suggests a higher ice saturation within this zone.

Moving toward the southeast, Fig. 5.6 illustrates seismic velocities and attenuation values associated with lines ARA05C-17 (a), ARA04C-10 (b), and ARA04C-02 (c). These lines

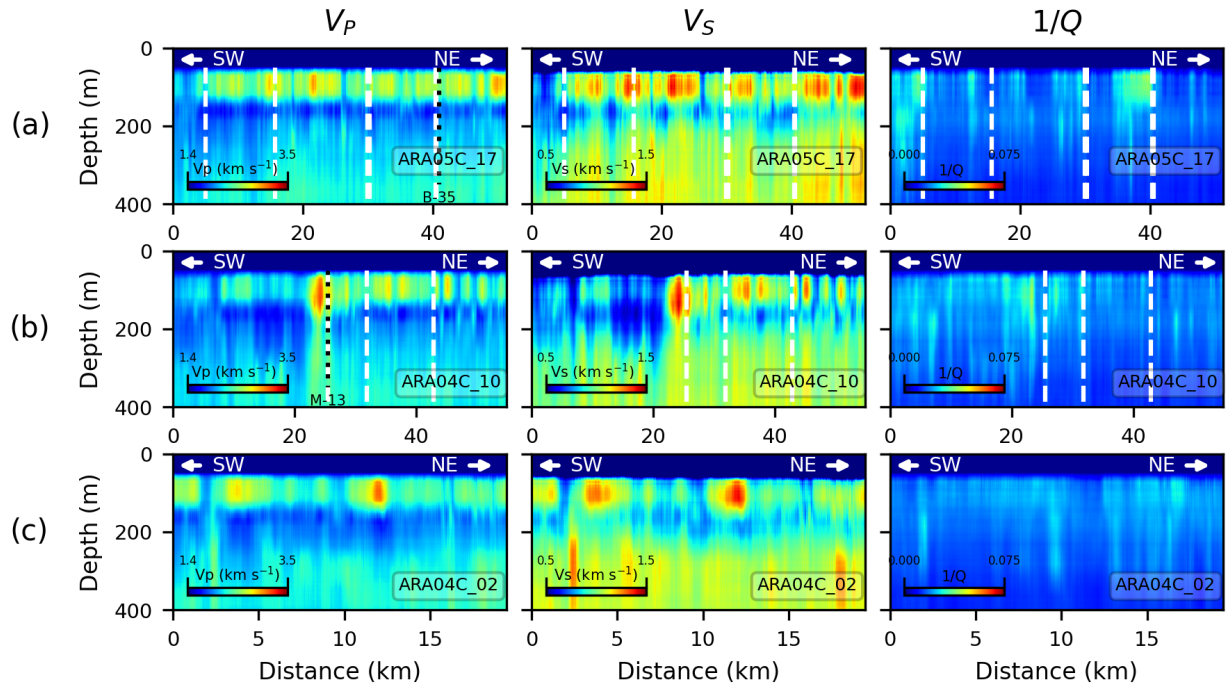


Figure 5.6 Inverted V_P (first columm), V_S (second column), and $1/Q$ (third column) for the seismic lines ARA05C-17 (a), ARA04C-10 (b), and ARA04C-02 (c). White dashed lines show the position of the intersections of the line with other seismic lines. Comparisons between inverted parameters at line intersections are presented in Fig. 5.8. Black dotted lines show the closest position (580 m and 975 m) from wells B-35 in (a) and M-13 in (b), where velocity logs are available.

are parallel to the shelf and are located in the central to southeast section of the research area. Figure 5.6 shows seismic velocity and attenuation values associated with permafrost transitioning from ice-bearing to ice-bonding conditions and from discontinuous to continuous conditions ($V_P \leq 3.5$ km s^{-1}, $V_S \leq 1.5$ km s^{-1} and $1/Q \geq 2.52e^{-2}$ at sections). The trend of increasing ice content and permafrost continuity towards the coast, observed in line ARA05C-07, remains consistent in line ARA05C-17. The inverted V_P, V_S, and $1/Q$ parameters correspond to ice-bonded permafrost transitioning from discontinuous to continuous permafrost across most surveyed areas. A nearly continuous permafrost layer is distinguishable from 5 km to the northeastern end of the ARA05C-17 line. Likewise, the inverted V_P, V_S, and $1/Q$ parameters for line ARA04C-10 also indicate a pattern of increasing ice content and permafrost continuity toward the coast. An almost continuous permafrost layer is also evident from 22 km to the northeastern end of the line. Lastly, the V_P, V_S, and $1/Q$ parameters suggest ice-bonded permafrost conditions for line ARA04C-02, with occasional lateral discontinuities.

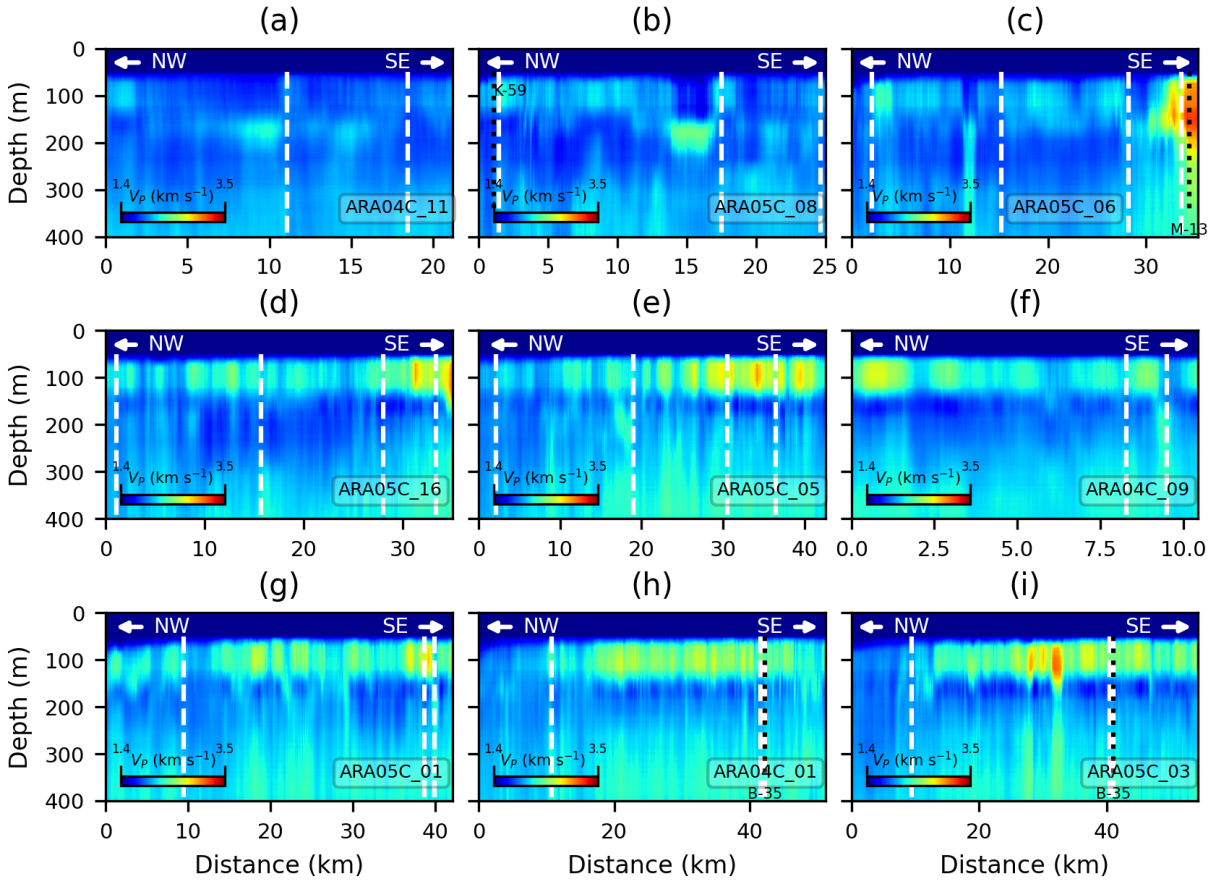


Figure 5.7 Inverted V_P for the seismic lines ARA04C-11 (a), ARA05C-08 (b), ARA05C-06 (c), ARA05C-16 (d), ARA05C-05 (e), ARA04C-09 (f), ARA05C-01 (g), ARA04C-01 (h), and ARA05C-03 (i). White dashed lines show the position of the intersections of the line with other seismic lines. Comparisons between inverted parameters at line intersections are presented in Fig. 5.8. Black dotted lines show the closest position (360 m, 75 m, 445 m and 440 m) to the wells K-59 (b), M-13 (c) and B-35 in (h, i), where velocity logs are available.

Figure 5.7 illustrates the V_P velocities for the seismic lines recorded perpendicular to the shelf. The first, second and third rows of Fig. 5.7 display the inverted velocities for the lines recorded in the southwestern section (ARA04C-11, ARA05C-08 and ARA05C-06), the centre section (ARA05C-16, ARA05C-05 and ARA04C-09), and the northeast section of the survey (ARA05C-01, ARA04C-01, ARA05C-03), respectively. For each row, the lines are arranged from SW to NE. Between adjacent lines, the inverted seismic velocities reveal a clear pattern of rising ice content and increasing continuity of the permafrost towards the northwest. Similarly, within each line, there is a trend of increasing ice content and permafrost continuity towards the southeast, indicating an augmentation in ice content and permafrost continuity towards the coastline. This trend is observed in the southwest portion of the

survey area, where lines ARA04C-11 (Fig. 5.7(a)), ARA05C-08 (Fig. 5.7(b)), and ARA05C-06 (Fig. 5.7(c)) exhibit a noticeable increase in seismic velocities when transitioning from one line to another. The velocities observed in ARA04C-11 suggest the presence of sediments ranging from ice-free to low ice content. For this line, we cannot identify a permafrost layer. The velocities in ARA05C-08 indicate a comparable ice-free to low ice content range. Nevertheless, within the ARA05C-08 line, a layer containing ice-bearing sediment becomes apparent, with a discernible discontinuity occurring around the 14-17 km mark. As we move closer to the central area of the survey, the seismic velocities observed along line ARA05C-06 suggest the presence of a discontinuous layer of ice-bearing permafrost that stretches from the northwestern end up to 30 km. Beyond this point, both the seismic velocities and attenuation experience a substantial change displaying values that align with what is commonly encountered in ice-bonded permafrost and reaching up to 350 m depth. This trend continues as we approach the southeastern end of the line.

In the central region of the survey area, seismic lines ARA05C-16, ARA05-05, and ARA04C-09 show changes in seismic velocities typically associated with permafrost, ranging from ice-bearing to ice-bonded conditions. These changes in seismic velocities may occur as the permafrost shifts from discontinuous to continuous, moving from the northwest to the southeast, as illustrated in Fig. 5.7(d), (e), and (f). For line ARA05C-16 (Fig. 5.7(d)), the gradual increase in seismic velocities suggests a transition from ice-bearing to ice-bonded permafrost as this layer becomes more continuous towards the southeastern end of the line. The increase in velocity also correlates with a thickening of the permafrost layer. Over a distance of 26 km, inverted velocities suggest the presence of continuous ice-bonded permafrost conditions until the southeastern end of the seismic line. Velocities on line ARA05C-05 (Fig. 5.7(e)) also display increasing values, improved continuity, and an apparent thickening of the permafrost layer as we move towards the southeast. From 24 km to the southeastern end of the line, inverted velocities again indicate values characteristic of continuous ice-bonded permafrost conditions. For ARA04C-09 (Fig. 5.7(f)), the velocities suppose the presence of a discontinuous permafrost layer that transitions from ice-bearing to ice-bonded conditions. Because of the limited length of the line, it is difficult to discern a clear trend in the inverted velocities within this layer.

In the northeastern section of the study area, the velocities derived from lines ARA05C-01 (Fig. 5.7(g)), ARA04C-01 (Fig. 5.7(h)), and ARA05C-03 (Fig. 5.7(i)) suggest the presence of a permafrost layer. This layer maintains a consistent pattern of increasing velocities, continuity, and apparent thickness as it extends towards the southeastern end of the lines. For example, a continuous permafrost layer can be identified starting at 12 km along line ARA05C-01, 18 km along line ARA04C-01, and 16 km along line ARA05C-03 towards the

southeastern end of the seismic profiles. The velocity data strongly suggest the existence of a permafrost layer containing ice, transitioning from ice-bearing to ice-bonded conditions.

Overall, the inverted seismic parameters presented in Figs. 5.5-5.7 show an evident variation towards the northeast and southeast sections of the study area, indicating an increase in ice content and permafrost continuity towards the coast. These findings are consistent with the permafrost distribution maps presented in previous studies (Hu et al., 2013; Pullan et al., 1987).

5.5.2 Line intersections

To evaluate the capability of our approach to process seismic parameters regardless of the recorded direction, we assessed the consistency of our results at the intersection point between the seismic lines recorded parallel and perpendicular to the shelf. Figure 5.8 displays the inverted V_P models of the CMPs at line intersections. The profiles are displayed in a grid replicating their relative position in the research area. In the figure, the columns represent the lines recorded perpendicular to the shelf, while the rows correspond to parallel lines. We only plotted the inverted models at the locations where a perpendicular line intersects a parallel line. Line ARA04C-09 is omitted because it only intersects line ARA05C-17 at a location very close to its intersection with line ARA05C-03.

In Fig. 5.8, V_P values agree at most intersection points, regardless of the line orientation. However, we observe a larger difference at the intersection between lines ARA05C-06 and ARA04C-10, i.e. the only location where the permafrost thickness differs between the parallel and perpendicular lines. We suspect this discrepancy could be attributed to the rapid increase in V_P close to the intersection point. More specifically, the intersection falls within a high-velocity zone in line ARA05C-06 (last dashed line in Fig. 5.7(c)). In contrast, in line ARA04C-10, the intersection is located in a zone transitioning from high velocities in the northwest to moderate velocities in the northeast (first dashed line in Fig. 5.6(b)).

