UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

A COMPARISON OF COHERENCE ENHANCEMENT, PROBABILISTIC NEURAL NETWORK, AND CONVOLUTIONAL NEURAL NETWORK FAULT IMAGING WITH MANUAL INTERPRETATION IN THE TARANAKI BASIN AREA, NEW ZEALAND.

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

JOSE PEDRO MORA

Norman, Oklahoma

2021

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/329553

A COMPARISON OF COHERENCE ENHANCEMENT, PROBABILISTIC NEURAL NETWORK, AND CONVOLUTIONAL NEURAL NETWORK FAULT IMAGING WITH MANUAL INTERPRETATION IN THE TARANAKI BASIN AREA, NEW ZEALAND.

A THESIS APPROVED FOR THE

SCHOOL OF GEOSCIENCES

BY THE COMMITTEE CONSISTING OF

Dr. Heather Bedle, Chair

Dr. Kurt J. Marfurt

Dr. Brett M. Carpenter

© Copyright by JOSE PEDRO MORA 2021

All Rights Reserved.

Acknowledgments

I would like to extend my deepest gratitude to the Saldarriaga concha foundation and the Fulbright program because they provide the means and help to people with different capacities. They gave me the needed connections to be at the University of Oklahoma and assemble immense efforts to help professionals reach their academic purposes.

I like to express gratitude and profound appreciation to my advisor Dr. Heather Bedle, for her guidance and unconditional support during my time here; her smile and positive energy helped me to focus on my priorities and objectives during my research. Thank you for teaching me how to be an outstanding geoscientist and an extraordinary human being.

I also want to express my appreciation to Dr. Kurt Marfurt; thank you for giving me invaluable contributions, suggestions, and guidance in this project. One does not always have the chance to come across a living legend and also an extraordinary person; thank you, Cacique.

I want to thank Dr. Brett Carpenter for his constructive advice during this semester and for being part of my committee. Dr. Brett Carpenter and Dr. Michael Behm, thank you for teaching me that learning is not a one-way communication practice but an interactive process with the student; thank you for your teachings and bits of advice.

Many thanks to the OU geosciences faculty and staff, Dr. Lyn Soreghan, Rebecca Fay, Leah Moser, Ginger Leivas, and Ashley Tullius, for their kindness and guide during my time here. To the SEG chapter and geoscience students for their company and good moments, thanks to Roberto

Clairmont, Julian Chenin, Carl Buist, Clayton Silver, Abidin Caf, David Lubo, Thang Ha, Kurt Crandall, and all of you who were part of this learning process.

To the Colombian student association and to Yoana Walschap, Javier Tellez, Karelia La Marca, David Duarte, Laura Ortiz, Jhouly Osorio for their support and friendship, I am proud to be part of this small part of Colombia here at OU; thank you for letting me be part of this vibrant community during the past two years.

I want to thank all the friends who share their time and affection with my family and me: Karen Leopoldinho, Edimar Perico, Claudia Boza, Ahmet Murat, Alexandro Vera, Erika Comparan, Cesar Vivas, Lina Perdomo and family; thank you for being here as a company and for your friendship; I don't have the words to express how grateful I am to you, you always worried about me, and I give thanks to God for knowing you.

Thank you to all my friends in Colombia, Nestor Moreno, Laura Ortega, Paola Florez, and Carlos Zambrano, for their support and help before my travel to Oklahoma. I also want to express my gratitude to Jorge Eduardo Pinto; he passed away a couple of days ago while writing this manuscript; he marked my professional life, I called him 'Profe', not because he was an actual teacher of mine, but he was someone who taught me to be a professional and human being and earned that title, thank you.

And finally, I want to express my love to my family and relatives because they are the energy to keep me moving and the reason to move forward. To my sister, who helped me at the beginning of this dream, and to her family because they lent her to be here with me. To my dad because his love gives me reasons to work harder. To my brother, who joined me here and gave us his

precious company and love during these difficult COVID-19 times. To my grandparents because they pray daily for us, and I hope to make them proud of me. Finally, to my mom, because she supports me and believes that I can reach new highs and fulfill my dreams. Thank you for being here and for your strength and love that gave me the reasons to believe and keep my fight during my graduate studies.

Table of Contents

ACKNOWLEDGMENTS ..... **IV**

TABLE OF CONTENTS ..... **VII**

LIST OF TABLES ..... **X**

LIST OF FIGURES ..... **XI**

ABSTRACT ..... **XVIII**

CHAPTER 1. INTRODUCTION **1**

MOTIVATION AND OBJECTIVES 1

FUNDAMENTALS 3

Coherence Enhancement *5*

Probabilistic neural networks (PNN) *6*

Convolutional neural networks *6*

Fault object extraction *7*

Active contour method *7*

GEOLOGICAL SETTING 8

SUMMARY 10

REFERENCES 11

CHAPTER 2. FAULT ENHANCEMENT USING PROBABILISTIC NEURAL NETWORKS AND ATTRIBUTE ENHANCEMENT IN SEISMIC IMAGES; A CASE STUDY IN THE GREAT SOUTH BASIN, NEW ZEALAND. **16**

ABSTRACT 16

INTRODUCTION 17

GEOLOGICAL SETTING 20

DATASETS 23

METHODOLOGY 23

Data preconditioning *25*

Probabilistic neural networks *26*

PNN training and classification *27*

Fault enhancement *28*

Attributes *28*

RESULTS 29

DISCUSSION 34

CONCLUSIONS 37

ACKNOWLEDGMENTS 39

REFERENCES 40

CHAPTER 3. FAULT ENHANCEMENT COMPARISON BETWEEN COHERENCE ENHANCEMENT, PROBABILISTIC NEURAL NETWORKS, AND CONVOLUTIONAL NEURAL NETWORKS IN THE TARANAKI BASIN AREA, NEW ZEALAND. 46

ABSTRACT 46

INTRODUCTION 47

GEOLOGICAL SETTING 51

Cape Egmont Fault Zone *53*

DATASETS 54

METHODOLOGY 55

Coherence enhancement 56

Probabilistic neural networks 57

Convolutional neural network 58

Fault extraction and comparison 60

RESULTS 60

Coherence enhancement 65

Probabilistic neural networks 66

Convolutional neural network 66

Fault extraction and comparison 68

DISCUSSION 69

CONCLUSIONS 75

ACKNOWLEDGMENTS 78

REFERENCES 79

CHAPTER 4. CONCLUSIONS 86

REFERENCES 89

APPENDIX A. PROBABILISTIC NEURAL NETWORKS DATA SCALING, SAMPLING SELECTION AND OPTIMUM ATTRIBUTE SELECTION 90

APPENDIX B. COMPARISON OF FAULT PREDICTION USING COHERENCE ENHANCEMENT, PROBABILISTIC NEURAL NETWORKS, AND CONVOLUTION NEURAL NETWORKS 93

List of Tables

TABLE 1.1. SEISMIC ATTRIBUTE CATEGORIES AND CORRESPONDING TYPES AND INTERPRETIVE USES, MODIFIED FROM RODEN AND SACREY (RODEN ET AL., 2015). 4

TABLE 3.1. BASIC STATISTIC MEASUREMENTS AFTER HAUSDORFF DISTANCE CALCULATIONS ON THREE DIFFERENT FAULTS, ALL MEASURED VALUES ARE POSITIVE DISTANCES FROM THE FAULT TO ANY POSITION ON THE OTHER SURFACE, DISTANCE IN METERS WITH A CONVERSION TIME TO DISTANCE USING A 3000 M/S VELOCITY. 68

TABLE 3.2. COMPUTATIONAL TIME EMPLOYED IN THE ENTIRE WORKFLOW EMPLOYING 60 PROCESSORS AND 256 GB RAM. PNN IS THE MOST COMPUTATIONALLY INTENSIVE METHOD TEN TIMES LONGER THAN THE CNN ANALYSIS. CNN IS THE SECOND TIME-CONSUMING, BUT THE ENERGY RATIO SIMILARITY CALCULATION ADDS 1.5 HOURS TO THE CLASSIFICATION, REACHING A TOTAL TIME OF 3 HOURS. THESE VALUES ARE ESTIMATED ON A 3.5 GB VOLUME ANALYSIS. 71

TABLE A.1. A LIST OF CANDIDATE PNN INPUT ATTRIBUTES AND SELECTED ATTRIBUTES USED AND THE CORRESPONDING SMOOTHING PARAMETER, R, USED IN EQUATION 1 OBTAINED FOR THE THREE NORMALIZATION STRATEGIES EVALUATED. 92

List of Figures

FIGURE 1.1. PROPOSED QUESTIONS FOR THIS RESEARCH. 3

FIGURE 1.2. GENERAL MAP OF NEW ZEALAND SHOWING PLATE TECTONIC BOUNDARIES AS THEY RELATE TO THE SOUTH AND NORTH ISLANDS. THE ANALYSIS AREAS ARE LOCATED IN THE GREAT SOUTH BASIN (CHAPTER 2) AND THE TARANAKI BASIN (CHAPTER 3). 9

FIGURE 2.1. GENERAL MAP OF (A) NEW ZEALAND AND (B) GREAT SOUTH BASIN. THE GENERAL MAP SHOWS PLATE TECTONICS BOUNDARIES INCLUDING THE STRIKE-SLIP FAULT CUTTING THE SOUTH ISLAND. (B) MAP OF THE SOUTH ISLAND SHOWING THE LOCATION OF THE GSB SEISMIC VOLUME. 21

FIGURE 2.2 REPRESENTATIVE (A) VERTICAL AND (B) TIME SLICES THROUGH THE 3D GSB SEISMIC AMPLITUDE VOLUME. BLACK ARROWS INDICATE SEVERAL FAULTS ON BOTH SLICES. THE ORANGE ARROW INDICATES A MASS TRANSPORT COMPLEX. (C) LEGEND SHOWING THE EXTENT AND LOCATION OF THE VERTICAL AND TIME SLICE IN THE LARGER GSB SURVEY. (DATA COURTESY OF NZP&M). 22

FIGURE 2.3. PNN AND FAULT ENHANCEMENT WORKFLOW. 24

FIGURE 2.4. CANDIDATE ATTRIBUTES USED AS INPUT IN THE PNN FAULT ENHANCEMENT PROCESS PRIOR TO DETERMINING THE BEST COMBINATION AS SHOWN IN FIGURE 2.3: (A) ENERGY RATIO SIMILARITY, (B) ABERRANCY MAGNITUDE, (C) DIP MAGNITUDE, (D) GLCM HOMOGENEITY, (E) GLCM ENTROPY, (F) ENVELOPE, (G) CURVEDNESS, (H) SOBEL FILTER SIMILARITY, AND (I) TOTAL ENERGY. TIME SLICES ARE AT T= 1.28 S. 30

FIGURE 2.5. THE IMPACT OF TRAINING DATA ON PNN FAULT CLASSIFICATION. (A) AMPLITUDE SECTION AND KEY REPRESENTATIVE TRAINING SELECTED, BLUE LINES ARE HAND-PICKED FAULT SAMPLES, ORANGE AND BLACK ARROWS SHOW NO-FAULT SELECTED AREAS AS REPRESENTATIVE TRAINING/VALIDATION IN THIS SECTION. (B) THE FAULT PREDICTION FOR A VERTICAL SLICE WITH TRAINING SAMPLES WITH LITTLE SEISMIC NOISE AND NO

significant MTD facies. Note the artifacts and no clean faults in the section. The image below (b) shows a representative sample training set, fewer artifacts, and better classification across the section (yellow arrow); still, some areas show mass transport complexes (red arrow). 31

Figure 2.6. The impact of data normalization on the resulting PNN fault prediction. (a) The robust scaler normalization shows an increased amount of noise and misclassification of low reflectivity areas as faults. (b) The logarithm scaler works well, and a cleaner result is achieved. (c) Counterintuitively, the Z-score normalization method shows the best image, consistent with findings by Ha et al., (2021) use of PNN to delineate salt facies. Yellow arrows indicate faults, and red arrows the absence of faults. Appendix Table 1.1 shows the input and selected attributes. 33

Figure 2.7. PNN fault classification results. (a) shows fault enhancement (Machado et al., 2016; Qi et al., 2019b) using the Z-score normalization result as input (Figure 2.6c). (b) Result after skeletonization described by Qi et al. (2019b). (c) Amplitude corendered with PNN results. Yellow arrows indicate areas of improved fault continuity. Red arrows indicate suppressed artifacts. 36

Figure 2.8 PNN classification comparison and RGB blended attributes. (a) PNN classification section and key representative areas showing faults as yellow arrows and mass transport complexes marked with blue arrows. (b) A blended image using the RGB color model. R corresponds to multispectral energy ratio similarity, G to aberrancy magnitude, and B to dip magnitude. The image below (c) shows a) A blended image using the RGB color model. R corresponds to multispectral energy ratio similarity, G to Sobel filter similarity, and B to dip magnitude. PNN

shows some enhanced features in the red arrow area compared to the RGB blending in the two other cases. 38

Figure 3.1. (a) General map of the Taranaki Basin showing major structural areas, faults, and sediment thickness. (After Strogen et al., 2014). The Western stable platform is mainly layer cake depositional environment, and the Eastern mobile belt is a more complex and active setting with active, reverse faults, grabens, and inversion structures (King and Thrasher, 1996; Kroeger et al., 2013; Haque et al., 2016; Reilly et al., 2016; Coleman, 2018). (b) Map of the Southern Taranaki Basin showing faults and petroleum fields showing sediment thickness near the analysis area (After Reilly et al., 2016). 52

Figure 3.2. (a) Merged 3D seismic surveys included in the 3D Trestles program (After Gong and Fell, 2016). The red square marks the area analyzed in this paper. (b) Representative vertical slice AA' through the merged seismic data volume showing an intricate faulting pattern to the east and additional faulting to the west. 55

Figure 3.3. Summary of our proposed workflow. We begin with postmigration data conditioning to suppress crosscutting noise and balance the spectrum and manually pick the main faults to form a baseline human interpretation. Then we evaluate three different fault delineation workflows. The results of each workflow are enhanced to provide skeletonized fault probability, fault dip magnitude, and fault dip azimuth; then, we use active contours to construct fault objects. Finally, the three fault predictions are compared with the manual baseline to determine which method provides the best image. We expect the best method to be close to the interpreted surface. 59

FIGURE 3.4. COMPARISON BETWEEN ATTRIBUTES COMPUTED FROM THE SEISMIC AMPLITUDE VOLUME (A) BEFORE AND (B) AFTER DATA CONDITIONING. THE CONDITIONED DIP MAGNITUDE IS PIECEWISE SMOOTHER AND SHOWS LESS FOOTPRINT (YELLOW ARROWS). THE CONDITIONED ENERGY RATIO SIMILARITY EXHIBITS FEWER INCOHERENT FEATURES (GREEN ARROWS) WHILE ENHANCING THE STRONGER FAULT STRATIGRAPHIC EDGES; LIKEWISE, THE CONDITIONED TOTAL ABERRANCY EXHIBITS FEWER ARTIFACTS (RED ARROWS). 62

FIGURE 3.5. FAULT ENHANCEMENT RESULTS USING NON-CONDITIONED DATA SHOWING COHERENCE ENHANCEMENT (A), PNN (B), AND CNN (C) RESULT IN A SEISMIC SECTION (IL 2400) AND TIME SLICE (0.76 S) WITH FAULT AMPLITUDE DATA AND CORENDERED FAULT SKELETONIZATION. COHERENCE ENHANCEMENT RESULTS (A) SUGGEST THE METHOD CLASSIFIES MINOR FRACTURES AND FEATURES, BUT STAIRS-STEPS AND DISCONTINUITIES BETWEEN FAULTS ARE STILL EVIDENT. PNN (B) RESULTS SHOW MORE SUBTLE FEATURES THAN COHERENCE ENHANCEMENT WITH A SALT-AND-PEPPER-LIKE RESULT. CNN (C) EXTRACTS LARGE FAULTS AND RELEVANT STRUCTURES, AVOIDING SMALL FRACTURES AND FAULTS. 63

FIGURE 3.6. FAULT ENHANCEMENT RESULTS USING CONDITIONED DATA SHOWING COHERENCE ENHANCEMENT (A), PNN (B), AND CNN (C) RESULTS IN A SEISMIC SECTION (IL 2400) AND TIME SLICE (0.76 S) WITH FAULT AMPLITUDE DATA AND CORENDERED FAULT SKELETONIZATION. COHERENCE ENHANCEMENT RESULTS (A) SUGGEST THE METHOD CLASSIFIES MINOR FRACTURES AND FEATURES, BUT STAIRS-STEPS AND DISCONTINUITIES BETWEEN FAULTS ARE STILL EVIDENT. PNN (B) RESULTS SHOW MORE SUBTLE FEATURES THAN COHERENCE ENHANCEMENT WITH A SALT-AND-PEPPER-LIKE RESULT. CNN (C) EXTRACTS LARGE FAULTS AND RELEVANT STRUCTURES, AVOIDING SMALL FRACTURES AND FAULTS. THE CONDITIONED RESULTS SHOW LESS NOISE AND SMALL FEATURES IN THE THREE METHODS. PNN CLASSIFICATION SHOWS ABERRANCY MAGNITUDE, MULTISPECTRAL ENERGY RATIO, AND SOBEL FILTER SIMILARITY (0-34 HZ) AS THE BEST ATTRIBUTE COMBINATION. 64

FIGURE 3.7. COMPARISON BETWEEN EXTRACTED SURFACES AND MANUAL INTERPRETATION. THE HAUSDORFF DISTANCE (CIGNONI ET AL., 1998) CALCULATES THE MAXIMUM DISTANCE BETWEEN TWO SURFACES IN EVERY DIRECTION. COLOR-CODED DIFFERENCES HELP TO IDENTIFY DISCREPANCIES BETWEEN MANUAL AND FAULT OBJECTS FROM CLASSIFICATION METHODS. THE IMAGE SHOW DISTANCE IN METERS WITH A CONVERSION TIME TO DISTANCE USING A 3000 M/S VELOCITY. 67

FIGURE 3.8. A SEISMIC SECTION NEAR THE MAUI-3 WELL. DIFFERENCES IN SEISMIC REFLECTIONS INFLUENCE THE FINAL RESPONSE OF FAULT ENHANCEMENT METHODS AND CAN BE BROADLY RELATED TO THESE FOUR ZONES; NOTICE THE BLACK ARROW SHOWING A DISTORTED VISUALIZATION OF THE FOOTWALL. 70

FIGURE 3.9. TIME SLICES FROM MULTISPECTRAL COHERENCE ANALYSIS, ENERGY RATIO (A), AND SOBEL FILTER SIMILARITY (B). HIGHER FREQUENCIES SHOW LESS COHERENT FEATURES, AND LOW FREQUENCIES RESEMBLE BETTER AND RELEVANT FAULTS. WE USE MULTISPECTRAL, 0 TO 34 HZ., 5 TO 63 HZ. AND 33 TO 91 HZ. VOLUMES IN THE PNN ANALYSIS TO FIND THE BEST COMBINATION. 73

FIGURE 3.10. DETAILED FAULT ENHANCEMENT FROM DIFFERENT WORKFLOWS (A) AND SEISMIC SECTION AND TIME SLICE LOCATION OF THE DETAILED AREA (B). NOTICE THE COMPLEX SECTION WITH CHAOTIC REFLECTOR AND SEMI-CONTINUOUS SUBPARALLEL REFLECTORS ON AMPLITUDE DUE TO CHANNELIZATION, (A) COHERENCE ENHANCEMENT SHOW VALUES CLOSE TO THE FAULT SUCH AS PNN, CNN HAS CLEANER RESULT BUT PLENTY OF ANOMALIES AND ARTIFACTS. PNN IS THE METHOD THAT SHOWS THE MOST INFERIOR RESULT DUE TO HIGH NOISE AND VERTICAL ARTIFACTS. 74

FIGURE 3.11. COMPARISON BETWEEN REVIEWED METHODS. THIS FIGURE SHOWS THE MAIN DIFFERENCES BETWEEN ANALYSIS TYPE, INTERPRETATION USE, PROCESSING TIMES, ATTRIBUTE CALCULATION PRODUCTS, AND MAIN DRAWBACKS. ALL MEASURED VALUES ARE POSITIVE DISTANCES FROM THE FAULT TO ANY POSITION ON THE

OTHER SURFACE, DISTANCE IN METERS WITH A CONVERSION TIME TO DISTANCE USING A 3000 M/S VELOCITY; THE INITIAL COLUMN SHOWS THE INITIAL APPROACH TO MAP FAULTS MANUALLY. 77

FIGURE A.1 HISTOGRAMS OF THE TRAINING DATASET AFTER Z-SCORE NORMALIZATION. 90

FIGURE A.2 HISTOGRAMS OF THE TRAINING DATASET AFTER ROBUST NORMALIZATION SCALING. 91

FIGURE A.3 HISTOGRAMS OF THE TRAINING DATASET AFTER LOGARITHMIC NORMALIZATION SCALING. 91

FIGURE B.1 THE MULTISPECTRAL COHERENCE ENHANCEMENT WORKFLOW. FIRST, WE APPLY A SPECTRAL BALANCING ALGORITHM TO FLATTEN THE SPECTRUM, THEREBY INCREASING THE BANDWIDTH AND VERTICAL RESOLUTION. NEXT, WE APPLY A STRUCTURAL ORIENTED FILTER TO REMOVE UNWANTED CROSSCUTTING COHERENT AS WELL AS RANDOM NOISE. THEN WE COMPUTE MULTISPECTRAL SIMILARITY (LI ET AL., 2018) VOLUMES USING FIVE OVERLAPPING FILTER BANKS. THEN, FOLLOWING QI ET AL. (2018), THE FAULTS ARE THEN ENHANCED USING A DIRECTIONAL LAPLACIAN OF A GAUSSIAN (LOG) FILTER PRIOR TO A FINAL SKELETONIZATION PROCESS. 93

FIGURE B.2 CNN FAULT PREDICTION AND ENHANCEMENT WORKFLOW. AS WITH THE WORKFLOW IN FIGURE B3, WE SPECTRALLY BALANCE THE DATA AND THEN APPLY A STRUCTURE-ORIENTED FILTER. THE CNN ALGORITHM EMPLOYS SYNTHETIC DATA AND FAULT LABELS DESCRIBED BY WU (2019) TO TRAIN A CONVOLUTIONAL NEURAL NETWORK, WHICH WE THEN USE TO PREDICT FAULTS. THE CNN-PREDICTED FAULTS ARE THEN FURTHER ENHANCED USING THE SAME WORKFLOW AS IN FIGURE B1. 94

FIGURE B.3 PNN FAULT PREDICTION AND ENHANCEMENT WORKFLOW. AS WITH THE WORKFLOW IN FIGURE B3, WE APPLY DATA CONDITIONING BEFORE ANY FAULT ENHANCEMENT PROCESS. ATTRIBUTE CALCULATION FOLLOWS DATA PRECONDITIONING. THEN, FAULT DIGITALIZATION TO SAMPLE RELEVANT FAULT AND NO-FAULT FEATURES TO SPLIT THEM INTO TRAINING AND VALIDATION DATASETS. PNN ALGORITHM ANALYZES THE INPUT DATA AND FINDS OPTIMAL PARAMETERS IN THE TESTED ATTRIBUTES. FINALLY, AN ATTRIBUTE COMBINATION IS DEFINED TO

classify the seismic dataset (Lubo-Robles et al., 2020). The last steps include a medium filter, and then the PNN faults are then further enhanced using the same workflow as in Figure B1. 95

Figure B.4 Vertical slices along crossline 2450 through corendered seismic amplitude and CNN fault prediction computed from (a) the original amplitude data volume (b) the data after structure-oriented filtering, and (c) the data after structure-oriented filtering and footprint suppression. The blue arrow indicates improvement in a complex area and the black arrow points features' continuity improvement with data conditioning. 96

Figure B.5 Time slices at t=0.76 s through co-rendered seismic amplitude and CNN fault prediction computed from (a) the original amplitude data volume (b) the data after structure-oriented filtering, and (c) the data after structure-oriented filtering and footprint. Green arrows indicate improved areas due to preconditioning. 97

Abstract

Fault identification is a key aspect in seismic interpretation, and model generation is an essential task to achieve any accurate, precise reservoir modeling or structural framework that solves problems in the subsurface. Manual interpretation is a common approach to map faults in seismic amplitude datasets; in addition, interpreters typically use specialized seismic attribute volumes, such as variance or coherence, to enhance fault visualization and ease their work of identifying faults. In addition to these attributes, specialized, enhanced attributes and machine learning techniques are recently being employed to enhance faults and structural features. Methods such as ant tracking, coherence enhancement, self-organized maps, convolutional neural networks, and additional data analytics methods are used to enhance faults; nonetheless, there is additional interpretation work to be done to create a fault object and differentiate faults from other complex features and noise inside seismic volumes.

Recent advances in probabilistic neural networks (PNN) have shown promise in delineating seismic facies, including salt and mass transport deposits from the surrounding sedimentary matrix. In the first part of this research, I evaluate the ability of PNN to delineate faults. Probabilistic neural networks (PNN) are feedforward neural networks often used in identification problems. Although almost all seismic attributes are in some way sensitive to faults, a much smaller subset highlight faults with respect to the non-faulted background geology. For this reason, I employ an exhaustive PNN search to identify the optimal set of attributes to create a fault probability volume. For a Great South Basin, New Zealand seismic search I found that the

best attributes to use were aberrancy magnitude, GLCM entropy and homogeneity, Sobel filter similarity, and envelope.

While time-consuming, handpicking of faults on vertical slices through the seismic amplitude volume to generate fault “sticks” and then interpolating those fault sticks to generate a fault surface is still the most commonly used fault interpretation workflow. In the second part of this research, I use such carefully hand-picked faults as the standard, and compare these faults to those computed using coherence enhancement, probabilistic neural networks (PNN), and convolutional neural networks (CNN) for a Taranaki Basin data volume. Coherence enhancement, PNN, and CNN all provide a voxel-by-voxel likelihood of a fault, not a fault surface. To address this issue, I use active contours to convert voxel estimates of fault probability into a significantly smaller set of samples that define a named fault object. This method is a semiautomatic approach that scans high probability fault locations and moves from voxels with low to high probability to fit the fault surface's shape. Although each fault enhancement method produces fault probability volumes, differences in noise, artifacts, and fault locations are found. I compute the Hausdorff difference between the three fault delineation workflows and the hand-picked baseline and find that multispectral coherence enhancement produces the most accurate results for the Taranaki Basin dataset analyzed. I also find that data conditioning prior to attribute calculation improves the fault enhancement result. I also show that data conditioning, including structure-oriented filtering, facilitates human interpretation and provides computer-generated faults that exhibit fewer artifacts.

Chapter 1. Introduction

Fault identification is a critical task in geologic modeling and is the initial step for creating a stratigraphic analysis in seismic volumes and provides the structural framework for subsequent interpretation. Although not always followed, a best practice is to first pick the major faults prior to picking any stratigraphic horizons to warn either the autopicker or the interpreter that a discontinuity will be encountered. Many tools are used to aid seismic interpreters; some focused on horizons, but nowadays, specialized attributes and machine learning techniques are being implemented to detect faults or salt bodies. This thesis aims to compare fault enhancement methods, including specialized and machine learning attributes, by extracting fault objects and comparing them to each other, in order to identify the differences and advantages for each technique.

