Naturally fractured reservoir characterization

Advanced workflows for discrete fracture network modeling

Prabhakaran, R.

DOI
10.4233/uuid:d5d43ee7-10bc-4924-beec-1040dba4ac12

Publication date
2021

Document Version
Final published version

Citation (APA)
Prabhakaran, R. (2021). *Naturally fractured reservoir characterization: Advanced workflows for discrete fracture network modeling.* [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:d5d43ee7-10bc-4924-beec-1040dba4ac12

Naturally fractured reservoir characterization

R. Prabhakaran

2021

ISBN 978-94-6384-257-0

NATURALLY FRACTURED RESERVOIR CHARACTERIZATION

Advanced Workflows for Discrete Fracture Network Modeling

RAHUL PRABHAKARAN

Invitation

It is my pleasure to invite you to attend the public defence of my doctoral Thesis:

Naturally Fractured Reservoir Characterization: Advanced Workflows for DFN Modeling

Monday, October 11th 2021 at 12:30 in the Senaatszaal of the Aula of TU Delft, Mekelweg 5, Delft

Prior to the defence, at 12:00, I will give a short presentation on the topic of my Thesis

Rahul Prabhakaran

NATURALLY FRACTURED RESERVOIR CHARACTERIZATION

ADVANCED WORKFLOWS FOR DISCRETE FRACTURE NETWORK MODELING

NATURALLY FRACTURED RESERVOIR CHARACTERIZATION

ADVANCED WORKFLOWS FOR DISCRETE FRACTURE NETWORK MODELING

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus prof.dr.ir. T.H.J.J. van der Hagen
chair of the Board of Doctorates
to be defended publicly on
Monday 11 October 2021 at 12:30 o'clock

by

Rahul PRABHAKARAN

Master of Science in Applied Earth Sciences,
Delft University of Technology, the Netherlands
born in Manama, Bahrain

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. G. Bertotti	Delft University of Technology, promotor
Prof. dr. ir. D.M.J. Smeulders	Eindhoven University of Technology, promotor
Independent members:	
Prof. dr. D.F Bruhn	Delft University of Technology; GFZ Helmholtz-Zentrum, Germany
Dr. A. Barnhoorn	Delft University of Technology
Dr. C.W. Nixon	University of Bergen, Norway
Dr. U. Kelka	Commonwealth Scientific and Industrial Research Organisation, Australia
Prof. dr. W.R. Rossen	Delft University of Technology, reserve member
Other members:	
Prof. dr. J.L. Urai	Rheinisch-Westfälische Technische Hochschule Aachen, Germany

Prof. dr. J.L. Urai contributed significantly to the realization of this dissertation.

Keywords: Naturally fractured reservoirs, automatic fracture detection, graph theory, spatial networks, spatial network heterogeneity

Printed by: Gildeprint

Front & Back: Spatial graphs from the Bristol Channel, UK

ISBN 978-94-6384-257-0

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

The beauty of a living thing is not the atoms that go into it, but the way those atoms are put together.

Carl Sagan

I would rather have questions that can't be answered than answers that can't be questioned.

Richard Feynman

CONTENTS

SUMMARY

Natural fractures in subsurface rocks are a source of heterogeneity that impacts flow and transport behaviour. The presence of fracture discontinuities needs to be modelled explicitly due to observed deviations from the continuum assumption of porous media. The departures are due to both individual properties (such as aperture, infill, and roughness) and global network properties (such as topological summary and length distribution). Understanding flow patterns due to effects of rock fractures networks is essential for many applications such as exploiting hydrocarbons, geothermal heat extraction, subsurface nuclear waste storage, and water aquifer development. Assessing the impact of fractures in modelling studies requires fracture network data which is difficult to sample from seismic data (due to image resolution issues) and borehole data (owing to sparse sampling). Outcrop analogue data provide a means to sample networks while honouring both spatial position and topological relationships.

This Thesis develops advanced methods by which outcrop-derived realistic fracture networks can be rapidly extracted and used to derive insights into fracturing patterns. We first introduce a method to extract fracture networks from images automatically in Chapter 2. The technique based on the complex shearlet transform is combined with an image post-processing and vectorizing workflow to convert image tiles of fractured rocks into digitized traces. The technique is applied to two case studies of drone imagery in fractured limestone pavements - the Parmelan Anticline, France and Brejões, Brazil. In both these cases, the UAV imagery has been shot at image resolutions of 18-20 mm/pixel and yields fracture datasets with millions of fractures. The workflow aids the process of natural fracture characterization from outcrops by replacing the time-consuming manual interpretation.

In Chapter 3, we extend the automatic extraction technique to a different case study: the Lilstock limestone pavements, Bristol Channel, UK. We treat fracture networks as spatial graph data structures consisting of nodes and edges. We design and apply a set of graph routines that utilize geometric and topological criteria to convert automatically traced segments to geologically significant fractures. The derived Bristol Channel dataset comprising more than 350,000 fractures is unique, forming a major fracture geology benchmark. The graph abstraction in the form of primal and dual representations allows the computation of network properties that can be used to compare networks. We find that the major topological type in the large-scale networks is Y-nodes or nodes with degree 3. We also find positive correlation between fracture lengths and number of topological connections indicating the disassortativity of natural fracture networks.

Fracture networks display a natural spatial variation in their organization. Previous methods within the structural geology literature to quantify spatial heterogeneity often ignore topological organization. Spatial variations are generally associated with high-deformation drivers such as folding, faulting, and diapirism. In Chapter 4, we quantify spatial variations in weakly-deformed rocks using examples of fracture networks in the

Lilstock pavements. We introduce a novel graph-based approach by which graph distance measures are used to compare sub-graphs with a gridded, moving-window approach. The pair-wise sub-graph similarities are analyzed using unsupervised hierarchical clustering on the distance matrices. We utilize graph similarity measures such as the fingerprint distance, D-measure, NetLSD, and Portrait Divergence to compute differences between graphs. The results highlight spatial clusters possessing autocorrelation, which is not otherwise detectable through spatial fracture persistence measures. The method is applicable to any 2D fracture trace network and can aid in deriving variation trends within analogues. These insights can then be applied in geostatistical discrete fracture network extrapolation techniques to impart realistic variation in reservoir models.

By their greater propensity to dissolve in acidic fluids, carbonate rocks can develop subsurface karstic networks through hypogenic processes. Similar to rock fractures, the presence of these karstic discontinuities implies that flow and transport do not adhere to a continuum description where the creeping flow regime is alone dominant. Instead in vuggy porous media, there is a free flow regime that needs to be considered. Geological models of karstic cave systems can include specification of 3D cave geometrical features explicitly. However, there is a paucity of realistic cave passage data. Chapter 5 presents a novel workflow based on mobile SLAM Lidar mapping by which one can rapidly acquire 3D point cloud data of cave analogues and generate accurate 3D renditions of intricate cave geometries. The workflow is applied to a few caves from Brazil.

The workflow presented in the Thesis is aimed at maximizing the potential of network discontinuity data that is ubiquitous owing to the recent proliferation of UAV-based photogrammetry. This Thesis is mainly focussed on opening-mode joints, although other discontinuities such as veins, shear fractures, and stylolites also impact subsurface flow and transport. However, the present framework can be expanded to such datasets with minimal degrees of adaptation. The approaches presented can be extended to quantify and compare fracture patterns in different tectonic settings and lithologies. Such a typology of fracture networks can be used to reduce uncertainty in fracture modelling further.

SAMENVATTING

Natuurlijke breuken in ondergrondse gesteenten zijn een bron van heterogeniteit die het stromings- en transportgedrag beïnvloedt. De aanwezigheid van breuk-discontinuïteiten moet expliciet worden gemodelleerd vanwege waargenomen afwijkingen van de continuüm aanname van poreuze media. De afwijkingen zijn te wijten aan zowel individuele eigenschappen (zoals opening, opvulling en ruwheid) als globale netwerkeigenschappen (zoals topologische samenvatting en lengteverdeling). Het begrijpen van stromingspatronen als gevolg van de netwerkeffecten van steenbreuken is essentieel voor veel toepassingen, zoals de exploitatie van koolwaterstoffen, geothermische warmte-extractie, ondergrondse opslag van kernafval en de ontwikkeling van watervoerende lagen. Het beoordelen van de impact van breuken in modelstudies vereist breuknetwerkgegevens die moeilijk te bemonsteren zijn uit seismische gegevens (vanwege problemen met de beeldresolutie) en boorgatgegevens (vanwege schaarse bemonstering). Analoge gegevens uit de ontsluiting bieden een manier om netwerken te bemonsteren, waarbij zowel de ruimtelijke positie als de topologische relaties worden geëerbiedigd.

Dit proefschrift ontwikkelt geavanceerde methoden waarmee uit de ontsluiting afkomstige realistische breuknetwerken snel kunnen worden geëxtraheerd en gebruikt om inzichten te verkrijgen in breukpatronen. We introduceren eerst een methode om breuknetwerken automatisch uit afbeeldingen te extraheren in Chapter 2. De techniek op basis van de complexe shearlet-transformatie wordt gecombineerd met de nabewerking van afbeeldingen en de vectoriserende workflow om beeldtegels van gebroken gesteenten om te zetten in gedigitaliseerde sporen. De techniek wordt toegepast op twee casestudies van dronebeelden in gebroken kalksteenverhardingen: de Parmelan Anticline in Frankrijk en Brejões, Brazilië. In beide gevallen zijn de UAV-beelden opgenomen met beeldresoluties van 18-20 mm/pixel en leveren breukgegevensreeksen op met miljoenen breuken. De workflow ondersteunt het proces van karakterisering van natuurlijke breuken uit ontsluitingen door de tijdrovende handmatige interpretatie te vervangen.

In Chapter 3, breiden we de automatische extractietechniek uit naar een andere dataset - de Lilstock kalksteenverhardingen, Bristol Channel, VK. We behandelen breuknetwerken als ruimtelijke grafische datastructuren die bestaan uit knooppunten en randen. We ontwerpen en passen een reeks grafiekroutines toe die geometrische en topologische criteria gebruiken om automatisch getraceerde segmenten om te zetten in geologisch significante breuken. De Bristol Channel dataset bevat meer dan 350,000 breuken en vormt een unieke dataset die een belangrijke benchmark voor breukgeologie vormt. De abstractie van de grafiek in de vorm van primaire en dubbele representaties maakt het mogelijk netwerkeigenschappen te berekenen die kunnen worden gebruikt om netwerken te vergelijken. We vinden dat het belangrijkste topologische type in de grootschalige netwerken Y-knooppunten zijn of knooppunten met graad 3. We vinden ook een positieve correlatie tussen breuklengtes en het aantal topologische verbindingen, wat wijst op de disassortativiteit van natuurlijke breuknetwerken.

Breuknetwerken vertonen een natuurlijke ruimtelijke variatie in hun organisatie. Eerdere methoden binnen de structurele geologieliteratuur om ruimtelijke heterogeniteit te kwantificeren, negeren vaak de topologische organisatie. Ruimtelijke variaties worden over het algemeen geassocieerd met factoren met een hoge vervorming, zoals vouwen, breuken en diapirisme. We laten zien dat dit niet het geval is in Chapter 4 door ruimtelijke variaties in zwak vervormde rotsen te kwantificeren met behulp van voorbeelden van breuknetwerken in de Lilstock-trottoirs. We introduceren een nieuwe op grafieken gebaseerde benadering waarbij afstandsmetingen worden gebruikt om subgrafieken te vergelijken met een gerasterde benadering met bewegend venster. De paarsgewijze overeenkomsten van de subgrafieken worden geanalyseerd met behulp van ongecontroleerde hiërarchische clustering op de afstandsmatrices. We gebruiken maatstaven voor gelijkenis van grafieken, zoals de vingerafdrukafstand, D-maat, NetLSD en Portrait Divergence om verschillen tussen grafieken te berekenen. De resultaten kunnen ruimtelijke clusters met autocorrelatie benadrukken, die anders niet detecteerbaar is door middel van ruimtelijke breukpersistentiemaatregelen. De methode is van toepassing op elk 2D-breukspoornetwerk en kan helpen bij het afleiden van variatie-trends binnen analogen. Deze inzichten kunnen vervolgens worden toegepast in geostatistische discrete breuknetwerk-extrapolatietechnieken om realistische variatie in reservoirmodellen te geven.

Door hun grotere neiging om op te lossen in zure vloeistoffen, ontwikkelen carbonaatgesteenten ondergrondse karstnetwerken door hypogene processen. Net als bij rotsbreuken, impliceert de aanwezigheid van deze karstische discontinuïteiten dat stroming en transport niet voldoen aan een continuumbeschrijving waarin alleen het sluipende stromingsregime dominant is. In plaats daarvan is er in vuggy poreuze media een regime van vrije stroom waarmee rekening moet worden gehouden. Geologische modellen van karstgrotsystemen kunnen expliciet specificatie van geometrische kenmerken van 3D-grotten bevatten. Er is echter een gebrek aan realistische gegevens over grotpassage. Chapter 5 presenteert een nieuwe workflow op basis van mobiele SLAM Lidar-mapping waarmee men snel 3D-puntenwolkgegevens van grotanalogen kan verkrijgen en nauwkeurige 3D-weergaven van i kan genereren.

De workflow die in het proefschrift wordt gepresenteerd, is gericht op het maximaliseren van het potentieel van 2D-breukgegevens die alomtegenwoordig zijn dankzij de recente proliferatie van op UAV-gebaseerde fotogrammetrie. Dit proefschrift is voornamelijk gericht op verbindingen in de openingsmodus, hoewel andere discontinuïteiten zoals aders, schuifbreuken en stylolieten ook de ondergrondse stroming en het transport beïnvloeden. Het huidige raamwerk kan echter met minimale aanpassingen worden uitgebreid tot dergelijke datasets. De gepresenteerde benaderingen kunnen worden uitgebreid om breukpatronen in verschillende tektonische omgevingen en lithologieën te kwantificeren en te vergelijken. Een dergelijke typologie van breuknetwerken kan worden gebruikt om de onzekerheid in breukmodellering verder te verminderen.

1

INTRODUCTION

The Discrete Fracture Network (DFN) abstraction in subsurface models provides explicit fracture network architecture renditions exhibiting complex spatial network-like patterns. Representing this geometry of fracture networks explicitly in subsurface models is necessary as spatio-temporal physical processes in the subsurface exhibit strong thermal-hydro-mechanical-chemical (THMC) coupling with structural heterogeneities. Naturally fractured rock outcrops are often used as geological analogues to guide subsurface modelling. The recent proliferation of UAV-based photogrammetry has resulted in significantly large volumes of image data that can yield 2D fracture network information. 2D fracture networks digitized from pavements contain spatial information on the organization of fractures. Hence, they are useful to bridge the gap between sparse borehole data and guide stochastic techniques in DFN modelling. There is a need for new tools and methods to utilize this wealth of data into a coherent DFN modeling workflow. This Thesis addresses related questions that arise when geomodellers need to take stock of new data pipelines and integrating them within fracture modelling workflows.

1

1.1. RATIONALE

Natural fractures in rocks form networks of discontinuities that are spatially variable [Laubach et al., 2019]. The spatial character of networks exerts an influence on bulk material continuum properties such as rock strength, stability, and permeability to fluid flow [Bourbiaux, B., 2010, National Research Council, 1996, Nelson, 2001]. Anthropogenic uses of fractured rock reservoirs, both shallow and deep, include hydrocarbon & geothermal energy exploitation [Berkowitz, 2002, Vidal et al., 2017], water aquifer management [Witherspoon, 1986], subsurface storage of gases such as CO_2, H_2, methane, and nuclear waste disposal [Wang and Hudson, 2015]. In all these applications, the natural fracture pattern plays an important role in model predictions when fractures are represented explicitly.

Computational models that deal with various aspects of fractured rock behaviour require geometric inputs of fracture patterns [Berkowitz, 2002, Bourbiaux et al., 2005a,b, Bourbiaux, B., 2010]. Subsurface data is not sufficient to fully characterise fracture networks as the features fall below seismic resolution, and both wellbore data and core data are too sparse [Maerten et al., 2006, Wu and Pollard, 2002]. Even when high-quality seismic is available, seismic attributes can only infer relative trends in fracturing intensity and orientation, with limited pattern information. Well test pressure data can provide insight into flow potential, but pressure signals are often intermixed with other heterogeneities within the subsurface [Aguilera, 1987, Aljuboori et al., 2015]. It is, therefore, difficult to ascertain or quantify how fracture networks spatially vary in space from subsurface data [Laubach et al., 2018a, 2019].

Considering the above-mentioned limitations associated with subsurface data, outcropping analogues can be used as a resource to understand multi-scale discontinuity patterns and relate them to deformation via geological rules [Awdal et al., 2016, Lamarche et al., 2012, Laubach et al., 2009, Shackleton et al., 2005]. Outcrops that are relatively free from exhumation and weathering effects are useful to study variations in fracturing. Spatial variation in fracture network geometry (or non-stationarity) is generally attributed to high-deformation drivers such as folding and faulting. However, natural variation is also qualitatively observed in weakly-deformed rocks and this can be attributed to variations in lithology and spatially-varying mechanical unit thicknesses.

Spatial variation in outcropping fractures is often analyzed using scanlines [Mauldon and Mauldon, 1997] and many statistical procedures have been developed for analysing scanline data. However, scanlines have similar deficiencies as well data, and there is no cognisance of spatial organization inherent to fracture networks. The spatial character of fracture networks is better studied from 2D trace maps derived from fractured outcropping analogues [Bisdom et al., 2017, Lamarche et al., 2012]. Digital outcrop models derived from UAV photogrammetry quickly generates large volumes of image data that can yield rich, quantitative information on 2D fracture network patterns [Bemis et al., 2014, Bisdom et al., 2017, Hodgetts, 2013]. However, a bottleneck in obtaining vectorized geometries from images is that 2D trace maps are usually interpreted by hand. This is a time-consuming process that is prone to interpreter bias and is subjective in terms of results. Such resulting hand-traced fracture maps are often too small to identify the non-stationary spatial character of fracture networks. Therefore, there is a dual need for methods that can efficiently extract 2D fracture trace data from imagery and robust

quantitative metrics that can quantify spatial variability in realized fracture trace maps.

1.2. SCOPE OF THIS THESIS

In this Thesis, we develop a workflow to automatically extract multi-scale fracture network datasets from fractured rock images derived from UAV-derived, Structure-from-Motion (SfM) photogrammetry. The workflow consisting of a series of image processing steps and is applied to three datasets from different carbonate settings; Parmelan (France), Brejões (Brazil), and Lilstock (UK) forming the largest fractured rock datasets to date with over a million fracture traces, enabling detailed quantification and study of spatial network variation. The Parmelan plateau in France is an example of fracturing in a pop-up fold structure. The Brejões and Lilstock pavements represent weakly deformed carbonate rocks with background fracturing.

The fracture networks are converted into graphs with node positioning, enabling the use of spatial graph theory-based methods. The automatic extraction results, in the form fracture segments, are converted to geologically significant fractures using graph-based routines. The size of these realized fracture graphs using the spatial graph abstraction enables a systematic analysis of intra-network and inter-network spatial variability. Using unsupervised clustering algorithms, we are able to discover patterns of variability within fracture networks. We utilise graph distance metrics based on various graph properties such as block geometry probability distributions and graph spectra to analyse spatial variation in fracture patterns. Recognizing the challenge of geomodelling in karst reservoirs, where in addition to flow through matrix and fractures a free-flow regime exists in vuggy pathways, we present an innovative data acquisition technique based on SLAM-lidar and point-cloud processing to rapidly capture the 3D geomorphology of caves.

1.3. THESIS OUTLINE

This Thesis starts with Chapter 2 devoted to the methodology and application of an automated fracture tracing workflow to convert UAV photogrammetry of fractured rock into DFNs. The method is validated using two carbonate outcrops from Parmelan, France, and Brejões, Brazil, as an example and benchmarked against a volcanic dyke dataset from the literature.

Chapter 3 deals with extending the automated tracing method in Chapter 2 to the famous Bristol Channel fractured limestones. This outcrop is unique, with large exposures of fractured rock with very little noise. This chapter presents a method to embed fracture networks within a spatial graph-theoretic framework and uses graph-based measures to characterize natural fractures.

Chapter 4 focusses on quantifying spatial variations in network geometry. Using fractures automatically extracted in Chapter 3 from the Lilstock pavements represented as spatial graphs, we use unsupervised hierarchical clustering combined with graph distance metrics to identify spatial clusters.

In Chapter 5, a novel method for the fast acquisition of cave patterns is introduced. The technique combines portable lidar in combination with point-cloud processing to rapidly generate high-resolution 3D cave system models.

2

AUTOMATIC FRACTURE TRACE DETECTION

Representing fractures explicitly using a discrete fracture network (DFN) approach is often necessary to model the complex physics that govern thermo-hydro-mechanical-chemical processes (THMC) in porous media. DFNs find applications in modelling geothermal heat recovery, hydrocarbon exploitation, and groundwater flow. It is advantageous to construct DFNs from photogrammetry of fractured outcrop analogues as the DFNs would capture realistic, fracture network properties. Recent advances in drone photogrammetry have greatly simplified the process of acquiring outcrop images, and there is a remarkable increase in the volume of image data that can be routinely generated. However, manually digitizing fracture traces is time-consuming and inevitably subject to interpreter bias. Additionally, variations in interpretation style can result in different fracture network geometries, which may then influence modelling results depending on the use-case of the fracture study. In this chapter, an automated fracture trace detection technique is introduced. The method consists of ridge detection using the complex shearlet transform coupled with post-processing algorithms that threshold, skeletonize, and vectorize fracture traces. The technique is applied to the task of automatic trace extraction at varying scales of rock discontinuities, ranging from 10^0-10^2 *m. We present automatic trace extraction results from three different fractured outcrop settings. The results indicate that the automated approach enables extraction of fracture patterns at a volume beyond what is manually feasible. Comparative analysis of automatically extracted results with manual interpretations demonstrates that the method can eliminate the subjectivity that is typically associated with manual interpretation. The proposed method augments the process of characterizing rock fractures from outcrops.*

This chapter is based on Prabhakaran et al. [2019a], An automated fracture trace detection technique using the complex shearlet transform, *Solid Earth*, 10(6), 2137-2166, https://doi.org/10.5194/se-10-2137-2019.

2.1. INTRODUCTION

Naturally fractured reservoir (NFR) modelling requires an explicit definition of fracture network geometry to accurately capture the effects of fractures on the overall reservoir behaviour. The National Research Council [1996] suggested the idea of using geologically realistic outcrop fracture patterns to guide subsurface fracture modelling. In recent work, the use of deterministic discrete fracture networks (DFNs) based on trace digitization from photogrammetry of outcrop analogues was investigated by Bisdom et al. [2017] and Aljuboori et al. [2015] for reservoir fluid flow simulation and well testing. Outcrop derived DFNs encapsulate 2D fracture network properties at a scale that cannot be characterized using either standard surface approaches (scanlines and satellite imagery) or subsurface techniques (seismic imaging/borehole imagery/core sampling). Ukar et al. [2019] suggested a comprehensive set of protocols to select fractured outcrops that are representative of the subsurface. Stochastic and geomechanical DFNs are alternatives to outcrop derived DFNs for fractured reservoir modeling. Stochastically generated DFNs have the disadvantage that they cannot replicate the spatial organization of fracture network patterns observed in nature [Thovert et al., 2017]. Geomechanically derived DFNs are based on the physics of fracture propagation [for e.g. Olson et al., 2009, Thomas et al., 2018] and can reproduce realistic fracture patterns provided the complex paleostress field and paleorock properties are known; however, they are also computationally intensive and hence have limited applicability. A carefully chosen fractured outcrop that is relatively free of noise (fractures resulting from exhumation and weathering and not too much hidden by vegetation) may be used to interpret realistic fracture networks which are geometrical inputs used in simulating various subsurface thermo-hydro-mechanical-chemical (THMC) processes.

Recent advances in Unmanned Aerial Vehicles (UAVs) and stereo-photogrammetry has dramatically simplified the acquisition of georeferenced datasets of fractured outcrop images [for e.g. Bemis et al., 2014, Harwin and Lucieer, 2012, Turner et al., 2012]. Photogrammetry using the Structure from Motion (SfM) principle is a relatively inexpensive and rapid technique by which 3D outcrop models are built by identifying, extracting, and positioning common points in georeferenced outcrop images [Donovan and Lebaron, 2009]. Images are captured using a camera-equipped UAV that is capable of following pre-programmed flight missions where flight path, altitude, velocity, and overlap are specified. The images undergo further processing steps that include generating sparse point clouds of common points, aligning the images, generating dense point clouds (3D representation of outcrop geometry), and generating meshed surfaces [Bisdom et al., 2017]. Interpreting fractures on the image orthomosaics with conventional Geographic Information System (GIS) software completes the outcrop-based DFN workflow.

Manually interpreting fractures is time-consuming and forms a bottleneck in outcrop-based DFN workflows. A manual interpretation has a fair degree of associated subjectivity, and interpreter bias may take the form of specific scales of features being inadvertently omitted or deliberately ignored [Bond et al., 2007, Scheiber et al., 2015]. Manual interpretation also suffers from a lack of repeatability owing to differing expertise levels of interpreters, and the interpretation criteria followed [Hillier et al., 2015, Sander et al., 1997]. Reproducibility may not be guaranteed even with the same interpreter in

multiple trials [Mabee et al., 1994]. According to Bond et al. [2015], quantifying the magnitude and impact of subjective uncertainty is difficult. Long et al. [2018] conducted a study on variability of fracture interpretation in which geologists with varying levels of expertise interpreted a single image. They found considerable variation in fracture topology, orientation, intensity, and length distributions in the interpretations. Andrews et al. [2019] made a detailed quantification of subjective bias in scanline-based fracture data collection, the associated effects on derived fracture statistics and suggested protocols for managing the variations. Peacock et al. [2019] delved into the multiple reasons for bias and the resulting implications for modelling. Given the amount of data generated in short UAV flight missions, man-hours spent in interpretation, and the need to de-bias interpretation as much as possible, automatic feature detection techniques may be considered. Automated approaches can speed up the process, improve accuracy, and exploit the acquired data to the fullest possible extent.

In this chapter, we apply an automated method to extract digitized fracture traces from images of fractured rocks. The method utilizes the complex shearlet transform measure to extract fracture ridge realizations from images. Post-processing image analysis algorithms are coupled with the ridge extraction process to vectorize fracture traces in an automated manner. The complex shearlet transform was introduced by Reisenhofer [2014] and King et al. [2015] and previously applied to problems such as detecting coastlines from Synthetic Aperture Radar (SAR) images [King et al., 2015] and propagating flame fronts from planar laser-induced fluorescence (PLIF) images [Reisenhofer et al., 2016]. We present automatic fracture extraction results from drone images of two carbonate outcrops (Parmelan, France & Brejões, Brazil) and station scale images of igneous dyke swarms.

2.2. BACKGROUND

2.2.1. REVIEW OF AUTOMATED AND SEMI-AUTOMATED FRACTURE DETECTION APPROACHES

Rapid digitization of geological features from photogrammetry is challenging owing to issues like spatially varying image resolution, inadequate exposure, the presence of shadows due to effects of topography on illumination conditions, and chromatic variations of essential features. False positives are non-geological features (such as trees, shrubbery, and human-made structures) that are detected using semi-automated/automated approaches [Vasuki et al., 2014]. Removal of false positives is time-consuming. On the other hand, essential features that are not detected at all (referred to as false negatives) by an algorithm, further complicates the task of automated feature extraction. Automated methods, in general, detect more features than what is present in the image [Abdullah et al., 2013]. In this section, we review some approaches for automatic fracture detection based on the class of algorithm used.

Automated fracture detection utilizing higher dimensional data such as point clouds, digital elevation models (DEMs) and digital terrain models (DTMs) have an advantage in that depth variations are captured and can be used to extract features. Thiele et al. [2017] presented an approach based on a least cost function algorithm applicable to orthophotographs of jointed fracture sets and 3D point cloud data. Masoud and Koike [2017]

introduced a software package to detect lineaments from composite grids derived from gravity, magnetic, DEMs, and satellite imagery. Bonetto et al. [2015] and Bonetto et al. [2017] presented semi-automatic approaches that extract lineaments from DTMs utilizing the curvature of geological features. Hashim et al. [2013] presented an edge detection and line linking method using Enhanced Thematic Mapping (ETM).

Colorimetry of an image can be used to detect features. By partitioning features in the image, e.g., matrix rock as lighter shades of gray and fractures as darker shades of gray, fracture pixels may be extracted separately from matrix rock using pixel values. Vasuki et al. [2017] developed an interactive colour-based image segmentation tool using superpixels [Ren and Malik, 2003] which are groupings of pixels that are perceptually similar.

Edge detection techniques identify points in images where sharp changes in image intensity occur. Some of commonly used edge detection techniques in image processing are Canny, Sobel, Prewitt, Robert, Kuwahara, and Laplacian of Gaussian filters. Alternatively, edges may be detected using methods that are invariant to contrast and illumination in images. Phase symmetry and phase congruency algorithms [Kovesi, 1999, 2000] fall under this category. Phase symmetry is an edge detection technique that is invariant to local signal strength. The method identifies axis of a feature, isolating pixels symmetric along profiles that are sampled from all orientations except parallel to the feature. The axes of symmetry are regions where frequency components either approach a maximum or minimum. The phase congruency method is another edge detection method that detects features by identifying points where Fourier components are maximally in phase. This approach is also invariant to the magnitude of the signal. The property of invariance enables the identification of structures within the image even in the presence of noise. Vasuki et al. [2014] utilized an edge detection algorithm using the phase congruency principle coupled with a multi-stage linking algorithm for detection of fault maps.

The Hough transform [Duda and Hart, 1972] is another technique that has been used to detect lineaments in images. The Hough transform identifies pixels in binary images that are likely to represent rock fractures using a voting procedure. Each pixel in a binary image is represented as a sinusoidal curve in a 2D parametric space (or a Hough space). The voting procedure accumulates a vote for each curve in the parametric space corresponding to each non-zero pixel in the binary image. The curves with the highest votes are selected as probable fractures since they correspond to the largest number of non-zero pixels. Results by Callatay [2016] using the Hough transform for fracture detection report the following limitations. Firstly, the detection is limited to a given fracture orientation set owing to the definition of the Hough transform parameter space. Secondly, the issues of false positive detection and discontinuities persisted. The method is also limited by the fact that it needs a binarized image to start.

The development of wavelet theory in the field of harmonic analysis have led to applications in edge detection [Daubechies, 1992, Heil et al., 2006]. Mallat and Hwang [1992] proposed wavelet-based approaches for edge detection. Wavelet-based methods differ from gradient-based edge detection methods that search for local maxima of the absolute value of the gradient. Felsberg and Sommer [2001] introduced monogenic wavelets for the purpose. Tu et al. [2005] considered the use of magnitude response of

complex wavelet transforms. Wavelets, owing to their isotropic properties, cannot extract curve-like features due to the lack of directional information [Labate et al., 2005]. A number of wavelet-based approaches that have been proposed to overcome this lack of directional information such as curvelets [Candès and Donoho, 2005], ridgelets [Candès and Guo, 2002], contourlets [Do and Vetterli, 2005], bandlets [Le Pennec and Mallat, 2005] , wedgelets [Donoho, 1999], shearlets [Guo et al., 2005], and band-limited shearlets [Yi et al., 2009].

2.2.2. THE COMPLEX SHEARLET TRANSFORM

In images of fractured outcrops, the presence of discontinuous gaps due to rupture within the rock mass, which occur naturally and which maybe enlarged through weathering processes, are commonly used as defining criteria by interpreters to digitally trace and classify as fractures within the rock mass. Fractures may also be partially or completely sealed by the presence of infilling material that may be mineralogically different from the adjacent rock material. In such a case, the contrast in colour and texture of the infill material provides an interpretative criterion for classification of these material regions as fractures. The presence of such prominent discontinuities within otherwise smooth regions of rock images, can be exploited by the complex shearlet transform to precisely identify position in the form of edges and ridges.

The basis of the complex shearlet transform applied to fracture extraction from images emanates from wavelet theory. Wavelets are rapidly decaying wavelike oscillations possessing a finite duration. Wavelet transforms are routinely used in digital signal processing applications which are often time-domain signals. They can also be applied to image data which can be considered as 2D functions. Wavelet transforms are not able to detect directionality of structural features in image data since they may only be dilated or translated. Shearlets that were introduced by Labate et al. [2005] as a new class of multidimensional representation systems, overcame a major shortcoming of wavelets by enabling dilation, shear transformation, and translation operations. The isotropic dilation of wavelets was replaced with anisotropic dilation and shearing in the case of shearlets. These modifications have resulted in shearlets possessing a number of properties that make them better suited to handle sparse, geometric features in 2D image data compared to traditional wavelets [Kutyniok and Labate, 2012].

The complex shearlet transform is a complex-valued generalization of the shearlet transform that was developed by Labate et al. [2005] to handle geometric structures in 2D data. Reisenhofer [2014] and King et al. [2015] proposed the idea of creating complex shearlets by modifying the shearlet construction so that real parts of the generating function are even-symmetric and imaginary parts of the generating function are odd-symmetric. They used the Hilbert transform to convert even-symmetric functions into odd-symmetric and vice versa. The complex shearlet measure for ridge and edge detection implemented in King et al. [2015], Reisenhofer [2014], and Reisenhofer et al. [2016] merged the ideas of phase congruency [Kovesi, 1999] and complex shearlets.

The complex shearlet transform is a complex-valued generalization of the shearlet transform that was developed by Labate et al. [2005] to handle geometric structures in 2D data. Reisenhofer [2014] and King et al. [2015] proposed the idea of creating complex shearlets by modifying the shearlet construction so that real parts of the generat-

ing function are even-symmetric and imaginary parts of the generating function is odd-symmetric. They used the Hilbert transform to convert an even-symmetric function into an odd-symmetric function and vice versa. The complex shearlet measure for ridge and edge detection implemented in King et al. [2015], Reisenhofer [2014], and Reisenhofer et al. [2016] merged the ideas of phase congruency and complex shearlets.

The complex shearlet measure first introduced by Reisenhofer [2014] and improved by King et al. [2015] was used for applications like coastline detection [King et al., 2015], flame front detection [Reisenhofer et al., 2016], and feature extraction from terrestrial LIDAR inside tunnels [Bolkas et al., 2018]. Karbalaali et al. [2018] used the complex shearlet transform for channel edge detection from synthetic and real seismic slices. Reisenhofer et al. [2016] presented a comprehensive comparison of CoShREM with Canny, Sobel, phase congruency, and another shearlet-based edge detector [Yi et al., 2009]. Bolkas et al. [2018] also makes specific comparisons between the performance of Canny, Sobel, Prewitt edge detection methods versus space-frequency transform methods such as wavelets, contourlets, and shearlets. A detailed overview of the complex shearlet transform is provided in Appendix A.

2.3. METHODS

2.3.1. THE AUTOMATIC DETECTION PROCESS

The automated fracture trace detection method that we present has five main steps (see Fig. 2.1). The 1st step uses the Complex Shearlet-Based Ridge and Edge Measure (CoShREM), a MATLAB implementation by Reisenhofer et al. [2016] that utilizes functions from Shearlab3D developed by Kutyniok et al. [2016] and Yet Another Wavelet Toolbox [Jacques et al., 2011]. The 1st step, namely the ridge detection, is dependent on a number of input parameters tabulated in Tables 2.1-2.2. Equation (A.27) gives the expression for the ridge measure.

An optimal set of deterministic parameter values which can extract features on all scales is not known *a priori*. Therefore, we vary the input parameters corresponding to shearlet system construction and ridge detection within user-defined ranges to compute multiple ridge realizations. A ridge ensemble map is obtained by superposing the ridge images and normalizing. A simple sigmoid function applied on the normalized ridge ensemble, non-linearly scales and thereby isolates higher image intensities. A user-defined threshold is then applied to intensity values of this non-linearly scaled, normalized ridge ensemble image to extract a highly probable, binarized, ridge network. The threshold is set by a visual comparison of the input image with the extracted ridges. The range for each parameter in Tables 2.1-2.2 is ascertained by first testing the effect of variation of each parameter with respect to a chosen base case image. This approach to automated detection captures features of multiple scales and highlights regions of uncertain feature extraction within the image.

The 2nd step is segmentation of ridges using Otsu thresholding [Otsu, 1979]. This operation removes small, disconnected, and isolated ridge pixel clusters. The 3rd step is skeletonization where clusters of pixels representing segmented ridges are thinned into single pixel representations. For intersecting fractures, the skeletonization procedure preserves topology by recognizing and splitting the frame at branch points. This ensures

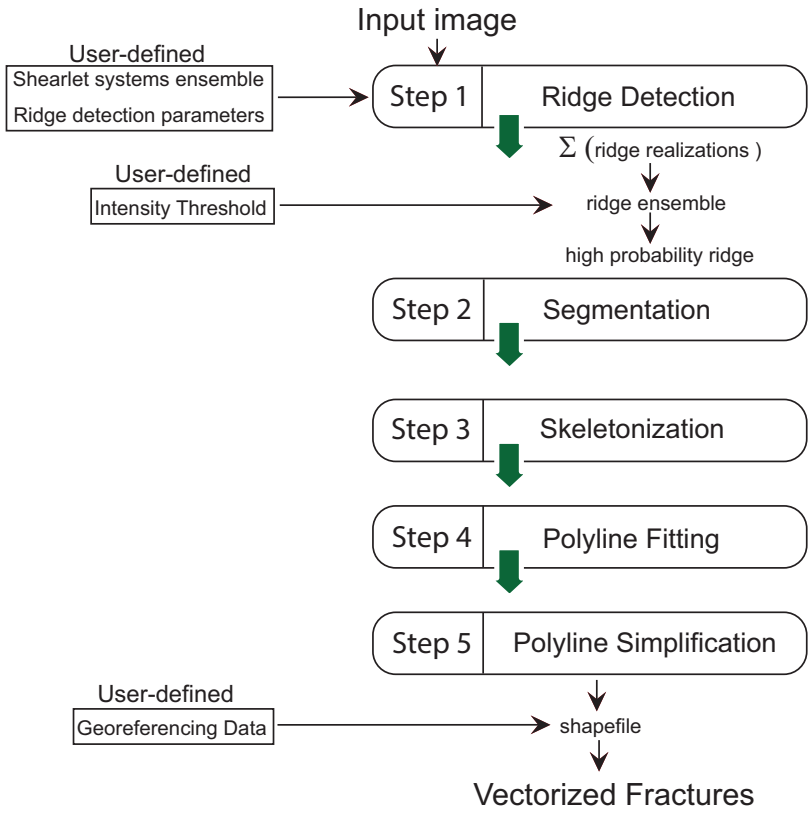

Figure 2.1: The Automated Fracture Trace Detection Workflow

Table 2.1: Shearlet system and detection parameters used to extract ridges for the base case

Base Case Parameters	
waveletEffSupp	125
gaussianEffSup	63
scalesPerOctave	2
shearLevel	3
alpha	0.5
octaves	3.5
minContrast	10
detection	negative ridges
even/odd offset	1

Table 2.2: Detection Parameters

Parameter	**Description**
DetectionType	Specification of detection measure (edge/ridge)
minContrast	Specification of the level of contrast for edge/ridge detection
offset	Scaling offset between the even- and odd- symmetric shearlets

that in subsequent DFN representation, no further effort is expended to manually connect detected segments.

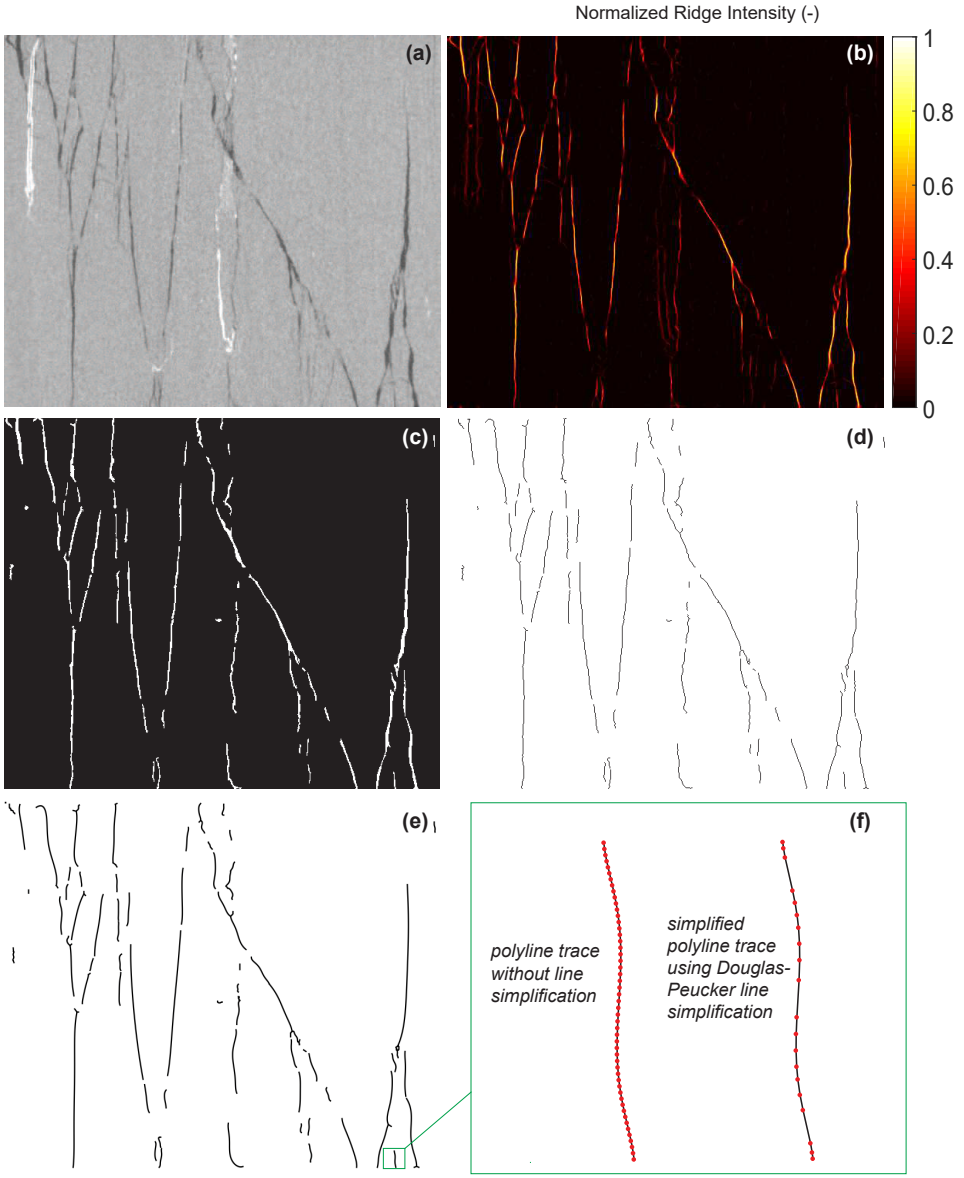

Figure 2.2: Illustration of the steps involved in the automatic fracture extraction using a 40 x 34 mm fractured shale core image (a) CT scan core image from Dwarkasing [2016] (b) Normalized ridge ensemble (c) Segmentation applied on the ridge ensemble (d) Skeletonization applied to the segmented ridge (e) Vectorized polylines fitted to the skeletonized clusters (f) Effect of line simplification applied to a single vectorized segment

The 4th step involves piecewise linear polyline fitting to the skeletonized clusters.

By default, we fit polylines rather than lines to the pixel clusters. Polyline fitting retains geologically realistic veering and curvature of fractures in the vectorized result. The 5th step is a line simplification procedure applied to the piecewise linear polyline clusters. A large number of polyline points would increase the size of vectorized files; hence, we use the Douglas-Peucker line simplification algorithm [Douglas and Peucker, 1973] that simplifies piecewise linear polylines. The number of polyline points assigned to each skeletonized cluster is set constant, but this may be modified to be a linear function of the cluster size measured in pixels. If images are georeferenced or image scale is known, the simplified polylines can be written to a vectorized shapefile format. The DFN in the vectorized shapefile format may now be used for any application that requires explicit fracture network geometry. An example of a fractured Posidonia shale micro CT image slice from Dwarkasing [2016] (see Fig. 2.2) illustrates the effects of each of the steps involved.

2.3.2. SENSITIVITY ANALYSIS OF PARAMETERS ON EXTRACTION RESULTS

Since the detection results may vary owing to different parameter combinations, we conducted a sensitivity analysis to investigate the ridge extraction output with variation in parameter input. An example of a fractured image sample representing Mesoproterozoic sandstone from the Tomkinson Province, Northern Territory, Australia (Fig. 2.3.a) is chosen to study the effect of shearlet parameter variation. The image dimensions are 1313 x 1311 pixels and has four prominent fractures with two of them forming an intersection. A subtler fracture is present towards the top-left and a thick fracture located at the bottom-left of the image. A base case set of parameters for constructing a shearlet system and for ridge identification is set up in the table adjoining Fig. 2.3(a). We vary all parameters one by one with respect to this base case. Ridge extraction using the base case shearlet system shows that the major intersecting fracture system is identified; however, the largest fracture is detected only partially and that too, only at the peripheries. (see Fig. 2.3.b). The subtle fracture is detected but disconnected. A large amount of noise is also present.

The complex shearlet system is constructed by the tensorial product of a Mexican hat wavelet and a gaussian filter. The first two parameters *waveletEffSupp* and *gaussianEffSup* refer to the pixel widths over which the wavelet amplitudes sharply change from zero. The even- and odd-symmetric elements of constructed shearlet system using the base case parameters for the siliclastic example are depicted in Fig. 2.4(i)-(ii). We chose to maintain a ratio of two between *waveletEffSupp* and *gaussianEffSup*. The effect of increasing the effective support on the complex shearlet system is shown in Fig. 2.4(xvii)-(xix). Fig. 2.4(xx)-(xxi) indicate the effects of large ratios between the wavelet effective support and gaussian support. The second parameter is the *scalesperOctave* which determines the number of intermediate scales per octave. An octave is the interval between two frequency peaks. For example, we may consider a wavelet that is scaled by a factor of 2. Physically, this means a stretching of the wavelet thereby decreasing the frequency. The base-2 logarithmic ratio of the reduced frequency with respect to the original frequency, is the number of octaves by which the frequency has reduced. We set the number of octaves as a constant value of 3.5. This implies that there are 7 scales for the complex shearlet system as can be seen in Fig. 2.4(iii)-(ix). The *shearLevel* param-

eter indicate the discrete number of orientations that the complex shearlet system can assume. The selected value of 3 indicates that there are 2^3+2 (ten) orientations possible for the complex shearlet system (see Fig. 2.4.x-xvi) and $2(2^3+2)$ (or 20 shearlets). For large images and large number of shearlets, computational effort is quite expensive. The *alpha* parameter is the degree of anisotropy induced by scaling with a null value of *alpha* maximizing the degree of anisotropy. We vary *alpha, shearLevel*, and the *scalesperOctave* but the effects on the constructed complex shearlet system are minimal as can be seen from Fig. 2.4(xxii)-(xxx).

The effects of variation of the parameters on ridge extraction is depicted in Fig. 2.3(c)-(p). Decreasing the value of the support by half identifies finer features, but then the largest fracture is completely missed (Fig. 2.3.e). When the support is doubled, the emphasis on larger features is more pronounced (Fig. 2.3.f). The effects of increasing and decreasing *scalesperOctave* is depicted in Fig. 2.3(g)-(h) with a higher value resulting in a finer ridge map. The effect of increase and decrease in the number of shear levels on the final ridge map is quite minimal as can be seen from Fig. 2.3(i)-(j). The effect of anisotropy parameter *alpha* is depicted in Fig. 2.3(k)-(l) with minimal anisotropy resulting in a finer ridge map. The *minContrast* parameter is a grayscale threshold (values from 0 - 255) applied to Eq. (A.27) to extract ridges. A larger value suppresses noisy features as can be seen from the comparison between Fig. 2.3(m)-(n). The *offset* parameter is a scaling offset between odd-symmetric and even-symmetric shearlets quantified in octaves. Reducing the value of this parameter results in a coarser ridge map with enhanced connectivity (Fig. 2.3.o) compared to the larger value which results in a finer map (Fig. 2.3.p).

From an interpreter's point of view, three different scales of fracturing need to be identified and false features also need to be suppressed. From the sensitivity analysis, the parameters that are most important to generate high-probability ridge maps, are the wavelet supports (required to capture multiple scales of fracture), grayscale contrast (suppressing noise and thereby false features), and even-odd offset (which suppresses ridge detachments). This example illustrates the necessity of computing a ridge ensemble instead of searching for an ideal parameter combination.

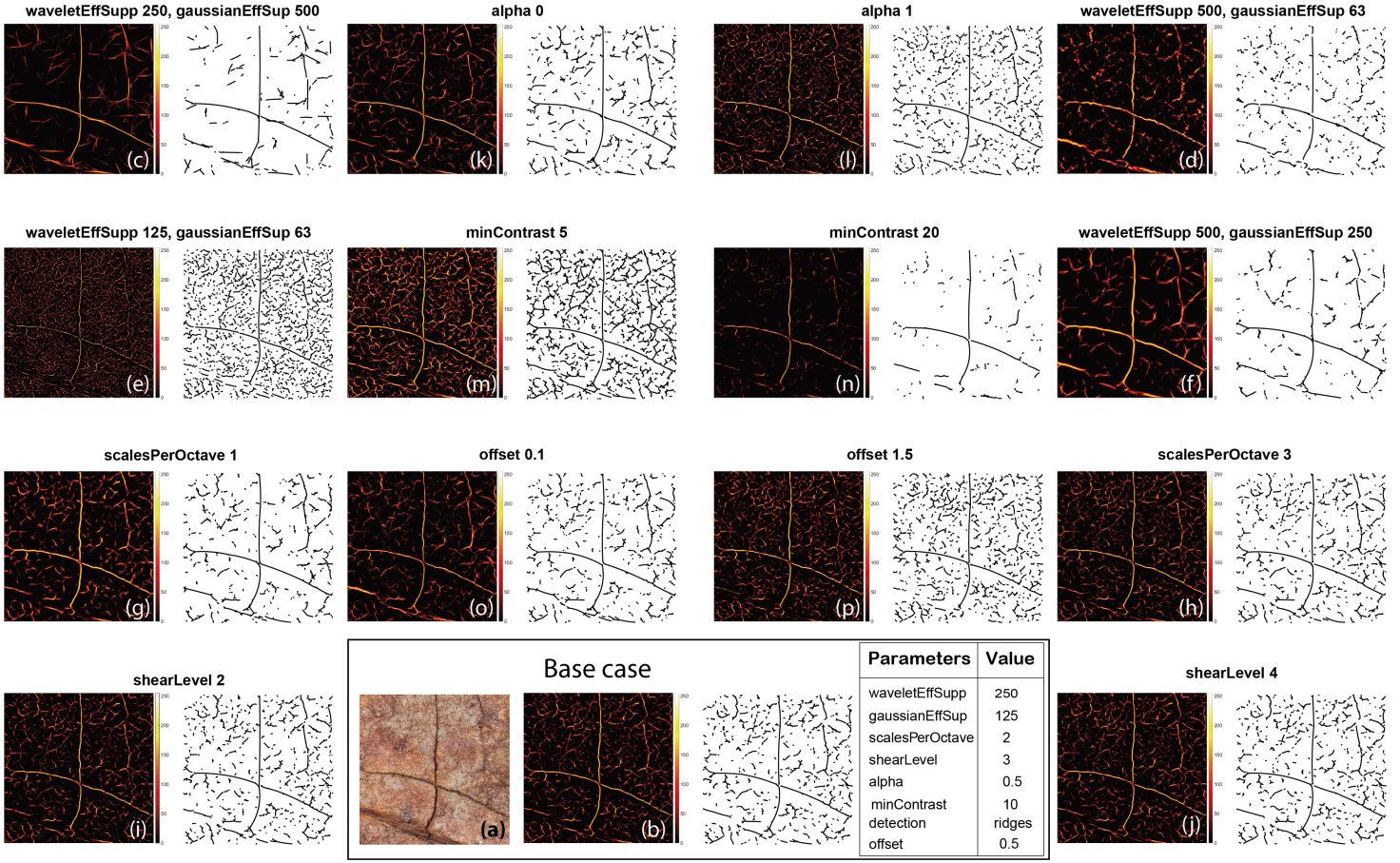

Figure 2.3: Effects of variation of ridge parameters on extracted ridges and the corresponding vectorizations using a fractured siliclastic example. A constant grayscale threshold is applied to the ridge map and all other parameters with respect to post processing are kept constant (a) 5 x 5 cm fractured rock image and base case parameters in table (b) ridge maps and vectorized traces for base case (c) effect of a higher gaussian effect support compared to wavelet support (d) effect of a large difference in wavelet effective support with respect to gaussian support (e),(g),(i),(k),(m),(o) lower bounds of parameters with respect to base case, corresponding ridge maps and traces (f),(h),(j),(l),(n),(p) upper bounds of parameters with respect to base case, corresponding ridge maps and traces

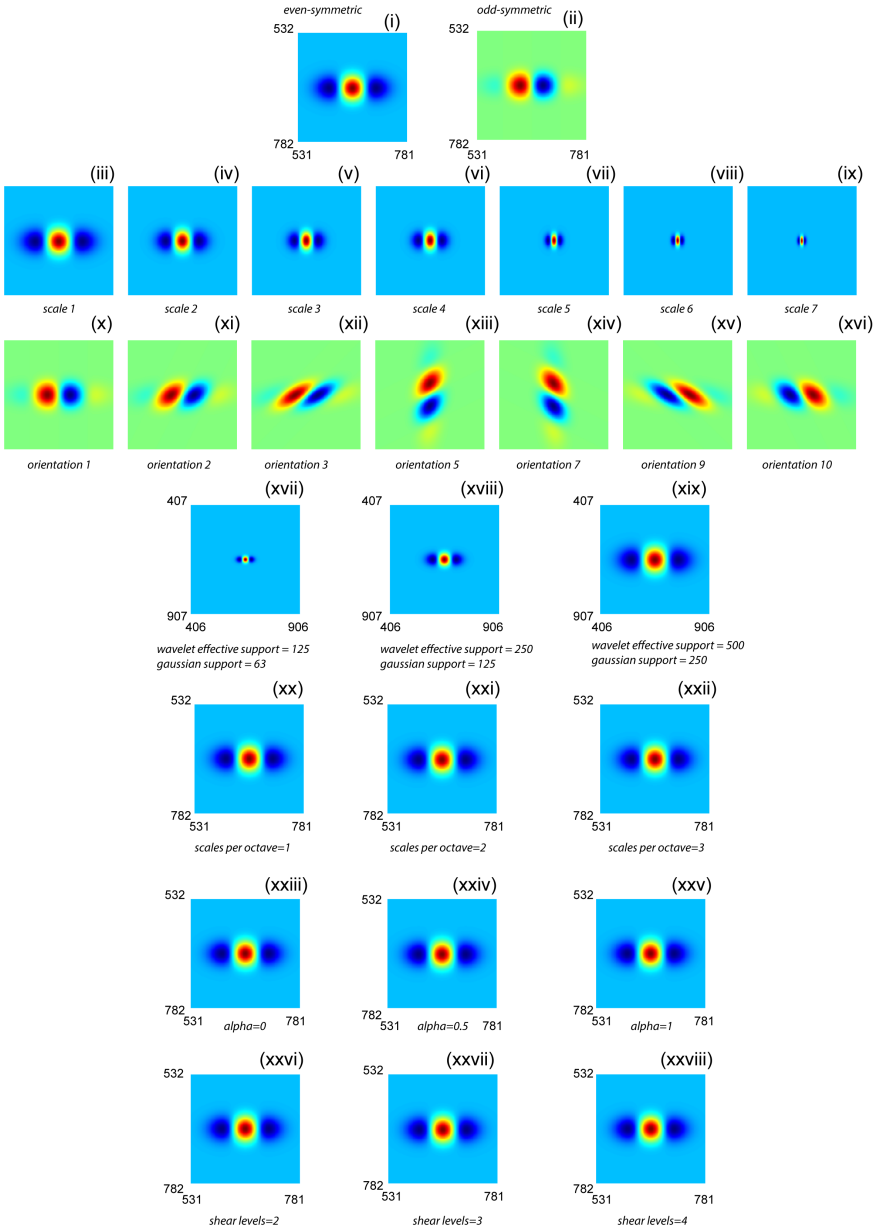

Figure 2.4: Effects of parameter variation on the constructed complex shearlet system for the fractured siliclastic example (i) Even-symmetric elements of the complex shearlet system constructed using the base case parameters in Fig. 2.3. Full system is 1313 x 1318 pixels. (ii) Odd-symmetric elements of the complex shearlet system using the base case parameters in Fig. 2.3 (iii) - (ix) depiction of seven scales (x) - (xvi) depiction of seven orientations (out of possible 10) for the odd-symmetric elements of the complex shearlet system (xvii) - (xix) effect of wavelet effective support and gaussian effective support on the even-symmetric elements of the complex shearlet system (xx) effect of gaussian effective support double that of wavelet effective support (xxi) effect of wavelet effective support very large than gaussian effective support (xxii) - (xxiv) effect of scales per octave on the even-symmetric elements (xxv) - (xxvii) effect of anisotropy parameter on the even-symmetric elements (xxviii) - (xxx) effect of shear levels on the even-symmetric elements

2.3.3. SHEARLET PARAMETER SELECTION

To decide upon the shearlet parameter space to generate multiple ridge realizations, we chose one sample image (see Fig. 2.6.a). Base case parameters are chosen based on recommendations underlined in Reisenhofer et al. [2016] for shearlet construction and ridge detection and these are tabulated in Table. 2.1. The use of the base case parameters results in the overlay depicted in Fig. 2.6(b). As can be observed from visual inspection of the overlay of the detected ridges over the original image, the automatic method can extract a large number of fractures. However, there are still some false positives (features detected on the trees and inside the large karstic cavities) and false negatives (undetected small scale fractures).

Table 2.3: Ensemble for Parameter Variation

Parameter	**Values**
waveletEffSupp	Original image size in pixels divided by 5, 8, 10, 12 and 15
gaussianEffSupp	Original image size in pixels divided by 5, 8, 10, 12 and 15
scalesPerOctave	1,2,3 and 4
shearLevel	2,3 and 4
alpha	0, 0.25, 0.5, 0.75, 1
minContrast	1, 5, 10, 15, 20
even/odd offset	0.001, 0.01, 0.1, 1, 2

To select the parameter ranges, we vary parameters with respect to the base case ridge image, thereby generating multiple ridge images. We use the structural similarity measure [Wang et al., 2004] to quantify the difference between the base case ridge image and other ridge images. Structural similarity (SSIM) is a measure commonly used in image quality assessment that returns one value as a measure of similarity between two images, where one image is the reference image. The SSIM is calculated for each ridge realization image corresponding to each parameter with respect to the base case ridge image. The SSIM for variation in scaling offset, anisotropy scaling α, Mexican hat wavelet support, gaussian filter support scales, minimum contrast, scales per octave, and number of shear levels are depicted in Fig. 2.5 according to the range of parameters in Table 2.3. From the analysis of the effects of parameters, we decided to vary the shearlet construction parameters so that we have 70 complex shearlet systems (see Table A.1 for the parameters used to construct the 70 complex shearlet systems).

The total number of stochastic runs for the ridge detection is the number of combinations of shearlet systems and ridge specification parameters. Using such an approach, a probability map of detected features may be obtained based on which cut-off thresholds can be defined to remove false positives. The result of such a stochastic run with 1050 realizations is depicted in Fig. 2.6. From this result, the utility of the method is evident wherein the features that are obscured by shadows and the shrubbery have a low strength signal which can then be filtered away thus reducing the number of false positives. Another advantage is that both large scale and fine features are captured which may not be possible using a single set of shearlet parameters.