In addition, we observe a general increasing trend of V_P values that can be associated with permafrost conditions, particularly ice content and bonding conditions. The increasing velocity values and apparent permafrost thickening towards the coast (NE and SE) are noticeable in lines ARA04-08 at the northwest-to-centre section of the surveyed area, ARA05C-07 and ARA05C-17 at the southwest-to-centre section of the survey area, ARA05C-06, ARA05C-16, ARA05C-05 and ARA05C-01 at the centre of the surveyed area, and ARA05C-03, and ARA05C-01 at the northeast section of the surveyed area. Moreover, the presence of discontinuous ice-bearing permafrost along line ARA05C-11 can be associated with the increase and decrease of the velocity values all along the line. These trends are consistent with the

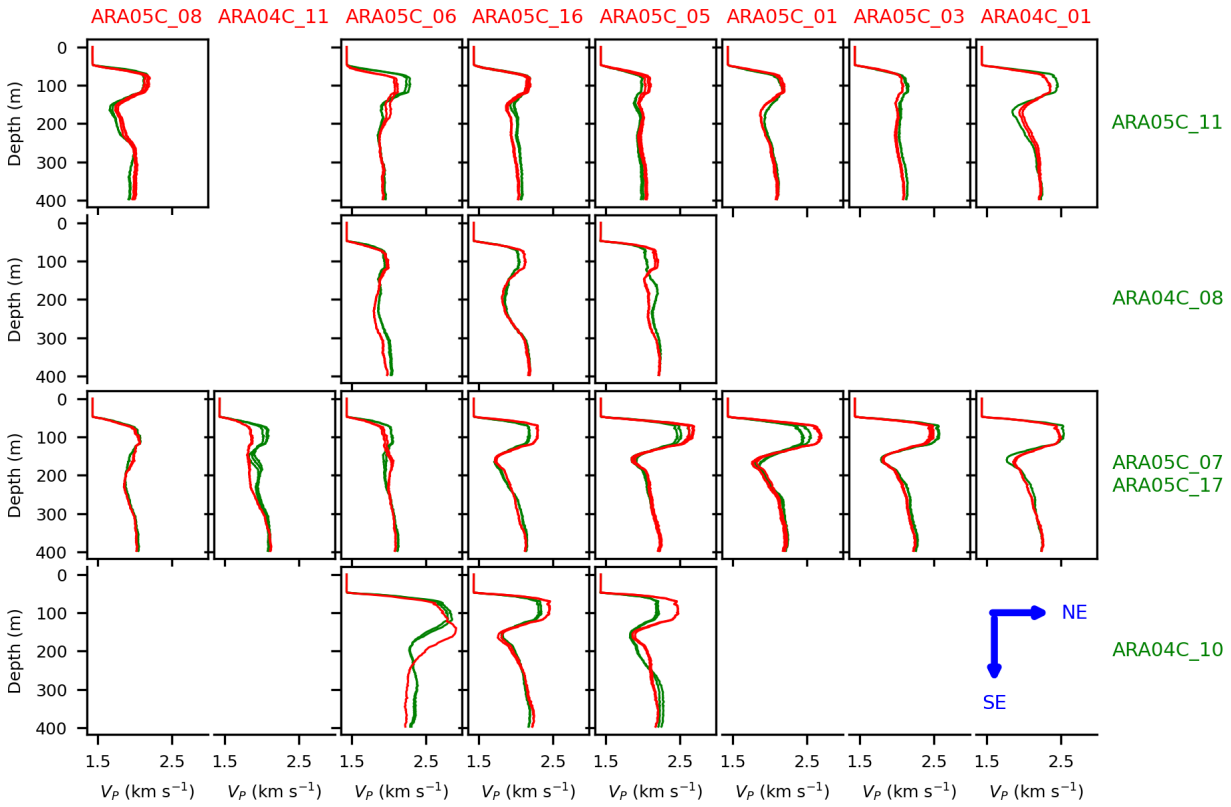


Figure 5.8 Inverted V_P (km s^{-1}) at the intersections of the seismic lines moving perpendicular (red) and parallel (green) to the shelf for the ARA04C and ARA05C surveys. Empty space in the figure means there is no intersection between lines.

existent geological knowledge of the region and support the validity of our inversion approach (Hunter et al., 1976; Pullan et al., 1987).

5.5.3 Well log comparisons

Figure 5.9 compares the well-logs on wells Nektoralik K-59, Kopanoar M-13 and Irkaluk B-35 with the closest 500 m of the inverted V_P velocities for seismic lines ARA05C-11, ARA05C-08, ARA05C-06, ARA04C-10, ARA05C-17, ARA05C-03, and ARA04C-01. Black dotted lines in Figs 5.5 to 5.7 indicate the positions of the well-logs.

Nektoralik K-59, is located at approximately 245 m from line ARA05C-11 and 360 m from line ARA05C-08 within a zone with expected ice-free to low ice-content permafrost. The crystal cable survey was conducted within the depth range of 231 m to 734 m. On Figs. 5.9(a) and 5.9(b), it is apparent that the calculated velocities derived from the crystal cable survey are correlated with the inverted velocities obtained for lines ARA05C-08 and ARA05C-11,

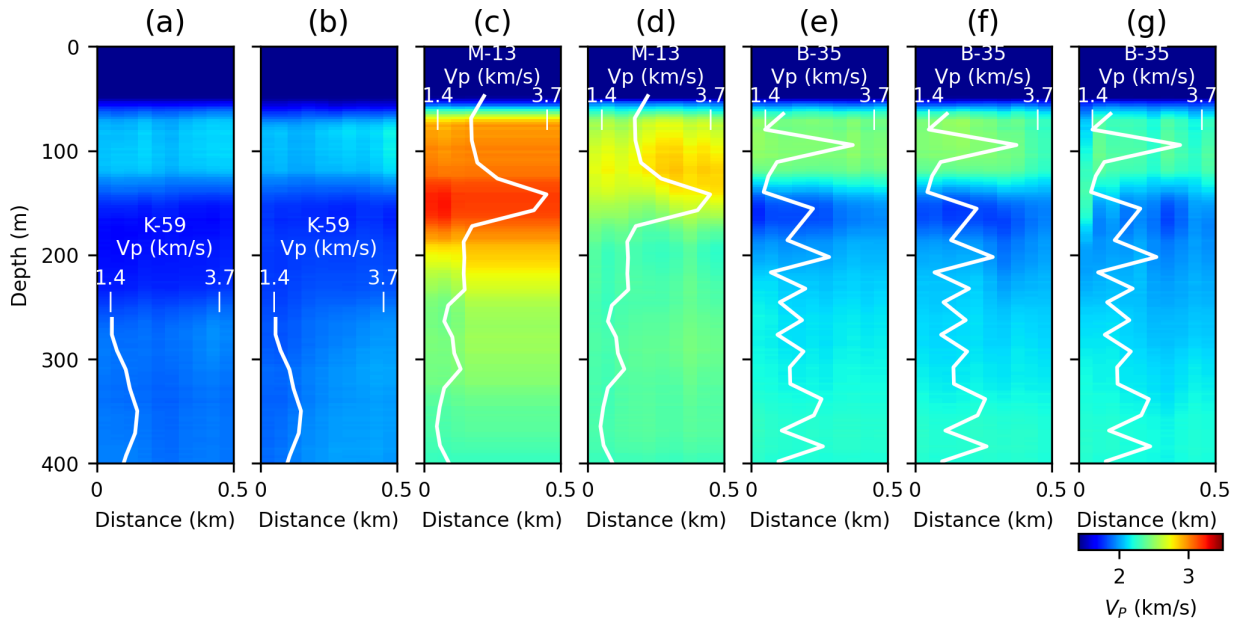


Figure 5.9 Comparison of well logs K-59, M-13, and B-35 (represented by black dotted lines in Figures 5.5-5.7) with the corresponding V_P velocity models processed along the 500 m section closest to the wells. Seismic lines ARA05C-11 (a), ARA05C-08 (b), ARA05C-06 (c), ARA04C-10 (d), ARA05C-17 (e), ARA05C-03 (f), and ARA04C-01 (g) are used for comparisons. Note that the color scale of the sections matches the minimum and maximum values of the velocity scale of the well logs.

respectively. Interestingly, near the uppermost part (around 231 m), the velocities obtained from the crystal cable correspond to the low V_P values derived from the processed model. Moreover, the velocity values exhibit a smooth increase with depth, culminating at 400 meters.

Eastward, well Kopanoar M-13 is located 75 m from line ARA05C-06 and 975 m from line ARA04C-10. Within this zone, the permafrost conditions are anticipated to range from low ice content to an intermittently ice-bonded state. The crystal cable survey conducted between 65 m and 598 m captures a significant portion of the depth where permafrost is expected. Notably, the velocities agree with the crystal cable survey in the uppermost section of the log (e.g., below 200 m depth), reflecting velocities associated with ice-bonded conditions. Additionally, the apparent lower boundary of the ice-bonded sediments, estimated to be around 180 m, aligns well with the velocity variations observed in the well log. Below this apparent ice-bonded boundary, the seismic velocity slowly reduces in both the well logs and the processed model, suggesting a decreasing ice content with depth. Beneath the 250-meter mark, both the well log and the processed velocities gradually decline until reaching velocity

values characteristic of unfrozen sediments.

Finally, well Irkaluk B-35 is positioned at the easternmost part of the survey close to the intersections of line ARA05C-17 with lines ARA05C-03 and ARA04C-01. More precisely, it is situated 580 m from line ARA05C-17 and 440 m and 445 m from lines ARA05C-03 and ARA04C-01, respectively. In this region, we anticipate permafrost conditions to range from low ice content to discontinuous ice-bonded conditions. The crystal cable survey at this well covers depths ranging from 65 m to 598 m. In Figs. 5.9(e)-(g), the velocities are in good agreement with the V_P velocities obtained from the NN. More specifically, velocities associated with ice-bearing permafrost conditions are observed within the top 100 m, followed by a sudden decrease leading to lower velocities associated with unfrozen sediments beyond the 100-meter depth mark.

In general, the velocities obtained from crystal cable surveys are in good agreement with the velocities derived from the NN. However, the NN results in a smoother velocity field when compared to the well velocity logs. The most significant differences between the crystal cable and NN velocities are observed at the location of well M-13, i.e. in the zone characterized by velocity values associated with discontinuous ice-bonded conditions. Within this zone, the crystal cable velocities indicate ice-bonded conditions at greater depths than the indicated by the processed velocities from line ARA04C-10. However, there is a good agreement between the well logs and the processed velocities from line ARA05C-06. This discrepancy can be attributed to the larger distance (975 m) between well Kopanoar M-13 and the seismic line ARA04C-10. Moreover, around the location of well M-13, the processed seismic parameters suggest rapid variations in permafrost conditions towards the southeast (SE) and northeast (NE) portions of the surveyed lines. This high level of heterogeneity likely contributes to the observed discrepancies. Other factors that could contribute to this mismatch include the sensitivity of the NN, which is limited to a depth resolution of 50 m as defined in the training dataset. Challenges associated with the interpretation of lithologies below the permafrost layer due to strong interference from different arrivals (e.g., reflections, refraction, guided waves and multiples) also represent a possible cause of differences between well and inverted velocities.

5.5.4 Permafrost distribution in the Beaufort Sea

In general, contrasts in V_P values can be observed down to the top 350 m of the inverted parameters, pointing at the potential depth of the base of the permafrost. However, interpreting the base of the permafrost solely based on the processed V_P values presents challenges. Firstly, the presence of remaining vertical artifacts, resulting from the neural network being

primarily trained on a 1D dataset, hampers the identification of the base of the permafrost in certain areas. Secondly, the inferred depth of the permafrost base varies depending on the inverted parameter used for interpretation. Additionally, in some locations, the V_P results do not exhibit the expected velocity inversion that is observed at the base of the permafrost.

A more robust approach can be adopted to overcome these challenges by combining the results obtained from the three inverted parameters: V_P, V_S, and $1/Q$. This integrated analysis allows us to characterize the presence of ice within the pore space effectively. It is important to note that seismic velocities (V_P and V_S) demonstrate increased values in correlation with the ice content and bonding conditions in the permafrost sediments. Conversely, $1/Q$ reaches its peak values when at intermediate ice saturation, where pore fluids, ice and gas can be present in the pore space, while lower values are associated with ice-free and ice-bonded conditions (Dou et al., 2016; Matsushima et al., 2016).

From the seismic parameters for different permafrost conditions outlined in Table 5.1, we can define three distinct zones: Ice-free conditions can be identified when $1/Q \leq$ 2.5e-2 and either $V_P \leq 2.25$ km s^{-1} or $V_S \leq 0.8$ km s^{-1}. Ice-bonded conditions can be inferred when $1/Q \leq$ 2.5e-2 and either $V_P \geq 3.0$ km s^{-1} or $V_S \geq 1.1$ km s^{-1}. Finally, permafrost with ice-bearing conditions can be defined for all other scenarios included in between the conditions defined above.

Figure 5.10 illustrates a fence diagram representation of the interpreted permafrost distribution across the study area. We applied additional adjustments to the threshold derived from the seismic parameters to ensure accuracy. Firstly, we eliminated vertical artifacts originating from the 1D training at the bottom of the inversion by masking them to zero. Furthermore, we employed a horizontal Gaussian filter with a sigma value of 3 (150 m) to achieve smooth transitions between the defined thresholds. Figure 5.10 clearly displays a discernible transition of permafrost conditions. Ice-free sediments dominate the southwestern (SW) and northwestern (NW) regions, gradually shifting towards discontinuous ice-bonded conditions in the northeastern (NE) and southeastern (SE) directions. Notably, our analysis does not reveal evidence of continuous ice-bonded permafrost within the research area. In cases where the model suggests the presence of ice within the pore space, its vertical and lateral distribution was easily traceable.

Our model suggests that the permafrost is primarily concentrated within an upper layer ranging from 60 m to 150 m thick. However, the continuity of the permafrost layer diminishes, and the presence of open taliks increases as we approach the outer shelf, reaching a region where ice saturation becomes too low to affect seismic velocities, particularly towards the NW and SW directions. Note that there is a significant increase in permafrost thickness at the

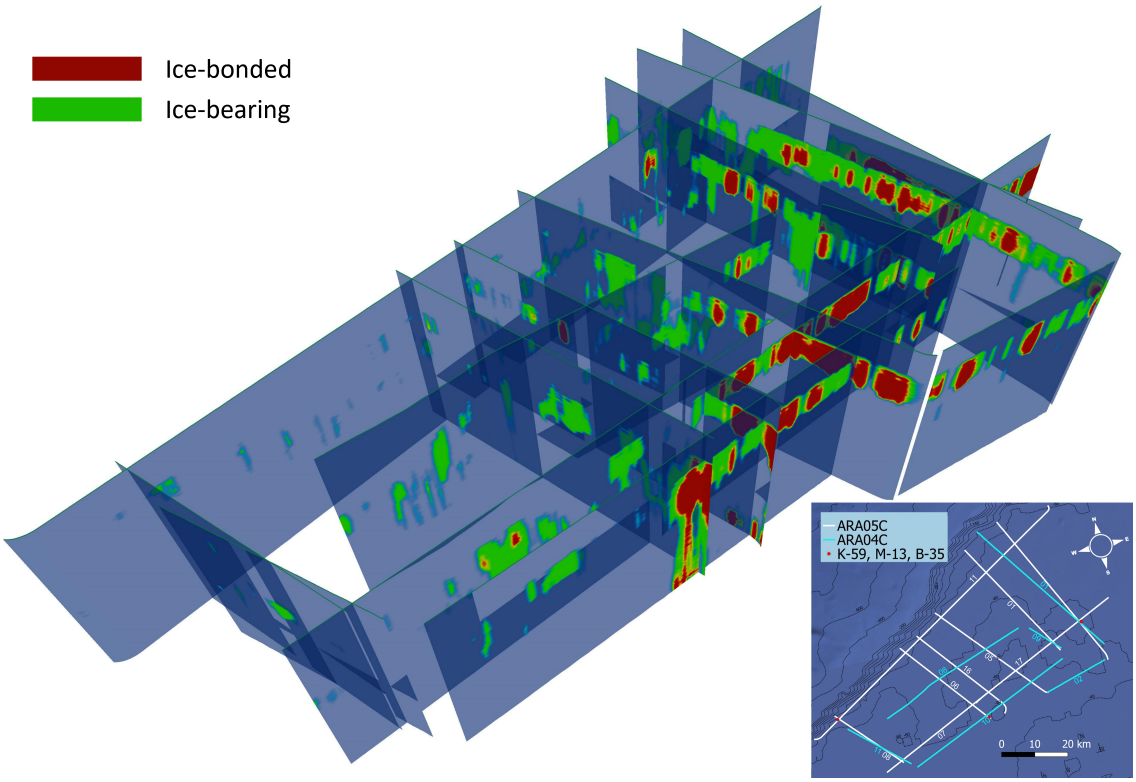


Figure 5.10 Interpretation of subsea permafrost distribution based on the defined thresholds from the inverted V_P, V_S and $1/Q$ values. The inset shows lines locations.