Motivation and objectives

It has long been established that fault interpretation is improved by using geometric attributes such as coherence, similarity, etc. (Bahorich and Farmer, 1995; Marfurt et al., 1998; Gersztenkorn and Marfurt, 1999; Chopra and Marfurt, 2005; Barnes, 2016). Additional research into fault identification has calculated probability volumes using image processing techniques (Barnes, 2006; Machado et al., 2016; Qi et al., 2018, 2019b). Recent data analytic methods use probabilistic neural networks (PNN), convolutional neural networks (CNN), or self-organized maps (SOM) to identify facies and faults (Zhao et al., 2015; Wu et al., 2018; Xiong et al., 2018;

Chopra et al., 2019; Lubo-Robles et al., 2019; Hussein et al., 2020). There are several specialized attributes and workflows commonly used to enhance fault visualization; one of my objectives is to better define the commonalities and differences between these methods. While there are many tools and methodologies to compare these methods, this research uses fault object extraction as a basis for comparison.

Object extraction is a method of generating objects from a seismic image. For this study, fault objects are extracted and compared to a manually mapped fault. Admasu et al. (2006), Cohen et al. (2006), and Hale (2013) describe several ways of extracting such a fault object from attributes computed from 3D seismic amplitude volumes, and some software tools include these capabilities.

To further establish the motivation behind this thesis, I pose a few questions (Figure 1.1). It is established that computer-intensive tasks such as attribute calculation, data analysis, and CNN can be used to aid the interpreter, but are these tools capable of detecting faults as completely as a human interpreter? What differences can be found between fault enhancement methods and a manual interpreter? How do fault enhancement methods compare in computational time and accuracy? Is fault object extraction an effective way to compare alternative fault enhancement methods? This thesis aims to address these questions by comparing fault enhancement methods by extracting fault objects from each method, and comparing them to the baseline human-interpreted result to identify differences and advantages for each technique.

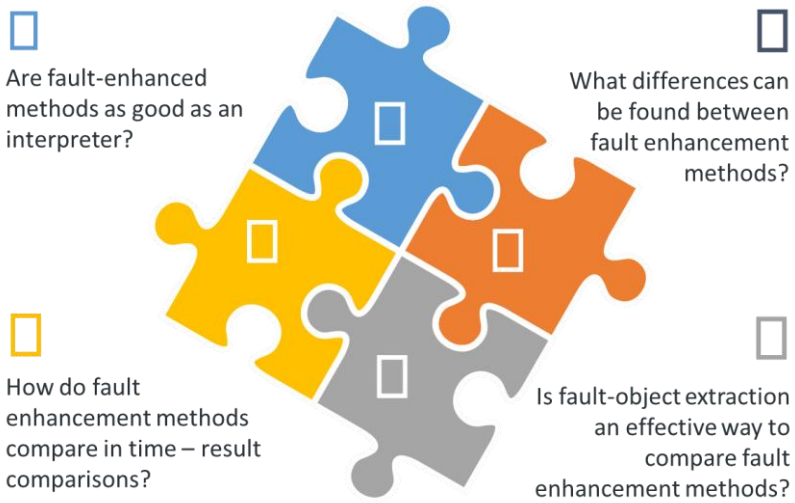


Figure 1.1. Proposed questions for this research.

Fundamentals

Seismic attributes are a quantitative measure of a seismic characteristic of interest, based on various calculations of the seismic trace (Chopra and Marfurt, 2005; Barnes, 2016). The seismic trace is a measured seismic signal which represents the changes in acoustic and shear impedances of a boundary between rocks, from which geologic subsurface interpretations are derived. This seismic response can be recorded and mapped in two, three, or more dimensions (frequency, incident angle, azimuth) and analyzed in a spatial context to create new derived products and estimations (Chopra and Marfurt, 2005; Barnes, 2016). These newly derived volume products are referred to as seismic attributes.

Several attributes are used as tools to interpret geology, structures, stratigraphy, rock and fluid properties, geometry, etc. Faults and structural features are usually enhanced using geometric

attributes such as semblance, coherence, entropy, curvature, and others (Chopra and Marfurt, 2005; Cohen et al., 2006; Barnes, 2016).

Table 1.1. Seismic attribute categories and corresponding types and interpretive uses, Modified from Roden and Sacrey (Roden et al., 2015).

Interpretive use	Type	Category
Lithology Contrasts, Bedding Continuity, Porosity, DHIs, Stratigraphy, Thickness	Reflection Strength, Instantaneous Phase, Instantaneous, Frequency, Quadrature, Instantaneous Q.	Instantaneous Attributes
Faults, Fractures, Folds, Anisotropy, Regional Stress fields	**Semblance and Eigen-Based Coherency/Similarity, Curvature (Maximum, Minimum, Most Positive, Most Negative, Strike, Dip)**	**Geometric Attributes**
Porosity, Stratigraphic and Lithologic Variations, DHI's	RMS Amplitude, Relative Acoustic Impedance, Sweetness, Average Energy	Amplitude Accentuating Attributes
Pore fluid, Lithology, DHI's	Intercept, Gradient, Intercept/Gradient Derivatives, Fluid Factor, Lambda-Mu-Rho, Far-Near, (Far-Near) Far	AVO Attributes
Lithology, Porosity, Fluid Effects	Colored inversion, Sparse Spike, Elastic Impedance, Extended Elastic Impedance, Prestack Simultaneous Inversion, Stochastic Inversion	Seismic Inversion Attributes
Layer Thicknesses, Stratigraphic Variations	Continuous Wavelet Transform, Matching Pursuit, Exponential Pursuit	Spectral Decomposition

Table 1.1 presents several of the more commonly used attributes, and includes the category of geometric attributes that are particularly helpful in visualizing faults and fractures. Since there are numerous attributes today, a general recommendation is to understand and discern seismic attributes before blindly employing them (Barnes, 2016). Not included in Table 1.1 are more recently developed specialized attributes that enhance faults, some of which are used to construct fault probability volumes, including fault dip azimuth, and dip magnitude (Barnes, 2016; Xiong et al., 2018; Qi et al., 2019b).

This research tests different algorithms to improve fault continuity and artifact removal through fault enhancement methods. I use data analysis methodologies to enhance the expression of faults measured in the seismic amplitude and attribute volumes.

Coherence Enhancement

Seismic interpreters often employ seismic attributes to enhance faults and accelerate manual fault interpretation. However, more modern image processing and ML tools promise to facilitate automatic fault detection. Although interpreters often pick faults on individual vertical and time slices through coherence volumes, Barnes (2016) showed that these initial images can be improved using a multi-line three-dimensional approach. Specifically, he showed a 3D eigenvector method to find the fault orientation, followed by alternative dilation and erosion image processing steps to reject noise and join disjoint components of the same fault.

Using a similar approach, Machado et al. (2016) applied a Laplacian of Gaussian operator to enhance coherence attributes by constructing fault probability, fault dip magnitude, and fault dip azimuth volumes. Qi et al. (2018 and 2019) adjusted the Machado et al. (2016) approach by developing an analysis window to trace features along with the fault dip and azimuth, using interpolation in candidate fault points, thereby creating a skeletonized volume of fault probability. Pedersen et al. (2005) used a swarm intelligence or 'ant tracking' algorithm to simultaneously smooth and connect nearby fault anomalies. Hussein et al. (2020) showed how self-organizing maps can be used to highlight faults using input attributes that included dip magnitude, multispectral energy ratio similarity, total aberrancy magnitude, and curvedness. Because SOM is an unsupervised classification algorithm, the interpreter needs to recognize which clusters represent faults.

Probabilistic neural networks (PNN)

PNN is a feedforward neural network and is well-known in pattern recognition problems outside and within geosciences.(Maurya and Singh, 2018; Lubo-Robles et al. (2019) used probabilistic neural networks to detect and classify seismic facies. A supervised approach to facies classification includes a set of seismic attributes and a training dataset; PNN evaluates attributes and estimates the best characteristics for further analysis. Lubo-Robles et al. (2019) create two different results, a probability volume and a classified volume for selected features. In the application of PNN, a set of attributes defined for fault recognition are tested. In Chapter 2 I evaluate the ability of PNN to enhance faults a Great South Basin, New Zealand survey.

Convolutional neural networks

Fault enhancement like that described by Barnes (2006) and Machado (2016) also enhances stratigraphic edges based on edge detection analysis in seismic volumes, which may overprint the desired fault image. Data analysis techniques, such as CNN, are heavily used today in image analysis in a wide variety of applications to recognize elements or patterns. CNN uses different image filters and compares several results to extract feature characteristics automatically. This principle is more commonly being applied to solve problems in geophysics and particularly in exploration geophysics. Seismic facies detection is the most common use of CNN, in identifying salt differentiation, mass transport deposits, etc. (Zhao et al., 2015; Zhao and Mukhopadhyay, 2018; Qi et al., 2019a), but it is also now being used to identify faults (Wu et al., 2018, 2019; Xiong et al., 2018). One of the main issues with CNN is related to the training stages, which can influence the results as it has been noted that specific training models increase mistakes and reduce precision (Zhao and Mukhopadhyay, 2018)

Fault object extraction

Fault sticks and three-dimensional meshes are a method to present and visualize 3D surfaces in most digital settings. Typically, additional processing is needed to obtain these feature objects, such as fault sticks. Previous efforts to obtain objects from seismic volumes are well documented (Admasu et al., 2006; Cohen et al., 2006; Hale, 2013), and a few geologic interpretation software include these capabilities. Different algorithms that use fault extraction are focused on mapping regions with high fault probability, and after that point, fault objects are merged into larger features (Hale, 2013). All these approaches can be used to define seed points to aid in the creation of these surfaces. As an example, Hale (2013) extracted meshes using fault likelihood algorithms, calculating strikes and dips from seismic volumes. An initial approach to extract fault objects (Mora et al., 2020) used a method that creates a cloud of candidate voxels around a seeded fault sample; then, it performs a comparative analysis of energy, orientation, and probability volumes in the candidates. Finally, it creates a fault object as a plane file format.

Active contour method

In Chapter 3, an object extraction method based on active contours (Kass et al., 1988; Admasu et al., 2006) is used to extract objects. This algorithm changes the shape of the polyline and fits the line accordingly to the maximum values of an image; it also can create illusory contours fitting values where boundary data are absent. I use a skeletonized fault enhanced volume following the method of Qi et al. (2019b) as an input for the active contour extraction method. The volume has probability values from zero (no-fault) to one (fault), and the active contour method follows the high values in the seismic volume.

Faults are picked every 10 to 20 lines in the regions in order to determine the fault trend. As a second step, interpolation methods are employed on the picked samples to create an approximate location in every line from top to bottom, populating locations near the fault with XYZ coordinates. The extracted coordinates are then used to define the mapped fault's primary trend and define the active contour analysis's initial direction.

The active contours process begins by scanning one line with samples from one fault; then, the algorithm creates a connection between the points to create a spline; by doing so, the algorithm then starts to find high probability values and smooths the result to match the fault shape. The process continues in all sections, and a list of adjusted points to fault probability is found. The final step in the process is exporting the sample values to a stick file format to plot the result and analyze the accuracy of the results.

Geological setting

This research focuses on the fault method analysis within two different regions of New Zealand. Chapter 2 presents an analysis in the Great South Basin, near the southeastern portion of New Zealand's South Island. The Great South Basin area is used in the initial PNN analysis due to its structural and sedimentary history and complexity, dating back to the Mesozoic. This region originated as an unsuccessful Cretaceous rift, but it later developed into a subsiding basin during the Late Cretaceous and Paleocene (Ballance et al., 2017; Li et al., 2020).

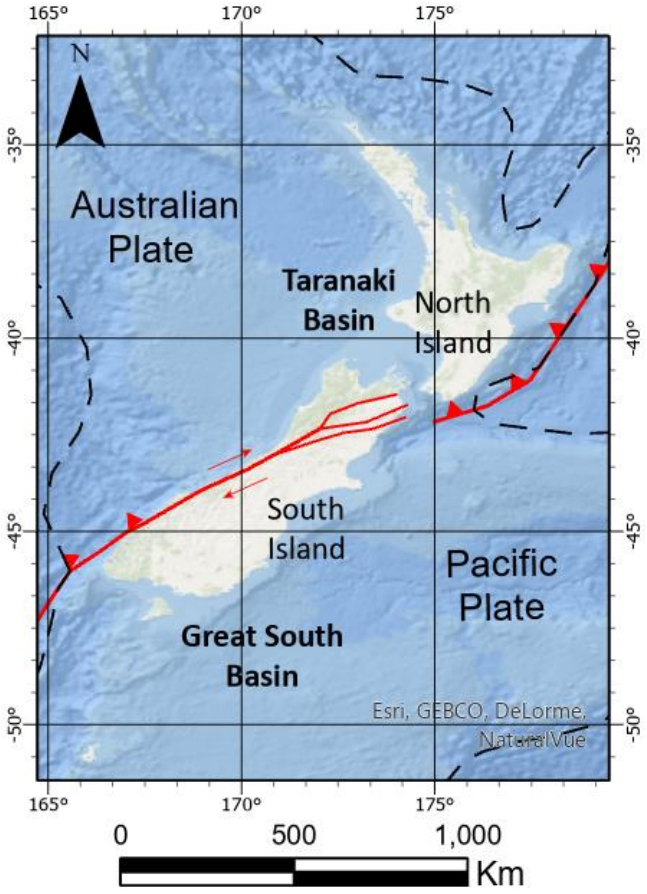


Figure 1.2. General map of New Zealand showing plate tectonic boundaries as they relate to the South and North Islands. The analysis areas are located in the Great South Basin (chapter 2) and the Taranaki Basin (chapter 3).

Chapter 3 moves to a more northern area, the Taranaki Basin. This basin has been the focus of many studies, and its geologic evolution is well documented (King and Thrasher, 1996; Strogen et al., 2014; Haque et al., 2016, 2016; Reilly et al., 2016; Coleman, 2018). The Taranaki Basin can be subdivided into two main structural areas (Nodder, 1993; King and Thrasher, 1996; Reilly et al., 2016). The first area is a relatively stable platform to the west, characterized by layer cake progradational deposition (Figure 1.2). The second area is the Eastern Mobile Belt, a more

complex and active setting consisting of active and reverse faults, grabens, and inversion structures (Figure 1.2) (King and Thrasher, 1996; Kroeger et al., 2013; Haque et al., 2016; Reilly et al., 2016; Coleman, 2018). These structures exhibit the basin history and current forces acting in northern New Zealand, and include a vast array of faults ideal for testing fault enhancement techniques.

Summary

This thesis is divided into several sections; this first chapter established the motivation for this research, including a brief introduction to the following chapters' methods.

Chapter 2 shows how the PNN method can enhance fault features using a seismic volume in the Great South Basin in New Zealand. I follow Lubo-Robles et al. (2020) and use an exhaustive search to test alternative attribute combinations in order to select the optimal attribute set to distinguish between faults and other seismic features. The initial set of attributes is based on previous research of fault enhancement in the Taranaki Basin (Hussein et al., 2020) as well as other selected attributes. A final step to enhance faults (Qi et al., 2019b) uses the PNN fault probability volume as the input for a fault enhancement process to increase the sharpness, erasing irrelevant and subparallel noise.

Chapter 3 presents a fault comparison in a seismic volume in the Taranaki Basin in New Zealand. This research starts with a review of the geologic setting and the dataset; then, it briefly reviews the three fault enhancement workflows and shows the effect of post-migration data conditioning on each of them. Next, the research presents the method to construct fault objects and measures

the differences between methods as compared to fault objects manually picked by a human interpreter.

References

Admasu, F., S. Back, and K. Toennies, 2006, Autotracking of faults on 3D seismic data: Geophysics, **71**, A49–A53.

Bahorich, M., and S. Farmer, 1995, 3-D seismic discontinuity for faults and stratigraphic features: The coherence cube: The Leading Edge, **14**, 1053–1058.

Ballance, P. F., L. Cotterall, and Geoscience Society of New Zealand, 2017, New Zealand Geology: An Illustrated Guide:

Barnes, A. E., 2016, Handbook of Poststack Seismic Attributes: SEG Books.

Chopra, S., and K. J. Marfurt, 2005, Seismic attributes — A historical perspective: Geophysics, **70**, 3SO-28SO.

Chopra, S., K. Marfurt, and R. Sharma, 2019, Unsupervised machine learning facies classification in the Delaware Basin and its comparison with supervised Bayesian facies classification: 89th Annual International Meeting of the SEG Expanded Abstracts, 2619–2623.

Cohen, I., N. Coult, and A. A. Vassiliou, 2006, Detection and extraction of fault surfaces in 3D seismic data: Geophysics, **71**, P21–P27.

Coleman, R. D., 2018, Tectonic evolution of Taranaki Basin, offshore New Zealand. M.S. Thesis, Colorado School of Mines.

Gersztenkorn, A., and K. J. Marfurt, 1999, Eigenstructure-based coherence computations as an aid to 3-D structural and stratigraphic mapping: Geophysics, **64**, 1468–1479.

Hale, D., 2013, Methods to compute fault images, extract fault surfaces, and estimate fault throws from 3D seismic images: Geophysics, **78**, O33–O43.

Haque, A. E., A. Islam, and M. R. Shalaby, 2016, Structure modeling of the Maui Gas Field, Taranaki Basin, New Zealand: Petroleum Exploration and Development, 12.

Hussein, M., R. Stewart, and J. Wu, 2020, Unsupervised Machine Learning Techniques for Subtle Fault Detection: 82nd EAGE Annual Conference & Exhibition Workshop Programme, **2020**, 1–5.

Kass, M., A. Witkin, and D. Terzopoulos, 1988, Snakes: Active contour models: International Journal of Computer Vision, **1**, 321–331.

King, P. R., and G. P. Thrasher, 1996, Cretaceous-Cenozoic Geology and Petroleum Systems of the Taranaki Basin, New Zealand: Institute of Geological & Nuclear Sciences.

Kroeger, K. F., R. H. Funnell, A. Nicol, M. Fohrmann, K. J. Bland, and P. R. King, 2013, 3D crustal-scale heat-flow regimes at a developing active margin (Taranaki Basin, New Zealand): Tectonophysics, **591**, 175–193.

Li, J., S. Mitra, and J. Qi, 2020, Seismic analysis of polygonal fault systems in the Great South Basin, New Zealand: Marine and Petroleum Geology, **111**, 638–649.

Lubo-Robles, D., T. Ha, S. Lakshmivarahan, and K. J. Marfurt, 2019, Supervised seismic facies classification using probabilistic neural networks: Which attributes should the interpreter use? 89th Annual International Meeting of the SEG, Expanded Abstracts, 2273–2277.

Lubo-Robles, D., T. Ha, S. Lakshmivarahan, K. J. Marfurt, and M. J. Pranter, 2020, Exhaustive Probabilistic Neural Network for attribute selection and supervised seismic facies classification: Interpretation, 1–67.

Machado, G., A. Alali, B. Hutchinson, O. Olorunsola, and K. J. Marfurt, 2016, Display and enhancement of volumetric fault images: Interpretation, **4**, SB51–SB61.

Marfurt, K. J., R. L. Kirlin, S. L. Farmer, and M. S. Bahorich, 1998, 3-D seismic attributes using a semblance-based coherency algorithm: Geophysics, **63**, 1150–1165.

Maurya, S. P., and N. P. Singh, 2018, Application of LP and ML sparse spike inversion with probabilistic neural network to classify reservoir facies distribution - A case study from the Blackfoot field, Canada: Journal of Applied Geophysics, **159**, 511–521.

Mora, J. P., H. Bedle, and Kurt. J. Marfurt, 2020, Constructing fault-surface objects from fault-sensitive attributes: 90th Annual International Meeting of the SEG, Expanded Abstracts, 1160–1164.

Nodder, S. D., 1993, Neotectonics of the offshore Cape Egmont Fault Zone, Taranaki Basin, New Zealand: New Zealand Journal of Geology and Geophysics, **36**, 167–184.

Pedersen, S. I., T. Randen, L. Sonneland, and O. Steen, 2002, Automatic 3D fault interpretation by artificial ants: 64th EAGE Conference & Exhibition, cp-5.

Qi, J., B. Zhang, B. Lyu, and K. Marfurt, 2019a, Seismic attribute selection for machine-learning-based facies analysis: Geophysics, **85**, O17–O35.

Qi, J., B. Lyu, A. AlAli, G. Machado, K. Marfurt, and Y. Hu, 2018, Image processing of seismic attributes for automatic fault extraction: 88th Annual International Meeting of the SEG Expanded Abstracts, 1644–1648.

Qi, J., B. Lyu, A. AlAli, G. Machado, Y. Hu, and K. Marfurt, 2019b, Image processing of seismic attributes for automatic fault extraction: Geophysics, **84**, O25–O37.

Reilly, C., A. Nicol, J. J. Walsh, and K. F. Kroeger, 2016, Temporal changes of fault seal and early charge of the Maui Gas-condensate field, Taranaki Basin, New Zealand: Marine and Petroleum Geology, **70**, 237–250.

Roden, R., T. Smith, and D. Sacrey, 2015, Geologic pattern recognition from seismic attributes: Principal component analysis and self-organizing maps: Interpretation, **3**, SAE59–SAE83.

Strogen, D. P., K. J. Bland, A. Nicol, and P. R. King, 2014, Paleogeography of the Taranaki Basin region during the latest Eocene–Early Miocene and implications for the 'total drowning' of Zealandia: New Zealand Journal of Geology and Geophysics, **57**, 110–127.

Wu, X., Y. Shi, S. Fomel, and L. Liang, 2018, Convolutional neural networks for fault interpretation in seismic images: 88th Annual International Meeting of the SEG, Expanded Abstracts, 1946–1950.

Wu, X., L. Liang, Y. Shi, and S. Fomel, 2019, FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation: Geophysics, **84**, IM35–IM45.

Xiong, W., X. Ji, Y. Ma, Y. Wang, N. M. AlBinHassan, M. N. Ali, and Y. Luo, 2018, Seismic fault detection with convolutional neural network: Geophysics, **83**, O97–O103.

Zhao, T., and P. Mukhopadhyay, 2018, A fault-detection workflow using deep learning and image processing: 88th Annual International Meeting of the SEG, Expanded Abstracts, 1966–1970.

Zhao, T., V. Jayaram, A. Roy, and K. J. Marfurt, 2015, A comparison of classification techniques for seismic facies recognition: Interpretation, **3**, SAE29–SAE58.

Chapter 2. Fault enhancement using probabilistic neural networks and attribute enhancement in seismic images; a case study in the Great South Basin, New Zealand.

Jose Pedro Mora, Heather Bedle, and Kurt J. Marfurt

Abstract

Seismic interpretation includes fault identification as an essential task to achieve any accurate, precise reservoir modeling or structural framework assessment and interpretation. Manual fault interpretation on vertical slices through the seismic amplitude volume is the traditional means to map faults, and sometimes the best data to display when quality controlling the fault picks. Interpreters routinely use edge-sensitive attributes such as coherence to accelerate the manual picking process, where the actual choice of a particular edge-sensitive attribute varies with the seismic data quality and with the reflectivity response of the faulted geologic formations. When more than one attribute is sensitive to faults, CMY color blending provides an effective way to combine the information content of two or three edge-sensitive attributes. In this work, we evaluate whether combining the information content of more than three attributes using probabilistic neural networks (PNN) provides any additional uplift.

We employ an exhaustive PNN to identify the optimal set of attributes to create a fault probability volume for a 3D survey acquired over the Great South Basin, New Zealand. The training data consists of manually picked faults on a coarse grid of 3D seismic lines. We construct a suite of candidate attributes using our understanding of the attribute response to faults seen in the data, as well as examples extracted from the published literature. Using a subset of this

training data, we then evaluate which subset of attributes and hyperparameters exhibit the highest validation on the remaining training data and find that when used together, aberrancy magnitude, GLCM homogeneity and entropy, Sobel filter similarity, and envelope best predict the faults. We also find that applying a directional Laplacian of a Gaussian and skeletonization filters further improves the result.

Introduction

As interpreters, we understand the importance of fault identification, although a very detailed analysis requires a manual and time-consuming process. Faults are used to understand the tectonic history of a basin, to improve horizon correlations, to determine the effect faults have on accommodation space, and to define potential hydrocarbon traps. Given both the amount of work and their importance, seismic interpreters are looking to machine learning and data analytics tools to find new approaches that can provide faster, more complete, yet accurate fault networks.

One standard method to identify geologic features in a seismic dataset is to employ algorithms and derive attributes that quantify seismic characteristics related to structural, stratigraphic, and discontinuity properties (Chopra and Marfurt, 2005; Barnes, 2016). Examples of such attributes are seismic discontinuity estimations obtained through cross-correlation, semblance, Sobel filters, and eigenanalysis methods (Bahorich and Farmer, 1995; Marfurt et al., 1998; Luo et al., 1996; Gersztenkorn and Marfurt, 1999; Barnes, 2006). These initial fault images can be improved through further filtering techniques, including swarm intelligence (Randen et al., 2001; Pedersen

et al., 2002), eigenstructure-based filters (Barnes, 2014; Machado et al., 2016; Qi et al., 2019b), or prediction error filters (Hale, 2012) that result in fault probability volumes.

In addition to attributes and filters, machine learning approaches provide new tools to solve seismic analysis problems such as fault detection (Zhao and Mukhopadhyay, 2018; Wu et al., 2019) and seismic facies identification (Zhao et al., 2015; Chopra et al., 2019; La Marca-Molina et al., 2019; Lubo-Robles et al., 2019). Machine learning algorithms can be either unsupervised or supervised (Zhao et al., 2015; Chopra et al., 2019). Unsupervised methods include K-means, principal component analysis (PCA), self-organizing maps (SOM), and generative topographic mapping (GTM) (Taner et al., 2001; Chopra and Marfurt, 2014; Zhao et al., 2015). These approaches identify clusters or patterns across multiple data volumes; interpreters then use their geologic insight to associate a given pattern to specific geologic features. Depending on the geology, many attributes are highly correlated and provide similar results. PCA and ICA are "dimensionality reduction" techniques that find a smaller set of "composite attributes" that more efficiently represent the geological variation in the data (Roden et al., 2015; Lubo-Robles et al., 2019; Hussein et al., 2020).

The unsupervised GTM and SOM algorithms can also be seen as a projection technique, where now the projection is onto a deformed lower-dimensional manifold that lives in the original N-dimensional attribute space (Chopra and Marfurt, 2014; Roden et al., 2015; Zhao et al., 2015). Taner et al. (2001) and Chopra et al. (2019) used SOM to identify and map different seismic Hussein et al. (2020) used SOM classification to map faults which enabled them to map small faults that were difficult to map using conventional methods.

In contrast to unsupervised classification, supervised classification is a guided process that uses well, production, or human experience to teach the algorithm what features or patterns need to be extracted (Bishop, 1995). In this work, we will use probabilistic neural networks to use the attribute expression of hand-picked faults to predict fault locations in the unmapped seismic volume. Convolutional neural networks (CNN) are based trained as a supervised classification approach to recognize elements or patterns in the same way as image analysis applications. Zhao et al. (2015), Zhao and Mukhopadhyay (2018), and Qi et al. (2019a) used CNN algorithms to map salt and mass transport deposits, whereas Wu et al. (2018, 2019), Xiong et al. (2018), and Qi et al. (2020) used CNN to map faults. One of the main issues with CNN relates to training stages that influence results because limited training could increase mistakes and reduce precision in features enhanced by this method (Zhao and Mukhopadhyay, 2018; Qi et al., 2020).

In addition to CNN, feedforward neural networks are used in pattern recognition problems to detect and classify facies (Maurya and Singh, 2018; Lubo-Robles et al., 2019; 2020). Lubo et al. (2020) employed probabilistic neural networks (PNN) coupled to an exhaustive algorithm to analyze a set of seismic attributes and a training dataset to estimate the best attribute combination for salt differentiation. This classification algorithm creates two different results, a probability volume and a classified volume for selected features (Lubo-Robles et al., 2019).