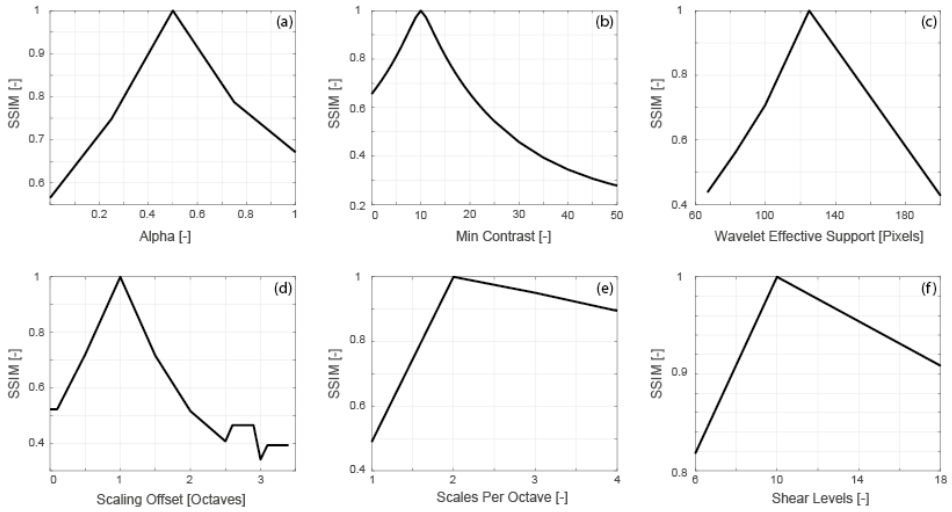

Figure 2.5: Variation of SSIM of of base case ridges with shearlet & detection parameters (a) SSIM vs Anisotropy Exponent (b) SSIM vs MinContrast (c) SSIM vs Wavelet Effective Support (d) SSIM vs Scaling Offset (e) SSIM vs Scales / Octave (f) SSIM vs Shearlevels

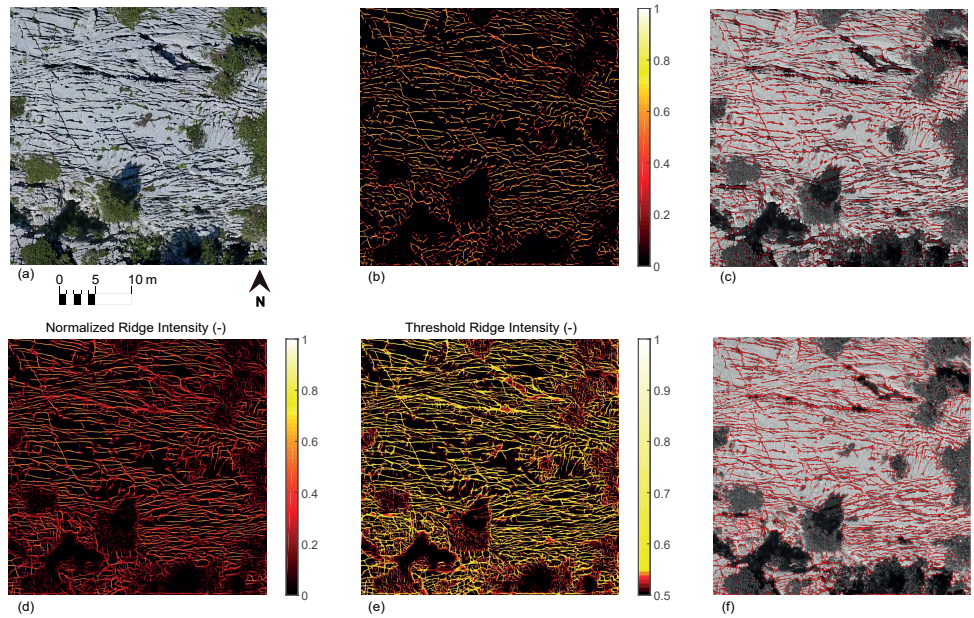

Figure 2.6: Effect of multiple ridge realizations (a) test image (b) Ridge map using base case shearlet parameters in Table. 2.1 (c) Overlay of ridges using base case shearlet parameters (d) Normalized ridge intensity ensemble map obtained after 1050 ridge realizations (e) Thresholded ridge intensity map (f) Overlay of ridges with threshold ridge intensities

2.4. RESULTS

2.4.1. TRACE EXTRACTION RESULTS FROM PARMELAN, FRANCE

GEOLOGICAL SETTING OF THE PARMELAN PLATEAU

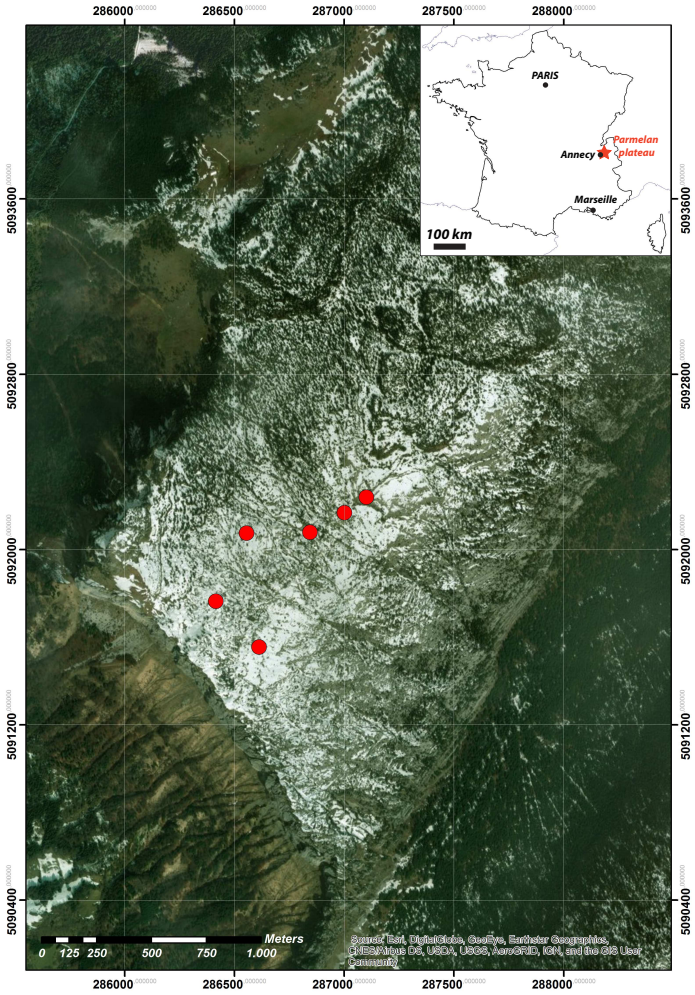

Figure 2.7: Location of the Parmelan plateau in France within the Bornes Massif depicting drone flight base points for six drone missions

We tested the automated fracture extraction method on an example from a carbonate outcrop from the Parmelan plateau in the Bornes Massif, France. The Bornes Massif is a northern subalpine chain in the western French Alps. The method was applied on a photogrammetric orthomosaic derived from a 3D outcrop model. The outcrop model was built from source photos acquired using a DJI Phantom 4 UAV. The image resolution

is 18.6 mm/pixel. Processing of the drone images and generating the orthomosaic was done using AgiSoft PhotoScan Professional (Version 1.2.6) [2016*] software. The Parmelan Anticline in France (see Fig. 2.7) is situated in the frontal part of the Bornes Massif and consists of Upper Jurassic to Cretaceous rocks of the European passive margin [Berio et al., 2018, Gidon, 1996, 1998, Huggenberger and Wildi, 1991].

This NE – SW trending anticline consists of a wide, flat crestal plateau bounded by steeply dipping limbs. Carbonates form the roof of a kilometre- scale box fold formed during the Alpine orogeny [Bellahsen et al., 2014]. On the crestal plateau, a 1.7 km by 2.3 km large pavement of flat-lying shallow-water carbonates is exceptionally well exposed. The Parmelan outcrop is a good example of fracture patterns formed in a fold-and-thrust setting. We applied the automatic fracture detection technique on an orthomosaic that has been stitched together from drone photogrammetry over six different drone missions over the Parmelan. The combined extent of the six orthomosaics is depicted in Fig. 2.8(a), and the areal extent of each orthomosaic is depicted in Fig. 2.8(b).

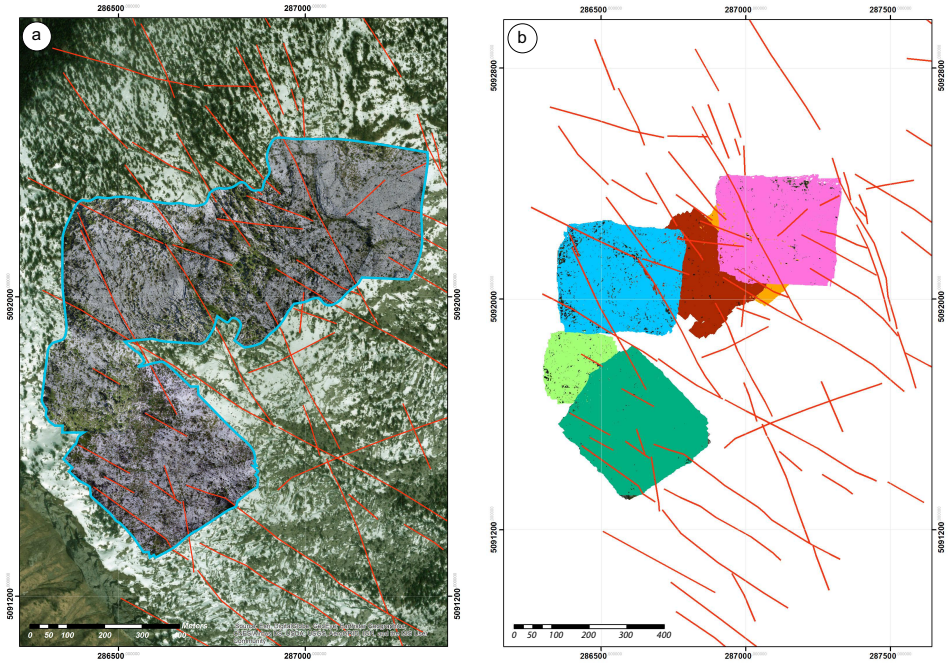

Figure 2.8: Drone photogrammetry coverage area from the Parmelan (a) Region within the Parmelan plateau highlighting the areal extent of the drone photogrammetric orthomosaics which are projected over the base map. Manually traced large scale faults are depicted in red. (b) Spatial extent of the drone coverage of each of the six UAV flight missions in different colours

AUTOMATIC EXTRACTION RESULTS ON THE PARMELAN ORTHOMOSAIC

Considering memory requirements and for faster computation, the image domain was divided into georeferenced sub-tiles using the Grid Splitter plugin in QGIS software. Visual filtering was carried out to remove tiles that did not have exposed rock, had a large

degree of shrubbery, and which were at the orthomosaic edges where image resolution is poor. A total of 1000 tiles were chosen for the automated interpretation process. The areal extent of the orthomosaic covered 0.589 km^2, and this region is depicted in Fig. 2.8. The region covered by the tiles amounts to 0.379 km^2 and this is shown as an overlay of the selected tiles in Fig. 2.9(a). Structural measurements were recorded at four small scale stations (around 2-5 sq.m per station) depicted in Fig. 2.9(c)-(f).

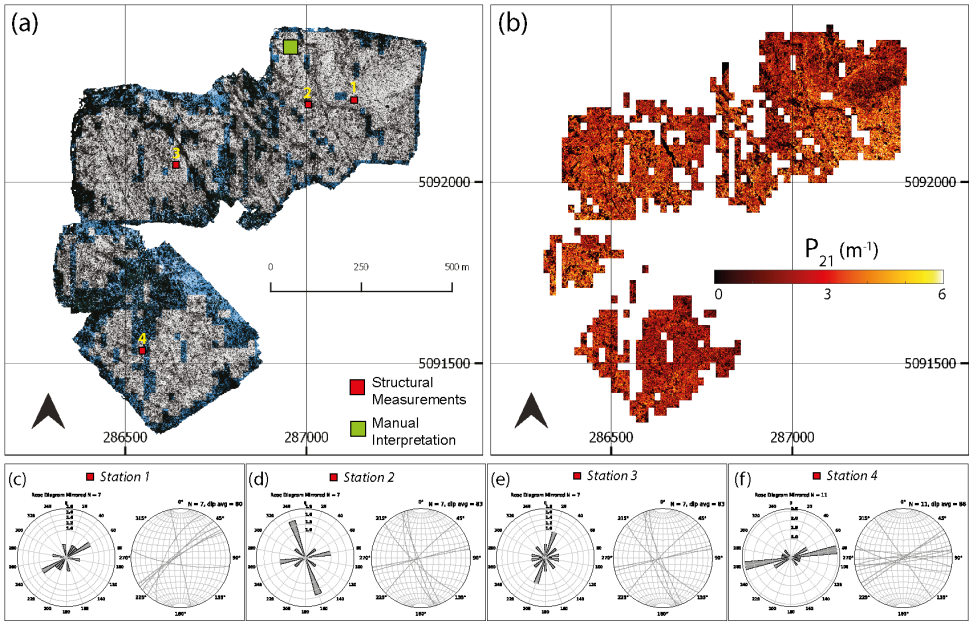

Figure 2.9: Trace extraction results from the Parmelan (a) Selected tiles from the orthomosaic spatial extent are highlighted (b) Spatial variation of the fracture intensity depicted as a P_{21} plot computed using the box counting method (c) Rose and stereoplot of Station 1 (d) Rose and stereoplot of Station 2 (e) Rose and stereoplot of Station 3 (f) Rose and stereoplot of Station 4

An ensemble of 1050 ridges was computed using a set of shearlet parameters. A threshold for the ridge intensity was chosen to filter out the false positives. The threshold was determined by a visual examination of the overlay of detected ridges over the original images. The subsequent post-processing steps yielded features in each tile. These were geo-referenced and stitched back into a single vectorized file representation. Around 3 million features were extracted from the Parmelan orthomosaic. The P_{21} fracture intensity was computed using the box-counting method by dividing the tile into a 25 x 25 (pixels) regular grid. The P_{21} fracture intensity plot highlights the spatial variation of fracturing over the Parmelan plateau (see Fig. 2.9.b). The vectorized fracture shape files along with the Parmelan basemap are presented as a public dataset.

COMPARISON WITH MANUAL INTERPRETATION

To compare results of the automated approach to a manual interpretation, we chose a sub-region within the Parmelan orthomosaic. The selected subregion depicted in Fig.

2.10(a) consists of a 24 m x 24 m tile of the Parmelan orthomosaic. The image indicates fractures that seem to be isolated, without a well-connected topology, and which are predominantly aligned along NW-SE direction. The fracturing intensity is variable across the tile. The contrast between fractures and the host rock fabric is intensified by the karstification of the fractures, which can be attributed to weathering and dissolution. Figure 2.10(b) depicts an overlay of the automatically interpreted fractures overlain over the original tile. A total of 2910 features was extracted in this tile. This example highlights some of the technical challenges associated with automated fracture trace detection. Shrubbery is present in the image which obscures certain relevant features. The north-western corner of the image is blurred since it forms the extent of the orthomosaic.

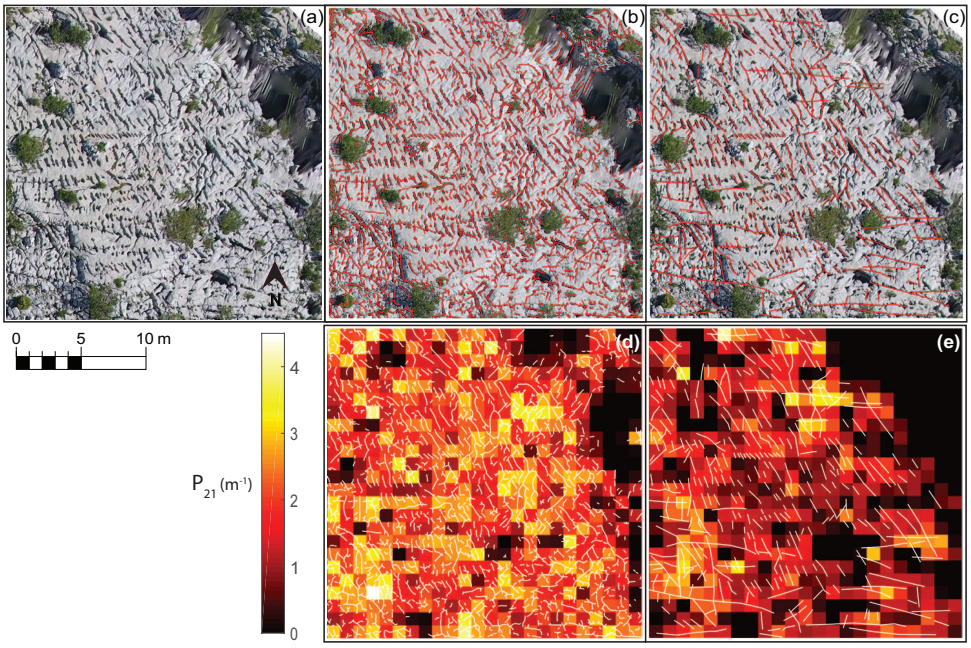

Figure 2.10: Comparison between automatic and manual interpretation on a tile from the Parmelan (a) Tile from the Parmelan orthomosaic depicting intense fracturing with an organization along the NW-SE corridors (b) Overlay of fractures traced using the automatic detection method (c) Overlay of fractures manually traced for the tile at a zoom of 1:2000 (d) P_{21} Fracture intensity for automatic extracted fractures (e) P_{21} Fracture intensity for manually extracted fractures

The image also depicts open cavities or blobs, which could be the result of localized weathering. The effect of the cavities on the feature extraction is that only an edge is detected. Overall the fracture extraction efficiency is quite dependent on the resolution and quality of images. In the case of the Parmelan data acquisition, the UAV was flown at an altitude of 50-70 metres above the pavement; therefore, features such as closed veins, and slightly open fractures are below the resolution of the drone camera. A higher image resolution is necessary to extract such features. In our specific case study, good lighting and exposure during the UAV flight mission prevented shadows from obscuring the im-

agery. Fig. 2.10(c) depicts a manually performed interpretation at a zoom level of 1:2000 on the raster image with a total of 341 features. P_{21} fracture intensity comparisons of both automatic and manual traces are shown in Fig. 2.10(d)-(e). The difference between the automatic and manual interpretation highlights the inclination of the interpreter to neglect small scale features. Based on geological experience and prior knowledge of the field area, there is a tendency to interpret and link together disconnected features from the original raster image. The closest small-scale station to the sub-tile depicted in Fig. 2.10(a) is station 2. There is agreement between the rose plots of station 2 (see Fig. 2.9.c) and the interpretations (Fig. 2.10.e-f). The observed fractures in both cases are predominantly sub-vertical.

APPLICATION TO MINERALIZED FRACTURES

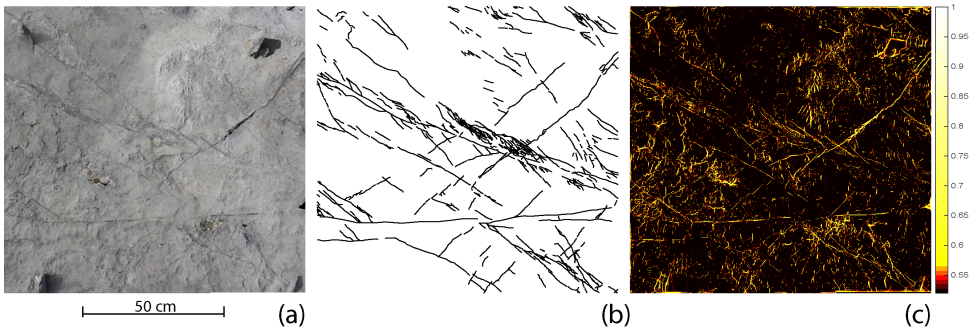

Figure 2.11: Extension of the automated method to extract mineralized fractures (a) Image from Parmelan depicting mineralized fractures (b) Manual interpretation of mineralized fractures (c) Ridge ensemble

We now showcase an example of a close-range image containing mineralized veins that are invisible to photogrammetry at altitudes of 40 - 70 m . The resolution of this image is 0.18 mm/pixel and was taken using a handheld DSLR camera. In this high-resolution image, the fracture infill has similar colour as the host rock as can be seen in Fig. 2.11(a). A manual interpretation of the veins (at a zoom of 1:750) is depicted in Fig. 2.11(b). Using a well-tuned set of parameters with reduced wavelet effective supports, it is possible to extract the much thinner and subtle features as depicted in Fig. 2.11(c). It can be observed from comparison between Fig. 2.11(b)-(c), that a large number of false features are also highlighted alongside the features of interest. The main contributors to the extraction of these non-fracture features are the natural rugosity of the rock face, presence of pebbles, pockmarks, and erosion features. The arrangement of these artefacts display a very different pattern; small lines with random direction compared to the fractures which are consistently oriented and quite continuous. The veins are also of different thicknesses, with a few veins anastomosing and some branching in a horsetail manner. Some of the thicker veins also exhibit microstructure within the mineral infill. Further tuning of parameters in order to capture all the veins while also suppressing false features is quite challenging and hence we do not explore this in further detail. Despite the noise, the automated method is not limited to capturing only open fractures but can also extract closed fractures.

2.4.2. TRACE EXTRACTION RESULTS FROM BREJÕES, BRAZIL

GEOLOGICAL SETTING OF THE BREJÕES PAVEMENT

The second case study for the automated extraction method is a carbonate outcrop from the Irecê Basin, Central Bahia, Brazil (see Fig. 2.12.a-b). The Irecê Basin is located within the northern region of the São Francisco Craton. The Brejões pavement study area is within the Irecê Basin and consists of Neoproterozoic platform carbonates of the Salitre Formation (750-650 Ma). The Neoproterozoic cover was affected by the Brasiliano Orogeny (750-540 Ma) in two separate folding events resulting in fold belts around edges of the São Francisco Craton [Ennes-Silva et al., 2016]. The Brejões pavement UAV imagery that we used for our analysis was acquired by Boersma et al. [2019]. Structural measurements from Boersma et al. [2019] is shown in Fig. 2.12(c). The orthomosaic covers an area of 0.81 km^2 and consists of fractured, black oolitic limestones that correspond to Unit A1 of the Salitre stratigraphy [Guimarães et al., 2011]. The resolution of the Brejões orthomosaic is 20.3 mm/pixel.

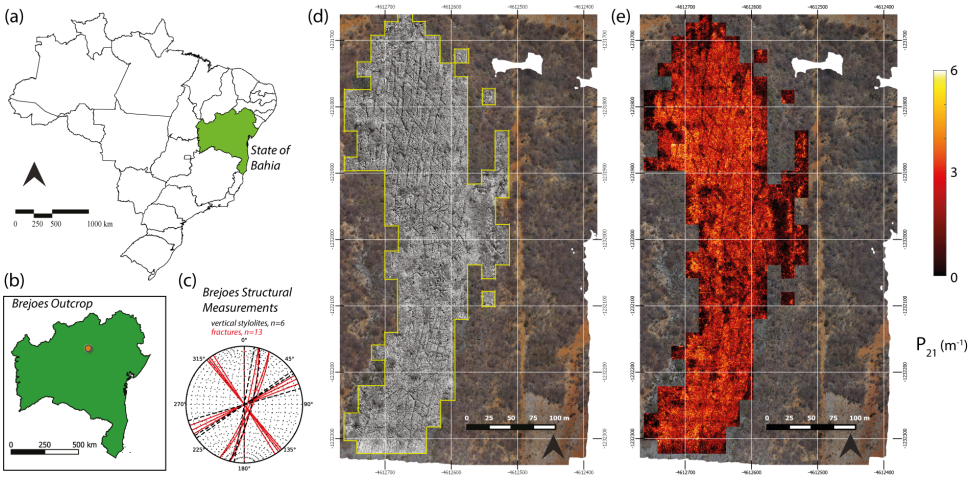

Figure 2.12: Trace extraction results from the Brejões outcrop (a) Bahia state in NE Brazil (b) Location of the Brejões outcrop in the state of Bahia (c) Selected tiles from the Brejões orthomosaic for the automated extraction (d) Spatial variation of the fracture intensity depicted as a P_{21} plot computed using the box counting method

AUTOMATIC EXTRACTION RESULTS ON THE BREJÕES ORTHOMOSAIC

The Brejões orthomosaic is split into 222 tiles for the analysis and this region is shown in Fig. 2.12(d). The Brejões example has a different fracturing style than the Parmelan and consists of an intricate pattern of multi-scale conjugate fractures. The shearlet combinations utilized in the case of the Parmelan was insufficient to capture this variation in scales. Specifically, in the Brejões case, the large scale features were not captured. A visual inspection of the ridges was necessary to identify the shearlet combinations that amplified the large scale features. The contribution of these ridges was increased (factor of 8) in the ridge ensemble to highlight these large deformation features. Fig. 2.12(e) depicts the P_{21} fracturing intensity computed using the box-counting method by dividing

each tile into a 25 x 25 (pixel) regular grid. The vectorized fracture shape files along with the Brejões basemap are presented as a public dataset (see Prabhakaran et al., 2019b).

COMPARISON WITH MANUAL INTERPRETATION

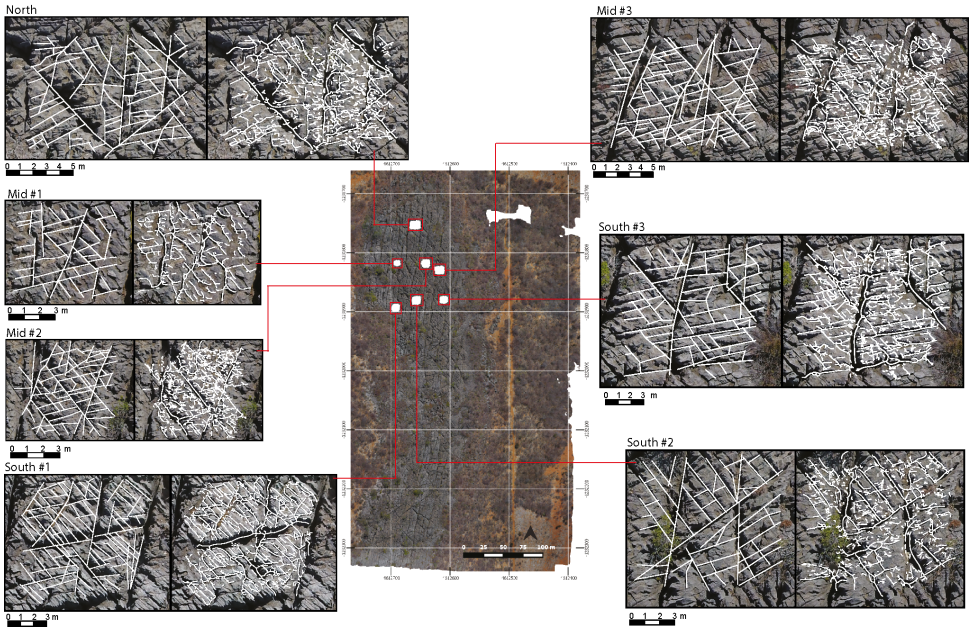

Figure 2.13: Comparison between manual (left) and automatic (right) interpretation on seven stations within the Brejões outcrop. The manual interpretations were obtained from Boersma et al. [2019].

The automatically extracted features from the Brejões image data were compared with manual interpretations performed by and obtained from Boersma et al. [2019] at seven stations. The automatic interpretations were trimmed to the peripheries of the manual interpretations for a fair comparison between both vectorizations. The location of these stations alongside the automatic versus manual interpretations are shown in Fig. 2.13. A comparison of the rose plots and cumulative length distributions of the manual and automatic interpretations is depicted in Fig. 2.14. A few observations can be made from the comparison. Firstly, similar to the Parmelan case, the interpreter picks a lesser number of features. Secondly, there is a tendency to extend fractures across image regions where there is no real evidence of rock failure. Thirdly, there is an inconsistency in specifying the connecting topologies between the interpreted traces. In some stations (see Mid #2, Mid #3 and North in Fig. 2.13), the automated interpretation suffers from a large number of false positives. A close examination indicates that the presence of shadows and eroded, undulating topography of the rocks are the main reasons for these false positives. In the Brejões case, the drone was flown at around 10.00 AM, and hence the exposure of the outcrop face was not optimal. The inclined illumination enhances shadows on the rugged topography, which are then seen as false positives in the automatic interpretation. False positives due to shrubbery are minimal in the station regions

considered.

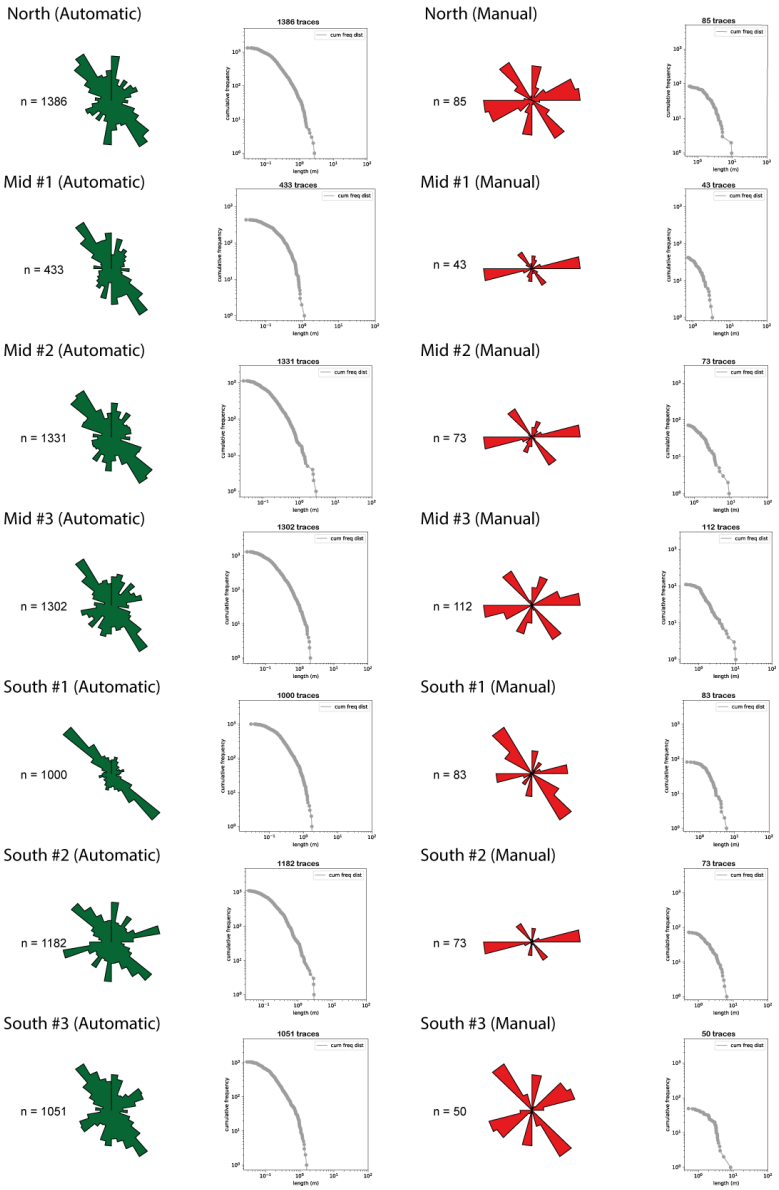

Figure 2.14: Comparison of trace length weighted rose plots and cumulative trace length distributions for automatic and manual trace interpretations from Brejões outcrop stations.

2.4.3. BENCHMARKING WITH DATA FROM THIELE ET AL. [2017]

We further tested the automated trace detection on a recently published case study from Thiele et al. [2017]. The images selected are orthophotographs of two 10 x 10 m areas from Bingie Bingie Point, New South Wales, Australia (Fig. 2.15.a, Fig. 2.16.a). The exposed rocks are Cretaceous to Paleogene dykes, intruding diorites, and tonalities cross-cut by joint sets (Thiele et al. 2017). The images are complex as they contain both open and closed fractures of different scales, distributed between multiple lithological layers. The images also contain water, shadows, and debris, which makes it even more challenging. We chose this dataset to benchmark the quality of our results with those presented using the semi-automatic cost function-based trace mapping approach of Thiele et al. [2017].

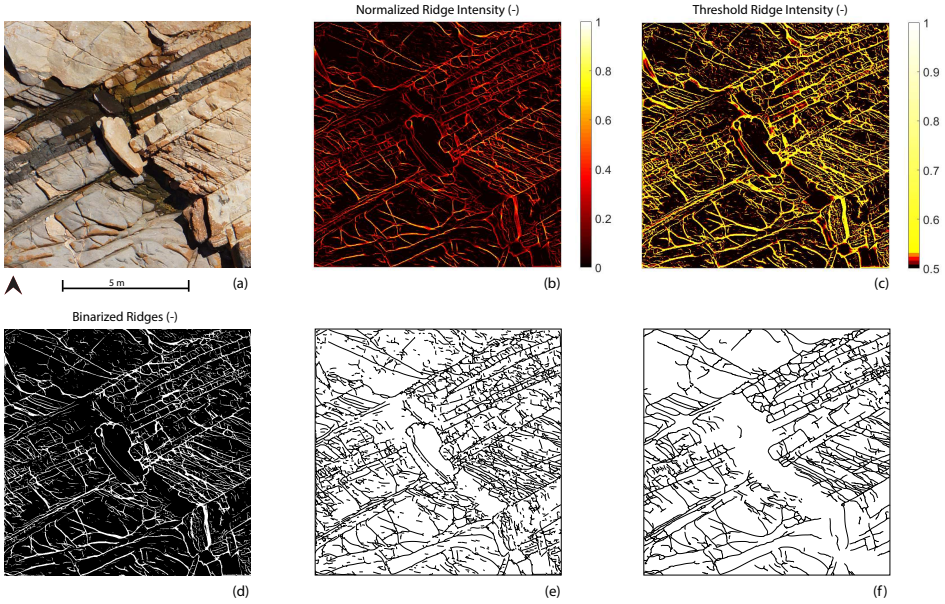

Figure 2.15: **Comparison of Benchmark Image 1 (a) Bingie Bingie Area 2 from Thiele et al. [2017] (b) Normalized ridge map using complex shearlet automatic extraction (c) Threshold applied to the normalized ridges (d) Binarized ridges map (e) Vectorized traces (f) Assisted cleaned up trace map for Area 1**

The variation in fracture scales implied that similar to Brejões, a different set of shearlet combinations were needed. We generated 2700 ridge realizations which were used to construct a normalized ridge ensemble map for both images (Fig. 2.15.b, Fig. 2.16.b). A simple, non-linear sigmoid function was applied to the normalized ridge intensity to enhance ridge strength (Fig. 2.15.c, Fig. 2.16.c) and a threshold was chosen based on visual comparison with the source image to yield highly probable, binarized ridge images (Fig. 2.15.d, Fig. 2.16.d). The subsequent workflow steps, as described in Sect. 2.3.1 were followed to obtain vectorized traces (Fig. 2.15.e, Fig. 2.16.e). The vectorized traces were used to render assisted interpretations depicted in Fig. 2.15(f) and Fig. 2.16(f) which are comparable in quality to the assisted interpretation of Thiele et al. [2017].

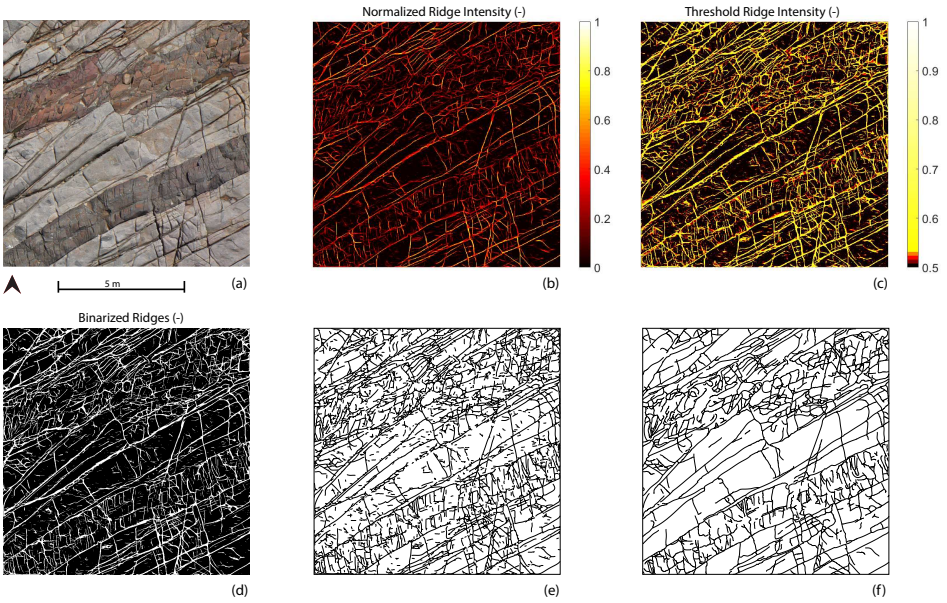

Figure 2.16: **Comparison of Benchmark Image 2 (a) Bingie Bingie Area 2 from Thiele et al. [2017] (b) Normalized ridge map using complex shearlet automatic extraction (c) Threshold applied to the normalized ridges (d) Binarized ridges map (e) Vectorized traces (f) Assisted cleaned up trace map for Area 2**

In the published results of Thiele et al. [2017], assisted interpretations of both areas are achieved in 37 minutes and 34 minutes, respectively. We can report better performances of 27 and 32 minutes for the same areas. The time does not include computing of the ridge realizations. Once the high probability trace map was generated, the subsequent steps of the automated detection workflow took around 3 minutes. The remainder of the time was used to perfect the assisted interpretation. The post-processing tasks performed in this second step were the removal of false positives owing to shadows, water, and debris and joining of segments which were disjointed due to poor resolution within the image. Though we have performed a benchmarking exercise with the data from Thiele et al. [2017] and also compared our results with manual interpretation, it would be useful to compare with more manual interpretations to further validate the accuracy of the technique. Such comparison, however, can be done only on networks which are either limited in their spatial extent or in the number of features interpreted. For large orthomosaics, a benchmarking exercise can be challenging as few manually rendered datasets are comparable in network size.

2.5. DISCUSSION

Extraction of fracture traces from photogrammetric data is a necessary processing step to construct DFN representations. DFNs created using fracture patterns that are directly extracted from rock images, are advantageous as they honour the spatial architecture of fracture networks. Automated extraction methods reduce the human component in

data processing, and we have achieved this using the complex shearlet transform ridge detection method accompanied by post-processing steps. The complex shearlet method can detect both edges as well as ridges in fractured rock images. We find that the ridge measure works very well for extraction of fractures, and we use the ridge measure in all our case studies. Though the method performs very well and can extract much more traces than is possible manually while reducing interpreter bias, there are some issues that need to be mentioned. In this section, we discuss on the validity and limitations of the technique, areas where there is scope for further development, and also describe some potential extended applications of the method.

2.5.1. VALIDITY AND LIMITATIONS

Detection of mineralized features: The method works well when the features of interest are barren and prominent. When fractures are closed and filled then they are generally harder to detect and require high resolution images (< 1 mm/pixel) which can be recorded only at very close ranges at very-low UAV flight altitudes. Recent outcrop studies [Ukar et al., 2019] indicate that many of the barren features in outcrop are absent within the same subsurface lithological unit while maintaining good correspondence between mineralized features in both outcrop and subsurface. When mineral fill has a marked colour contrast with respect to the host rock (as in vein data published recently by Meng et al., 2019), then superpixel segmentation algorithms can be successful [Vasuki et al., 2017]. In the case of poor contrast, the complex shearlet transform would require a great deal of manual tuning of detection parameters to extract reliable results. At such close ranges, as is needed for veins extraction, it is also likely that many more noisy features un-related to fracturing would arise. Since mineral-fill of fractures can provide a clearer picture into evolution, timing, and stress history of fractures, identifying them on an outcrop scale is important. This is doubly significant, when the goal is to directly extrapolate fracture patterns from a particular outcropping to the same subsurface target. In such a case, close range UAV-mounted hyperspectral data acquisition would be better suited than conventional imaging and image processing methods. With hyperspectral imaging, data is collected in near-continuous spectral bands. The spectral response of minerals constituting the rock, owes to atomic-molecular level processes triggered on interaction with a light source (active or passive) and this may be utilized to identify mineral composition. Since mineral fill of veins are likely to have a different spectral response from the mineralogy of the host rock, this variation may be used to isolate the pixels that correspond to veins. A recent review on close range hyperspectral imaging for mineral identification identifies various previous studies performed for specific minerals [Krupnik and Khan, 2019]. It would be interesting to observe, identify, and distinguish between mineralized sequences based on the differences in spectral response of the fracture infill material. Since hyperspectral data is much more voluminous and with significantly more complex image processing than conventional photogrammetry, such analysis could be confined to selected regions within the outcrop. In conjunction with conventional UAV photogrammetry that covers larger spatial area, laboratory-based geochemical studies, and outcrop observations (scanline sampling, abutting relations etc.), a more detailed fracture characterization may be conducted.

Detection of large cavities and false features: Both the Parmelan and Brejões pavements exhibit karstification with the Parmelan containing many more collapsed karstic regions. The presence of such low-aspect ratio discontinuities are quite rare in siliciclastic and volcanic outcrops but can prove problematic to the application of the method in carbonate outcrops where karstification is severe. Both the ridge and edge measures would fail in identifying such blobs or would at best, extract the periphery of the cavity. In recent work by Reisenhofer and King [2019], blob detection measures have been developed within the shearlet framework and could potentially solve this issue. Another issue is the effect of undulating topography and shrubbery in generating false positives. False positives generally appear when there is shrubbery, shadows, very rugged terrain, and non-fracture bedding planes. In the case of the Parmelan, the use of multiple ridges was successful in suppressing the false positives owing to shrubbery. However, in Brejões, false positives due to underbrush were more difficult to suppress because they shared the same scale as that of the fractures. In Brejões, shrubbery was also present within some of the wider fractures causing false negatives. In such cases, manual interference is necessary to either mask the regions of shrubbery before the automatic extraction or to remove (or connect) the vectorized traces after the automated extraction. Additionally, carbonate outcrops are prone to widespread erosion owing to exposure to meteoric water from precipitation cycles and air corrosion. Geomorphological features owing to these erosive processes may also play a role in generation of false positives.

Parameter selection A significant difference in fracture scales within an image of interest can prove problematic for the method. In such a case, a vast number of ridge detection runs and associated increase in computational time is needed to construct a ridge ensemble that takes into account all scales of discontinuities and yields a satisfactory result. When such variation is localized, the image could be segmented into regions that correspond to varying fracture intensities and processed separately. This may be difficult to assess *a priori* and in such cases, would require trial runs. In the Brejões outcrop example and the close range Parmelan vein example, this difference in fracture scales was ubiquitous throughout the exposure and more pronounced than the Parmelan outcrop. Using visual comparison with the original image, the effect of ridges resulting from certain shearlet parameter combinations was enhanced, so that the ridge ensemble is improved. In Brejões, it was the large scale features that needed to be strengthened while in the case of the Parmelan vein example, the smaller features needed sharpening. Since parameter selection is still done manually, a more comprehensive way of arriving at the optimal shearlet combination is desirable. An algorithm that automatically optimizes for shearlet parameters corresponding to each individual scale of fracture is worthy of attention.

Artificial fragmentation of traces Manual fracture interpretation from images often involves the step of classifying fracture traces into separate sets based on ground truth observations or with respect to fracture strike. The automated method described here in its current form can only extract traces and cannot distinguish/classify traces as belonging to separate sets. When fractures intersect each other, the issue of artificial fragmentation of seemingly continuous traces arises. If an image consists of two orthogonally intersecting fractures, the automated method would result in four traces intersecting at

a single branch point, even though a manual interpretation would only identify two fracture traces belonging to two different geometric sets. This type of fragmentation would result in different length distributions; however, such fragmentation is not an issue if the outcrop DFN is used for geometric input for flow/geomechanics simulation. This is because the process of meshing models with explicitly specified DFN geometry would, in any case, require the specification of all intersection points (or forced fragmentation of long intersecting fractures). Therefore, the practitioner must exercise caution when using cumulative length distributions derived from outcrop DFNs that are automatically extracted. A single fracture could also be fragmented without being cut by other intersecting fractures. This may happen in the case of false negatives (due to shadows falling over part of fracture, debris or shrubbery within an open fracture, and when fracture opening is very thin at some regions along fracture length) that cause fragmentation of fractures with gaps in between them. This kind of fragmentation affects the topology of the network in addition to depressing the height cumulative length distribution. It maybe noted that manual specification of fracture endpoints is also fraught with bias [Peacock et al., 2019]. A solution would be to use a range of linking thresholds to connect traces and study the effects of threshold values on network topology and length distribution.

2.5.2. RECOMMENDATIONS FOR FUTURE WORK

Link between extractable P_{21}, drone flying altitude, and camera resolution: From the P_{21} analysis on the Parmelan and the Brejões automatically extracted fractures, the maximum value P_{21} was around eight m^{-1}. The same drone model was used in both cases (DJI Phantom 4), and the flying altitude was also similar (between 40 and 70 metres). Although such a conjecture needs further verification, there could be a relation between the resolution of imagery and maximum extractable fracture intensity. Often flight altitudes are chosen by drone pilots depending upon considerations such as local topography, weather conditions, and presence of impediments (such as trees, electricity poles, and telecommunications towers). A detailed analysis of the relation between flying altitude (and consequently image resolution) and extracted fracture intensity could provide drone pilots with insights and guidelines for UAV-based outcrop analysis. The ideal flying resolution to identify features of interest may be ascertained by carrying out a series of acquisitions at a location where ground truth is known.

Generating data for fractured reservoir modelling workflows: Fractured reservoir characterization workflows in the oil and gas industry have traditionally used stochastic techniques that attempt to extrapolate averaged fracture statistics (either from borehole imagery, core data, or outcrop analysis) to reservoir volumes. The use of Multiple Point Statistics (MPS) for fracture network generation was highlighted by Bruna et al. [2019c] as an alternative approach to DFN modelling. MPS uses training images of realistic fracture networks to learn patterns and then generate non-stationary fractured reservoir models. Corrected for false positives and noise, the automated method can produce accurate, geologically realistic, and unbiased training images that can feed into the MPS workflow. Since our method can extract large scale fracture networks (millions of features from sub- square kilometre regions), it is also well suited to provide training data for deep learning workflows. Recently, the use of Generative Adversarial Networks (GANs)

for geological modelling at the reservoir scale was proposed by Dupont et al. [2018] and Zhang et al. [2019] as an alternative to conventional geostatistics, MPS, and object-based modelling. GANs form a subset of deep learning architectures that are used for generative modelling [Goodfellow et al., 2014]. GANs that are trained on realistic data can then generate geologically realistic, non-stationary models.

Deep learning methods for trace extraction: Deep learning methods have revolutionized computer vision applications. Various neural architectures have documented high degrees of accuracy in machine vision tasks such as overall image classification, identification and classification of objects within an image, localization of objects, extraction of regions of interest (semantic segmentation), and extraction of regions corresponding to individual objects (instance segmentation). The problem of fracture trace extraction falls within the problem category of region extraction of individual objects and hence may be attempted using techniques such as mask Regional Convolutional Neural Networks [He et al., 2017]. Deep learning methods, in general, require large amounts of labelled data to train. In the case of a mask RCNN, the library of training images must contain marked regions (or overlays) indicating pixels of interest that correspond to individual fractures. The automated method described in this manuscript can be used to rapidly generate a large number of overlay images that can be used as training data for mask RCNN architectures.

Using manual tracing vector data to clean-up traces: The automatic interpretation can generate false positives even in the best outcrops free from noise. This can happen even with deep learning methods with large amounts of training data. In such situations, we may consider using graph-based methods to clean up the vectorized polylines that are generated using any type of image processing method (shearlet-based or otherwise). Available prior geological knowledge in the form of manual trace maps can be used to generate plausible association scenarios between the extracted vectorized traces. In recent work by Godefroy et al. [2019] the Bron-Kerbosch graph algorithm was used to identify possible structures in 2D seismic intepretation. Cordonnier and Loukas [2019] introduced a graph neural network-based method in which missing paths in a network are inferred and assigned based on training data. Such approaches should also be considered in resolving the accuracy of fracture trace extraction to the levels necessary for geomodelling purposes.

2.6. CONCLUSIONS

This chapter presents a method to automatically detect and digitize fracture traces from images of rock fractures using the complex shearlet transform. The technique replaces the task of manually interpreting fractures, which is time-consuming, prone to interpreter bias, and which suffers from a lack of repeatability. The case studies that are presented highlight the utility of the complex shearlet-based measure for automatically detecting fracture traces from 2D images. The automatic trace detection method combines the complex shearlet ridge measure with a series of post-processing steps that include image segmentation, skeletonization, polyline fitting, and polyline simplification. We tested the method at different scales of rock displacement, at outcrop scale (~

10^2 m) and station-scale (< 10 m), using two orthomosaics reconstructed from drone photogrammetry and two rock pavement images. We have considered carbonate and igneous rock lithologies in the case studies. Using the method, we have extracted millions of 2D features from outcrop-scale drone orthophotos. The processing time of the technique depends upon the intensity of fracturing and the complexity of the fracture networks contained within the image. The automatic trace extraction results are quantitatively compared with manually interpreted fractures on selected sub-samples of the image domain using fracture trace density metrics. The automated technique is capable of extracting a much larger number of features, with a marked reduction in bias. The method outlined in this chapter greatly simplifies the process of generating deterministic, outcrop-based DFNs. The automatically extracted fracture patterns can be used by structural geologists to link deformation features to tectonic history and by geomodellers in sub-surface NFR modelling.

3

THE BRISTOL CHANNEL DATASET

The Lilstock outcrop in the southern Bristol Channel provides exceptional exposures of several limestone beds displaying stratabound fracture networks, providing the opportunity to create a very large, complete, and ground-truthed fracture model. In this chapter we present the result of automated fracture extraction of high-resolution photogrammetric images (0.9 cm/pixel) of the full outcrop, obtained using an unmanned aerial vehicle, to obtain a spatially extensive, full-resolution map of the complete fracture network with nearly 350,000 ground-truthed fractures. We developed graph-based functions to resolve some common issues that arise in automatic fracture tracing such as incomplete traces, incorrect topology, artificial fragmentation, and linking of fracture segments to generate geologically significant trace interpretations. The fracture networks corresponding to different regions within the outcrop are compared using several network metrics and the results indicate both inter- and intra-network (layer to layer) structural variabilities. The dataset is a valuable benchmark in the study of large-scale natural fracture networks and its extension to stochastic network generation in geomodelling. The dataset also highlights the intrinsic spatial variation in natural fracture networks that can occur even in weakly-deformed rocks over relatively short length scales of tens of metres.

This chapter is based on Prabhakaran et al. [2021b], Large-scale fracture network patterns: Insights from automated mapping in the Lilstock (Bristol Channel) limestone outcrops, *Journal of Structural Geology*, 150, 104405, https://doi.org/10.1016/j.jsg.2021.104405

3.1. INTRODUCTION

Fractures in rocks can form networks with fracture tips forming abutting or crossing physical interactions with other fractures or remaining isolated within rock matrix. The evolution into a final cumulative network pattern is history-dependent. The pattern evolves depending upon how loading paths and rock properties governing growth affect propagation rates under episodic conditions when failure criteria are met, in addition to the superposed effects of pre-existing sets [Laubach et al., 2019]. The spatial arrangements of fracture networks can be a significant geomorphic agent, influencing landscape evolution processes [Scott and Wohl, 2019], serve as dissolution pathways for karstic cave formation [Bertotti et al., 2020, Boersma et al., 2019], and influence subsurface fluid flow patterns that are relevant for hydrogeological, geo-energy and waste disposal applications [Berkowitz, 2002, National Research Council, 1996]. Given such non-trivial influences, it is important to be able to characterize large-scale fracture patterns from a network perspective.

Mechanistic numerical modelling of fracture propagation and subsequent fracture network formation can include complex physics pertaining to individual fractures such as fracture tip behaviour, fluid driven fracturing, interaction of propagating fractures with pre-existing discontinuities and other propagating fractures [Laubach et al., 2019]. Such mechanistic models can be based on extended finite element methods [such as Remij et al., 2015, Valliappan et al., 2019 etc], discrete element methods [such as Guo et al., 2017, Virgo et al., 2016 etc], and phase-field methods [such as Lepillier et al., 2020, Yoshioka and Bourdin, 2016 etc], and differ in the way rock substrate and propagating fracture are numerically treated. Such complex models are computationally intensive and are limited in the spatial extent of network evolution. Recent developments include approaches in which fracture networks genetically evolve from flaws without resorting to rigorous geomechanical treatment [such as Lavoine et al., 2020, Welch et al., 2019] but large-scale network development is still difficult to realize.

In such a context, outcrop-derived networks hold relevance. The advantage of outcrops is that they implicitly encode spatial organization of networks and network properties can be observed and sampled when outcrop quality permits. The proliferation of Unmanned Aerial Vehicle (UAV) photogrammetry has lead to an increase in both volumes and speed of acquisition of digital outcrop data [Bemis et al., 2014, Hodgetts, 2013]. Coupled with automatic image processing tools, it is now possible to obtain outcrop-derived 2D discrete fracture networks (DFNs) at large enough scales to enhance our understanding of geometrical organization and spatial heterogeneity in natural fracture networks [Palamakumbura et al., 2020].

Outcrop-based characterization of natural fractures typically involves fracture sampling methods such as the use of scanlines [for e.g. Mauldon et al., 2001, Priest and Hudson, 1981, Rohrbaugh et al., 2002], fracture traces from orthorectified fractured rock images [for e.g. Bisdom et al., 2017, Bistacchi et al., 2015], fracture traces from LIDAR [for e.g. Lamarche et al., 2011, Wilson et al., 2011], and geophysical imaging such as ground penetrating radar [Day-Lewis et al., 2017, Molron et al., 2020]. Recent advances in fracture characterization utilize data-fusion techniques in which multi-spectral, hyperspectral, gravity, and magnetic remote sensing are combined in outcrop studies. The combination of these techniques incorporates high-degrees of geological realism. In this chap-

ter, we focus on automated mapping of 2D fracture traces using UAV photogrammetry at the Lilstock outcrop, Bristol Channel, UK which exposes multiple fractured limestone beds.

The complex shearlet transform method described in Chapter 2 is used to automatically extract fracture traces from high resolution photogrammetric data published by Weismüller et al. [2020]. A critical comparison between automatic and manual tracing was presented in Weismüller et al. [2020] using topological relationships, fracture intensity, and fracture density measures, and showed that the quality of automatic tracing is consistent with the interpretations of a proficient interpreter. Weismüller et al. [2020] covered five regions of 140 sq. m each within the Lilstock pavement while Passchier et al. [2021] has mapped the different fracture generations but incompletely.

In this work, the automatic tracing is extended to an area that is 20 times larger resulting in a rich dataset that amounts to nearly 800,000 fracture segments. We develop a set of graph-based programmatic routines that convert automatically extracted fracture segments into geologically significant fracture traces. The areal extent, detail, and spatial continuity of the resultant fracture network data, comprising of around 350,000 fractures, is unique. The dataset is an important benchmark which provides a means to overcome sampling and truncation effects typically experienced with small fractured outcrops and which have so-far limited efforts at fracture characterization.

3.2. FRACTURES AS SPATIAL GRAPHS

Graph theory concerns the study of mathematical structures, *graphs,* that model pairwise relations between objects. The use of graph theory and spatial graph representations to represent fracture networks was suggested by authors such as Adler and Thovert [1999], Manzocchi [2002], Valentini et al. [2007a], Valentini et al. [2007b], Santiago et al. [2016], and Sanderson et al. [2019]. Such a graph representation maintains topological relationships between fracture segments and spatial relationships between fracture edges. Topology serves as a means to quantify connectivity of a fracture network. Network connectivity parameters have important implications for fractured hydrogeologic and subsurface modelling [Berkowitz, 2002]. In addition to topological relationships, fracture networks are also spatial in nature. In this regard, fracture networks are similar to roads and power grids in that the networks are embedded in space. Such families of spatial networks are observed to have steric constraints that impose limitations on the maximum degree of a node [Barthelemy, 2018]. Many specialized methods and techniques developed for spatial graphs can, therefore, be applied to fracture network data.

Graphs are typically represented in matrix forms, allowing the use of computational linear algebra techniques and algorithms to investigate properties of the network structure, and derive insights into spatial and spectral properties. Within the structural geology literature, such approaches are not widespread as data pipelines that can deliver sufficient volumes of fracture data in the form of spatial graphs face several challenges in data acquisition and processing. The advent of UAV-based data acquisition and automatic fracture trace extraction opens up new avenues to use graph algorithms to extract insights from large-scale fracture patterns.

From graph theory, a *graph* is a pair $G = (V, E)$ with V being a set of *vertices* and E, a set of *edges*. The abstraction that connects mathematical graph theory to fracture

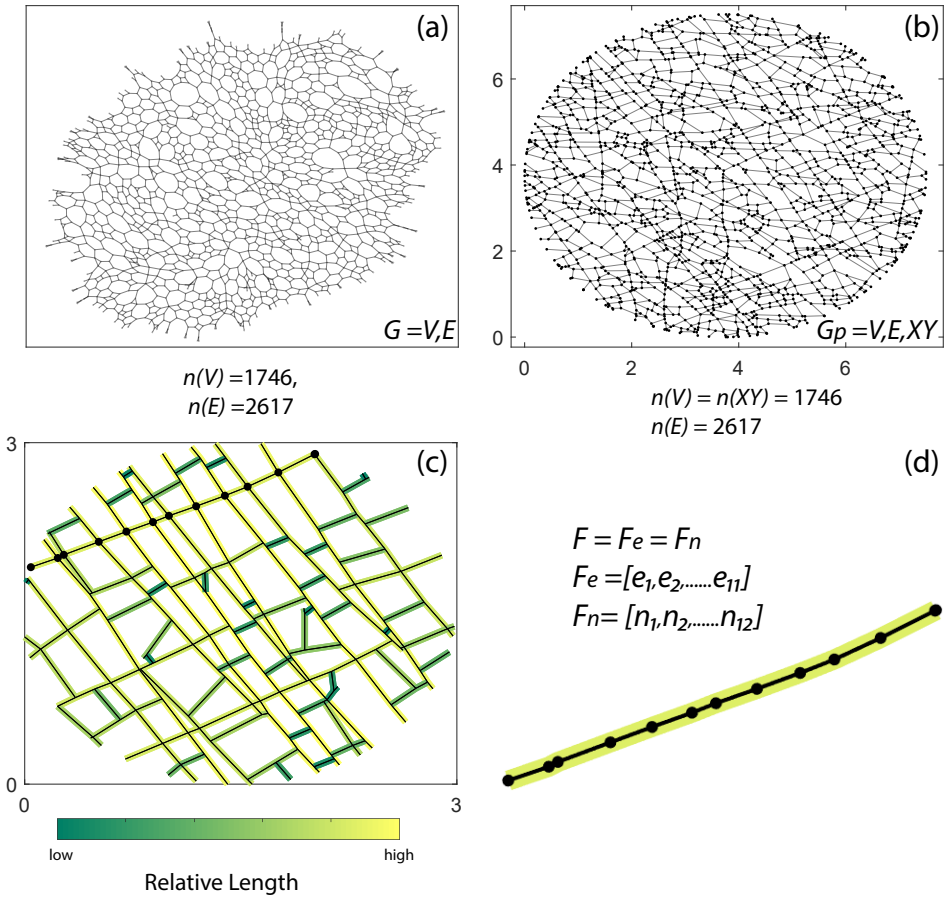

Figure 3.1: (a) A graph with no spatial positioning can be simply depicted as nodes and edges with a method of planar drawing [Nishizeki and Rahman, 2004]. Here a fracture network is converted to a graph and drawn in a "gravity" layout. (b) The fracture graph with spatial positioning applied to each of its nodes (dimensions in metres) (c) An example of a fracture network plotted as a spatial graph with individual fractures from tip-to-tip colour coded based on fracture length (dimensions in metres). One fracture is highlighted with enlarged nodes (d) enlarged view of a single fracture 'F' within a spatial graph, defined as a set of 'n' nodes or '$n-1$' edges

networks is that fracture intersections form the vertex set, V and fracture segments linking the vertex set V form the edge set, E. When a spatial positioning data structure is additionally specified to represent position of each fracture intersection in 2D cartesian space, the fracture network forms the planar graph, G_p. An example of a fracture network in the form of a graph, with edges representing fracture segments and vertices denoting fracture intersections is depicted in Fig. 3.1(a). This is a simple planar drawing where nodes are positioned in such a way that there are no edge crossings. The graph has 1746 nodes and 2617 edges. The same graph with spatial positioning is depicted in Fig. 3.1(b).