SE end of lines ARA05C-06 and ARA05C-16 and at the intersection with line ARA04C-10. The crystal cable profile obtained from the Kopanoar M-13 well corroborates this gradual augmentation in thickness. However, further investigations and data are required to obtain a more accurate understanding of the permafrost distribution in this particular zone towards the SE.

5.6 Discussion

5.6.1 Comparisons with previous studies

By integrating the three inverted models (V_P, V_S, and $1/Q$) and applying the ice condition threshold, we have achieved unprecedented detailed mapping of the distribution of subsea permafrost at the regional scale in the Canadian Beaufort Sea. Our permafrost distribution model offers valuable insights into the lateral and vertical distribution of three distinct permafrost conditions: ice-free sediments, ice-bearing permafrost, and ice-bonded permafrost. This represents a significant advancement compared to the findings of Pullan et al. (1987),

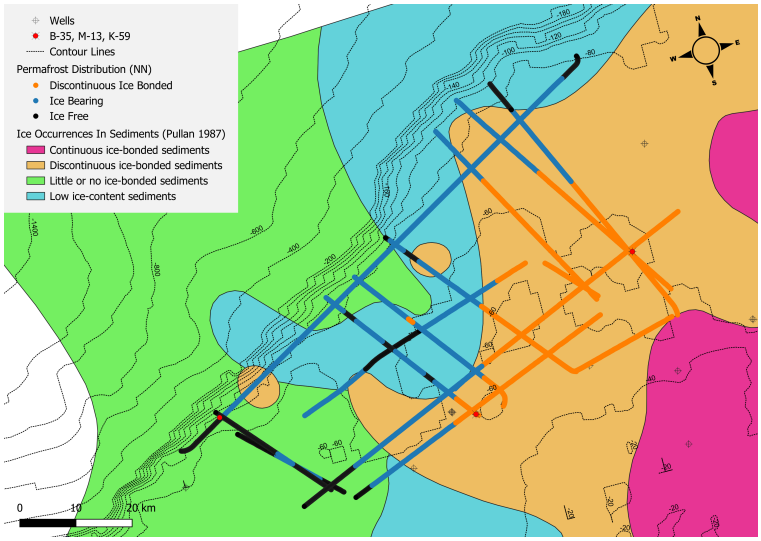


Figure 5.11 Comparison between the permafrost conditions inferred by using NN versus ice occurrence in sediments as interpreted in Pullan et al. (1987)

as we successfully provide permafrost distribution data with depth in a laterally continuous fashion. Furthermore, our work benefited from the research conducted by Hu et al. (2013) by discretizing these distinct permafrost conditions and enhancing the resolution of permafrost distribution. Notably, our model provides enhanced capabilities for mapping permafrost distribution by offering extensive information on the depth of the permafrost base at a larger scale compared to other inversion schemes tested in the region (Kang et al., 2021).

Figure 5.11 presents a comparison between the permafrost distribution inferred using our NN and the ice occurrence map in sediments from Pullan et al. (1987). We observe a general agreement between both models. Most sediments exhibiting ice-bonded conditions inferred using the NN align with the region identified as discontinuous ice-bonded sediments in Pullan et al. (1987). However, minor differences exist within the transition from ice-bonded sediments to low-ice or ice-bearing sediments. These differences can be attributed to the varying thresholds employed by each method.

In Pullan et al. (1987), the thresholds for sediment classification are defined as follows: ice-free sediments for $V_P \leq 1.8$ km s^{-1}, low-ice content sediments for 1.8 km s^{-1} $\leq V_P \leq$ 2.5 km s^{-1}, and ice-bonded sediments for $V_P \geq 2.5$ km s^{-1}. These limits are conservative and may exclude certain lithologies from their corresponding groups. For instance, unfrozen shale may be excluded from the ice-free conditions ($V_P \sim$ 2.1-2.4 km s^{-1}), or partially ice-bonded sands ($V_P \sim$ 3.0-3.5 km s^{-1}) may not be considered within the low-ice sediments category. In contrast, our approach utilizes a multi-parameter threshold to characterize permafrost

conditions using all CMPs of the seismic survey. By combining the inverted models (V_P, V_S, and $1/Q$), we capture a broader range of lithological properties associated with permafrost. Consequently, our method offers a more refined representation of the permafrost distribution, covering a wider spectrum of sediment compositions and ice content variations.

5.6.2 Schematic permafrost distribution in the Canadian Beaufort Sea.

Based on our interpretation of permafrost distribution across the seismic lines, we present a conceptual model of subsea permafrost within the Canadian Beaufort Sea (Fig. 5.12). This model also integrates data from nearby well logs (Fig. 5.9) and previous investigations of permafrost distribution within the area (Hu et al., 2013; Hunter et al., 1976; Pullan et al., 1987).

According to Osterkamp (2001), subsea permafrost degradation occurs mainly due to heat flux from the seabed and geothermal heat from below. This results in thawing occurring at the top and base of the permafrost. This has led to the widely accepted concept of a wedge-shaped permafrost distribution that thins out towards the shelf-edge. (Boswell et al., 2020; Gwiazda et al., 2018; Overduin et al., 2019; Paull et al., 2022; Ruppel et al., 2016; Taylor et al., 2013). Nevertheless, our findings reveal a more complex permafrost distribution in the study area. Particularly, our analysis identifies distinct zones of discontinuous ice-bearing permafrost, highlighting the complexity and variability of permafrost conditions in the Beaufort Sea.

Subsea permafrost thickness gradually diminishes, moving towards the open sea. Reductions in ice content and permafrost thickness towards the shelf edge are evident in processed parameters (V_P, V_S, and $1/Q$) presented in Figs 5.5-5.7, well logs (Fig 5.9), and intersection point comparisons in Fig 5.8. As shown in Fig. 5.10, permafrost continuity decreases towards

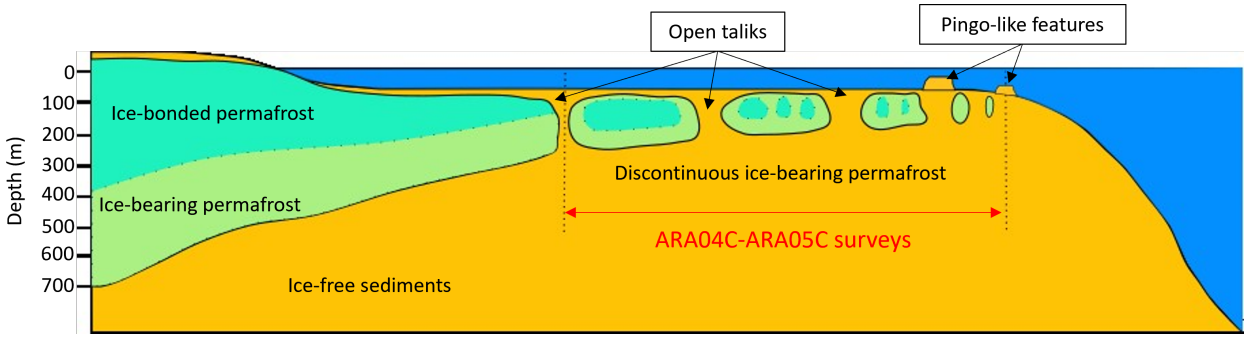


Figure 5.12 Schematic of the general distribution of subsea permafrost in the Canadian Beaufort Sea as observed from the interpreted model. The left side of the figure is inspired by the conventional wedge-shaped permafrost distribution presented in Osterkamp (2001).

the shelf, resulting in diminished volumes of subsea permafrost that progressively shrink in size and ice content until permafrost becomes imperceptible.

Distinct discontinuities observed in ice-bearing permafrost conditions within the Canadian Beaufort Sea correspond to open taliks (as indicated in Fig 5.12). These taliks have earned significant attention from the environmental community due to their role in providing a pathway for greenhouse gases trapped within and beneath the subsea permafrost to reach the ocean layer (Harris et al., 1988; Zhang, 2011). Previous studies that included changes in heat transfer through geological time have considered open taliks for modelling permafrost distribution (Angelopoulos et al., 2019; Romanovskii and Hubberten, 2001). However, it is essential to note that these studies have primarily assumed the occurrence of open taliks in regions characterized by high geothermal heat flux, such as under faults and fault zones, which may not accurately reflect the conditions observed in the surveyed area of the Canadian Beaufort Sea.

One explanation for creating subsea open taliks involves heat advection, which can accelerate the thawing of permafrost layers (Angelopoulos et al., 2020; Shakhova et al., 2017a). In the Canadian Beaufort Sea, Gwiazda et al. (2018) documented the influx of freshwater along the shelf edge and slope. This phenomenon was attributed to a combination of seawater and freshwater filtration, degradation of subsea permafrost, and the discharge of groundwater beneath the seafloor. It is plausible that similar flow pathways may have increased permafrost permeability in specific areas due to heat advection, leading to the development of open taliks.

Furthermore, another factor that likely contributes to talik formation is the movement of water within a thawed permafrost layer. This movement is driven by convection processes (Semenov et al., 2020). However, the investigation of the groundwater system in the Canadian Beaufort Sea has been limited, and submarine groundwater discharge beneath and within subsea permafrost is still poorly understood (Charkin et al., 2017).

The subsea permafrost volumes delimited by open taliks show increasing ice content towards the centre (Fig. 5.10). Frequently, these sections display ice-bearing characteristics at their outer edges while preserving ice-bonding properties towards the central core. In some cases, the ice-bonding conditions are visible across various cores within a given subsea permafrost volume. As the thickness of the water layer increases towards the outer shelf, subsea permafrost volumes undergo a size reduction. Near the shelf-edge, subsea permafrost volumes become scarcer and smaller, before disappearing or becoming undetectable by seismic methods.

The schematic model of subsea permafrost distribution in the Canadian Beaufort Sea, as depicted in Figure 5.12, incorporates certain features associated with permafrost structures that

have been documented in the literature but are not distinctly observed on the inverted seismic data such as pingo-like features. Two factors may account for the absence of discernible pingo-like features within the inverted region. Firstly, it is plausible that pingos, being isolated geomorphological entities, do not intersect with any of the seismic lines analyzed in our study. Secondly, it is conceivable that the presence of pingos may elude detection due to limitations in the resolving power of our inversion algorithm, primarily contingent upon the depth of the training dataset, which is primarily constrained to a depth of 50 meters.

Conclusion

In this study, we evaluated subsea permafrost distribution on 15 seismic lines collected in two separated surveys on the outer shelf in the Canadian Beaufort Sea. Seismic inversion was performed on the lines using a multi-input, multi-output NN trained on synthetic data from 1D and 2D scenarios. The NN input seismic information in four distinct domains and outputs inverted V_P, V_S and $1/Q$. We identified a consistent pattern in the processed seismic velocity and attenuation indicating a gradual increase in ice content and bonding effects towards the coastline, specifically, towards the northeast and southeast areas of the studied region.

To validate the robustness of the model, we conducted a comparative analysis of the inverted parameters at the points where the seismic lines intersect. This test demonstrated the capability of the NN to handle the seismic dataset effectively, regardless of the line azimuth. Furthermore, we carried out model validation near three well locations by tying the V_P inversion results with well-velocity surveys. This comparison allowed us to ground truth the accuracy and reliability of our models in capturing the subsurface characteristics close to these wells.

We comprehensively characterized the permafrost distribution in the study region by establishing three distinct thresholds based on the inverted parameters V_P, V_S, and $1/Q$. These thresholds allow us to differentiate ice-free sediments from ice-bearing permafrost, and ice-bonded permafrost. The resulting interpretation of the permafrost distribution exhibits unprecedented vertical and lateral resolution, over laterally continuous segments, significantly advancing subsea permafrost investigations in the region. This will hopefully trigger further research interest to study the geological processes that can lead to the subsea permafrost distribution observed in the Canadian Beaufort Sea. The observations made on seismic inversion are summarized in a conceptual model.

Our method achieves a higher resolution and better applicability to large-scale interpretations compared to other methods that have been recently reported in the literature. Still, it is

limited by the representativeness of the training dataset with the recorded data. The high resolution of our method reveals subtle details and variations in permafrost conditions across the study area. The level of accuracy achieved opens up new possibilities for understanding the complex dynamics and processes associated with subsea permafrost on a broader scale. For example, within the study region, we have observed a reduction in both the continuity and thickness of subsea permafrost as we extend farther from the coastline. However, the precise mechanisms explaining the subsea permafrost distribution we observe remain unknown.

Further investigations using the method could be conducted to assess permafrost distribution over a larger area of the Canadian Beaufort Sea. Of particular interest is the understanding of transition zones from ice-free to ice-bonded permafrost conditions towards the coastal region. Moreover, additional advancements in training the NN could involve improving its adaptability to various survey geometries, thereby facilitating the evaluation of additional seismic lines already acquired within the study area. Those studies can also include the analysis of the effect of anisotropy induced by the permafrost layers.

In addition, our work paves the way for large-scale automated reprocessing of existing seismic data in polar regions, which has the potential to enhance the accuracy and precision of submarine permafrost mapping. This automated approach presents a valuable opportunity to refine our understanding of the distribution of submarine permafrost in the Canadian Beaufort Sea.