This study aims to apply PNN networks to enhance fault features in a seismic volume in the Great South Basin in New Zealand. We use the PNN exhaustive approach (Lubo-Robles et al., 2020) and test several attribute combinations towards selecting the optimal attribute set to distinguish between faults and other seismic features. First, we use a preconditioning workflow to remove unwanted noise and increase the frequency content. Second, we map several faults inside the

analysis area to sample different characteristics across the data; in this step, we divide the data into training and validation to quantify the proposed method's error. Third, we analyze attribute combinations to identify the minimum error and optimal parameters to use in the final classification. The initial set of attributes is based on the Hussein et al. (2020) study in the Taranaki Basin and additional user-defined attributes. After the error is defined, a PNN model is then created to classify the entire seismic volume. A final step to enhance faults (Qi et al., 2019b) uses the fault probability volume as the input for a fault enhancement process to increase the sharpness, erasing irrelevant and subparallel noise. Then a skeletonized volume is calculated from the fault enhancement process.

Geological setting

The Great South Basin has been the focus of several studies to understand its geological evolution (Sahoo et al., 2014; Ballance et al., 2017; Li et al., 2020); due to its configuration near a plate tectonic limit (), it has been affected by complex structural processes resulting in several faulting systems and made the area a perfect candidate for this research. Located on the southern side of the Southern Island in New Zealand, the Great South Basin is a tectonically active area due to its closeness to the Alpine Fault and the strike-slip plate boundary (). This basin's structural and sedimentary history dates back to the Mesozoic, originated as an unsuccessful Cretaceous rift, and developed into subsiding basins during the Late Cretaceous and Paleocene (Ballance et al., 2017; Li et al., 2020).

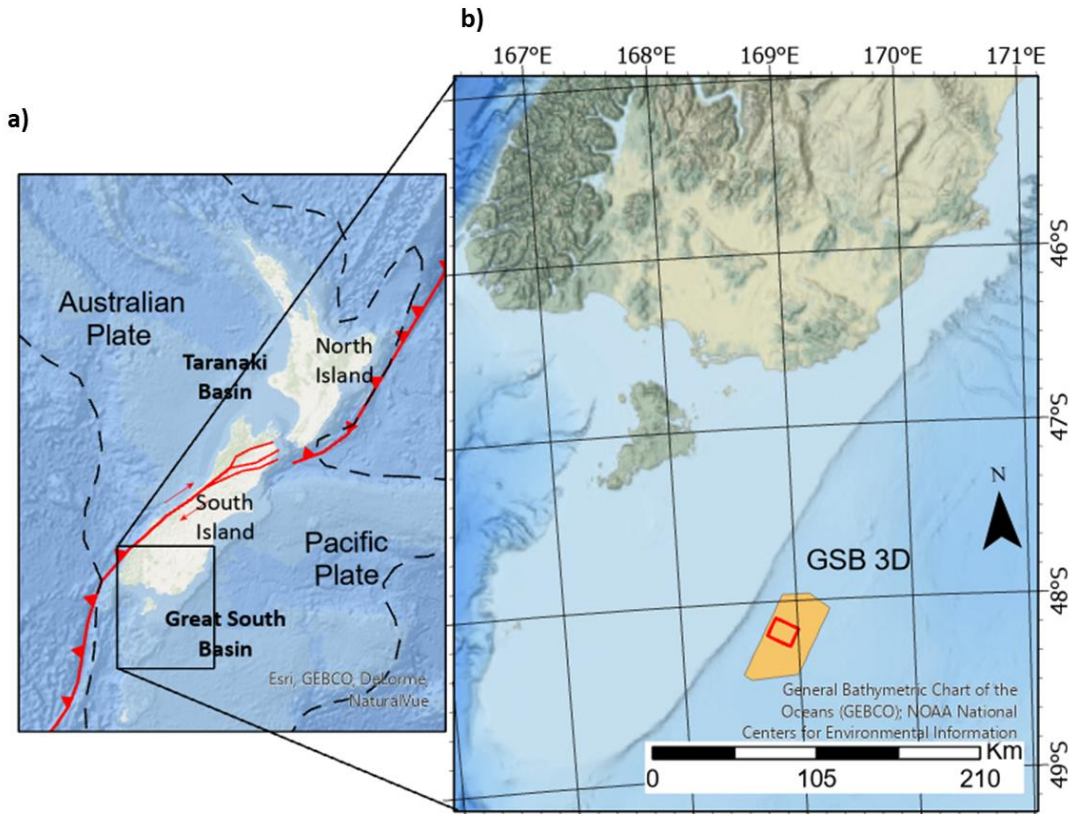


Figure 2.1. General map of (a) New Zealand and (b) Great South Basin. The general map shows plate tectonics boundaries including the strike-slip fault cutting the South Island. (b) Map of the South Island showing the location of the GSB seismic volume.

This region is reported to host a 9 km sediment succession filled mainly with sediments from Cretaceous to the recent. The main stages are interpreted as a back-arc and forearc close to a subduction margin (Cook, 1999; Li et al., 2020). A second stage latest Jurassic to the earliest Cretaceous and consists mainly of conglomerate, sandstone, mudstone, and coals deposited in deltaic to terrestrial (Cook, 1999; Li et al., 2020). Eocene strata have been affected in local areas

by polygonal faulting consisting of normal faults and diverse slope patterns (Morley et al., 2017; Li et al., 2020).

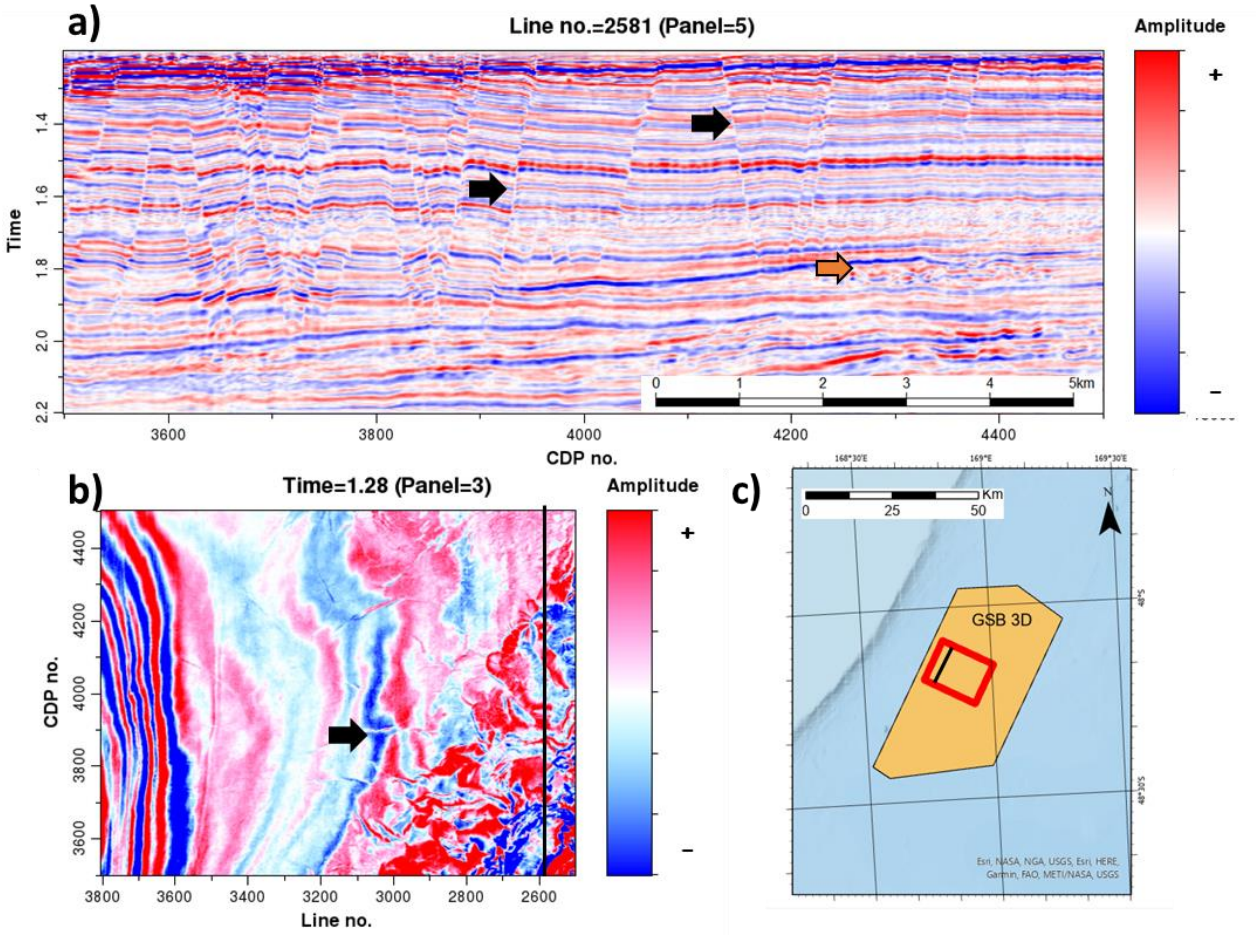


Figure 2.2 Representative (a) vertical and (b) time slices through the 3D GSB seismic amplitude volume. Black arrows indicate several faults on both slices. The orange arrow indicates a mass transport complex. (c) Legend showing the extent and location of the vertical and time slice in the larger GSB survey. (Data courtesy of NZP&M).

Datasets

ExxonMobil and WesternGeco acquired the dataset in the Great South Basin (GSB), New Zealand; the seismic dataset is a post-stack time migrated volume acquired between 2007 and 2008, obtaining a 1344 km2 survey to evaluate a potential reservoir presence in the southern end of the Parahaka Graben. The chosen volume is cropped to 1200 ms - 2200 ms in time (Figure 2.2). We use commercial software for seismic display and conventional fault interpretation, wherein the faults are exported as a surface.

Methodology

We follow the workflow presented in Figure 2.3; first, we apply a spectral balancing algorithm to increase the frequency content, followed by a structural oriented filtering to remove unwanted noise and artifacts subparallel to seismic reflectors. Secondly, attribute calculation follows data preconditioning and fault digitalization to extract relevant samples and split them into training and validation datasets. PNN algorithm analyzes the input data, searches the optimal attribute combination, and quantifies the error and optimal parameters in the tested groups. Finally, attribute combination is defined to create the PNN model and classify the seismic dataset (Lubo-Robles et al., 2020). The last steps include a medium filter to remove outliers to smooth the fault probability; then, we use a fault enhancement process to remove subparallel features and enhance continuity (Qi et al., 2019b); then, we employ skeletonization.

Manual interpretation and preconditioning are the first tasks to complete to create attributes training and classification. We follow a manual interpretation workflow in commercial software

by selecting individual faults on a time slice and mapping the entire feature over inline or crossline sections. A relevant number of faults across the analysis area are needed to train the algorithm to identify fault characteristics across the computed seismic attributes. The fault interpretation is then converted to a fault surface and exported as the primary input for the sampling stage. A second manual interpretation takes place to extract non-fault seismic examples; in this step, we sample clean reflectors, noise, MTDs, and other features. All the mapped samples must represent faults, anomalies, clean reflectors, or noise inside the volume to provide a good object differentiation in the classification stage.

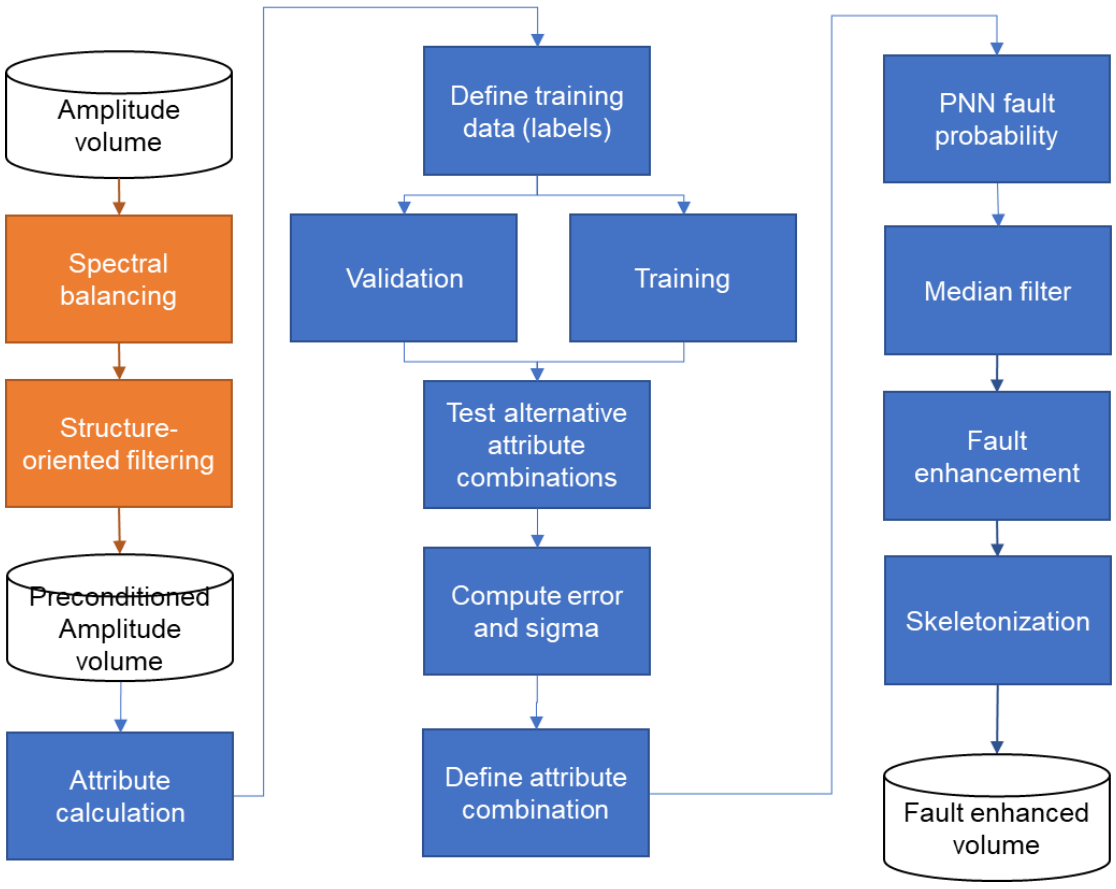


Figure 2.3. PNN and fault enhancement workflow.

Data preconditioning

We apply data preconditioning to increase fault detectability, remove unwanted noise and minimize stratigraphic features (Ha et al., 2019). First, we apply spectral balancing to increase the bandwidth and frequency content; we improve the signal-to-noise ratio and feature detectability inside the dataset (Chopra and Marfurt, 2016). A second preconditioning step includes two structural-oriented filters that favor edge detection and preservation by using a smoothing operator in parallel reflectors, avoiding seismic reflection terminations (Fehmers and Höcker, 2003a).

Next, we calculate seismic attributes to act as the PNN analysis input. PNN analyzes numerous seismic attributes and identifies the relevant ones to classify seismic facies (Lubo-Robles et al., 2020). However, this is an unrealistic approach due to computational constraints and time/effort investment. We define a small set of attributes by prior knowledge and analysis of previous studies on the subject. Hussein et al., (2020) defined a small set of critical seismic attributes for unsupervised classification methods, and we added additional user-defined attributes such as envelope, multispectral Sobel filter similarity, and total energy since these products show faults as sharp features.

The sampling process involves two main stages, sample definition, and sample extraction. Firstly, we transform fault surface and no-fault areas into sample coordinates to obtain sample locations; then, we extract several samples to be used on the PNN training stage.

Probabilistic neural networks

The PNN method uses a Parzen window and Bayes' criteria to estimate a probability density function (PDF) from training data; the PDF calculated from the training is used to assign a class to each voxel where the PDF is the largest (Specht, 1990; Lubo-Robles et al., 2020). We follow Lubo-Robles et al. (2020) and define PNN using Gaussian functions. The estimated probability of the attribute vector $\mathbf{x}_q$ at the q^{th} voxel being a member of the k^{th} class is

$$g_k(\mathbf{x}_q) = \frac{1}{N_k} \sum_{n=1}^{N_k} \exp\left[-\sum_{m=1}^{M} \frac{\left(x_{qm} - a_{nmk}\right)^2}{r^2} \right] \quad (1)$$

where

M is the length of the attribute vector (the number of attributes),

N_k is the number of vectors used to define the kth class,

a_{nmk} are the attribute values at the nth voxel of the kth class, and

r is a smoothing parameter used to measure the Gaussian distance between the current (q^{th}) vector and the n^{th} training vector. In the machine learning vocabulary r is a "hyperparameter" that will be estimated using an optimization procedure. By normalizing the various attribute volumes, we can limit the search for smoothing parameter, r, to be a single scalar, instead of a different smoothing parameter for each attribute.

Examining Figure x, the PNN consists of an input layer, a pattern layer, a summation layer, and an output layer (Specht, 1990; Lubo-Robles et al., 2020). The input layer selects a new sample to be analyzed as a vector. The pattern layer calculates the difference between the input sample

and the training data vectors. These differences are then used in the Gaussian (or "activation function") seen in equation 1. In the summation layer, the algorithm finds the average estimated density function for each possible class. Finally, the output layer assigns the most probable class, *c*, to the unknown sample (Lubo-Robles et al., 2020): using

$$c(\mathbf{x}_q) = \mathrm{ARG}\left\{\underset{k}{\mathrm{MAX}}\left[g_k(\mathbf{x}_q)\right]\right\}. \quad (2)$$

In contrast to SOM, K-means, and several other clustering algorithms, PNN provides a numerical confidence in the classification step (Lubo-Robles et al., 2020).

PNN training and classification

We test the sampled candidate attributes to identify each combination's error and identify the suite of attributes to use. First, we divide the sampled dataset into two different groups as training and validation sets. We try to maintain 50% samples on faults and no faults and to have at least a 10% in the validation dataset (fault and no-fault samples). Second, we apply data normalization to transform attribute values to a common scale, retaining differences in values ranges; Z-score normalization, robust scaling, and logarithmic normalization methods are tested, but Z-score is applied in the final analysis to estimate all attribute characteristics.

The exhaustive PNN algorithm tests all possible combinations of seismic attributes and provides the associated error from the sampled dataset, and includes an exhaustive search on different r parameters (0 to 3.0) for each combination. We obtain a list of error estimates and the optimum parameter r; from this list, we select the optimal set of parameters to create the PNN classification. The final step in the PNN analysis is to create a model to classify the seismic volume

into a probability volume and a fault/no-fault volume. The fault probability product is the one used as input for the fault enhancement process.

Fault enhancement

The last step in our workflow is the skeletonization process to sharpen the faults in two steps (Machado et al., 2016; Qi et al., 2019b). We use a median filter to remove anomalies and outliers in the volume. Then, a fault enhancement filter searches the closest probability within a 3D small window in each voxel; this step identifies fault planes, enhances those features to estimate a continuous fault probability, and creates additional dip magnitude calculation (Machado et al., 2016; Qi et al., 2019b). Skeletonization searches faults and unconformities along planes, discarding lower candidates in the analysis window to produce a sharp image from the fault volume using the computed fault dip magnitude and fault dip azimuth.

Attributes

Our initial set of attributes includes aberrancy magnitude, multispectral energy ratio similarity, Sobel filter similarity, envelope, dip magnitude, curvedness, GLCM homogeneity, and entropy. We use multispectral similarity because it can be used a measure of the waveform coherence, which improves fault images (Gersztenkorn and Marfurt, 1999; Chopra and Marfurt, 2007; Li et al., 2017)**.** Aberrancy is also helpful because it measures the curvature's lateral change along the surface and might delineate subtle faults where coherence anomalies might not be present (Qi and Marfurt, 2018). We also consider curvedness because it measures the total deformation of a seismic horizon given by maximum and minimum curvatures and is a complimentary attribute to coherence in tectonic deformation studies (Chopra and Marfurt, 2015); curvedness must be used with caution because the fault anomalies are not aligned with the curvature anomalies

(Bhattacharya and Verma, 2019)**.** The dip magnitude is also considered to define a high dip in the seismic but avoiding dip azimuth since it is not realistic to include azimuthal trends as a general characteristic for all fault systems. In addition to this set of attributes, we also use the envelope that shows the complex trace magnitude (Chopra and Marfurt, 2007). Gray-level co-occurrence matrix (GLCM) homogeneity and entropy are also tested; the first one is a measure of the seismic smoothness, and entropy measures the degree of disorder or complexity of the image (Gao, 2003; Chopra and Marfurt, 2007); additionally, a separate test includes individual multispectral volumes as inputs.

Results

The initial preconditioning provides the data for attribute calculations;

Figure 2.4a shows a time slice on selected attributes tested in the PNN classification. Multispectral energy and Sobel similarity show clear seismic discontinuities and crisp anomalies (yellow arrows); other enhanced areas are depositionally related anomalies, not low enough to match the incoherent faulting (black arrow). Aberrancy magnitude (

Figure 2.4b) shows the same uncoherent features, but events are not so crisp as the coherence attribute (yellow arrow); features that appeared in the previous figure are now absent (black arrow). Dip magnitude shows reflector magnitudes similar to aberrancy; fault areas are not sharp as coherence (yellow arrow); the area to the left (black arrow) is enhanced as in the coherence case, which shows that the area's dip is higher than surrounding places. GLCM homogeneity shows acceptable results, particularly in blue areas of low homogeneity, but several non-fault-

related features are also enhanced. Entropy seems the inverted case as homogeneity with high entropy values related to faults, but non-fault-related features are also enhanced. The envelope shows significant sharp features on cross-sections, but deep blue colors related to low amplitudes also appear in different areas with non-fault-related anomalies. Curvedness shows some resemblance to the dip and aberrancy magnitude, but the events' position is related to the fault's vicinity. Finally, total energy shows low values in fault-related events (yellow arrow), but the contrast between faults and surroundings is not strong in some areas.

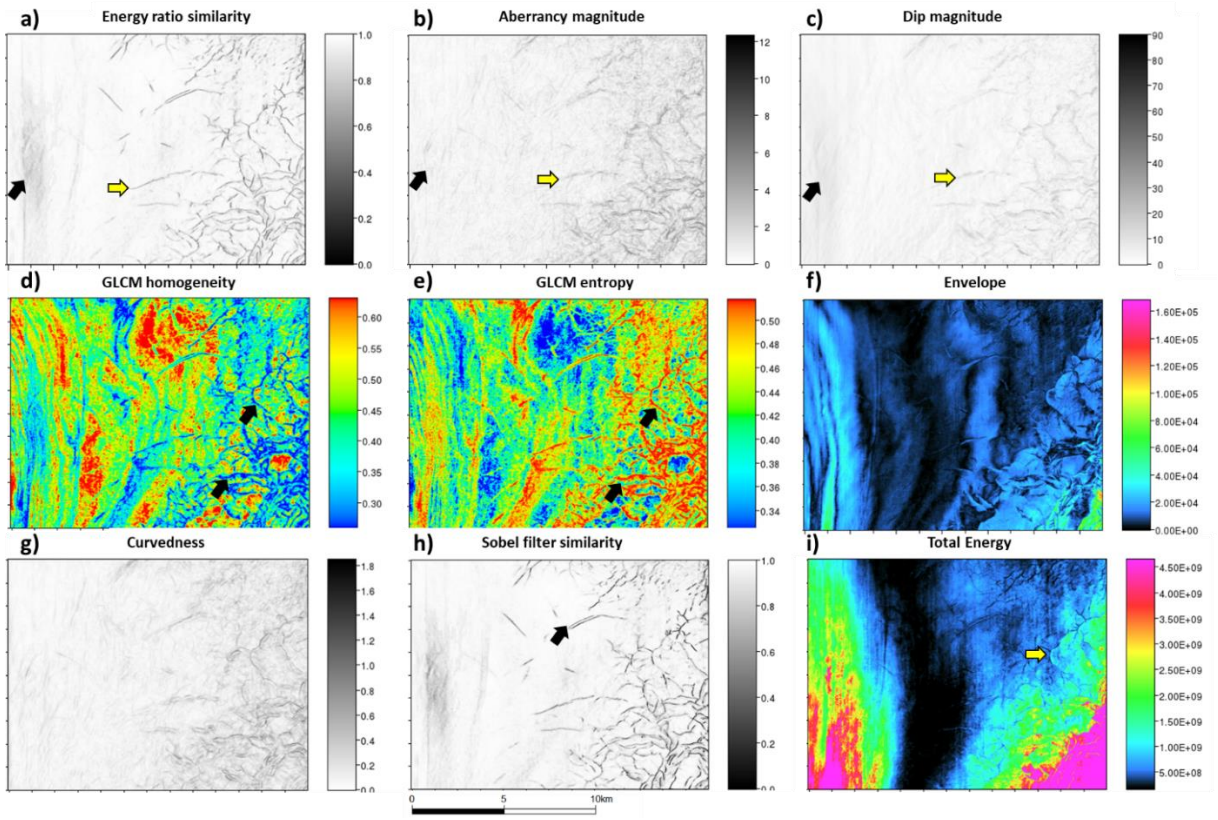


Figure 2.4. Candidate attributes used as input in the PNN fault enhancement process prior to determining the best combination as shown in Figure 2.3: (a) Energy ratio similarity, (b) aberrancy magnitude, (c) dip magnitude, (d) GLCM homogeneity, (e) GLCM entropy, (f)

envelope, (g) curvedness, (h) Sobel filter similarity, and (i) total energy. Time slices are at t= 1.28 s.