In this representation, the definition of a geological fracture 'F', is simply a subset

of 'n' nodes within the graph. This is also equivalent to a subset of '$n-1$' connected edges which are contained within the edge set that forms a walk or path within the graph (see Fig. 3.1.c and Fig. 3.1.d). The entire fracture network is a list of paths which are specific sequences of nodes (and edges). A *weighted* graph is one in which the edge set is associated with weights that can represent, for instance, the relative importance of edges within the complete edge list. In case of fracture networks, this may simply be the euclidean distance between the end nodes of the particular edge. A graph may be *directed* and referred to as a *digraph* which implies that an edge has a source node and a target node. In case of fracture networks, an *undirected* graph representation is sufficient.

The graph representation where fracture intersections form vertices and fracture segments form edges, as depicted in Fig. 3.1(b), is called the *primal* form [Barthelemy, 2018]. There is also a *dual* form of a graph in which fractures from tip-to-tip form graph nodes and interconnections between fractures form the edges [Barthelemy, 2018]. Such *dual* representations have been used by Adler and Thovert [1999], Valentini et al. [2007b], Andresen et al. [2013], and Vevatne et al. [2014] for fracture networks. To illustrate the difference between the two representations, an example network from Bisdom et al. [2017] is depicted in the primal form in Fig. 3.2(a) with fracture intersections being the vertices and fracture segments forming the edges. The dual form where fractures from tip-to-tip are nodes and intersections between fractured are edges is depicted in Fig. 3.2(b). It can be observed that the longest fracture striking NW-SE has the maximum number of intersections with smaller fractures abutting on to or crossing it. The longest fracture is therefore the node with the highest degree in the dual graph. Since the dual representation considers only topological connections between fractures from tip-to-tip, we do not associate any spatial position to the nodes in Fig. 3.2(b). Figures 3.2(c)-(d) depict adjacency matrices of the primal and dual graphs respectively. The degree of a node denotes the number of edges that intersect the particular node. The node degree distributions of the primal and dual are depicted in Fig. 3.2(e) and Fig. 3.2(f), respectively. The node degrees in the primal are subject to geometric constraints with a maximum degree of 6 (a hexa type joint). The dual graph degree distribution is more spread out with 64 being the largest degree.

By converting fracture network shapefiles to primal graphs, we can then use graph algorithms and metrics to analyze the networks. Various network metrics can be used to quantify inter- and intra-network variability in fracture networks using the graph representation. This is a novel approach in fracture network analysis in the Geosciences, made possible by the large amount of fractures. Our results form a valuable benchmark dataset for future fracture mapping and characterisation methods.

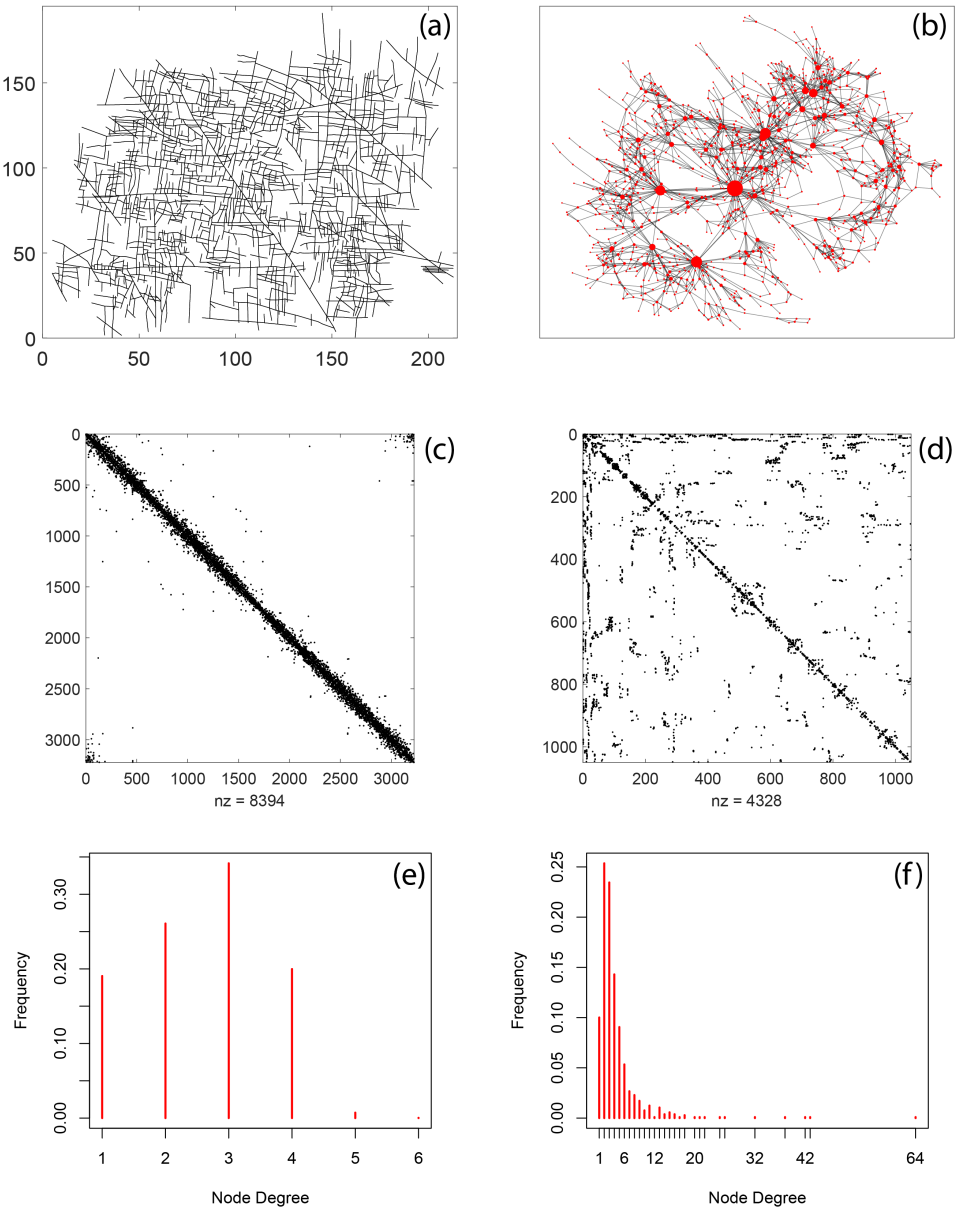

Figure 3.2: (a) Primal graph representation of a manually interpreted fracture network, Apodi-4, from Bisdom et al. [2017] in the Jandaira formation of the Potiguar Basin, Brazil having 3309 nodes and 4258 edges. Only the largest connected component of the network is depicted after removing all isolated fractures. (b) Dual graph representation of the Apodi-4 fracture network using a 'force' layout. Fracture traces from tip-to-tip are represented as graph nodes and intersections between fractures are considered as edges. The dual representations has 2172 edges and 1082 nodes. Node size is plotted proportional to the node degrees and highlights the centrality of the relatively few long fractures (c) Adjacency matrix of primal graph (d) Adjacency matrix of dual graph (e) Degree histogram representing node topology of primal graph (f) Degree histogram representing node topology of dual graph

3.3. GEOLOGY OF THE STUDY AREA

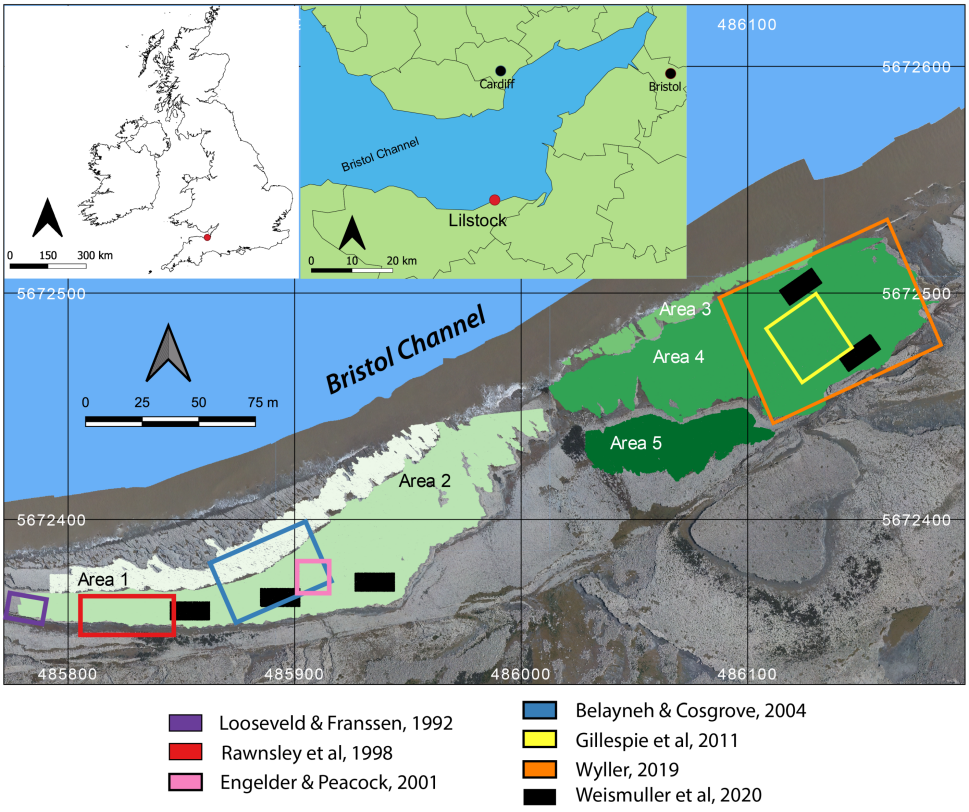

Figure 3.3: Overview of the study area located at Lilstock, Bristol Channel, UK generated from UAV photogrammetry at an altitude of 100 m. The orthomosaic is available as an open dataset [Weismüller et al., 2020]. Spatial extent of the five areas within the Bristol Channel outcrop where fracture networks are automatically extracted is depicted in sequential shades of green. Approximate areas where previous studies done within the same outcrop are also marked.

The outcrops are located off the southern coast of the Bristol Channel in West Somerset, UK, close to the hamlet of Lilstock (see Fig. 3.3). The area is within a 7.428 sq.km geological Site of Special Scientific Interest (SSSI), referred to as the Blue Anchor to Lilstock Coast SSSI, due to the exposures ranging from Hettangian to Pliensbachian within the Early Jurassic. Deformation features such as faults, fractures, and joints are exposed within the study area [Spruženiece et al., 2021]. The site is well-studied in terms of structural features, tectonic history, sedimentology, stratigraphy, and basin evolution [for e.g., Belayneh et al., 2006, Engelder and Peacock, 2001, Glen et al., 2005, Kamerling, 1979, Kelly et al., 1999, Nemčok et al., 1995, Peacock, 2004, Rawnsley et al., 1998, Van Hoorn, 1987]. The regions of interest and the focus of this contribution are three fractured limestone pavements referred to as *benches* by Loosveld and Franssen [1992].

The fractured pavements correspond to the northern limb of a single E-W trending

anticline formed due to N-S compression during the Tertiary [Dart et al., 1995]. We focus on five fractured pavements the extent of which is depicted in Fig. 3.3. The fractured regions are designated as Areas 1-5. Areas 1 & 3 and Areas 2 & 4 belong to the same stratigraphic layer. The particular areas were chosen as they are largely devoid of vegetation and weathering, and contain joints belonging to different stages in the tectonic history forming a well-connected spatial network. Additionally, the studied regions contain sub-regions which were the focus of previous work by Loosveld and Franssen [1992], Rawnsley et al. [1998], Engelder and Peacock [2001], Belayneh and Cosgrove [2004b], Belayneh [2004], and Gillespie et al. [2011a]. The relationship between joints described in the above-mentioned works is discussed by Passchier et al. [2021].

3.3.1. STRUCTURAL HISTORY

The structural history of the region may be classified into several tectonic phases. Beginning with N-S extension in the Early Jurassic to Early Cretaceous and again in the Late Cretaceous to Oligocene [Rawnsley et al., 1998], these events are evidenced by E-W striking normal faults [Brooks et al., 1988]. These extension events were followed by N-S Alpine compression during the late Oligocene to Miocene resulting in inversion of normal faults and gentle folding, followed by progressive relaxation during the Late or post-Miocene [Rawnsley et al., 1998]. Normal faults and conjugate strike slip faults indicate this event [Dart et al., 1995, Glen et al., 2005, Kelly et al., 1999, Nemčok et al., 1995]. This was followed by burial of up to 1.5 km and exhumation with features such as small folds, faults, veins, and joints [Hancock and Engelder, 1989, Rawnsley et al., 1998].

3.3.2. PREVIOUS DESCRIPTIONS OF JOINTING

The joints exposed in the Lilstock are bedding-perpendicular & largely stratabound with apertures enhanced by tide-induced dissolution, ranging from sub-millimeter at the bottom to an order of centimetres at the bed top [Gillespie et al., 2011a]. The decimeter thick limestone layers are intercalated with claystone layers of the order of $10^0 - 10^2$ cm thicknesses. A striking feature of the jointing is the network that is formed due to joints abutting or crossing each other. The presence of small displacement faults within the bench cause visibly identifiable variations in fracture patterns and intensities. The Lilstock outcrop also contains several long, fan-shaped joints that emanate from asperities on faults [Rawnsley et al., 1998]. These joint fans have also been described in other outcrops near the Bristol Channel in similar lithologies [Bourne and Willemse, 2001].

The joints are believed to be due to minor tectonic events that post-dated the stress inversion. Various authors have interpreted jointing histories and number of joint sets based on observations within sub-regions of the outcrop. Loosveld and Franssen [1992] identified six joint sets based on orientation. Rawnsley et al. [1998] identified four main joint sets using characteristics such as orientation, length, and spacing. Engelder and Peacock [2001] identified six jointing sets based on orientation and abutting criteria. Belayneh [2004] identified six joint sets based on orientation, length, and aperture. More recent work by Wyller [2019] distinguished ten jointing generations using abutting relationships, length, and orientation. These above-mentioned attempts at delineating jointing generations are limited to certain regions within the entire outcrop (see Fig. 3.3). Passchier et al. [2021] utilized the same image dataset as ours and was able to iden-

tify eight generations of joints from manually traced fractures that include all regions covered by the previous studies. The criteria used by Passchier et al. [2021] to partition individual fractures into jointing generations consisted of combination of length, orientation, and abutting criteria. The results highlighted considerable spatial variability in jointing with some regions containing just 2-3 generations while other areas achieved saturation with the maximum eight sets.

Rawnsley et al. [1998] associate the earliest joint sets as forming sub-parallel to regional Alpine compression, with subsequent jointing sets being perturbed by faults and influenced by anticlockwise shift of maximum horizontal stress during basin-wide relaxation of Alpine compression. The youngest joints were proposed to be correlated with relaxation or contracting of rock. Engelder and Peacock [2001] suggested that joint formation is linked to minor tectonic events postdating the basin inversion. The youngest joints are proposed to be correlated with the contemporary stess field [Engelder and Peacock, 2001] or due to exhumation in a late stage of the Alpine stress field [Hancock and Engelder, 1989]. Dart et al. [1995] proposed that the jointing patterns involve overprinting of joint generations.

3.4. METHODS

3.4.1. PHOTOGRAMMETRIC DATASET

The image data we consider in this work are extracted from UAV-derived orthoimagery published as a dataset [Weismüller et al., 2020]. The full dataset comprises of orthomosaics generated from UAV flights at 10 m, 20 m, 25 m, and 100 m. We utilize the orthomosaics acquired between 20-25 m flight altitude resulting in imagery of 0.9 cm/pixel. Weismüller et al. [2020] used this value of resolution to manually interpret fractures in five 140 sq.m regions within Areas 2 and 4 (see Fig. 3.3). Weismüller et al. [2020] also performed automatic interpretation in the five regions using the complex shearlet transform method described in Chapter 2 and quantitatively compared with manual interpretations. The validation of manual with respect to automatic mapping indicated closely similar fracture patterns, generating confidence in an endeavour to extend the automatic interpretation to larger regions of the outcrop over multiple layers. Passchier et al. [2021] used the same image dataset with similar resolution to identify jointing generations from manual interpretations within Areas 2 and 4.

3.4.2. AUTOMATIC TRACING WORKFLOW

We extend the workflow based on the complex-shearlet transform from Chapter 2 to the photogrammetric dataset. The image data are divided into sub-tiles of 1000 x 1000 pixels for efficient computation and considering memory requirements. The processing steps are then applied to each tile separately. This splitting of the images therefore enables processing on multiple workstations. The realized vector geometries are combined into shapefiles. The number of image tiles that correspond to each bench is summarized in Table 3.1 along with approximate areal extent.

Since quality of automatic fracture detection depends on enlarged discontinuities owing to weathering or otherwise and given that the degree of weathering is spatially variable, a single set of parameters is insufficient to efficiently extract all exposed traces.

Table 3.1: Study areas and approximate area covered

Region	Image tiles	Approx. area (sq.m)
Area 1	58	2034
Area 2	128	6017
Area 3	25	714
Area 4	107	6749
Area 5	34	1473

Therefore, three different sets of shearlet parameters are used for ridge detection yielding three different ridge image ensembles (E_1, E_2, E_3) that capture fractures both subtle and well-eroded. The three shearlet system parameters used are listed in the data supplement. Various linear combinations (a, b, c) are applied to E_1, E_2, E_3 to obtain an optimal E_{final} for each image tile as per

$$E_{final} = a.E_1 + b.E_2 + c.E_3. \tag{3.1}$$

This combined ensemble, E_{final} is then used for further image processing as per the workflow in Section 2.3.1. The traces extracted from each image tile are then merged as a single shapefile. An example of an image tile with a ridge ensemble and the corresponding vectorized shapefile is depicted in Fig. 3.4. Though the Lilstock outcrop is a high-quality exposure, there are still sources of false positives owing to erosion, water puddles, shrubbery, and rubble. These artefacts are removed manually using interactive GIS tools. The total time taken for automatic mapping for all tiles was 384 hours CPU time (using an Intel Xeon processor with 4 cores, 3.5 GHz, and 32 GB RAM). The time taken to clear the artefacts varies between 1-2 hours per image tile depending upon the image.

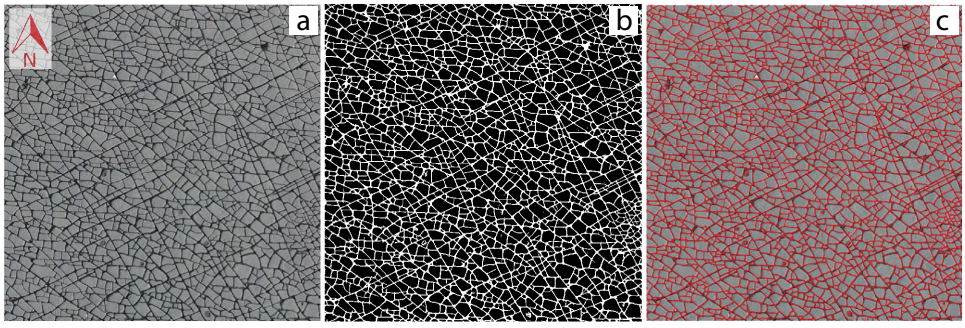

Figure 3.4: (a) An image tile (9.3 x 9.3 m) from the Bristol Channel dataset (b) computed ridge ensemble (c) the vectorized shapefile overlain on the image

3.4.3. SHAPEFILES TO GRAPHS

The automatic traces are in the form of shapefiles. We developed MATLAB routines to enable conversion of shapefiles of fracture networks into graph data structures and vice-versa. The conversion results in a primal graph, which can then be converted to a dual

graph if the sequence of primal graph edges that correspond to a complete fracture from tip-to-tip can be specified. The graph representations can then be exported in various graph formats that are readable by graph visualization software and packages such as Gephi [Bastian et al., 2009], iGraph [Csardi and Nepusz, 2006], and NetworkX [Hagberg et al., 2008].

3.4.4. MAKING GRAPH REPRESENTATIONS GEOLOGICALLY MEANINGFUL

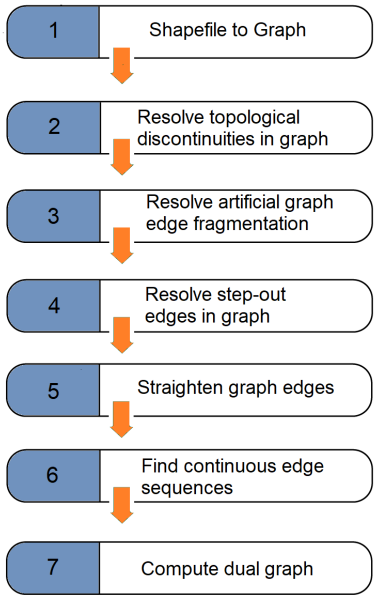

Figure 3.5: Sequence of graph manipulation routines to convert shapefiles of automatically traced fracture segments to geologically significant fracture traces and dual graph representations

The use of automatic tracing may produce fractures that deviate from a manual interpretation. When interpreting by hand, an interpreter utilizes multiple cues to trace a fracture from tip-to-tip and identify fracture tip topologies. Therefore, using ubiquitous network metrics such as cumulative length distributions, rose plots, topological summaries on automatically extracted traces can result in skewed results. To this end, we developed a series of graph manipulation routines that take the raw graph data input generated from the automatic traces into geologically meaningful data. This workflow is summarized in Fig. 3.5 and further described in the following sections. The code implementations of the functions is published as an open repository at Prabhakaran [2021a] with algorithmic description in Appendix B.

TOPOLOGICAL DISCONTINUITIES

Automatically traced interpretations can contain topological discontinuities. By analysing automatically-traced networks and comparing them with manual interpretations, we classify connectivity issues and design specific routines to resolve these discontinuities.

The three most common topological errors are depicted in Fig. 3.6. These include situations when

- a degree-1 node is in close proximity to a degree-2 node with near orthogonal angles
- a degree-3 (or Y-node) is present as three closely spaced degree-1 nodes
- two degree-2 nodes with sharp orthogonal angles are in close proximity

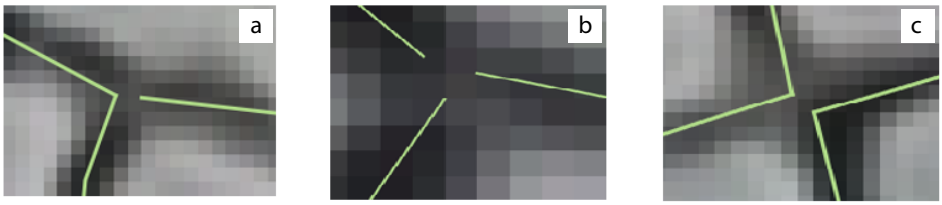

Figure 3.6: Common topological errors caused by automatic detection (a) a degree-3 connection inaccurately traced as a degree-2 node with two nearly orthogonal edges in close proximity to a degree-1 node (b) a degree-3 connection incorrectly traced as three degree-1 nodes in close proximity (c) two degree-2 nodes with nearly orthogonal edges that are disconnected

In order to resolve these topological errors in connectivity, we perform a delaunay triangulation [De Berg et al., 2000] on the fracture spatial graphs using the nodes as control points. The triangulation creates tri-elements around the fracture traces. By inspecting the histograms of tri-element areas, anomalous elements with very small areas can be isolated. These small tri-elements are formed at the regions of topological errors or with very high aspect ratios. Using a suitable cut-off area that is determined by visual inspection of the small tri-element areas, graph manipulations are performed on the graphs that resolve the loss of connectivity depending upon the node types and edge properties involved. The manipulations involve adding / removing edges and nodes and updating the fracture graph. The three types of manipulations that are done to rectify topological discontinuities are illustrated in Fig. 3.7 and the sequence of operations are summarized in Appendix B, Algorithms 2-4.

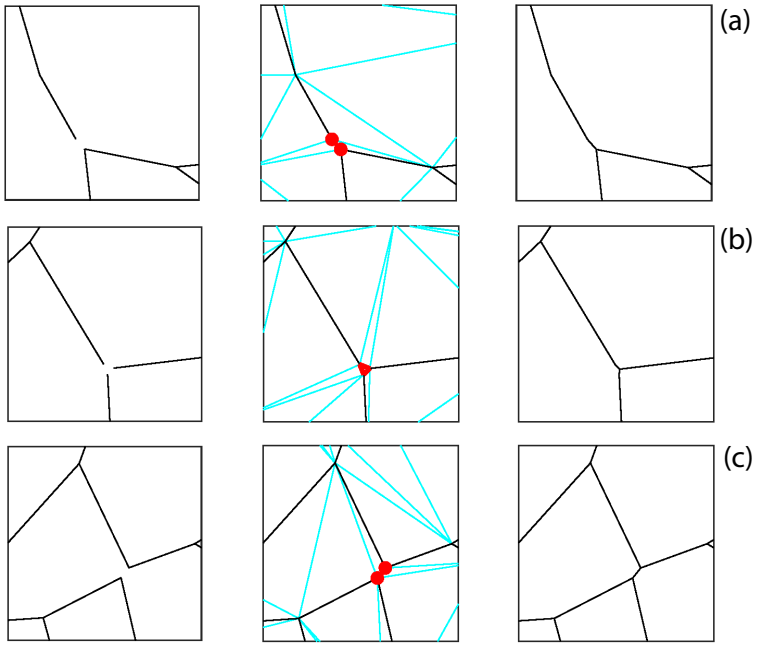

Figure 3.7: Detail of rectification of the three types of topological discontinuities using Delaunay triangulation (a) type-1 discontinuity with degree-1 node in close proximity to a sharp-angled degree-node (b) type-2 discontinuity with three degree-1 nodes in close proximity (c) type-3 discontinuity with two degree-2 nodes having sharp angles in close proximity

RESOLVING ARTIFICIAL FRAGMENTATION OF FRACTURE SEGMENTS

Artificial fragmentation of fracture trace happens when traces appear to be connected and topologically correct to visual inspection but split and saved separately within the shapefile attribute tables. This kind of situation can happen due to tile-wise image processing where fracture polylines that are otherwise continuous, are fragmented and saved as a cascade of isolated segments. Other reasons are due to the way polylines are fitted to skeletonized, binary pixel clusters as per the workflow in Fig. 2.1. The skeletonization procedure specifies branch points between intersecting fractures. However, due to varying ridge thickness within the image, it is sometimes possible that segments are connected but incorrectly labelled from a geological perspective. Such a situation is depicted in Fig. 3.8(a).

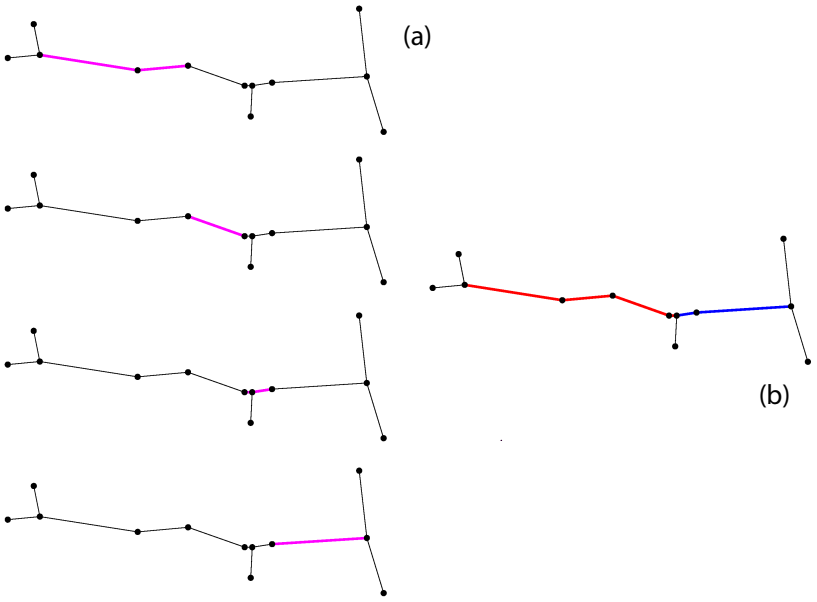

Figure 3.8: Resolving artificial fragmentation (a) an example of an artificially segmented fracture is shown which is saved as four polyline entries within the shapefile. These are highlighted in magenta. The first segment (top) is of topology type Y-V-V (where V used to denote a degree-2 node and Y a degree-3 node), second is a V-V segment, third is a V-Y-V segment, and the last one at the bottom is a V-Y segment. (b) The graph edge linking converts the fragmented four segments into two segments which are both of Y-Y topology type. The routine does both merge and split operations to ensure that there are no attribute table entries in the shapefile that begin or terminate in degree-2 nodes

In order to be geologically consistent, the visually continuous but disconnected segments have to be combined into a single polyline entity. We develop a graph edge linking function that first identifies all degree-2 nodes within the graph. For these nodes, node neighbours with degree 2 are identified and appended into a preliminary node path. The end nodes of the node path are queried again for further neighbour nodes having degree-2 and repeated till there are no more such nodes in either direction of the node path. The resulting node path is now a single connected polyline representing a fracture segment. Algorithm 5 in Appendix B summarizes the procedure. The effect of the edge linking is depicted in Fig. 3.8(b).

RESOLVING STEP-OUTS

Automatically identifying fracture edges that belong to a single, continuous fracture from tip-to-tip is a task that can face complications due to the presence of step-outs or edges that have degree-3 (or Y-nodes) on either ends. Such Y-Y motifs often form *step-outs* which impede continuous path finding as they may strike in a different direction as that of longer adjacent edges. They turn out to be bottlenecks when we seek to identify long and continuous paths using segment strike as a search attribute. Examples of such step-out edges are shown in Figs. 3.9(a) and 3.9(c). To resolve the issue, we specifically filter for graph edges that are below a certain length threshold that have a degree of 3 on both

start and terminating ends. Below a certain length threshold corresponding to the resolution of the image, a *merge* operation can be carried out deleting the step-out and creating a degree-4 node (see Fig. 3.9.b) after adding three edges and removing one node. In our case we found that a value of around 1-2 times the image resolution was sufficient for the merge operations.

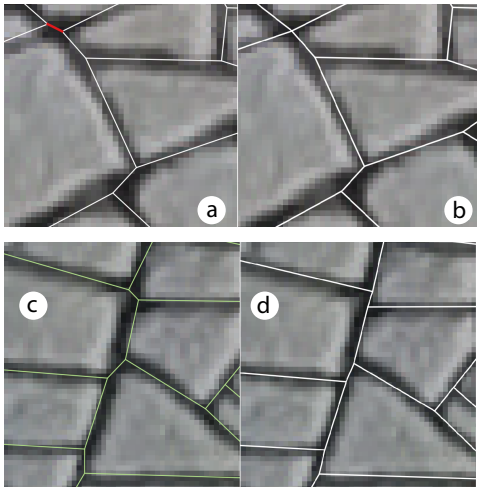

Figure 3.9: An example of automatically resolving a stepout by a merge operation (a) stepout Y-Y segment depicted in red (b) Y-Y segment removed and edges merged to form an X node. An example of automatically resolving a stepout by a flatten operation from Area 4 (c) stepout segments with varying strike that can cause loss in continuity when parsing for possible walks (d) stepout segments flattened

Above this length threshold, it is likely that the topology at either end of the step-out is correct, but the Y-Y edge needs to be *flattened* to correspond with the strike angle of one pair of edges on either side (see Fig. 3.9.d). In this case, merging of the step-out may incorrectly displace some edges of the spatial graph. In this procedure, the edges that are connected to the start and terminating nodes of each step-out are identified. A walk is identified for each of these edges. Though the step-out is a geometric feature that impedes the possibility of a walk, there are still possibilities of walks looking upstream on both directions away from the step-out. A decision is made as to which direction alongside the step-out provides the best increase in walkability based on length of walk. Once this is identified, the node of the step-out that causes the bottleneck is moved to a more preferable alignment. The sequence of graph manipulations involved in this flattening operation consists of adding three edges, removing three edges, adding a node, and removing one node (see Algorithm 6 in Appendix B). The step-out flattening procedure therefore improves the walkability in one direction.

STRAIGHTENING FRACTURE SEGMENTS

During piecewise polyline fitting as performed when vectorizing fracture traces (see Fig. 3.10.a), a large number of points are inserted to represent the natural sinuosity of fracture traces. Within the graph representation these points are degree-2 nodes and are

the predominant topology type. Although these nodes provide useful local information, in terms of overall macro-network topology, it maybe useful to *straighten* or *flatten* the graph edges by removing these degree-2 nodes and replacing them by single edges between the non-degree 2 nodes. This type of graph manipulation involves removal of all edges that either start or end in degree-2 nodes (or both) and addition of single edges between the non-degree 2 nodes (see Algorithm 7 in Appendix B). The effect of such an edge straightening operation is depicted in Fig. 3.10(b).

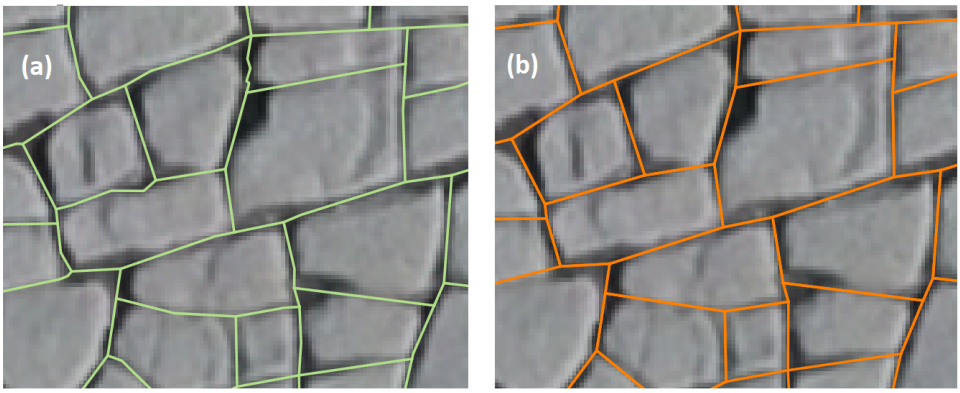

Figure 3.10: An example of straightening of fracture segments (a) original fracture network with piece-wise linear segments and degree-2 nodes (b) fracture segments which are straightened removing the degree-2 nodes

FROM FRACTURE TRACES TO GEOLOGICALLY SIGNIFICANT FRACTURES

The geological identification of a fracture in the outcrop or from image data is that of a discontinuity feature that is geometrically continuous with the tip extremities either abutting another fracture, cutting across another fracture, or terminating within rock matrix. In a typical manual interpretation using UAV-derived imagery, the interpreter draws polylines in a digitizing software (eg. Adobe Illustrator, Coreldraw, QGIS, ArcGIS etc) tracing across image pixels that seemingly correspond to a perceived fracture using visual cues within the image coupled with specific knowledge of the particular outcrop and general training in structural geology. There are many ways in which such a GIS-derived interpretation may be biased and lacking repeatability as discussed in Andrews et al. [2019] and Peacock et al. [2019]. Given these considerations, it is useful to have an automated method of obtaining geologically significant fractures (or fracture sets) rather than just fracture segments. A simple way to assign segments to individual sets is to sort based on striking angles as is done in popular tools such as FracPaQ [Healy et al., 2017], and NetworkGT [Nyberg et al., 2018]; however, this may be difficult when fractures are very sinuous. Additionally as pointed out by Peacock et al. [2018] and Andrews et al. [2020], orientation sets need not always equate to an age sets and there are drawbacks involved in only considering strike.

The graph representation of a fracture network is complete when we have list of nodes, spatial positioning data corresponding to each node, a list of edges with start and terminating points indexed as per node numberings, and a list of edge sequences to

represent each fracture. Automatic tracing cannot yield the edge sequences so that they represent sets of fractures (tip-to-tip). To this end, a function is developed to automatically identify continuous paths along graph edges based on twin rules of connectedness and small strike variation. The routine considers each edge individually and checks if adjacent edges fall within the threshold of edge strike, on either ends of the edge. Sequences of edges (or walks) are assigned as fractures. The routine is described in Algorithm 8 in Appendix B. An example of a continuous and sinuous fracture automatically combined from graph segments are shown in Fig. 3.11.

Figure 3.11: Continuous and sinuous fracture from Area 4 automatically joined from graph segments with a strike threshold of 20 degrees. Note that the strike of the start and end segment of the fracture vary by more than 50 degrees

In a related publication based on the same dataset as ours, Passchier et al. [2021] manually interpret and classify continuous edges as belonging to a single generation. We have compared the results of the automated function described in this section to the manually assigned joint generations of Passchier et al. [2021] and there is generally a good agreement.

COMPUTING DUAL GRAPHS

A dual graph can be computed from a primal graph if the edges sequences corresponding to individual fractures (tip-to-tip) are known or is computed using function described in Section 3.4.4. The dual graph depicted in Fig. 3.2(b), was computed from a shapefile in which fracture id's of manually interpreted fractures were already been listed. Given the edge sequence information, obtained either from manual interpretation or automati-

cally, the procedure to compute the dual is by initializing an adjacency matrix whose size is equal to number of fractures (A_{adj} is an $n \times n$ matrix where 'n' is the number of tip-to-tip fractures). By parsing through the intersections made by each fracture with others, the sparse adjacency matrix is then built up by filling in rows and columns corresponding to fracture intersection. The routine is described in Appendix B, Algorithm 9.

3.5. RESULTS

The methods in Section 3.4 are applied to image tiles corresponding to the five selected areas and based on these we generate five large networks. The created fracture data are in the form of spatial graphs and shapefiles. A summary of the number of nodes, edges, and tip-to-tip fractures (or walks) for each area is tabulated in Table 3.2. Edge/node and edge/walk ratios are also shown as they give an indication as to the connectedness of the networks. In order to illustrate the level of detail within the generated network data, zoomed cut-out regions from Area 2 (see Figs. 3.12.a-c) and Area 4 (see Figs. 3.12.e-f) are depicted. From the cut-outs of Area 2 in Figs. 3.12(a)-(c), there are clear visual differences in fracturing even though the orientations of fractures are quite consistent among all three samplings. This is however, not the case in the cut-outs from Area 4 shown in Figs. 3.12(d)-(f). In Fig. 3.12(e), a radial NW-SE trending fracture pattern that is orthogonally cut by NE-SW fractures can be observed. The fracturing style is very different in Fig. 3.12(d) with no radial fracturing, higher fracture intensity, and smaller bounded blocks. In Fig. 3.12(f), the fracturing intensity is highest with even smaller bounded blocks.

Table 3.2: Summary of primal graph structure

Region	**Edges** (e)	**Nodes** (n)	e/n	**Walks** (w)	e/w	**Polygons**
Area 1	42301	30299	1.39	18078	2.34	11992
Area 2	364703	228661	1.59	123592	2.95	136053
Area 3	40243	26372	1.52	16900	2.38	13874
Area 4	365333	235089	1.55	141344	2.58	129690
Area 5	78151	49771	1.57	28892	2.7	27220

3.5.1. LENGTH DISTRIBUTIONS AND FRACTURE SET DIRECTIONS

Trace length distributions corresponding to the five areas are depicted in Fig. 3.13. Trace length distributions show the lengths from fracture tip-to-tip. These are affected by boundaries of the sampled regions which may be observed by comparing the plots of largest areas, 2 and 4, with the other three. In Fig. 3.14(a) and Fig. 3.14(b) we depict fractures plotted by their length classified into three logarithmic bins for Areas 1 & 3 which are stratigraphically the same layer. Similarly, the length-binned fractures are depicted for Areas 2, 4 & 5 in Fig. 3.14(c), Fig. 3.14(d), and Fig. 3.14(e) respectively.

The rose plots depicted in Fig. 3.13 are computed from strike data that is a length-weighted average of the strike of edges that sum up to a tip-to-tip fracture. The rose plots highlight differences in fracture orientation between the layers. Orientation of the fractures do not vary significantly in Areas 1 & 3. However, Areas 2 & 4 from the same

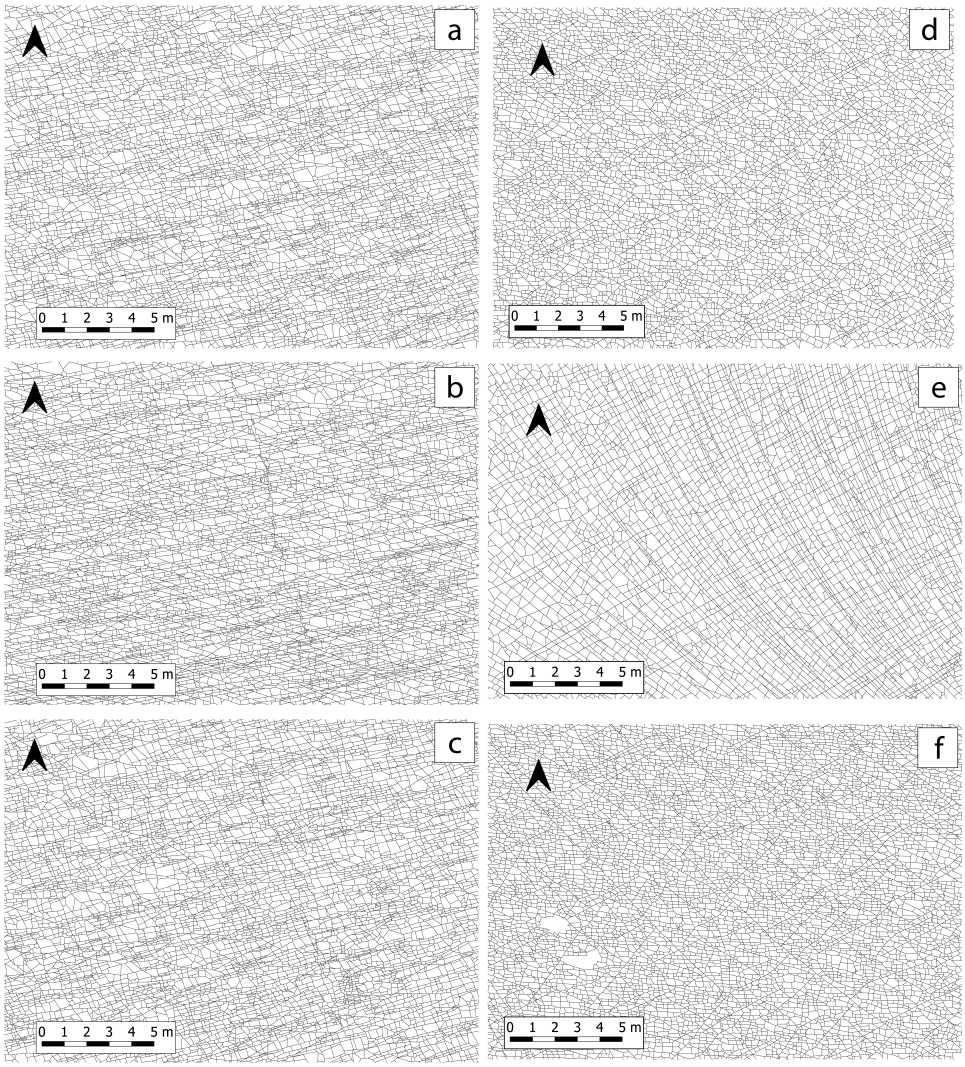

Figure 3.12: Samples of fracture networks from a single stratigraphic layer across Area 2 and 4 highlighting the differences in fracture network organization. Samples (a), (b), and (c) are from Area 2 and (d), (e), and (f) are from Area 4

stratigraphic layer have considerably different fracture orientations. This is illustrated in Fig. 3.14(d) with Area 4 containing curved and radial fractures. However, Area 2 does not have any curved fractures (see Fig. 3.14.c). Similar to Area 4, Area 5 also has curved fractures as can be seen in Fig. 3.14(e). The scatter in rose-plots corresponding to Areas 4 & 5 is related to the presence of the curved joints.

From Fig. 3.14(c), Fig. 3.14(d), and Fig. 3.14(e), spatial variations in the distribution of fractures in Areas 2, 4, and 5 can be observed. The longest joints in Area 2 display a

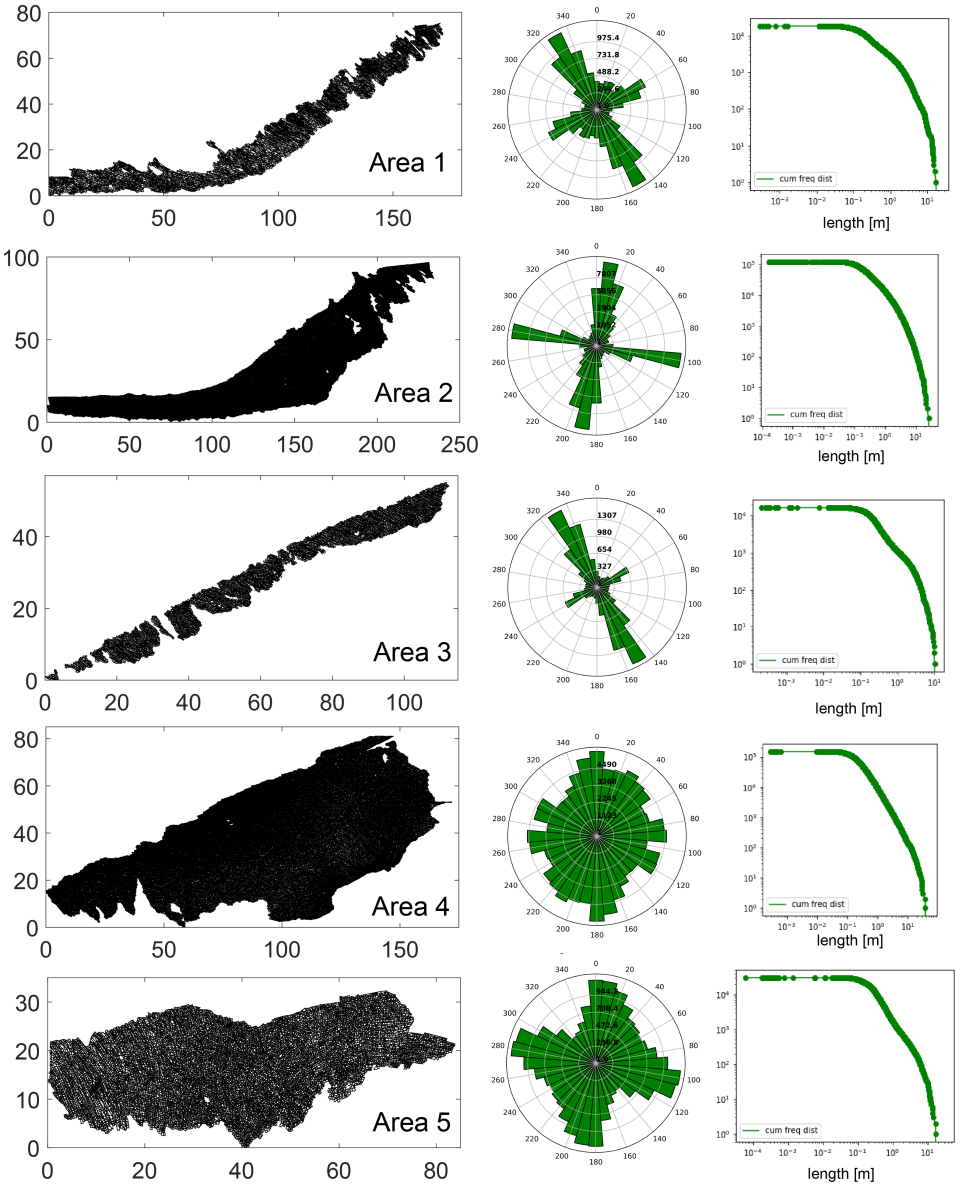

Figure 3.13: Fracture network trace maps for all areas with corresponding rose plots and cumulative trace length distributions

spatial variation with a larger concentration to the SW (see Fig. 3.14.c). In case of Area 4, the radial and curved fractures which are also the longest are located in the western part of Area 4 (see Fig. 3.14.d). The occurrence of these long, radial joints diminishes to the east of Area 4. In the case of Area 5, the long fractures has strikingly different curvature

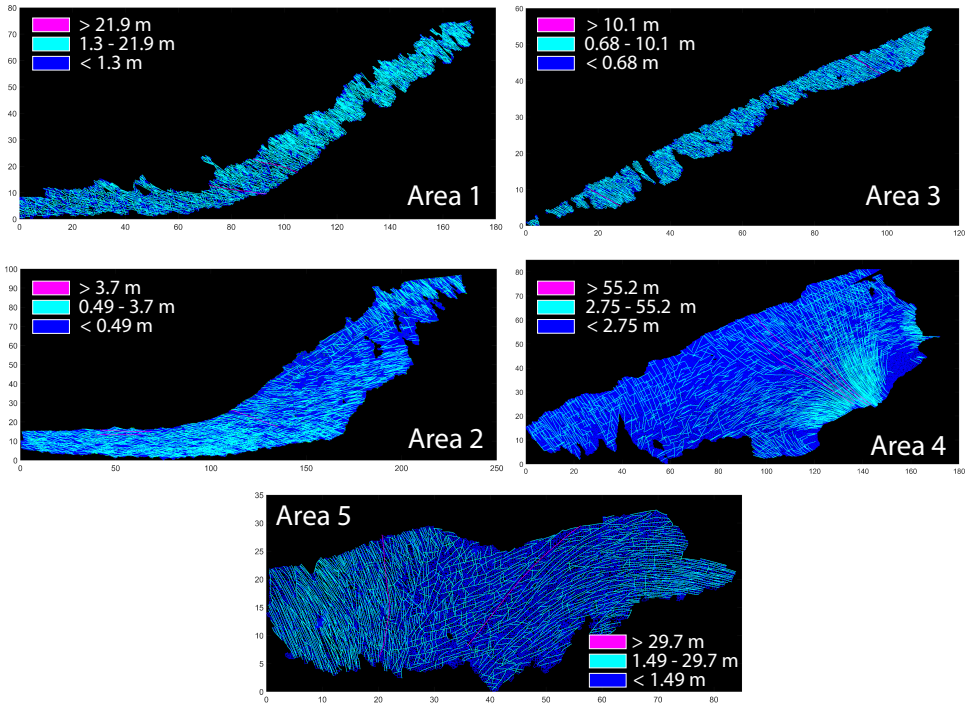

Figure 3.14: Plotting fractures by logarithmically spaced length bins

directions towards its east compared to its west (see Fig. 3.14.e).

3.5.2. NETWORK TOPOLOGICAL SUMMARY

From Manzocchi [2002], Sanderson and Nixon [2015], and others, an I-node corresponds to a fracture tip that is isolated, a Y-node is analogous to fracture tip that has abutting interactions with other fractures (or splaying fractures), and an X-node represents an intersection where one fracture crosses another fracture. The proportions of each node type can be summarized in an I-Y-X ternary diagram. To quantify network topology, we use node degree histograms instead of I-Y-X ternary plots. This is because of the need to depict node degrees greater than four which are not unusual in large-scale networks as is observed in the Lilstock pavement. Additionally, in the case of dual graph representations, where fractures are represented as nodes, the node degree can be larger [Valentini et al., 2007a]. The node degree distribution of the primal graphs corresponding to the five networks is depicted in Fig. 3.15. The node degree distribution of the dual graphs corresponding to the five networks is depicted in Fig. 3.16(a)-(e). Degree distributions of all the primal graphs indicate that the predominant node topology are Y-nodes with a 70-80 % contribution followed by X-nodes. A summary of primal graph edge types based on topology is shown in Table 3.3.

The dual graph degree distributions provide insight into the connectivity behaviour of each network. The topological summary of the dual graphs are tabulated in Table 3.4.

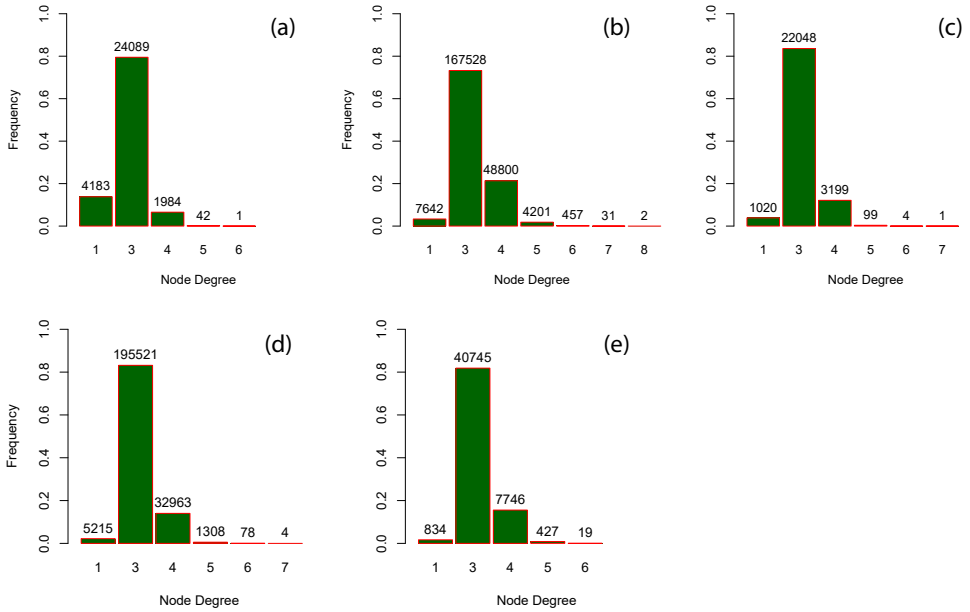

Figure 3.15: Degree distributions for the primal graphs with number of nodes corresponding to each topology type (a) Area 1 (b) Area 2 (c) Area 3 (d) Area 4 (e) Area 5

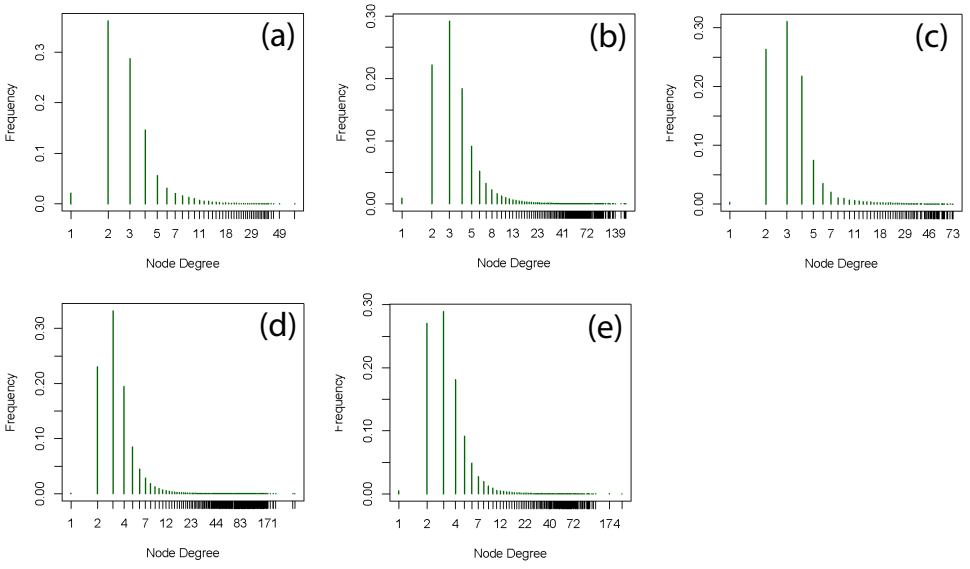

Figure 3.16: Degree distributions for the dual graphs (a) Area 1 (b) Area 2 (c) Area 3 (d) Area 4 (e) Area 5

The node degree value indicates the number of connections that a fracture makes with other fractures within a network. Maximum node degrees in dual graphs are observed

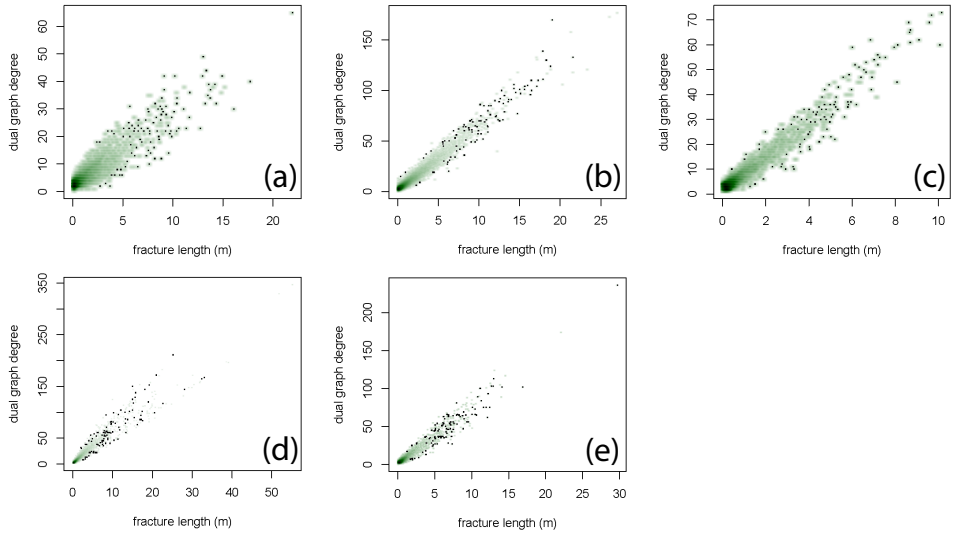

Figure 3.17: Correlation between dual graph degree and trace length (a) Area 1 (b) Area 2 (c) Area 3 (d) Area 4 (e) Area 5

Table 3.3: Summary of primal graph edges based on topology

Edge type	Area 1	Area 2	Area 3	Area 4	Area 5
1-1		4			
1-3	4041	7048	1007	5127	783
1-4	139	552	12	87	43
1-5	3	27	1		8
1-6		7		1	
3-3	30612	176360	27186	238130	47983
3-4	6815	127218	10355	99922	23793
3-5	182	13740	386	4902	1610
3-6	5	1708	18	329	83
3-7		141	6	23	
3-8		9			
4-4	478	30074	1161	15094	3327
4-5	25	6328	100	1522	465
4-6	1	884	6	129	29
4-7		63	1	4	
4-8		7			
5-5		392	4	53	25
5-6		115		9	2
5-7		11		1	
6-6		13			
6-7		2			
Total	42301	364703	40243	365333	78151

from Areas 4 and 5 which contain continuous and long, radial fractures. The correlation between dual graph degree (number of intersections made a fracture) and the fracture length is also plotted in Fig. 3.17 depicting a positive correlation between fracture length

and number of intersections. The number of connections is least in Areas 1 and 3. This is possibly an effect of sample size as these regions are the smallest and their spatial extent in the N-W direction is quite thin. Area 2, despite covering more area than Area 5, has a lesser maximum dual degree.