Open Research Section

The raw seismic dataset used in this study can be accessed online via the Korea Polar Data Center (KPDC) `https://dx.doi.org/doi:10.22663/KOPRI-KPDC-00002217.1`. Similarly, the code for generating the data results from this article are available in `https://github.com/GeoCode-polymtl/Mapping_SSP_NN.git`.

5.7 Acknowledgments

The authors would like to thank the Korea Polar Research Institute (KOPRI) and the members of the ARA04C and ARA05C expeditions for providing access to the recorded data.

5.8 Addendum A: Limitations of the NN

Like any neural network approach, the accuracy of our inverted seismic models is closely tied to the representativeness of the training dataset with the recorded data. Previous stud-

ies (Duchesne et al., 2022; Fabien-Ouellet et al., 2020) have established that most features observed in seismic data from permafrost zones can be effectively replicated by generating synthetic data based on tabular velocity models. Building upon these findings, our NN is initially trained using synthetic 1D models. Subsequently, a second training phase uses synthetic 2D layers to mitigate artifacts that may arise when applying 2D data to the NN trained solely on 1D data.

Still, the capacity of the NN is limited to the training dataset. For instance, as the minimum thickness of the layers used for the training is 50 m, we can expect that the processed models have a maximum resolution of 50 m in depth. Moreover, although the order of the geological sequence in the training dataset is well justified in chapters 3 and 4, any deviation from it can lead to additional errors in the inverted models.

The characteristics of the 2D training dataset also play a critical role in constraining the lateral resolution of our neural network (NN). Specifically, the maximum slope of 15^o present in the training dataset limits our ability to interpret permafrost distribution around the shelf edge. However, it is essential to note that this restriction was considered during the design of the 2D NN, given that permafrost is expected to occur within the 100 m isobath.

Figure 5.13 demonstrates an example of the inverted V_P, V_S and $1/Q$ models using the NN in a zone with features not included in the training dataset. The seismic data was collected in the northwest section of line ARA05C-16, surveyed over the Beaufort continental slope. The inverted parameters correctly display the seafloor depth until reaching 400 meters from 21 km towards the southeast end of the line. However, as the slope increases, the models show a layer with increasing velocities and attenuation, which is not expected at this location as the permafrost boundary is limited by the 100 m isobath. We believe that the presence of this layer is due to artefacts generated by the NN, as the training dataset only includes slopes lower than 15^o.

From 21 km towards the northwest end of the line, the NN cannot accurately locate the seafloor because the NN is limited to a maximum of 400 m seafloor depth. In this section, the NN shows an irregular seafloor with a depth lower than 400 meters instead of the slope. Moreover, the inverted velocity and attenuation indicate increased values arranged in a series of layers, which suggest the presence of clinorforms. However, as the shelf slope continues in the surveyed interval (Figure 5.13a)), clinoforms are not expected at this location.

Generating a training dataset that resembles the features observed in the recorded data was a crucial task for this thesis. Throughout the project, the training dataset underwent various transformations. Initially, we created a simple tabular 5-layer dataset with a water layer and random lithologies. This dataset helped us test the base components of the NN. After creating

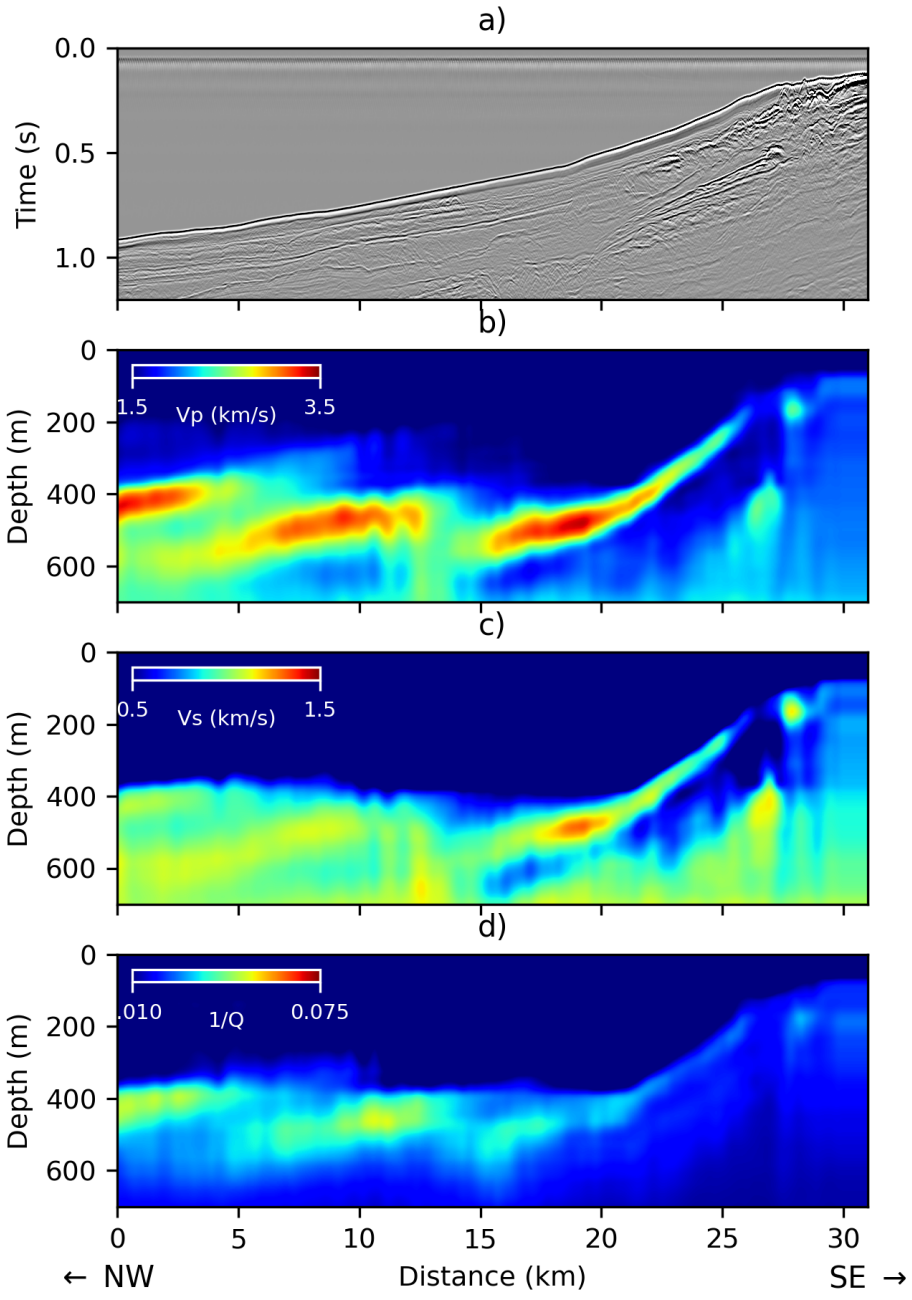


Figure 5.13 Stacked image (a) and Inverted V_P (b), V_S (c), and $1/Q$ (d) on 30 km of line ARA05C-16 recorded on the NW section of the Beaufort continental slope.

a raw version of the NN, we modified the training dataset to include velocities typically found in permafrost sediments. We then performed a sensitivity analysis to evaluate the effect of changing permafrost parameters such as depth, thickness, seismic velocity, and attenuation on different domains. Once we completed the sensitivity analysis, we modified the training

dataset again to match the lithology found in the Beaufort Sea. At this point, we also varied the number of layers in the models. To account for wavelengths in a fast layer of 4.0 km/s, 3.0 k/s, and in unfrozen sediments with 2.0 km/s, we designed three subsets with minimum thicknesses of 100 m, 75 m, and 50 m. Although we initially intended to provide sequential training resembling the sequential inversion applied in FWI strategies, the NN was able to generalize for the 3 subsets without the need for sequential training, which improved training times. Finally, we designed a similar approach for creating a 2D dataset that resembles the features found in the outer continental shelf in the Canadian Beaufort Sea. This approach allowed our NN to learn different characteristics expected in the survey area, but it also limited its application to this particular zone. In the end, although the proposed approach is limited to seismic data collected in the Beaufort Sea, it provided enhanced results than the obtained using conventional seismic processing techniques Grob et al. (2023) or FWI approaches Kang et al. (2021) in the same zone.

One more limitation of our approach stems from the inherent characteristics of the recorded data itself. Specifically, the significant interference between different arrivals in the CMP gathers can persist even after data transformation into different domains. Extensive research conducted by Bustamante et al. (2023b) shows that using NN for inverting seismic parameters (e.g. seismic velocity and attenuation) in permafrost zones can lead tp better results than other methods such as travel time tomography or physics-based inversion. However, if the interference between different arrivals is sufficiently strong that our NN cannot accurately identify the relevant training features, it may result in a misfit of the inverted models.

CHAPTER 6 CONCLUSION AND FUTURE WORK

This thesis aimed to assess the potential of seismic data combined with deep learning techniques to improve subsurface mapping and characterization of subsea permafrost. We developed a structured methodology consisting of three progressive steps to achieve this objective. The first step was a sensitivity analysis to assess the viability of using seismic surveys for discerning changes in permafrost properties. Next, we designed a deep neural network to invert seismic properties from field data collected in areas where the occurrence of permafrost is known. Finally, we evaluated the proposed NN using two regional-scale surveys collected in the Canadian Beaufort Sea.

The viability of using seismic surveys to detect changes in permafrost properties is explored in Chapter 3 through a sensitivity analysis. This analysis evaluates the variations in geometrical and seismic parameters of a permafrost layer, such as depth, thickness, seismic velocity, and attenuation, in the $t-x$, $f-c$, $\tau-p$ and $f-p$ domains. The sensitivity curves indicate that the $t-x$ domain exhibits constructive and destructive interference of multiple arrivals, resulting in several sensitivity peaks. In contrast, the sensitivity of the $f-c$, $\tau-p$ and $f-p$ domain shows a consistent monotonic decrease for all tested parameters. Despite being influenced by constructive/destructive interference, the $f-c$, $\tau-p$ and $f-p$ domains still provide valuable information about changes in permafrost layer geometry and seismic parameters. Notably, the sensitivity of the $\tau-p$ domain concerning thickness displays an interesting behaviour where a stable curve is visible for the evaluated depths. These results suggest that a combined approach that integrates all four domains can significantly enhance the capabilities of the inversion approach in determining changes in permafrost distribution.

The development of the deep-learning methodology for the inversion of seismic properties from recorded data is detailed in Chapter 4. This work introduces a novel multi-input, multi-output NN architecture. This architecture takes as input seismic images from the four distinct domains discussed in the sensitivity analysis (Chapter 3). The NN produces inverted vectors that represent models of V_P, V_S, and $1/Q$ at specific positions. A two-stage training strategy is implemented for the NN to address high computational complexities. Initially, the network undergoes training using synthetic seismic images generated from 1D layered models. Subsequently, a transfer learning approach is employed to retrain the NN, using lateral information extracted from 2D synthetic seismic models. The effectiveness of the proposed NN architecture is evaluated through its application to two seismic lines acquired from the Canadian Beaufort Sea. The results demonstrate the capability of the NN to image

the subsurface velocity model in permafrost zones accurately.

Assessing the potential of the proposed NN architecture for large-scale inversion of subsea permafrost distribution constitutes the focal point of Chapter 5. Making use of the computing performance of the NN, we analyzed data from 15 seismic lines acquired during two surveys on the outer continental shelf of the Canadian Beaufort Sea. The inversion results reveal a trend with increasing seismic velocities and more continuous permafrost layers towards the coast. While the inverted models are representative of the current knowledge of the permafrost distribution and are confirmed by well logs located near the research area, inferring permafrost distribution solely from the models presents challenges. Therefore, our approach uses predetermined thresholds on the inverted parameters, which allowed us to effectively outline the permafrost distribution into three categories: ice-free, ice-bearing, and ice-bonded permafrost.

To summarize, this thesis includes three distinct and original contributions. Firstly, using seismic data displayed in various domains, our study underlines the capacity to extract valuable insights about permafrost distribution and its seismic characteristics (e.g., velocity and attenuation). Secondly, in the absence of a standard processing methodology to use seismic information for subsea permafrost mapping and characterization, we introduce a multi-input, multi-output neural network architecture designed for the inversion of seismic data related to permafrost zones. Finally, our proposed neural network framework was used to assess the permafrost distribution in the Canadian Beaufort Sea, using 15 seismic lines from the ARA04C and ARA05C surveys.

The findings presented in this thesis mark the initial step towards a broader perspective. Ultimately, our work aimed at providing a comprehensive monitoring of subsea permafrost degradation at the regional scale, thereby contributing to forecast the impacts of its degradation on global warming. Further work may include implementing a time-lapse strategy to monitor the progression of permafrost degradation systematically.

The main challenge to such large-scale monitoring and mapping remains the accessibility of seismic data. For instance, vast zones of the Canadian Beaufort Sea have been extensively covered by industry seismic surveys since the 1970s. Similar to the work of Hunter et al. (1976) and Pullan et al. (1987), the proposed NN could be tailored to invert existing seismic datasets. Furthermore, additional seismic expeditions would be required to assess temporal changes through a time-lapse approach (e.g., every 10 years), facilitating a comparative analysis to provide insights into subsea permafrost degradation. Finally, our methodology can be extended to other difficult subsurface conditions where the interference of different seismic wave arrivals can impede seismic processing and interpretation. Some examples of

such subsurface conditions include shallow-water carbonates or other hard-bottom formations that are prone to generating guided waves.

6.1 Future work

The results of this dissertation open the door to a better understanding of the complex processes associated with the distribution of subsea permafrost. Within this context, two promising avenues for future investigation can be outlined. Firstly, there is potential for enhancing the capabilities of the NN to improve the resolution of the inverted models and facilitate interpretation of associated hydrological and geological processes. Secondly, conducting additional studies and surveys in the Beaufort Sea region can substantially augment our knowledge about permafrost distribution and its associated processes.

6.1.1 Enhancing the capabilities of the NN

On one side, enhancing the NN capabilities can be achieved by improving the quality of the training dataset. Although 2D features have already been incorporated during the training phase, adopting supplementary transfer learning methodologies is a viable avenue to teach the NN additional features that might have been overlooked in the testing dataset. The presence of taliks prominently observed in the inverted permafrost distribution model are of particular significance. Moreover, to further augment the capacity of the NN, features such as PLFs (pingo-like features), gas channels, or pockmarks, which have been extensively documented in the existing literature, should be included in the synthetic models of the training set.

To enhance the capabilities of neural networks, leveraging emerging architectures can be advantageous. One promising approach involves incorporating the proposed neural network as a discriminator within a Generative Adversarial Network (GAN) framework (Marano et al., 2023; Wang et al., 2022; Zhang and Lin, 2020). The generator in this context can be constructed using either a physics-based forward model or an inverted variant of our proposed architecture. Our NN can discriminate real data from synthetic data built using the generator. For example, we adjust our predefined loss function to discriminate real data if the match is better than a specific value (e.g., 75%) and synthetic data otherwise. Then competitive training between the generator to reproduce such an accurate synthetic model to cheat the discriminator and the discriminator to discern between real and synthetic data could help our NN to further improve the capacities of inverting seismic properties in permafrost zones. However, successfully implementing the adversarial methodology necessitates

additional research to define real data that the discriminator can discern precisely.