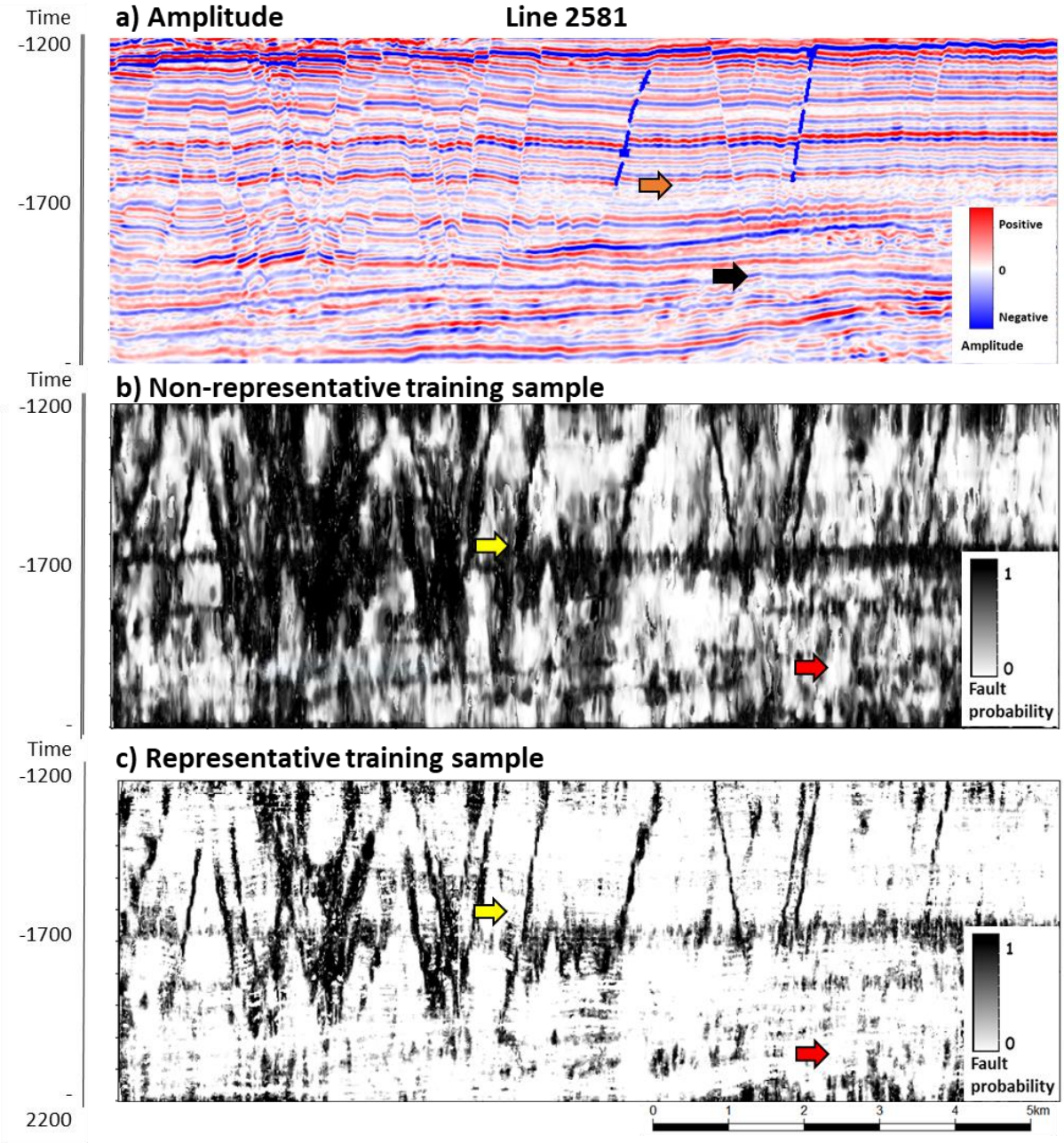


Figure 2.5. The impact of training data on PNN fault classification. (a) Amplitude section and key representative training selected, blue lines are hand-picked fault samples, orange and black

arrows show no-fault selected areas as representative training/validation in this section. (b) The fault prediction for a vertical slice with training samples with little seismic noise and no significant MTD facies. Note the artifacts and no clean faults in the section. The image below (b) shows a representative sample training set, fewer artifacts, and better classification across the section (yellow arrow); still, some areas show mass transport complexes (red arrow).

We carefully manually mapped ten faults inside the GSB volume; then, we sample areas with clean and continuous reflectors; Then, we completed the entire workflow with this initial set of samples and a Z-score normalization approach; the results are presented in Figure 2.5a

Figure 2.5b shows the same parameters and normalization, but we used a different sampling set; we mapped complex areas with noise, low amplitude areas, and mass transport complexes to accurately depict complicated areas.

We tested three different normalization approaches using the improved sample dataset; we followed the proposed methodology and completed the classification process. Figure 2.6a and Figure 2.6b show the comparison between Z-score normalization, logarithmic scaler, and robust scaler; even though there is fault presence, the results are noisy, with a high amount of intra-reflector anomalies. In contrast, the Z-score normalization approach in Figure 2.6c shows cleaner fault probabilities and provides a good insight into this section's faulting system. Still, there are noisy areas found at 1700 ms that exhibit several irregularities in the final classification. Appendix Table A.1 shows the input and selected attributes.

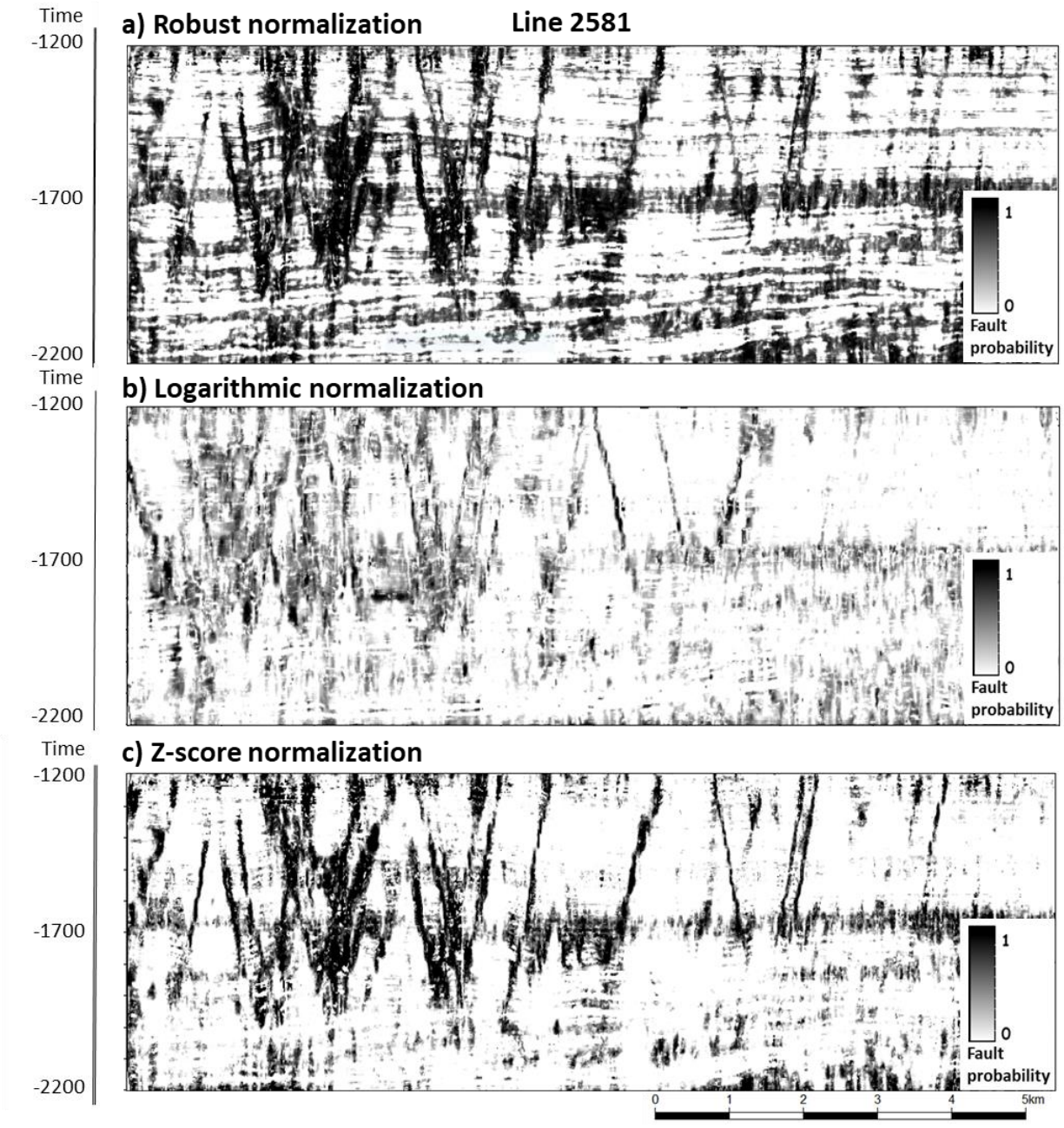


Figure 2.6. The impact of data normalization on the resulting PNN fault prediction. (a) The robust scaler normalization shows an increased amount of noise and misclassification of low reflectivity areas as faults. (b) The logarithm scaler works well, and a cleaner result is achieved. (c) Counterintuitively, the Z-score normalization method shows the best image, consistent with

findings by Ha et al., (2021) use of PNN to delineate salt facies. Yellow arrows indicate faults, and red arrows the absence of faults. Appendix Table 1.1 shows the input and selected attributes.

The exhaustive PNN and Z-score normalization in Figure 2.6c show that aberrancy magnitude, GLCM entropy, GLCM homogeneity, Sobel filter similarity, and envelope are the best combinations to map fault features in the analysis area with an optimal parameter r, or sigma of 0.4. The fault probability result is used in the next step to create an enhanced fault probability volume; complete results are presented in the appendix Table A.1.

The fault enhancement process creates an initial sharpened volume with several faults enhanced across the section in Figure 2.7a. Numerous noisy areas are removed, and continuous faults are now enhanced; also, stair-step artifacts are reduced compared to the initial PNN classification. The final step calculates a skeletonized volume.

Discussion

Figure 2.5 shows a critical factor in the PNN and supervised classification and lies in the fact that sampling and user mapping influence the subsequent classification. In Figure 2.5a, noise, mass transport complexes, and other stratigraphic features are classified closely because those events are more similar to faults than perfectly aligned reflectors. To overcome sampling problems, we need to select samples on problematic areas, such as complex stratigraphic areas, noise, and near-fault spatially and characteristic relevant.

Z-score normalization is employed in the final classification approach for attribute normalization, while previous PNN classification presents their results using robust scaler or log-normalization (Lubo-Robles et al., 2020) in our analysis, the final result is not good enough to be selected. Z-score is the right decision based on the results, and some authors find advantages in this normalization type (DeLaughter et al., 2005); still, the user needs to be careful when selecting normalization approaches due to the impact on final results.

Energy ratio similarity or Sobel filter similarity are expected to be selected because they show discontinuities in seismic datasets; the envelope is also selected probably due to the sharp differences between faults and background seismic but not all the low envelope values are faults. GLCM shows a good coupling due to the sharp difference between faults and surroundings. PNN provides a way to use different attributes and couple them to enrich a classification compared to one attribute. Some examples of multiple attribute classification can be found to identify subtle faults in complex environments (Hussein et al., 2020, 2021; Roden, in press).

The PNN itself can create a classified volume; however, the result (Figure 2.6c) is not clean enough to be used directly use in seismic interpretation, so we use the PNN fault probability as the input for a fault enhancement workflow. The fault enhancement process sharpens fault probabilities, and the results show its effectiveness.

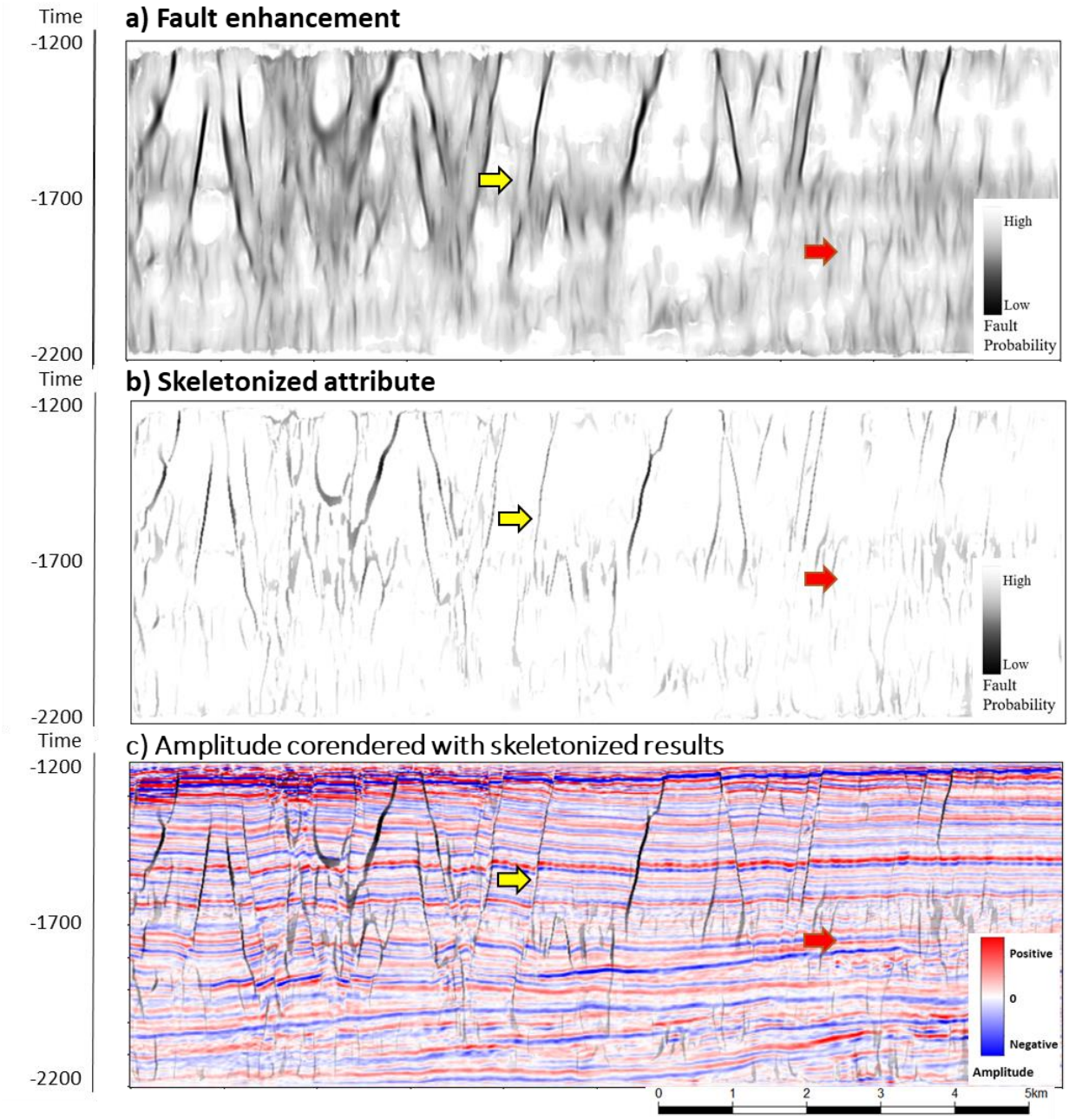


Figure 2.7. PNN fault classification results. (a) shows fault enhancement (Machado et al., 2016; Qi et al., 2019b) using the Z-score normalization result as input (Figure 2.6c). (b) Result after

skeletonization described by Qi et al. (2019b). (c) Amplitude corendered with PNN results. Yellow arrows indicate areas of improved fault continuity. Red arrows indicate suppressed artifacts.

Conclusions

We present a workflow to use exhaustive PNN classification coupled with the fault enhancement process; interpreters can use the result as an additional tool to identify faults in seismic interpretation. We selected initial geometric and instantaneous attributes, such as aberrancy magnitude, dip magnitude, energy ratio similarity, GLCM entropy and homogeneity, Sobel filter similarity, envelope, and curvedness.

In our analysis, the initial set of attributes was selected employing the Z-score normalization due to better results than other scaling methods. Even though robust and logarithmic scalers provide a smaller number of attributes to create the classified volume, the fault results do not present well-defined faults, and Z-score results are better compared to the mentioned methods. The PNN process defined aberrancy magnitude, GLCM entropy and homogeneity, Sobel filter similarity, and envelope as the best combination in the fault enhancement classification process. The additional step that involves fault enhancement (Machado et al., 2016; Qi et al., 2019b) sharpens the fault probability and skeletonizes the result to improve the fault volume following Qi et al. (2019b).

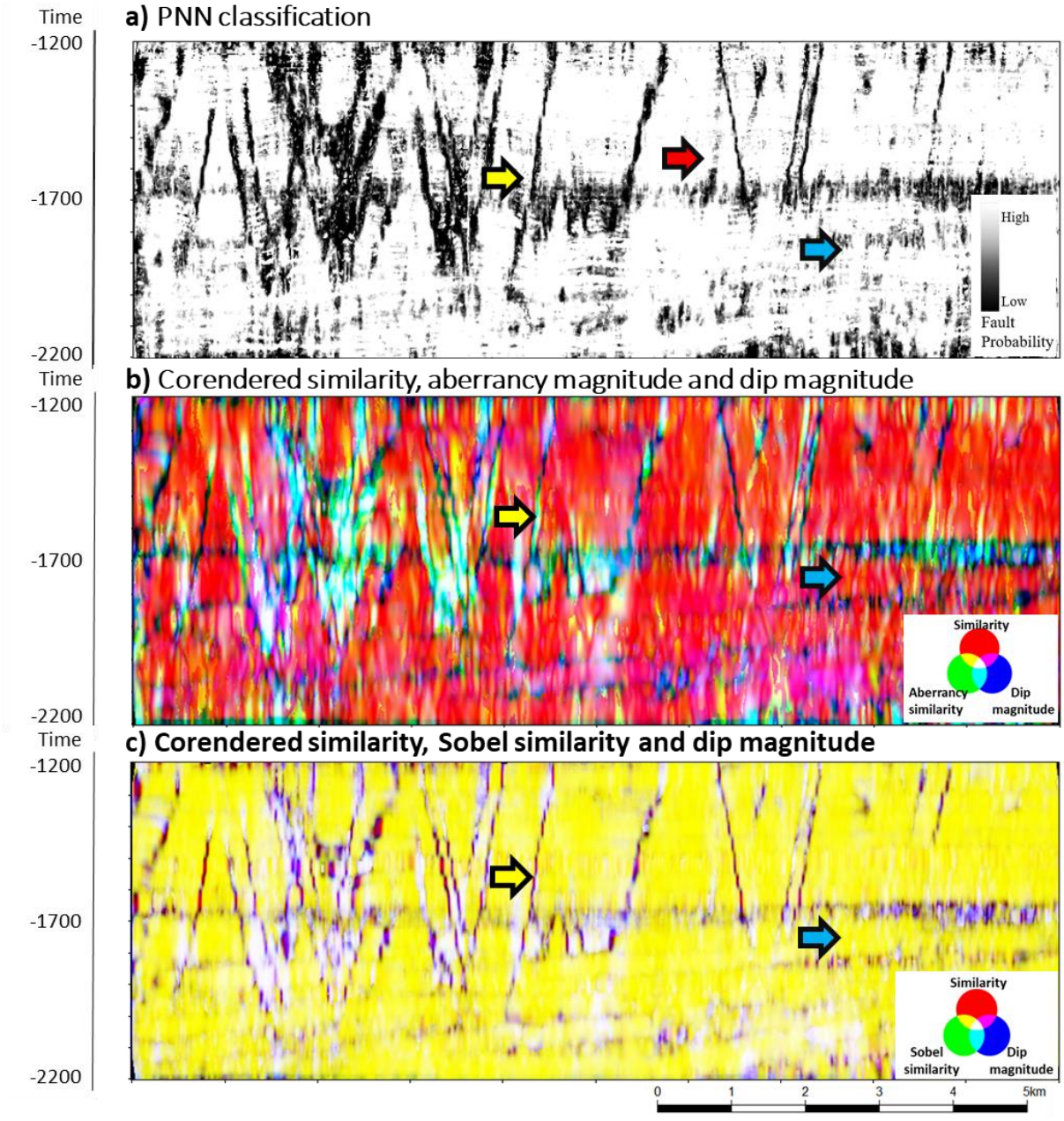


Figure 2.8 PNN classification comparison and RGB blended attributes. (a) PNN classification section and key representative areas showing faults as yellow arrows and mass transport complexes marked with blue arrows. (b) A blended image using the RGB color model. R corresponds to multispectral energy ratio similarity, G to aberrancy magnitude, and B to dip magnitude. The image below (c) shows a) A blended image using the RGB color model. R

corresponds to multispectral energy ratio similarity, G to Sobel filter similarity, and B to dip magnitude. PNN shows some enhanced features in the red arrow area compared to the RGB blending in the two other cases.

Compared to coherence enhancement or CNN fault enhancement, PNN will take more time to get a complete result due to attribute calculation and individual PNN voxel analysis; however, results might be richer than individual attributes and have the potential to be used in subtle fault analysis. Sampling selection is a critical step in all supervised classification processes; undersampling can be a source of artifacts and misclassification; a representative sample dataset can influence the result depending on user input. Several attributes can be tested, but the process needs to be subjected to meaningful attributes to optimize time and computation resources in PNN classification. As Figure 2.8 shows, some additional fault-like features are enhanced (Figure 2.8a, red arrow); we hypothesize this method potential should be tested in small volumes to analyze its performance in reservoir analysis and subtle fault identification.

Acknowledgments

We want to express our gratitude to David Lubo-Robles and Thang Ha for their valuable comments and help achieve these results. We also want to thank New Zealand Petroleum and Minerals for providing the data used in this project available at the New Zealand Petroleum and Minerals administration website. We also thank the Attribute Assisted Seismic Processing and Interpretation consortium (AASPI) for their continuous and invaluable support. We also acknowledge Schlumberger, which provided Petrel licenses to the University of Oklahoma for

research and education use. We want to thank the SDA research group at the Oklahoma University for its observations. The first author also wants to express his gratitude to the Fulbright program and the Saldarriaga-Concha Foundation because they made this research possible.

References

Bahorich, M., and S. Farmer, 1995, 3-D seismic discontinuity for faults and stratigraphic features: The coherence cube: The Leading Edge, **14**, 1053–1058.

Ballance, P. F., L. Cotterall, and Geoscience Society of New Zealand, 2017, New Zealand Geology: An Illustrated Guide:

Barnes, A. E., 2006, A filter to improve seismic discontinuity data for fault interpretation: GEOPHYSICS, **71**, P1–P4.

Barnes, A. E., 2016, Handbook of Poststack Seismic Attributes: SEG Books.

Bhattacharya, S., and S. Verma, 2019, Application of volumetric seismic attributes for complex fault network characterization on the North Slope, Alaska: Journal of Natural Gas Science and Engineering, **65**, 56–67.

Bishop, C., 1995, Neural Networks for Pattern Recognition: Oxford University Press.

Chopra, S., and K. J. Marfurt, 2005, Seismic attributes — A historical perspective: Geophysics, **70**, 3SO-28SO.

Chopra, S., and K. J. Marfurt, 2007, Seismic Attributes for Prospect Identification and Reservoir Characterization: Society of Exploration Geophysicists and European Association of Geoscientists and Engineers.

Chopra, S., and K. J. Marfurt, 2014, Seismic facies analysis using generative topographic mapping: 84th Annual International Meeting, SEG, Expanded Abstracts, 1390–1394.

Chopra, S., and K. J. Marfurt, 2015, Is curvature overrated? No, it depends on the geology: First Break, **33**.

Chopra, S., and K. J. Marfurt, 2016, Spectral decomposition and spectral balancing of seismic data: The Leading Edge, **35**, 176–179.

Chopra, S., K. Marfurt, and R. Sharma, 2019, Unsupervised machine learning facies classification in the Delaware Basin and its comparison with supervised Bayesian facies classification: 89th Annual International Meeting of the SEG Expanded Abstracts, 2619–2623.

Cook, R. A., 1999, Cretaceous-Cenozoic Geology and Petroleum Systems of the Great South Basin, New Zealand: Institute of Geological & Nuclear Sciences.

DeLaughter, J., B. Meltz, S. Smith, J. Yun, and M. Murat, 2005, Use of z-score to rescale amplitudes: The Leading Edge, **24**, 698–701.

Fehmers, G. C., and C. F. W. Höcker, 2003, Fast structural interpretation with structure-oriented filtering: Geophysics, **68**, 1286–1293.

Gao, D., 2003, Volume texture extraction for 3D seismic visualization and interpretation: GEOPHYSICS, **68**, 1294–1302.

Gersztenkorn, A., and K. J. Marfurt, 1999, Eigenstructure-based coherence computations as an aid to 3-D structural and stratigraphic mapping: Geophysics, **64**, 1468–1479.

Ha, T. N., K. J. Marfurt, B. C. Wallet, and B. Hutchinson, 2019, Pitfalls and implementation of data conditioning, attribute analysis, and self-organizing maps to 2D data: Application to the Exmouth Plateau, North Carnarvon Basin, Australia: Interpretation, 7, SG23–SG42.

Ha, T. N., D. Lubo-Robles, K. J. Marfurt, and B. C. Wallet, 2021, An in-depth analysis of logarithmic data transformation and per-class normalization in machine learning: Application to unsupervised classification of a turbidite system in the Canterbury Basin, New Zealand and supervised classification of salt in the Eugene Island mini-basin, Gulf of Mexico: Interpretation, 1–109.

Hale, D., 2012, Fault surfaces and fault throws from 3D seismic images: SEG Technical Program Expanded Abstracts 2012, 1–6.

Hussein, M., R. Stewart, and J. Wu, 2020, Unsupervised Machine Learning Techniques for Subtle Fault Detection: 82nd EAGE Annual Conference & Exhibition Workshop Programme, **2020**, 1–5.

Hussein, M., R. R. Stewart, and J. Wu, 2021, Which seismic attributes are best for subtle fault detection? Interpretation.

La Marca-Molina, K., C. Silver, H. Bedle, and R. Slatt, 2019, Seismic facies identification in a deepwater channel complex applying seismic attributes and unsupervised machine learning techniques. A case study in the Taranaki Basin, New Zealand.: 89th Annual International Meeting of the SEG, Expanded Abstracts, 2059–2063.

Li, F., J. Qi, B. Lyu, and K. J. Marfurt, 2017, Multispectral coherence: Interpretation, **6**, T61–T69.

Li, J., S. Mitra, and J. Qi, 2020, Seismic analysis of polygonal fault systems in the Great South Basin, New Zealand: Marine and Petroleum Geology, **111**, 638–649.

Lubo-Robles, D., T. Ha, S. Lakshmivarahan, and K. J. Marfurt, 2019, Supervised seismic facies classification using probabilistic neural networks: Which attributes should the interpreter use? 89th Annual International Meeting of the SEG, Expanded Abstracts, 2273–2277.

Lubo-Robles, D., T. Ha, S. Lakshmivarahan, K. J. Marfurt, and M. J. Pranter, 2020, Exhaustive Probabilistic Neural Network for attribute selection and supervised seismic facies classification: Interpretation, 1–67.

Luo, Y., W. G. Higgs, and W. S. Kowalik, 1996, Edge detection and stratigraphic analysis using 3-D seismic data: 66th Annual International Meeting, SEG, Expanded Abstracts, 324–327.

Machado, G., A. Alali, B. Hutchinson, O. Olorunsola, and K. J. Marfurt, 2016, Display and enhancement of volumetric fault images: Interpretation, **4**, SB51–SB61.

Marfurt, K. J., R. L. Kirlin, S. L. Farmer, and M. S. Bahorich, 1998, 3-D seismic attributes using a semblance-based coherency algorithm: Geophysics, **63**, 1150–1165.

Maurya, S. P., and N. P. Singh, 2018, Application of LP and ML sparse spike inversion with probabilistic neural network to classify reservoir facies distribution - A case study from the Blackfoot field, Canada: Journal of Applied Geophysics, **159**, 511–521.

Morley, C. K., A. Maczak, T. Rungprom, J. Ghosh, J. A. Cartwright, C. Bertoni, and N. Panpichityota, 2017, New style of honeycomb structures revealed on 3D seismic data indicate widespread

diagenesis offshore Great South Basin, New Zealand: Marine and Petroleum Geology, **86**, 140–154.

Pedersen, S. I., T. Randen, L. Sonneland, and O. Steen, 2002, Automatic 3D fault interpretation by artificial ants: 64th EAGE Conference & Exhibition, cp-5.

Qi, J., B. Zhang, B. Lyu, and K. Marfurt, 2019a, Seismic attribute selection for machine-learning-based facies analysis: Geophysics, **85**, O17–O35.

Qi, J., B. Lyu, X. Wu, and K. Marfurt, 2020, Comparing convolutional neural networking and image processing seismic fault detection methods: 90th Annual International Meeting of the SEG, Expanded Abstracts, 1111–1115.

Qi, J., B. Lyu, A. AlAli, G. Machado, Y. Hu, and K. Marfurt, 2019b, Image processing of seismic attributes for automatic fault extraction: Geophysics, **84**, O25–O37.

Qi, X., and K. Marfurt, 2018, Volumetric aberrancy to map subtle faults and flexures: Interpretation, **6**, T349–T365.

Randen, T., S. I. Pedersen, and L. Sønneland, 2001, Automatic extraction of fault surfaces from three-dimensional seismic data: 71st Annual International Meeting of the SEG, Expanded Abstracts, 551–554.