Table 3.4: Summary of dual graph structure

Region	Nodes (n)	Edges (e)	e/n	Max degree
Area 1	18078	34077	1.88	65
Area 2	124006	301077	2.42	177
Area 3	16900	36320	2.14	73
Area 4	141344	314537	5.27	347
Area 5	28892	65867	2.28	236

3.5.3. BOUNDED AREA DISTRIBUTION

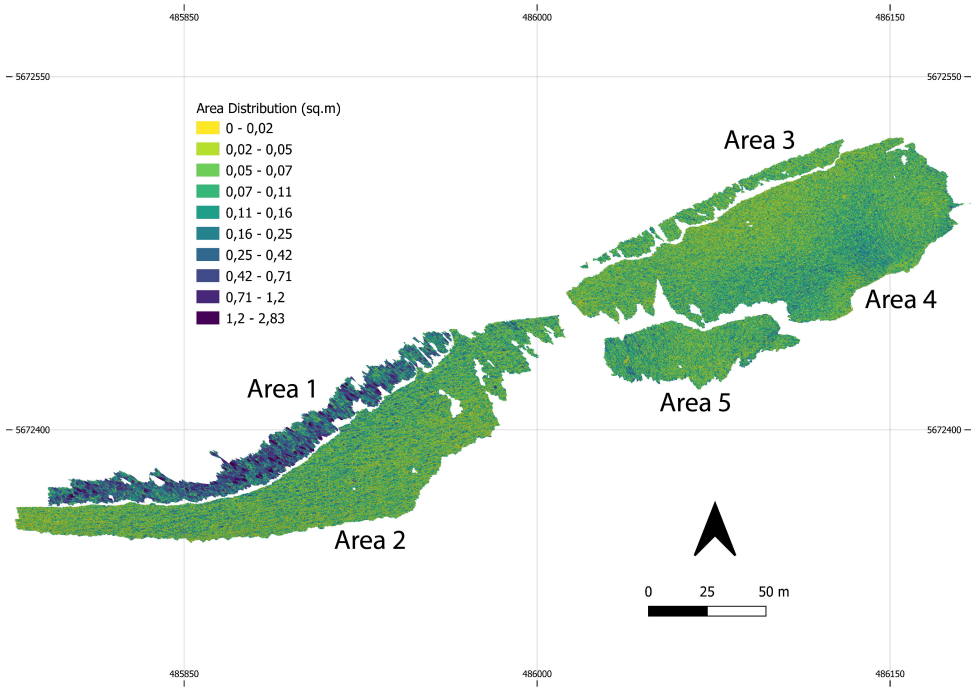

Figure 3.18: Spatial distribution of polygonal regions highlighting the variation in fracturing across different areas

The fracture patterns develop and enclose bounded regions of unfractured rocks. These enclosed polygonal areas are extracted from the spatial graphs by identifying the primary cycles that are created by edges. The spatial distribution of areas corresponding to these polygonal regions is depicted in Fig. 3.18 as a chloropleth and depicts the variation across the layers. Histograms of the area distributions of each layer is depicted in

Fig. 3.19. Area 1 appears to have the largest block areas, followed by similar distributions for Areas 3 and 5. The largest Areas 2 and 4 have smaller block areas with visibly more intensive fracturing.

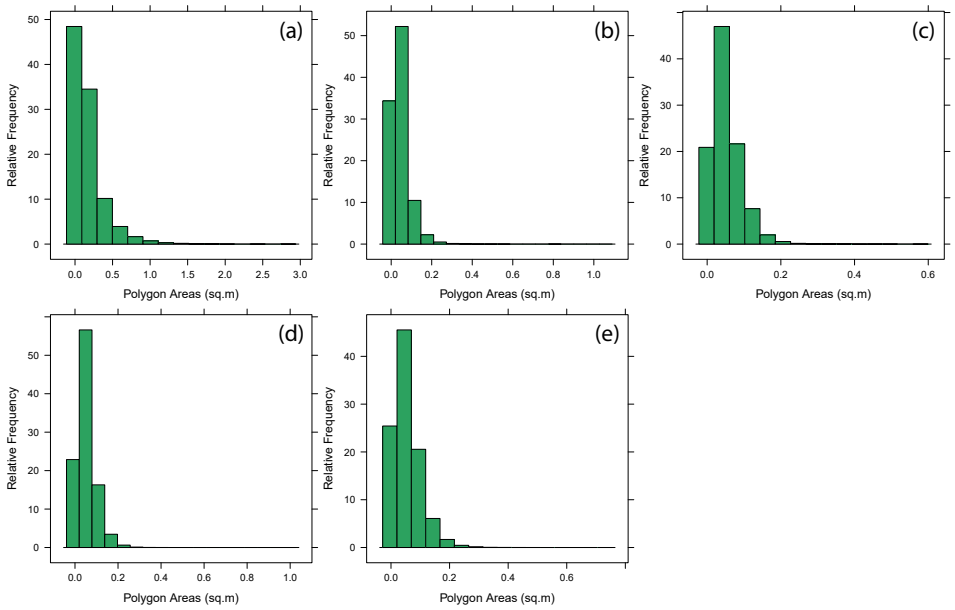

Figure 3.19: Bounded area distributions with relative frequency in percentages (a) Area 1 (b) Area 2 (c) Area 3 (d) Area 4 (e) Area 5

3.5.4. SPATIAL P_{20} AND P_{21}

Fracture persistence measures (P_{ij}) formulated by Dershowitz and Herda [1992] are used to investigate the spatial differences in fracturing. The fracture intensity, P_{21} and fracture density P_{20} metrics are computed using the box-counting method by overlaying the networks with a cartesian grid of box size of 2.5 x 2.5m. Fracture intensity (m/m^2) involves computing 2D trace length per area for each grid box. This is depicted for all areas in Fig. 3.20(a). Fracture density (m^{-2}) computes the number of segments within each grid box and this is depicted in Fig. 3.20(b). The persistence results reveals regions within the outcrop with different fracturing motifs. Area 1 has the least fracturing intensity and density which is uniform in the spatial distribution. Area 3 also is homogenous in the type of networks present. The greatest variation is in Area 4 which has clear regions of low and high P_{21} and P_{20} with a demarcable boundary. Area 2 has the greatest fracturing intensity among all regions especially concentrated in the south and western parts of Area 2. Similar intense fracturing regions can also be seen in the north-western peripheral parts of Area 4. These are not fracture corridors but progressively intense fracturing with smaller block areas.

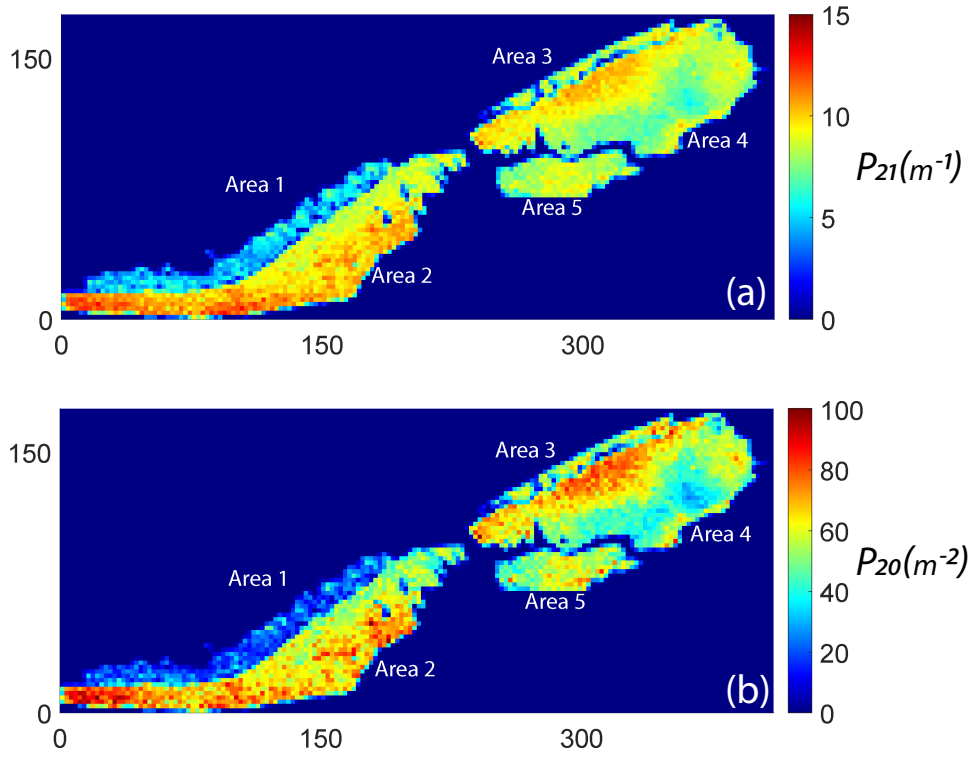

Figure 3.20: (a) Fracture intensity, P_{21} (m/m^2) for all areas (b) Fracture density, P_{20} (m^{-2}) for all areas

3.6. DISCUSSION

Manually tracing fracture networks from image data is time-consuming and can introduce various types of biases depending upon skill, style, and perseverance of the interpreter. These challenges are evident from the observed networks in the structural geology literature which are not large and not continuous enough to study spatial network heterogeneity or do not have sufficient resolution to correctly identify topology. Automatic tracing affords rapid and reproducible network results which can be applied to large image datasets. In case of the Lilstock pavement, high image resolution, enlarged apertures due to erosion, high contrast in imagery between the wet apertures and dry surface, and lack of vegetation, aided in easily applying automatic mapping. One major drawback associated with automatic interpretations which precludes direct usability by a structural geologist and which were evident from the results in Chapter 2 is that the detected segments were not yet organized into geologically meaningful, tip-to-tip fractures.

The treatment of fracture networks as graph data structures with spatial positioning allows us to perform various sequences of graph manipulations to rectify these issues and convert the data into geologically realistic fractures. The combined use of automatic tracing and application of such specific routines have resulted in a spectacular, large-

scale fracture network dataset with unprecedented spatial coverage and resolution. The network data is of great relevance as it can be used to obtain valuable insights into spatial arrangements of fracture networks and network morphogenesis. In this section, we delve into possible reasons for the observed spatial variations in network geomorphology. Issues regarding the applicability of automatic mapping and how large-scale network data can be leveraged are also considered.

3.6.1. SPATIAL HETEROGENEITY

One of the interesting results of our fracture maps is the layeral differences in patterns. Areas 1 and 3 have relatively less spatial variation as can be quantified from spatial plots of fracturing intensity, density, and polygonal areas (see Fig. 3.20.a-b and Fig. 3.18). They are also the smallest regions with long and thin strips of exposed rock. Area 1 corresponds to regions with the least fracture intensity and density, and highest bounded areas. The most spatially extensive layer, comprising of Area 2 and 4 depict the most striking variations. From previous work by Gillespie et al., 2011a, Hancock and Engelder, 1989, Rawnsley et al., 1998 and many others, the long radial, fan-like fracture sets are hydraulically-driven and originate from stress concentrations on the small fault. This region in the SE of Area 4 also has the least fracturing intensity with wide spacing between the radial fractures. The interference of small low-displacement faults can also be seen in the NE region of Area 2 which again has a low-fracture intensity. Similar to Area 4, Area 5 also contains highly sinuous fractures that can be linked to the NE trending regional fault. In Area 5, the long, radial fractures have strikingly different curvature directions towards its east as compared to its west (see Fig. 3.14.e). These effects totally disappear in Areas 1,2, and 3 which have mostly straight fractures. Within Area 2, a trend of high fracturing intensity can be observed towards the SW which progressively decreases towards the NE. Area 5 has the largest fracturing intensity in its centre and this progressively decreases to its east-west peripheries. Passchier et al. [2021] highlighted spatial variations in the presence of joints in the regions covered by Areas 2 and 4. From a total of eight identified jointing generations, only two are distributed evenly across both areas. Three sets of joints exclusively appear in Area 2 but are absent in Area 4. Another three sets are found in both Areas 2 and 4, but they are restricted to certain localized regions. The spatial variation of the polygonal area distributions (Fig. 3.18) follows a similar trend as the fracture persistence plots (Fig. 3.20a-b). The spatial variation in block areas is likely to depend upon the thickness variation of limestone and underlying shale layers [Belayneh and Cosgrove, 2004b].

The reasons behind spatial variation may also originate from factors not observable from simple photogrammetric data. For example, differences in fracturing may emanate from local variations in layer thickness, due to changes in mineralogical composition of the host-rock, and orientation/strengths of pre-existing structures. Our image resolution does not include vein or stylolite networks which are also present in the outcrop and whose spatial variation may have an influence on the development and of the barren fracture networks that we have mapped. Spatial layer thickness can be estimated by methods such as ground penetrating radar (GPR) and mineralogical variation can be explored using UAV-based sensors such as magnetic and hyperspectral imaging. Therefore, it is suggested that further investigations into spatial variation in fracture networks

should incorporate data fusion from a variety of UAV-mounted sensors, coupled with field-level and ground truthing of features that are below remote sensing spatial resolutions.

3.6.2. FROM TRACES TO TIMING

Previous work on the Bristol Channel summarized in Section 3.3 focussed on relationship between structural history of the region, exposed fractures, and other large deformation features. Identifying fracture generations and sequences of network evolution is routinely done based on geometric criteria and topological relationships of fracture tips, sometimes supported by geochemical analysis of cement within fractures. The problem of identifying fracture timing from the automatically traced fractures was not in the scope of this contribution. Using the same dataset as we have used, Passchier et al. [2021] identified eight generations of fractures traced segments without resorting to a fully detailed network interpretation. The oldest generations were considered to be the most continuous and longest which do not abut against others. Subsequent generations were then identified based on strike and abutting criteria w.r.t each older joints generation. In their study, a correlation between length and age seemed probable with only few exceptions. In the same work, there are also highlighted cases where sequential rule-based joint identification results in *Escherian* paradoxes.

Another study by Wyller [2019] focussed on an area that roughly conforms to the western parts of Area 4 and was able to identify ten sets of joints using statistical analysis of joint lengths, orientations, and topology. In this study as well, assigning hierarchies based on abutting relations result in paradoxes which Procter and Sanderson [2018] and Wyller [2019] refer to as *backcycling* between joint generations.

The above studies are based on the assumption that abutting and cross-cutting relationships are a sufficient criteria, if not necessary, to be able to delineate fracture sets into a hierarchy of fracturing episodes. Such approaches may not always suffice, for instance, if fracturing drivers are due to high-deformation episodes or if there is evidence of complex structural inheritance. In outcrops such as the Lilstock pavement, where fractures are mostly formed in low-deformation settings, simple geometric criteria as proposed by Passchier et al. [2021] may be programmed to automatically assign fractures into hierarchical episodes. Given large networks and well-defined criteria, if might be more prudent to use statistical strategies such as Markov chains to automatically assign generations [Snyder and Waldron, 2018]. The combinatorial approach of Potts and Reddy [2000] in the form of younging tables to identify deformation histories may also be applied in an automated manner. In future work, such automated approaches may be applied to the full-detailed fracture networks presented in this chapter to compare automatically-assigned generations to those that have been manually-assigned in previous literature relevant to the Lilstock pavement.

3.6.3. EXTENT OF APPLICABILITY OF AUTOMATIC METHODS

We have been able to extract a very large number of geologically relevant fracture traces focussing only on the opening-mode fractures that are visible from a flying altitude of 20-25 m. The quality of the interpretations are comparable to the work of a manual interpreter and this is attained in much less time [Weismüller et al., 2020]. Often, the

error in automatic tracing results are within the limits of subjectivity associated with even a well-trained interpreter. The largest variation in interpretation between manual and automatic is the creation of stepped-out segments. This is due to the fact that unlike manual interpretation where the interpreter can make a decision on a possible fracture intersection considering the full outcrop image, automatic methods make use of local information in the image which leads to uncertainty in regions which are more eroded than normal. The presence of step-outs sections was observed by Weismüller et al. [2020] when comparing topological differences between the two approaches and revealed that manual interpretations result in topological distributions skewed to higher node degrees. From a network connectivity point-of-view, such a configuration may be correct but this can result in shorter length distributions. Despite the fact that multiple interpreters may differ in the choice of a fracture tip start and end (as shown by Andrews et al., 2019 in experiments with multiple participants), the issue of curtailed length distributions are less likely to arise in manual tracing as interpreters use multiple global cues available within an image to decide the continuity of a trace. We addressed these issues using the step-out fixing functions. The methods developed here are extendable to other photogrammetric datasets.

3.6.4. EXTENSION OF OUTCROP FRACTURE NETWORK DATA

In subsurface applications, geomodelers often have to contend with sparse borehole fracture data as the only available ground-truth. Since geophysical imaging resolution are often too coarse to resolve subsurface fractures, outcropping fractures have long been considered as analogues to guide subsurface discrete fracture network models. In a typical subsurface situation, it is required to be able to extrapolate away and interpolate between points of well control where fracture data exists in the form of cores, formation micro-images (FMI), and resistive/acoustic logging. This is a highly ill-posed problem as the naturally heterogeneous behaviour of fracture patterns are typically underrepresented. This is due to inherent sampling bias within each well data point and well as uncertainty in relationship between large-scale geological drivers.

The commonly used methods for subsurface fracture network modelling are based on stochastic point processes that use 1D well data input such as fracture size, type, intensity, number of sets, and cumulative length distributions [Thovert et al., 2017]. Stochastically generated DFNs that utilize such sparse data to extrapolate, are often limited in their ability to represent fracture clustering effects, spatial variations in fracture orientation, and topological connections, all of which are observable in natural fracture networks.

Alternative methods to stochastic point process-based methods such as the semivariogram approach of Hanke et al. [2018] applied to areal fracture intensity and fracture intersection density maps, and the multipoint statistics approach of Bruna et al. [2019a,b] which use training images of user-defined outcrops can help in incorporating more geologically-realistic fracture networks into geological models.

In this respect, one needs to assess the fracture network properties to be replicated and for which 2D fracture trace maps can provide additional value. From our analysis of the large-scale Lilstock fracture networks, we would suggest that DFN generating methods should also be able to replicate bounded area distributions. This may be justi-

fied by the fact that fracture networks influence effective rock permeability also through time-dependent diffusive effects from the matrix. Since matrix block area distributions contributes to the matrix-fracture fluid exchange, it needs to be represented as a parameter. A second useful parameter that arises from 2D trace maps is the correlation between fracture length and number of intersections. From our analysis of dual graphs, (Fig. 3.16) we find this to be positively-correlated.

In the work of Andresen et al. [2013] and Vevatne et al. [2014] where fractures are represented using dual graphs, the networks display the property of *disassortativity* in which nodes of larger degree (longer fractures) share coordination with nodes of a smaller degree. This is also referred to as *small-world behaviour* by Watts and Strogatz [1998], a property shared by many other classes of networks. A DFN generating technique must also be able to replicate disassortativity in network realizations.

At this juncture, we revisit the point on applicability of outcrop-derived fracture networks. Recent work by Laubach et al. [2019] have raised questions on the use of fracture network data that has no provable correlation to subsurface fractures. Ukar et al. [2019] and Laubach et al. [2019] proposed protocols to identify suitable analogues based on mineral-lining evidence in fractures. In the case of network data presented in this article, we repeat this caveat that though the data is useful in studying the fracture network properties and their spatial distribution, caution needs to be exerted when extrapolating to subsurface conditions.

Although the results presented in this contribution have a very clear network structure with a very low proportion of I-nodes, many outcrop fracture studies and subsurface data indicate the prevalence of sub-parallel fractures where the isolated type of topology is the most common. Some examples of such systems are documented by Odling [1997], National Research Council [1996], Wüstefeld [2010] etc. Such fractured systems which have poor connectivity among the fractures still have an impact on fluid flow behaviour [Philip et al., 2005], and there is a need for large-scale data pertaining to such systems.

3.7. CONCLUSION

We present automatically extracted, large-scale fracture networks from limestone pavements the Bristol Channel, UK using photogrammetric data previously published by Weismüller et al. [2020]. The automatic extraction process combines methods from Chapter 2 and a series of programmatic routines described in this chapter. The functions developed receive fracture network input in the form of a graph data structure, perform node/edge manipulations on the graph so as to rectify issues such as lack of connectivity, artificial segmentation, and linking of segments. The resultant graphs can then be converted into geologically significant fracture traces amenable for further analysis. In summary, the main findings are listed as:

- Fracture networks from five fractured limestone pavements spread over approximately 17,000 sq.m are automatically extracted using the complex shearlet transform method from UAV-borne photogrammetric imagery. From a spatial graph perspective, the number of fracture segments or edges is nearly 800,000.
- A set of programmatic functions is designed to perform topological manipulations on fracture segments, resolve discontinuities, resolve artificial fragmentation, and

combine segments into geologically significant fractures. The programmatic routines are applied to the automatically extracted fracture segments and a large-scale fracture dataset comprising around 350,000 fractures is presented.

- Length distributions of fracture networks corresponding to the five regions follow a power-law scaling. Fracture orientations in two regions show considerable scatter owing to presence of sinuous fractures, while fractures in the other three regions are organized into well-defined orientation clusters.

- Analysis of node degree distributions of primal graphs indicate that the most common topology type is the degree-3 node or Y-node indicating the probable sequential development of the networks in each of the five studied outcrops with younger and shorter fractures abutting on to older and longer fractures.

- In all the mapped areas, degree distributions of dual graphs positively correlate to the total fracture lengths highlighting the fact that longest fractures are likely to have the largest topological length. All the five networks display this property of disassortativity where fractures possessing smaller degree attach on to fractures possessing a higher degree.

- The networks possess both inter-network and intra-network variability despite belonging to similar stratigraphic layers and in weakly deformed settings. The variation is quantified using spatial maps of block area distributions, fracture density, and fracture intensity, and reveal that fracturing patterns in the Lilstock pavement are heterogeneous over distances of tens of metres.

4

Quantifying Spatial Variation in Fracture Networks

Rock fractures organize as networks, exhibiting natural variation in their spatial arrangements. Therefore, identifying, quantifying, and comparing variations in spatial arrangements within network geometries are of interest when explicit fracture representations or discrete fracture network models are chosen to capture the influence of fractures on bulk rock behaviour. By treating fracture networks as spatial graphs, in this chapter we introduce a novel approach to quantify spatial variation. The method combines graph similarity measures with hierarchical clustering and is applied to investigate the spatial variation within large-scale 2D fracture networks digitized from the well-known Lilstock limestone pavements, Bristol Channel, UK. We consider three large, fractured regions, comprising nearly 300,000 fractures spread over 14,200 sq.m. from the Lilstock pavements. Using a moving-window sampling approach, we first subsample the large networks into subgraphs. Four graph similarity measures: fingerprint distance, D-measure, NetLSD, and portrait divergence, that encapsulate topological relationships and geometry of fracture networks, are then used to compute pair-wise sub-graph distances serving as input for the statistical hierarchical clustering technique. In the form of hierarchical dendrograms and derived spatial variation maps, the results indicate spatial autocorrelation with localized spatial clusters that gradually vary over distances of tens of metres with visually discernable and quantifiable boundaries. Fractures within the identified clusters exhibit differences in fracture orientations and topology. The comparison of graph similarity-derived clusters with fracture persistence measures indicate an intra-network spatial variation that is not immediately obvious from the ubiquitous fracture intensity and density maps. The proposed method provides a quantitative way to identify spatial variations in fracture networks, guiding stochastic and geostatistical approaches to fracture network modelling.

This chapter is based on Prabhakaran et al. [2021a], Investigating spatial heterogeneity within fracture networks using hierarchical clustering and graph distance metrics, *Solid Earth*, In Press

4.1. INTRODUCTION

Fracture networks in rocks develop due to loading paths that vary over geological timescale [Laubach et al., 2019]. The evolution of the network exhibits characteristics of a complex system. There is feedback between the evolving spatial structure and the rock substrate in which the networks are positioned [Laubach et al., 2018b]. The resulting spatial arrangement that emerges after cumulative network evolution is of considerable interest as it influences flow, transport, and geomechanical stability in multiple anthropogenic subsurface applications such as geothermal energy [Vidal et al., 2017], nuclear waste disposal [Wang and Hudson, 2015], aquifer management [Witherspoon, 1986], and hydrocarbon exploitation [Nelson, 2001]. Systematically documenting near-surface fracture patterns is essential, for example, in mining applications where fracture patterns often provide clues to ore deposit patterns [Jelsma et al., 2004], and in geotechnical engineering, where fractures influence stability in human-made structures such as tunnels [Lei et al., 2017].

An important property of natural fracture networks is that of *spatial organization*, which means that the arrangements are not random but follow a statistically discernable pattern. One can view the spatial arrangement of fractures as a set of objects within a geographical reference system. Within such a framework, fracture objects are either regularly spaced, irregularly spaced with statistically significant regions of close spacing, and irregularly spaced with statistically insignificant regions of close spacing [Laubach et al., 2018b]. An alternate framework is a network, where fracture objects are described in relation to one another [Andresen et al., 2013, Sanderson and Nixon, 2015, Valentini et al., 2007b]. Spatial variations in fracture network organization are quite common. The physical phenomena commonly used to explain spatial variation in fracture arrangements are stress shadowing, layer thickness differences, host rock lithology, layered mechanical anisotropy, high-strain events such as faulting/folding, and diagenesis. However, it is generally not easy to associate a type of spatial arrangement to any unique set of input boundary conditions as similar loading paths can lead to diverging patterns, and dissimilar loading paths can lead to converging patterns [Laubach et al., 2019].

Quantifying variations in spatial arrangements of fractures involves the sampling of fracture data. Such quantifications can be in the form of 1D (using scanline methods, borehole sampling), in 2D (fracture trace maps from outcrop imagery), or 3D (ground-penetrating radar, microseismic). 1D scanlines provide a method to quantify arrangements and variation, and several statistical measures have been proposed, such as fracture spacing [Priest and Hudson, 1976], fracture intensity [Dershowitz and Herda, 1992], coefficient of variation [Gillespie et al., 1993], normalized correlation count [Marrett et al., 2018], and cumulative spacing derivative [Bistacchi et al., 2020]. These measurements, however, only indicate the variation of fracture arrangements on the scanline and fail to depict the variation in directions away from the scanline direction. Scanlines do not provide information on properties such as fracture length, spatial arrangements, and relationships with other fractures.

2D fracture trace maps are especially useful as this type of data combines both geometric and topological information in the form of a network. Recent advances in UAV-photogrammetry [Bemis et al., 2014, Bisdom et al., 2017] and automated image processing algorithms [Prabhakaran et al., 2019a] have led to large datasets of 2D fracture

traces that reveal much more about network attributes than is possible from 1D sampling. Given such large datasets with rich information, it is pertinent to directly quantify spatial variation from the network structure. Spatial fracture persistence [Dershowitz and Herda, 1992] can quantify 2D spatial variation but only considers some aspects of the network (such as the sum of trace lengths, number of traces, etc., within a sampling region). Thus, there is a need for more advanced techniques specific to 2D fracture trace data and which can use the combined geometric and topological structure.

From a geostatistical perspective, the concept of spatial variability describes how a measurable attribute varies across a spatial domain [Deutsch, 2002]. Quantifying magnitude and directional dependence of the variability can also be done using geostatistical tools, provided there is a means to measure variability across multiple spatial samples. The variability in fracture data has typically been reduced to variability in attributes (such as fracture length by sampling area, number of intersections, number of sets, and orientations), and attribute variability used to make decisions of *stationarity*. The identification of representative element volumes (REVs) then follows from the choice of stationarity. However, given that natural fracture networks display spatial heterogeneity, the suitability of such REVs based on stationarity assumptions needs to be re-examined. Therefore, it is interesting to compare network variation (rather than attribute variation) across the spatial domain. Any comparative method must retain topological and geometric structures encoded within the spatial samples.

4.2. GRAPH THEORY IN FRACTURE NETWORK ANALYSIS

4.2.1. FRACTURE NETWORKS AS GRAPHS

Many authors have suggested using graph theory for the characterization of fracture networks [such as Andresen et al., 2013, Sanderson and Nixon, 2015, Sanderson et al., 2019, Valentini et al., 2007b, Vevatne et al., 2014]. In graph theory and network science, graphs are structures that comprise a set of edges and vertices representing relationships between data. In fracture networks, the vertices are intersections between fractures, and the edges represented by fracture segments connecting the vertices [Sanderson and Nixon, 2015]. By assigning positional information to the vertices (also called nodes), fractures in the form of graphs encapsulate both topological and spatial information [Sanderson et al., 2019]. An alternate graph representation is when fractures from tip-to-tip are vertices, and intersections with other fractures are edges. Barthelemy [2018] refers to these types of representations as to *primal* and *dual* forms, respectively. Others, such as Doolaeghe et al. [2020], call the two representations as *intersection graphs* and *fracture graphs*.

We had previously depicted an example of a fracture network in its primal form (see Fig. 3.2.a) and in its dual form (see Fig. 3.2.b). The degree of a graph node is simply the number of edges that are incident at a particular node. As seen in the primal graph in Fig. 3.2.(e), the maximum node degree is 6, with the most common degree value being 3. This type of degree distribution is typical for a spatial graph in which physical constraints limit the maximum possible node degree. We may note that node degrees in spatial graph representations of fracture networks are most likely to be 1,3, or 4. For fracture networks interpreted from outcrop images as depicted in Fig. 3.2(a), eroded

fractures and enlarged apertures may lead to higher degrees due to issues in resolving closely spaced nodes.

In the case of the alternate representation, referred to as dual graphs by Barthelemy [2018] and depicted in Fig. 3.2(f), the maximum degree can be much higher, and the longest fractures that have the highest number of intersections also have the highest degree. Andresen et al. [2013] and Vevatne et al. [2014] suggested that fracture networks are *disassortative* in that shorter fractures preferentially attach on to the longer fractures. The property of disassortativity is quantitatively defined using assortativity coefficients [Newman, 2002] with disassortative networks having negative assortativity coefficients. Andresen et al. [2013] and Vevatne et al. [2014] report negative assortativity coefficients for fracture networks that are represented in the dual form. From Fig. 3.17, for the Bristol Channel fracture data, such a correlation between dual graph node degree and length is observed.

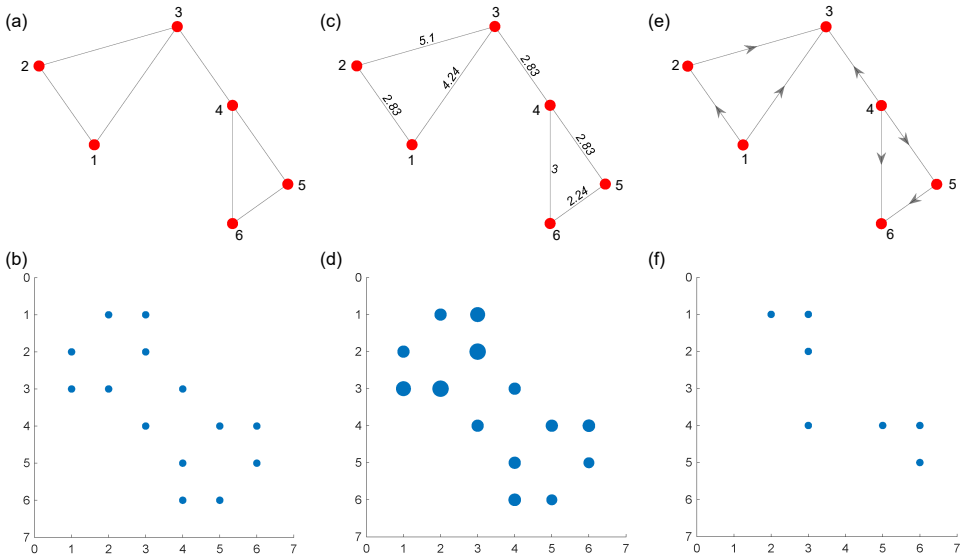

Figure 4.1: (a) An unweighted planar graph with six nodes and seven edges (b) adjacency matrix of unweighted graph (c) a weighted planar graph with edge weights proportional to euclidean distances between connecting nodes (d) weighted sparse adjacency matrix for weighted planar graph (e) a directed, unweighted graph (f) adjacency matrix of directed graph

In graph representations, weights can be assigned to edges that are proportional to the importance of that edge. In the case of fracture networks in the primal form, this can be the euclidean distance between the nodes (or fracture edge intersections). The weight may also be the direction cosine of the particular edge that indicates orientation. In the dual graph representation, intersections between fractures represent the edges. Therefore the edge weight may be specified in terms of intersection angle. Graphs may also be *directed* with a specific direction to edges. In the case of spatial graphs derived from fracture networks, an undirected but weighted representation is sufficient. Fig. 4.1(a), Fig. 4.1(c), and Fig. 4.1(e) depicts examples of unweighted, weighted, and directed planar

graphs, respectively. The corresponding adjacency matrices are depicted in Fig. 4.1(b), Fig. 4.1(d), and Fig. 4.1(f).

4.2.2. GRAPH DISTANCE MEASURES TO QUANTIFY NETWORK SIMILARITY

Several graph similarity measures exist within the graph theory literature to compare graphs [see Emmert-Streib et al., 2016, Hartle et al., 2020, Tantardini et al., 2019 for recent reviews]. Graph comparisons are a challenging, non-trivial problem in terms of computing complexity [Schieber et al., 2017]. Still, various measures exist that can capture and highlight useful aspects of the graph structure that facilitate comparisons. Graph *isomorphism* between two graphs implies that there exists a series of necessary conditions such as an equal number of nodes, edges, degree sequences, and sufficient conditions such as equal adjacency matrices [Van Steen, 2010]. An isomorphism test on two graphs G_1 and G_2, can only yield two results, either isomorphic or not. Graph *similarity* can therefore be differentiated from graph *isomorphism* in that the latter comparison can only return a binary outcome. Graph similarity on G_1 and G_2, on the other hand, returns a real-valued quantity that converges to zero when the two graphs approach isomorphism (or complete similarity).

Tantardini et al. [2019] classify distance measures based on whether the metric is capable of comparing graphs with an unequal number of nodes or not. The metrics may also be classified based on whether they can also handle weighted and directed graphs. Using a graph-similarity measure on a fracture network, we can explore spatial variations in network structure by comparing multiple sampling points.

4.2.3. COMBINING DISSIMILARITY MEASURES WITH CLUSTERING ALGORITHMS

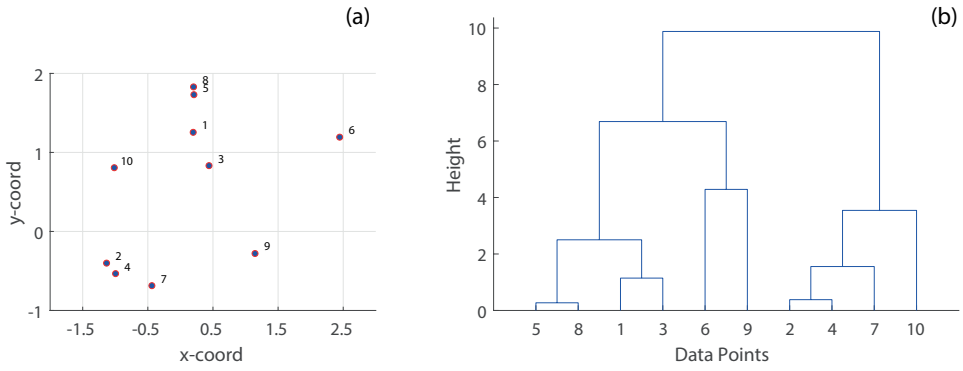

Figure 4.2: A simple example of hierarchical clustering using euclidean distance (a) 10 randomly positioned points in 2D space (b) dendrogram computed from hierarchical clustering using the euclidean distance depicting clusters of the 10 individual points at different levels organized into a hierarchy. The procedure of hierarchical clustering is shown in Algorithm 1.

Since we are interested in quantifying spatial variability, we may recast the problem as that of identifying clusters within the network. Clustering is also referred to as unsupervised classification and is a process of finding groups within a set of objects with an assigned measurement [Everitt et al., 2011]. If we consider a dataset, $D = [X_1, X_2, ...X_n]$,

containing 'n' data samples, clustering then implies arranging the elements of D into 'm' distinct subsets, $C = [C_1, C_2, ...C_m]$, where $m \leq n$. From a statistical perspective, the clustering task is different from classification because the former is *exploratory*, whereas the latter is *predictive*, although both attempt to assign labels. Therefore clustering must precede classification.

In the existing literature on fracture networks, assigning labels to specific perceived archetypal networks (or end-members) is standard. These typologies include terms such as orthogonal, nested, ladder-like, conjugated, polygonal, corridors, etc. [Bruna et al., 2019a,b, Peacock et al., 2018]. However, when faced with the reality of outcrop-derived 2D fracture trace data, it is not easy to assign such labels. Therefore, clustering is a significant and necessary step in exploratory fracture data analysis.

Hierarchical clustering (HC) is an unsupervised statistical clustering method [Kaufman, 1990] that can identify clusters within a set of observations given a distance matrix computed by applying a well-defined distance function, pair-wise on each observation. In contrast to other clustering methods such as *k-means* or *k-medoids*, which require an a priori known number of clusters as input arguments, HC re-organizes observations into hierarchical representations from which the user can pick a level of granularity. At the lowest level, there is just one cluster containing all the observations. At the highest level, the number of clusters is equal to the observations. HC algorithms are referred to as *agglomerative* or *divisive* depending upon whether they begin from a lower level or from the highest level. The clustering then organises the discrete data into a hierarchical dendrogram structure that positions the clusters based on the magnitude of similarity. By combining graph distance computations across spatially distinct samplings with unsupervised HC, cluster detection automatically leads to quantified spatial variation. A simple example of HC is illustrated on a set of randomly distributed points in space (see Fig. 4.2.a). The result is the hierarchical dendrogram structure depicted in Fig. 4.2(b).

4.3. FRACTURE DATASETS

To validate the proposed approach based on graph distance metrics and hierarchical clustering, we utilize three large-scale fracture networks from Chapter 3 (see Fig. 4.3). There is considerable spatial variation in the jointing. From previous literature documenting joints within the Lilstock pavements, the spatial variation is attributed to multiple reasons. The proposed explanations include proximity and influence of faults explained by fluid-driven radial-jointing emanating from asperities within fault [e.g., Gillespie et al., 1993, Rawnsley et al., 1998], spatial variation of thicknesses of intercalated limestone and shale layers (e.g., Belayneh, 2004), proximity to high-deformation features such as folding [e.g., Belayneh and Cosgrove, 2004a], the interplay between regional and local stresses resulting in complex stress fields [e.g., Whitaker and Engelder, 2005], inheritance from the spatial distribution of pre-existing vein/stylolite networks that influenced later joint network development [e.g., Dart et al., 1995, Wyller, 2019], and synkinematic cementation in veins affecting later development of joints [Hooker and Katz, 2015]. Recent work on fractures at the Kilve outcrop [Procter and Sanderson, 2018], exposing the same geological units as those considered in this work, conclude that anomalous fracture intensity exists in fracturing at various locations and suggest that variability in fracture intensity cannot be fully explained by variations in thickness, compositional,

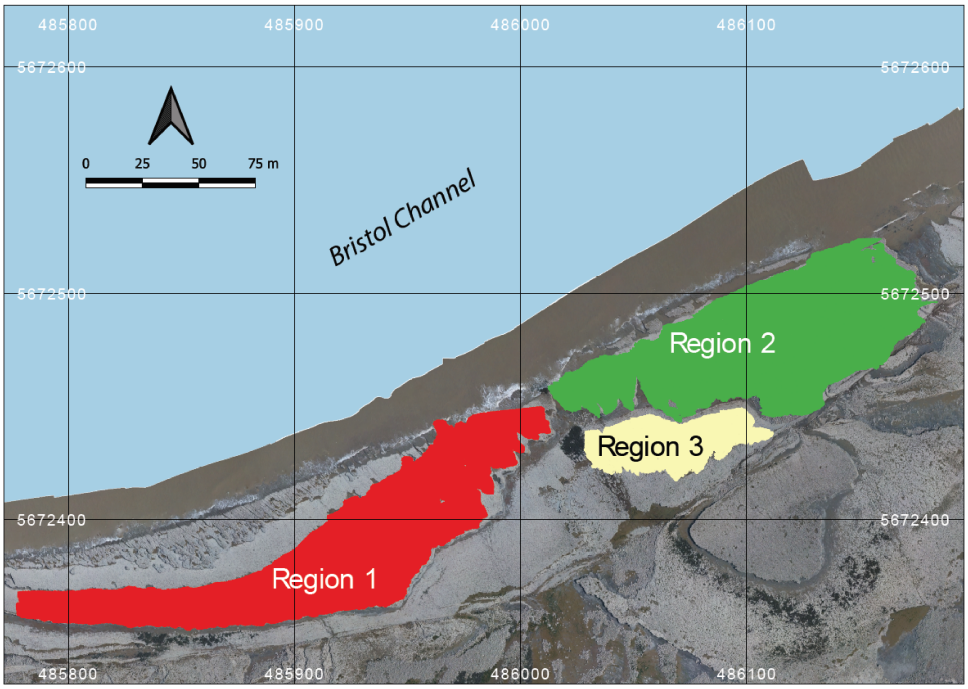

Figure 4.3: Overview of fracture networks corresponding to the three considered regions. This map is derived from an open image dataset published by Weismüller et al. [2020] and available for download with a CC-BY license

or textural variations.

Table 4.1: Summary statistics for the three regions

Region	Approx. area (sq.m)	Fractures	Edges	Nodes
Region 1	6017	124006	364703	228661
Region 2	6749	141344	365333	235089
Region 3	1473	28892	78151	49771

From this dataset, we utilize fracture networks corresponding to three contiguous regions. Figure. 4.3 depicts the three areas' spatial extent labelled as Regions 1 to 3. The intensity of fracturing is such that the spatial graphs corresponding to each region have a single connected component. Table 4.1 tabulates summary statistics for the three networks. The number of edges and nodes correspond to the primal graph representation. What is referred to as *fractures* in Table 4.1 are sequences of graph edges that are clubbed together based on continuity and a strike direction threshold (or number of dual graph nodes). Regions 1 and 2 correspond to a single stratigraphic layer but, owing to erosion, they are not contiguous within the outcrop. We treat them separately in our analysis of spatial variation.

The detailed resolution, topological accuracy, and spatial extent of the traced net-

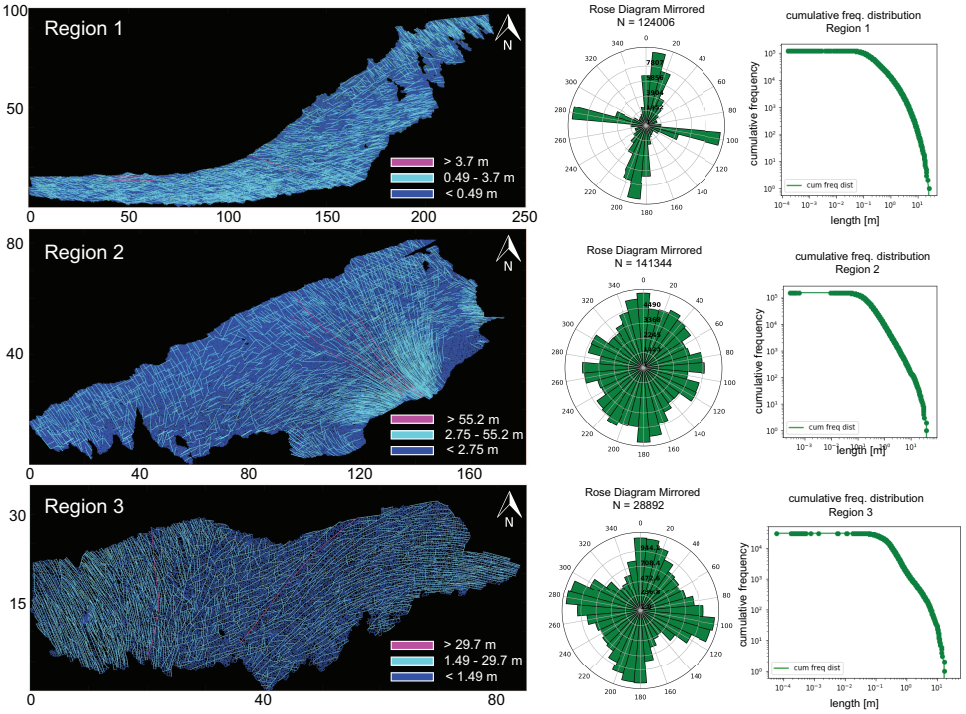

Figure 4.4: Comparison of the three regions in terms of networks, orientations and length distributions. Map dimensions are in metres.

works make the dataset appropriate for a detailed analysis of spatial variation in fracturing. The networks have significant intra- and inter-network variability in fracturing. Figure. 4.4 and Fig. 4.5 illustrate these differences. From Fig. 4.4, the fracture orientations of Region 1 depict discernable angular bins of fracture orientations. On the other hand, rose plots of Regions 2 and 3 show considerable scatter owing to the presence of long and curved fractures. Fracture length distributions are different, with Region 2 having the longest fractures and Region 1 the shortest. The distribution of joints within a particular length bin is also highly variable. In Fig. 4.5, the cumulative variation in the strike along individual fracture edges that comprise a tip-to-tip fracture is plotted as a function of the total length. The slope of the scatter plots give an indication of the fracture curvature. The slope of the scatter plot is higher in Regions 2 and 3 than in Region 1. We interpret the curvature to, therefore, be the least in Region 1.

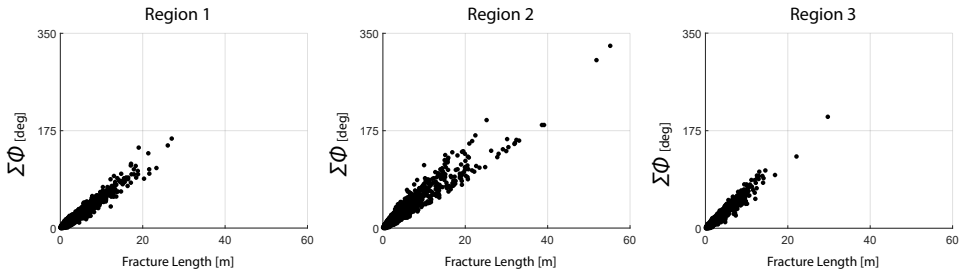

Figure 4.5: Correlation between sum of strike differences of fracture segments constituting tip-to-tip fractures versus total fracture length for the three regions

4.4. METHODS

4.4.1. SUB-SAMPLING THE NETWORK DATA

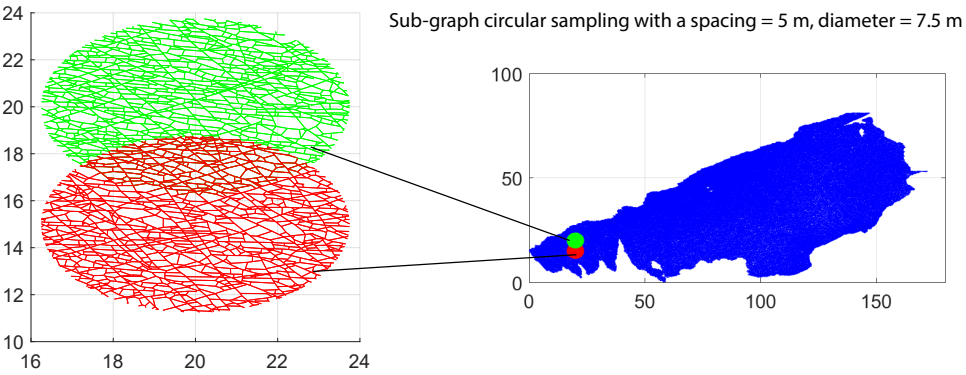

Figure 4.6: Sub-sampling of a fracture graph corresponding to full region into sub-graphs of 7.5 m diameter and spacing of 5 m

We circularly sample the fracture networks on a cartesian grid with a sub-graph extracted within a circular region centered at each grid point. The grid spacing-to-circle diameter is maintained such that neighboring sub-graphs share some portion of the area (see Fig. 4.6). Near the networks' boundaries, the sub-graphs are either too small or result in disconnected graph components. We neglect these samples so that they do not affect the clustering results. The process of circular sampling creates edge nodes with degree 1, which has the effect of altering node topology by introducing isolated, degree-1 nodes. To prevent this from impacting clustering results, we remove all edges from the sub-graphs emanating from degree-1 nodes that contact the periphery of the circular sample. This effect is illustrated in Fig. 4.7. Each sub-graph can now be compared to every other sub-graph using a graph distance metric to compute a pair-wise distance matrix. The distance matrix serves as the input to the hierarchical clustering algorithm.

For N sub-graphs, the number of comparisons necessary are $\frac{N(N-1)}{2}$. The computational complexity of graph comparison increases polynomially with the size of sub-

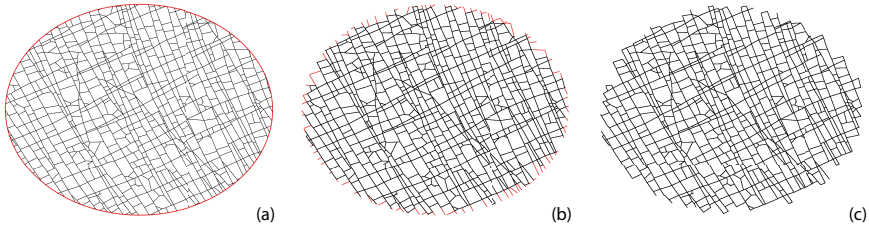

Figure 4.7: Treating isolated nodes and dangling edges that arise due to circular-sampling (a) circularly sampled subgraph with a diameter of 7.5 m (b) edges connected to isolated nodes intersected by circle (c) subgraph after removing isolated nodes and corresponding dangling edges

Table 4.2: Number of subgraphs obtained per region

Region	No. of sub-graphs
Region 1	219
Region 2	212
Region 3	117

graphs in terms of node sizes. Since the number of comparisons increases quadratically with the number of sub-graphs, we seek to balance grid spacing and sampling diameter. For Regions 1 and 2, we choose a spacing of 5 metres for circularly sampled subgraphs with a diameter of 7.5 m. For Region 3, which is also the smallest region, a spacing of 5 metres would lead to quite a smaller number of sub-graphs. Therefore, we use a more dense spacing of 3 metres with a diameter of 7.5 m. Table 4.2 tabulates the number of sub-graphs pertaining to each region.

4.4.2. GRAPH SIMILARITY MEASURES

We use the following four graph similarity measures to compare the sub-graphs.

- Fingerprint Distance [Louf and Barthelemy, 2014]
- D-measure [Schieber et al., 2017]
- NetLSD [Tsitsulin et al., 2018]
- Portrait Divergence [Bagrow and Bollt, 2019]

The performance of these similarity measures have been validated previously by Hartle et al. [2020] and Tantardini et al. [2019] for a variety of benchmark graph datasets. Each similarity measure is described briefly in the following subsections. The reader is referred to the references above for further details on the similarity measures.

FINGERPRINT DISTANCE

The fingerprint distance introduced by Louf and Barthelemy [2014] is purely geometric and combines statistics of block faces and shape factors in computing a probability distribution of a spatial graph. Louf and Barthelemy [2014] formulated the measure in the context of quantifying differences in street patterns. A *block* denotes the 2D region

enclosed by graph edges. For any given spatial graph, this corresponds to the number of bounded sub-graphs or primary cycles. We neglect isolated fractures and those having dead ends when computing these blocks. Given the network intensity in our dataset, such isolated fractures are minimal. Every block has an associated *shape factor*, 'ϕ' which is expressed in terms of block area 'A' and circumscribing circle area, 'A_c' ,

$$\phi = \frac{A}{A_c} \tag{4.1}$$

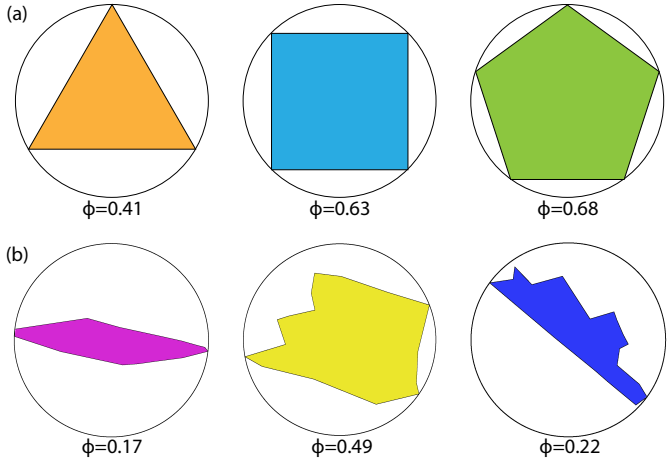

Figure 4.8: (a) shape factors for regular block shapes with equal edge lengths (b) shape factors for polygonal blocks resulting from real fracture networks in Region 1 (dimensions are relative)

The value of ϕ is always smaller than 1, with larger values meaning that the block face shape is closer to that of a regular polygon. Figure. 4.8(a) depicts shape factors of regular polygons versus that of polygons derived from spatial networks in Fig. 4.8(b). No unique correspondence exists between a particular shape and a magnitude of ϕ; however, the overall distribution of ϕ indicates reveals block shape distribution patterns and highlights differences between spatial graphs. The shape factor alone does not fully serve as a similarity measure as blocks can have similar shapes but different face areas. The distribution of the block-face areas is binned logarithmically to integrate information from the shape factor and block area distributions. A conditional probability distribution, $P(\phi|A)P(A)$, is then defined representing the contribution of $P(\phi)$ for each area bin and the summation of which yields the fingerprint curve, $P(\phi)$,

$$P(\phi) = \sum_{A} P(\phi|A)P(A). \tag{4.2}$$

An example of a *fingerprint*, so named by Louf and Barthelemy [2014], is depicted in Fig. 4.9(e) and Fig. 4.9(j), with the distribution curves for three area bins, for two fracture networks derived from image tiles (see Figs. 4.9.b,c,f,g) corresponding to Region 1 (Fig. 4.9.a). The curves in Fig. 4.9(e) and Fig. 4.9(j) encapsulates information based on shape

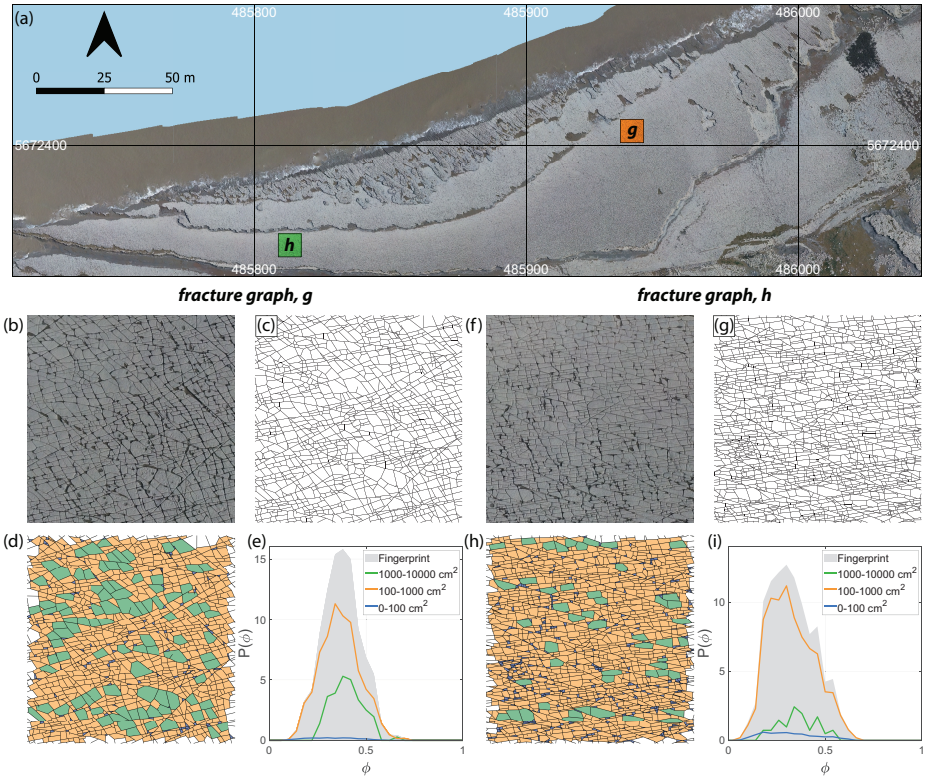

Figure 4.9: (a) Overview of Region 1 with two selected 1000 x 1000 pixel image tiles (b) enlarged view of first image tile (c) fracture network corresponding to first tile as a spatial graph with dimensions of 8.6 m x 6.75 m and having 3583 edges and 2382 nodes (d) block face areas coloured as per three area bins, 0-100 cm^2, 100-1000 cm^2, and 1000-10000 cm^2 (e) $P(\phi)$ or fingerprint of the sub-graph depicting the combined effects of area and shape factor, ϕ pertaining to the three area bins. (f) enlarged view of second image tile (g) fracture network corresponding to second image tile as a spatial graph with 5418 edges and 3539 nodes (h) block face areas binned logarithmically (i) fingerprint of second spatial graph. (a),(b), and (f) are derived from images contained in the open dataset (CC-BY license) published by Weismüller et al. [2020]

factors and block areas (see Figs. 4.9.d,h), including the proportional contribution from all logarithmic area bins considered.

Denoting $f_\alpha(\phi)$ as the ratio of the number of faces with a shape factor 'ϕ' that lie in a bin 'α' over the total number of faces for that graph, a distance d_α between two graphs G_a and G_b is computed by integrating over $f_\alpha(\phi)$ for the two different graphs. The distance based on $f_\alpha(\phi)$ of the two graphs for a single area bin is defined as:

$$d_\alpha(G_a, G_b) = \int_0^1 |f_\alpha^a(\phi) - f_\alpha^b(\phi)|^n d\phi \tag{4.3}$$

As per Louf and Barthelemy [2014], the value of n can either be 1 or 2. We choose $n = 1$ in our computation. The global fingerprint distance D_{FP} between G_a and G_b can then be computed summing over all area bins α,

$$D_{FP}(G_a, G_b) = \sum_{\alpha} d_{\alpha}(G_a, G_b)^2 \tag{4.4}$$

A MATLAB implementation of the fingerprint distance is presented in Prabhakaran [2021b] and we computed the distance matrix for all sub-graphs corresponding to the three regions using this implementation.

D-MEASURE

The D-measure introduced by Schieber et al. [2017] is a three-component distance metric with weighting constants for each component. The three properties of graphs compared are the network node dispersion (NND), node distance distribution (μ), and the alpha centrality (α). The dissimilarity measure, D_{DM}, is the weighted sum:

$$D_{DM}(g,h) = w_1\sqrt{\frac{\mathscr{J}(\mu_g,\mu_h)}{log2}} + w_2\left|\sqrt{NND(g)} - \sqrt{NND(h)}\right| + \frac{w_3}{2}\left(\sqrt{\frac{\mathscr{J}(P_\alpha(g),P_\alpha(h))}{log2}} + \sqrt{\frac{\mathscr{J}\left(P_\alpha(g^c),P_\alpha(h^{c'})\right)}{log2}}\right), \tag{4.5}$$

where $\mathscr{J}$ indicates the Jensen-Shannon divergence. The constants w_1, w_2, and w_3 in Eq. (4.5) are real and non-negative weights such that $w_1 + w_3 + w_3 = 1$.