An alternative strategy involves redefining the NN through a diffusion architecture (Atwood and Towsley, 2016). This approach trains the neural network on "clean" data instead of incorporating noise into the training dataset. Subsequently, a separated neural network is trained to discern noise from a contaminated data version (Durall et al., 2023). In this case, a significant portion of the research effort can be devoted to understanding the specific types of noise that the disruptive neural network can accurately identify. Such investigations are essential for advancing the effectiveness and robustness of the proposed approach.

Moreover, the implementation of confidence metrics could be useful in enhancing the capabilities of the NN. However, validating the accuracy of the inverted models using metrics such as well logs is challenging, given the sparsity of well logs in the Beaufort Sea. Most of the available well logs do not survey the depth of interest, (e.g., casing starting over 500 m depth) making it difficult to use them as confidence metrics. One way to overcome this challenge is by reproducing the seismic data using the modelled data and comparing it with the recorded data. However, our NN is not trained to reproduce seismic data, and we cannot expect it to perform well in this regard.

Instead, we can use our NN in a transfer learning approach for a physics-guided NN. For example, Dhara and Sen (2022) proposed to tie a forward computation of the wave equation at the output of the NN. This approach allows the NN to provide a velocity model, and the forward models calculate the corresponding seismic data. As in an FWI methodology, Dhara and Sen (2022) proposed to minimize the difference between recorded and synthetic data, by modifying the weights of the NN instead of the velocity model itself. We can use this methodology by tying the forward model to our NN and minimizing the error between observed and recorded seismic data from each seismic line on the ARA04C and ARA05C surveys. However, the number of iterations needed for convergence could be a limitation of this approach. Although we suspect that our pre-trained NN could significantly improve the convergence rate, further studies are required in this context.

6.1.2 Additional studies and surveys in the Beaufort Sea

The evaluation of the NN can be extended to incorporate additional surveys conducted in the Beaufort Sea. This endeavour can be accomplished through two approaches: acquiring new data or using pre-existing seismic surveys. Notably, previous studies, such as those conducted by Hunter et al. (1976) and Pullan et al. (1987), have proven that industry data can be used to determine permafrost distribution. However, access to such information is generally limited due to its confidential nature.

Incorporating additional surveys into the proposed methodology for determining permafrost distribution can provide valuable insights and pave the way for in-depth geological investigations. For instance, building upon the proposed permafrost distribution outlined in Chapter 5, it is suggested that taliks may become more developed towards the shelf-edge, which could influence fluid flow patterns. Consequently, additional insights on offshore groundwater flow can likely help to understand better the distribution of subsea permafrost on the outer continental shelf. Moreover, analyzing the hydrogeological conditions can provide an enhanced understanding of the formation of PLFs. This study may shed light on whether PLFs arise due to the freezing of freshwater flowing from underneath the shelf (Paull et al., 2022) or may result from overpressured gas accumulation (Bogoyavlensky, 2015; Serov et al., 2015). Furthermore, incorporating data on ice distribution obtained through our deep learning approach can aid in investigating large-scale permafrost degradation in future hydrogeological studies.

REFERENCES

Abbas, A., Vantassel, J. P., Cox, B. R., Kumar, K., and Crocker, J.: A Frequency-Velocity CNN for Developing Near-Surface 2D Vs Images from Linear-Array, Active-Source Wavefield Measurements, arXiv preprint arXiv:2207.09580, 2022.

Adler, A., Araya-Polo, M., and Poggio, T.: Deep learning for seismic inverse problems: Toward the acceleration of geophysical analysis workflows, IEEE Signal Processing Magazine, 38, 89–119, 2021.

Agarap, A. F.: Deep learning using rectified linear units (relu), arXiv preprint arXiv:1803.08375, 2018.

Alfarraj, M. and AlRegib, G.: Semisupervised sequence modeling for elastic impedance inversion, Interpretation, 7, SE237–SE249, 2019.

Allard, M., Lemay, M., Barrette, C., L'Hérault, E., Sarrazin, D., Bell, T., and Doré, G.: Permafrost and climate change in Nunavik and Nunatsiavut: Importance for municipal and transportation infrastructures, Nunavik and Nunatsiavut: From science to policy. An Integrated Regional Impact Study (IRIS) of climate change and modernization, pp. 171–197, 2012.

Ammon, C., Velasco, A., Thorne Lay, T., and Wallace, T.: An overview of earthquake and seismic-wave mechanics, Foundations of modern global seismology, 2, 39–63, 2020.

Angelopoulos, M., Westermann, S., Overduin, P., Faguet, A., Olenchenko, V., Grosse, G., and Grigoriev, M. N.: Heat and salt flow in subsea permafrost modeled with CryoGRID2, Journal of Geophysical Research: Earth Surface, 124, 920–937, 2019.

Angelopoulos, M., Overduin, P. P., Miesner, F., Grigoriev, M. N., and Vasiliev, A. A.: Recent advances in the study of Arctic submarine permafrost, Permafrost and Periglacial Processes, 31, 442–453, 2020.

Anthony, K. M. W., Anthony, P., Grosse, G., and Chanton, J.: Geologic methane seeps along boundaries of Arctic permafrost thaw and melting glaciers, Nature Geoscience, 5, 419, 2012.

Araya-Polo, M., Jennings, J., Adler, A., and Dahlke, T.: Deep-learning tomography, The Leading Edge, 37, 58–66, 2018.

Aster, R. C., Borchers, B., and Thurber, C. H.: Parameter estimation and inverse problems, Elsevier, 2018.

Atwood, J. and Towsley, D.: Diffusion-convolutional neural networks, Advances in neural information processing systems, 29, 2016.

Badrinarayanan, V., Kendall, A., and Cipolla, R.: SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 39, 2481–2495, https://doi.org/10.1109/TPAMI.2016.2644615, 2017.

Bartlett, P. L. and Maass, W.: Vapnik-Chervonenkis dimension of neural nets, The handbook of brain theory and neural networks, pp. 1188–1192, 2003.

Barzilai, J. and Borwein, J. M.: Two-point step size gradient methods, IMA journal of numerical analysis, 8, 141–148, 1988.

Batchelor, C., Dowdeswell, J., and Pietras, J.: Evidence for multiple Quaternary ice advances and fan development from the Amundsen Gulf cross-shelf trough and slope, Canadian Beaufort Sea margin, Marine and Petroleum Geology, 52, 125–143, 2014.

Baydin, A. G., Pearlmutter, B. A., Radul, A. A., and Siskind, J. M.: Automatic Differentiation in Machine Learning: a Survey, Journal of Machine Learning Research, 18, 1–43, URL `http://jmlr.org/papers/v18/17-468.html`, 2018.

Bellefleur, G., Riedel, M., Brent, T., Wright, F., and Dallimore, S. R.: Implication of seismic attenuation for gas hydrate resource characterization, Mallik, Mackenzie Delta, Canada, Journal of Geophysical Research-Solid Earth, 112, https://doi.org/ArtnB1031110.1029/2007jb004976, 2007.

Bishop, C. M.: Neural networks and their applications, Review of scientific instruments, 65, 1803–1832, 1994.

Blasco, S., Jenner, K., Davies, E., Michel, F., Pollard, W., Graham, C., and Ruffel, P.: Origin and Evolution of Subsea Ice-bearing Permafrost on the Canadian Beaufort Shelf: Implications from a 500m Deep Borehole, Extended Abstracts of the Tenth International Conference on Permafrost, 4, 2012.

Bogoyavlensky, V.: Gas blowouts on the Yamal and Gydan Peninsulas, GEO Expro Magazine, 2015.

Boiero, D., Wiarda, E., and Vermeer, P.: Surface-and guided-wave inversion for near-surface modeling in land and shallow marine seismic data, The Leading Edge, 32, 638–646, 2013.

Boswell, R., Hancock, S., Yamamoto, K., Collett, T., Pratap, M., and Lee, S.-R.: Natural gas hydrates: Status of potential as an energy resource, Future energy, pp. 111–131, 2020.

Bottou, L. and Bousquet, O.: The tradeoffs of large scale learning, Advances in neural information processing systems, 20, 2007.

Bouchard, F., MacDonald, L. A., Turner, K. W., Thienpont, J. R., Medeiros, A. S., Biskaborn, B. K., Korosi, J., Hall, R. I., Pienitz, R., and Wolfe, B. B.: Paleolimnology of thermokarst lakes: a window into permafrost landscape evolution, Arctic Science, 3, 91–117, 2016.

Bridle, J.: Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimation of parameters, Advances in neural information processing systems, 2, 1989.

Bring, A., Fedorova, I., Dibike, Y., Hinzman, L., Mård, J., Mernild, S., Prowse, T., Semenova, O., Stuefer, S. L., and Woo, M.-K.: Arctic terrestrial hydrology: A synthesis of processes, regional effects, and research challenges, Journal of Geophysical Research: Biogeosciences, 121, 621–649, 2016.

Brothers, L. L., Hart, P. E., and Ruppel, C. D.: Minimum distribution of subsea ice-bearing permafrost on the US Beaufort Sea continental shelf, Geophysical Research Letters, 39, 2012a.

Brothers, L. L., Kelley, J. T., Belknap, D. F., Barnhardt, W. A., Andrews, B. D., Legere, C., and Clarke, J. E. H.: Shallow stratigraphic control on pockmark distribution in north temperate estuaries, Marine Geology, 329, 34–45, 2012b.

Brothers, L. L., Herman, B. M., Hart, P. E., and Ruppel, C. D.: Subsea ice-bearing permafrost on the US Beaufort Margin: 1. Minimum seaward extent defined from multichannel seismic reflection data, Geochemistry, Geophysics, Geosystems, 17, 4354–4365, 2016.

Buckingham, M. J.: Theory of acoustic attenuation, dispersion, and pulse propagation in unconsolidated granular materials including marine sediments, The journal of the acoustical society of America, 102, 2579–2596, 1997.

Buddo, I., Sharlov, M., Shelokhov, I., Misyurkeeva, N., Seminsky, I., Selyaev, V., and Agafonov, Y.: Applicability of transient electromagnetic surveys to permafrost imaging in Arctic West Siberia, Energies, 15, 1816, 2022.

Burn, C.: Permafrost distribution and stability, Changing Cold Environments: A Canadian Perspective, pp. 126–146, 2011.

Bustamante, J., Fabien-Ouellet, G., Duchesne, M. J., and Ibrahim, A.: Filtering guided waves for characterizing subsea permafrost, in: SEG/AAPG/SEPM First International Meeting for Applied Geoscience & Energy, OnePetro, 2021.

Bustamante, J., Fabien-Ouellet, G., Duchesne, M. J., and Ibrahim, A.: Understanding seismic velocity variations of subsea permafrost: A sensitivity study, GEOPHYSICS, 88, R655–R669, https://doi.org/10.1190/geo2022-0568.1, 2023a.

Bustamante, J., Fabien-Ouellet, G., Duchesne, M. J., and Ibrahim, A.: Deep-learning viscoelastic seismic inversion for mapping subsea permafrost, submitted, 2023b.

Charkin, A. N., Rutgers van der Loeff, M., Shakhova, N. E., Gustafsson, Ö., Dudarev, O. V., Cherepnev, M. S., Salyuk, A. N., Koshurnikov, A. V., Spivak, E. A., Gunar, A. Y., et al.: Discovery and characterization of submarine groundwater discharge in the Siberian Arctic seas: a case study in the Buor-Khaya Gulf, Laptev Sea, The Cryosphere, 11, 2305–2327, 2017.

Cheng, F., Lindsey, N. J., Sobolevskaia, V., Dou, S., Freifeld, B., Wood, T., James, S. R., Wagner, A. M., and Ajo-Franklin, J. B.: Watching the cryosphere thaw: Seismic monitoring of permafrost degradation using distributed acoustic sensing during a controlled heating experiment, Geophysical Research Letters, 49, e2021GL097 195, 2022.

Cheverda, V., Bratchikov, D., Gadylshin, K., Golubeva, E., Malakhova, V., and Reshetova, G.: Subsea Methane Hydrates: Origin and Monitoring the Impacts of Global Warming, Applied Sciences, 12, 11 929, 2022.

Chuvilin, E., Stanilovskaya, J., Titovsky, A., Sinitsky, A., Sokolova, N., Bukhanov, B., Spasennykh, M., Cheremisin, A., Grebenkin, S., Davletshina, D., and Badetz, C.: A Gas-Emission Crater in the Erkuta River Valley, Yamal Peninsula: Characteristics and Potential Formation Model, Geosciences, 10, https://doi.org/10.3390/geosciences10050170, 2020.

Clemmensen, L. H. and Kjærsgaard, R. D.: Data representativity for machine learning and ai systems, arXiv preprint arXiv:2203.04706, 2022.

Clevert, D.-A., Unterthiner, T., and Hochreiter, S.: Fast and accurate deep network learning by exponential linear units (elus), arXiv preprint arXiv:1511.07289, 2015.

Creel, R. C., Miesner, F., Wilkenskjeld, S., Austermann, J., and Overduin, P. P.: Glacial Isostatic Adjustment reduces past and future Arctic subsea permafrost, EarthArXiv, 2023.

Dallimore, S. and Collett, T.: Intrapermafrost gas hydrates from a deep core hole in the Mackenzie Delta, Northwest Territories, Canada, Geology, 23, 527–530, 1995.

Dallimore, S., Paull, C., Taylor, A., Riedel, M., MacAulay, H., Côté, M., and Jin, Y.: Geohazard investigations of permafrost and gas hydrates in the outer shelf and upper slope of the Canadian Beaufort Sea, in: Proceedings of a symposium to commemorate the contributions of J. Ross Mackay (1915–2014) to permafrost science in Canada, 7th Canadian permafrost conference, pp. 37–42, 2015.

D'Exelle, B.: Representative Sampling, in: Encyclopedia of quality of life and well-being research, Springer, 2014.

Dhara, A. and Sen, M. K.: Physics-guided deep autoencoder to overcome the need for a starting model in full-waveform inversion, The Leading Edge, 41, 375–381, 2022.

Dietrich, J., Dixon, J., and McNeil, D.: Sequence analysis and nomenclature of Upper Cretaceous to Holocene strata in the Beaufort-Mackenzie Basin, Current Research, Part A, Geological Survey of Canada, Paper, 85, 613–628, 1985.

Dix, C.: Seismic prospecting for oil: Harper & Brother, Seismic velocities from surface measurements: Geophysics, 20, 68–86, 1952.