Roden, R., 2021, Unsupervised Machine Learning Techniques for Subtle Fault Detection - Geophysical Insights: Https://Www.Geoinsights.Com/.

Roden, R., T. Smith, and D. Sacrey, 2015, Geologic pattern recognition from seismic attributes: Principal component analysis and self-organizing maps: Interpretation, **3**, SAE59–SAE83.

Sahoo, T., P. King, K. Bland, D. Strogen, R. Sykes, and F. Bache, 2014, Tectono-sedimentary evolution and source rock distribution of the mid to Late Cretaceous succession in the Great South Basin, New Zealand: The APPEA Journal, **54**, 259–274.

Specht, D. F., 1990, Probabilistic neural networks: Neural Networks, **3**, 109–118.

Taner, M. T., J. D. Walls, M. Smith, G. Taylor, M. B. Carr, and D. Dumas, 2001, Reservoir characterization by calibration of self-organized map clusters: 71st Annual International Meeting, SEG, Expanded Abstracts, 1552–1555.

Wu, X., Y. Shi, S. Fomel, and L. Liang, 2018, Convolutional neural networks for fault interpretation in seismic images: 88th Annual International Meeting of the SEG, Expanded Abstracts, 1946–1950.

Wu, X., L. Liang, Y. Shi, and S. Fomel, 2019, FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation: Geophysics, **84**, IM35–IM45.

Xiong, W., X. Ji, Y. Ma, Y. Wang, N. M. AlBinHassan, M. N. Ali, and Y. Luo, 2018, Seismic fault detection with convolutional neural network: Geophysics, **83**, O97–O103.

Zhao, T., and P. Mukhopadhyay, 2018, A fault-detection workflow using deep learning and image processing: 88th Annual International Meeting of the SEG, Expanded Abstracts, 1966–1970.

Zhao, T., V. Jayaram, A. Roy, and K. J. Marfurt, 2015, A comparison of classification techniques for seismic facies recognition: Interpretation, **3**, SAE29–SAE58.

Chapter 3. Fault enhancement comparison between coherence enhancement, probabilistic neural networks, and convolutional neural networks in the Taranaki Basin area, New Zealand.

Jose Pedro Mora, Heather Bedle, and Kurt J. Marfurt

Abstract

Fault identification is a key aspect in seismic interpretation and model generation; sometimes, accurate and precise fault models are needed to identify reservoirs or solve complex problems in the subsurface. Nowadays, specialized attributes and machine learning techniques are used extensively to enhance faults and structural features. Coherence enhancement, convolutional neural networks, and plenty of data analytics methods are employed to map faults, but subtle differences in results from these techniques can be found when comparison arises.

We tested three fault enhancement methods in the Taranaki Basin using coherence enhancement probability neural networks (PNN) and convolutional neural networks (CNN) by using fault object extraction to estimate differences between workflows. Although all approaches produce enhanced fault volumes, differences due to artifacts and spatial locations are found. We also use two different datasets, one volume with non-conditioned amplitude values and a second volume with a cleaner and adjusted amplitude conditioning before all workflows are applied; this shows that data preconditioning will help the interpreter remove noise and unrelated features.

Introduction

Seismic interpretation is a powerful way to understand complex geologic structures, including sedimentation patterns and level changes in the past; in order to generate this knowledge, structural interpretation and horizon mapping need to be completed. The structural framework is a key aspect in the interpretation task, and the geoscientist needs to be accurate and precise to identify fault systems; but fault identification is a time-consuming, manual, sometimes uncertain, yet critical task in geologic modeling. Many tools are available as an aid to seismic interpreters (Pedersen et al., 2002; Goldner et al., 2015); some of those are focused on horizons, but, nowadays, specialized attributes and machine learning techniques are being implemented to enhance faults and structural features.

Structural features, such as faults or folds, are better represented in seismic datasets with seismic attributes. Attributes quantify seismic characteristics related to structural, stratigraphic, and discontinuity properties (Chopra and Marfurt, 2005; Barnes, 2016). Seismic discontinuity estimations are obtained through energy ratios and derivatives, and results include covariance, semblance, and coherence (Bahorich and Farmer, 1995; Marfurt et al., 1998; Gersztenkorn and Marfurt, 1999; Barnes, 2016). Dip, azimuth, and curvature for structural and stratigraphic features are estimated using dip scanning, complex seismic trace analysis, and gradient squared tensor (Chopra and Marfurt, 2005; Barnes, 2016). Still, additional features such as channels, mass transport deposits, salt, and other geologic features might be sharpened using seismic attributes due to their incoherent nature and noise.

Coherence enhancement processes start with fault feature enhancing, and additional post-processing creates fault probability volumes from initial inputs such as coherence or variance.

Innovative approaches to fault probability mapping include new attributes and voxel connectivity analysis (Hale, 2013; Wu and Hale, 2016). Barnes (2016) proposed eigenvector methods to find a way to enhance faults and carried out a coherence analysis using dilation to find voxel distortions in neighborhoods. Using a similar approach, Machado et al. (2016) applied a Laplacian of Gaussian operator to enhance coherence attributes constructing a fault probability, fault dip magnitude, and fault dip azimuth volumes. Qi et al. (2018 and 2019) adjust the Machado et al. approach by developing an analysis window to trace features along with fault dip and azimuth, using interpolation in candidate fault points; they finally created a skeletonized volume of fault probability. Other approaches use Swarm intelligence algorithms to improve fault detection (Randen et al., 2001; Pedersen et al., 2002; Hale, 2013; Wu and Hale, 2016).

Data analysis techniques bring new light to solve fault identification by using convolutional neural networks (CNN) to recognize elements or patterns in a similar aspect as image analysis applications. CNN estimates different image filters and compares several results to extract feature characteristics automatically. In seismic interpretation, facies detection is the most common use of CNN identifying salt differentiation, mass transport deposits, etc. (Zhao et al., 2015; Zhao and Mukhopadhyay, 2018; Qi et al., 2019a), but it is also used to identify faults (Wu et al., 2018, 2019; Xiong et al., 2018; Qi et al., 2020). Some supervised CNN classification methods are based on initial training samples; they influence the final fault detection; if the initial training is biased to a specific dataset or model, results will show reduced precision and artifacts in fault detection (Zhao and Mukhopadhyay, 2018; Qi et al., 2020)

In addition to the previous approach, probabilistic neural networks (PNN), a feedforward neural network, are well-known in pattern recognition problems and can be employed to detect and

classify facies (Maurya and Singh, 2018; Lubo-Robles et al., 2019). A supervised approach to facies classification includes a set of seismic attributes and a training dataset such as PNN that evaluates seismic attributes and estimates the best characteristics for further analysis (Maurya and Singh, 2018; Lubo-Robles et al., 2020). Lubo-Robles et al.'s (2020) classification workflow generates two different results, a probability volume and a classified volume for selected features (Lubo-Robles et al., 2019, 2020). Hussein et al. (2020) performed a self-organized map (SOM) classification coupled with principal component analysis (PCA) to highlight subtle faults in the Taranaki Basin. The faults require significant time and experience to be mapped using conventional methods due to small throws (Hussein et al., 2021); they defined key attributes to classify faults as dip magnitude, multispectral energy ratio similarity, total aberrancy magnitude, and curvedness.

However, one of the seismic interpretation goals is to create interpreted faults and horizons as digital objects. A standard traditional workflow to map faults in seismic datasets could use an amplitude volume co-rendered with a coherence volume to provide a means to visualize both values at the same time. Fault mapping begins with digitalization in a time slice to provide the main fault direction and extension; then, the user can switch to an arbitrary vertical section using an appropriate direction perpendicular to the fault trend; by doing this, the user can see the extension of the fault in the vertical plane, and pick samples usually named fault sticks. The number of points picked determines the detail in the mapped features; consequently, a highly accurate analysis will require several points, but regional mapping will require a small number of those points. High amplitude wavelets provide an accurate position to pick fault positions since sometimes attributes provide non-precise fault positions due to stair-step artifacts (Lin and Marfurt, 2017). Fault mapping continues by moving to the next section until mapping the entire

fault. A smooth surface is obtained using corrections due to structural styles identified in the seismic dataset.

Of course, automatic tools are also employed in fault mapping. Automation in large 3D data sets has been achieved in horizon picking by auto trackers (Dorn, 1998); nonetheless, manual and semi-automatic digitizing is still a typical method to create fault objects as fault sticks or three-dimensional meshes. Previous efforts to obtain fault objects by semi-automatic methods (Admasu et al., 2006; Mora et al., 2020) and automatic techniques are documented (Admasu et al., 2006; Cohen et al., 2006; Hale, 2013; Wu and Hale, 2016). Diverse algorithms used in fault extraction are focused on ant tracking following regions with high probability fault objects, and then fault objects are merged into larger features. As an example, Hale (2013) extracted meshes using fault likelihood algorithms, calculating strikes and dips from seismic volumes. All these approaches could use defining seed points to create these surfaces.

Computer-intensive processes such as attribute calculation, PNN, and CNN can be used as an aid to the interpreter. This research aims to identify how effective these approaches are to detect faults compared to standard manual workflows. Moreover, we identify differences between coherence enhancement processes, PNN, and CNN fault detection methods employing fault object generation from a seismic image.

We begin our study with a review of the geologic setting and the data used in the interpretation. Then we briefly review the three alternative fault delineation workflows and show the effect of post-migration data conditioning on each of them. Next, we describe how we construct fault objects and define a metric that compares the computer-generated fault objects to those

manually picked by a human interpreter. We conclude with a discussion of our results and a summary of the relative interpreter effort and computational cost.

Geological setting

Located on the western side of the North Island in New Zealand (Figure 3.1); the Taranaki Basin is tectonically active due to its closeness to the Hikurangi subduction margin (Ballance et al., 2017; Bull et al., 2018; Coleman, 2018). Due to its location near an active margin, the Taranaki Basin has been affected by complex structural processes resulting in a plethora of normal and reverse faults (King and Thrasher, 1996; Kroeger et al., 2013; Haque et al., 2016; Reilly et al., 2016; Coleman, 2018). The abundance of faulting in this area made it a prime candidate for this research. This basin's structural and sedimentary history dates back to the Paleozoic when the Gondwana continent was formed by terrane accretion; these processes led to the basin basement development (King and Thrasher, 1996; Kroeger et al., 2013; Coleman, 2018).

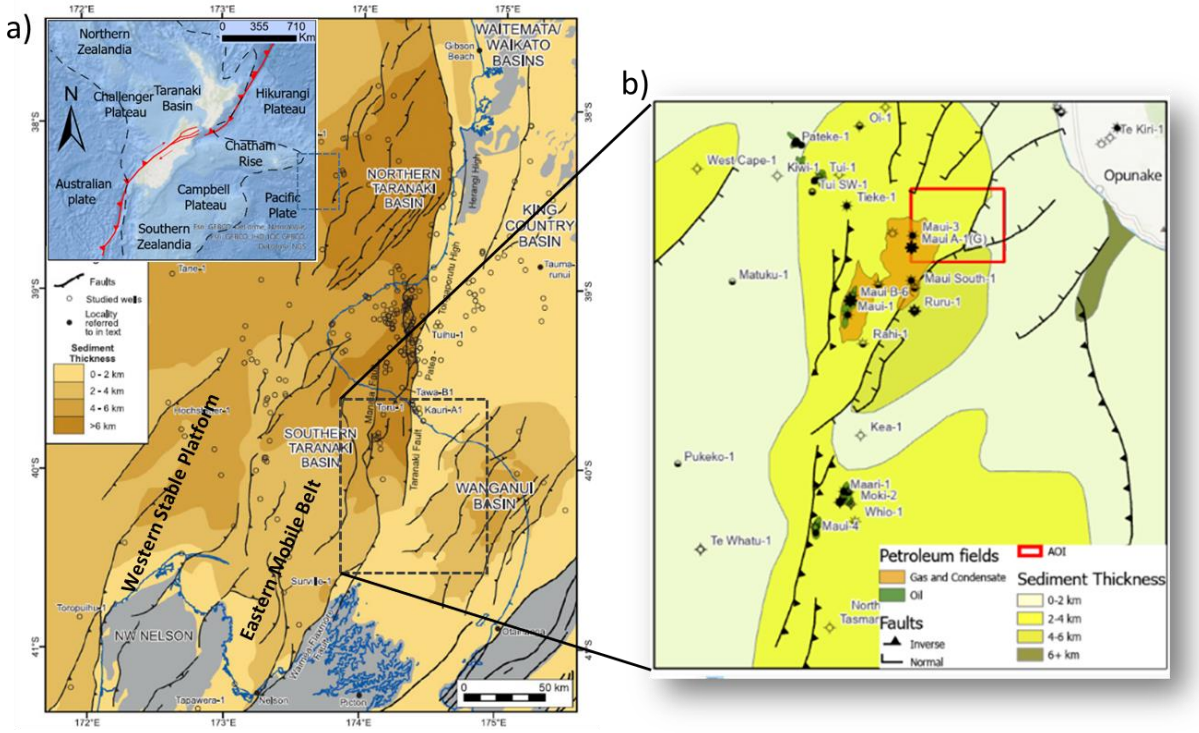


Figure 3.1. (a) General map of the Taranaki Basin showing major structural areas, faults, and sediment thickness. (After Strogen et al., 2014). The Western stable platform is mainly layer cake depositional environment, and the Eastern mobile belt is a more complex and active setting with active, reverse faults, grabens, and inversion structures (King and Thrasher, 1996; Kroeger et al., 2013; Haque et al., 2016; Reilly et al., 2016; Coleman, 2018). (b) Map of the Southern Taranaki Basin showing faults and petroleum fields showing sediment thickness near the analysis area (After Reilly et al., 2016).

Significant events in the Taranaki Basin region mark the current state of the basin starting from Gondwana break-up and the subsequent Tasman Sea rifting (~80-~55 my) to create a sedimentary deposition center controlled by normal faulting, whereas passive margin

configuration continued through the Early Eocene in the southern Taranaki Basin (King and Thrasher, 1996; Kroeger et al., 2013; Strogen et al., 2014; Reilly et al., 2016; Coleman, 2018).

Basin shortening (~43-~5 my) creates a structural setting where compressional forces create a foreland basin under a marine transgression period (Kroeger et al., 2013; Strogen et al., 2014), Middle Miocene shortening resulted in fault inversion, and fold-thrust belt propagation is found across the southern basin (Bierbrauer et al., 2008; Strogen et al., 2014; Franzel and Back, 2019). A final extensional configuration started in the northern area and migrated to the south, triggering the opening of northern and Southern grabens (King and Thrasher, 1996; Bierbrauer et al., 2008; Kroeger et al., 2013; Strogen et al., 2014; Reilly et al., 2016).

Cape Egmont Fault Zone

The research area is located to the Central Graben's western side and close to the Cape Egmont Fault Zone (CEFZ) (Figure 3.1). Several authors mark the CEFZ as the boundary between the Central Graben area and the Western Platform. CEFZ extends over 200 km in a north-south trend and 15 to 20 km wide, the region displays en echelon normal and reverse faults in a continuous downthrown to the southeast (King and Thrasher, 1996; Giba et al., 2012; Reilly et al., 2016).

The Cape Egmont Fault had an initial normal and left lateral movement during Cretaceous, then reversed back in Miocene to have a last right lateral oblique movement from 2 to 3 my. (Nodder, 1993; King and Thrasher, 1996; Giba et al., 2012; Reilly et al., 2016).

The area of analysis near the CEFZ has several fault planes, and some of these features might be mapped easily; nevertheless, detailed manual mapping is an overwhelming task for an

interpreter due to fault extension and density. Normal faulting related to CEFZ and additional fractures are present in the area. (Nodder, 1993; Giba et al., 2012; Reilly et al., 2016; Perico and Bedle, 2020).; consequently, plenty of enhancement and testing can be achieved.

Datasets

The dataset used in this research was obtained in the southern Taranaki Basin, New Zealand, a merged volume created from the Trestles, Maui, and Tui 3D surveys. This dataset includes the Maui Field and geomorphologic features, such as channels, and active structural features, such as the Cape Egmont Fault system.

WesternGeco completed the Trestles 3D Merge PSDM project for Todd Exploration Ltd., with the processing involving Trestles, vintage Maui, and Tui surveys from field tapes. Trestles, Maui, and Opunake surveys were acquired in the N-S direction, and Tui Survey was acquired in a NE-SW direction (Gong and Fell, 2016). Post-imaging processing was completed in February 2016. All the survey data was updated to match the NZGD2000 datum (Figure 3.2).

This dataset is zero phase with a bin size 25 x 12.5 m with a 2 ms sample time. SEG reverse polarity convention is used in the result (acoustic impedance increase by a zero-phase wavelet centered on a trough and a negative number on tape). The pre-stack Kirchhoff time migrated full angle stack volume is cropped to inline range 1750 to 2450, crossline range 1900 to 300, and vertical range 200 ms to 2500 ms.

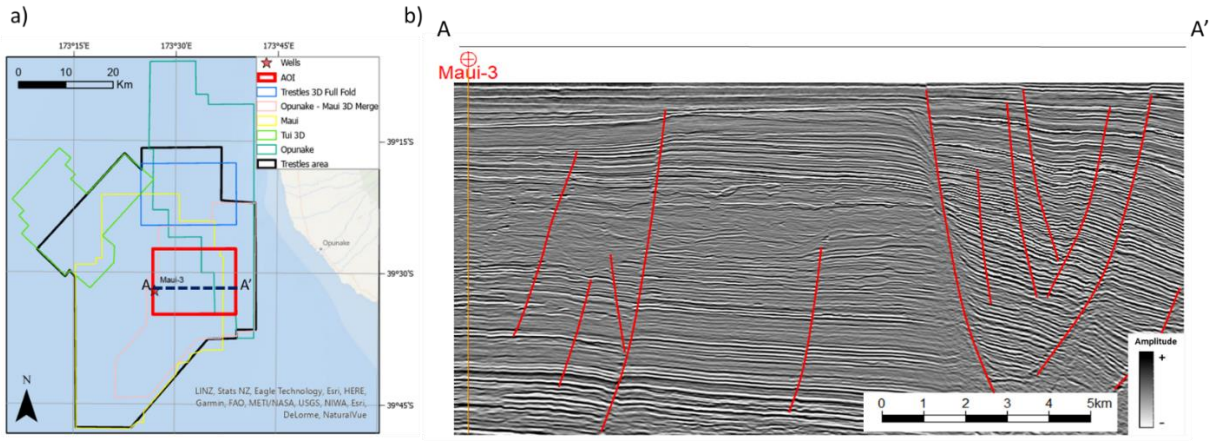


Figure 3.2. (a) Merged 3D seismic surveys included in the 3D Trestles program (After Gong and Fell, 2016). The red square marks the area analyzed in this paper. (b) Representative vertical slice AA' through the merged seismic data volume showing an intricate faulting pattern to the east and additional faulting to the west.

Methodology

We define an analysis area focusing on fault presence in different regions to start a manual interpretation of relevant features as a second step. We map relevant faults following a standard workflow, such as the one presented earlier, in two stages. First, we did an initial interpretation every twenty lines to identify general trends using seismic amplitudes and attributes such as variance and coherence. We use the initial interpretation in subsequent steps to extract faults objects from fault probability volumes. The second stage includes a detailed digitalization of the same features to compare fault object extraction. A visual schematic of the method is presented in Figure 3.3.

Data conditioning is employed to remove the acquisition footprint and unwanted features by applying footprint removal, spectral balancing and a structural oriented filtering (Ha et al., 2019). Spectral balancing increases bandwidth and frequency content, improves the signal-to-noise ratio, leading to a better definition of structures in the dataset by removing the limitation on bandwidth and signal-to-noise ratio (Chopra and Marfurt, 2016).

A structural filter is applied to remove noise, small-scale features and simplify structural events by performing a smoothing operation parallel to clear reflectors, avoiding seismic reflection terminations (Fehmers and Höcker, 2003b). Two different tests are applied, an initial approach by enhancing faults using a non-conditioned amplitude volume and a second workflow using a preconditioned dataset.

Coherence enhancement

The Coherence enhancement workflow is based on a 3D fault directional filtering of seismic attributes implemented by Machado and Qi (Machado et al., 2016; Qi et al., 2019b) and creates fault probability, dip magnitude, and dip azimuth volumes.

After data conditioning, we compute multispectral coherence (Li et al., 2017; Marfurt, 2017) as the primary input dataset to apply a directional analysis to sharpen fault features in a seismic attribute volume. In the next step, the coherence enhancement filter searches the closest coherent values within a 3D window in each voxel. Eigenvector analysis defines the importance of discontinuities, identifies fault planes, and enhances those planes to estimate a continuous fault probability feature and create additional dip magnitude calculation (Qi et al., 2019b).

An additional step includes a fault sharpening using the previously computed fault dip magnitude and azimuth. Skeletonization searches faults and unconformities along planes, discarding lower candidates in the analysis window to produce a sharp image from the fault volume (Qi et al., 2019b).

Probabilistic neural networks

PNNs are neural networks with hidden layers connected in the output layer's direction; in this approach, a probability density function (PDF) uses a Parzen window and a non-parametric function to approximate the initial class (Specht, 1990). By using this method, the probability is better estimated, and misclassification is minimized. This classification method has four layers; the first input layer calculates the sample distance to the training vectors and produces a measurement of how close the sample is to the initial training dataset. The second layer then adds all the calculated contributions and creates a probability vector converted to binary classification in the final output layer (Specht, 1990; Lubo-Robles et al., 2019).

Lubo-Robles et al. classify facies using training from initial relevant attributes (Lubo-Robles et al., 2019, 2020); but our approach is focused on fault classification; then, we use the Hussein et al., (2020) analysis of critical seismic attributes for unsupervised classification methods in the Taranaki Basin. The initial attribute analysis includes multispectral energy ratio similarity, dip magnitude, total aberrancy magnitude, curvedness, and gray-level co-occurrence matrices (GLCM) homogeneity and entropy and additional attributes such as multispectral volumes from energy ratio and Sobel filter similarity. We then create a training dataset to analyze the data and generate a model to classify the entire volume.

Initially, we use the initial set of faults as a training and validation dataset. Next, we use a Z-score calculation and normalization to compare faults and no-fault features across different seismic attributes. PNN analysis estimates the error and best sigma estimation in all attribute combinations to choose the optimal set of seismic attributes to use in the classification stage.

The classification process generates a fault probability volume, which is sharpened by using fault enhancement processes to obtain a skeletonized fault probability, dip magnitude, and dip azimuth volumes (Qi et al., 2019b).

Convolutional neural network

CNN is a deep neural network class focused on pattern recognition and is employed extensively in image analysis (Yamashita et al., 2018). CNN models learn from input features by dividing all the features into hierarchies or pyramids to understand the features as a whole by moving from low to high-level characteristics. In general, CNNs are constructed by an input layer, hidden layers, and an output layer (Qi et al., 2020). The convolutional layer convolves or filters small pieces of information using mathematical operations and down samples the data by summarizing how the convolution modifies the information. Convolution and pooling identify patterns and perform a feature extraction; meanwhile, the connected layer maps all these extracted features in a final classification (Yamashita et al., 2018).

The CNN classification method we use is based on U-Net architecture (Ronneberger et al., 2015) that optimizes the number of sample training. However, the U-Net basic architecture is modified by Qi et al., (2020) to estimate fault probabilities. This model training uses synthetic seismic amplitude volumes to accurately provide several samples and a synthetic faulting model (Wu et

al., 2019; Qi et al., 2020). This classification method uses an amplitude volume as input and provides a fault probability volume as the output.

As the last step, a fault enhancement process (Qi et al., 2019b) is applied to create a clear and sharp image with fault probability, dip magnitude, and dip azimuth volumes.

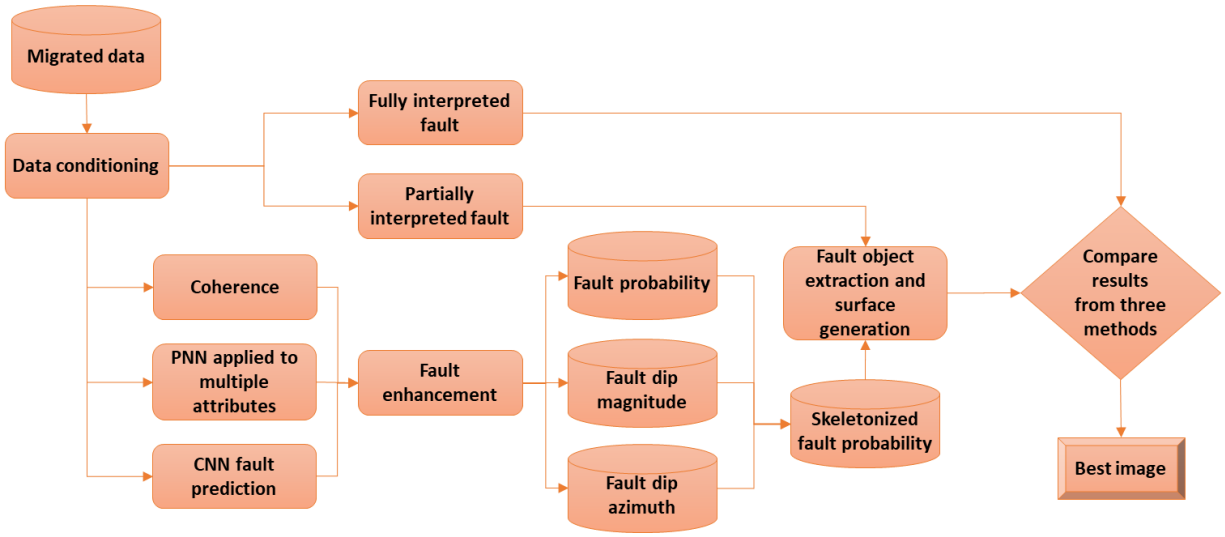


Figure 3.3. Summary of our proposed workflow. We begin with postmigration data conditioning to suppress crosscutting noise and balance the spectrum and manually pick the main faults to form a baseline human interpretation. Then we evaluate three different fault delineation workflows. The results of each workflow are enhanced to provide skeletonized fault probability, fault dip magnitude, and fault dip azimuth; then, we use active contours to construct fault objects. Finally, the three fault predictions are compared with the manual baseline to determine which method provides the best image. We expect the best method to be close to the interpreted surface.

Fault extraction and comparison

Fault probability extraction aims to create fault objects from an initial mapped object close to the fault location and fit the object to the most probable location from each volume to compare and estimate differences with a manually interpreted fault.

Active contour is a well-known algorithm in computer vision (Kass et al., 1988) and tries to solve boundaries of shapes in images when the approximated shape is known. This algorithm changes the shape of the polyline and fits the fault probability by searching the image's maximum values; it also can create illusory contours fitting values where boundary data is absent (Kass et al., 1988). We use the active contour method to scan line by line using a coarse grid of fault sticks around a manually interpreted fault to serve as the analysis's seed points. The active contour scans the proposed fault location and moves from low probability values to high probability values to fit the shape close to the enhanced fault volume.

The final step consists of a surface generation and comparison with a manual interpreted fault. Hausdorff distance (Cignoni et al., 1998) is employed widely in surface and mesh analysis; generally, it is the greatest of all distances from one point in a surface to any location on the compared surface. The manual interpreted fault presents the distance between surfaces as a color-coded attribute.

Results

There are two different sources of what we call noise when analyzing fault probability volumes. Acquisition noise and ‘geologic noise’, the first type is related to random and coherent noise, and

sources such as ambient, wave propagation, data acquisition, and processing are normally what we can find in this category. The 'geologic noise' is related to features mapped in the seismic volume but not related to faults; an example of these is coherence enhancement of channel fill, oxbows, levees, volcanic related, and other small non-structural features. We call these artifacts 'noise' because our main objective is fault enhancement avoiding all non-fault enhanced features.