As per Schieber et al. [2017] the first term in Eq. (4.5) compares averaged connectivity node's patterns as per node distance distribution. Schieber et al. [2017] define NND, within the second term, as a measure of the heterogeneity of a graph with respect to connectivity distances that capture global topological differences. The NND is computed as:

$$NND(G) = \frac{\mathscr{J}(P_1,....,P_N)}{log(d+1)}, \tag{4.6}$$

where the numerator in Eq. (4.6) is the Jensen Shannon divergence of N connectivity distance distributions $[P_1, P_2.....P_N]$. P_i is constructed as $P_i = p_i(j)$ where $p_i(j)$ is the fraction of nodes connected to node i at distance j. The Jensen-Shannon divergence of $[P_1, P_2.....P_N]$ is expressed as:

$$\mathscr{J}(P_1,....,P_N) = \frac{1}{N}\sum_{i,j} p_i \log\left(\frac{p_i(j)}{\mu_j}\right). \tag{4.7}$$

μ_j in Eq. (4.7) is the average of N distributions and can be written as,

$$\mu_j = \frac{1}{N}\sum_{i=1}^{N} p_i(j). \tag{4.8}$$

The third term in Eq. (4.5) is based on probability density functions associated with alpha centrality of graph $P_\alpha(g)$ and alpha centrality of the graph complement $P_\alpha(g^c)$. The value of weights was suggested by Schieber et al. [2017] as $w_1 = w_2 = 0.45$ and $w_3 = 0.1$. We use the implementation provided by Schieber et al. [2017] with these sets of weights to build the distance matrices for all sub-graphs within the three regions of interest. We depict in Fig. 4.10 for the two example fracture networks, the three properties that are used in computing the D-measure.

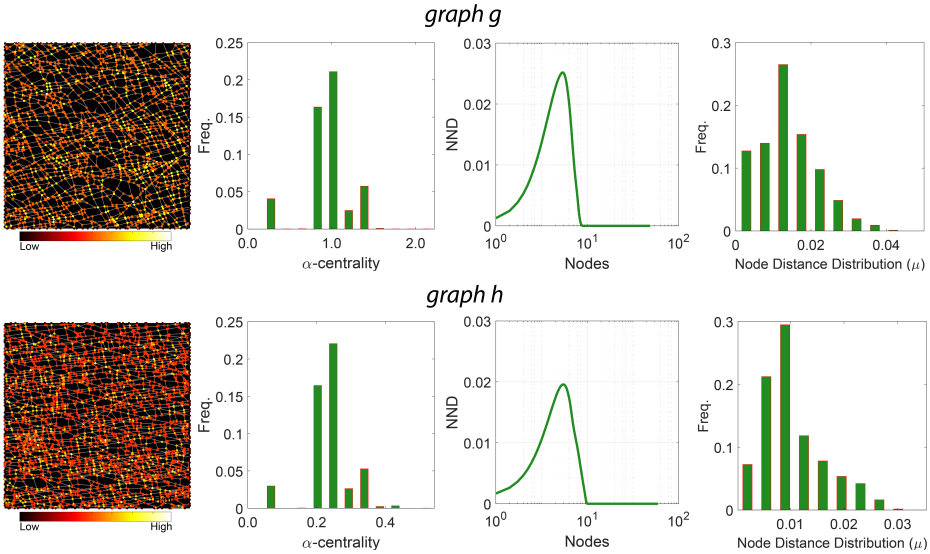

Figure 4.10: D-measure components for the two example fracture graphs comparing α-centrality of nodes, distributions of α-centrality, NND distributions, and node distance distributions

PORTRAIT DIVERGENCE

The Portrait Divergence similarity score derives from *network portraits* introduced by Bagrow et al. [2008] for unweighted graphs and extended to weighted graphs by Bagrow and Bollt [2019]. For a graph g with N nodes, the network portrait is defined as a matrix B_{lk} where each entry is the number of nodes with k nodes at l distance. The limits of l and k are $0 \leq l \leq d$ and $0 \leq k \leq N-1$, with d being the diameter of the graph. The row entries of the network matrix B_{lk} are probability distributions of a random node having k nodes at a distance l:

$$P(k|l) = \frac{B_{lk}}{N} \tag{4.9}$$

For a second graph h, if the network matrix is B'_{lk} with a corresponding probability distribution of $Q(k|l)$ and diameter d', the Kullback Leibler (KL) divergence between $P(k|l)$ and $Q(k|l)$ is expressed as:

$$KL(P(k|l)||P(k|l)) = \sum_{l=0}^{max(d,d')} \sum_{k=0}^{N} P(k,l) log \frac{P(k,l)}{Q(k,l)} \tag{4.10}$$

The portrait divergence $D_{PD}(g,h)$ is computed by the Jensen Shannon divergence between $P(k|l)$ and $Q(k|l)$:

$$D_{PD}(g,h) = JSD(P(k|l), Q(k|l). \tag{4.11}$$

This can be expressed in terms of Kullback Leibler divergences and mixture distributions as:

$$D_{PD}(g,h) = \frac{1}{2}(KL(P||M) + KL(Q||M)) \tag{4.12}$$

where the mixture distribution M of $P(k|l)$ and $Q(k|l)$ is given by:

$$M = \frac{1}{2}(P(k|l) + Q(k|l)) \tag{4.13}$$

The portrait divergence measure provides a single value $0 \leq D_{PD}(g,h) \leq 1$ for any pair of graphs. Bagrow and Bollt [2019] applied the portrait divergence measure to both synthetic and real world networks. The code implementation of portrait divergence attached with Bagrow and Bollt [2019] is used to construct the distance matrices for all sub-graphs within the three regions of interest. The network portrait or the B_{lk} matrix for the example fracture graphs are depicted as heatmaps in Fig. 4.11.

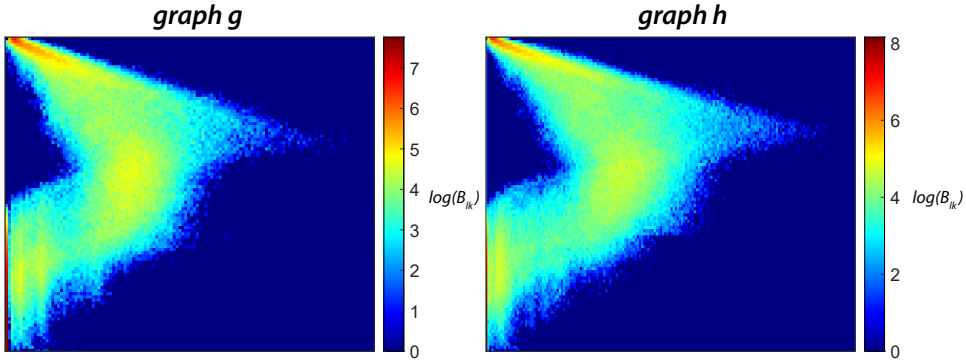

Figure 4.11: Heatmap representations of network portrait sparse matrices (B_{lk}) for the two example fracture graphs

LAPLACIAN SPECTRAL DESCRIPTOR

The NetLSD distance was introduced by Tsitsulin et al. [2018]. It is based on a Frobenius norm computed between heat trace signatures of normalized Laplacian matrices of two graphs. For a graph g with a normalized Laplacians L and n nodes, a heat kernel matrix is defined as:

$$H_t = e^{-tL} = \sum_{j=1}^{n} e^{-t\lambda_j} \phi_j \phi_j^T \tag{4.14}$$

Using the heat kernel matrix H_t, a heat trace h_t is defined as:

$$h_t = \sum_{j=1}^{n} e^{-t\lambda_j} \tag{4.15}$$

For a second graph g' with a heat trace signature of h'_t, the NetLSD distance D_{LSD} is then the Frobenius norm of the two heat signatures as:

$$D_{LSD} = ||h_t, h'_t||_{Frobenius} \tag{4.16}$$

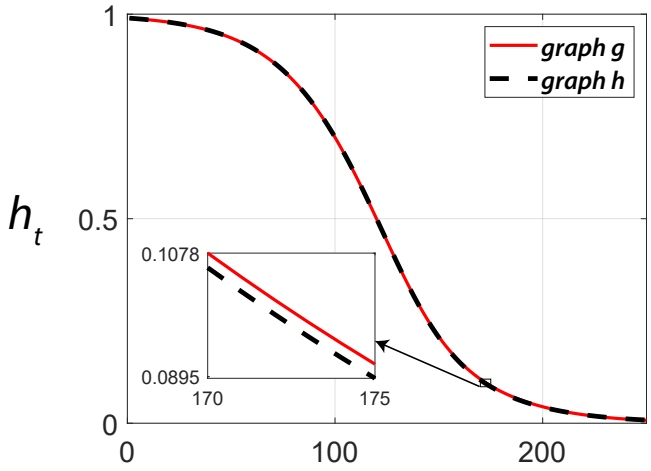

Figure 4.12: Comparing heat trace signature vectors for the two example fracture graphs computed using NetLSD

Figure. 4.12 depicts heat trace signatures computed using the NetLSD python package implemented by Tsitsulin et al. [2018] for the two example fracture graphs. We use this package to populate the distance matrices associated with sub-graphs from each region.

The values of graph similarity computed using the four metrics described by Equations (4.4), (4.5), (4.12), and (4.16) for the two example fracture graphs depicted in Fig. 4.9(c) and Fig. 4.9(g) are summarized in Table. 4.3.

Table 4.3: Summary of graph similarities computed for example fracture networks

Graph Similarity	**Value**
Fingerprint Distance [D_{FP}]	0.1414
D-measure [D_{DM}]	0.1244
Portrait Divergence [D_{PD}]	0.2926
NetLSD [D_{LSD}]	0.0147

4.4.3. HIERARCHICAL CLUSTERING

After sub-sampling the fracture networks (see Section 4.4.1) and using the graph distance metrics described in Section 4.4.2 to construct distance matrices, we apply hierarchical clustering. HC can be done in an agglomerative versus divisive manner [Hennig et al., 2016]. We utilize the agglomerative approach, which generally follows the steps described in Algorithm 1. Based on how linking of clusters is done as per Algorithm 1(iii), HC can be classified into methods such as single linkage, complete linkage, unweighted pair-group average, weighted pair-group average, unweighted pair-group cen-

Algorithm 1 Agglomerative Clustering

Input: Data $D = [X_1, X_2, ...X_n]$
Output: Dendrogram $C = [C_1, C_2...C_m]$
(i). *Initialization.* m clusters of one element each with pair-wise distances computed and stored in symmetric square distance matrix D_{dist}
(ii). form pair C_i and C_j that are closest within C
(iii). form cluster $C_k = C_i \cup C_j$ and generate a new dendrogram node
(iv). update D_{dist} after computing distance between C_k and $C - C_k$
(v). delete rows and columns corresponding to C_i, C_j from D_{dist} and add rows and columns pertaining to C_k
(vi). repeat (i) - (v) till only a single cluster remains

troid, weighted pair group centroid, and Ward's method [Wierzchoń and Kłopotek, 2018]. Ward's method performs the linkage by minimizing the sum-of-squares of distances between objects and cluster centres. We use Ward's method implemented within the R statistical programming environment to apply the HC to all the sub-graph distance data.

4.5. RESULTS

We first show region-wise results of graph property computations. Intra-region spatial clustering resulting from the combined application of graph similarity measures with HC is then discussed. We use the following abbreviations for brevity throughout the section: FP - fingerprint distance, DM - D-measure, LSD - NetLSD, PD - portrait divergence.

4.5.1. REGION-WISE GRAPH CHARACTERISTICS

Fingerprints pertaining to the regions is depicted in Fig. 4.13(a). The peak of the fingerprint plot is highest at a shape factor of 0.4 for Region 1 and increases to above 0.5 for Regions 2 and 3. Histograms in Fig. 4.13(a) depict the number of polygons within each area bin pertaining to fracture networks in each region. The network portraits or B_{lk} matrices of each sub-graph within the three regions are combined to create ensemble region-wise network portraits depicted as heatmaps in Fig. 4.13(b). The non-zero entries in the B_{lk} matrices, indicated by warmer colours in the heatmaps, have visibly different patterns. Heat traces for the sub-graphs in each region are shown in Fig. 4.13(c). Figure. 4.14 depicts the variation of the network properties that are components of the D-measure distance i.e., α- centrality, NND, and μ for sub-graphs for the three regions.

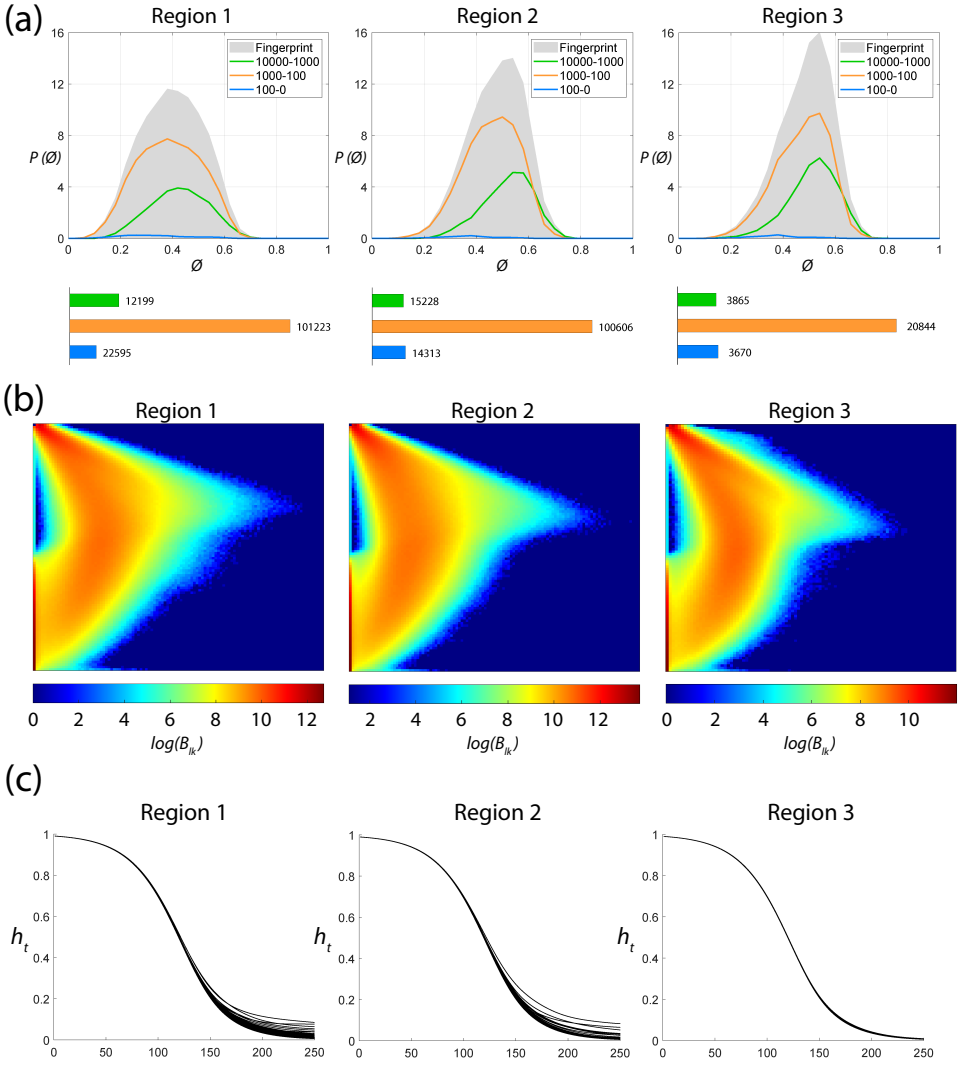

Figure 4.13: Region-wise graph properties (a) Fingerprints (b) Network Portrait Ensembles (c) Heat trace vectors

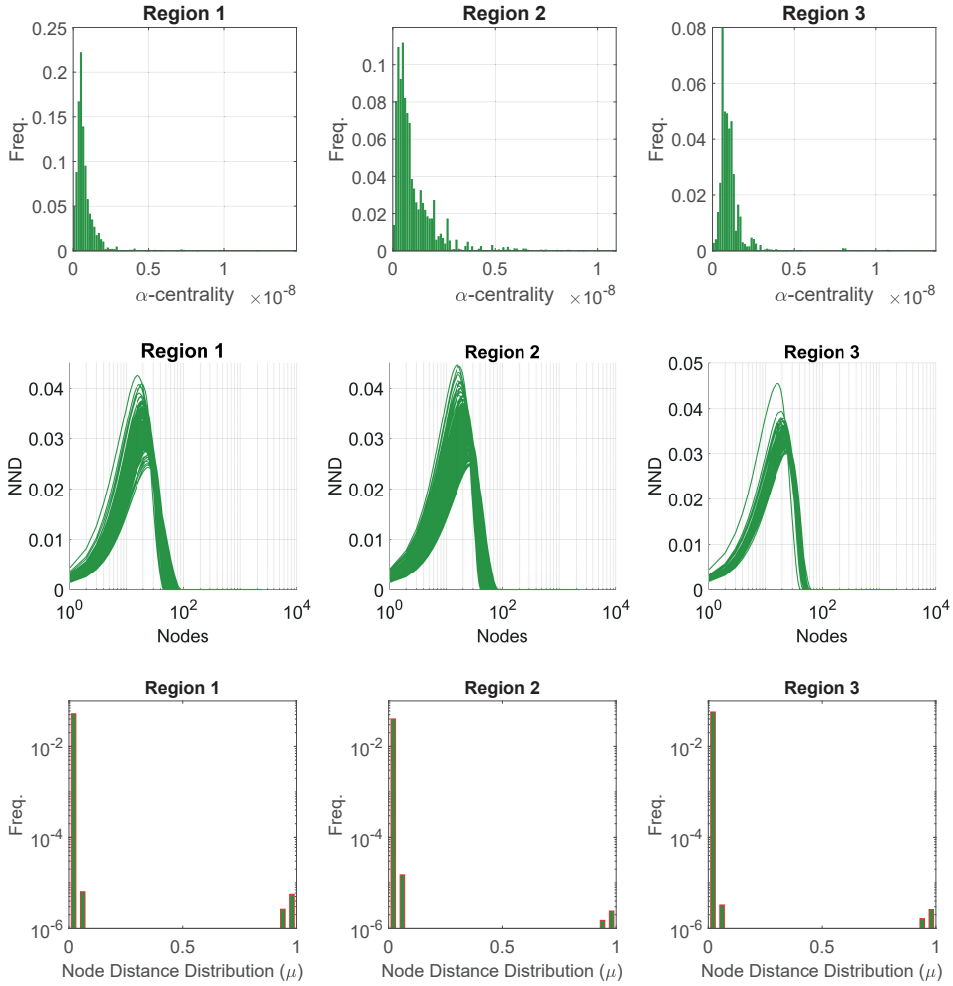

Figure 4.14: Region-wise properties used to compute the D-measure represented as ensemble plots of α-centrality, network node dispersion (NND), and node distance (μ) distributions for subgraphs

4.5.2. INTRA-REGION SPATIAL VARIATION

Intra-region spatial variation results can be presented as distance matrix heatmaps corresponding to each graph similarity metric. Dendrograms depict the hierarchical organization of the sub-graphs corresponding to similarity entries within the distance matrix entries. The intra-regional variation is more intuitively illustrated spatially by showing sub-graphs using an appropriate colour scheme that groups similar clusters under colours picked within a linear spectrum. This section presents the clustering results for all three regions using a combination of dendrograms, spatial cluster maps, and heatmaps.

REGION 1

The spatial distribution of clusters pertaining to the four distance metrics overlain over the network is shown in Fig. 4.15(a)-(d) along with the associated dendrograms for the top 10 clusters. The sub-graphs are represented by coloured discs that follow a diverging colour scheme. The number of sub-graphs within each of the top 10 clusters is also listed under the dendrogram branches. It may be noted that the top 10 clusters are shown to depict, analyse, and compare the spatial variation across distance measures. A complete, uncut dendrogram and associated heatmaps of the similarity measures are depicted in Appendix Fig. C.1. We can cut the dendrogram at different heights guided by slope changes in the weighted sum of squares plots shown in Fig. C.1. The boundaries of spatial clusters vary with the dendrogram cut height, with sub-regions emerging by traversing deeper into the dendrogram. This variation is depicted in Appendix Figs. C.4 - C.5 for a range of clusters varying from 5-10. The number of sub-samples for the four similarity measures pertaining to a dendrogram cut of $k = 10$ is tabulated in Table. 4.4.

Table 4.4: Summary of sub-graphs within each cluster of Region 1 for $k = 10$

Metric ↓	clst 1	clst 2	clst 3	clst 4	clst 5	clst 6	clst 7	clst 8	clst 9	clst 10
FP	3	36	24	47	2	24	41	16	20	6
DM	24	25	5	15	17	19	40	39	10	25
LSD	12	17	21	23	16	28	11	30	13	48
PD	38	17	3	5	24	25	79	13	6	9
Total					219					

We can observe that spatial autocorrelation exists for the FP (Fig. 4.15.a), DM (Fig. 4.15.b), and PD (Fig. 4.15.d) similarity measures. The LSD yields a speckled pattern with no obvious spatial autocorrelation (Fig. 4.15.c). In order to compare clustering results derived from the graph similarity measures, the spatial fracture persistence P_{20} and P_{21} computed using box-counting (box size of 0.5 × 0.5 m) is depicted in Fig. 4.15(e) and Fig. 4.15(f), respectively. Comparing clusters derived from graph similarity measures to the fracture persistence plots reveals boundaries within the network that are not easily discernable from the latter. Since LSD does not show spatial autocorrelation, we do not analyse it further.

Fig. 4.16(a)-(c) depicts topology histograms and rose plots of the clusters pertaining to the remaining three similarity measures. The orientation rose plots and topological summaries are generated by combining all circular samples identified under a cluster into ten clusters sub-graphs from the larger region fracture graph. It can be ob-

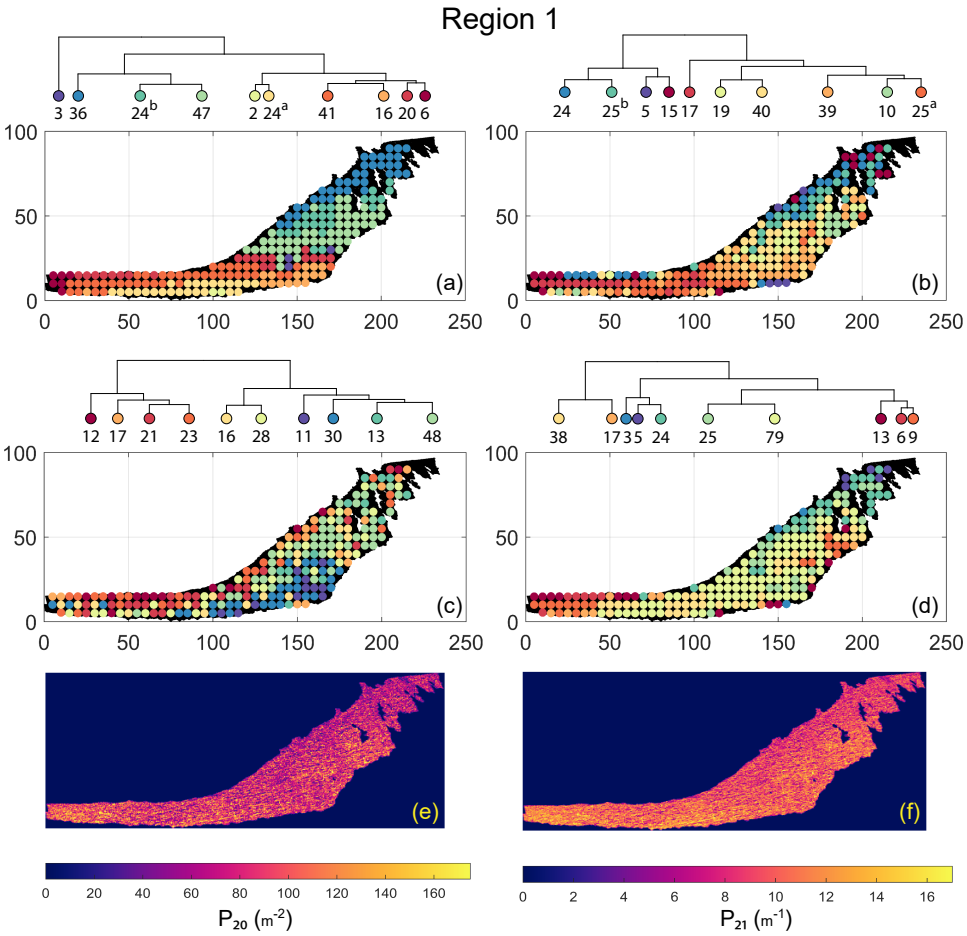

Figure 4.15: Hierarchical clustering results for Region 1 depicting the top 10 clusters using (a) Fingerprint distance (b) D-measure distance (c) NetLSD distance (d) Portrait Divergence distance (e) spatial P_{20} (f) spatial P_{21}

served from the rose plots that the clusters have varying fracture orientations that transitions across the hierarchy identified by the dendrograms. The topological summaries of the clusters do not vary significantly. Appendix Figs. C.10-C.12 depicts zoomed-in sub-graphs corresponding to each of the top 10 clusters that visually confirm the intra-regional variation.

We briefly describe the characteristics of the clustering results prefixing 'n' to the number of subsamples within a cluster to refer to a particular cluster at a k=10 dendrogram cut. From Fig. 4.15(a) and the zoomed-in archetypal examples in Fig. C.10, the clustering derived from FP seems to have a N-S variation trend. The trend is corroborated by observing the dendrogram, which splits into a northern branch comprising of clusters $n36$, $n24^b$, $n47$ and a southern branch with clusters $n6$, $n20$, $n41$, $n16$, $n2$, $n24^a$.

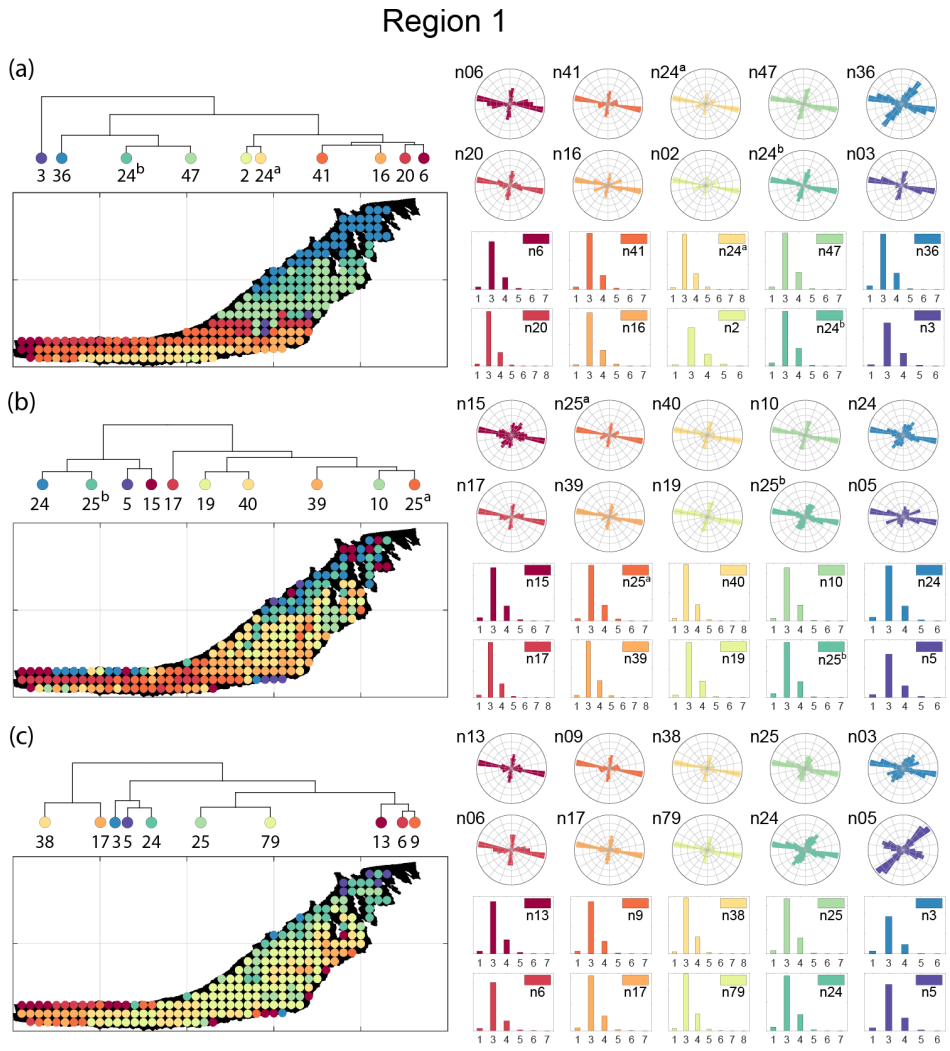

Figure 4.16: Variation in fracture orientations and topological summary for Region 1 corresponding to (a) Fingerprint (b) D-Measure (c) Portrait divergence

An outlier branch $n3$ exists at the boundary between northern and southern branches.

A similar variation is observable from the result of DM (see Fig. 4.15.b). However, the cluster demarcations are less stark than with FP with a notable stippled pattern. A major dendrogram division is a branch consisting of a thin sliver in the N-E (clusters $n24$, $n25^b$, $n5$, $n15$) which also include some boundary periphery samplings in the wast and south of Region 1. The south-western sliver is mainly contained in a branch containing cluster $n17$. The central parts of Region 1 fall under the dendrogram branch containing clusters $n39$, $n10$, $n25^a$. The remainder of the Region 1 is covered by branch containing clusters

$n19$ & $n40$. Appendix Fig. C.11 depicts archetypal examples of sub-graphs relevant to each cluster for DM.

The results of PD also depict N-S variation (see Fig. 4.15.d) in the clustering. Similar to DM, PD is also sensitive to the sub-graph completeness with peripheral clusters represented under $n13$. The branches comprising $n13$, $n6$, $n9$, and $n38$, $n17$ closely correspond to the trend of high fracture persistence (compare with Fig. 4.15.e and Fig. 4.15.f). Similar to results from FP and DM, the thin sliver in the N-E of Region 1 is captured under the branch with clusters $n3$, $n5$, $n24$. The remainder of Region 1 falls under clusters $n25$, $n79$. Fig. C.12 depicts archetypal examples of sub-graphs corresponding to each cluster for PD.

REGION 2

Spatial distribution along with dendrograms of top ten clusters pertaining to the four graph similarity measures for Region 2 is depicted in Fig. 4.17. The full dendrograms and heatmaps are placed in Appendix Fig. C.2. The variation of spatial clusters with different choices of dendrogram cut-heights is shown in Appendix Figs. C.6 and C.7. The number of sub-samples for the four similarity measures pertaining to a dendrogram cut of $k = 10$ is tabulated in Table. 4.5. Similar to Region 1, there is marked spatial autocorrelation with FP (Fig. 4.17.a), DM (Fig. 4.17.b), and PD (Fig. 4.17.d), whereas the LSD (Fig. 4.17.c) shows a speckled pattern. The spatial clustering results can be compared with the fracture persistence plots in Fig. 4.17(e) and Fig. 4.17(f).

Table 4.5: Summary of sub-graphs within each cluster of Region 2 for $k = 10$

Metric ↓	clst 1	clst 2	clst 3	clst 4	clst 5	clst 6	clst 7	clst 8	clst 9	clst 10
FP	20	22	41	52	9	17	36	8	3	4
DM	19	20	14	25	10	16	23	38	23	24
LSD	2	6	31	28	53	9	15	38	5	25
PD	17	30	17	8	24	15	28	24	15	34
Total					212					

Node degree histograms and rose plots depict the differences in network topology and fracture orientations between the identified clusters pertaining to FP (Fig. 4.18.a), DM (Fig. 4.18.b), and PD (Fig. 4.18.c). For all three measures, the shape of rose plots indicates a transition of principal orientations smoothly across clusters. For example in Fig. 4.18(a) for FP, the more complex fracturing in the west of Region 2 is depicted by cluster $n20$ with a very diffuse rose plot, changing orientations to a predominantly orthogonal pattern in cluster $n03$. The DM (clusters $n16$ and $n10$ in Fig. 4.18.b) and PD (clusters $n08$ and $n17^b$ in Fig. 4.18.c) also identify this region of orthogonal fracturing. The corresponding topological summaries also depict an increased proportion of degree-4 nodes as compared to the histograms of other clusters.

From FP clustering results (see Fig. 4.17.a), the dendrogram identifies a western branch with clusters $n20$ and $n22$. The branch comprising of clusters $n8$, $n3$, $n4$ correspond to the radial fracturing region identified by Gillespie et al. [1993] that originates from the fault in the SE of Region 2. Clusters $n9$, $n17$, $n36$ all under a branch covering parts of Region 2 further away from the radial fracturing region. Clusters $n41$, $n52$ originate under a branch forming the northern and eastern boundaries of Region 2. Fig. C.13

depicts archetypal sub-graphs under each cluster in detail for FP. The clustering results of DM (Fig. 4.17.b) and PD (Fig. 4.17.d) appear to be similar and with dendrograms roughly splitting into three main branches that correspond to specific portions of Region 1. First is the radial fracturing area represented by branch forming clusters $n10$, $n16$, $n23^b$ for D-measure (Fig. 4.17.b) and branch forming clusters $n17^b$, $n8$ for the portrait divergence (Fig. 4.17.d). The area to the N-W periphery of Region 2, farthest away from the fault, is represented by branch forming clusters $n19$, $n20$, $n14$, $n25$ for DM and by branch forming clusters $n24^a$, $n15^a$, $n34$ for PD. The transition region branch is represented within the DM dendrogram by clusters $n38$, $n23^a$, $n24$ and within the PD dendrogram by clusters $n24^b$, $n15^b$, $n28$. Appendix Figs. C.14 and C.15 depict detailed sub-graph examples for DM and PD, respectively.

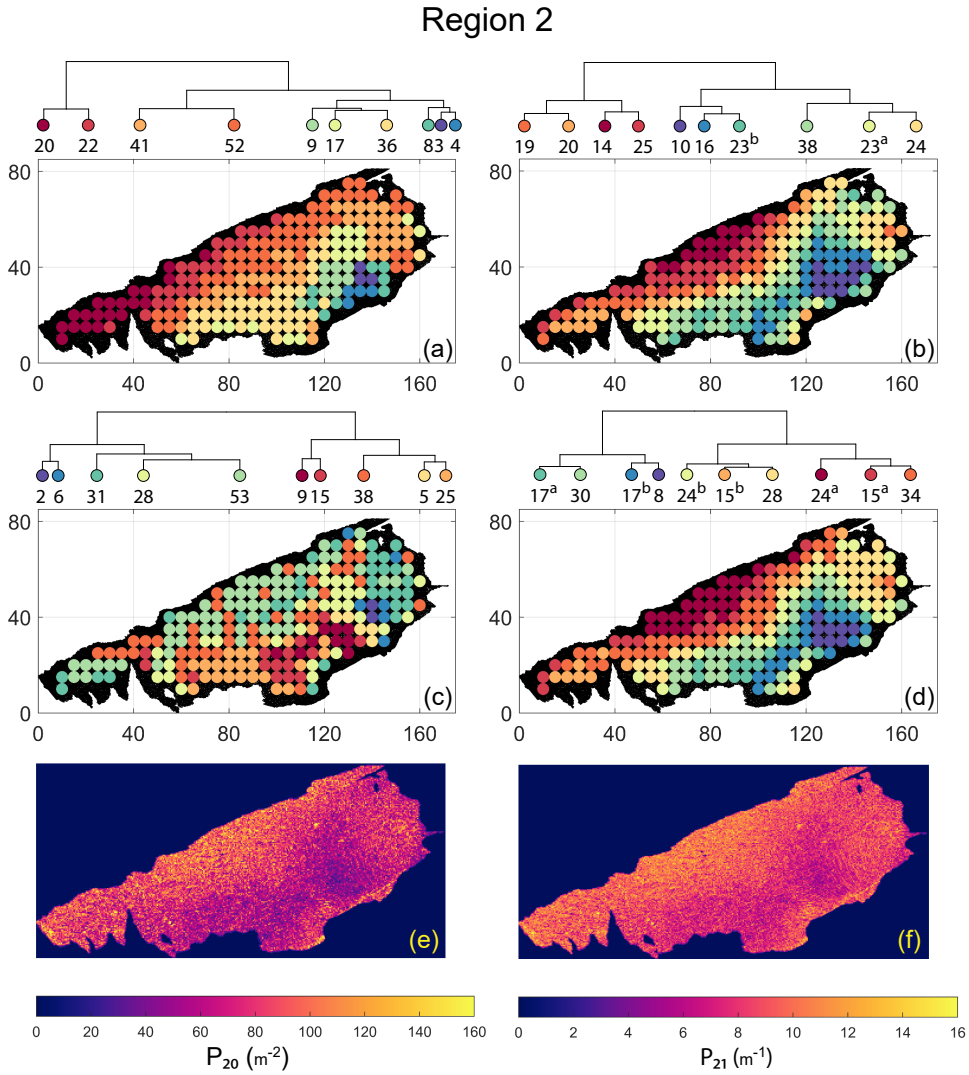

Figure 4.17: Hierarchical clustering results for Region 2 depicting the top 10 clusters using (a) Fingerprint distance (b) D-measure distance (c) NetLSD distance (d) Portrait Divergence distance. (e) spatial P_{20} (f) spatial P_{21}

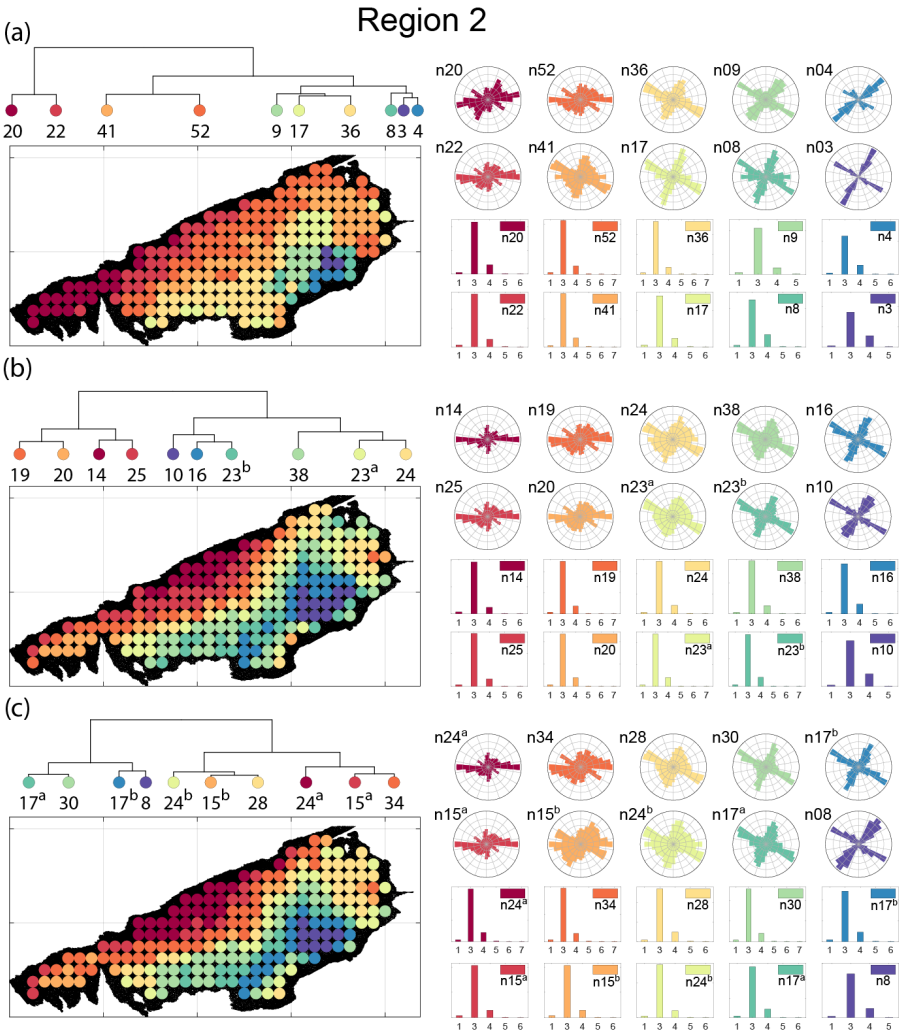

Figure 4.18: Variation in fracture orientations and topological summary for Region 2 corresponding to (a) Fingerprint (b) D-Measure (c) Portrait divergence

REGION 3

The spatial distribution along with dendrograms of the top 10 clusters pertaining to the four graph similarity measures for Region 3 is depicted in Fig. 4.19. The full dendrograms and heatmaps are placed in Appendix Fig. C.3. The variation of spatial clusters with different choices of dendrogram cut-heights (and number of clusters) is shown in Appendix Figs. C.8 and C.9. Similar to the Region 1 and 2 results, there is marked spatial autocorrelation with FP (Fig. 4.19.a), DM (Fig. 4.19.b), and PD (Fig. 4.19.d), whereas the LSD (Fig. 4.19.c) shows a stippled pattern. The spatial clustering results can be compared with the fracture persistence plots in Fig. 4.19(e) and Fig. 4.19(f). The number of sub-samples for the four similarity measures associated with a dendrogram cut of $k = 10$ is tabulated in Table 4.6.

Table 4.6: Summary of sub-graphs within each cluster of Region 3 for $k = 10$

Metric ↓	clst 1	clst 2	clst 3	clst 4	clst 5	clst 6	clst 7	clst 8	clst 9	clst 10
FP	2	5	28	11	27	6	9	4	11	14
DM	24	15	6	16	1	5	5	25	3	17
LSD	9	25	4	9	6	5	5	10	16	28
PD	7	23	25	1	5	21	4	16	3	12
Total					117					

Node degree histograms and rose plots depict the differences in network topology and fracture orientations between the identified clusters relating to FP (Fig. 4.20.a), DM (Fig. 4.20.b), and PD (Fig. 4.20.c). For all three measures, the shape of rose plots indicates a transition of principal orientations smoothly across clusters. For example, in Fig. 4.20(a) for the fingerprint measure, the cluster $n06$ in the west of Region 3 has three main sets that become orthogonal in cluster $n09$, the nearest cluster eastwards. Cluster $n05$ at the eastern extremity of Region 3 has an orthogonal pattern that has rotated almost 80 degrees clockwise compared to the western boundary. Orientations of fractures clusters between the eastern-most and western-most clusters show transitions between the extremal archetypes.

From the FP clustering results (Fig. 4.19.a), the spatial variation appears to have an E-W trend. From the dendrogram, an eastern branch comprising clusters $n6$, $n9$, 4, 11^a, 14 and a western branch consisting of clusters $n5$, $n28$, $n11^b$, $n27$ can be identified. An outlier branch with Cluster $n2$ appears at the interface between the eastern and western branches. Detailed visualization of archetypal sub-graphs relating to each of the FP clusters is presented in Appendix Fig. C.16. The dendrogram structure and spatial clustering for the DM (Fig. 4.19.b) depicts a central region represented by a branch containing clusters $n24$, $n15$, $n6$, $n16$. The eastern and western peripheries organize as clusters $n1$, $n5^a$, $n5^b$, $n25$, $n3$, $n17$ under a second branch. Underneath this branch, clusters $n1$, $n5^a$, $n5^b$ correspond to extremities of the Region 3, which are not fully sampled. The dendrogram structure for the PD (Fig. 4.19.d) is similar with clusters $n7$, $n23$, $n25$ organizing under the branch representing the central region and clusters $n1$, $n5$, $n21$, $n4$, $n16$, $n3$, $n12$ forming the eastern and western peripheral regions. Clusters $n1$ and $n5$ pertain to extremities of Region 3, which are not fully sampled. Appendix Figs. C.17 and C.18 depict zoomed-in sections of the sub-graphs relating to each of the top clusters that confirm the detected intra-regional variation for DM and PD, respectively.

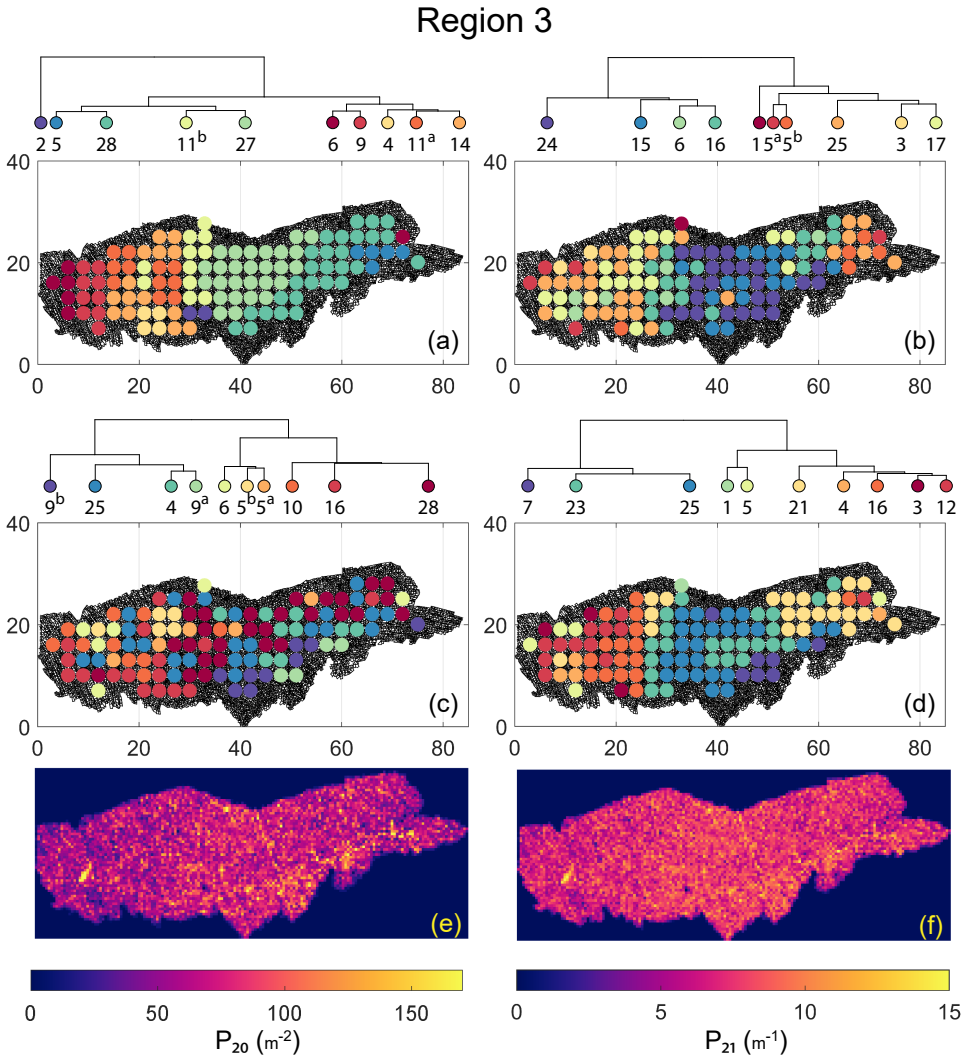

Figure 4.19: Hierarchical clustering results for Region 3 depicting the top 10 clusters using (a) Fingerprint distance (b) D-measure distance (c) NetLSD distance (d) Portrait Divergence distance (e) spatial P_{20} (f) spatial P_{21}

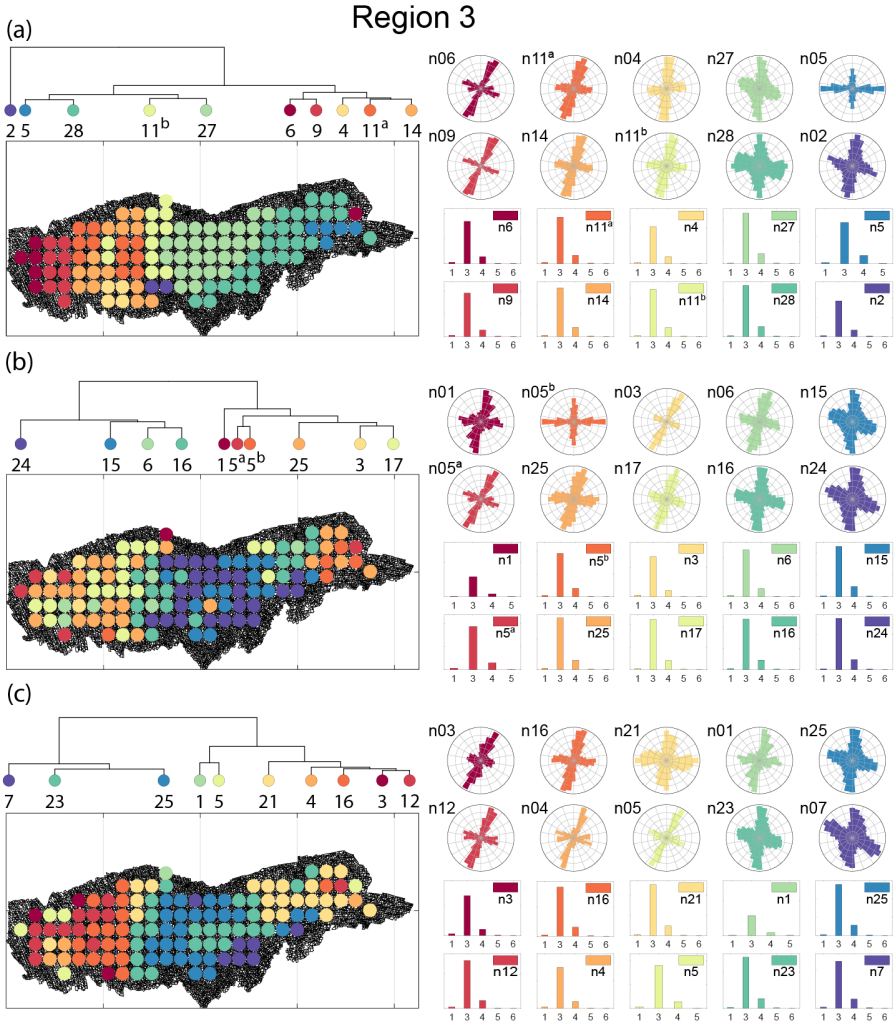

Figure 4.20: Variation in fracture orientations and topological summary for Region 3 corresponding to (a) Fingerprint (b) D-Measure (c) Portrait divergence

4.6. DISCUSSION

Within the structural geology literature, fracture persistence measures of Dershowitz and Herda [1992], the topological approach of Sanderson and Nixon [2015], and qualitative descriptions are most commonly to compare 2D fracture networks. The lack of quantitative measures for spatial network data is partially due to the lack of extensive 2D fracture trace data. Using fully mapped, extensive fracture networks, it is possible to systematically investigate network variations. By treating 2D fracture networks as planar graphs, applying graph similarity measures to quantitatively compare sub-samplings, and using hierarchical clustering, we are able to organize data into a hierarchy of patterns. The resulting dendrogram, derived from similarity scores, can be examined at multiple levels. One can argue that variation exists at multiple length scales, and more granular inquiry would lead to different clusters. In this section, we discuss some additional perspectives related to our methodology and results.

Linking spatial variation patterns to fracturing drivers: The results indicate that spatial variation in fracture networks is not always evident from the ubiquitously used fracture persistence measures, such as P_{20} and P_{21}. The proposed method highlights variations in network structure which can then help draw inferences into possible drivers for the spatial differences. In the case of Regions 2 and 3, the proximity to the fault influences network development. Such a model has been proposed by Peacock and Sanderson [1995], Gillespie et al. [2011b], and Wyller [2019], where the oldest fractures are long and radial, emanating from local asperities within the fault. These older fractures then influence the development of younger fractures. This is observed in Region 2, where clusters form roughly parallel to the E-N-E striking fault with the direction of variation to the N-W. Region 3 is positioned between two such asperity epicentres. There are long, radial fractures on the eastern and western extremities with a transition region in-between. The direction of cluster variation trends E-W. Fracture pattern variation in Region 1 is not affected by faulting. Since Regions 1 and 2, pertain to a single layer, the N-E regions of Region 1 show visual similarities between the westernmost extremities of Region 2. The intraregional variations in Region 1 could be due to layer thickness variation, although we do not have sufficient thickness data to confirm this. The analysis of spatial variation can assist in deciphering fracture timing. Given the temporal nature of network formation, it is desirable to delineate network evolution into relative episodes of fracturing. In previous analyses specific to the Lilstock dataset used in this contribution, Passchier et al. [2021] identified jointing sets with timing history based on fracture length, strike, and topological relationships. Although the temporal history is identified from joints that were picked manually but not wholly by Passchier et al. [2021], there is still a discernable spatial variation where some jointing sets are localized while others occur throughout the outcrop. Identifying spatial clustering in complete networks provides a basis by which joint sets can then be arranged in a hierarchy of temporal development.

On the choice of a graph distance metric: We have restricted our investigation scope to four state-of-the-art graph similarity distances from the recent graph theory literature. Many more graph distances applicable to spatial graphs exist [Hartle et al., 2020, Tantardini et al., 2019]; furthermore, the best means remain an open problem in network science research. Some novel distance measures are not graph-based but derive from

persistent homology (such as Feng, 2020). In this approach that considers the *shape of data*, persistence diagrams are generated from spatial graphs, and bottleneck distances are combined with hierarchical clustering to discover clusters. The results from Feng [2020] compared favourably to that of Louf and Barthelemy [2014] when applied to patterns of cities. As may be observed from our results, the metrics highlight certain aspects of the fracture network while not considering others. For instance, the fingerprint distance only considers block area and shape factor distributions of the blocks and neglects orientations. The other three distances use graph properties directly, and hence orientation information (or the lack of it) is a consequence of how the spatial graph is defined. We used weighted graphs that incorporate euclidean distance between nodes as edge weights for the similarity computations. However, each edge also has a striking attribute to completely describe its position in 2D space (in the case of 3D, it needs a dip). Ideally, the edge weight should then be a vector, $w = [l, \theta]$ incorporating both lengths, 'l' and orientation, 'θ', but the distance metrics we use do not allow the use of non-scalar weights.

Do REV's exist for fracture networks? In the context of fractured reservoir modelling, identification of a representative elemental volume (REV) aids continuum-based simulation approaches. However, the complexities of fluid flow and transport through fractured porous media require an explicit representation of fractures. Given the difficulties associated with obtaining realistic network geometries, stochastic-process-based methods derived from sparse fracture data are commonplace. However, these methods are often unable to represent inherent non-stationarity in spatial variation [Thovert et al., 2017], and work by Andresen et al. [2013] find that DFNs from nature exhibit disassortativity, which is not a property of generated networks. Other techniques based on multipoint statistics [Bruna et al., 2019b] attempt image-based approaches to modelling non-stationary networks. Estrada and Sheerin [2017] present a different approach in which DFNs are directly generated as spatial graphs (referred to as *random rectangular graphs*). Such a method can incorporate insights from outcrop-derived NFRs. Regardless of the method used to extrapolate, stationarity decisions have to be made based on hard data, and this is where our approach is helpful. We can use outcrop-derived networks to define and delineate stationarity's spatial boundaries and assign a particular type of network with due cognition of the inherent graph structure. Much literature exists on linking fracture patterns to high-deformation drivers such as folding, faulting, and diapirism, with the goal being to identify and correlate appropriate outcrop analogues to particular subsurface conditions. As our clustering results indicate, at the dimensional scales of sampling we have used, Tobler's first law of geography applies to fracture networks. Therefore, a representative network based on network similarity can be derived. The method can be applied to analogues for which data already exists. Further work is required to differentiate fluid-flow and transport responses of the identified cluster type.

Other clustering methods: We have used a combination of HC and graph distance metrics to delineate regions within a spatial graph and arrange them in a hierarchy of similarities. Within the graph theory literature, there are other non-HC methods based on graph properties such as modularity [Blondel et al., 2008, Newman and Girvan, 2004, Traag et al., 2019] or by graph spectral partitioning [Fiedler, 1973, Spielman and Teng,

2007]. Recent developments using graph neural networks and graph machine learning include modifications on the concept of modularity [Tsitsulin et al., 2020] and spectral methods [Bianchi et al., 2020] towards the goal of partitioning graphs into clusters.

4.7. CONCLUSION

This chapter presents a method to automatically identify spatial clusters and quantify intra-network spatial variation within 2D fracture networks. We test the technique on 2D trace data from a prominent limestone outcrop within the Lilstock pavements, located off the southern coast of the Bristol Channel, UK. The fracture network data that span three separate regions and cover over 14,200 sq.m are converted to the form of planar graph structures, spatially sampled into sub-graphs, and then compared using four different graph-distance measures. The pair-wise similarities in the form of distance matrices are used to arrange region-wise sub-graphs into a hierarchical relationship structure, also referred to as a dendrogram, using the statistical technique of hierarchical clustering. Positional order information from the dendrogram is used to render maps depicting the spatial variation within the fracture networks. The delineations of these intra-network sub-patterns provide a way to identify representative elemental volumes that preserve fracture networks' topological and geometric properties. The presence of these sub-regions can also serve as a guide to making decisions on stationarity with respect to geostatistical modelling. The main findings from this chapter are summarized as:

- representing fracture networks as graphs enable combining hierarchical clustering and graph-distance metrics to reveal interesting intra-network spatial similarity patterns not otherwise discernable from existing global or local fracture network descriptors.
- organization of fracture network sub-graphs based on pair-wise similarities into a hierarchical tree enables identification of spatial clustering at different dendrogram heights with newer and more granular cluster boundaries emerging at successively deeper levels of enquiry.
- spatial autocorrelation is more apparent with the fingerprint, D-measure, and the portrait divergence distances than the NetLSD, which yields speckled patterns with little or no spatial autocorrelation.
- spatial variation maps deriving from hierarchical clustering using the D-measure and portrait divergence identify similar spatial clusters and cluster boundaries. However, with the fingerprint distance, the cluster boundaries are different.
- fracture segment orientations show gradual variation in segment strikes across the identified clusters despite orientation not being explicitly considered and only euclidean distance being used to weight spatial graph edges.

5

SLAM LIDAR FOR 3D CAVE GEOMETRY DATA ACQUISITION

This chapter describes a novel approach to data acquisition in karstic caves using Simultaneous Localization and Mapping (SLAM) LIDAR technology. Many carbonate reservoirs are to some degree karstified, and there is evidence from well-bore logs, drilling data, seismic signatures, and well test data to indicate that large vugular cavities with dimensions of 10s of meters exist in the subsurface. The subsurface cavities are hypothesized to be either due to epigenetic karst created by meteoric water and then buried as paleokarst, or hypogenic karst developed due to subsurface mixing of dynamic porous fluid streams. Similar to fracture discontinuities in the subsurface, dissolution features also form spatial networks. Understanding the 3D structure of karstic cave networks is essential for dynamic modeling of karstic reservoirs where shape influences fluid transport and wellbores' geomechanical stability. SLAM Lidar is well suited for subterranean caves that suffer from conditions such as poor illumination and non-availability of GPS-signals, making georeferenced photogrammetry impossible. We showcase the application of SLAM LIDAR caves from Bahia, NE Brazil, where 3D point clouds are acquired and converted to 3D shape models. The methodology is useful for rapidly acquiring 3D shapes of speleogenetic features. These 3D models may be used as physical benchmark models for further developments in coupled Darcy-Stokes reservoir simulation and hydrodynamically coupled geomechanical stability analysis.

5.1. KARST AND ITS RELEVANCE IN RESERVOIR SYSTEMS

Carbonate reservoir rocks in the deep subsurface host around 60% and 40% of the world's oil and gas reserves respectively, with field data indicating that they are all karstified to some extent [Ford, 2007]. The hydrogeological phenomenon of karstification is a highly complex process which depends on several factors, such as regional geology, fluid-rock interaction, and flow- stress conditions. Surface karst systems are characterized by caves, enclosed depressions and opened fractures and their global distribution has been estimated around 10-20 % of the Earth's ice-free land area, roughly following the distribution of carbonate and gypsum rocks over the Earth's surface [Ford and Williams, 2007]. Understanding and monitoring geometries and behaviour of karsts acquires urgency owing to their other anthropogenic applications. Near sub-surface karst aquifers form drinking water sources for 25% of the world population. Karst also poses major hazards for constructions due to high probability of collapse and subsidence caused by sinkholes.