Dixon, J., Lane, L., Dietrich, J., McNeil, D., and Chen, Z.: Geological History of the Late Cretaceous to Cenozoic Beaufort-Mackenzie Basin, Arctic Canada, in: The Sedimentary Basins of the United States and Canada, pp. 695–717, Elsevier, 2019.

Dobinski, W.: Permafrost, Earth-Science Reviews, 108, 158–169, 2011.

Dobiński, W.: Permafrost active layer, Earth-Science Reviews, 208, 103 301, 2020.

Dou, S.: Field and Laboratory Investigations on Seismic Properties of Unconsolidated Saline Permafrost, University of California, Berkeley, 2015.

Dou, S., Nakagawa, S., Dreger, D., and Ajo-Franklin, J.: A rock-physics investigation of unconsolidated saline permafrost: P-wave properties from laboratory ultrasonic measurements, Geophysics, 81, WA233–WA245, 2016.

Duchesne, M. J., Brake, V. I., Claprood, M., Vosoughi, E., Huot-Vézina, G., Kang, S. G., Jang, U., Jin, Y. K., King, E. L., Côté, M. M., and Dallimore, S. R.: Shallow seismic reflection mapping of the Mackenzie Trough, Beaufort Sea, offshore Yukon and Northwest Territories, in: Geological Survey of Canada, Open File 8657, 2020, 19p. https://doi.org/10.4095/321488, 2020.

Duchesne, M. J., Fabien-Ouellet, G., and Bustamante, J.: Detecting subsea permafrost layers on marine seismic data: an appraisal from forward modeling., Near Surface Geophysics, 2022.

Duchi, J., Hazan, E., and Singer, Y.: Adaptive subgradient methods for online learning and stochastic optimization., Journal of machine learning research, 12, 2011.

Dumoulin, V. and Visin, F.: A guide to convolution arithmetic for deep learning, arXiv preprint arXiv:1603.07285, 2016.

Dunbar, M. J., Carmack, E. C., and Williams, W. J.: Beaufort Sea, in: The Canadian Encyclopedia, Historica Canada, https://doi.org/10.1002/9781118705766, 2015.

Durall, R., Ghanim, A., Fernandez, M. R., Ettrich, N., and Keuper, J.: Deep diffusion models for seismic processing, Computers & Geosciences, 177, 105 377, 2023.

Dvornikov, Y. A., Leibman, M. O., Khomutov, A. V., Kizyakov, A. I., Semenov, P., Bussmann, I., Babkin, E. M., Heim, B., Portnov, A., Babkina, E. A., et al.: Gas-emission craters of the Yamal and Gydan peninsulas: A proposed mechanism for lake genesis and development of permafrost landscapes, Permafrost and Periglacial Processes, 30, 146–162, 2019.

Ellis, R. and Oldenburg, D.: Applied geophysical inversion, Geophysical Journal International, 116, 5–11, 1994.

Fabien-Ouellet, G. and Sarkar, R.: Seismic velocity estimation: A deep recurrent neural-network approach, Geophysics, 85, U21–U29, 2020.

Fabien-Ouellet, G., Gloaguen, E., and Giroux, B.: Time domain viscoelastic full waveform inversion, Geophysical Journal International, 209, 1718–1734, 2017.

Fabien-Ouellet, G., Duchesne, M. J., and Bustamante Restrepo, J.: Seismic characterization of subsea permafrost: Insights from viscoelastic modeling, in: SEG International Exposition and Annual Meeting, OnePetro, 2020.

Feng, L., Xue, Y., Chen, C., Guo, M., and Shen, H.: De-aliased high-resolution Radon transform based on the sparse prior information from the convolutional neural network, Journal of Geophysics and Engineering, 19, 663–680, 2022.

Fish, A. M., Zaretsky, Y. K., et al.: Ice strength as a function of hydrostatic pressure and temperature, CRREL Report, 1997.

Forte, E., French, H. M., Raffi, R., Santin, I., and Guglielmin, M.: Investigations of polygonal patterned ground in continuous Antarctic permafrost by means of ground penetrating radar and electrical resistivity tomography: Some unexpected correlations, Permafrost and Periglacial Processes, 33, 226–240, 2022.

Frederick, J. and Buffett, B.: Taliks in relict submarine permafrost and methane hydrate deposits: Pathways for gas escape under present and future conditions, Journal of Geophysical Research: Earth Surface, 119, 106–122, 2014.

Frederick, J. and Buffett, B.: Effects of submarine groundwater discharge on the present-day extent of relict submarine permafrost and gas hydrate stability on the Beaufort Sea continental shelf, Journal of Geophysical Research: Earth Surface, 120, 417–432, 2015.

French, H. M.: The periglacial environment, John Wiley & Sons, 2017.

Fukushima, K.: Cognitron: A self-organizing multilayered neural network, Biological cybernetics, 20, 121–136, 1975.

Fukushima, K.: Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, Biological cybernetics, 36, 193–202, 1980.

Fyfe, J. C., von Salzen, K., Gillett, N. P., Arora, V. K., Flato, G. M., and McConnell, J. R.: One hundred years of Arctic surface temperature variation due to anthropogenic influence, Scientific Reports, 3, 2645, 2013.

Geldart, L. P. and Sheriff, R. E.: Problems in exploration seismology and their solutions, Society of Exploration Geophysicists, 2004.

Géron, A.: Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow, " O'Reilly Media, Inc.", 2022.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y.: Generative adversarial nets, Advances in neural information processing systems, 27, 2014.

Goodfellow, I., Bengio, Y., and Courville, A.: Deep Learning, MIT Press, http://www.deeplearningbook.org, 2016.

Greer, P., Ulinski, A., Khan, S. D., Bering, E., Vega, A., Chitturi, R., Panda, S., and Vandevort, D.: Understanding the change in permafrost by imaging the CRREL Permafrost Tunnel, Fairbanks, Alaska, in: SEG International Exposition and Annual Meeting, p. D011S073R004, SEG, 2022.

Grob, H., Riedel, M., Duchesne, M., Krastel, S., Bustamante, J., Fabien-Ouellet, G., Jin, Y., and Hong, J.: Revealing the extent of submarine permafrost and gas hydrates in the Canadian Arctic Beaufort Sea using seismic reflection indicators, Geochemistry, Geophysics, Geosystems, 24, e2023GC010 884, 2023.

Gruber, S. and Haeberli, W.: Permafrost in steep bedrock slopes and its temperature-related destabilization following climate change, Journal of Geophysical Research: Earth Surface, 112, 2007.

Gwiazda, R., Paull, C., Dallimore, S., Melling, H., Jin, Y., Hong, J., Riedel, M., Lundsten, E., Anderson, K., and Conway, K.: Freshwater seepage into sediments of the shelf, shelf edge, and continental slope of the Canadian Beaufort Sea, Geochemistry, Geophysics, Geosystems, 19, 3039–3055, 2018.

Harrington, P.: Formation of pockmarks by pore-water escape, Geo-Marine Letters, 5, 193–197, 1985.

Harris, S., French, H., Heginbottom, J., Johnston, G., Ladanyi, B., Sego, D., and van Everdingen, R.: Glossary of permafrost and related ground-ice terms, Associate Committee on Geotechnical Research, National Research Council of Canada, Ottawa, 156, 1988.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778, 2016.

Heagy, L. J., Cockett, R., and Oldenburg, D.: GeoSci: Practices to Collaboratively Build Online Resources for Geophysics Education, in: AGU Fall Meeting Abstracts, vol. 2016, pp. ED51A–0786, 2016.

Heffernan, L., Estop-Aragonés, C., Knorr, K.-H., Talbot, J., and Olefeldt, D.: Long-term impacts of permafrost thaw on carbon storage in peatlands: Deep losses offset by surficial accumulation, Journal of Geophysical Research: Biogeosciences, 125, e2019JG005 501, 2020.

Hilbich, C.: Time-lapse refraction seismic tomography for the detection of ground ice degradation, The Cryosphere, 4, 243, 2010.

Hochreiter, S. and Schmidhuber, J.: Long short-term memory, Neural computation, 9, 1735–1780, 1997.

Hovland, M., Gardner, J. V., and Judd, A.: The significance of pockmarks to understanding fluid flow processes and geohazards, Geofluids, 2, 127–136, 2002.

Hu, K., Issler, D., Chen, Z., and Brent, T.: Permafrost investigation by well logs, and seismic velocity and repeated shallow temperature surveys, Beaufort-Mackenzie Basin, Geol. Surv. Can. Open File Rep, 6956, 228, 2013.

Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q.: Densely connected convolutional networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700–4708, 2017.

Hunter, J., Judge, A., and MacAuley, H.: The Occurrence of Permafrost and Frozen Sub-sea Bottom Materials in the Southern Beaufort Sea. Beaufort Sea Project, Department of the Environment, 1976.

Ibrahim, A. and Fabien-Ouellet, G.: Imaging dispersion curves using sparse hybrid Radon transform, in: First International Meeting for Applied Geoscience & Energy, pp. 2839–2843, Society of Exploration Geophysicists, 2021.

Ibrahim, A., Terenghi, P., and Sacchi, M. D.: Simultaneous reconstruction of seismic reflections and diffractions using a global hyperbolic Radon dictionary, Geophysics, 83, V315–V323, 2018.

James, S., Knox, H., Abbott, R., and Screaton, E.: Improved moving window cross-spectral analysis for resolving large temporal seismic velocity changes in permafrost, Geophysical Research Letters, 44, 4018–4026, 2017.

Jensen, F. B., Kuperman, W. A., Porter, M. B., and Schmidt, H.: Computational ocean acoustics, Springer Science & Business Media, 2011.

Jin, Y. and Dallimore, S.: ARA05C Marine Research Expedition—Canada-Korea-USA Beaufort Sea Geoscience Research Program: Summary of 2014 Activities, Geological Survey of Canada, Open File, 7999, 107, 2016.

Jin, Y., Riedel, M., Hong, J., Nam, S., Jung, J., Ha, S., Lee, J., Conway, K., Kim, G., Standen, G., et al.: Overview of field operations during a 2013 research expedition to the southern Beaufort Sea on the RV Araon. Geological Survey of Canada, Open File 7754, 180, 2015.

Johansen, T. A., Digranes, P., van Schaack, M., and Lønne, I.: Seismic mapping and modeling of near-surface sediments in polar areas, Geophysics, 68, 566–573, 2003.

Jong, D., Bröder, L., Tanski, G., Fritz, M., Lantuit, H., Tesi, T., Haghipour, N., Eglinton, T. I., and Vonk, J. E.: Nearshore zone dynamics determine pathway of organic carbon from eroding permafrost coasts, Geophysical Research Letters, 47, e2020GL088 561, 2020.

Jordan, M. I.: Serial order: A parallel distributed processing approach, in: Advances in psychology, vol. 121, pp. 471–495, Elsevier, 1997.

Judd, A. and Hovland, M.: Seabed fluid flow: the impact on geology, biology and the marine environment, Cambridge University Press, 2009.

Kang, S.-G., Jin, Y. K., Jang, U., Duchesne, M. J., Shin, C., Kim, S., Riedel, M., Dallimore, S. R., Paull, C. K., Choi, Y., et al.: Imaging the P-wave velocity structure of Arctic subsea permafrost using Laplace-domain full-waveform inversion, Journal of Geophysical Research: Earth Surface, 126, e2020JF005 941, 2021.

Kaur, H., Pham, N., and Fomel, S.: Separating primaries and multiples using hyperbolic Radon transform with deep learning, in: SEG Technical Program Expanded Abstracts 2020, pp. 1496–1500, Society of Exploration Geophysicists, 2020.

Kazei, V., Ovcharenko, O., Plotnitskii, P., Peter, D., Zhang, X., and Alkhalifah, T.: Mapping full seismic waveforms to vertical velocity profiles by deep learning, Geophysics, 86, R711–R721, 2021.

Keating, S. and Innanen, K. A.: Parameter crosstalk and modeling errors in viscoacoustic seismic full-waveform inversion, Geophysics, 84, R641–R653, 2019.

Kelley, H. J.: Method of gradients, in: Mathematics in Science and Engineering, vol. 5, pp. 205–254, Elsevier, 1962.

King, L. H. and Maclean, B.: Pockmarks on the Scotian shelf, Geological Society of America Bulletin, 81, 3141–3148, 1970.

King, M.: The influence of clay-sized particles on seismic velocity for Canadian Arctic permafrost, Canadian Journal of Earth Sciences, 21, 19–24, 1984.

King, M., Pandit, B., Hunter, J., and Gajtani, M.: Some seismic, electrical, and thermal properties of sub-seabottom permafrost from the Beaufort Sea, in: Proceedings of the Fourth Canadian Permafrost Conference, pp. 268–273, National Research Council of Canada, 1982.

King, M., Zimmerman, R., and Corwin, R.: SEISMIC AND ELECTRICAL PROPERTIES OF UNCONSOLIDATED PERMAFROST 1, Geophysical Prospecting, 36, 349–364, 1988.

Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980, 2014.

Klein, G., Bohlen, T., Theilen, F., Kugler, S., and Forbriger, T.: Acquisition and inversion of dispersive seismic waves in shallow marine environments, Marine Geophysical Researches, 26, 287–315, 2005.

Kneisel, C., Hauck, C., Fortier, R., and Moorman, B.: Advances in geophysical methods for permafrost investigations, Permafrost and periglacial processes, 19, 157–178, 2008.

Krizhevsky, A., Sutskever, I., and Hinton, G. E.: Imagenet classification with deep convolutional neural networks, Advances in neural information processing systems, 25, 2012.

Kugler, S., Bohlen, T., Forbriger, T., Bussat, S., and Klein, G.: Scholte-wave tomography for shallow-water marine sediments, Geophysical Journal International, 168, 551–570, 2007.

Kvenvolden, K. A., Lilley, M. D., Lorenson, T. D., Barnes, P. W., and McLaughlin, E.: The Beaufort Sea continental shelf as a seasonal source of atmospheric methane, Geophysical Research Letters, 20, 2459–2462, 1993.

Lantuit, H., Overduin, P., and Wetterich, S.: Recent progress regarding permafrost coasts, Permafrost and periglacial processes, 24, 120–130, 2013.

Lecher, A. L., Kessler, J., Sparrow, K., Garcia-Tigreros Kodovska, F., Dimova, N., Murray, J., Tulaczyk, S., and Paytan, A.: Methane transport through submarine groundwater discharge to the North Pacific and Arctic Ocean at two Alaskan sites, Limnology and Oceanography, 61, S344–S355, 2016.

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P.: Gradient-based learning applied to document recognition, Proceedings of the IEEE, 86, 2278–2324, 1998.

LeCun, Y., Bengio, Y., and Hinton, G.: Deep learning, nature, 521, 436–444, 2015.