We use two data conditioning steps to remove both types of noises; first, a footprint removal process to remove coherent, periodic noise. Second, we use a structural oriented filtering to remove other noises such as geomorphology features as channel, oxbows, levees, or bars. Additional post-enhancement filtering can be removed using the fault enhancement step in the Qi et al. workflow.

Overall, data conditioning reduces seismic noise and removes uncoherent reflections related to uncoherent features such as channels. Figure 3.4 shows six seismic attributes calculated from non-conditioned data and preconditioned data. Dip magnitude shows extreme values reduction and footprint removal. In the similarity case, calculations showed a stark difference between the initial dataset and preconditioned data; a high amount of noise and non-fault features are minimized in the south and center of the area, and the acquisition footprint to the east is also minimized.

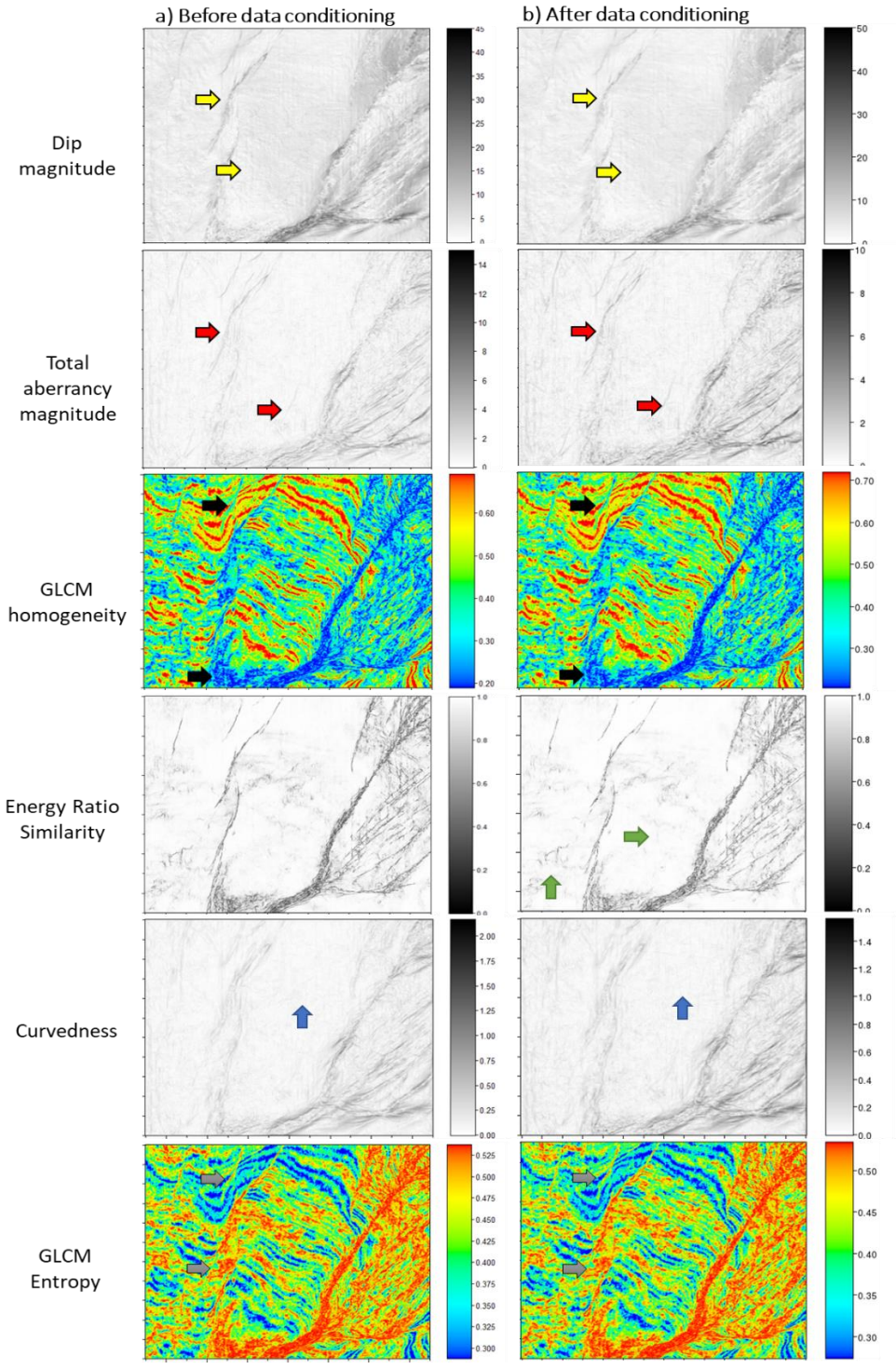


Figure 3.4. Comparison between attributes computed from the seismic amplitude volume (a) before and (b) after data conditioning. The conditioned dip magnitude is piecewise smoother and shows less footprint (yellow arrows). The conditioned energy ratio similarity exhibits fewer

incoherent features (green arrows) while enhancing the stronger fault stratigraphic edges; likewise, the conditioned total aberrancy exhibits fewer artifacts (red arrows).

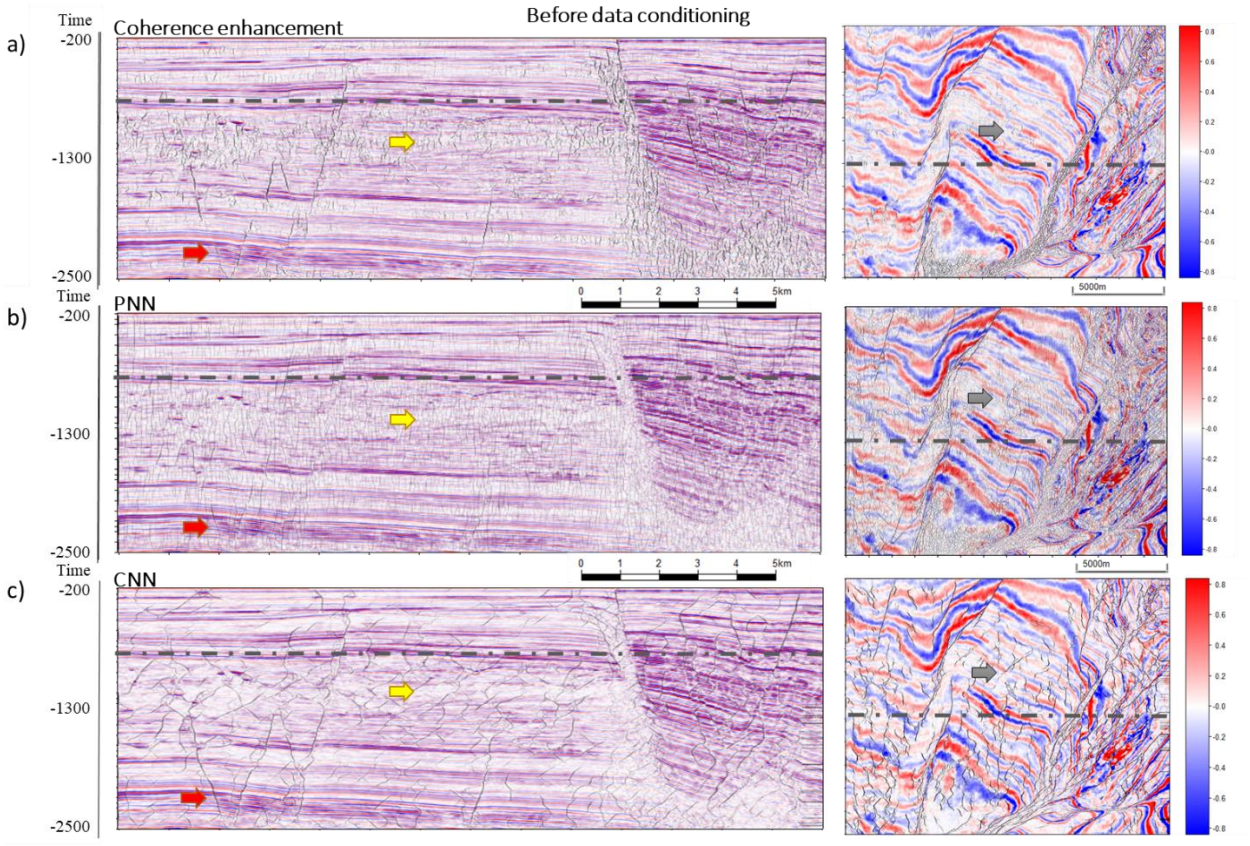


Figure 3.5. Fault enhancement results using non-conditioned data showing Coherence Enhancement (a), PNN (b), and CNN (c) result in a seismic section (IL 2400) and time slice (0.76 s) with fault amplitude data and corendered fault skeletonization. Coherence enhancement results (a) suggest the method classifies minor fractures and features, but stairs-steps and discontinuities between faults are still evident. PNN (b) results show more subtle features than coherence enhancement with a salt-and-pepper-like result. CNN (c) extracts large faults and relevant structures, avoiding small fractures and faults.

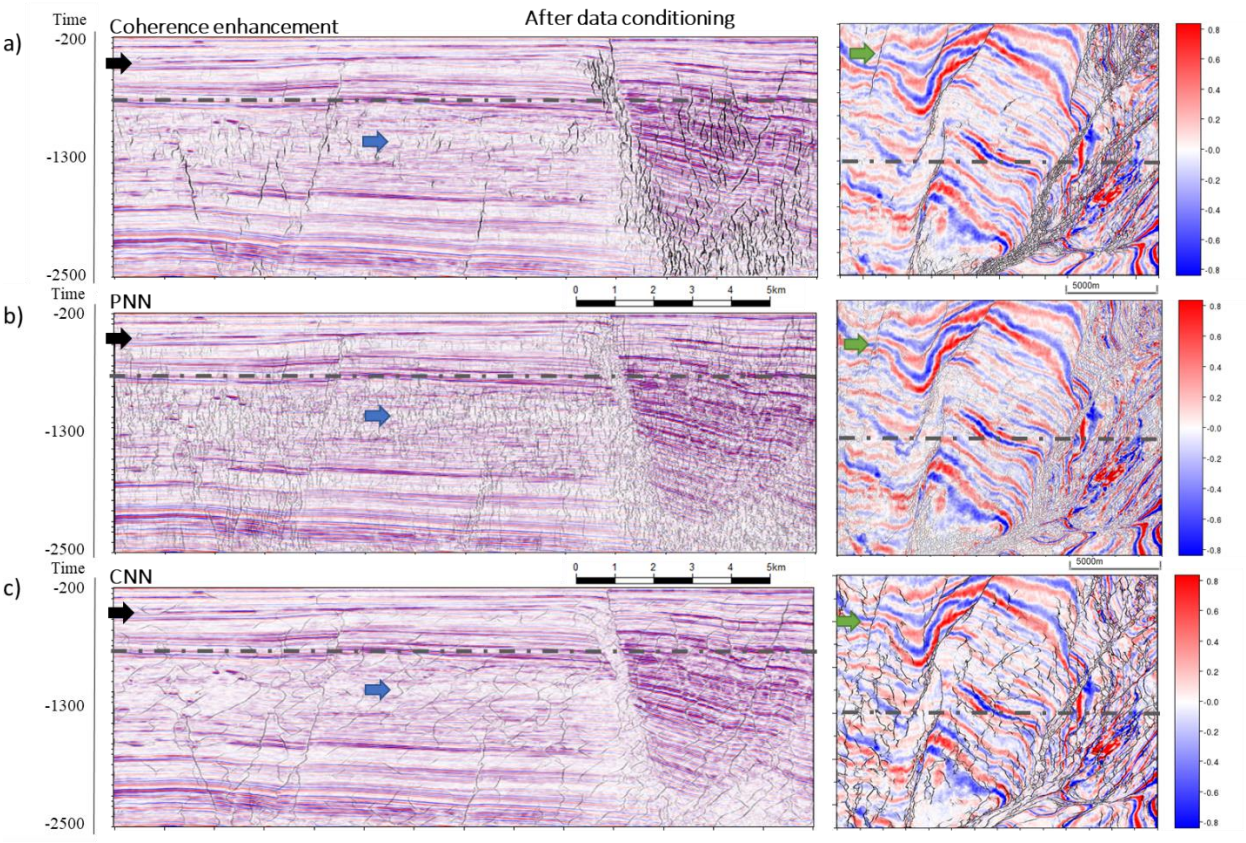


Figure 3.6. Fault enhancement results using conditioned data showing Coherence Enhancement (a), PNN (b), and CNN (c) results in a seismic section (IL 2400) and time slice (0.76 s) with fault amplitude data and corendered fault skeletonization. Coherence enhancement results (a) suggest the method classifies minor fractures and features, but stairs-steps and discontinuities between faults are still evident. PNN (b) results show more subtle features than coherence enhancement with a salt-and-pepper-like result. CNN (c) extracts large faults and relevant structures, avoiding small fractures and faults. The conditioned results show less noise and small

features in the three methods. PNN classification shows aberrancy magnitude, multispectral energy ratio, and Sobel filter similarity (0-34 Hz) as the best attribute combination.

Total aberrancy magnitude and curvedness show small changes between raw and conditioned data in non-fault anomalies removal, but there is an extreme data reduction due to smoothing, as we see the comprised data range. In the GLCM attributes case, homogeneity and entropy arrows mark increased differentiation between faults and surrounding geology, and the upper and lower limit range also increased due to preconditioning.

Coherence enhancement

Figure 3.5a and Figure 3.6a show a seismic section and time slice with coherence enhancement results from non-conditioned and preconditioned data. All images have four distinct areas in all line sections: a shallow section from 200 ms to approx. 900 ms shows continuous reflections with enhanced faulting. A second section starts from 900 ms to 1600 ms and offers an increased noise and artifact presence (yellow arrow, blue arrow). The more in-depth section from 1600 to 2500 ms shows similar results as the shallow section, with continuous reflections and localized faulting. Some faults are not apparent enough in the raw dataset, but preconditioning helps to improve fault detection (red arrow). Eastward in the seismic line, a faulted area shows heavy fracturing and stair-step anomalies. The time section presents areas with well-defined high fault probability values. While high fault probability values dominate the east part, the center and west sides are clean with a few well-defined features. Non-conditioned data show more noise, subtle features, and less sharpness in the fault enhancement result, and the preconditioned dataset shows more consistent and continuous features.

Probabilistic neural networks

Figure 3.5b and Figure 3.6b show PNN results with similar area distribution as the previous workflow. Shallower and deeper areas exhibit more fault probabilities due to the absence of noise and artifacts. In contrast, the middle section has more noisy artifacts (yellow arrow, blue arrow), but some mapped features are found in the seismic section. On the eastern side, plenty of fault-enhanced features fill the CEFZ, but the salt and pepper behavior persists. In contrast to the previous method, weak fault probabilities appear across the section; again, non-conditioned data is noisier than preconditioned attributes.

Convolutional neural network

Figure 3.5c and Figure 3.6c show fault enhancement using the CNN approach. This method produces sharp and crisper fault probability features than previous approaches. The critical area between 900 ms to 1600 ms (yellow arrow, blue arrow) shows less noise than previous methods, but several unrealistic features appear in the analysis (black arrow). Time slice shows the same trend as the section, apparent fault probabilities, and several new unrecognized events are enhanced. Preconditioned data show less unrealistic features, so the result improves in comparison to non-conditioned data.

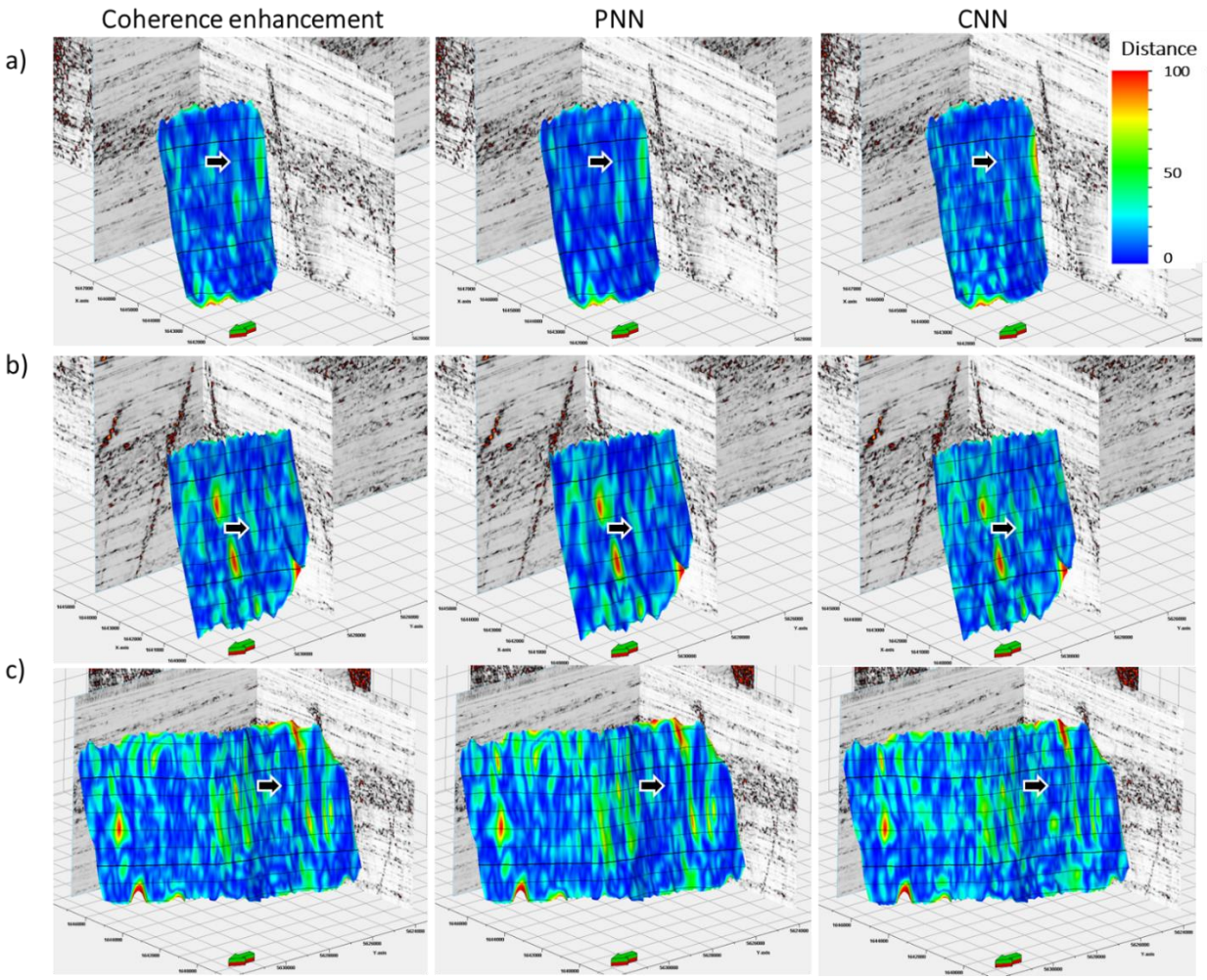


Figure 3.7. Comparison between extracted surfaces and manual interpretation. The Hausdorff distance (Cignoni et al., 1998) calculates the maximum distance between two surfaces in every direction. Color-coded differences help to identify discrepancies between manual and fault objects from classification methods. The image show distance in meters with a conversion time to distance using a 3000 m/s velocity.

Fault extraction and comparison

The fault extraction algorithm generates a point cloud used as input for an interpolation process that creates a continuous surface employed to calculate the Hausdorff distance to manual interpreted faults. After an initial analysis of fault probability results, we decided to use fault probability volumes from the preconditioned dataset because they have less noise and spurious probabilities. We assumed the correct fault position as the manually interpreted surface. Figure 3.7 presents the distance measurement stored as a color bar to analyze differences between the methods easily on three faults inside the analysis area. Visually, the comparison with coherence enhancement workflows shows less salt-and-pepper-like results, and all differences are between 0 and 100m with locally abnormal results. Statistically, the results are not conclusive since the standard deviation from all the workflows is similar; for example, the measured SD in fault one is CNN 15.05, and coherence enhancement is 15.85 (Table 3.1).

	Fault 1			Fault 1			Fault 2		
Fault enhancement Type	Coherence	CNN	PNN	Coherence	CNN	PNN	Coherence	CNN	PNN
Min:	0	0	0	0	0	0	0	0	0
Max:	114.93	121.65	108.18	138.69	132.61	133.77	190.38	181.22	189.73
Delta:	116.63	123.51	108.87	138.69	132.69	133.7	190.39	181.21	189.73
Mean:	14.89	15.62	13.31	18.45	18.76	18.02	20.12	20.58	20.78
Std. dev.	15.06	15.85	13.99	17.39	17.13	17.30	20.33	19.72	20.78
Variance:	226.77	251.25	195.86	302.45	293.28	299.46	413.32	388.89	431.75

Table 3.1. Basic statistic measurements after Hausdorff distance calculations on three different faults, all measured values are positive distances from the fault to any position on the other surface, distance in meters with a conversion time to distance using a 3000 m/s velocity.

Discussion

The seismic volume we analyzed covers a heterogeneous history of deposition and fault geometries, and results on fault enhancement are tied to these geologic conditions; consequently, a simple answer on which method is best is not straightforward. To understand the faults enhancement results in different areas, we present a brief explanation of the seismic reflections found in the area (Figure 3.8).

In Figure 3.8, the Cape Egmont Fault Zone, the hanging wall shows a clean reflector succession with an apparent high layer dip eastward than the footwall; the fault plane is distinguished due to the amplitude reduction in strata reflection near the fault, folding adjacent to both fault sides and poorly organized apparent reflections. The area near and beneath the fault described as a normal fault (Reilly et al., 2016), is surrounded by low amplitude and incoherent seismic traces and distorted visualization of the footwall fault block. This section is located northward the main CEFZ bend showing all the multiple fractures present in the section

Section 1 (Figure 3.8) is a Pliocene-Pleistocene package (Franzel and Back, 2019) found at 200 ms to approximately 900 ms and can be divided into three broad subdivisions. The upper section of this portion (1a) exhibits semi-parallel and semi-continuous reflectors with medium amplitude reflectors on top. The middle part of section 1 displays semi-parallel discontinuous to semi-continuous reflectors with low amplitude. The lower part of section 1 is continuous to semi-continuous oblique reflections with high amplitude thinning toward the section's central structure. Well tops in the Maui 3 borehole show this section relates to Giant Foresets formation, mainly formed by sandy siltstone to mudstone and deposited in clinoforms (Roncaglia et al., 2013; Franzel and Back, 2019).

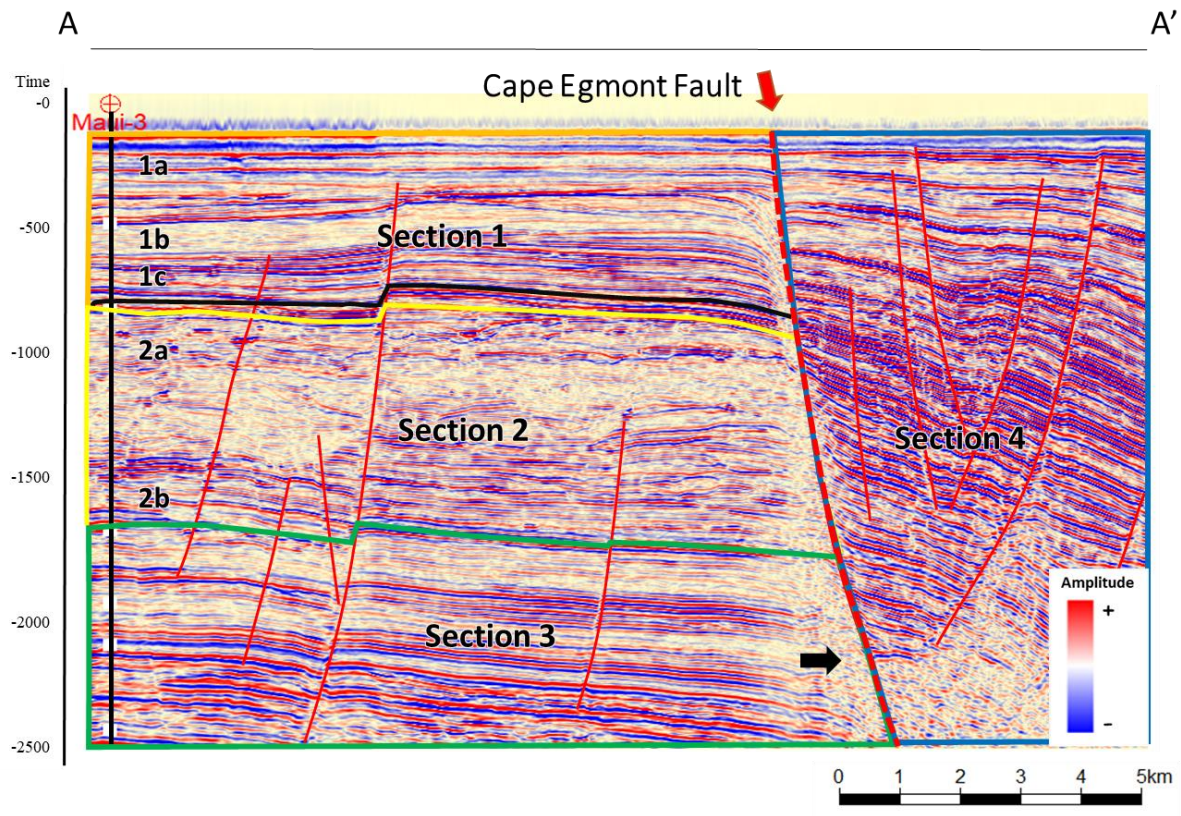


Figure 3.8. A seismic section near the Maui-3 well. Differences in seismic reflections influence the final response of fault enhancement methods and can be broadly related to these four zones; notice the black arrow showing a distorted visualization of the footwall.

Section 2 (Figure 3.8) is a middle-upper Miocene (Franzel and Back, 2019) from 900 ms to 1600 ms; this package shows chaotic and mostly discontinuous reflection low and medium amplitudes (Section 2a). Section 2b is locally mounded with semi-parallel and semi-continuous reflectors with medium amplitude. This section is more disordered than surrounding areas. Well tops show very thick patchy and discontinuous reflectors that match sandstones from the upper Manganui formation (Roncaglia et al., 2013; Franzel and Back, 2019).

Section 3 in Figure 3.8 (1600 ms to 2500 ms) is mostly parallel continuous with moderate to high amplitude reflectors and small semi-parallel areas. Well tops indicate a correlation to medium and lower Manganui formation (Roncaglia et al., 2013; Franzel and Back, 2019). Finally, Section 4 (Figure 3.8) shows semi-continuous and discontinuous reflections, high amplitude to low amplitude, localized oblique, chaotic reflections, and several cut reflectors due to heavy faulting CEFZ influence. Also, acquisition footprint noise heavily contaminates the eastern part of the dataset.

Spectral enhancement improves seismic reflector continuity increasing the detection of faults. However, increasing the frequency might sharpen faults and other discontinuities, so a subsequent structural filter is needed. Data conditioning removes several unwanted features such as acquisition footprint, geomorphologic features, small unconformities, and small discontinuities with low dips.