In the deep subsurface, which is relevant for oil and gas, karst is problematic due to the dissolution features forming complex shapes, organizing as spatial network structures, and possessing lateral variability. The combined effects result in poor 3D seismic quality and, therefore, poor predictability of karst occurrences, hampering the development of reservoir geo-models. Concerning reservoir development, drilling through karstic features causes significant problems like drilling fluid losses, drillstring failures, improper casing placement, and cementation integrity issues due to crooked hole geometry. Additionally, palaeo-karsts form complex pore networks resulting from collapsed karst caves coalescing within single or multiple karst cave systems. Karstic reservoirs are challenging to model as their multi-scale nature necessitates the mathematical implementation and coupling of different governing flow equations for porous continua, fractures, and vugular cavities.

The formation of karst, commonly referred to as karstification, is triggered by water-rock interactions, hydrogeological mass transport, and destruction of permeable soluble rocks [Dreybrodt et al., 2005]. Typical karst morphologies include pavements, karren, dolines (or sinkholes), poljes, cones (or towers), and caves. In cave formation (or speleogenesis), a significant distinction has been made based on genetic settings and is thus classified broadly as epigenic or hypogenic. The former develops in intimate interaction with the landscape, with both surface and underground components (under un-confining conditions) formed by downward-moving CO2, consuming meteoric water. Hypogene caves evolve without any direct genetic linkage with the surface (under confining conditions). Such caves form due to an upward-moving, cross-formational fluid flow that intersects with contrasting water chemistry, gas composition, and temperature. Hypogenic speleogenesis is defined as the formation of solution enhanced permeability structures by water that recharges the soluble formation from below [Klimchouk, 2011]. The fluids are driven either by hydrostatic pressure or other sources of energy like thermal or concentration gradients. The conventional thinking in epigenic speleogenesis does not explain many features routinely observed in karstic caves, such as dead ends, bedrock partitions, and the absence of elements like scallops and flutes specific to lateral flow. Therefore, discriminating between epigenic and hypogenic origins is possible using passage morphology, sedimentological, and mineralogical criteria [Klimchouk, 2009].

Hypogenic speleogenesis results in complex patterns guided by the original permeability structure of the soluble host rock, heavily influenced by the fracture patterns in the pre-speleogenetic rock matrix. Network and sponge work mazes, irregular chambers, isolated passages, and rising shafts are some of the elementary patterns observed in hypogenic caves. The combination of pre-existing fracture stratigraphy with hypogenic speleogenesis leads to the formation of multiple storeys of layered mazes with inlet feeders at the deepest levels and outflow cupolas at the shallowest levels. Many hypogenic systems that are uplifted and brought up to the epigenic realm acquire epigenic overprint due to the newly established un-confining flow patterns; however, there is a limited degree of integration due to the different organization, driving forces, and functioning of respective parent flow systems [Klimchouk et al., 2017].

Reservoir characterization and modelling of subsurface karst systems require a genetic interpretation of void-conduit systems and an understanding of their hydraulic functioning. The void conduit systems could have a positive influence when they enhance effective reservoir permeability and negatively influence reservoir development strategies in case of water breakthrough or short-circuiting due to insufficient oil sweep of the injected water. They could also compromise seal integrity, which may impact applications like geological sequestration of CO2. Given such challenges, there is a need for real-world geological analogues of karstic cave systems to guide reservoir modelling. Detailed analysis of accessible hypogene cave systems can serve as analogues for karst reservoirs and provide insight into the resultant porosity and permeability distribution. Models currently in use assume extremely simplified circular or oblong geometries for the vuggy cavities and often neglect the relationship between pre-existing fractures and the arrangement of cavities. Three-dimensional morphological data can bridge this gap.

Challenges unique to the cave environment hamper the mapping of cave systems for analogue studies. These include difficulties in illumination and the absence of geopositioning satellite signals. Acquiring quality datasets within cave environments has been attempted using stationary Lidar. Combined with artificial illumination, stationary LiDAR can generate 3D point clouds of cave systems fused with image data. The main disadvantage with this approach is the unwieldy nature of stationary LiDAR equipment, which is generally bulky and requires precise set-up. Such techniques are ill-suited to map narrow passages with relevant features. Mobile LiDAR is an alternative to stationary LiDAR to map cave geometry. Simultaneous Location and Mapping (SLAM) LiDAR is a compact and mobile version of the conventional stationary LiDAR, which a single operator can hold. The device continuously records point clouds of the cave geometry while the operator moves around regions of interest and stores the data onto portable data storage that the operator also carries. Accurately positioning the entire point cloud is done using georeferenced station points acquired outside the cave system.

3D mapping of hypogene caves' interior provides the unique intricate geometries useful in populating triple porosity (rock matrix, fractures, and void conduits) reservoir models. This workflow aims to rapidly obtain quantitative cave mapping data, which can extract geometric properties relevant for cave characterization. The digital geometries can serve either as direct input for geomechanical and reservoir fluid flow modelling or as input to populate stochastic karst cave models.

5.2. 3D POINT CLOUD DATA ACQUISITION USING SLAM LIDAR

2D cave maps provide limited quantitative information on caves' volumetrics, internal geometries, and associated length scales. Additional measurements inside the cave systems are necessary using 3D mapping. Illumination issues, limited space inside caves make photogrammetric reconstruction of 3D geometry difficult. LiDAR technology does not have this disadvantage. It relies on transmitting a beam of light from a source, which is reflected off the object(s) of interest, followed by a recording and calculation of the exact travel time. Thus the measurements are invariant to illumination.

However, conventional lidar equipment is cumbersome to set up and operate within confined spaces peculiar to cave settings. Mobile lidar mapping also referred to as Simultaneous Location and Mapping (SLAM), is a recent innovation in which data measurements are made with a portable moving sensor in a fraction of time compared with conventional terrestrial lidar. SLAM lidar is well-suited to cave settings due to the ease and speed at which the data is acquired. The measurements are obtained by a single operator holding the handheld device while walking through areas of interest.

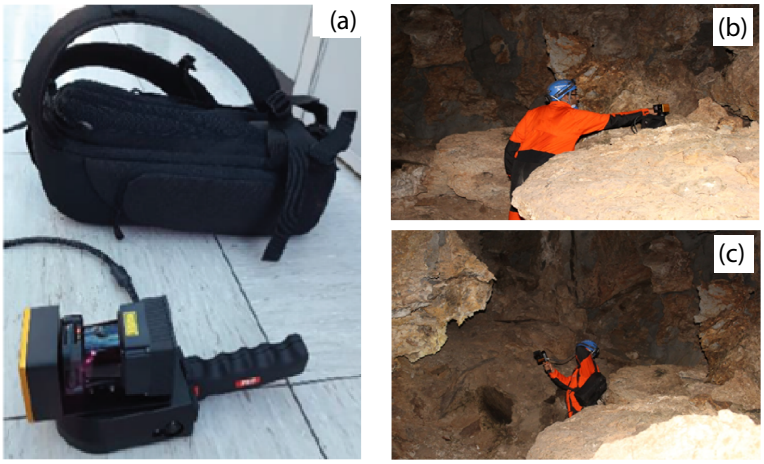

Figure 5.1: (a) The Geoslam Zeb Revo SLAM lidar device with sensor unit connected to battery and recording unit in a backpack (b) start of a mission at a fixed location (b) movement through the cave starting the data capture

We utilize the GeoSLAM ZebREVO portable LiDAR to acquire cave geometries. The system consists of a rotating scanning head that houses a 2D time-of-flight laser scanner and an inertial measurement unit. The scanning head connects to the data storage unit and battery through a short cable. Typically during operation, the user holds the sensor unit in one hand with the battery-pack and data storage units placed in a wearable backpack. The scanner unit weighs only around 1 kg. The lithium-polymer battery weighs around 600g and has a life of approximately 4 hours in continuous use. The configuration is depicted in Fig. 5.1(a) with operations depicted in Fig. 5.1(b)-(c).

The maximum range of the ZEB-REVO is 30 m with effective range in indoor conditions being 15-20 m. The scanner is able to record around 43,200 points/sec and has an

angular field of view of 270° × 360° ×. The maximum data storage capacity is 55 GB with scan data acquired at a rate of approximately 10 MB/min. Absolute position accuracy is between 3-30 cm.

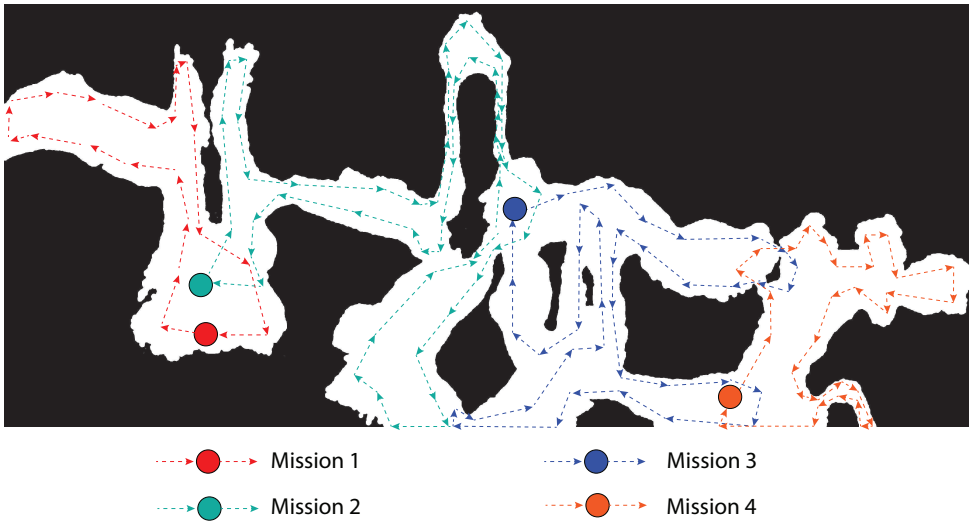

Figure 5.2: An example of four mission trajectories chosen for mapping a region of maze-like cave. The plan view is extraced from Calisto Cave. Brazil. Each trajectory closes a loop that starts and ends at a single point.

A typical data acquisition routine would consist of an initial reconnaissance of the cave passages of interest. The region to be covered is then divided into missions, with a pre-determined start-point for each mission. The data acquisition begins from such a start-point, and an operator then walks through the pre-decided trajectory at a constant slow speed. The walking pace is moderate so that there are enough repeat scans for the SLAM algorithm to convert laser data into point clouds. The trajectories are pre-planned such that the mission starts and ends at the same point, and each cave passage is covered twice in advancing and returning strides. When the routes are extensive and in the presence of intricate speleothems, loopy and zigzag paths may be chosen so that there are more points for correct geometry capture. Figure 5.2 depicts an example of a plan view of a maze-like cave with four mission trajectories. Each mission trajectory has some degree of overlap with that of other missions so that each mission's point-clouds can then be properly aligned and connected.

5.3. POINT CLOUD PROCESSING

Converting raw laser scan data captured using the ZebRevo®is done using GeoSLAM Hub®software. The data storage device is connected to a conventional laptop by a USB cable, and GeoSLAM Hub's interface can access and process the raw datasets. The conversion from raw scan to point clouds can be standard point cloud formats such as .las, .laz, .ply, .e57, and .ascii. The time taken for conversion to point clouds depends on the dataset's size (typically the length of a mission) and processing options. The point clouds

may be viewed within the interface and examined for completeness and coverage. This step can be performed while in the field on a laptop.

5.3.1. POINT CLOUD REGISTRATION AND ALIGNMENT

The point clouds within a single cave are initially disparate datasets. Merging of separate point-cloud datasets corresponding to each mission within a cave into a single point-cloud is done using CloudCompare®, an open-source point-cloud processing software. Identifying overlap between individual point sets is done using mission start-positions. By choosing the trajectories to produce such overlaps deliberately, this identification is made simple. The merging of two datasets is then done by manually picking a minimum of four common-points. Correctly orienting the merged point-cloud is done using GNSS information from GCPs recorded on the cave exterior.

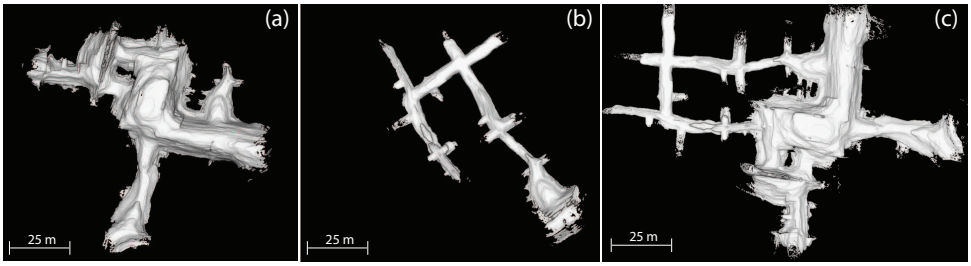

Figure 5.3: An example of merging and aligning two point clouds from Lapinha Cave, Brazil (a) plan view of point cloud data from first mission (b) plan view of point cloud data from second mission (c) merged and aligned point cloud

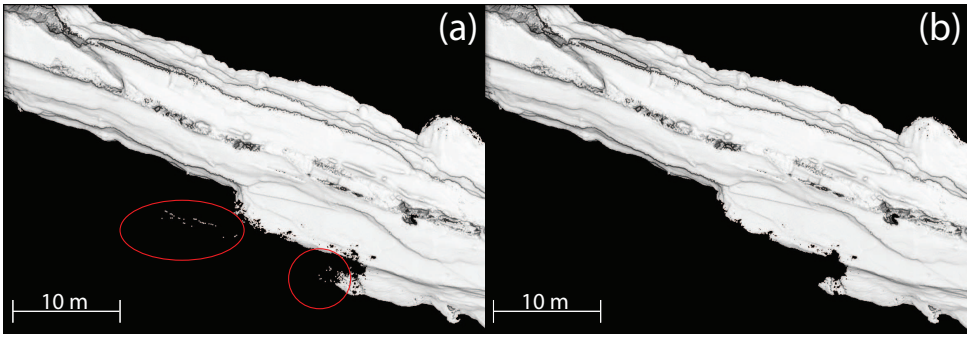

Figure 5.4: An example of de-noising from point cloud data of Ioio Cave, Brazil (a) point cloud prior to de-noising with isolated points highlighted (b) de-noised point cloud

The raw point cloud data may contain stray points which are spatially very far away from the main point clusters. A de-noising procedure can be carried out which removes such isolated data points and shrinks the dataset. An example of such a de-noising procedure is depicted in Fig. 5.4.

5.3.2. SURFACE WRAPPING OVER POINT CLOUDS

Using Geomagic Wrap®, we convert the merged, aligned, and de-noised point cloud data into 3D meshes. We use the *wrap* functionality within Geomagic Wrap® to fit a triangular mesh surface through the point cloud. Fitting the meshed surface in a free-form manner honours the natural shapes represented by the 3D point data. Depending upon point cloud size and the complexity of geometry, we can set the desired number of triangular polygons with a maximum of up to 100 million. Additional inputs are the number of holes to be retained within the final polygon object. Once a basic mesh is obtained, several options are available to rectify geometric issues such as self-intersections, small holes, spikes, and creased edges. We may also decimate the mesh to reduce the number of triangles if the mesh is too detailed.

Figure 5.5 depicts an example of a point-cloud converted to a 3D surface object. This meshed surface, also referred to as a polygon object, can be exported in various CAD formats such as .stl, .dxf, .iges that can be used as input for fluid-flow or geomechanical modelling. 3D models of caves serve as digital twins may be re-used in future cave expeditions targetting certain speleological features of interest.

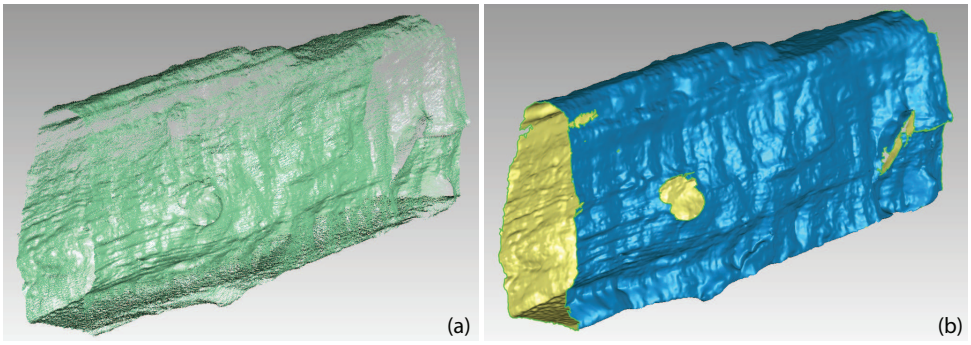

Figure 5.5: (a) a 3D point cloud data set sub-sampled from Cave Torrinha, Brazil consisting of 966798 points (b) point-cloud converted to a 3D surface object consisting of around 1.5 million triangles

5.4. DISCUSSION

5.4.1. ACCURACY AND RESOLUTION OF POINT CLOUDS

The digital cave models constructed from point clouds have a high degree of accuracy compared with manual mapping. Due to the absence of differential GNSS inside caves, accurate positioning needs a combination of cave data lidar measurements with digital outcrop models (UAV-borne photogrammetry or lidar), which are calibrated and ground-truthed with ground control points (GCPs). Mapping of entire caves may not be possible owing to data storage limitations but just small sections. Suppose the region of interest lies deep within the cave interiors. In that case, accurate positioning of the point clouds and generated 3D surfaces may not be possible without referring to manual measurements or 2D cave maps. In such a scenario, we may induce errors in the absolute positioning of 3D cave models. Therefore, recording a GCP at the cave entry location and performing a starter lidar survey that covers this georeferenced location

is recommended. Linking additional surveys within the cave to the first georeferenced survey can then be done.

5.4.2. COMPARISON WITH TERRESTRIAL LIDAR

Before the advent of portable SLAM-based lidar, terrestrial lidar was the preferred method to obtain 3D geometry of caves. However, its usage is limited to easily accessible and sufficiently spacious cave passages. The flexibility in use is limited by bulky equipment, challenging to position and move within complex cave passages. Terrestrial lidar can, however, yield much more accurate and detailed point clouds. The additional detail in terrestrial lidar is often unnecessary in terms of computations related to point cloud processing and conversion to polygonal meshes. Even with point clouds generated using SLAM-based lidar, de-noised and downsampled point clouds are often sufficient to build high-quality surface meshes. With SLAM-based lidar, once the data is acquired, the raw scan data can be converted into 3D point clouds on a conventional laptop in a matter of hours. Further processing and conversion to structural surfaces require more computational resources, which one cannot do in a field environment. Visualizing the raw point cloud data using open-source software and interesting speleological features can be identified for further investigations while at the fieldwork site.

5.5. CONCLUSION

The presented workflow combines acquiring point-cloud data from cave analogues using SLAM Lidar followed by point cloud processing in GeoMagic to generate realistic 3D cave geometry. SLAM lidar technology and point cloud processing have been used for indoor mapping and in digital manufacturing. By combining these, we present a workflow aimed at rapid cave digitization and generation of geometric datasets. Compared to alternative techniques of speleological fieldwork, this workflow has several advantages:

- the SLAM lidar is compact, wearable by an operator, and thereby ensures a very high degree of mobility within tortuous cave passages with minimal need for device setups and adjustments
- SLAM lidar can overcome the twin disadvantages of poor illumination and lack of GPS signals that make comprehensive data acquisition in cave environments difficult
- the point cloud data generated by SLAM lidar systems are near-perfect representations of cave geometry that eliminates and makes redundant the need for manual mapping
- the availability of real-time visualization of recorded points enable the operator to plan walk trajectories and selectively capture regions of interest within the cave environment. This feature also allows the user to ensure proper coverage of structurally intricate geometries such as speleothems
- the point cloud processing workflow generates water-tight shape features which is helpful for geometric data analysis, 3D modelling using realistic cave geometry

6

CONCLUSION

This Thesis advances state-of-the-art natural fracture characterization by utilizing automatic techniques for fracture tracing and introducing graph theory in analysing spatial variation in fracture networks. The key contributions are methodological and generally applicable to any fracture dataset. We have applied the automatic tracing method to three example carbonate outcrops in Parmelan (France), Brejões (Brazil), and Lilstock (UK) producing large-scale datasets that contain millions of fracture segments. By representing fracture networks as spatial graphs, we developed graph manipulation routines that convert the trace data into geologically meaningful fracture information. Using the spatial graph representation of fracture networks, we further utilize graph similarity measures and hierarchical clustering to analyse and quantify the natural spatial variability in fracture networks. This chapter summarizes the main findings and critically dwells upon the relevance and applications of the developed workflows and results.

6.1. AUTOMATING THE PROCESS OF TRACE EXTRACTION FROM IMAGES

Studying fracture networks from outcrops requires datasets that are multiscale, incorporative of network structure, and reasonably large enough in spatial extent to derive lessons for subsurface modelling. To efficiently acquire such datasets, we have developed an automated trace extraction workflow that can rapidly extract 2D fracture traces from UAV-derived photogrammetry (Chapter 2). The workflow is novel and consists of ridge detection using the complex shearlet transform followed by a series of image processing and vectorization steps that are flexible and need little or no manual intervention. The obtained fracture traces are unbiased and enable complete coverage of available image data hitherto underutilized due to manual interpretation bottlenecks.

Even though outcrops provide detailed and quantitative fracture geometry, we recognize that outcrop-derived networks may not always represent subsurface conditions. Our workflow in Chapter 2 primarily targets open-mode fractures that are easily visible on images acquired with the typical resolution of UAV cameras. In previous studies that

compared discontinuity patterns observed in the subsurface with corresponding outcroppings, mineralized vein and stylolite networks in both surface and subsurface have concurrence. We can adapt our workflow to such situations, provided there are images of sufficient resolution consistent with spatial scales of discontinuity networks of interest. One limitation is that the automated workflow requires prior shearlet construction to detect multiscale discontinuities in images. An improvement is by combining a few samples of derived traces in the form of training image data with a deep learning workflow outlined in Appendix F.

6.2. GRAPH-BASED REPRESENTATIONS AND CHARACTERIZATION

Applying the methodology in Chapter 2 to different outcrop settings, the results provide fracture segments of more than a million traces. Chapter 3 presents large-scale fracture networks automatically traced from five regions within the Lilstock pavement into spatial graphs. The graph representation treats fracture data not as objects in space but as relationships. Such a representation enables the use of graph theory. We developed a series of routines to represent fracture networks as primal and dual graphs, perform topological modifications, resolve issues such as artificial fragmentation, and automatically identify geologically significant fractures. Appendix B details these graph-based routines. Using spatial graph theory enables easy computation of many topological distributions such as node degree, cumulative length based on edge topology type, and block areas. We find that natural fracture networks are dominated globally by degree-3 nodes. The networks display disassortativity properties via a correlation between fracture length and several intersected fractures. In Appendix D, we further use the betweenness centrality metric corresponding to the Lilstock networks to illustrate network information's importance instead of sparse sampling. These 2D network properties provide essential information that one can implement within stochastic DFN generators based on point processes derived from sparse data.

6.3. SPATIAL VARIABILITY IN FRACTURE NETWORKS

The large spatial extent of automatically extracted fractures (made possible by methodology in Chapter 2) and the representation as spatial graphs (Chapter 3) allow the use of graph metrics to investigate intra- and inter-network spatial variability. In Chapter 4, we utilize both global graph properties and graph distances to identify spatial differences in fracture network organization. Graph distances provide heuristics to compare graphs based on inherent graph properties. By spatially sampling the large-scale networks obtained from Chapter 3, and by using the combination of graph distances with hierarchical clustering, we can uncover spatial patterns within fracture networks.

We use four graph distances, namely, fingerprint, D-measure, NetLSD, and portrait divergence. These measures convert a graph into probability distributions (fingerprint, D-measure, Portrait divergence) or signals (NetLSD), that encode some topological and geometric properties of the graphs. The distance measures then are a comparison of the extracted probability distribution or signal. We find that the selected graph metrics can highlight spatial patterns that are not discernable from other computed spatial properties such as fracturing intensity or density. The results indicate that a REV, which

is challenging to define otherwise for fractured media, can be approached using such a spatial variation investigation. Although the networks we consider are not representative of subsurface conditions, applying the methodology to properly identified and representative outcrops can provide useful rules for geostatistical or stochastic extrapolation to subsurface models.

The core of the approach to comparing graphs in Chapter 4 is based on the extraction of a suitable graph property representing a key geometric and topological feature underlying the fracture graph. In Appendix E, we present an implementation of simplicity profiling of graphs. The simplicity profile compares statistics of simplest paths versus shortest paths within a graph sample and converts this information into a normalized distribution. The shape of this profile is also beneficial as a comparison metric.

Like fracture networks, karstic cavities that form due to dissolution in porous media also display a network structure. Karst is abundant in carbonate rocks and often co-exist with fracture networks providing additional storativity and enhanced connectivity. In any case, pre-existing fracture networks offer preferential pathways for karstic pathway evolution. There remains a paucity of realistic and accurate 3D representations of cave passages. In Chapter 5, we present a workflow to acquire 3D geometric data of caves efficiently. Using portable lidar based on SLAM-algorithms and combined with point-cloud processing, a fast workflow is presented to obtain detailed, quantitative datasets of karst cave 3D patterns and intricate geomorphology of speleological features. The caves that we have explored are within carbonate rocks in Bahia, Brazil and they display a mixture of epigenic and hypogenic footprint. The 3D geometries serve as practical input for geomechanical and fluid flow modelling that considers explicit karst cavity representation.

6.4. SYNOPSIS

The workflows presented in Chapters 2 and 3 help to obtain truly large-scale, realistic natural fracture patterns from outcrops. The workflow in Chapter 5, is similarly applicable to karst cavity networks. The datasets generated provided valuable insight into the structure and spatial variability of fracture networks as presented in Chapter 4. The case studies of fracture networks (Chapter 4) show both layeral and intra-network differences in fracturing patterns. Given that fracture systems have a potentially large impact on subsurface permeability, it is important to account for these spatial variations in the context of fractured reservoir modelling. The datasets generated have widespread value in this regard, as one can use them to further test flow, transport, and geomechanical behaviour within explicit discontinuum models. Although spatial variation in high-deformation settings such as folding and faulting is well-recognized with many analogue outcrops, quantitative information in weakly-deformed rocks is lacking.

Although the fracture datasets are limited to 2D, and we do not consider stylolite or vein networks, they still provide relevant information that can guide the decision of stationarity bounds in geostatistical extrapolation and to determine the size of REVs. This Thesis presents an integrated workflow that is generally applicable to any dataset. This data-driven framework based on automated data processing is expandable and improvable in many dimensions, such as the use of graph-based DFN generation and graph-based reservoir simulation.

A

OVERVIEW OF THE COMPLEX SHEARLET TRANSFORM

This appendix serves as supplement to Chapter 2.

SHEARLET SYSTEMS

A shearlet generating function consists of an anisotropic scaling matrix and a shear matrix. Let the shearlet generating function be:

$$\psi \in L^2\left(\mathbb{R}^2\right) \tag{A.1}$$

The admissibility criteria for the shearlet generating function is :

$$\int_{R^2} \frac{\left|\hat{\psi}(\xi_1\xi_2)\right|^2}{{\xi_1}^2}\, d\xi_2 d\xi_1 < \infty \tag{A.2}$$

where $\hat{\psi}$ is the 2D fourier transform of ψ.

A shearlet satisfying Eq.A.2 is an admissible shearlet or a continuous shearlet [Kutyniok and Labate, 2012]. The admissibility condition implies that a reconstruction formula exists for the associated continuous shearlet transform. In order to achieve an optimally sparse approximation of an image that possesses anisotropic singularities, the analysing elements must consist of waveforms that range over several scales, orientations, and locations with the ability to become very elongated. To this end, a combination of a scaling operator to generate elements at different scales, an orthogonal operator to change orientations, and a translation operator to displace elements over the 2D plane, is used. The scaling matrix A_a is defined as [Labate et al., 2005]:

$$A_a = \begin{pmatrix} a & 0 \\ 0 & a^\alpha \end{pmatrix},\ \alpha \in [0,1]$$

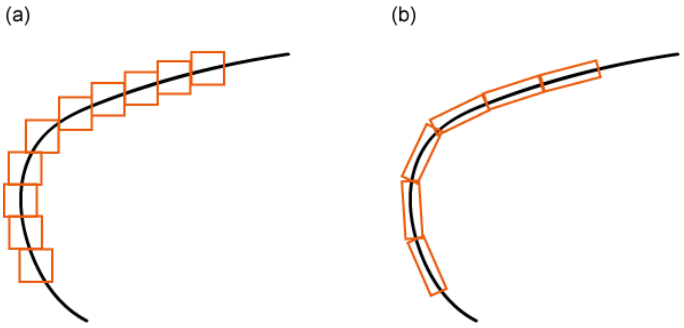

Figure A.1: (a). Isotropic elements capturing a discontinuity curve (b). Sheared, anisotropic elements capturing a discontinuity curve [modified from Kutyniok and Labate, 2012]

The value of α controls the degree of anisotropy [see Grohs et al. 2016 for information on anisotropy scaling molecules or α–molecules]. The scaling matrix is parabolic when $\alpha = \frac{1}{2}$.

An orthogonal transformation to change the orientations of waveforms. Rotation operators are not preferred as they destroy the structure of the integer lattice Z^2 whenever the rotation angle is different from $0, \pm\frac{\pi}{2}, \pm\pi, \pm\frac{3\pi}{2}$. Changes in the structure of integer lattice is problematic when transitioning from continuum to digital setting. Hence, a shearing transformation is used where the anisotropic shearing transformation matrix S_s are defined as:

$$S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \text{ where the parameters } a \in \mathbb{R}^+, s \in \mathbb{R}$$

The shearing matrix S_s preserves the structure of the integer grid for any $s \in \mathbb{N}$. The shearing matrix parametrizes orientations using the variable s associated with slopes rather than angles and leaves the integer lattice invariant, provided s is an integer. The difference between isotropic and anisotropic dilation with shearing is depicted in Fig. A.1a and Fig. A.1b).

A shearlet system is defined as [Kutyniok and Labate, 2012]:

$$SH(\psi) = \left\{ \psi_{a,s,t} = a^{-3/4} \psi \left(A_a{}^{-1} S_s{}^{-1} (\cdot - t) \right) a \in \mathbb{R}^+, s \in \mathbb{R}, t \in \mathbb{R}^2 \right\} \tag{A.3}$$

where $(\cdot - t)$ denotes the translation by a point t.
The corresponding shearlet transform for mapping a function $f \in L^2(\mathbb{R}^2)$ into coefficients, $SH_\psi f(a,s,t)$ specified by scaling a, shearing s and translation t is given by:

$$f \rightarrow SH_\psi f(a,s,t) = f, \psi_{a,s,t} \tag{A.4}$$

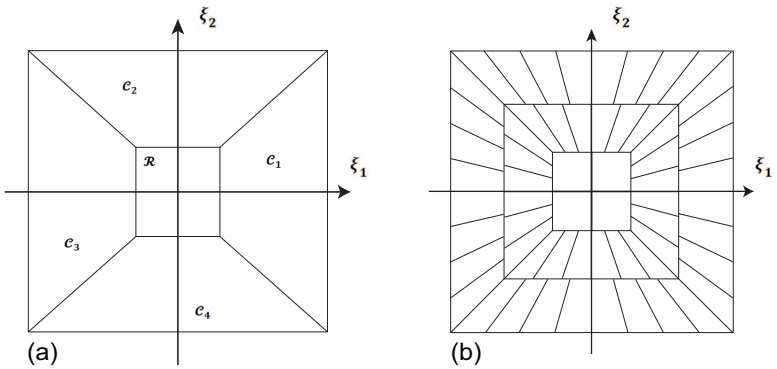

Figure A.2: The cone adapted continuous shearlet system (a) Bias in directions is handled by dividing the frequency plane into 4 cones $\mathscr{C}_1, \mathscr{C}_2, \mathscr{C}_3, \mathscr{C}_4$ and a square low frequency box region in the centre $\mathscr{R}$ (b) Trapezoidal shaped wedge tiling of the frequency induced domain induced by the shearlet transform [modified after Kutyniok and Labate 2012]

Cone Adapted Continuous Shearlet Systems

Equation (A.4) renders horizontal shearlets elongated at very fine scales, which is problematic in digital implementations. Because the shearing operator can range over a non-bounded interval, directions are not treated uniformly. To overcome this drawback of shearing, the cone adapted shearlet system was introduced in which the frequency plane is split into a horizontal and vertical cone that restricts the shear parameter to bounded intervals (see Fig. A.2.a). Dividing the frequency plane in such a manner ensures uniform treatment of directions [Guo et al., 2005, Kutyniok and Labate, 2012]. A cone adapted shearlet system can be tiled by further division of the frequency domain. Such a tiling configuration (see Fig. A.2.b) ensures that all directions are treated "almost equally" [Kutyniok and Labate, 2012]. There is still small, but controllable bias in the coordinate axes directions). The cone adapted shearlet systems can therefore be expressed as the union of a horizontal cone, a vertical cone, and a low-frequency centre component. The frequency plane is thus split into four horizontal and vertical cones with a low-frequency square region in the centre. The low-frequency region is given by the relation [Kutyniok and Labate, 2012]:

$$\mathscr{R} = \{(\xi_1, \xi_1) : |\xi_1|, |\xi_2| \leq 1\} \tag{A.5}$$

Inside each cone, the shearing variable s is only allowed to vary over a finite range. This produces elements with uniformly distributed orientations. The union of the generating functions for the horizontal cones $\psi \in L^2(\mathbb{R}^2)$, vertical cones $\tilde{\psi} \in L^2(\mathbb{R}^2)$ and for the square low frequency region $\varphi \in L^2(\mathbb{R}^2)$ is expressed as [Kutyniok and Labate, 2012]:

$$SH(\varphi, \psi, \tilde{\psi}) = \Phi(\varphi) \cup \Psi(\psi) \cup \tilde{\Psi}(\tilde{\psi}) \tag{A.6}$$

where

$$\Phi(\varphi) = \{\varphi_t = \varphi(\cdot - t) : t \in \mathbb{R}^2\}; \tag{A.7}$$

$$\tilde{\Psi}(\psi) = \left\{\tilde{\psi}_{a,s,t} = a^{-\frac{3}{4}} \tilde{\psi}\left(\tilde{A}_a^{-1} S_s^{-1}(\cdot - t)\right) : a \in (0,1], |s| \leq 1 + a^{\frac{1}{2}}, \ t \in \mathbb{R}^2\right\}; \tag{A.8}$$

$$\tilde{\Psi}(\psi) = \left\{\tilde{\psi}_{a,s,t} = a^{-\frac{3}{4}}\tilde{\psi}\left(\tilde{A}_a^{-1}S_s{}^{-1}(\cdot - t)\right) \colon a \in (0,1], |s| \leq 1 + a^{\frac{1}{2}},\ t \in \mathbb{R}^2\right\}. \tag{A.9}$$

Scaling matrix for vertical cones, $\tilde{A}_a$ is expressed as:

$$\tilde{A}_a = \begin{pmatrix} a^{\alpha} & 0 \\ 0 & a \end{pmatrix} \tag{A.10}$$

The cone adapted continuous shearlet transform is expressed as the mapping:

$$f \to SH_{\varphi,\psi,\tilde{\psi}}\, f\left(t', (a,s,t), (\tilde{a},\tilde{s},\tilde{t})\right) = \left(f,\ \varphi_{t'}, f,\ \psi_{a,s,t}, f,\ \tilde{\psi}_{\tilde{a},\tilde{s},\tilde{t}}\right) \tag{A.11}$$

THE DISCRETE CONE ADAPTED SHEARLET SYSTEM

A discrete version of the cone adapted shearlet system may be defined with scaling parameter j, shearing parameter k, and translation parameter m for a sampling factor of $c = (c_1, c_2) \in (\mathbb{R}_+)^2$. Similar to Eq. (A.6) this is a union of the generating functions for vertical, horizontal, and low frequency central region.

$$SH(\varphi,\psi,\tilde{\psi};c) = \Phi(\varphi;c_1) \cup \Psi(\psi;c) \cup \tilde{\Psi}(\tilde{\psi};c) \tag{A.12}$$

$$\Phi\left(\varphi;c_1\right) = \left\{\varphi_m = \varphi\left(\cdot - c_1 m\right) : m\epsilon\mathbb{Z}^2\right\}; \tag{A.13}$$

$$\Psi\left(\psi;c\right) = \left\{\psi_{j,k,m} = 2^{\frac{3}{4}j}\psi\left(S_k A_{2^j} \cdot - M_c m\right) \colon\ j \geq 0, |k| \leq \left\lceil 2^{\frac{j}{2}}\right\rceil,\ m \in \mathbb{Z}^2\right\}; \tag{A.14}$$

$$\tilde{\Psi}\left(\tilde{\psi};c\right) = \left\{\tilde{\psi}_{j,k,m} = 2^{\frac{3}{4}j}\tilde{\psi}\left(S_k{}^T \tilde{A}_{2^j} \cdot - \tilde{M}_c m\right) \colon\ j \geq 0, |k| \leq \left\lceil 2^{\frac{j}{2}}\right\rceil,\ m \in \mathbb{Z}^2\right\}; \tag{A.15}$$

with $M_c = \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}$; $\tilde{M}_c = \begin{bmatrix} c_2 & 0 \\ 0 & c_1 \end{bmatrix}$;

(M_c and $\tilde{M}_c$ are sampling matrices for horizontal and vertical cones)

$A_{2^j} = \begin{bmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{bmatrix}$; $\tilde{A}_{2^j} = \begin{bmatrix} 2^{j/2} & 0 \\ 0 & 2^j \end{bmatrix}$;

(A_{2^j} and $\tilde{A}_{2^j}$ are dyadic scaling matrices for horizontal and vertical cones)
, and $S_k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$ (shearing matrix).

The discrete cone adapted shearlet transform associated with ϕ, ψ and $\tilde{\psi}$ is given by the mapping,

$$f \to SH_{\varphi,\psi,\tilde{\psi}}\, f\left(m', (j,k,m), (\tilde{j},\tilde{k},\tilde{m}\right) = \left(f,\ \varphi_{m'}, f,\ \psi_{j,k,m}, f,\ \tilde{\psi}_{\tilde{j},\tilde{k},\tilde{m}}\right). \tag{A.16}$$

THE COMPLEX DISCRETE CONE ADAPTED SHEARLET SYSTEM

Taking the complex-valued wavelet of a real-valued even-symmetric wavelet generator $\psi^{even} \in L^2(\mathbb{R}^2)$, using the Hilbert transform operator ($\mathscr{H}$), a complex valued shearlet generator is obtained [from King et al., 2015, Reisenhofer, 2014]

$$\psi^c = \psi^{even} + i\,\psi^{odd}. \tag{A.17}$$

The complex-valued function can be written in terms of a Hilbert transform pair of an even-symmetric real valued shearlet and an odd-symmetric real valued shearlet: [from King et al., 2015, Reisenhofer, 2014]

$$\psi^c = \psi^{even} + i\,\mathscr{H}\psi^{even}. \tag{A.18}$$

The Hilbert transform operator is written as,

$$\mathscr{H}(f)(t) = \lim_{a\to\infty} \int_{-a}^{a} \frac{f(\tau)}{t-\tau} d\tau. \tag{A.19}$$

The discrete cone adapted complex shearlet system is given as [King et al., 2015, Reisenhofer, 2014]:

$$SH(\varphi,\psi,\tilde{\psi};c) = \Phi(\varphi;c_1) \cup \Psi(\psi;c) \cup \tilde{\Psi}(\tilde{\psi};c) \tag{A.20}$$

and

$$SH^c(\varphi,\psi,\tilde{\psi};c) = \Phi(\varphi;c_1) \cup \Psi^c(\psi;c) \cup \tilde{\Psi}^c(\tilde{\psi};c) \tag{A.21}$$

where,

$$\Phi(\varphi;c_1) = \left\{\varphi_t = \varphi(\cdot - c_1 m) : m \in \mathbb{Z}^2\right\}, \tag{A.22}$$

$$\Psi^c(\psi;c) = \left\{{\psi^c}_{j,k,m} = {\psi^c}_{j,k,m} + i\,(\mathscr{H}_{(1,0)}\mathrm{T}\,\psi)_{j,k,m} :\ j \geq 0, |k| \leq \lceil 2^{\frac{j}{2}} \rceil,\ m \in \mathbb{Z}^2\right\}, \tag{A.23}$$

$$\tilde{\Psi}^c(\tilde{\psi};c) = \left\{\tilde{\psi}^c_{j,k,m} = \tilde{\psi}^c_{j,k,m} + i\,(\mathscr{H}_{(0,1)}\mathrm{T}\,\tilde{\psi})_{j,k,m} :\ j \geq 0, |k| \leq \lceil 2^{\frac{j}{2}} \rceil,\ m \in \mathbb{Z}^2\right\}. \tag{A.24}$$

Correspondingly the discrete complex cone adapted shearlet transform is given by the mapping,

$$f \to {SH^c}_{\varphi,\psi,\tilde{\psi}}\, f(m',(j,k,m),(\tilde{j},\tilde{k},\tilde{m})) = \left(f,\ \varphi_{m'},\ f,\ {\psi^c}_{j,k,m},\ f,\ \tilde{\psi}^c_{\tilde{j},\tilde{k},\tilde{m}}\right). \tag{A.25}$$

EDGE AND RIDGE DETECTION USING THE COMPLEX SHEARLET TRANSFORM

The behavior of the coefficients of the even symmetric and odd symmetric shearlets can be used to detect edges and ridges. An edge measure for an image $f \in L^2(R^2)$, a location $x \in R^2$ and a shear parameter s is given as,

$$E_\psi(f,x,s) = \frac{\left|\sum_{a\epsilon A} Im\left(f,{\psi^c}_{a,s,x}\right)\right| - \sum_{a\epsilon A}\left|Re\left(f,{\psi^c}_{a,s,x}\right)\right|}{|A|\, max_{a\epsilon A}\left|Im(f,{\psi^c}_{a,s,x})\right| + \varepsilon}, \tag{A.26}$$

where $A \subset \mathbb{R}^+$ is a set of scaling parameters, ψ is a real valued symmetric shearlet and ϵ prevents division by zero [King et al., 2015, Reisenhofer, 2014]. The complex shearlet based edge measure can give approximations of the tangential directions of an edge. A line measure or ridge measure is obtained by interchanging the role of the even symmetric and odd symmetric shearlets [King et al., 2015, Reisenhofer, 2014],

$$L_\psi(f,x,s) = \frac{\left|\sum_{a\epsilon A} Re\left(f,{\psi^c}_{a,s,x}\right)\right| - \sum_{a\epsilon A}\left|Im\left(f,{\psi^c}_{a,s,x}\right)\right|}{|A|\, max_{a\epsilon A}\left|Re(f,{\psi^c}_{a,s,x})\right| + \varepsilon}. \tag{A.27}$$

Both the edge and ridge measures given above are inspired from the phase congruency measure of Kovesi [2000]. The edge and ridge measures are almost contrast invariant.

Table A.1: Shearlets

Shearlet System	waveletEffSupp	gaussianEffSupp	scalesPerOctave	shearLevel	scales	nShearlets	alpha
1	200	100	1	2	1	36	0
2	125	63	1	2	1	36	0
3	84	42	1	2	1	36	0
4	67	34	1	2	1	36	0
5	200	100	2	2	7	12	0
6	125	63	2	2	7	12	0
7	84	42	2	2	7	12	0
8	67	34	2	2	7	12	0
9	200	100	3	2	10.5	12	0
10	125	63	3	2	10.5	12	0
11	84	42	3	2	10.5	12	0
12	67	34	3	2	10.5	12	0
13	200	100	4	2	14	12	0
14	125	63	4	2	14	12	0
15	84	42	4	2	14	12	0
16	67	34	4	2	14	12	0
17	200	100	1	3	3.5	20	0
18	125	63	1	3	3.5	20	0
19	84	42	1	3	3.5	20	0
20	67	34	1	3	3.5	20	0
21	200	100	2	3	7	20	0
22	125	63	2	3	7	20	0
23	84	42	2	3	7	20	0
24	67	34	2	3	7	20	0
25	200	100	3	3	10.5	20	0
26	125	63	3	3	10.5	20	0
27	84	42	3	3	10.5	20	0
28	67	34	3	3	10.5	20	0
29	200	100	4	3	14	20	0
30	125	63	4	3	14	20	0
31	84	42	4	3	14	20	0
32	67	34	4	3	14	20	0
33	200	100	1	4	3.5	36	0
34	125	63	1	4	3.5	36	0
35	84	42	1	4	3.5	36	0
36	67	34	1	4	3.5	36	0
37	200	100	2	4	7	36	0
38	125	63	2	4	7	36	0
39	84	42	2	4	7	36	0
40	67	34	2	4	7	36	0
41	200	100	3	4	10.5	36	0
42	125	63	3	4	10.5	36	0
43	84	42	3	4	10.5	36	0
44	67	34	3	4	10.5	36	0
45	200	100	4	4	14	36	0
46	125	63	4	4	14	36	0
47	84	42	4	4	14	36	0
48	67	34	4	4	14	36	0
49	200	100	1	2	3.5	12	0.5
50	125	63	1	2	3.5	12	0.5
51	84	42	1	2	3.5	12	0.5
52	67	34	1	2	3.5	12	0.5
53	200	100	2	2	7	12	0.5
54	125	63	2	2	7	12	0.5
55	84	42	2	2	7	12	0.5
56	67	34	2	2	7	12	0.5
57	200	100	3	2	10.5	12	0.5
58	125	63	3	2	10.5	12	0.5
59	84	42	3	2	10.5	12	0.5
60	67	34	3	2	10.5	12	0.5
61	200	100	4	2	14	12	0.5
62	125	63	4	2	14	12	0.5
63	84	42	4	2	14	12	0.5
64	67	34	4	2	14	12	0.5
65	200	100	1	3	3.5	20	0.5
66	125	63	1	3	3.5	20	0.5
67	84	42	1	3	3.5	20	0.5
68	67	34	1	3	3.5	20	0.5
69	200	100	2	3	7	20	0.5
70	125	63	2	3	7	20	0.5

B

GRAPH ROUTINES

This appendix contains a list of algorithmic routines to modify graphs as described in Chapter 3. The shearlet parameters used to create ensembles of ridge maps are also listed.

Algorithm 2 Topological Graph Manipulation Type 1

$g, xy \leftarrow$ Input Graph and Spatial Positioning Matrix
$dt \leftarrow$ Perform Delaunay Triangulation around Graph Nodes (xy)
$A \leftarrow$ Compute tri-element areas dt
$a \leftarrow$ Input tri-element area threshold
for $i = 1$ to length $(A < a)$ **do**
  $m \leftarrow$ vertex attachment of tri-elements $(A < a)$ with degree 2 (if they exist)
  $s \leftarrow$ vertex attachment of tri-elements $(A < a)$ with degree 1 (if they exist)
  $t \leftarrow$ target node of edge in graph g containing node i
  Remove edge $s - t$ from graph g
  Remove node s from graph g
  Add edge $m - t$ from graph g
end for

Algorithm 3 Topological Graph Manipulation Type 2

$g, xy \leftarrow$ Input Graph and Spatial Positioning Matrix
$dt \leftarrow$ Perform Delaunay Triangulation around Graph Nodes (xy)
$A \leftarrow$ Compute tri-element areas dt
$a \leftarrow$ Input tri-element area threshold
for $i = 1$ to length $(A < a)$ **do**
  $c, i1, i2, i3 \leftarrow$ centroid and vertex elements of tri-element $(A < a)$
  $t1, t2, t3 \leftarrow$ target nodes of edges in graph g containing nodes $i1, i2, i3$
  Remove edges $i1 - i1, i2 - i2, i3 - i3$ from graph g
  Remove nodes $i1, i2, i3$ from graph g
  Add edges $c - t1, c - t2, c - t3$ to graph g
end for

Algorithm 4 Topological Graph Manipulation Type 3

$g, xy \leftarrow$ Input Graph and Spatial Positioning Matrix
$dt \leftarrow$ Perform Delaunay Triangulation around Graph Nodes (xy)
$A \leftarrow$ Compute tri-element areas dt
$ar \leftarrow$ Input tri-element aspect ratio
for $i = 1$ to length $(A < ar)$ **do**
  $m1, m2 \leftarrow$ vertex attachment of tri-elements $(A < ar)$ with degree 2 (if they exist)
  $t1m1, t2m1, t1m2, t2m2 \leftarrow$ target nodes of edges in graph g containing nodes $m1, m2$
  $p \leftarrow$ midpoint on line connecting $m1, m2$
  Remove edges $m1 - t1m1, m1 - t2m1, m2 - t1m2, m2 - t2m2$ from graph g
  Remove nodes $m1, m2$ from graph g
  Add edges $p - t1m1, p - t2m1, p - t1m2, p - t2m2$ to graph g
end for

Algorithm 5 Resolving Artificial Fragmentation

$g, xy \leftarrow$ Input Graph and Spatial Positioning Matrix
$N2 \leftarrow$ Nodes in graph g having degree 2
for $i = 1$ to length $(N2)$ **do**
  $n1, \leftarrow$ find neighbours of N2 having degree 2 looking forward
  $n2 \leftarrow$ find neighbours of N2 having degree 2 looking backward
  $PATH \leftarrow [n1 - N2_i - n2]$
  while exists$(n1)$ OR exists$(n2)$ **do**
    $n1, \leftarrow$ find neighbours of n1 having degree 2 looking only forward (if exists)
    $n2 \leftarrow$ find neighbours of n2 having degree 2 looking only backward (if exists)
    $PATH \leftarrow [n1 - PATH - n2]$
  end while
end for

Algorithm 6 Resolving step-outs

$[g], [xy] \leftarrow$ Input Graph and Spatial Positioning Matrix
$E_{3-3} \leftarrow$ sort for edges in graph g having degree 3 at both start and end nodes
$t \leftarrow$ threshold length
$E_{3-3m} \leftarrow E_{3-3} < t$: edges that can be merged
$E_{3-3f} \leftarrow E_{3-3} > t$: edges that can be flattened
for $i = 1$ to length E_{3-3m} **do**
  $f_1, f_2 \leftarrow$ neighbour nodes of edge $E_{3-3m}(i)$ looking forward
  $b_1, b_2 \leftarrow$ neighbour nodes of edge $E_{3-3m}(i)$ looking backward
  remove edges $E_{3-3m}(i)$, $f_1 - s$, $f_2 - s$ from $[g]$
  remove source node of edge $E_{3-3m}(i) : s$
  add edges $f_1 - E_{3-3m}(i) : t$, $f_2 - E_{3-3m}(i) : t$ to $[g]$
end for
$\phi \leftarrow$ strike threshold for flattening
for $i = 1$ to length E_{3-3f} **do**
  $w_{s1}, w_{s2} \leftarrow$ walks starting from source node $E_{3-3f}(i) : s$ and away from $E_{3-3f}(i)$
  $w_{t1}, w_{t2} \leftarrow$ walks starting from tail node $E_{3-3f}(i) : t$ and away from $E_{3-3f}(i)$
  $C \leftarrow combs(w_{s1}, w_{s2}, w_{t1}, w_{t2}, +/-\phi)$ combinations of walk pairs that are alignable
  $C_{max} \leftarrow max(length(C))$ longest combination that is within $+/-\phi$
  $E_{3-3f}(i) : s' \leftarrow$ find geometric point that is perpendicular from ($E_{3-3f}(i) : s$ or $E_{3-3f}(i) : t$) to line between $C_{max} : end_1$ and $C_{max} : end_2$
  remove edges $E_{3-3f}(i) : s - E_{3-3f}(i) : t$ and $C_{max} : end_1 - E_{3-3f}(i) : s$ from $[g]$
  remove node $E_{3-3f}(i) : s$ from $[g]$
  add node $E_{3-3f}(i) : s'$ to $[g]$
  add edges $E_{3-3f}(i) : s' - E_{3-3f}(i) : t$ and $E_{3-3f}(i) : s' - C_{max} : end_1$ to $[g]$
end for

Algorithm 7 Straightening fracture traces

$g, xy, PATHS \leftarrow$ Input Graph, Spatial Positioning Matrix, and Segmented Fracture Paths
$N2 \leftarrow$ nodes in graph with degree 2
for $i = 1$ to length $(PATHS)$ **do**
  $s, e \leftarrow$ source and end nodes of path
  add edge $s - e$ to graph
end for
remove nodes $N2$ from graph
$xy \leftarrow$ remove points pertaining to nodes $N2$ and update spatial positioning matrix

Algorithm 8 Geologically significant fracture traces

$g, xy \leftarrow$ Input Graph, Spatial Positioning Matrix
$alpha \leftarrow$ Set threshold for fracture strike
$E \leftarrow$ counter of all edges in graph, edges(g)
for $i = 1$ to length $(edges(g))$ **do**
  $s, e \leftarrow$ source and end nodes for edge
  $edges_s \leftarrow$ possible edges emanating from source node s excluding edge $s-e$ having a fracture strike within threshold $alpha$
  $edges_e \leftarrow$ possible edges emanating from end node e excluding edge $s-e$ having a fracture strike within threshold $alpha$
  $walks \leftarrow$ appending $edges_s$ and $edges_e$ to $walks$
  book keeping on E by removing edges within $walks$ to avoid repetition
end for

Algorithm 9 Primal to Dual

$[g_p], [xy], F \leftarrow$ Input Graph, Spatial Positioning Matrix, and edge sequence corresponding to geologically relevant fractures
$A_d \leftarrow$ initialize empty sparse adjacency matrix of length(F) Xlength(F)
for $i = 1$ to length (F) **do**
  $N \leftarrow$ all neighbour nodes connected to nodes in F_i
  $f_{id} \leftarrow$ indices of entries in F that either start or terminate with nodes in N
  fill $A_d[i.*length(f_{id})f_{id}]$ and $A_d[f_{id}i.*length(f_{id})]$ with ones
  create $[g_d]$ using A_d
end for

Table B.1: Shearlet System 1

Shearlet System	waveletEffSupp	gaussianEffSupp	scalesPerOctave	shearLevel	scales	nShearlets	alpha
1	200	100	1	2	1	36	0
2	125	63	1	2	1	36	0
3	84	42	1	2	1	36	0
4	67	34	1	2	1	36	0
5	200	100	2	2	7	12	0
6	125	63	2	2	7	12	0
7	84	42	2	2	7	12	0
8	67	34	2	2	7	12	0
9	200	100	3	2	10.5	12	0
10	125	63	3	2	10.5	12	0
11	84	42	3	2	10.5	12	0
12	67	34	3	2	10.5	12	0
13	200	100	4	2	14	12	0
14	125	63	4	2	14	12	0
15	84	42	4	2	14	12	0
16	67	34	4	2	14	12	0
17	200	100	1	3	3.5	20	0
18	125	63	1	3	3.5	20	0
19	84	42	1	3	3.5	20	0
20	67	34	1	3	3.5	20	0
21	200	100	2	3	7	20	0
22	125	63	2	3	7	20	0
23	84	42	2	3	7	20	0
24	67	34	2	3	7	20	0
25	200	100	3	3	10.5	20	0
26	125	63	3	3	10.5	20	0
27	84	42	3	3	10.5	20	0
28	67	34	3	3	10.5	20	0
29	200	100	4	3	14	20	0
30	125	63	4	3	14	20	0
31	84	42	4	3	14	20	0
32	67	34	4	3	14	20	0
33	200	100	1	4	3.5	36	0
34	125	63	1	4	3.5	36	0
35	84	42	1	4	3.5	36	0
36	67	34	1	4	3.5	36	0
37	200	100	2	4	7	36	0
38	125	63	2	4	7	36	0
39	84	42	2	4	7	36	0
40	67	34	2	4	7	36	0
41	200	100	3	4	10.5	36	0
42	125	63	3	4	10.5	36	0
43	84	42	3	4	10.5	36	0
44	67	34	3	4	10.5	36	0
45	200	100	4	4	14	36	0
46	125	63	4	4	14	36	0
47	84	42	4	4	14	36	0
48	67	34	4	4	14	36	0
49	200	100	1	2	3.5	12	0.5
50	125	63	1	2	3.5	12	0.5
51	84	42	1	2	3.5	12	0.5
52	67	34	1	2	3.5	12	0.5
53	200	100	2	2	7	12	0.5
54	125	63	2	2	7	12	0.5
55	84	42	2	2	7	12	0.5
56	67	34	2	2	7	12	0.5
57	200	100	3	2	10.5	12	0.5
58	125	63	3	2	10.5	12	0.5
59	84	42	3	2	10.5	12	0.5
60	67	34	3	2	10.5	12	0.5
61	200	100	4	2	14	12	0.5
62	125	63	4	2	14	12	0.5
63	84	42	4	2	14	12	0.5
64	67	34	4	2	14	12	0.5
65	200	100	1	3	3.5	20	0.5
66	125	63	1	3	3.5	20	0.5
67	84	42	1	3	3.5	20	0.5
68	67	34	1	3	3.5	20	0.5
69	200	100	2	3	7	20	0.5
70	125	63	2	3	7	20	0.5

Table B.2: Shearlet System 2

Shearlet System	waveletEffSupp	gaussianEffSupp	scalesPerOctave	shearLevel	scales	nShearlets	alpha
1	67	67	1	2	3.5	12	0
2	50	50	1	2	3.5	12	0
3	40	40	1	2	3.5	12	0
4	34	34	1	2	3.5	12	0
5	67	67	2	2	7	12	0
6	50	50	2	2	7	12	0
7	40	40	2	2	7	12	0
8	34	34	2	2	7	12	0
9	67	67	3	2	10.5	12	0
10	50	50	3	2	10.5	12	0
11	50	50	3	2	10.5	12	0
12	40	40	3	2	10.5	12	0
13	34	34	3	2	10.5	12	0
14	67	67	1	3	3.5	12	0
15	50	50	1	3	3.5	20	0
16	40	40	1	3	3.5	20	0
17	34	34	1	3	3.5	20	0
18	67	67	2	3	7	20	0
19	50	50	2	3	7	20	0
20	40	40	2	3	7	20	0
21	34	34	2	3	7	20	0
22	67	67	3	3	10.5	20	0
23	50	50	3	3	10.5	20	0
24	40	40	3	3	10.5	20	0
25	67	67	1	4	3.5	36	0
26	50	50	1	4	3.5	36	0
27	40	40	1	4	3.5	36	0
28	34	34	1	4	3.5	36	0
29	67	67	2	4	7	36	0
30	50	50	2	4	7	36	0
31	40	40	2	4	7	36	0
32	34	34	2	4	7	36	0
33	67	67	3	4	10.5	36	0
34	50	50	3	4	10.5	36	0
35	40	40	3	4	10.5	36	0
36	34	34	3	4	10.5	36	0
37	67	67	1	2	3.5	12	0.5
38	50	50	1	2	3.5	12	0.5
39	40	40	1	2	3.5	12	0.5
40	34	34	1	2	3.5	12	0.5
41	67	67	2	2	7	12	0.5
42	50	50	2	2	7	12	0.5
43	40	40	2	2	7	12	0.5
44	34	34	2	2	7	12	0.5
45	67	67	3	2	10.5	12	0.5
46	50	50	3	2	10.5	12	0.5
47	40	40	3	2	10.5	12	0.5
48	34	34	3	2	10.5	12	0.5
49	67	67	1	3	3.5	20	0.5
50	50	50	1	3	3.5	20	0.5
51	40	40	1	3	3.5	20	0.5
52	34	34	1	3	3.5	20	0.5
53	67	67	2	3	7	20	0.5
54	50	50	2	3	7	20	0.5
55	40	40	2	3	7	20	0.5
56	34	34	2	3	7	20	0.5
57	67	67	3	3	10.5	20	0.5
58	50	50	3	3	10.5	20	0.5
59	40	40	3	3	10.5	20	0.5
60	34	34	3	3	10.5	20	0.5
61	67	67	1	4	3.5	36	0.5
62	50	50	1	4	3.5	36	0.5
63	40	40	1	4	3.5	36	0.5
64	34	34	1	4	3.5	36	0.5
65	67	67	2	4	7	36	0.5
66	50	50	2	4	7	36	0.5
67	40	40	2	4	7	36	0.5
68	34	34	2	4	7	36	0.5
69	67	67	3	4	10.5	36	0.5
70	50	50	3	4	10.5	36	0.5

Table B.3: Shearlet System 3

Shearlet System	waveletEffSupp	gaussianEffSupp	scalesPerOctave	shearLevel	scales	nShearlets	alpha
1	29	145	1	2	3.5	12	0
2	25	125	1	2	3.5	12	0
3	23	115	1	2	3.5	12	0
4	20	100	1	2	3.5	12	0
5	29	145	2	2	7	12	0
6	25	125	2	2	7	12	0
7	23	115	2	2	7	12	0
8	20	100	2	2	7	12	0
9	29	145	3	2	10.5	12	0
10	25	125	3	2	10.5	12	0
11	23	115	3	2	10.5	12	0
12	20	100	3	2	10.5	12	0
13	29	145	1	3	3.5	20	0
14	25	125	1	3	3.5	20	0
15	23	115	1	3	3.5	20	0
16	20	100	1	3	3.5	20	0
17	29	145	2	3	7	20	0
18	25	125	2	3	7	20	0
19	23	115	2	3	7	20	0
20	20	100	2	3	7	20	0
21	29	145	3	3	10.5	20	0
22	25	125	3	3	10.5	20	0
23	23	115	3	3	10.5	20	0
24	20	100	3	3	10.5	20	0
25	29	145	1	4	3.5	36	0
26	25	125	1	4	3.5	36	0
27	23	115	1	4	3.5	36	0
28	20	100	1	4	3.5	36	0
29	29	145	2	4	7	36	0
30	25	125	2	4	7	36	0
31	23	115	2	4	7	36	0
32	20	100	2	4	7	36	0
33	29	145	3	4	10.5	36	0
34	25	125	3	4	10.5	36	0
35	23	115	3	4	10.5	36	0
36	20	100	3	4	10.5	36	0
37	29	145	1	2	3.5	12	0.5
38	25	125	1	2	3.5	12	0.5
39	23	115	1	2	3.5	12	0.5
40	20	110	1	2	3.5	12	0.5
41	29	145	2	2	7	12	0.5
42	25	125	2	2	7	12	0.5
43	23	115	2	2	7	12	0.5
44	20	100	2	2	7	12	0.5
45	29	145	3	2	10.5	12	0.5
46	25	125	3	2	10.5	12	0.5
47	23	115	3	2	10.5	12	0.5
48	20	100	3	2	10.5	12	0.5
49	29	145	1	3	3.5	20	0.5
50	25	125	1	3	3.5	20	0.5
51	23	115	1	3	3.5	20	0.5
52	20	100	1	3	3.5	20	0.5
53	29	145	2	3	7	20	0.5
54	25	125	2	3	7	20	0.5
55	23	115	2	3	7	20	0.5
56	20	100	2	3	7	20	0.5
57	29	145	3	3	10.5	20	0.5
58	25	125	3	3	10.5	20	0.5
59	23	115	3	3	10.5	20	0.5
60	20	100	3	3	10.5	20	0.5
61	29	45	1	4	3.5	36	0.5
62	25	125	1	4	3.5	36	0.5
63	23	115	1	4	3.5	36	0.5
64	20	100	1	4	3.5	36	0.5
65	29	145	2	4	7	36	0.5
66	25	125	2	4	7	36	0.5
67	23	115	2	4	7	36	0.5
68	20	100	2	4	7	36	0.5
69	29	145	3	4	10.5	36	0.5
70	25	125	3	4	10.5	36	0.5

C

SPATIAL VARIATION RESULTS

This appendix contains supplementary spatial variation results derived from the methodology presented in Chapter 4. Hierarchical clustering results for each region are presented in the form of heatmaps distance matrices and dendrograms in Figs. C.1-C.3. The variation of cluster boundaries for varying levels of dendrogram cuts or granularity are shown in Figs. C.4-C.9. Archetypal subgraph examples from the detected clusters are depicted in Figs. C.10-C.18.