Lewkowicz, A. G. and Harris, C.: Morphology and geotechnique of active-layer detachment failures in discontinuous and continuous permafrost, northern Canada, Geomorphology, 69, 275–297, 2005.

Li, J., Hanafy, S., and Schuster, G.: Wave-Equation Dispersion Inversion of Guided P Waves in a Waveguide of Arbitrary Geometry, Journal of Geophysical Research: Solid Earth, 123, 7760–7774, 2018.

Li, S., Vladimirsky, A., and Fomel, S.: First-break traveltime tomography with the double-square-root eikonal equation, Geophysics, 78, U89–U101, 2013.

Liebner, S., Welte, C. U., et al.: Roles of thermokarst lakes in a warming world, Trends in Microbiology, 28, 769–779, 2020.

Lim, M., Whalen, D., Martin, J., Mann, P., Hayes, S., Fraser, P., Berry, H., and Ouellette, D.: Massive ice control on permafrost coast erosion and sensitivity, Geophysical Research Letters, 47, e2020GL087 917, 2020.

Lind, S., Ingvaldsen, R. B., and Furevik, T.: Arctic warming hotspot in the northern Barents Sea linked to declining sea-ice import, Nature climate change, 8, 634–639, 2018.

Liu, B., Yang, S., Ren, Y., Xu, X., Jiang, P., and Chen, Y.: Deep-learning seismic full-waveform inversion for realistic structural models, Geophysics, 86, R31–R44, 2021.

Lowney, B., Lokmer, I., and O'Brien, G. S.: Multi-domain diffraction identification: A supervised deep learning technique for seismic diffraction classification, Computers & Geosciences, 155, 104 845, 2021.

Maas, A. L., Hannun, A. Y., Ng, A. Y., et al.: Rectifier nonlinearities improve neural network acoustic models, in: Proc. icml, vol. 30, p. 3, Atlanta, Georgia, USA, 2013.

Mackay, J. R.: Pingo growth and collapse, Tuktoyaktuk Peninsula area, western Arctic coast, Canada: A long-term field study, Géographie physique et Quaternaire, 52, 271–323, 1998.

Maier, A., Syben, C., Lasser, T., and Riess, C.: A gentle introduction to deep learning in medical image processing, Zeitschrift für Medizinische Physik, 29, 86–101, 2019.

Marano, G. C., Rosso, M. M., Aloisio, A., and Cirrincione, G.: Generative adversarial networks review in earthquake-related engineering fields, Bulletin of Earthquake Engineering, pp. 1–52, 2023.

Marques, C. R., dos Santos, V. G., Lunelli, R., Roisenberg, M., and Rodrigues, B. B.: Analysis of Deep Learning Neural Networks for Seismic Impedance Inversion: A Benchmark Study, Energies, 15, 7452, 2022.

Martin, G. S., Wiley, R., and Marfurt, K. J.: Marmousi2: An elastic upgrade for Marmousi, The leading edge, 25, 156–166, 2006.

Matsushima, J., Suzuki, M., Kato, Y., Nibe, T., and Rokugawa, S.: Laboratory experiments on compressional ultrasonic wave attenuation in partially frozen brines, Geophysics, 73, N9–N18, 2008.

Matsushima, J., Suzuki, M., Kato, Y., and Rokugawa, S.: Ultrasonic measurements of attenuation and velocity of compressional and shear waves in partially frozen unconsolidated sediment and synthetic porous rock, Geophysics, 81, D141–D153, 2016.

Mavko, G., Mukerji, T., and Dvorkin, J.: The rock physics handbook, Cambridge university press, 2020.

McCulloch, W. S. and Pitts, W.: A logical calculus of the ideas immanent in nervous activity, The bulletin of mathematical biophysics, 5, 115–133, 1943.

McLaughlin, F., Carmack, E., Macdonald, R., Melling, H., Swift, J., Wheeler, P., Sherr, B., and Sherr, E.: The joint roles of Pacific and Atlantic-origin waters in the Canada Basin, 1997–1998, Deep Sea Research Part I: Oceanographic Research Papers, 51, 107–128, https://doi.org/https://doi.org/10.1016/j.dsr.2003.09.010, 2004.

Meng, D., Wu, B., Wang, Z., and Zhu, Z.: Seismic impedance inversion using conditional generative adversarial network, IEEE Geoscience and Remote Sensing Letters, 19, 1–5, 2021.

Menke, W.: Geophysical data analysis: Discrete inverse theory, Academic press, 2018.

Meredith, M., Sommerkorn, M., Cassotta, S., Derksen, C., Ekaykin, A., Hollowed, A., Kofinas, G., Mackintosh, A., Melbourne-Thomas, J., Muelbert, M., et al.: Polar Regions. Chapter 3, in: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate, 2019.

Miner, K. R., Turetsky, M. R., Malina, E., Bartsch, A., Tamminen, J., McGuire, A. D., Fix, A., Sweeney, C., Elder, C. D., and Miller, C. E.: Permafrost carbon emissions in a changing Arctic, Nature Reviews Earth & Environment, 3, 55–67, 2022.

Mousavi, S. M. and Beroza, G. C.: Deep-learning seismology, Science, 377, eabm4470, 2022.

Murray, I.: Backpropagation of Derivatives - Class Notes: Machine Learning and Patern Recognition - University of Edinburgh, URL `https://www.inf.ed.ac.uk/teaching/courses/mlpr/2019/notes/w7c_backprop.html`, 2019.

Mustafa, A., Alfarraj, M., and AlRegib, G.: Joint learning for spatial context-based seismic inversion of multiple data sets for improved generalizability and robustness, Geophysics, 86, O37–O48, 2021.

Najafi, M. R., Zwiers, F. W., and Gillett, N. P.: Attribution of Arctic temperature change to greenhouse-gas and aerosol influences, Nature Climate Change, 5, 246–249, 2015.

Nesterov, Y.: A method for unconstrained convex minimization problem with the rate of convergence O (1/k2), in: Dokl. Akad. Nauk. SSSR, 3, p. 543, 1983.

Nielsen-Englyst, P., Høyer, J. L., Kolbe, W. M., Dybkjær, G., Lavergne, T., Tonboe, R. T., Skarpalezos, S., and Karagali, I.: A combined sea and sea-ice surface temperature climate dataset of the Arctic, 1982–2021, Remote Sensing of Environment, 284, 113 331, 2023.

Nikitenko, M., Glinskikh, V., Mikhaylov, I., and Fedoseev, A.: Mathematical Modeling of Transient Electromagnetic Sounding Signals for Monitoring the State of Permafrost, Russian Geology and Geophysics, 64, 488–494, 2023.

Nocedal, J. and Wright, S.: Numerical optimization, Springer Science & Business Media, 2006.

Notz, D. and Stroeve, J.: Observed Arctic sea-ice loss directly follows anthropogenic CO2 emission, Science, 354, 747–750, 2016.

Operto, S., Gholami, Y., Prieux, V., Ribodetti, A., Brossier, R., Metivier, L., and Virieux, J.: A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice, The leading edge, 32, 1040–1054, 2013.

Opitz, D. and Maclin, R.: Popular ensemble methods: An empirical study, Journal of artificial intelligence research, 11, 169–198, 1999.

Osterkamp, T.: Sub-sea permafrost, Elements of Physical Oceanography: A Derivative of the Encyclopedia of Ocean Sciences, 2, 2902–2912, 2001.

Overduin, P., Schneider von Deimling, T., Miesner, F., Grigoriev, M., Ruppel, C., Vasiliev, A., Lantuit, H., Juhls, B., and Westermann, S.: Submarine permafrost map in the Arctic modeled using 1-D transient heat flux (supermap), Journal of Geophysical Research: Oceans, 124, 3490–3507, 2019.

Overduin, P. P., Haberland, C., Ryberg, T., Kneier, F., Jacobi, T., Grigoriev, M. N., and Ohrnberger, M.: Submarine permafrost depth from ambient seismic noise, Geophysical Research Letters, 42, 7581–7588, 2015.

O'Neill, H. B., Roy-Leveillee, P., Lebedeva, L., and Ling, F.: Recent advances (2010–2019) in the study of taliks, Permafrost and Periglacial Processes, 31, 346–357, https://doi.org/https://doi.org/10.1002/ppp.2050, 2020.

Pachauri, R. K., Gomez-Echeverri, L., and Riahi, K.: Synthesis report: summary for policy makers, in: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2014.

Park, C. B., Miller, R. D., and Xia, J.: Imaging dispersion curves of surface waves on multi-channel record, in: SEG Technical Program Expanded Abstracts 1998, pp. 1377–1380, Society of Exploration Geophysicists, 1998.

Patton, A., Rathburn, S., Capps, D., McGrath, D., and Brown, R.: Ongoing landslide deformation in thawing permafrost, Geophysical research letters, 48, e2021GL092 959, 2021.

Patton, A. I., Rathburn, S. L., and Capps, D. M.: Landslide response to climate change in permafrost regions, Geomorphology, 340, 116–128, 2019.

Paull, C. K., Ussler III, W., Dallimore, S. R., Blasco, S. M., Lorenson, T. D., Melling, H., Medioli, B. E., Nixon, F. M., and McLaughlin, F. A.: Origin of pingo-like features on the Beaufort Sea shelf and their possible relationship to decomposing methane gas hydrates, Geophysical Research Letters, 34, 2007.

Paull, C. K., Dallimore, S., Hughes-Clarke, J., Blasco, S., Lundsten, E., Ussler, W., Graves, D., Sherman, A., Conway, K., Melling, H., et al.: Tracking the decomposition of submarine permafrost and gas hydrate under the shelf and slope of the Beaufort Sea, in: Proceedings of the 7th International Conference on Gas Hydrates, vol. 1, pp. 1689–1699, ICGH Edinburgh, Scotland, 2011.

Paull, C. K., Dallimore, S. R., Jin, Y. K., Caress, D. W., Lundsten, E., Gwiazda, R., Anderson, K., Hughes Clarke, J., Youngblut, S., and Melling, H.: Rapid seafloor changes associated with the degradation of Arctic submarine permafrost, Proceedings of the National Academy of Sciences, 119, e2119105 119, 2022.

Pham, H., Dai, Z., Xie, Q., and Le, Q. V.: Meta pseudo labels, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 11 557–11 568, 2021.

Pinet, N., Brake, V., and Duchesne, M. J.: Regional significance of normal fault systems in the western Canadian Beaufort Sea (Mackenzie delta), Tectonophysics, 850, 229 767, 2023.

Polyakov, I. V., Pnyushkov, A. V., Alkire, M. B., Ashik, I. M., Baumann, T. M., Carmack, E. C., Goszczko, I., Guthrie, J., Ivanov, V. V., Kanzow, T., et al.: Greater role for Atlantic inflows on sea-ice loss in the Eurasian Basin of the Arctic Ocean, Science, 356, 285–291, 2017.

Pullan, S., MacAulay, H., Hunter, J., Good, R., Gagne, R., and Burns, R.: Permafrost distribution determined from seismic refraction, Marine Science Atlas of the Beaufort Sea: Geology and geophysics, 1987.

Péwé, T. L.: Permafrost Definition Thawing Facts Britannica, `https://www.britannica.com/science/permafrost`, (Accessed 21/8/2023), 2015.

Qi, S., Li, G., Chen, D., Niu, F., Sun, Z., Wu, G., Du, Q., Chai, M., Cao, Y., and Yue, J.: Research on the Characteristics of Thermosyphon Embankment Damage and Permafrost Distribution Based on Ground-Penetrating Radar: A Case Study of the Qinghai–Tibet Highway, Remote Sensing, 15, 2651, 2023.

Ramachandran, K., Bellefleur, G., Brent, T., Riedel, M., and Dallimore, S.: Imaging permafrost velocity structure using high resolution 3D seismic tomography, Geophysics, 76, B187–B198, 2011.

Razavi, S. and Gupta, H. V.: What do we mean by sensitivity analysis? The need for comprehensive characterization of "global" sensitivity in Earth and Environmental systems models, Water Resources Research, 51, 3070–3092, 2015.

Rekant, P., Cherkashev, G., Vanstein, B., and Krinitsky, P.: Submarine permafrost in the nearshore zone of the southwestern Kara Sea, Geo-Marine Letters, 25, 183–189, 2005.

Riedel, M., Villinger, H., Asshoff, K., Kaul, N., and Dallimore, S.: Temperature measurements and thermal gradient estimates on the slope and shelf edge region of the Beaufort Sea, Canada, Geological Survey of Canada Open File, 2015.

Robbins, H. and Monro, S.: A stochastic approximation method, The annals of mathematical statistics, pp. 400–407, 1951.

Robinson, E. A. and Treitel, S.: Digital imaging and deconvolution: The ABCs of seismic exploration and processing, Society of Exploration Geophysicists, 2008.

Romanovskii, N., Hubberten, H.-W., Gavrilov, A., Eliseeva, A., and Tipenko, G.: Offshore permafrost and gas hydrate stability zone on the shelf of East Siberian Seas, Geo-marine letters, 25, 167–182, 2005.

Romanovskii, N. N. and Hubberten, H.-W.: Results of permafrost modelling of the lowlands and shelf of the Laptev Sea region, Russia, Permafrost and periglacial processes, 12, 191–202, 2001.

Ronneberger, O., Fischer, P., and Brox, T.: U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, pp. 234–241, Springer, 2015.

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

Ruder, S.: An overview of gradient descent optimization algorithms, arXiv preprint arXiv:1609.04747, 2016.

Ruffell, J., Murphy, T., and Graham, C.: Planning and execution of a 500 m corehole through offshore permafrost, in: Proceedings of the Fifth Canadian Conference on Permafrost, Quebec. Collect. Nord, vol. 54, pp. 271–282, 1990.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J.: Learning representations by back-propagating errors, nature, 323, 533–536, 1986.

Ruppel, C.: Permafrost-associated gas hydrate: Is it really approximately 1% of the global system?, Journal of Chemical & Engineering Data, 60, 429–436, 2015.

Ruppel, C. D., Herman, B. M., Brothers, L. L., and Hart, P. E.: Subsea ice-bearing permafrost on the US Beaufort Margin: 2. Borehole constraints, Geochemistry, Geophysics, Geosystems, 17, 4333–4353, 2016.

Sayedi, S. S., Abbott, B. W., Thornton, B. F., Frederick, J. M., Vonk, J. E., Overduin, P., Schädel, C., Schuur, E. A., Bourbonnais, A., Demidov, N., et al.: Subsea permafrost carbon stocks and climate change sensitivity estimated by expert assessment, Environmental Research Letters, 15, 124 075, 2020.

Sazli, M. H.: A brief review of feed-forward neural networks, Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering, 50, 2006.

Schmidhuber, J.: Deep learning in neural networks: An overview, Neural networks, 61, 85–117, 2015.