	Coherence enhancement	CNN	PNN
Attribute calculation	1.5 hours	1.5 hours	7 hours
Classification		1.5 hours	33 hours
Enhancement	0.6 hours	0.6 hours	0.6 hours
Total	**2.1 hours**	**3.6 hours**	**40,6 hours**

Table 3.2. Computational time employed in the entire workflow employing 60 processors and 256 Gb RAM. PNN is the most computationally intensive method ten times longer than the CNN analysis. CNN is the second time-consuming, but the Energy ratio similarity calculation adds 1.5

hours to the classification, reaching a total time of 3 hours. These values are estimated on a 3.5 Gb volume analysis.

Multispectral coherence method enhances minor faults and creates easily distinguishable features; the main drawback is high computational loads due to calculations on multispectral energy ratio and subsequent attribute calculations; some stair-step artifacts are also noticed, although it is not a typical pattern in our results.

CNN method creates distinctive fault probabilities, but several anomalous features appear. The CNN method removed small and uncertain features, enhancing relevant faults; also, it is a fast way to identify fault probabilities, as Qi et al. (2020) also pointed out. Drawbacks are undesired artifacts that might be solved using the skeletonization algorithm but not entirely; additional training might fix these undesired features using real amplitude values and noise to improve the classification model. Another possible source for this result is the previous spectral enhancement that sharpens anomalies all over the volume.

PNN enhances faults, but it shows several discontinuities around the volume (Figure 3.6); faults and fractures, artifacts, noise, and small feature enhancement might mislead the interpreter into mapping unrealistic features. One reason for this unreliable result might be related to the analysis's size; a large window includes more samples and different populations, leading to difficulties in defining subtle divisions in the data. A way to improve the result is to reduce the analysis area. Our analysis reduced the attribute number to four to optimize results and computational time. Initial analysis shows curvedness, GLCM homogeneity, energy ratio similarity, and total aberrancy magnitude enough to classify faults in the analysis area; dip

magnitude attribute was skipped in favor of total aberrancy. A second analysis to improve the PNN classification includes multispectral similarity and aberrancy volumes presented in Figure 3.9; after this test, aberrancy magnitude, multispectral energy ratio, and Sobel filter similarity (0-34 Hz) are selected as the best combination to enhance faults. A drawback in this methodology is the computational time to classify the entire volume in *Table 3.2*.*Table 3.2. Computational time employed in the entire workflow employing 60 processors and 256 Gb RAM. PNN is the most computationally intensive method ten times longer than the CNN analysis. CNN is the second time-consuming, but the Energy ratio similarity calculation adds 1.5 hours to the classification, reaching a total time of 3 hours. These values are estimated on a 3.5 Gb volume analysis.* Still, PNN and the optimized attribute combination did not match the coherence enhancement or CNN results.

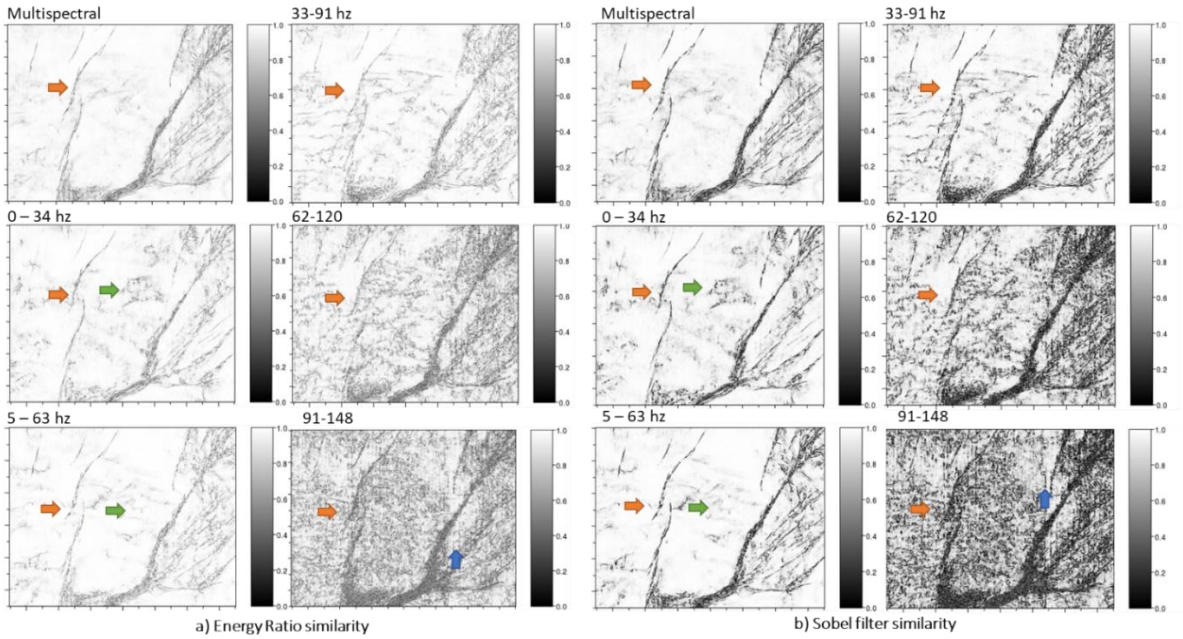


Figure 3.9. Time slices from multispectral coherence analysis, Energy ratio (a), and Sobel filter similarity (b). Higher frequencies show less coherent features, and low frequencies resemble

better and relevant faults. We use multispectral, 0 to 34 Hz., 5 to 63 Hz. and 33 to 91 Hz. volumes in the PNN analysis to find the best combination.

Fault object extraction works well to measure differences between classification methods (Figure 3.10). Dissimilarities using Hausdorff distance could be enough to estimate which classification is closer to the manual interpreted surface, but some areas inside the fault surface (Figure 3.7) show locally pronounced distances to the mapped fault. The initial reason for this lies in the fact that active contours search for high probability values close to values mapped by an interpreter, so high probability values not related to the fault are selected instead of the real position. A second possibility could be that the initial manual interpretation moves the real position away from the fault. These errors are not persistent in all faults.

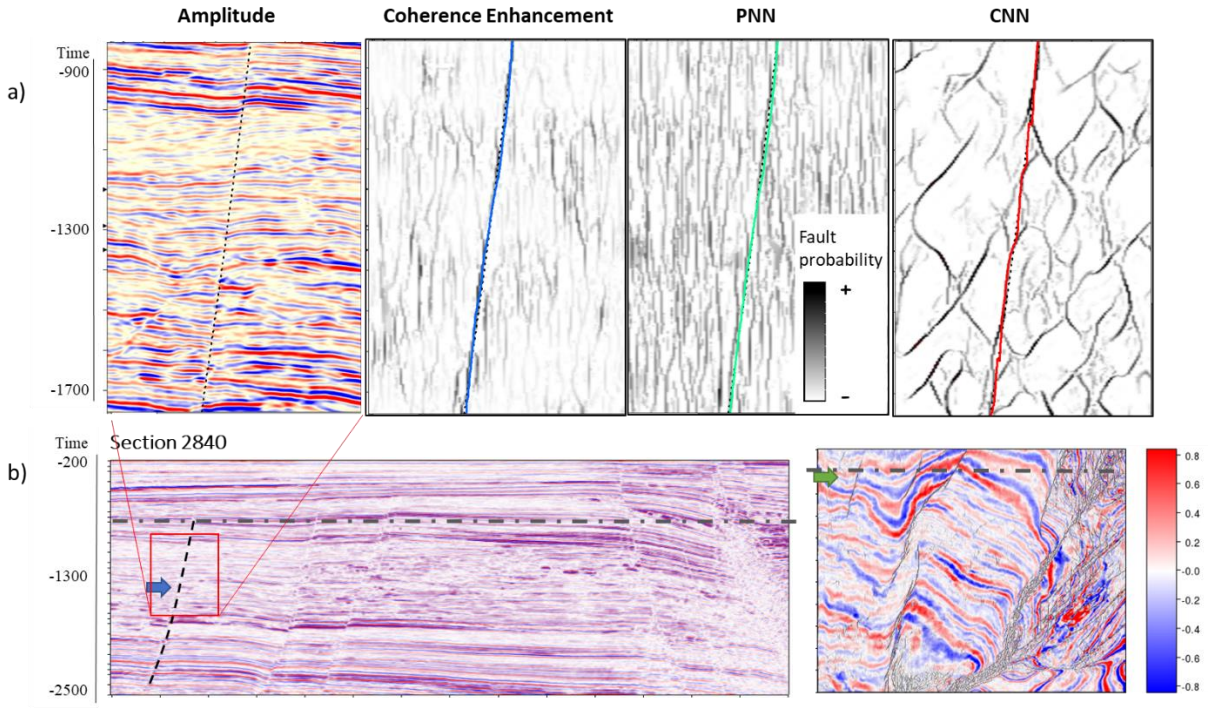


Figure 3.10. Detailed fault enhancement from different workflows (a) and seismic section and time slice location of the detailed area (b). Notice the complex section with chaotic reflector and

semi-continuous subparallel reflectors on amplitude due to channelization, (a) coherence enhancement show values close to the fault such as PNN, CNN has cleaner result but plenty of anomalies and artifacts. PNN is the method that shows the most inferior result due to high noise and vertical artifacts.

Conclusions

Advance fault enhancement methods are helpful to the interpreter and reduce time when working with seismic volumes. Still, a good interpreter should identify the best way to improve these faults accordingly to the time and detail required.

We tested an initial approach to reduce the impact of noise in a seismic fault probability estimation. Our test in the Taranaki Basin shows that seismic noise and non-fault geological features influence all three methods tested. The best result can be found using a preconditioning workflow before the start of each methodology. Although we can reduce artifacts, we must understand the initial approach; for instance, heavy conditioning such as applying a filter several times could erase small features or subtle faults. In that sense, intense data conditioning is suggested in regional analysis to increase the continuity and sharpness in large fault images.

On a small scale, reservoir analysis suggests a more cautious approach, minimizing conditioning by avoiding footprint removal and minimizing filtering to just one step or completely remove the conditioning, if the data quality is enough to enhance faults. Data conditioning removes noise and small features, but it also can remove small faults, so the interpreter must be conscious of the pros and cons of each method.

Figure 3.11 shows a brief conclusion on fault enhancement comparison, analysis type, interpretation use, time constraints, number of calculated attributes, and mean and standard deviation measurements of three selected faults. The CNN method can provide a fast way to classify a volume, but the results must be verified to avoid mapping anomalies or inexistent features. Also, CNN seems to classify larger faults and avoid minor features, desirable in regional analysis, but a drawback in detailed modeling. Classification results can be improved using an increased number of training samples with real data to provide a better fault estimation and fewer artifacts (Qi et al., 2020). CNN was trained using entire faults crossing inside a 128x128x128 sample window (Qi et al., 2020); this training might be the reason for small faults misclassification. In addition, CNN identifies channels and other features as faults; one hypothesis is that to have many artifacts inside the second section could be explained as a discontinuities lineament. The algorithm was trained with normal faults, and discontinuities lineament could be the source of these CNN artifacts.

Coherence enhancement is a precise method based on a classic analysis window approach; it works well since it is possible to minimize stair-step artifacts with the skeletonized process. Also, multispectral coherence improves fault probability estimation (Li et al., 2017). One of the main drawbacks is that plenty of fault probabilities can be enhanced due to noise or stratigraphic factors, and stair-step artifacts might be an issue in some analysis areas. Minor faults are also enhanced, so the method can be used to find those in detailed modeling.

PNN Fault enhancement could be used as a tool to enhance faults; however, as a supervised classification method, good training must be achieved to obtain satisfactory results. PNN classification might work well in small datasets when detail is needed, and integrated analysis of

several attributes is needed, but more research on this topic must be done. In this analysis, PNN results show no noticeable faults enhanced; moreover, salt-and-pepper-like results show a high amount of uncertainty in the classification. Although the fault enhancement process and skeletonization reduce noise still the initial PNN result shows poor results.

Each method shows different results, and additional processes might be applied to remove residual artifacts and non-connected fault probabilities. Some interpretation packages can define thresholds and identify fault probability clusters and connectedness to remove non-connected data and isolated sample probabilities. This additional post-classification process could lead to removing small geologic features and remaining noise mapped as fault probabilities.

In an actual interpretation workflow, occasionally, users extract faults with automatic methods.

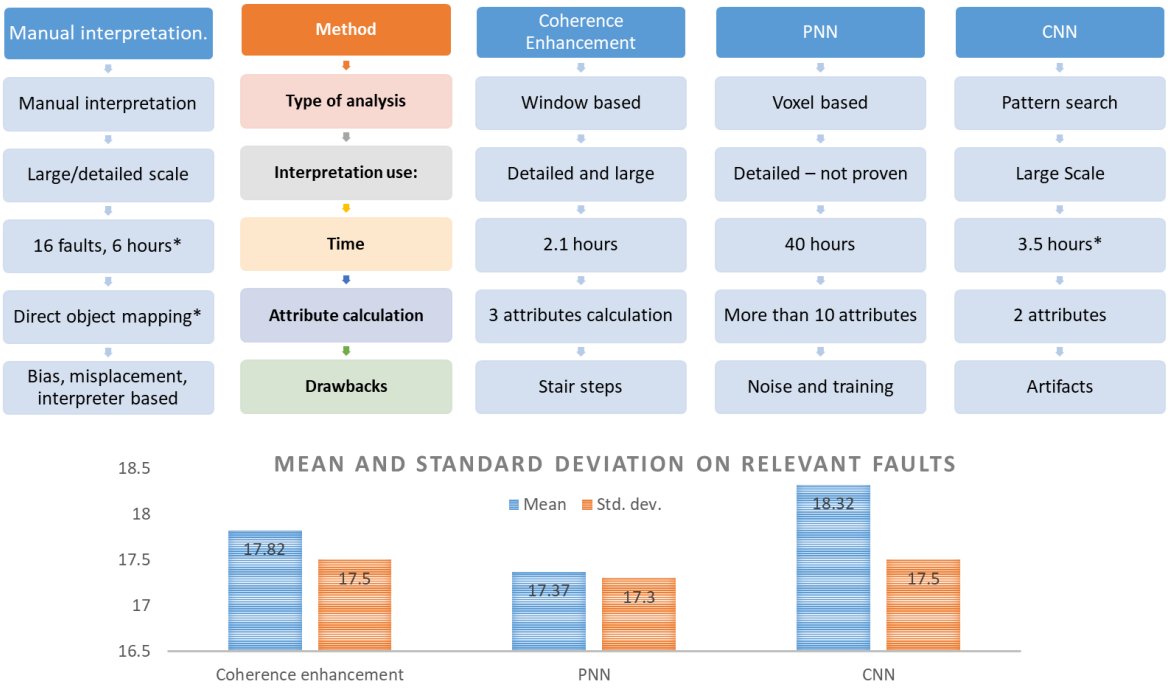


Figure 3.11. Comparison between reviewed methods. This figure shows the main differences between analysis type, interpretation use, processing times, attribute calculation products, and

main drawbacks. All measured values are positive distances from the fault to any position on the other surface, distance in meters with a conversion time to distance using a 3000 m/s velocity; the initial column shows the initial approach to map faults manually.

These fault enhancement workflows are a tool to the interpreter; although it can provide fault images at a very detailed scale, it is also true that geologic background and physical understanding are not included; consequently, an interpreter must validate and extract or map useful features. Images can provide a precise location to fault probabilities, and higher accuracy can be obtained, but user input is also needed to validate images and extracted objects from fault probability volumes. We need to understand that any fault enhancement method and object extraction have uncertainty, and we should be aware of the processes we use and understand their limitations.

Acknowledgments

We would like to thank New Zealand Petroleum and Minerals for providing the data used in this project; seismic surveys, wellbore information, and processing report data are available through the New Zealand Petroleum and Minerals administration website and used under its permission. We also thank the Attribute Assisted Seismic Processing and Interpretation consortium (AASPI) for their continuous and invaluable support. We also acknowledge Schlumberger, which provided Petrel licenses to the University of Oklahoma for research and education use. The first author also wants to express his gratitude to the Fulbright program and the Saldarriaga-Concha Foundation because they made this research possible.

References

Admasu, F., S. Back, and K. Toennies, 2006, Autotracking of faults on 3D seismic data: Geophysics, **71**, A49–A53.

Bahorich, M., and S. Farmer, 1995, 3-D seismic discontinuity for faults and stratigraphic features: The coherence cube: The Leading Edge, **14**, 1053–1058.

Ballance, P. F., L. Cotterall, and Geoscience Society of New Zealand, 2017, New Zealand Geology: An Illustrated Guide:

Barnes, A. E., 2016, Handbook of Poststack Seismic Attributes: SEG Books.

Bierbrauer, K., T. J. Herdy, A. Rek, and K. L. Mills, 2008, Exploring the Greater Taranaki Basin North of the Established Hydrocarbon Fairway: New Zealand Petroleum Conference Proceedings Crown Minerals.

Bull, C. Y. S., A. E. Kiss, E. van Sebille, N. C. Jourdain, and M. H. England, 2018, The Role of the New Zealand Plateau in the Tasman Sea Circulation and Separation of the East Australian Current: Journal of Geophysical Research: Oceans, **123**, 1457–1470.

Chopra, S., and K. J. Marfurt, 2005, Seismic attributes — A historical perspective: Geophysics, **70**, 3SO-28SO.

Chopra, S., and K. J. Marfurt, 2016, Spectral decomposition and spectral balancing of seismic data: The Leading Edge, **35**, 176–179.

Cignoni, P., C. Rocchini, and R. Scopigno, 1998, Metro: Measuring Error on Simplified Surfaces: Computer Graphics Forum, **17**, 167–174.

Cohen, I., N. Coult, and A. A. Vassiliou, 2006, Detection and extraction of fault surfaces in 3D seismic data: Geophysics, **71**, P21–P27.

Coleman, R. D., 2018, Tectonic evolution of Taranaki Basin, offshore New Zealand. M.S. Thesis, Colorado School of Mines.

Dorn, G. A., 1998, Modern 3-D seismic interpretation: The Leading Edge, **17**, 1262–1262.

Fehmers, G. C., and C. F. W. Höcker, 2003, Fast structural interpretation with structure-oriented filtering: Geophysics, **68**, 1286–1293.

Franzel, M., and S. Back, 2019, Three-dimensional seismic sedimentology and stratigraphic architecture of prograding clinoforms, central Taranaki Basin, New Zealand: International Journal of Earth Sciences, **108**, 475–496.

Gersztenkorn, A., and K. J. Marfurt, 1999, Eigenstructure-based coherence computations as an aid to 3-D structural and stratigraphic mapping: Geophysics, **64**, 1468–1479.

Giba, M., J. J. Walsh, and A. Nicol, 2012, Segmentation and growth of an obliquely reactivated normal fault: Journal of Structural Geology, **39**, 253–267.

Goldner, E. L., C. N. Vasconcelos, P. M. Silva, and M. Gattass, 2015, A shortest path algorithm for 2D seismic horizon tracking: Proceedings of the 30th Annual ACM Symposium on Applied Computing - SAC ’15, 80–85.

Gong, S., and D. Fell, 2016, Trestles 3D Merge PSDM Seismic Data Processing Report: Todd Exploration (WesternGeco).

Ha, T. N., K. J. Marfurt, B. C. Wallet, and B. Hutchinson, 2019, Pitfalls and implementation of data conditioning, attribute analysis, and self-organizing maps to 2D data: Application to the Exmouth Plateau, North Carnarvon Basin, Australia: Interpretation, 7, SG23–SG42.

Hale, D., 2013, Methods to compute fault images, extract fault surfaces, and estimate fault throws from 3D seismic images: Geophysics, **78**, O33–O43.

Haque, A. E., A. Islam, and M. R. Shalaby, 2016, Structure modeling of the Maui Gas Field, Taranaki Basin, New Zealand: Petroleum Exploration and Development, 12.

Hussein, M., R. R. Stewart, and J. Wu, 2020a, Which seismic attributes are best for subtle fault detection? Interpretation, T299–T314.

Hussein, M., R. Stewart, and J. Wu, 2020b, Unsupervised Machine Learning Techniques for Subtle Fault Detection: 82nd EAGE Annual Conference & Exhibition Workshop Programme, **2020**, 1–5.

Kass, M., A. Witkin, and D. Terzopoulos, 1988, Snakes: Active contour models: International Journal of Computer Vision, **1**, 321–331.

King, P. R., and G. P. Thrasher, 1996, Cretaceous-Cenozoic Geology and Petroleum Systems of the Taranaki Basin, New Zealand: Institute of Geological & Nuclear Sciences.

Kroeger, K. F., R. H. Funnell, A. Nicol, M. Fohrmann, K. J. Bland, and P. R. King, 2013, 3D crustal-scale heat-flow regimes at a developing active margin (Taranaki Basin, New Zealand): Tectonophysics, **591**, 175–193.

Li, F., J. Qi, B. Lyu, and K. J. Marfurt, 2017, Multispectral coherence: Interpretation, **6**, T61–T69.

Lin, T., and K. J. Marfurt, 2017, What Causes those Annoying Stair Step Artifacts on Coherence Volumes? AAPG EXPLORER.

Lubo-Robles, D., T. Ha, S. Lakshmivarahan, and K. J. Marfurt, 2019, Supervised seismic facies classification using probabilistic neural networks: Which attributes should the interpreter use? 89th Annual International Meeting of the SEG, Expanded Abstracts, 2273–2277.

Lubo-Robles, D., T. Ha, S. Lakshmivarahan, K. J. Marfurt, and M. J. Pranter, 2020, Exhaustive Probabilistic Neural Network for attribute selection and supervised seismic facies classification: Interpretation, 1–67.

Machado, G., A. Alali, B. Hutchinson, O. Olorunsola, and K. J. Marfurt, 2016, Display and enhancement of volumetric fault images: Interpretation, **4**, SB51–SB61.

Marfurt, K. J., 2017, Interpretational Aspects of Multispectral Coherence: 79th EAGE Conference and Exhibition 2017, **2017**, 1–5.

Marfurt, K. J., R. L. Kirlin, S. L. Farmer, and M. S. Bahorich, 1998, 3-D seismic attributes using a semblance-based coherency algorithm: Geophysics, **63**, 1150–1165.

Maurya, S. P., and N. P. Singh, 2018, Application of LP and ML sparse spike inversion with probabilistic neural network to classify reservoir facies distribution - A case study from the Blackfoot field, Canada: Journal of Applied Geophysics, **159**, 511–521.

Mora, J. P., H. Bedle, and Kurt. J. Marfurt, 2020, Constructing fault-surface objects from fault-sensitive attributes: 90th Annual International Meeting of the SEG, Expanded Abstracts, 1160–1164.

Nodder, S. D., 1993, Neotectonics of the offshore Cape Egmont Fault Zone, Taranaki Basin, New Zealand: New Zealand Journal of Geology and Geophysics, **36**, 167–184.

Pedersen, S. I., T. Randen, L. Sonneland, and O. Steen, 2002, Automatic 3D fault interpretation by artificial ants: 64th EAGE Conference & Exhibition, cp-5.

Perico, E., and H. Bedle, 2020, Seismic interpretation of structural features in the Kokako 3D seismic area, Taranaki Basin (New Zealand): 90th Annual International Meeting of the SEG, Expanded Abstracts, 1130–1134.

Qi, J., B. Zhang, B. Lyu, and K. Marfurt, 2019a, Seismic attribute selection for machine-learning-based facies analysis: Geophysics, **85**, O17–O35.

Qi, J., B. Lyu, X. Wu, and K. Marfurt, 2020, Comparing convolutional neural networking and image processing seismic fault detection methods: 90th Annual International Meeting of the SEG, Expanded Abstracts, 1111–1115.

Qi, J., B. Lyu, A. AlAli, G. Machado, Y. Hu, and K. Marfurt, 2019b, Image processing of seismic attributes for automatic fault extraction: Geophysics, **84**, O25–O37.

Randen, T., S. I. Pedersen, and L. Sønneland, 2001, Automatic extraction of fault surfaces from three-dimensional seismic data: 71st Annual International Meeting of the SEG, Expanded Abstracts, 551–554.

Reilly, C., A. Nicol, J. J. Walsh, and K. F. Kroeger, 2016, Temporal changes of fault seal and early charge of the Maui Gas-condensate field, Taranaki Basin, New Zealand: Marine and Petroleum Geology, **70**, 237–250.

Roncaglia, L., M. Milner, M. P. Crundwell, M. Fohrmann, and H. E. G. Morgans, 2013, Well Log Stratigraphy in the Central and Southern Offshore Area of the Taranaki Basin, New Zealand: GNS Science.

Ronneberger, O., P. Fischer, and T. Brox, 2015, U-Net: Convolutional Networks for Biomedical Image Segmentation: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, 234–241.

Specht, D. F., 1990, Probabilistic neural networks: Neural Networks, **3**, 109–118.

Strogen, D. P., K. J. Bland, A. Nicol, and P. R. King, 2014, Paleogeography of the Taranaki Basin region during the latest Eocene–Early Miocene and implications for the 'total drowning' of Zealandia: New Zealand Journal of Geology and Geophysics, **57**, 110–127.

Wu, X., and D. Hale, 2016, Automatically interpreting all faults, unconformities, and horizons from 3D seismic images: Interpretation, **4**, T227–T237.

Wu, X., Y. Shi, S. Fomel, and L. Liang, 2018, Convolutional neural networks for fault interpretation in seismic images: 88th Annual International Meeting of the SEG, Expanded Abstracts, 1946–1950.

Wu, X., L. Liang, Y. Shi, and S. Fomel, 2019, FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation: Geophysics, **84**, IM35–IM45.

Xiong, W., X. Ji, Y. Ma, Y. Wang, N. M. AlBinHassan, M. N. Ali, and Y. Luo, 2018, Seismic fault detection with convolutional neural network: Geophysics, **83**, O97–O103.

Yamashita, R., M. Nishio, R. K. G. Do, and K. Togashi, 2018, Convolutional neural networks: an overview and application in radiology: Insights into Imaging, **9**, 611–629.

Zhao, T., and P. Mukhopadhyay, 2018, A fault-detection workflow using deep learning and image processing: 88th Annual International Meeting of the SEG, Expanded Abstracts, 1966–1970.

Zhao, T., V. Jayaram, A. Roy, and K. J. Marfurt, 2015, A comparison of classification techniques for seismic facies recognition: Interpretation, **3**, SAE29–SAE58.

Chapter 4. Conclusions

These research findings are focused on fault enhancement comparisons, PNN classification results, and object extraction. I extracted large faults using CNN and enhanced coherence and PNN, but the complexity in areas like CEFZ is challenging to image, especially in the PNN case. Plenty of acquisition footprint noise and irrelevant features are found in both volumes, affecting fault enhancement processes and subsequent analysis, a problem that is slightly reduced in part with data preconditioning.

Firstly, PNN can be used to enhance fault features in seismic datasets. I used a PNN workflow from seismic facies analysis (Lubo-Robles et al., 2020) to identify and enhance faults in the Great South Basin and Taranaki Basin, New Zealand. Interpreters can use PNN as an additional tool to enhance faults in seismic interpretation using several attributes to increase the data combination variability and enrich the final classification process. Compared to processes such as coherence enhancement or CNN, PNN will take more time to complete; however, the results could be richer than an individual attribute by itself and have the potential to be used in subtle fault analysis. A drawback in this method might be found in sampling selection because undersampling can be a source of artifacts and misclassification, turning the sampling into a critical step in this approach. Attribute selection also needs to be subjected to meaningful attributes to optimize time and computation resources. As an example, a 3.5 Gb classification analysis with four attributes took 33 hours of processing time using 66 processors; although the method has inherent potential, its use in relatively huge volumes is not optimal. The analysis performed in a 500 Mb volume shows

reasonable times as much as 30 minutes, which is a stark difference from the previous 3.5 Gb. A recommendation is to test this method in subtle fault analysis in a small dataset.