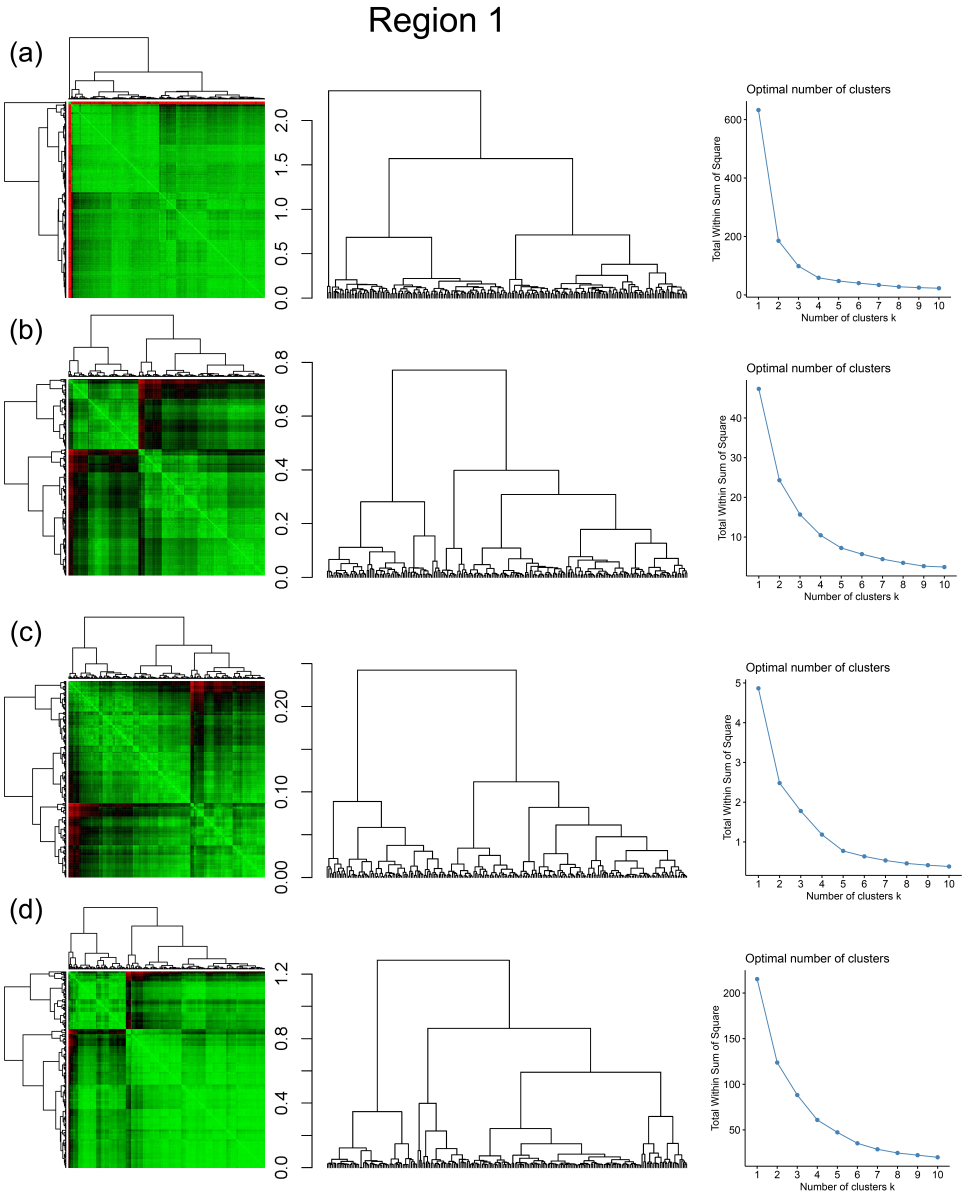

Figure C.1: Combined symmetric heatmap of distance matrix and dendrograms, dendrograms, and sum-of-squares elbow plots for Region 1 (a) Fingerprint (b) D-measure (c) NetLSD (d) Portrait Divergence

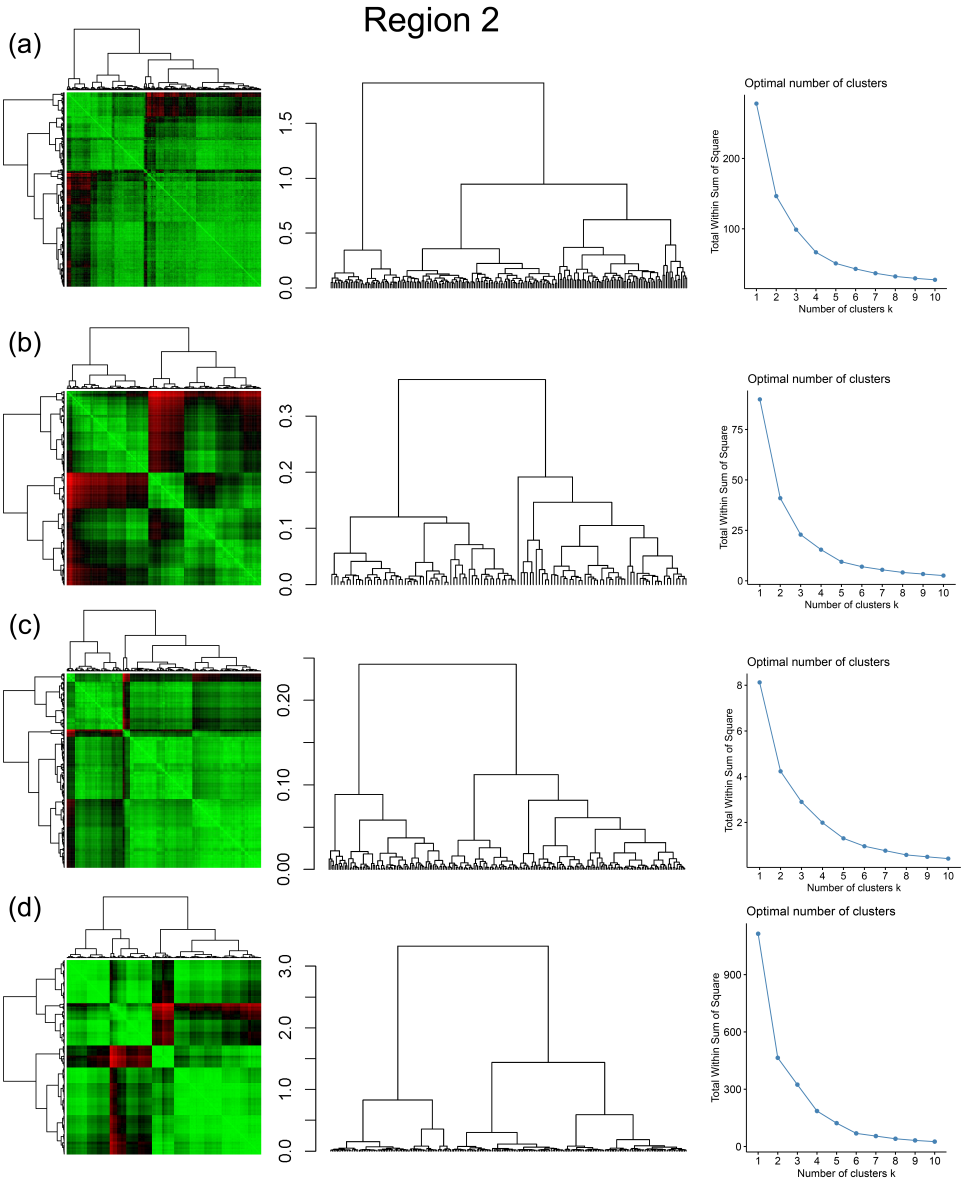

Figure C.2: Combined symmetric heatmap of distance matrix and dendrograms, dendrograms, and sum-of-squares elbow plots for Region 2 (a) Fingerprint (b) D-measure (c) NetLSD (d) Portrait Divergence

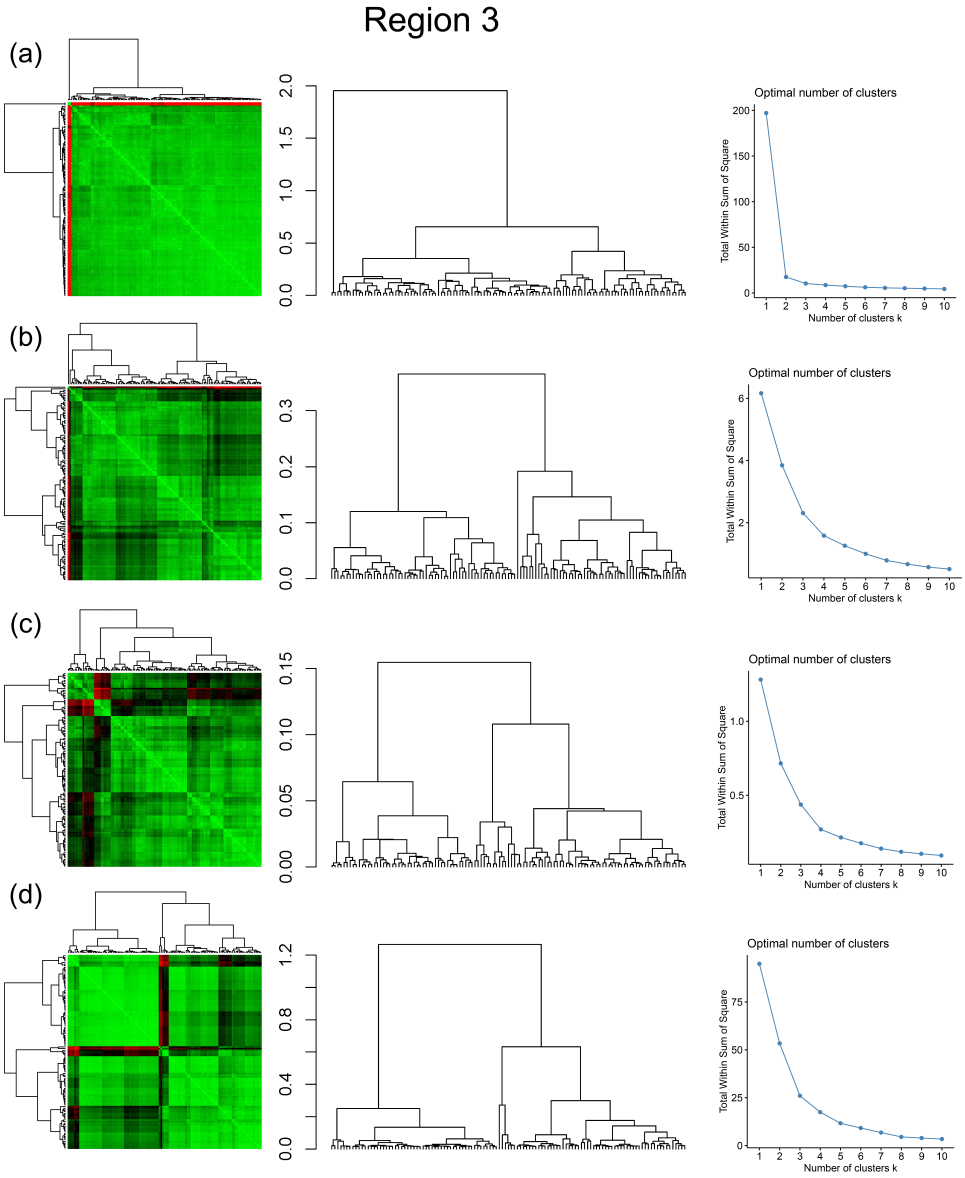

Figure C.3: Combined symmetric heatmap of distance matrix and dendrograms, dendrograms, and sum-of-squares elbow plots for Region 3 (a) Fingerprint (b) D-measure (c) NetLSD (d) Portrait Divergence

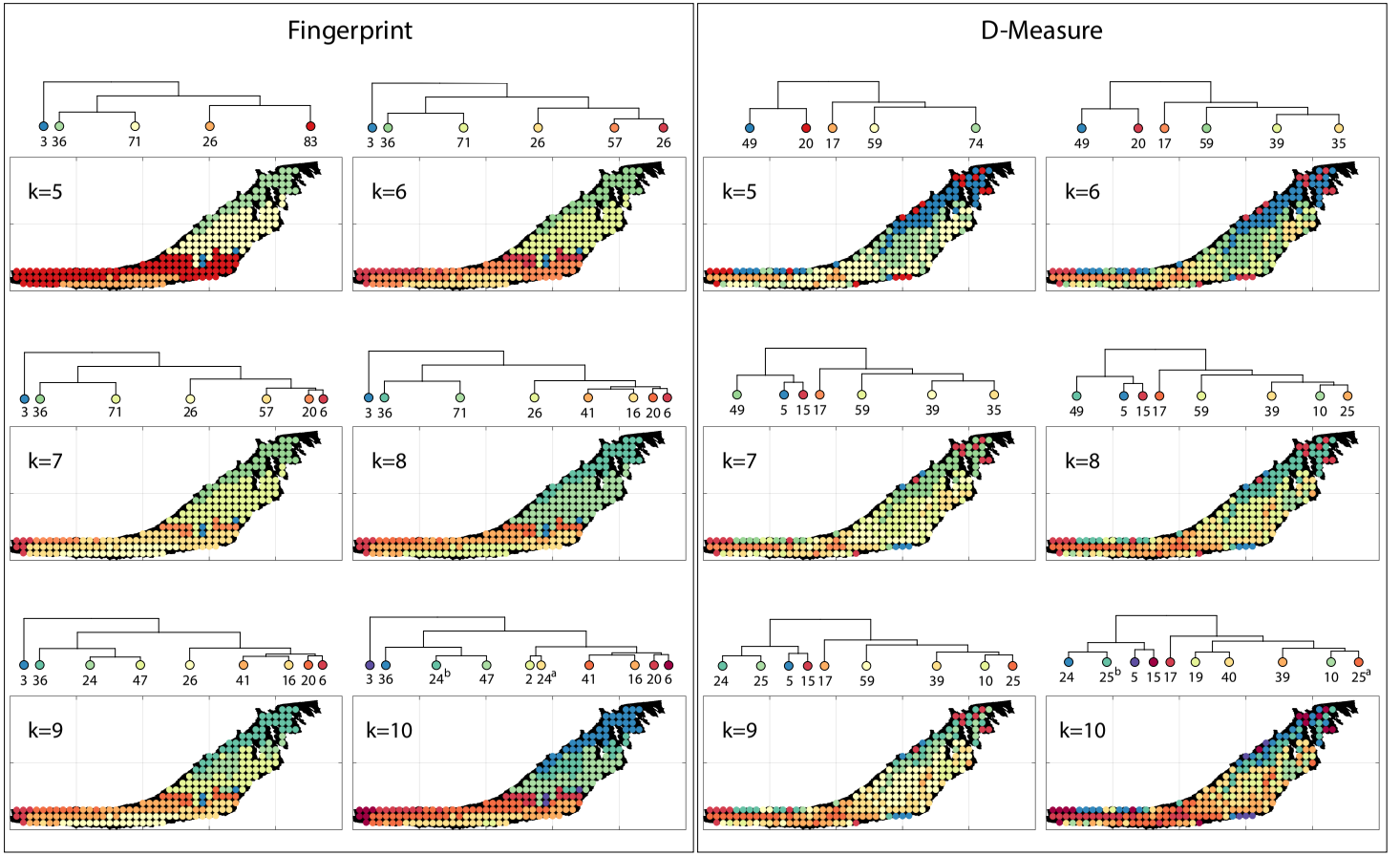

Figure C.4: Variation in cluster boundaries for k clusters in Region 1 using (left) Fingerprint (right) D-measure

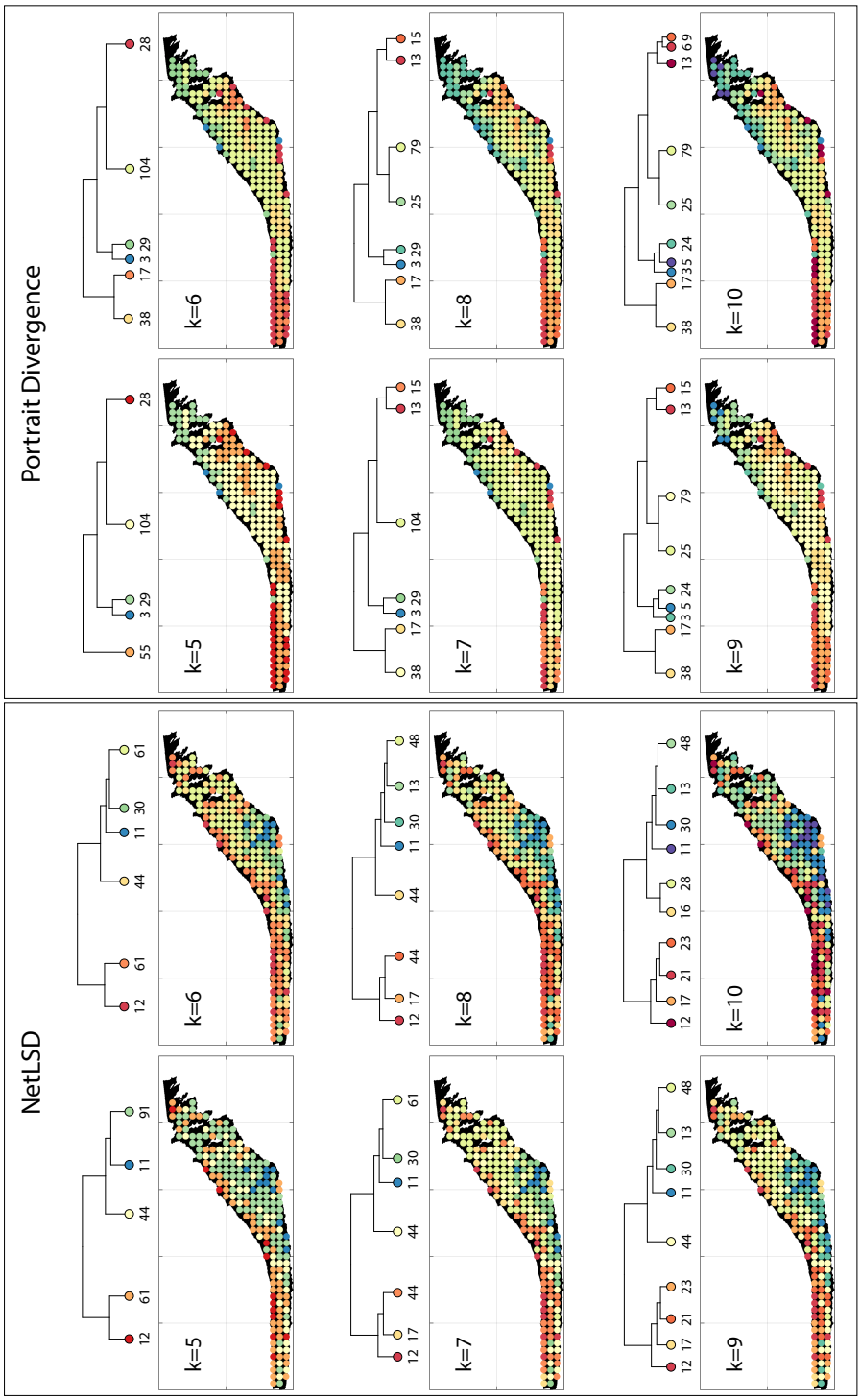

Figure C.5: Variation in cluster boundaries for k clusters in Region 1 using (left) NetLSD (right) Portrait Divergence

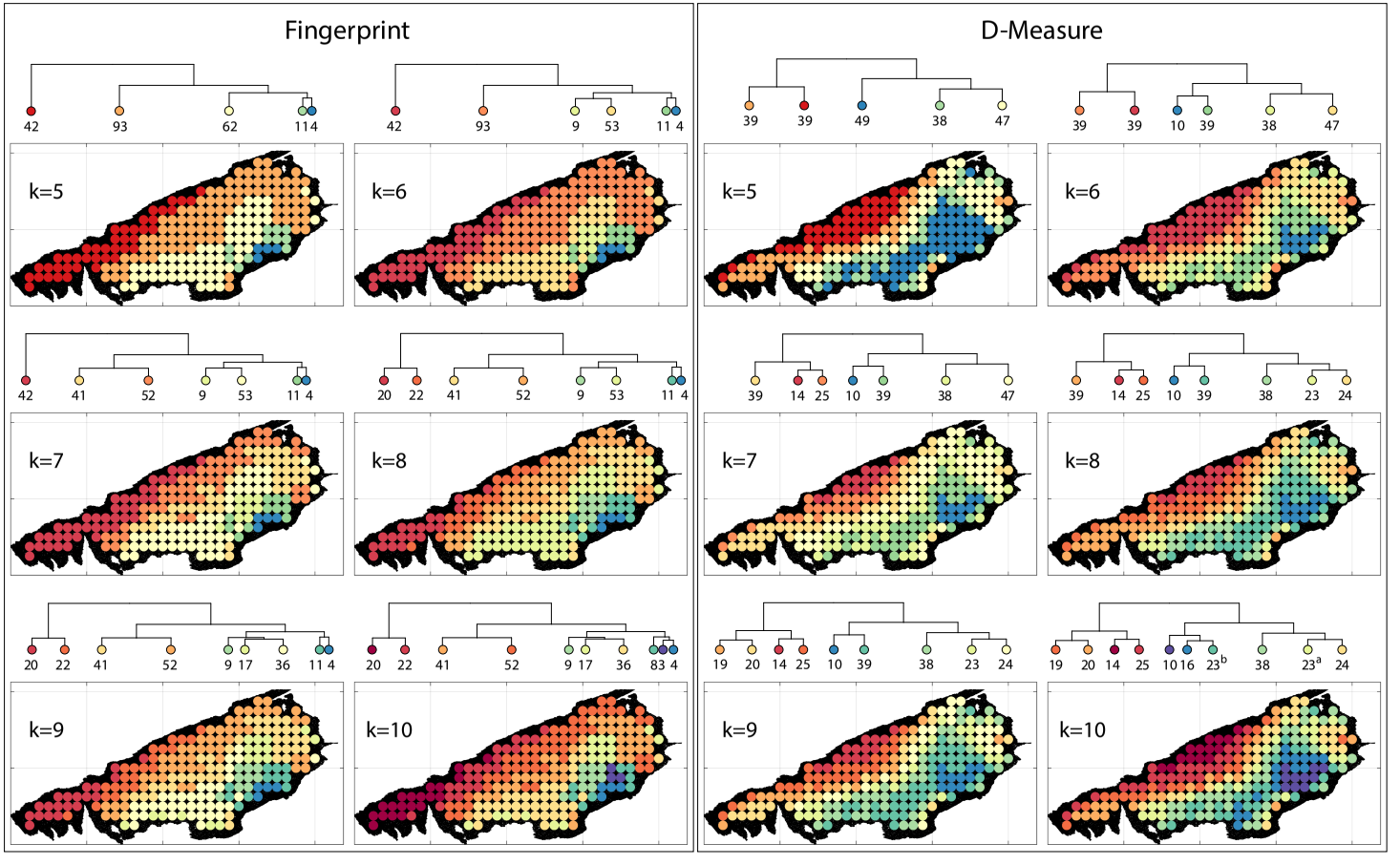

Figure C.6: Variation in cluster boundaries for k clusters in Region 2 using (left) Fingerprint (right) D-measure

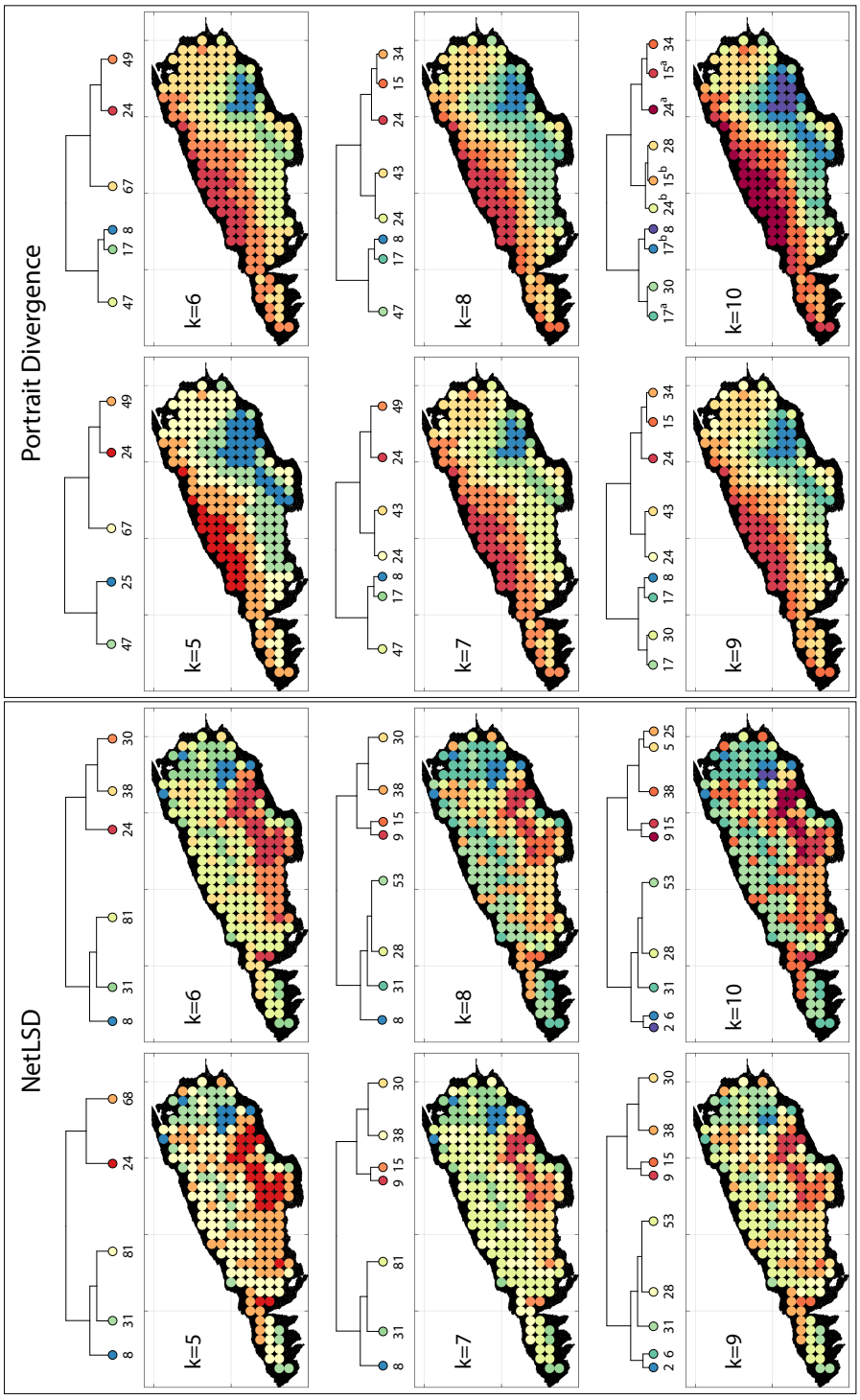

Figure C.7: Variation in cluster boundaries for k clusters in Region 2 using (left) NetLSD (right) Portrait Divergence

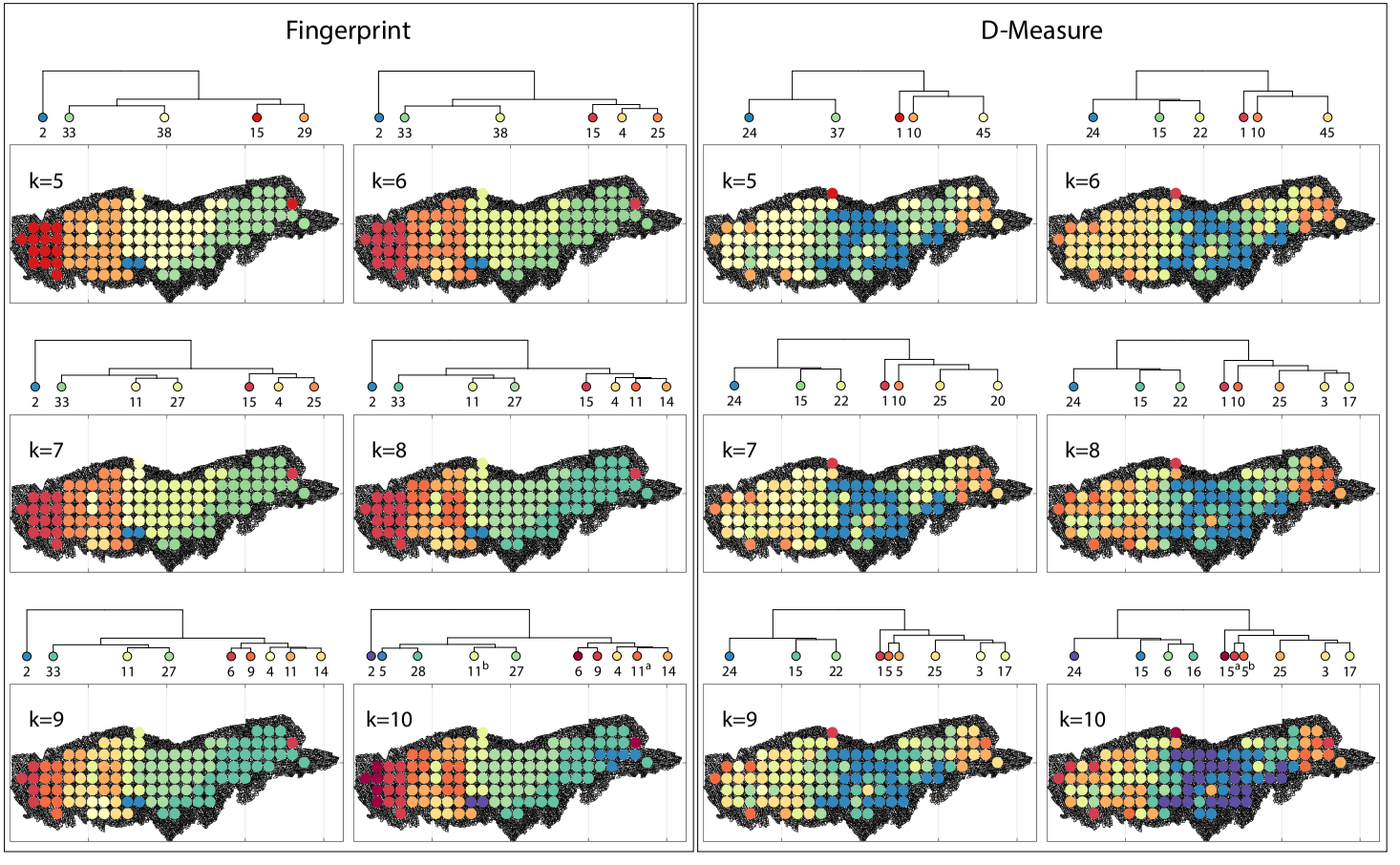

Figure C.8: Variation in cluster boundaries for k clusters in Region 3 using (left) Fingerprint (right) D-measure

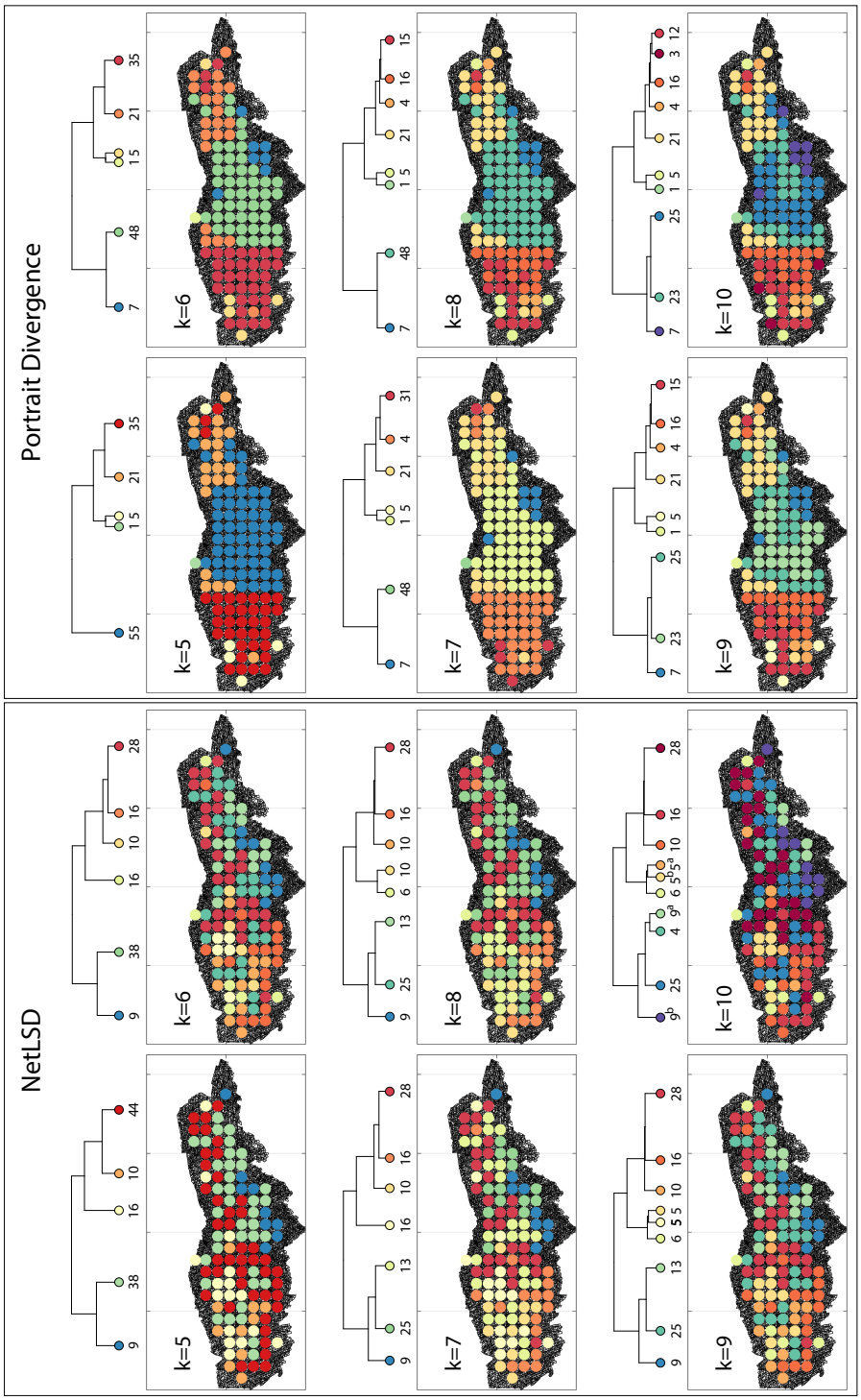

Figure C.9: Variation in cluster boundaries for k clusters in Region 3 using (left) NetLSD (right) Portrait Divergence

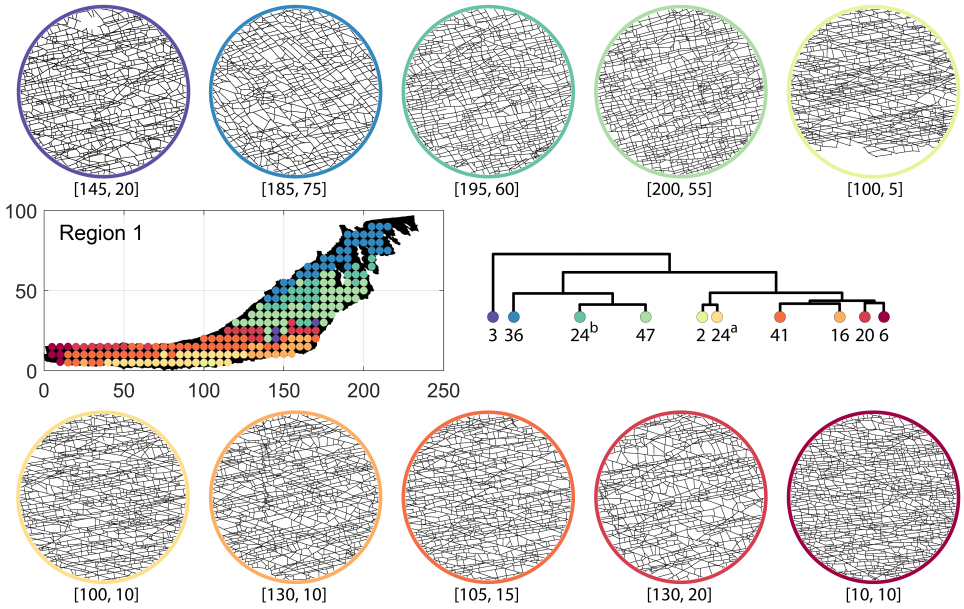

Figure C.10: Subgraph samples depicting variation in fracturing identified by fingerprint distance in Region 1

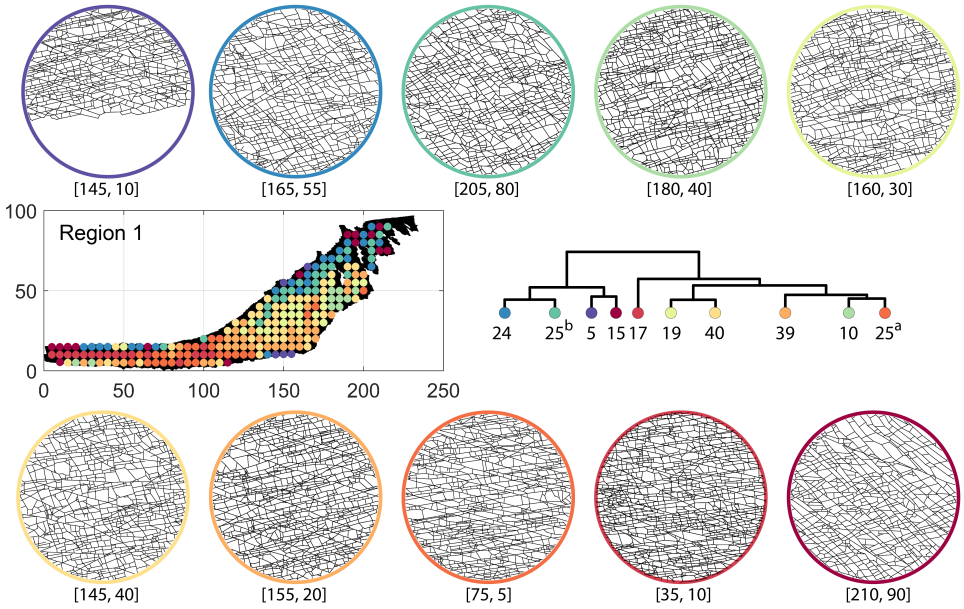

Figure C.11: Subgraph samples depicting variation in fracturing identified by D-measure in Region 1

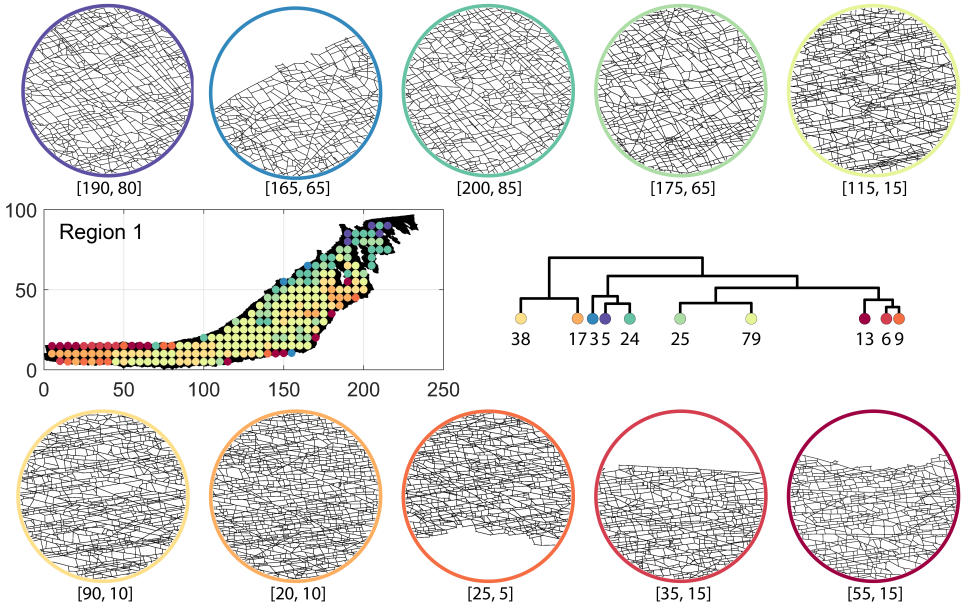

Figure C.12: Subgraph samples depicting variation in fracturing identified by portrait divergence in Region 1

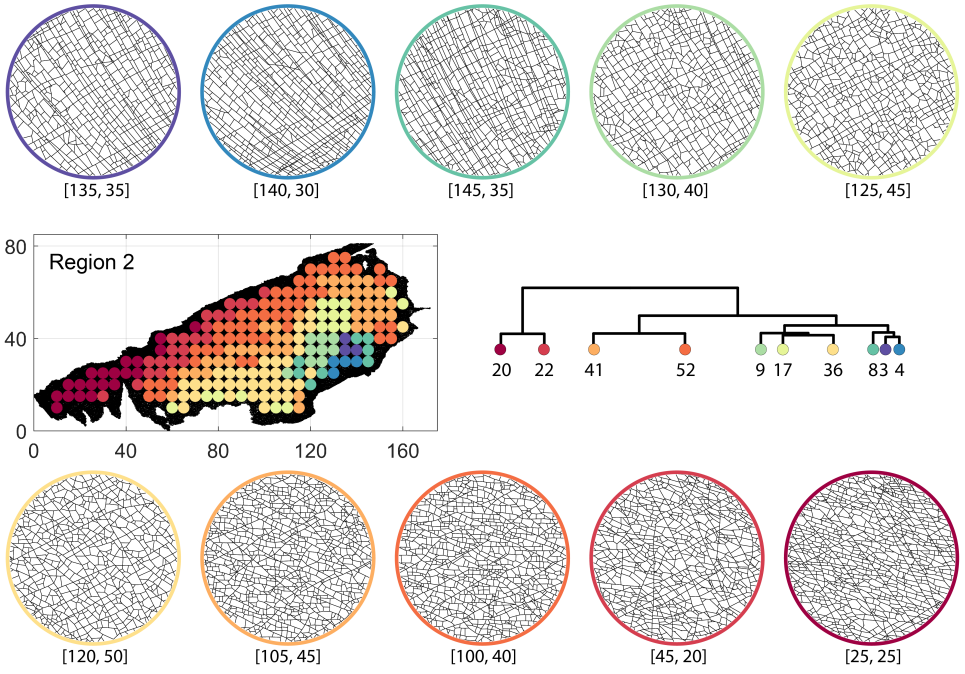

Figure C.13: Subgraph samples depicting variation in fracturing as identified by fingerprint distance in Region 2

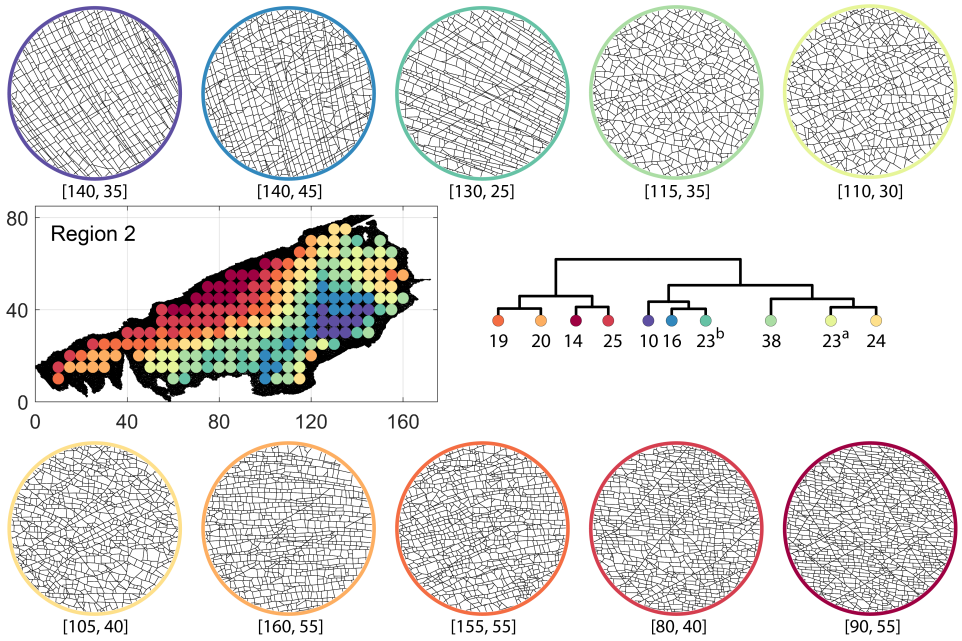

Figure C.14: Subgraph samples depicting variation in fracturing as identified by D-measure in Region 2

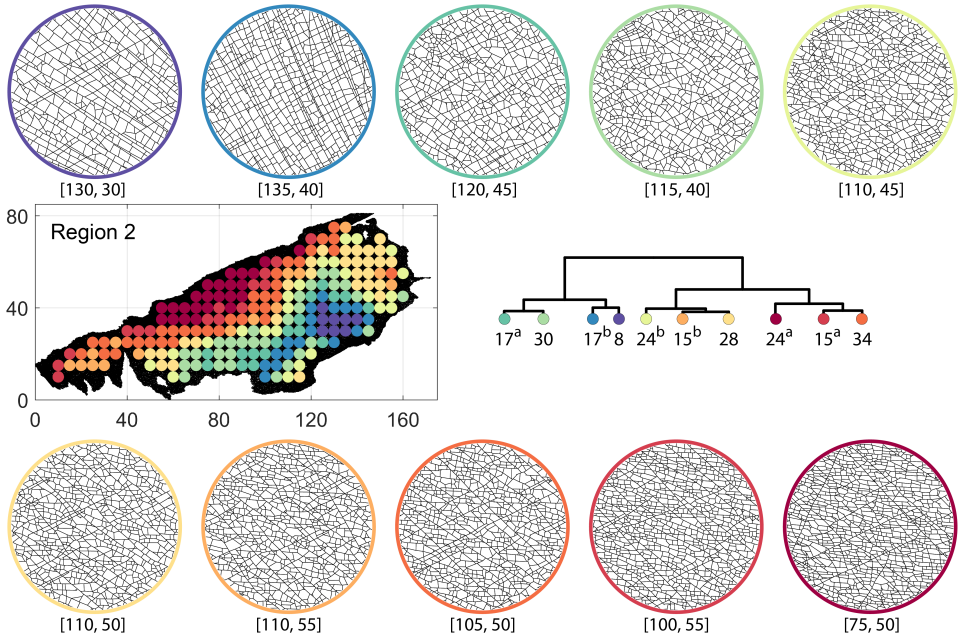

Figure C.15: Subgraph samples depicting variation in fracturing as identified by portrait divergence in Region 2

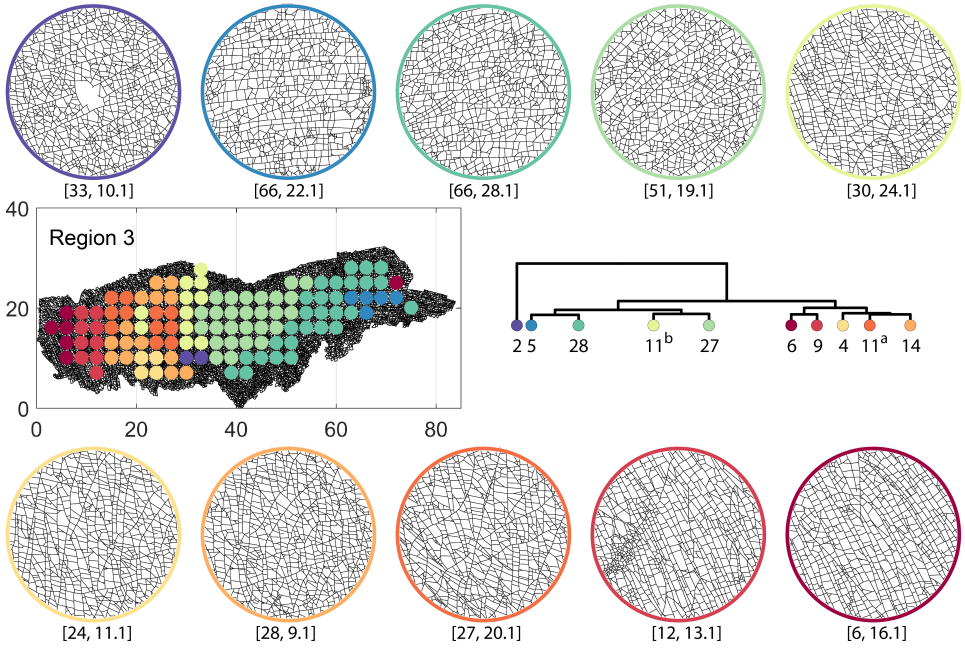

Figure C.16: Subgraph samples depicting variation in fracturing as identified by fingerprint distance in Region 3

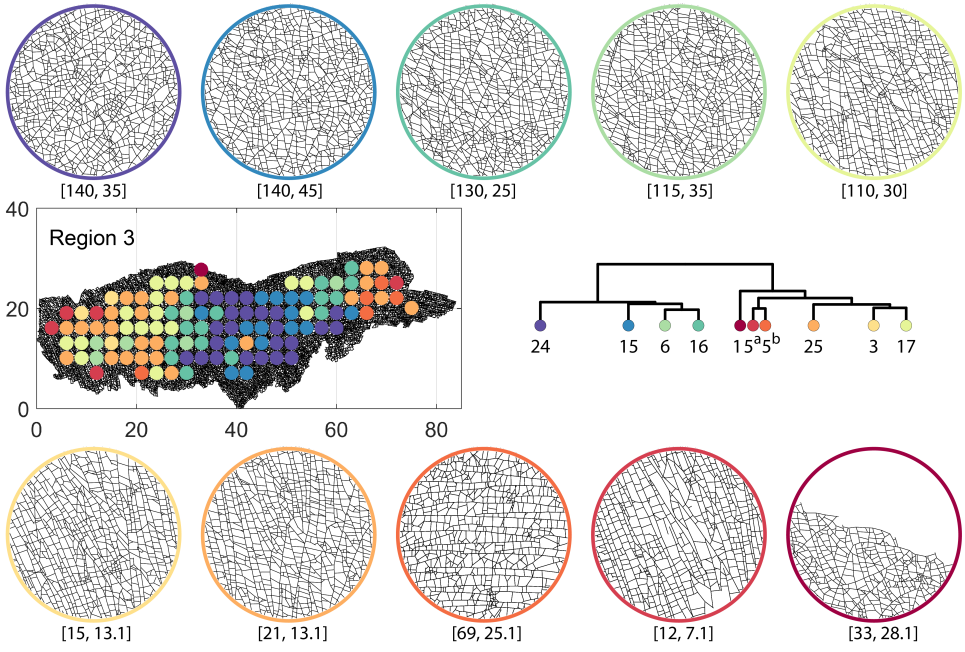

Figure C.17: Subgraph samples depicting variation in fracturing as identified by D-measure in Region 3

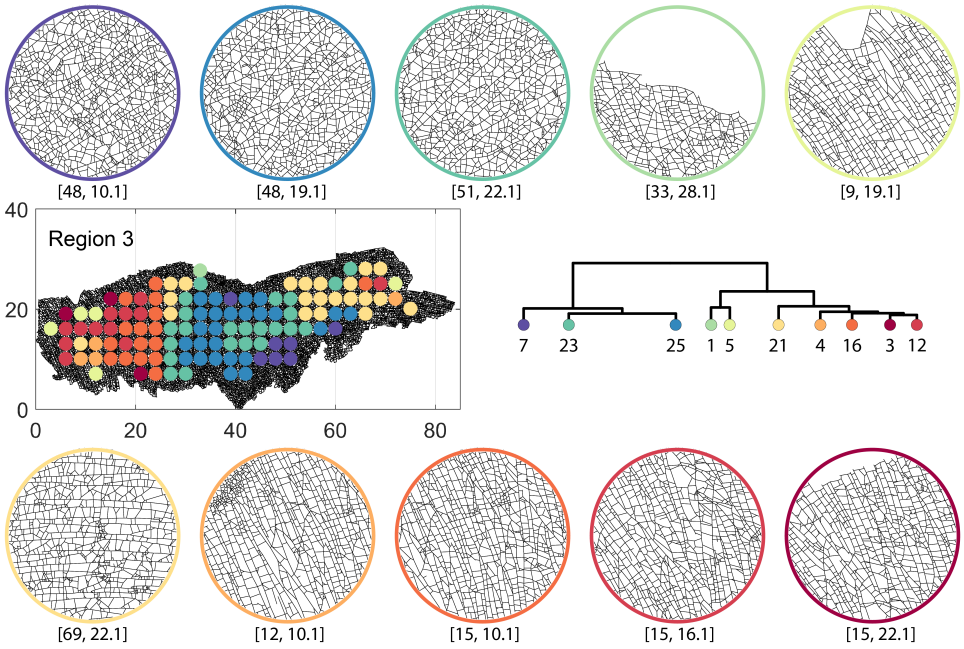

Figure C.18: Subgraph samples depicting variation in fracturing as identified by portrait divergence in Region 3

D

BETWEENNESS CENTRALITIES OF FRACTURE NETWORKS

This appendix deals with the application of betweenness centrality (BC) to spatial fracture networks. BC is a value assigned to a node in a spatial graph that is computed based on the ratio between the number of times a node falls on a geodesic path between all combination of nodes in the graph to the number of shortest paths between all node combinations of nodes in the graph [Newman, 2001]. It is given as::

$$g(i) = \frac{1}{\mathcal{N}} \sum_{s \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}, \tag{D.1}$$

where σ_{st} is the number of shortest paths from s to t and $\sigma_{st}(i)$ is the number of shortest paths from s to t passing through node i (S and t is usual graph convention for source and target nodes). $\mathcal{N}$ is the constant of normalization, which is equal to $\mathcal{N} = (N-1)(N-2)$ where N is the total number of nodes. The BC for a square lattice decreases with distance to the barycentre of nodes. In non-lattice spatial graphs, such as fracture networks, the relative disorder in node positions introduces fluctuations that create non-trivial patterns which develop due to an interplay of space and topology [Barthelemy, 2018].

The betweenness centrality scores of the nodes create specific arterial paths in the network that serves as a proxy for traffic flow through the graph Jonckheere et al. [2011]. If we assume that the travelling quantity chooses shortest paths and each node pair forms an origin-destination couple, BC corresponds to local traffic at a node [Barthelemy, 2018]. The flow can be any physical process such as fluid flow, electricity, traffic, or intangible such as data packets. These arterial paths are not evident by simple a priori observation of the network but only revealed after computation of BC.

Using Euclidean distance as edge weights, we compute the BC profiles for all five networks (fracture graphs) as presented in Chapter 3. A set of pathways is revealed by the BC computation, which is not evident from a traditional fracture network analysis. These are plotted in Fig. D.1(a) - D.3(a). For each area, we extract a subgraph that contains the

nodes with the top 12 percent of BC magnitudes. These are depicted in Fig. D.1(b) - D.3(b). The results indicate that boundary effects and the general shape of the periphery of the fracture networks influence the final pattern.