Schuur, E. A., Bockheim, J., Canadell, J. G., Euskirchen, E., Field, C. B., Goryachkin, S. V., Hagemann, S., Kuhry, P., Lafleur, P. M., Lee, H., et al.: Vulnerability of permafrost carbon to climate change: Implications for the global carbon cycle, BioScience, 58, 701–714, 2008.

Schuur, E. A., McGuire, A. D., Schädel, C., Grosse, G., Harden, J., Hayes, D. J., Hugelius, G., Koven, C. D., Kuhry, P., Lawrence, D. M., et al.: Climate change and the permafrost carbon feedback, Nature, 520, 171, 2015.

Semenov, P., Portnov, A., Krylov, A., Egorov, A., and Vanshtein, B.: Geochemical evidence for seabed fluid flow linked to the subsea permafrost outer border in the South Kara Sea, Geochemistry, 80, 125 509, 2020.

Serov, P., Portnov, A., Mienert, J., Semenov, P., and Ilatovskaya, P.: Methane release from pingo-like features across the South Kara Sea shelf, an area of thawing offshore permafrost, Journal of Geophysical Research: Earth Surface, 120, 1515–1529, 2015.

Shakhova, N., Semiletov, I., Gustafsson, O., Sergienko, V., Lobkovsky, L., Dudarev, O., Tumskoy, V., Grigoriev, M., Mazurov, A., Salyuk, A., et al.: Current rates and mechanisms of subsea permafrost degradation in the East Siberian Arctic Shelf, Nature Communications, 8, 1–13, 2017a.

Shakhova, N., Semiletov, I., Gustafsson, O., Sergienko, V., Lobkovsky, L., Dudarev, O., Tumskoy, V., Grigoriev, M., Mazurov, A., Salyuk, A., et al.: Current rates and mechanisms of subsea permafrost degradation in the East Siberian Arctic Shelf, Nature Communications, 8, 1–13, 2017b.

Sharma, S., Sharma, S., and Athaiya, A.: Activation functions in neural networks, Towards Data Sci, 6, 310–316, 2017.

Shearer, J., Macnab, R., Pelletier, B., and Smith, T.: Submarine pingos in the Beaufort Sea, Science, 174, 816–818, 1971.

Sheriff, R. E.: Encyclopedic dictionary of applied geophysics, Society of exploration geophysicists, 2002.

Sherman, D., Kannberg, P., and Constable, S.: Surface towed electromagnetic system for mapping of subsea Arctic permafrost, Earth and Planetary Science Letters, 460, 97–104, 2017.

Siahkoohi, A., Louboutin, M., and Herrmann, F. J.: The importance of transfer learning in seismic modeling and imaging, Geophysics, 84, A47–A52, 2019.

Signs, V.: Arctic Report Card: Update for 2022, Atlantic, 1970, 0, 2022.

Simon, J., Fabien-Ouellet, G., Gloaguen, E., Khurjekar, I., and Araya-Polo, M.: Velocity model building using transfer learning, in: SEG/AAPG/SEPM First International Meeting for Applied Geoscience & Energy, OnePetro, 2021.

Simon, J., Fabien-Ouellet, G., Gloaguen, E., and Khurjekar, I.: Hierarchical transfer learning for deep learning velocity model building, Geophysics, 88, R79–R93, 2023.

Simonyan, K. and Zisserman, A.: Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556, 2014.

Sippel, J., Scheck-Wenderoth, M., Lewerenz, B., and Kroeger, K. F.: A crust-scale 3D structural model of the Beaufort-Mackenzie Basin (Arctic Canada), Tectonophysics, 591, 30–51, 2013.

Sippel, J., Scheck-Wenderoth, M., Lewerenz, B., and Klitzke, P.: Deep vs. shallow controlling factors of the crustal thermal field–insights from 3D modelling of the Beaufort-Mackenzie Basin (Arctic Canada), Basin Research, 27, 102–123, 2015.

Sloan Jr, E. D. and Koh, C. A.: Clathrate hydrates of natural gases, CRC press, 2007.

Soviany, P., Ionescu, R. T., Rota, P., and Sebe, N.: Curriculum learning: A survey, International Journal of Computer Vision, 130, 1526–1565, 2022.

Spitz, S.: Seismic trace interpolation in the FX domain, Geophysics, 56, 785–794, 1991.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M. M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M.: Climate change 2013: the physical science basis, Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change, 1535, 2013.

Sullivan, T. D., Parsekian, A. D., Sharp, J., Hanke, P. J., Thalasso, F., Shapley, M., Engram, M., and Walter Anthony, K.: Influence of permafrost thaw on an extreme geologic methane seep, Permafrost and Periglacial Processes, 32, 484–502, 2021.

Szandała, T.: Review and comparison of commonly used activation functions for deep neural networks, Bio-inspired neurocomputing, pp. 203–224, 2021.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A.: Going deeper with convolutions, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1–9, 2015.

Taillandier, C., Noble, M., Chauris, H., and Calandra, H.: First-arrival traveltime tomography based on the adjoint-state method, Geophysics, 74, WCB1–WCB10, 2009.

Tan, M. and Le, Q.: Efficientnet: Rethinking model scaling for convolutional neural networks, in: International conference on machine learning, pp. 6105–6114, PMLR, 2019.

Tarantola, A.: Inverse problem theory and methods for model parameter estimation, SIAM, 2005.

Taylor, A. E., Dallimore, S., Hill, P., Issler, D., Blasco, S., and Wright, F.: Numerical model of the geothermal regime on the Beaufort Shelf, arctic Canada since the Last Interglacial, Journal of Geophysical Research: Earth Surface, 118, 2365–2379, 2013.

Temizel, A. et al.: Paired 3d model generation with conditional generative adversarial networks, in: Proceedings of the European Conference on Computer Vision (ECCV) Workshops, pp. 0–0, 2018.

Thomas, M., du Bois d'Aische, E., Moenaert, T., Villani, M., Hirst, C., Lundin, E., Jonard, F., Lambot, S., Van Oost, K., Vanacker, V., et al.: Combining geophysical data, microtopography, and very-high resolution UAV imagery to map lowland permafrost degradation in the Stordalen mire, Abisko, Sweden., in: 6th European Conference on Permafrost, 2023.

Tomaškovičová, S. and Ingeman-Nielsen, T.: Quantification of freeze–thaw hysteresis of unfrozen water content and electrical resistivity from time lapse measurements in the active layer and permafrost, Permafrost and Periglacial Processes, 2023.

Torrey, L. and Shavlik, J.: Transfer learning, in: Handbook of research on machine learning applications and trends: algorithms, methods, and techniques, pp. 242–264, IGI global, 2010.

Tourei, A., Ji, X., dos Santos, G. F. R., Czarny, R., Rybakov, S., Wang, Z., Hallissey, M., Martin, E., Xiao, M., Zhu, T., et al.: Mapping Permafrost Variability and Degradation Using Seismic Surface Waves, Electrical Resistivity and Temperature Sensing: A Case Study in Arctic Alaska, Journal of Geophysical Research: Earth Surface, 2023.

Valueva, M. V., Nagornov, N., Lyakhov, P. A., Valuev, G. V., and Chervyakov, N. I.: Application of the residue number system to reduce hardware costs of the convolutional

neural network implementation, Mathematics and computers in simulation, 177, 232–243, 2020.

Van Everdingen, R. O.: Multi-language glossary of permafrost and related ground-ice terms, International Permafrost Association, https://globalcryospherewatch.org/reference/glossary_docs/Glossary_of_Permafrost_and_Ground-Ice_IPA_2005.pdf, 2005.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I.: Attention is all you need, Advances in neural information processing systems, 30, 2017.

Virieux, J. and Operto, S.: An overview of full-waveform inversion in exploration geophysics, Geophysics, 74, WCC1–WCC26, 2009.

Vonk, J. E., Tank, S. E., Bowden, W. B., Laurion, I., Vincent, W. F., Alekseychik, P., Amyot, M., Billet, M., Canário, J., Cory, R. M., et al.: Reviews and syntheses: Effects of permafrost thaw on Arctic aquatic ecosystems, Biogeosciences, 12, 7129–7167, 2015.

Vosoughi, E., Giroux, B., Duchesne, M. J., and Christian Dupuis, J.: Simultaneous acoustic, electrical, and X-ray computed tomography laboratory measurements of partially saturated permafrost degradation, in: First International Meeting for Applied Geoscience & Energy, pp. 2258–2262, Society of Exploration Geophysicists, 2021.

Waldeland, A. U., Jensen, A. C., Gelius, L.-J., and Solberg, A. H. S.: Convolutional neural networks for automated seismic interpretation, The Leading Edge, 37, 529–537, 2018.

Walker, J. and Stuart, A.: Permafrost Investigations By Crystal Cable Surveys Mackenzie Delta, NWT, in: SPWLA Annual Logging Symposium, pp. SPWLA–1976, SPWLA, 1976.

Walvoord, M. A. and Kurylyk, B. L.: Hydrologic impacts of thawing permafrost—A review, Vadose Zone Journal, 15, 2016.

Walvoord, M. A. and Striegl, R. G.: Increased groundwater to stream discharge from permafrost thawing in the Yukon River basin: Potential impacts on lateral export of carbon and nitrogen, Geophysical Research Letters, 34, 2007.

Wang, J., Stewart, R. R., Dyaur, N. I., and Bell, M. L.: Marine guided waves: Subbottom property estimation and filtering using physical modeling data Marine guided waves, Geophysics, 81, V303–V315, 2016.

Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E.: Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing, 13, 600–612, https://doi.org/10.1109/TIP.2003.819861, 2004.

Wang, Z., Wang, S., Zhou, C., and Cheng, W.: Dual Wasserstein generative adversarial network condition: A generative adversarial network-based acoustic impedance inversion method, Geophysics, 87, R401–R411, 2022.

Wei, X., Zhang, C., Wang, H., Tan, C., Xiong, D., Jiang, B., Zhang, J., and Kim, S.-W.: Seismic Data Interpolation based on Denoising Diffusion Implicit Models with Resampling, arXiv preprint arXiv:2307.04226, 2023.

Winters, W. J., Pecher, I. A., Waite, W. F., and Mason, D. H.: Physical properties and rock physics models of sediment containing natural and laboratory-formed methane gas hydrate, American Mineralogist, 89, 1221–1227, 2004.

Winters, W. J., Waite, W. F., Mason, D., Gilbert, L., and Pecher, I.: Methane gas hydrate effect on sediment acoustic and strength properties, Journal of Petroleum Science and Engineering, 56, 127–135, 2007.

Wolfe, P.: Convergence conditions for ascent methods, SIAM review, 11, 226–235, 1969.

Wu, B., Meng, D., and Zhao, H.: Semi-supervised learning for seismic impedance inversion using generative adversarial networks, Remote Sensing, 13, 909, 2021.

Wu, G., Li, G., Cao, Y., Chen, D., Qi, S., Wang, F., Gao, K., Du, Q., Wang, X., Jing, H., et al.: Distribution and Degradation Processes of Isolated Permafrost near Buried Oil Pipelines by Means of Electrical Resistivity Tomography and Ground Temperature Monitoring: A Case Study of Da Xing'anling Mountains, Northeast China, Remote Sensing, 15, 707, 2023.

Wu, X., Geng, Z., Shi, Y., Pham, N., Fomel, S., and Caumon, G.: Building realistic structure models to train convolutional neural networks for seismic structural interpretation, Geophysics, 85, WA27–WA39, 2020.

Wu, Y., Nakagawa, S., Kneafsey, T. J., Dafflon, B., and Hubbard, S.: Electrical and seismic response of saline permafrost soil during freeze - Thaw transition, Journal of Applied Geophysics, 146, 16–26, https://doi.org/10.1016/j.jappgeo.2017.08.008, 2017.

Xie, Q., Luong, M.-T., Hovy, E., and Le, Q. V.: Self-training with noisy student improves imagenet classification, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 10 687–10 698, 2020.

Yamazaki, Y., Kubota, J., Ohata, T., Vuglinsky, V., and Mizuyama, T.: Seasonal changes in runoff characteristics on a permafrost watershed in the southern mountainous region of eastern Siberia, Hydrological Processes: An International Journal, 20, 453–467, 2006.

Yang, J., Price, B., Cohen, S., Lee, H., and Yang, M.-H.: Object contour detection with a fully convolutional encoder-decoder network, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 193–202, 2016.

Yang, J., Lin, N., Zhang, K., Zhang, C., Fu, C., Tian, G., and Song, C.: Reservoir characterization using multi-component seismic data in a novel hybrid model based on clustering and deep neural network, Natural Resources Research, 30, 3429–3454, 2021.

Yoshikawa, K.: 8.18 Pingos, in: Treatise on Geomorphology, edited by Shroder, J. F., pp. 274–297, Academic Press, San Diego, https://doi.org/https://doi.org/10.1016/B978-0-12-374739-6.00212-8, 2013.

Young, F. G., DW, M., and CJ, Y.: GEOLOGY OF THE BEAUFORT-MACKENZIE BASIN., in: Geological survey of canada, vol. 11, Supply and Services Canada, 1976.

Yuan, S., Jiao, X., Luo, Y., Sang, W., and Wang, S.: Double-scale supervised inversion with a data-driven forward model for low-frequency impedance recovery, Geophysics, 87, R165–R181, 2022.

Zeiler, M. D.: Adadelta: an adaptive learning rate method, arXiv preprint arXiv:1212.5701, 2012.

Zeiler, M. D. and Fergus, R.: Visualizing and understanding convolutional networks, in: Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13, pp. 818–833, Springer, 2014.

Zhang, J., Li, J., Chen, X., Li, Y., Huang, G., and Chen, Y.: Robust deep learning seismic inversion with a priori initial model constraint, Geophysical Journal International, 225, 2001–2019, 2021.

Zhang, T.: Talik, pp. 1143–1144, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-90-481-2642-2_563, 2011.

Zhang, T., Barry, R. G., Knowles, K., Heginbottom, J., and Brown, J.: Statistics and characteristics of permafrost and ground-ice distribution in the Northern Hemisphere, Polar Geography, 23, 132–154, 1999.

Zhang, T., Heginbottom, J., Barry, R. G., and Brown, J.: Further statistics on the distribution of permafrost and ground ice in the Northern Hemisphere, Polar geography, 24, 126–131, 2000.

Zhang, Z. and Lin, Y.: Data-driven seismic waveform inversion: A study on the robustness and generalization, IEEE Transactions on Geoscience and Remote sensing, 58, 6900–6913, 2020.

Zhang, Z.-D. and Alkhalifah, T.: Regularized elastic full-waveform inversion using deep learning, GEOPHYSICS, 84, R741–R751, https://doi.org/10.1190/geo2018-0685.1, 2019.

Zheng, Y., Zhang, Q., Yusifov, A., and Shi, Y.: Applications of supervised deep learning for seismic interpretation and inversion, The Leading Edge, 38, 526–533, 2019.