A good interpreter must be aware that each method has different characteristics, and the decision on which one to use is defined accordingly to the time and detail required in its task. The CNN method can provide a fast way to classify a volume, a 3,5 Gb volume classification using a 64 window took around 1.5 hours to achieve the task, but a smaller window could show a similar result in 15 minutes. I noticed that large faults are the ones detected and enhanced, and some small discontinuities are ignored in consequence; this method provides a reasonable way to map large faults in regional analysis. However, CNN results must be verified to avoid anomalies or inexistent features; some artifacts are found in this method, so interpreters must be aware of these problems. CNN training needs to be improved using real and complex seismic models to provide better fault enhancement results.

Coherence enhancement is based on common coherence attributes, the interpreter can modify its results, but sometimes this could increase the analysis time. Although it has more steps than CNN, results can be used with more confidence than the CNN method; but, stair-step artifacts might be an issue in this type of analysis (Lin and Marfurt, 2017). Dip and multispectral coherence calculation increase the time to 3 hours, but fewer artifacts can be noticed. Besides, small discontinuities might be enhanced by this method, making it a better choice than CNN for small fault analysis.

This study was done in a 2.5-second volume in an extensional setting inside the Taranaki Basin; depositional environments such as channels, lithology, and heavy faulting influenced the seismic

quality, causing fault enhancement methods to underperform. PNN method shows low-quality results on heavy faulted areas. Coherence enhancement and CNN behavior are changed in different settings; for example, coherence will perform better where small areas and under plenty of faulting are present because the analysis is less affected by surroundings, but CNN will produce a better continuous image of large faults. To prove this, an additional analysis could be done in real and complex scenarios, including a potential comparison in a compressive regime and using low-quality data to test fault enhancement. Also, since this analysis was done in a real setting, a potential analysis comparison on CNN and coherence might be carried out with synthetic volumes to quantify differences in a modeled environment; it is also possible to analyze listric faults and other complex settings. In the PNN case, a test area with a subtle fault case problem might be employed to find the potential use in small volumes.

The fault enhancement step (Qi et al., 2018) is crucial for creating a skeletonized fault volume. It reduces probabilities around the main faults since multispectral coherence, PNN, and CNN's unprocessed results are discontinuous and noisy.

In this research, object extraction is a tool to compare different fault probability volumes and quantify dissimilarities. In an actual interpretation workflow, users occasionally extract faults with automatic methods. The interpreter needs to understand that any fault enhancement method has uncertainty and should be aware of the methods and understand their limitations because a subsequent object extraction might have artifacts and noise.

The active contour approach follows high probability values that extract faults in a seismic section; this approach changes the line's shape to fit high probability voxels. Seismic fault images

sometimes do not show continuous fault values due to discontinuities or seismic resolution; however, active contours allow the interpreter to map continuous faults even if values do not appear in the processed attribute. This initial work can be used in a manual interactive picking to speed up manual interpretations. A drawback is that active contours follow strictly high probability values, and anomalies such as stair-steps in fault volumes might be mapped automatically. Also, if high probability anomalies are close to fault locations, the algorithm might follow the incorrect fault position. However, an additional smoothing process is possible to achieve in standard interpretation software.

References

Lin, T., and K. J. Marfurt, 2017, What Causes those Annoying Stair Step Artifacts on Coherence Volumes? AAPG EXPLORER.

Lubo-Robles, D., T. Ha, S. Lakshmivarahan, K. J. Marfurt, and M. J. Pranter, 2020, Exhaustive Probabilistic Neural Network for attribute selection and supervised seismic facies classification: Interpretation, 1–67.

Qi, J., B. Lyu, A. AlAli, G. Machado, K. Marfurt, and Y. Hu, 2018, Image processing of seismic attributes for automatic fault extraction: 88th Annual International Meeting of the SEG Expanded Abstracts, 1644–1648.

Appendix A. Probabilistic neural networks data scaling, sampling selection and optimum attribute selection

In this appendix, additional images and tables from the PNN analysis in chapter two includes images from data distribution, tested attributes, and fault enhancement using two different attribute combination.

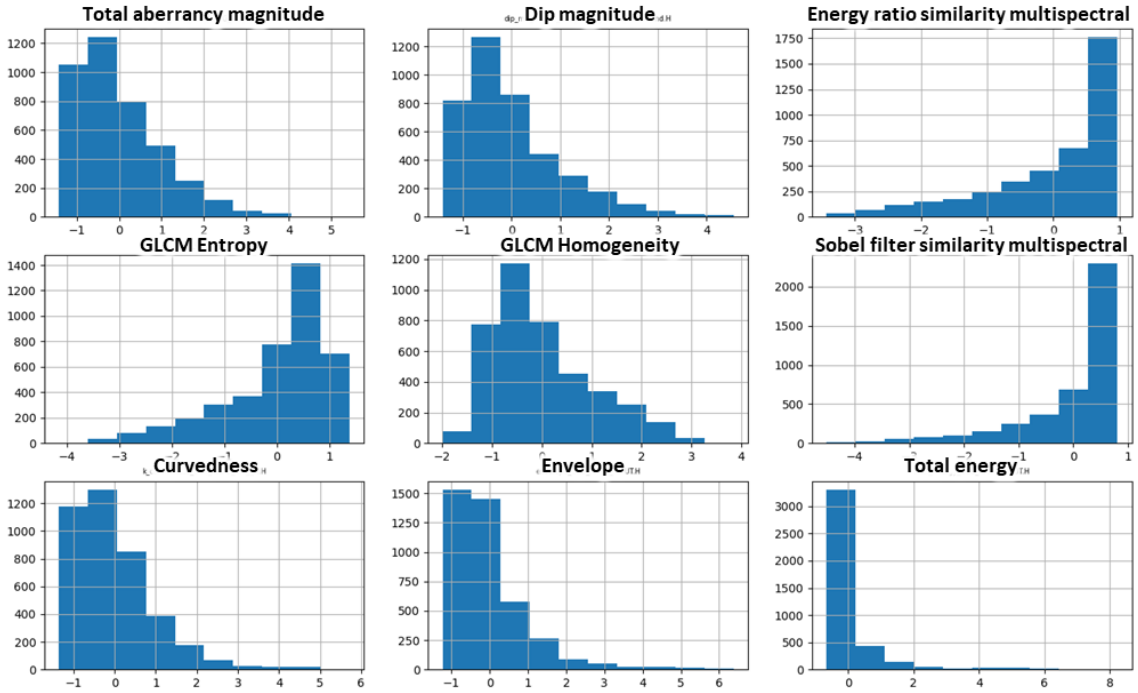


Figure A.1 Histograms of the training dataset after Z-score normalization.

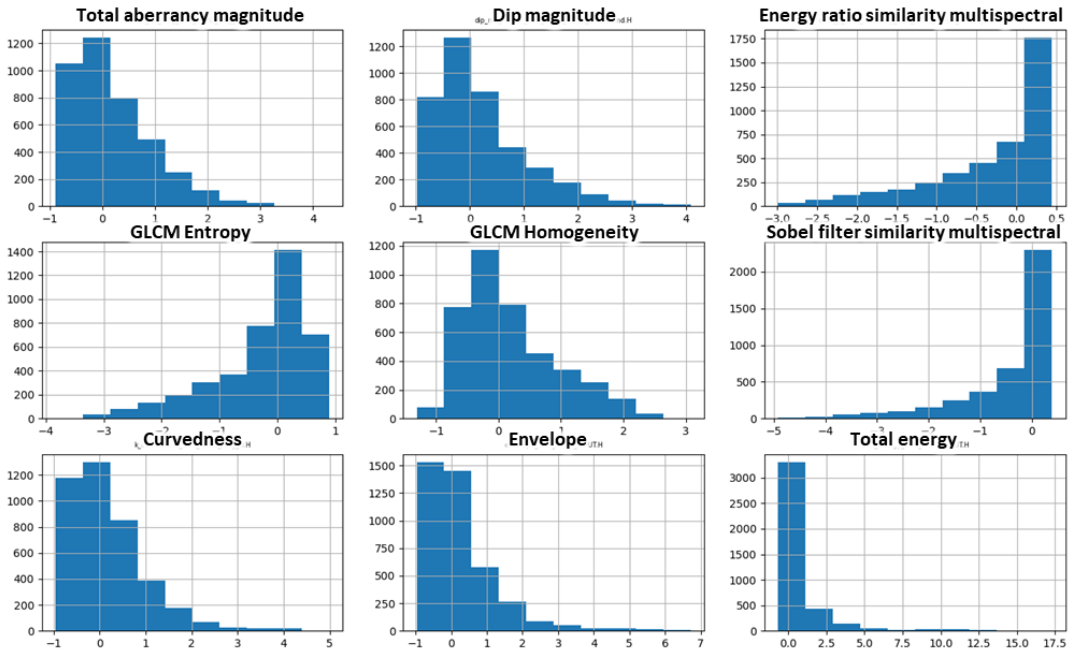


Figure A.2 Histograms of the training dataset after Robust normalization scaling.

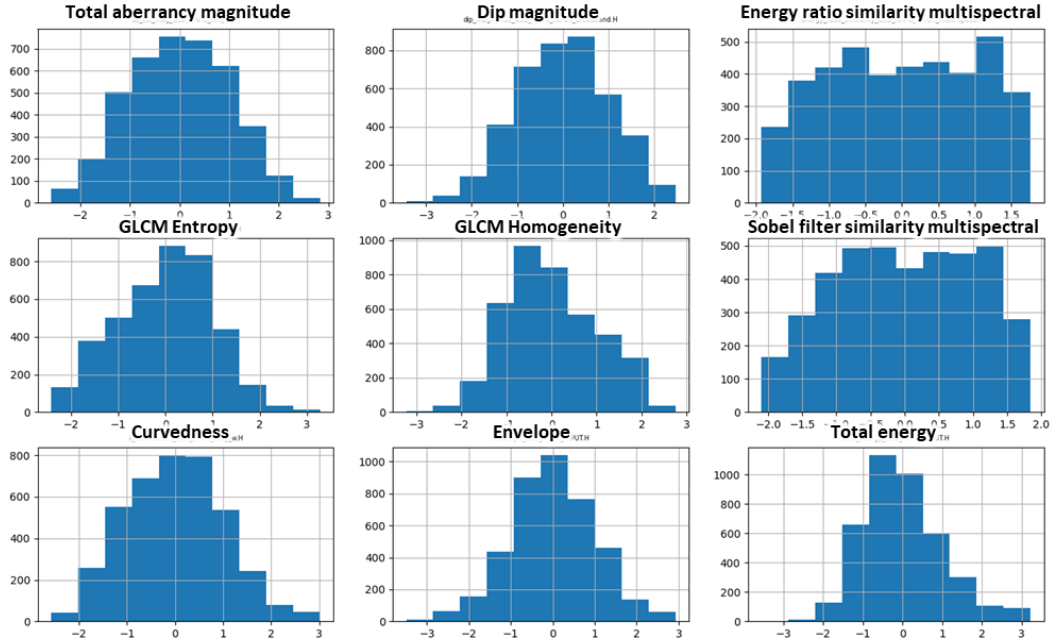


Figure A.3 Histograms of the training dataset after logarithmic normalization scaling.

Table A.1. A list of candidate PNN input attributes and selected attributes used and the corresponding smoothing parameter, r, used in equation 1 obtained for the three normalization strategies evaluated.

Candidate attributes	Selected attributes for different normalizations		
	Z-score	Logarithmic	Robust
Aberrancy magnitude Dip magnitude Energy ratio similarity GLCM entropy GLCM homogeneity Sobel filter similarity Envelope Curvedness Total energy	Aberrancy magnitude GLCMentropy GLCM homogeneity Sobel filter similarity Envelope	Aberrancy magnitude Dip magnitude GLCM homogeneity Sobel filter similarity	Aberrancy magnitude Dip magnitude GLCM homogeneity Envelope
Smoothing parameter *r*	0.4	0.7	0.5

Appendix B. Comparison of fault prediction using coherence enhancement, probabilistic neural networks, and convolution neural networks

In this appendix, I provide additional images and tables that compare PNN fault prediction described in Chapter 3 with those obtained by image directional Laplacian of Gaussian filtering of coherence and convolutional neural networks. In all cases, the original data have been undergone footprint suppression and structure-oriented filtering to remove artifacts hamper fault identification.

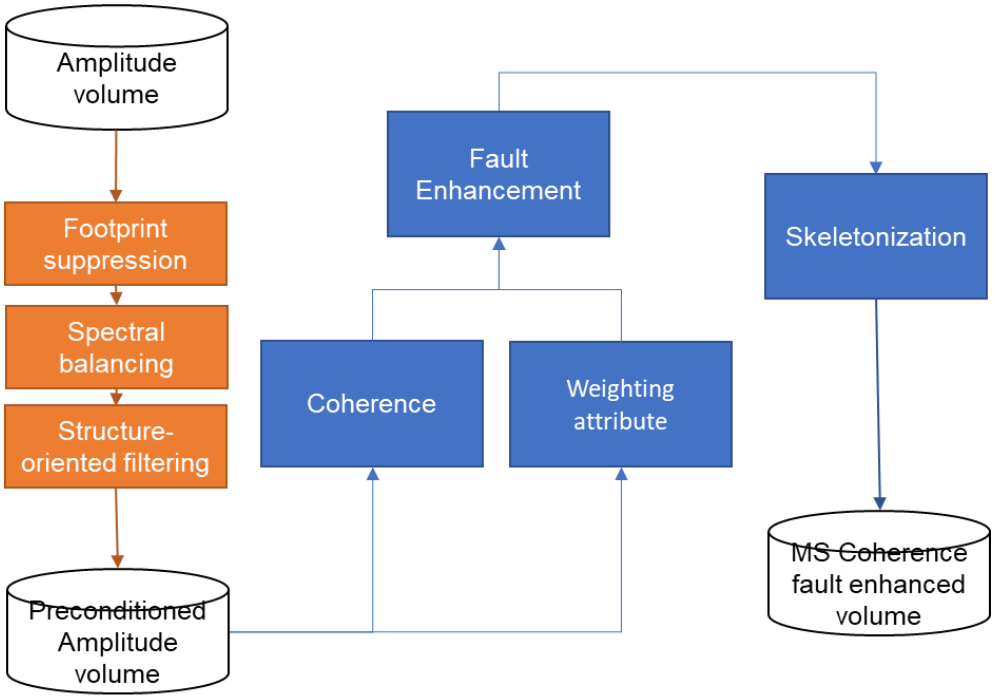


Figure B.1 The multispectral coherence enhancement workflow. First, we apply a spectral balancing algorithm to flatten the spectrum, thereby increasing the bandwidth and vertical resolution. Next, we apply a structural oriented filter to remove unwanted crosscutting coherent as well as random noise. Then we compute multispectral similarity (Li et al., 2018) volumes

using five overlapping filter banks. Then, following Qi et al. (2018), the faults are then enhanced using a directional Laplacian of a Gaussian (LoG) filter prior to a final skeletonization process.

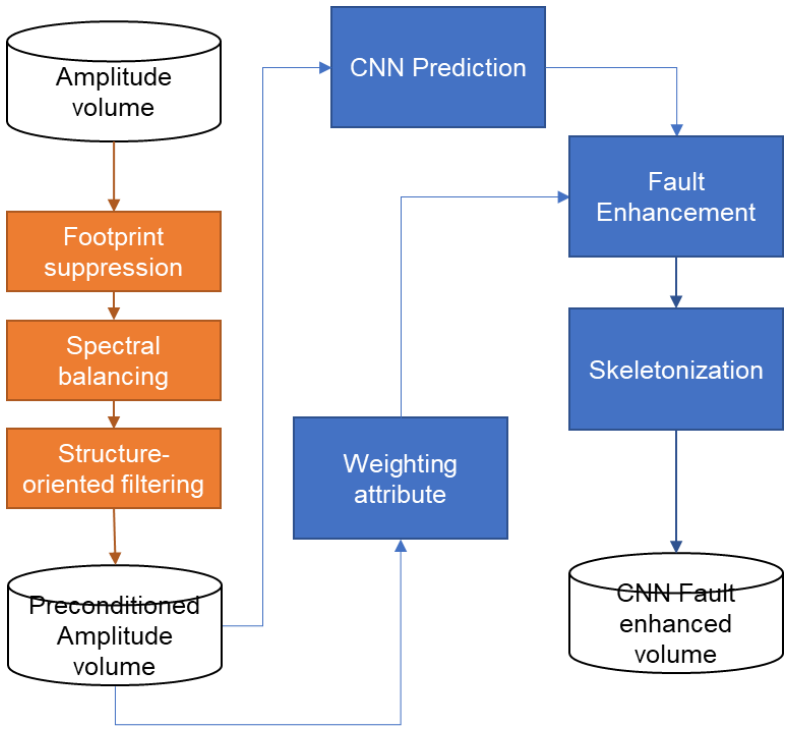


Figure B.2 CNN fault prediction and enhancement workflow. As with the workflow in Figure B3, we spectrally balance the data and then apply a structure-oriented filter. The CNN algorithm employs synthetic data and fault labels described by Wu (2019) to train a convolutional neural network, which we then use to predict faults. The CNN-predicted faults are then further enhanced using the same workflow as in Figure B1.

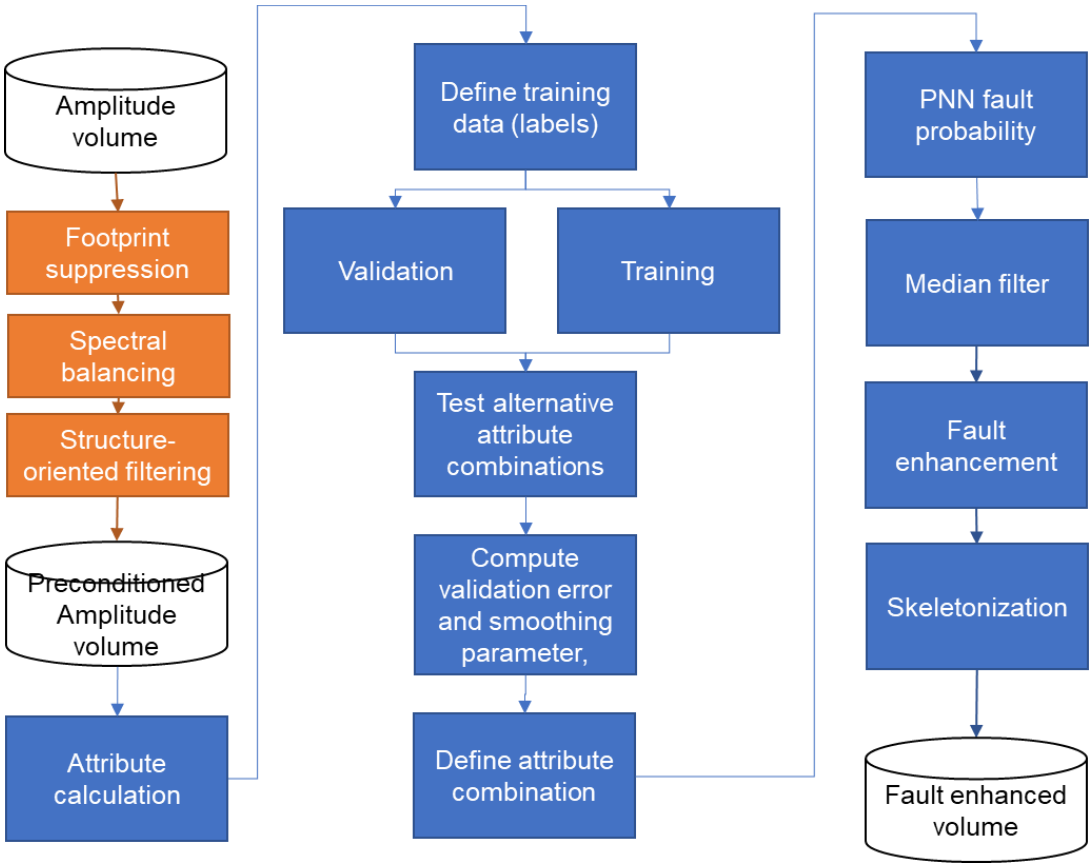


Figure B.3 PNN fault prediction and enhancement workflow. As with the workflow in Figure B3, we apply data conditioning before any fault enhancement process. Attribute calculation follows data preconditioning. Then, fault digitalization to sample relevant fault and no-fault features to split them into training and validation datasets. PNN algorithm analyzes the input data and finds optimal parameters in the tested attributes. Finally, an attribute combination is defined to classify the seismic dataset (Lubo-Robles et al., 2020). The last steps include a medium filter, and then the PNN faults are then further enhanced using the same workflow as in Figure B1.

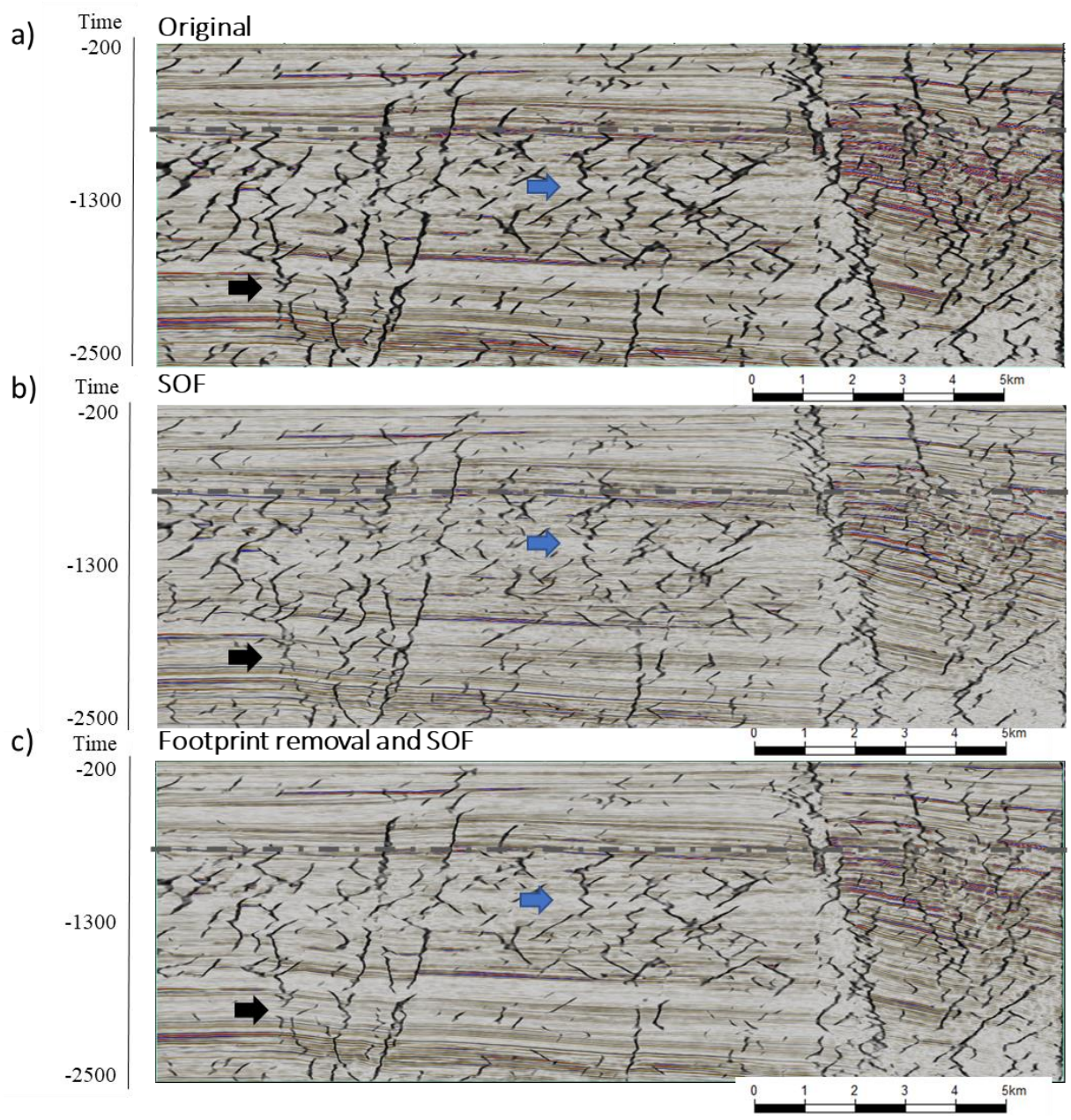


Figure B.4 Vertical slices along crossline 2450 through corendered seismic amplitude and CNN fault prediction computed from (a) the original amplitude data volume (b) the data after structure-oriented filtering, and (c) the data after structure-oriented filtering and footprint

suppression. The blue arrow indicates improvement in a complex area and the black arrow points features' continuity improvement with data conditioning.

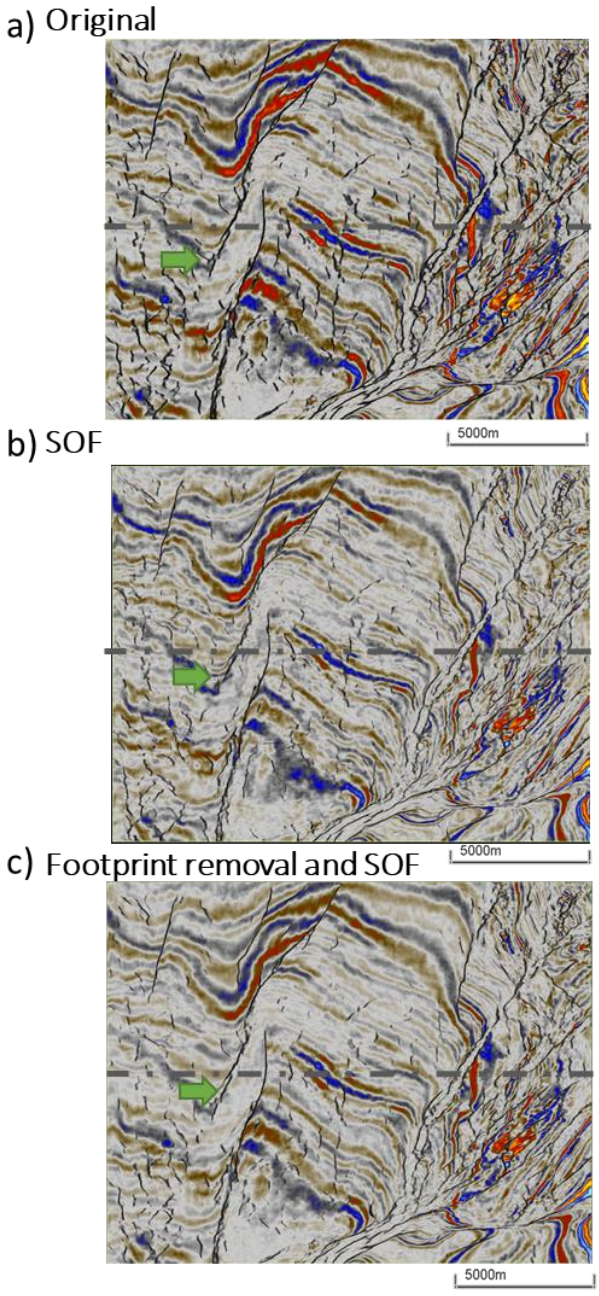


Figure B.5 Time slices at t=0.76 s through co-rendered seismic amplitude and CNN fault prediction computed from (a) the original amplitude data volume (b) the data after structure-

oriented filtering, and (c) the data after structure-oriented filtering and footprint. Green arrows indicate improved areas due to preconditioning.