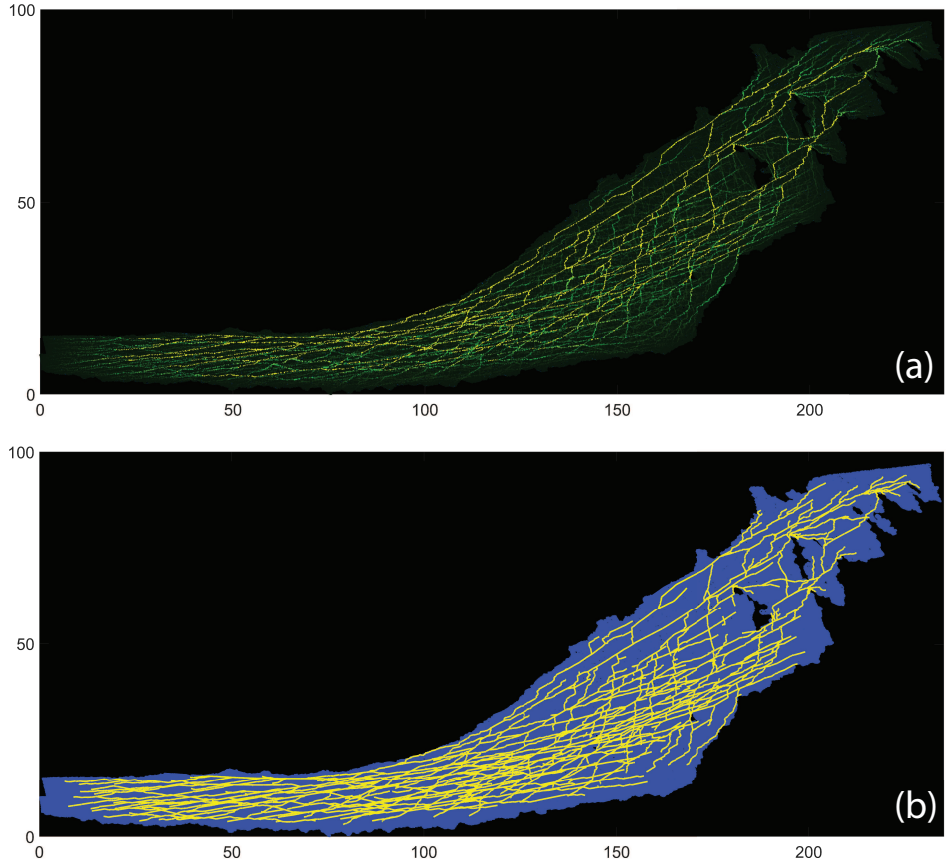

Figure D.1: (a) BC profile for fracture network in Area 2. Warm colours indicate high BC and dark low BC (b) Highlighting edges connecting nodes that are in the top 12 % w.r.t BC values

We computed the BC based on edge weights that are equivalent to Euclidean distance. However, flow and transport in fracture networks are dependent upon fracture conductivity, F_c is a function of fracture aperture w and length l. Approximately, we can write,

$$F_c \sim \frac{w^3}{l}. \tag{D.2}$$

F_c can now be used as an edge weight. Fracture aperture distributions are stress-sensitive. To study the effect of stress-sensitivity on BC, we use the geometric aperture method of Bisdom et al. [2016] to define an effective hydraulic aperture. The variation of BC w.r.t maximum stress directions is depicted in Fig. D.4 for Areas 4 and 5

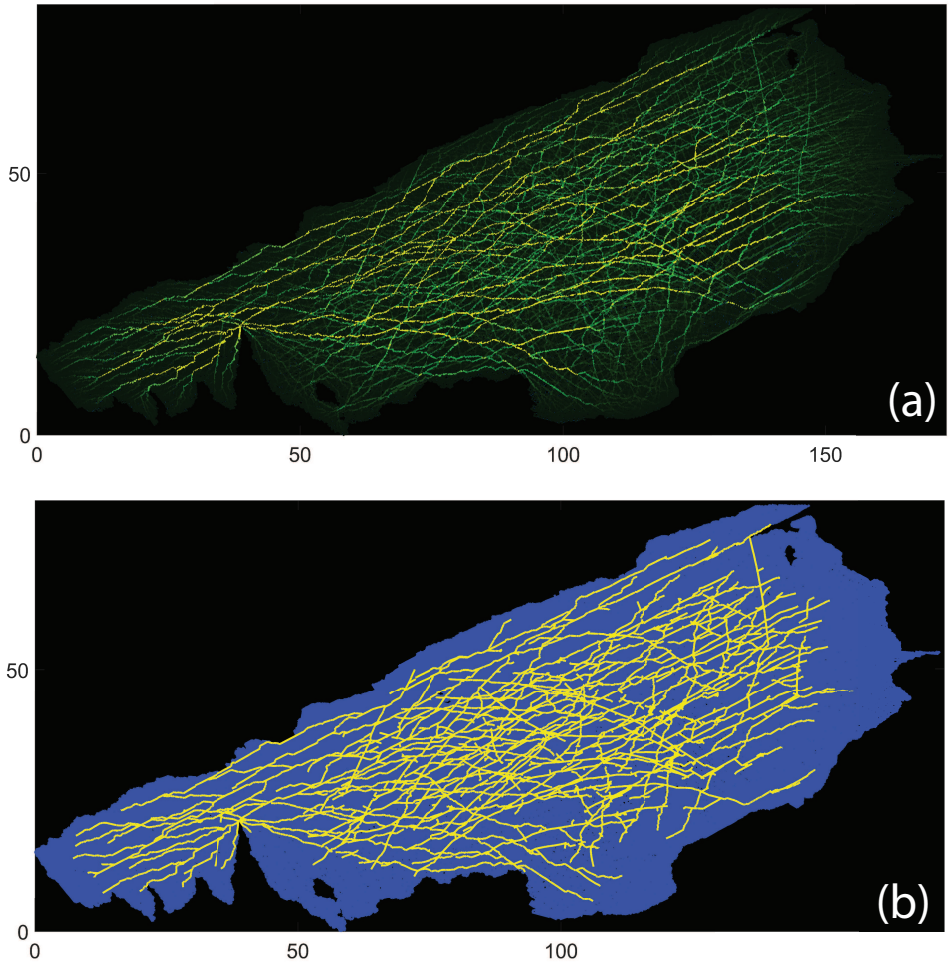

Figure D.2: (a) BC profile for fracture network in Area 4. Warm colours indicate high BC and dark low BC (b) Highlighting edges connecting nodes that are in the top 12 % w.r.t BC values

We have not yet tested the large-scale networks' flow behavior to verify the high-BC nodes' role as playing an essential role in flow and transport in fractured media. However, there is evidence from the work of Srinivasan et al. [2018] and Srinivasan et al. [2019], that use machine learning coupled with physics-based flow simulations, that a pruned DFN (or backbone) is extractable from a more extensive network while retaining transport characteristics. It remains to be verified whether one can use the BC for model order reduction of fracture networks.

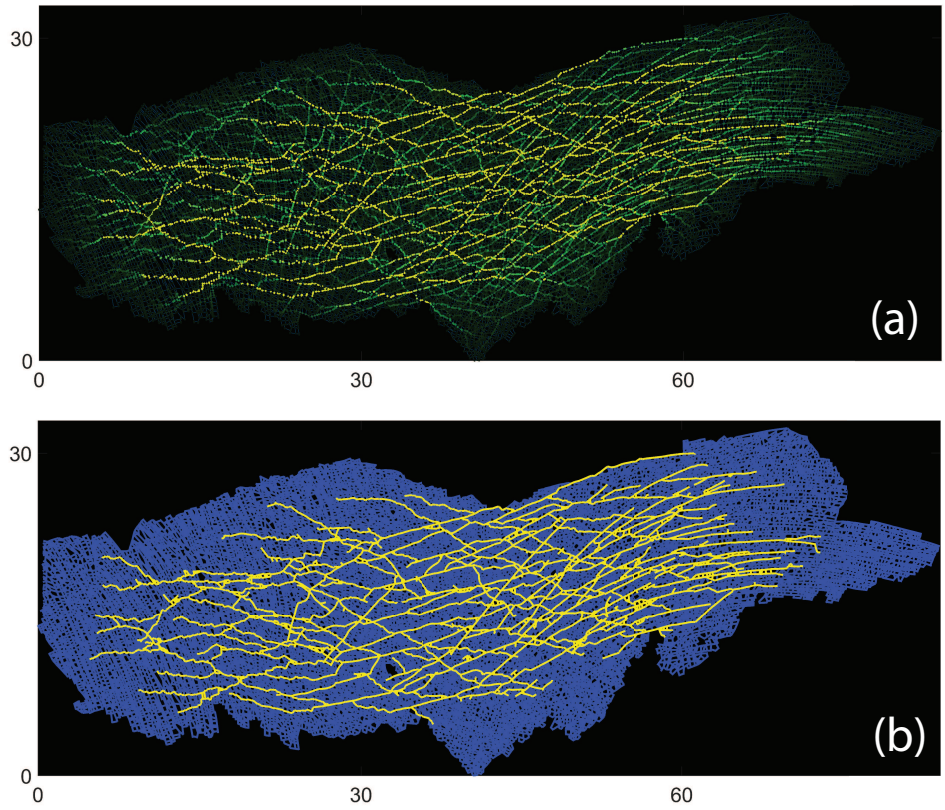

Figure D.3: (a) BC profile for fracture network in Area 5. Warm colours indicate high BC and dark low BC (b) Highlighting edges connecting nodes that are in the top 12 % w.r.t BC values

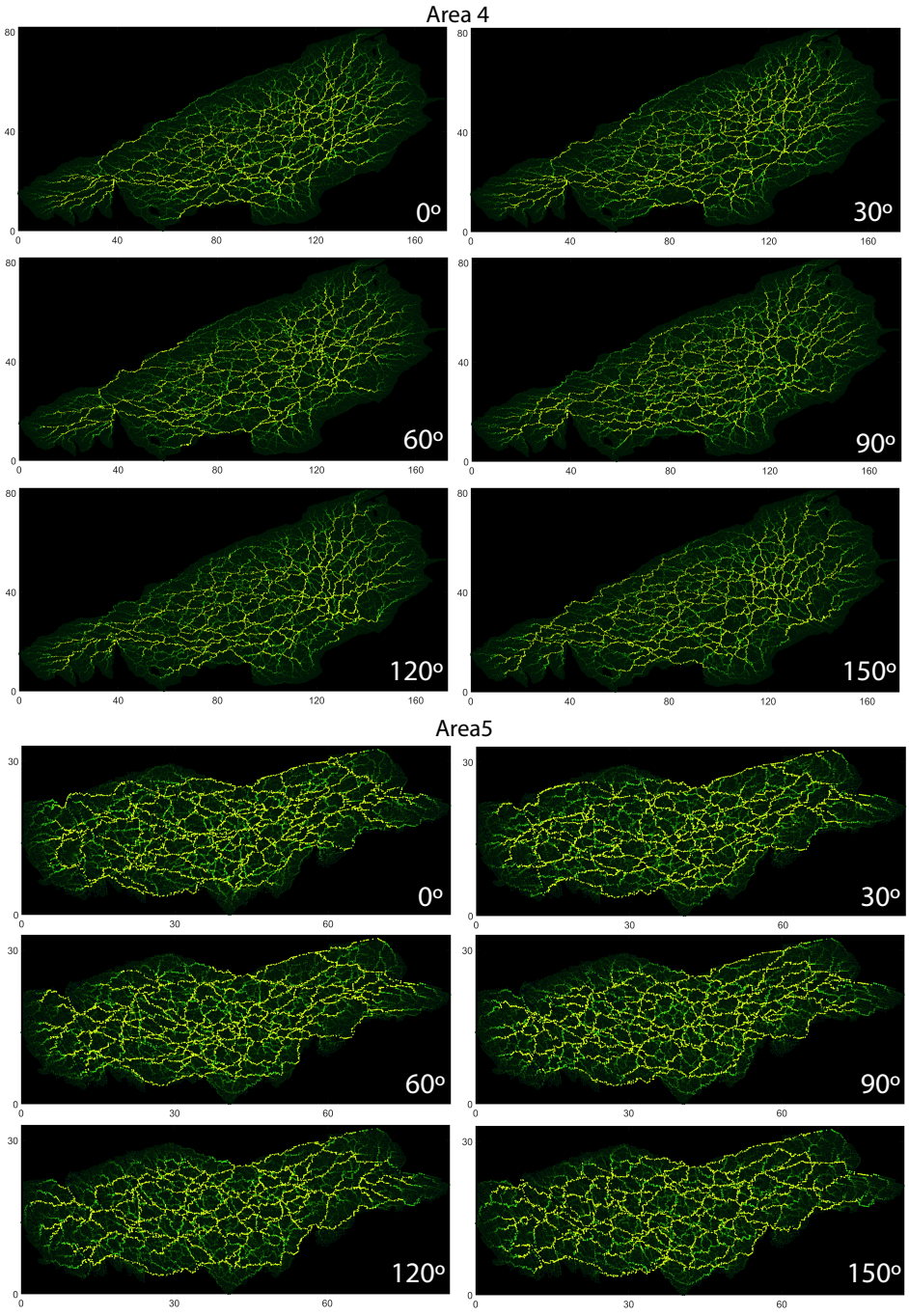

Figure D.4: BC profiles corresponding to Areas 4 and 5 with edge weights varying as an edge conductivity that is a function of geometric aperture

E

NETWORK SIMPLICITY

The simplicity profile of a spatial graph proposed by Viana et al. [2013] is a high-level metric to characterize a spatial graph. The concept is based on the statistics of shortest and simple paths within a planar, spatial graph. For any pair of nodes in a graph (i, j), the shortest path between the two nodes is obtained by the euclidean length $l(i, j)$ through graph edges. In contrast, a simple path, $l^*(i, j)$, is defined as one which requires the least number of turns to connect the nodes (i, j) through the graph.

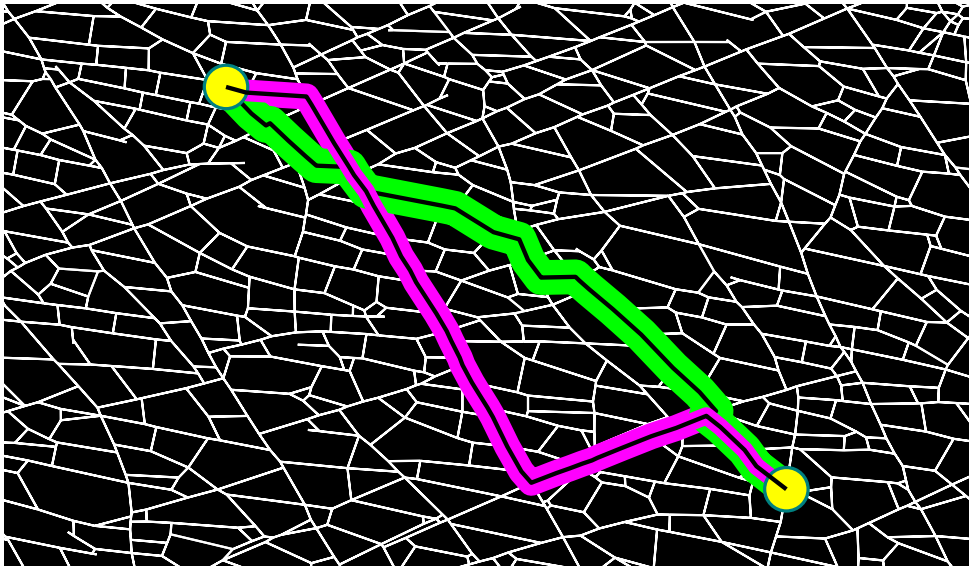

Figure E.1: An example depicting the difference between simple and shortest paths for a fracture network. Source and target nodes are depicted in yellow. The shortest path based on euclidean distance between source and target nodes is depicted in green. The simplest path requires minimum number of turns to traverse from source to target and is depicted in magenta

Simple paths equal the shortest path in lattice-type networks but diverge from unity in non-lattice networks. A statistical comparison between the ratio of lengths of the simplest and shortest paths yields information about a fracture network structure. This ratio $\frac{l^*(i,j)}{l(i,j)} \geq 1$ for a spatial graph. For lattices and trees, the ratio is 1. For every path in the primal, i.e., P(i,j), the nodes 'i' and 'j' are part of fractures F(i) and F(j). F(i) and F(j) correspond to nodes, N(i) and N(j) in the dual graph. We can count the shortest paths in the dual graph between N(i) and N(j), in terms of the number of dual graph nodes one must hop to reach from N(i) and N(j). This is equivalent to the number of turns taken. There can be multiple paths that have a similar number of turns (or node hops). L*(i,j) refer to the least euclidean distance. Figure E.1 depicts an example of the difference between the shortest and simplest paths.

The simplicity index S is defined as the normalized average for N nodes as

$$S = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{l^*(i,j)}{l(i,j)}. \tag{E.1}$$

The simplicity index S mixes various scales within the fracture network and we can define a simplicity profile in terms of length d,

$$S(d) = \frac{1}{N(d)} \sum_{i,j/d_E(i,j)=d} \frac{l^*(i,j)}{l(i,j)}. \tag{E.2}$$

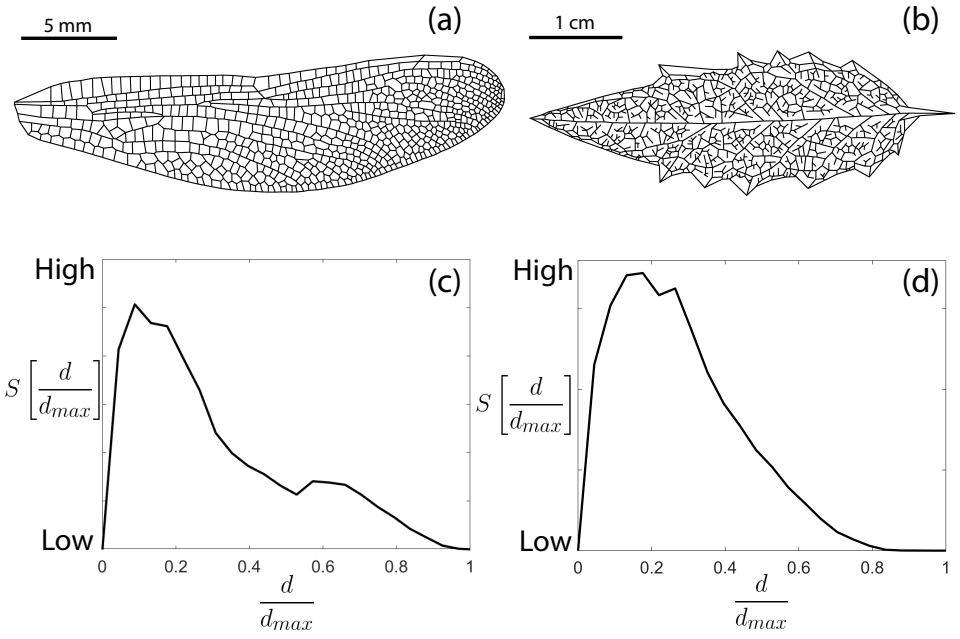

Figure E.2: Simplicity profiles for biological networks (a) Dragonfly wing (b) *Ilex aquifolium* (c) simplicity profile of dragonfly wing (d) simplicity profile of*Ilex aquifolium*

The simplicity profile $S(d)$ always has a value greater than 1. The variation of $S(d)$, with increasing and decreasing d, highlights structural variations within the graph. For $S(d \to 0) \sim 1$ and at smaller values of d, the paths are at the nearest neighbors' scale, and the simplest and shortest paths converge. At the other end of the length-scale spectrum, when $d \to d_{max}$, there are many more long and straight paths available, decreasing the difference between simple and straight paths. At an intermediate d^*, where $S(d)$ peaks, the divergence between the simplest and shortest paths are maximum. At this length scale, d^*, we have the graph domains that are not connected by straight paths. We can normalize $S(d)$ by dividing by d_{max}. Thus the simplicity profile makes a comparison of graphs across scales possible. Viana et al. [2013] compared various spatial networks such as roads, railways, waterways, leaves, and dragonfly wings.

We depict the simplicity profiles of two types of biological spatial networks, i.e., a dragonfly wing and a a leaf veination pattern in Fig. E.2. Both the graphs were digitized from image data published in Viana et al. [2013]. An interesting phenomenon is the peak of d^* within the ensemble of simplicity profiles. From the observations of Viana et al. [2013], different from man-made networks, biological networks such as leaf venation patterns and *physarum* have their peaks clustering around $\frac{d}{d_{max}} \to 0.2$ followed by fast decay. From the reconstructed networks of a dragonfly wing and a leaf veination pattern we are able to see such a peak. Our implementation of the computation of normalized simplicity broadly matches the curve shapes presented by Viana et al. [2013]. It may be noted that the profile might vary depending on the strike thresholds used to compute the dual graph as described in Chapter 3.

We now compute the simplicity profiles for circularly sampled sub-graphs from Area 4 as described in Chapter 3 using our implementation. The sub-graphs and their respective profiles are plotted in Fig. E.3. The profiles are quite different with varying peaks $S(\frac{d}{d_{max}})$, peak values of $S(\frac{d}{d_{max}})$, and decay behaviour. Sub-graph (b), has the most elevated profile while sub-graph (e) has the lowest height. Unlike the biological networks (see Fig. E.2) the peaks are shifted to the right to a range of between 0.4-0.6 for $\frac{d}{d_{max}}$.

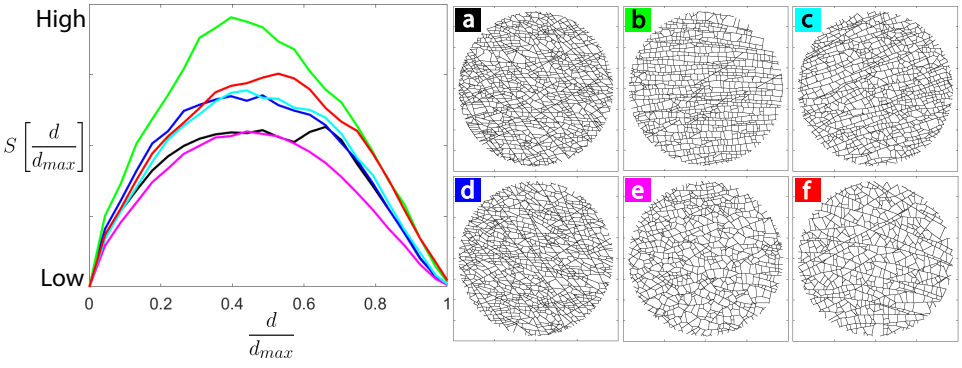

Figure E.3: Simplicity profiles for six different fracture network samples each of 7.5 m diameter. Curve colours correspond to the colours of the square label insets in the top-left corner of each fracture network sampling.

This can be explained by close scrutiny of the corresponding networks. Sub-graph (e) has a very diffused type of fracturing with short sets, and hence a lesser number of turns

is needed to traverse between any two pair of nodes. On the other hand, sub-graph (b) has long and persistent fracture sets that traverse the sampling, implying more detours and increasing the simplicity peak. The fracture intensity also plays a role with sub-graph (f) having the least fracture intensity having the second-largest peak of $S(\frac{d}{d_{max}})$ among all samplings. The simplicity profiles provide a unique, global indicator by which one can mean to compare fracture networks across scales as the index is normalized. The concept can be explored further connecting it to detailed comparisons of spatial heterogeneity in fracturing and the relation of network structure with respect to equivalent permeability.

F

DEEP LEARNING FOR AUTOMATIC FRACTURE TRACING

This appendix introduces a deep learning approach to automatic fracture extraction using Generative Adversarial Networks (GANs). In contrast to the method described in Chapter 2 where the complex shearlet transform was used to convert images of fractured rocks to vectorized fracture traces, deep-learning methods use a set of training data which a neural network architecture learns a mapping function between a set of inputs to outputs. Many such architectures have been successful in computer vision tasks such as image classification, feature extraction, image generation etc, such as U-Net, CNNs, etc. This chapter utilizes the vector traces generated by complex shearlet-based automatic tracing pertaining to the Bristol Channel outcrop data to train a conditional GAN. The trained model is then applied to a new region within the Bristol Channel pavement to extract fracture traces. The results indicate that such an approach using a pre-trained deep neural architecture is quite amenable to the problem of fracture detection.

Deep neural architectures for feature detection

Modelling of naturally fractured reservoirs (NFRs) using Discrete Fracture Network (DFN) approaches, in which spatial organization of fractures are represented explicitly within porous continuum, have become increasingly popular and relevant in NFR modeling. Digitized fracture traces from Digital Outcrop Models (DOMs) based on LIDAR/UAV photogrammetry provide critical information on fracture network properties and architecture at a sub-seismic scale. These include fracture length scaling, orientation, topology, and intensity. However, the extraction of fracture traces is a manual, time-intensive process and prone to interpreter bias, and the total value of large photogrammetric datasets is seldom achieved in practice.

The general goal in an automated fracture trace detection workflow can be generalized as an image pixel prediction task. The convolutional neural net (CNN), since its inception in 2012, has fast evolved and is a ubiquitous choice for many tasks in computer vision. CNNs have evolved from the basic architecture of convolutional and pooling lay-

ers into variations such as the U-Net and the RCNN. In these neural architectures, the CNN learns a loss function between target inputs and outputs. Generative Adversarial Networks (GANs) are a type of neural architecture that consist of a generator G and a discriminator D that participate in a minimax game so that the output indistinguishably converges to the input and the loss function is automatically learnt [Goodfellow et al., 2014].

Conditional GANs (cGANs) are a variation of a GAN in a conditional setting. The typical GAN learns a mapping from a random noise vector z to an output image y, $G : z \rightarrow y$. A cGAN learns a mapping from an observed image x and random noise vector z to y, such that $G : x, z \rightarrow y$.

Description of cGAN architecture

We utilize the cGAN architecture implementation referred to as *Pix2Pix* by Isola et al. [2017]. The generator architecture consists of a modified U-Net with an encoder and a decoder that retains the skip connections between both components. In a typical encoder-decoder configuration, input passes through multiple layers that progressively downsample till reaching the bottleneck. The process reverses at the bottleneck. There is a loss of low-level information due to the bottleneck, and adding skip connections prevents this. The blocks of the encoder consist of a Conv → Batchnorm → Leaky ReLU sequence, and the blocks of the decoder consist of a Transposed Conv → Batchnorm → Dropout sequence.

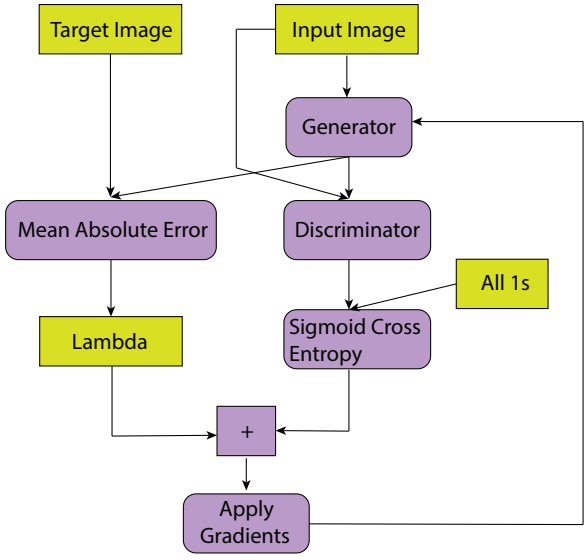

Figure F.1: Workflow of the pix2pix cGAN (modified from Isola et al, 2018)

The discriminator is a PatchGAN with each block having a Conv → Batchnorm → Leaky ReLU sequence.

Training Data

We create input training data from vectorized fracture traces as presented in Chapter 3. A total of 138 image tiles from the Bristol Channel dataset is extracted (see Fig. F.2). The vector trace data is rasterized and placed alongside the source to represent the target image output, as shown in example Fig. F.3.

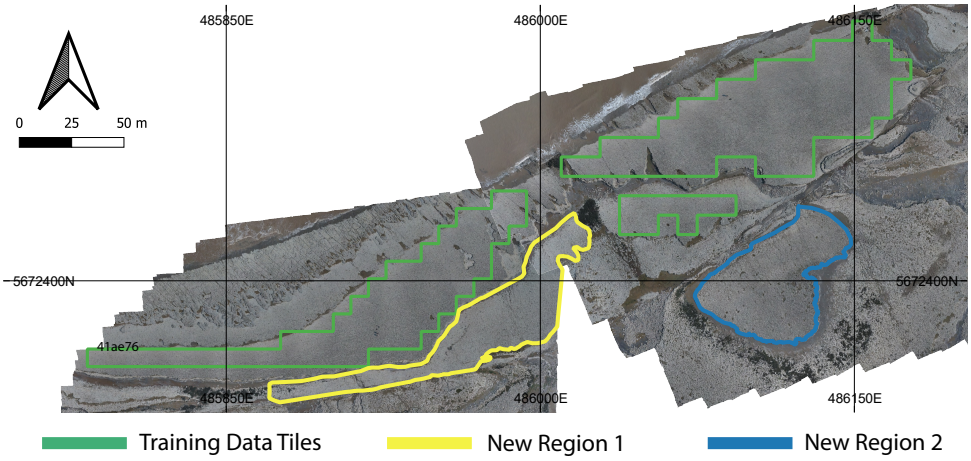

Figure F.2: Training image tiles and validation regions from the Bristol Channel dataset

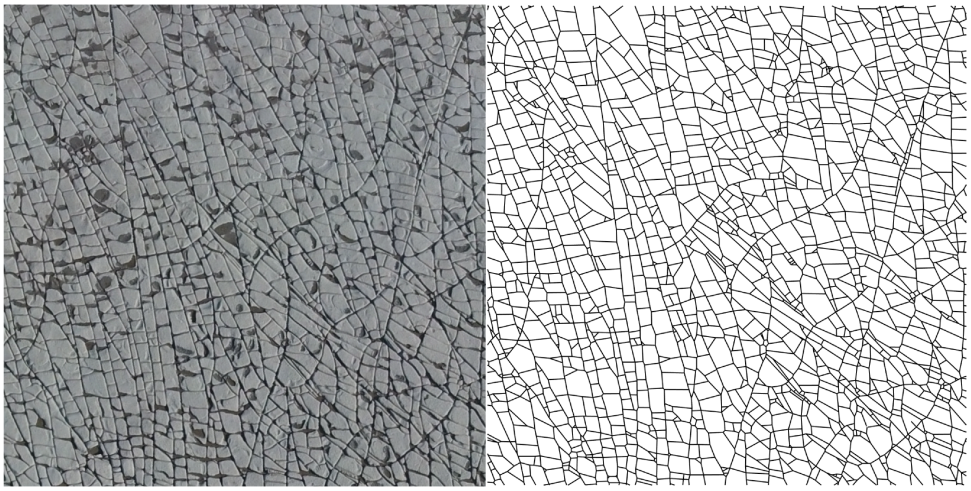

Figure F.3: An example of training data for the cGAN from Area 5. The Pix2Pix cGAN learns a pixel-to-pixel representation from source image on the left to the target image on the right

cGAN Results

The Pix2Pix cGAN is applied to two new regions within the Bristol Channel outcrop (see Fig. F.2). An example result on four images from the two new regions is shown in Fig. F.4 and Fig. F.5. The output is a cartoon-like image which still needs to be vectorized.

The trace vectorization workflow described in Chapter 2 is applied to the output images. The quality of results compare well with the complex shearlet with the added advantage of lesser false positives. Further training of the cGAN using augmented training data by mirroring and rotations can improve the trace detection quality.

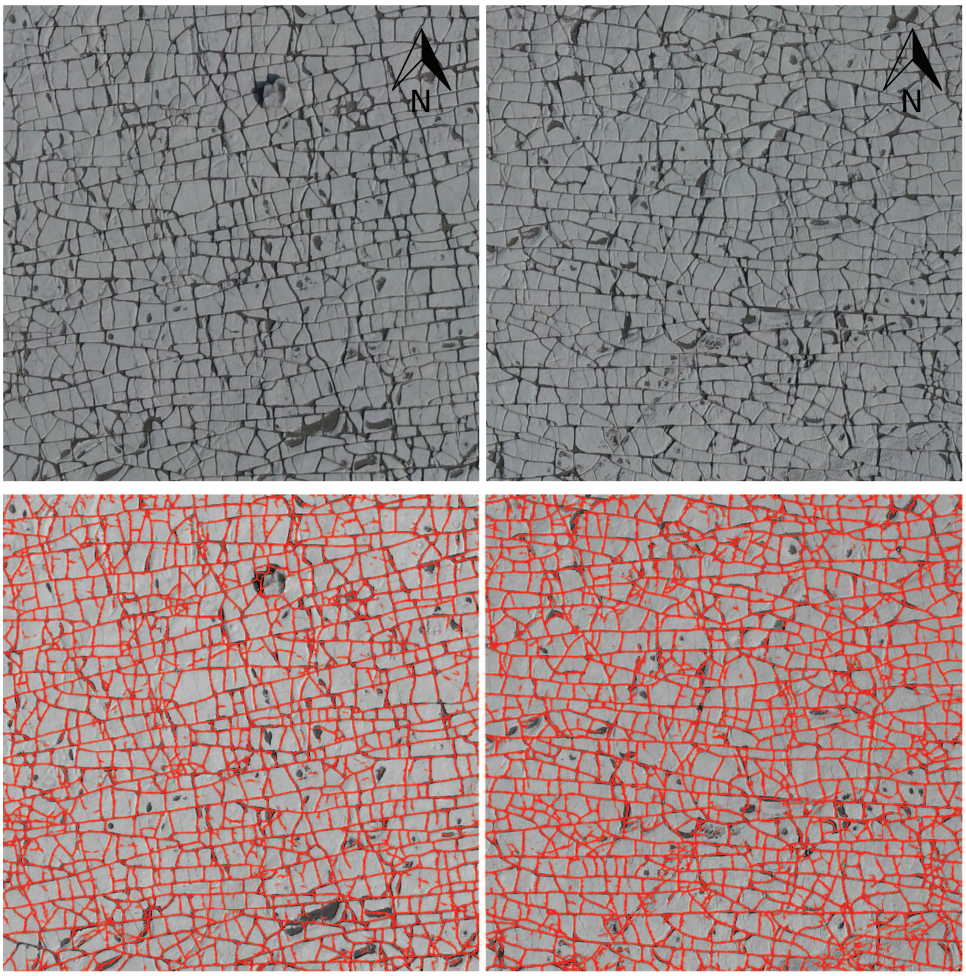

Figure F.4: Results of the Pix2Pix cGAN applied to an image tile from new Region 1

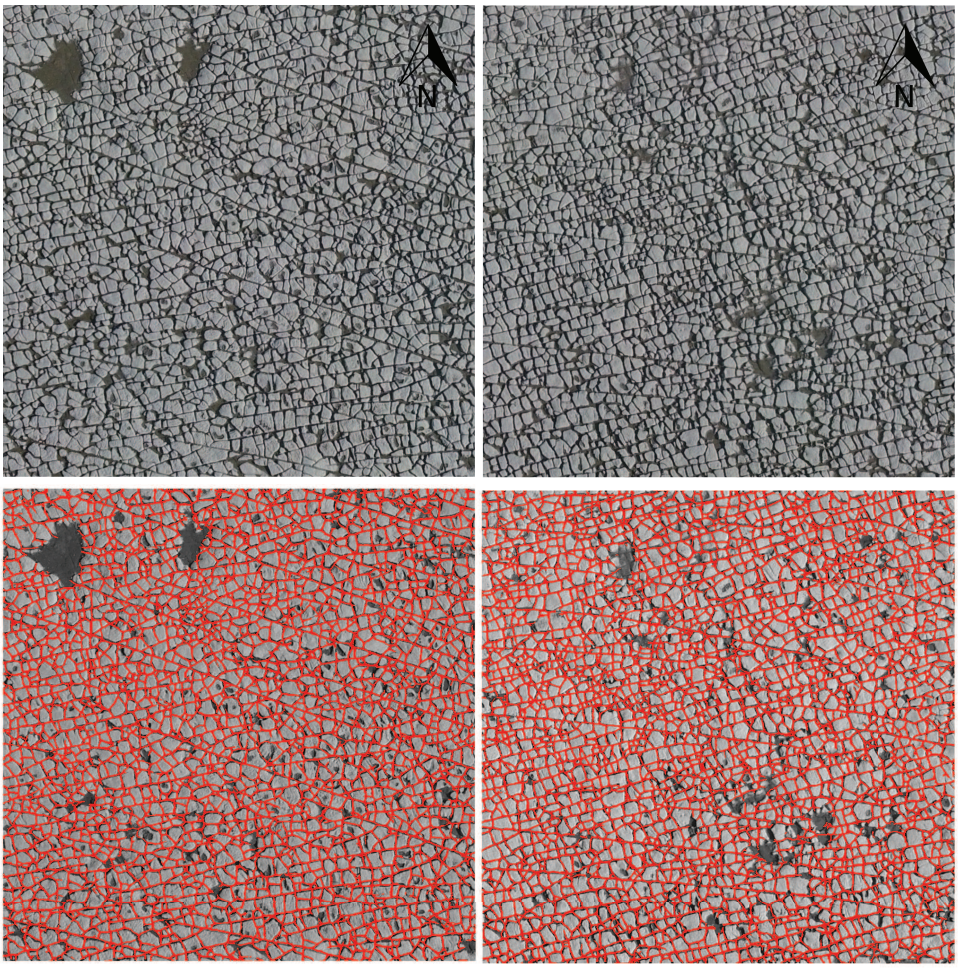

Figure F.5: Results of the Pix2Pix cGAN applied to an image tile from new Region 2

BIBLIOGRAPHY

A. Abdullah, S. Nassr, and A. Ghaleeb. Landsat etm-7 for lineament mapping using automatic extraction technique in the sw part of taiz area, yemen. *Global Journal of Human-Social Science Research (B)*, 13(3):35–38, May 2013. ISSN 2249-460X.

P. M. Adler and J.-F. Thovert. *Fractures and Fracture Networks*, volume 15 of *Theory and Applications of Transport in Porous Media*. Springer, Dordrecht, 1 edition, 1999. ISBN 978-90-481-5192-9. doi: 10.1007/978-94-017-1599-7.

AgiSoft PhotoScan Professional (Version 1.2.6). Photoscan professional (version 1.2.6) retrieved from http://www.agisoft.com/downloads/installer/, 2016*.

R. Aguilera. Well test analysis of naturally fractured reservoirs. *SPE Form Eval*, 2(03): 239–252, Oct. 1987. ISSN 0885-923X. URL https://doi.org/10.2118/13663-PA.

F. Aljuboori, P. W. M. Corbett, K. Bisdom, G. Bertotti, and S. Geiger. Using outcrop data for geological well test modelling in fractured reservoirs. In *77th EAGE Conference and Exhibition, Madrid, Spain, 1-5 June 2015*. We-N118-01, June 2015. doi: 10.3997/2214-4609.201413037.

C. Andresen, A. Hansen, R. Le Goc, P. Davy, and S. Hope. Topology of fracture networks. *Frontiers in Physics*, 1:7, 2013. ISSN 2296-424X. doi: 10.3389/fphy.2013.00007. URL https://www.frontiersin.org/article/10.3389/fphy.2013.00007.

B. J. Andrews, J. J. Roberts, Z. K. Shipton, S. Bigi, M. C. Tartarello, and G. Johnson. How do we see fractures? quantifying subjective bias in fracture data collection. *Solid Earth*, 10(2):487–516, 2019. doi: 10.5194/se-10-487-2019.

B. J. Andrews, Z. K. Shipton, R. Lord, and L. McKay. The growth of faults and fracture networks in a mechanically evolving, mechanically stratified rock mass: a case study from spireslack surface coal mine, scotland. *Solid Earth*, 11 (6):2119–2140, 2020. doi: 10.5194/se-11-2119-2020.

A. Awdal, D. Healy, and G. I. Alsop. Fracture patterns and petrophysical properties of carbonates undergoing regional folding: A case study from kurdistan, n iraq. *Marine and Petroleum Geology*, 71:149 – 167, 2016. ISSN 0264-8172. doi: 10.1016/j.marpetgeo.2015.12.017.

J. P. Bagrow and E. M. Bollt. An information-theoretic, all-scales approach to comparing networks. *Applied Network Science*, 4(1):45, 2019. ISSN 2364-8228. doi: 10.1007/s41109-019-0156-x.

J. P. Bagrow, E. M. Bollt, J. D. Skufca, and D. ben Avraham. Portraits of complex networks. *EPL (Europhysics Letters)*, 81(6):68004, feb 2008. doi: 10.1209/0295-5075/81/68004.

M. Barthelemy. *Morphogenesis of Spatial Networks*. Lecture Notes in Morphogenesis. Springer International Publishing, 2018 edition, 2018. ISBN 978-3-319-20564-9. doi: 10.1007/978-3-319-20565-6.

M. Bastian, S. Heymann, and M. Jacomy. Gephi: An open source software for exploring and manipulating networks, 2009. URL http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154.

M. Belayneh. *Palaeostress orientation inferred from surface morphology of joints on the southern margin of the Bristol Channel Basin, UK*, pages 243–255. Number 1. Geological Society of London, Special Publications, 2004. doi: 10.1144/GSL.SP.2004.231.01.14.

M. Belayneh and J. W. Cosgrove. Fracture-pattern variations around a major fold and their implications regarding fracture prediction using limited data: an example from the bristol channel basin. *Geological Society, London, Special Publications*, 231(1):89–102, 2004a. ISSN 0305-8719. doi: 10.1144/GSL.SP.2004.231.01.06.

M. Belayneh and J. W. Cosgrove. *Fracture-pattern variations around a major fold and their implications regarding fracture prediction using limited data: an example from the Bristol Channel Basin*, volume 231 (1), pages 89–102. Geological Society of London, Special Publications, 2004b. doi: 10.1144/GSL.SP.2004.231.01.06.

M. Belayneh, S. Geiger, and S. K. Matthai. Numerical simulation of water injection into layered fractured carbonate reservoir analogs. *AAPG Bulletin*, 90(10):1473–1493, 10 2006. ISSN 0149-1423. doi: 10.1306/05090605153.

N. Bellahsen, F. Mouthereau, A. Boutoux, M. Bellanger, O. Lacombe, L. Jolivet, and Y. Rolland. Collision kinematics in the western external alps. *Tectonics*, 33 (6)(6):1055–1088, June 2014. ISSN 0278-7407. doi: 10.1002/2013tc003453.

S. P. Bemis, S. Micklethwaite, D. Turner, M. R. James, S. Akciz, S. T. Thiele, and H. A. Bangash. Ground-based and uav-based photogrammetry: A multi-scale, high-resolution mapping tool for structural geology and paleoseismology. *Journal of Structural Geology*, 69(Part A):163–178, 2014. ISSN 0191-8141. doi: 10.1016/j.jsg.2014.10.007.

L. Berio, F. Balsamo, S. Mittempergher, M. Mozafari, F. Storti, A. Bistacchi, P.-O. Bruna, and G. Bertotti. Deformation pattern in the thrust-related Parmelan Anticline (Bornes Massif, Subalpine Chains, Haute-Savoie, France): preliminary results. In *20th EGU General Assembly, Vienna, Austria, 8–13 April 2018*. 2018EGUGA..20.8923B, Apr. 2018. URL http://resolver.tudelft.nl/uuid:259ea083-7bac-43d8-a4c5-34e79cc001e6.

B. Berkowitz. Characterizing flow and transport in fractured geological media: A review. *Advances in Water Resources*, 25(8):861–884, Aug. 2002. ISSN 0309-1708. doi: https://doi.org/10.1016/S0309-1708(02)00042-8.

G. Bertotti, P. Audra, A. Auler, F. H. Bezerra, S. de Hoop, C. Pontes, R. Prabhakaran, and R. Lima. The morro vermelho hypogenic karst system (brazil): Stratigraphy, fractures, and flow in a carbonate strike-slip fault zone with implications for carbonate reservoirs. *AAPG Bulletin*, 104 (10):2029–2050, 2020. doi: 10.1306/05212019150.

F. M. Bianchi, D. Grattarola, and C. Alippi. Spectral clustering with graph neural networks for graph pooling. In H. D. III and A. Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 874–883. PMLR, 13–18 Jul 2020. URL http://proceedings.mlr.press/v119/bianchi20a.html.

K. Bisdom, G. Bertotti, and H. M. Nick. A geometrically based method for predicting stress-induced fracture aperture and flow in discrete fracture networks. *AAPG Bulletin*, 100(7):1075–1097, Sept. 2016. ISSN 0149-1423. doi: 10.1306/02111615127.

K. Bisdom, H. M. Nick, and G. Bertotti. An integrated workflow for stress and flow modelling using outcrop-derived discrete fracture networks. *Computers & Geosciences*, 103:21–35, 2017. ISSN 0098-3004. doi: 10.1016/j.cageo.2017.02.019.

A. Bistacchi, F. Balsamo, F. Storti, M. Mozafari, R. Swennen, J. Solum, C. Tueckmantel, and C. Taberner. Photogrammetric digital outcrop reconstruction, visualization with textured surfaces, and three-dimensional structural analysis and modeling: Innovative methodologies applied to fault-related dolomitization (vajont limestone, southern alps, italy). *Geosphere*, 11 (6)(6):2031–2048, Nov. 2015. ISSN 1553-040X. doi: 10.1130/GES01005.1.

A. Bistacchi, S. Mittempergher, M. Martinelli, and F. Storti. On a new robust workflow for the statistical and spatial analysis of fracture data collected with scanlines (or the importance of stationarity). *Solid Earth*, 11(6):2535–2547, 2020. doi: 10.5194/se-11-2535-2020.

V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008 (10):P10008, oct 2008. doi: 10.1088/1742-5468/2008/10/p10008.

Q. Boersma, R. Prabhakaran, F. H. Bezerra, and G. Bertotti. Linking natural fractures to karst cave development: a case study combining drone imagery, a natural cave network and numerical modelling. *Petroleum Geoscience*, 2019. ISSN 1354-0793. doi: 10.1144/petgeo2018-151.

D. Bolkas, I. Vazaios, A. Peidou, and N. Vlachopoulos. Detection of rock discontinuity traces using terrestrial lidar data and space-frequency transforms. *Geotechnical and Geological Engineering*, 36 (3):1745–1765, Jun 2018. ISSN 1573-1529. doi: 10.1007/s10706-017-0430-6.

C. Bond, A. Gibbs, Z. Shipton, and S. Jones. What do you think this is?: "conceptual uncertainty" in geoscience interpretation. *GSA Today*, 17 (11):4–10, 12 2007. ISSN 1052-5173. doi: 10.1130/GSAT01711A.1.

C. Bond, G. Johnson, and J. Ellis. Structural model creation: the impact of data type and creative space on geological reasoning and interpretation. *Geological Society of London Special Publications*, 421 (1):83–97, 2015. doi: doi:10.1144/SP421.4.

S. Bonetto, A. Facello, A. M. Ferrero, and G. Umili. A tool for semi-automatic linear feature detection based on dtm. *Computers & Geosciences*, 75:1 – 12, 2015. ISSN 0098-3004. doi: 10.1016/j.cageo.2014.10.005.

S. Bonetto, A. Facello, and U. Gessica. A new application of curvatool semi-automatic approach to qualitatively detect geological lineaments. *Environmental and Engineering Geoscience*, 23(3):179–190, 2017. doi: 10.2113/gseegeosci.23.3.179.

B. Bourbiaux, R. Basquet, J. M. Daniel, L. Hu, S. Jenni, G. Lange, and P. Rasolofosaon. Fractured reservoirs modelling: a review of the challenges and some recent solutions. *First Break*, 23(9), 2005a. ISSN 0263-5046. doi: https://doi.org/10.3997/1365-2397.2005018.

B. Bourbiaux, R. Basquet, A. Lange, and L.-Y. Hu. Fractured reservoir management - recent progress and promising perspectives, Oct. 2005b.

Bourbiaux, B. Fractured reservoir simulation: a challenging and rewarding issue. *Oil Gas Sci. Technol. - Rev. IFP*, 65(2):227–238, 2010. doi: 10.2516/ogst/2009063.

S. Bourne and E. Willemse. Elastic stress control on the pattern of tensile fracturing around a small fault network at nash point, uk. *Journal of Structural Geology*, 23(11):1753–1770, 2001. ISSN 0191-8141. doi: 10.1016/S0191-8141(01)00027-X.

M. Brooks, P. M. Trayner, and T. J. Trimble. Mesozoic reactivation of variscan thrusting in the bristol channel area, uk. *Journal of the Geological Society*, 145(3):439–444, 1988. ISSN 0016-7649. doi: 10.1144/gsjgs.145.3.0439.

P. Bruna, R. Prabhakaran, G. Bertotti, J. Straubhaar, R. Plateaux, L. Maerten, G. Mariethoz, and M. Meda. The mps-based fracture network simulation method: Application to subsurface domain. *81st EAGE Conference and Exhibition, London 2019*, 2019(1):1–5, 2019a. ISSN 2214-4609. doi: 10.3997/2214-4609.201901679.

P.-O. Bruna, A. P. C. Lavenu, C. Matonti, and G. Bertotti. Are stylolites fluid-flow efficient features? *Journal of Structural Geology*, 125:270–277, Aug. 2019b. ISSN 0191-8141. doi: 10.1016/j.jsg.2018.05.018.

P.-O. Bruna, J. Straubhaar, R. Prabhakaran, G. Bertotti, K. Bisdom, G. Mariethoz, and M. Meda. A new methodology to train fracture network simulation using multiple-point statistics. *Solid Earth*, 10 (2)(2):537–559, 2019c. doi: 10.5194/se-10-537-2019.

G. Callatay. Hough transform application to natural fracture networks: Detection, characterization and simulation. Master's thesis, Delft University of Technology, 2016. URL http://resolver.tudelft.nl/uuid:e7bea512-6b64-455c-961b-ceeff573c79c.

E. J. Candès and D. L. Donoho. Continuous curvelet transform: I. resolution of the wavefront set. *Applied and Computational Harmonic Analysis*, 19(2):162 – 197, 2005. ISSN 1063-5203. doi: 10.1016/j.acha.2005.02.003.

E. J. Candès and F. Guo. New multiscale transforms, minimum total variation synthesis: applications to edge-preserving image reconstruction. *Signal Processing*, 82(11):1519–1543, Nov. 2002. ISSN 0165-1684. doi: 10.1016/S0165-1684(02)00300-6.

J.-B. Cordonnier and A. Loukas. Extrapolating paths with graph neural networks. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, pages 2187–2194. International Joint Conferences on Artificial Intelligence Organization, 7 2019. doi: 10.24963/ijcai.2019/303.

G. Csardi and T. Nepusz. The igraph software package for complex network research. *InterJournal*, Complex Systems:1695, 2006. URL https://igraph.org.

C. J. Dart, K. McClay, and P. N. Hollings. 3d analysis of inverted extensional fault systems, southern bristol channel basin, uk. *Geological Society, London, Special Publications*, 88(1):393–413, 1995. ISSN 0305-8719. doi: 10.1144/GSL.SP.1995.088.01.21.

I. Daubechies. *Ten Lectures on Wavelets, CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, USA, 1999 edition, 1992. doi: 10.1137/1.9781611970104.

F. D. Day-Lewis, L. D. Slater, J. Robinson, C. D. Johnson, N. Terry, and D. Werkema. An overview of geophysical technologies appropriate for characterization and monitoring at fractured-rock sites. *Journal of Environmental Management*, 204(Part 2):709–720, Dec. 2017. ISSN 0301-4797. doi: 10.1016/j.jenvman.2017.04.033.

M. De Berg, M. Van Kreveld, M. Overmars, and O. C. Schwarzkopf. *Computational Geometry. Algorithms and Applications*. Springer, Berlin, Heidelberg, 2000. ISBN 978-3-662-04247-2. doi: 10.1007/978-3-662-04245-8.

W. S. Dershowitz and H. H. Herda. Interpretation of fracture spacing and intensity. In *The 33rd U.S Symposium on Rock Mechanics (USRMS), 3-5 June, Santa Fe, New Mexico*, number ARMA-92-0757, Sept. 1992.

C. V. Deutsch. *Geostatistical Reservoir Modeling*. Oxford University Press, 1st edition edition, 2002. ISBN 0195138066.

M. N. Do and M. Vetterli. The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*, 14(12):2091–2106, Dec 2005. ISSN 1057-7149. doi: 10.1109/TIP.2005.859376.

D. L. Donoho. Wedgelets: Nearly minimax estimation of edges. *The Annals of Statistics*, 27(3):859–897, 1999. ISSN 00905364. doi: 10.1214/aos/1018031261.

J. Donovan and A. Lebaron. A comparison of photogrammetry and laser scanning for the purpose of automated rock mass characterization. In *43rd U.S. Rock Mechanics Symposium & 4th U.S. - Canada Rock Mechanics Symposium, 28 June-1 July 2009, Asheville, North Carolina, USA*. ARMA-09-122, 2009. URL https://www.onepetro.org/conference-paper/ARMA-09-122.

D. Doolaeghe, P. Davy, J. Hyman, and C. Darcel. Graph-based flow modeling approach adapted to multiscale discrete-fracture-network models. *PHYSICAL REVIEW E*, 102: 53312, Nov. 2020.

D. Douglas and T. Peucker. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 10(2):112–122, 1973. doi: 10.3138/FM57-6770-U75U-7727.

W. Dreybrodt, F. Gabrovsek, D. Romanov, S. Bauer, and S. Birk. *Processes of Speleogenesis: A Modeling Approach*. Zalozhba, 2005. ISBN 9616500910.

R. O. Duda and P. E. Hart. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, Jan. 1972. ISSN 0001-0782. doi: 10.1145/361237.361242.

E. Dupont, T. Zhang, P. Tilke, L. Liang, and W. Bailey. Generating realistic geology conditioned on physical measurements with generative adversarial networks. In *35th International Conference on Machine Learning, Stockholm, Sweden, 10 - 15 July 2018*, number 1802.03065 in 35th International Conference on Machine Learning, 2018. URL https://arxiv.org/abs/1802.03065v3.

A. Dwarkasing. 3d fracture analyses of various rock samples through x-ray microtomography. Master's thesis, Delft University of Technology, 2016. URL http://resolver.tudelft.nl/uuid:aefe6746-9788-4c93-ac4c-f80250e6a12c.

F. Emmert-Streib, M. Dehmer, and Y. Shi. Fifty years of graph matching, network alignment and network comparison. *Information Sciences*, 346-347:180 – 197, 2016. ISSN 0020-0255. doi: https://doi.org/10.1016/j.ins.2016.01.074.

T. Engelder and D. C. Peacock. Joint development normal to regional compression during flexural-flow folding: the lilstock buttress anticline, somerset, england. *Journal of Structural Geology*, 23(2):259 – 277, 2001. ISSN 0191-8141. doi: 10.1016/S0191-8141(00)00095-X.

R. A. Ennes-Silva, F. H. R. Bezerra, F. C. C. Nogueira, F. Balsamo, A. Klimchouk, C. L. Cazarin, and A. S. Auler. Superposed folding and associated fracturing influence hypogene karst development in neoproterozoic carbonates, são francisco craton, brazil. *Tectonophysics*, 666:244–259, Jan. 2016. ISSN 0040-1951. doi: 10.1016/j.tecto.2015.11.006.

E. Estrada and M. Sheerin. Random neighborhood graphs as models of fracture networks on rocks: Structural and dynamical analysis. *Applied Mathematics and Computation*, 314:360 – 379, 2017. ISSN 0096-3003. doi: 10.1016/j.amc.2017.06.018.

B. Everitt, S. Landau, M. Leese, and D. Stahl. *Cluster Analysis*. Wiley Series in Probability and Statistics. Wiley, 5 edition, 2011. ISBN 978-0-470-74991-3.

M. Felsberg and G. Sommer. The monogenic signal. *IEEE Transactions on Signal Processing*, 49(12):3136–3144, Dec. 2001. ISSN 1053-587X. doi: 10.1109/78.969520.

M. H. Feng. *Topological tools for understanding complex systems*. phdthesis, 2020. URL https://escholarship.org/uc/item/1t32m3z7.

M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2): 298–305, 1973. doi: 10.21136/CMJ.1973.101168.

D. Ford. Jovan cvijić and the founding of karst geomorphology. *Environmental Geology*, 51(5):675–684, Jan. 2007. ISSN 1432-0495. doi: 10.1007/s00254-006-0379-x.

D. Ford and P. Williams. *Karst Hydrogeology and Geomorphology*. Wiley, 2007. ISBN 9780470849965. doi: 10.1002/9781118684986.

M. Gidon. Vues nouvelles sur la structure des massifs des bornes et des bauges orientales. *Géologie Alpine*, 72:35–39, 1996. URL http://geologie-alpine.ujf-grenoble.fr/articles/GA_1996__72__35_0.pdf.

M. Gidon. Failles extensives antérieures au plissement dans les massifs subalpins: un exemple nouveau dans le massif des bornes (france). *Géologie Alpine*, 74:91–96, 1998. URL http://geologie-alpine.ujf-grenoble.fr/articles/GA_1998__74__91_0.pdf.

P. Gillespie, C. B. Howard, J. J. Walsh, and J. Watterson. Measurement and characterisation of spatial distribution of fractures. *Tectonophysics*, 226(1-4):113–141, Nov. 1993. doi: 10.1016/0040-1951(93)90114-Y.

P. Gillespie, E. Monsen, L. Maerten, D. Hunt, J. Thurmond, and D. Tuck. Fractures in Carbonates: From Digital Outcrops to Mechanical Models. In *Outcrops Revitalized: Tools, Techniques and Applications*. SEPM Society for Sedimentary Geology, 01 2011a. ISBN 9781565762930. doi: 10.2110/sepmcsp.10.137.

P. A. Gillespie, E. Monsen, L. Maerten, D. W. Hunt, J. Thurmond, and D. Tuck. *Fractures in Carbonates: From Digital Outcrops to Mechanical Models*, volume 10, pages 137–147. SEPM Society for Sedimentary Geology, 2011 edition, 2011b. ISBN 9781565763067. doi: 10.2110/sepmcsp.10.003.

R. Glen, P. Hancock, and A. Whittaker. Basin inversion by distributed deformation: the southern margin of the bristol channel basin, england. *Journal of Structural Geology*, 27(12):2113 – 2134, 2005. ISSN 0191-8141. doi: 10.1016/j.jsg.2005.08.006.

G. Godefroy, G. Caumon, G. Laurent, and F. Bonneau. Structural interpretation of sparse fault data using graph theory and geological rules. *Mathematical Geosciences*, 51(8): 1091–1107, 2019. ISSN 1874-8953. doi: 10.1007/s11004-019-09800-0.

I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, 08 - 13 December 2014*, NIPS'14, pages 2672–2680, Cambridge, MA, USA, 2014. MIT Press. URL http://dl.acm.org/citation.cfm?id=2969033.2969125.

P. Grohs, S. Keiper, G. Kutyniok, and M. Schäfer. Alpha-molecules. *Applied and Computational Harmonic Analysis*, 41(1):297–336, July 2016. ISSN 1063-5203. doi: 10.1016/j.acha.2015.10.009.

J. T. Guimarães, A. Misi, A. J. Pedreira, and J. M. L. Dominguez. *The Bebedouro Formation, Una Group, Bahia (Brazil), The Geological Record of Neoproterozoic Glaciations, Geological Society, London, Memoirs*, volume 36, chapter 47, pages 503–508. Geological Society of London, 2011 edition, Jan. 2011. doi: 10.1144/M36.47.

K. Guo, G. Kutyniok, and D. Labate. Sparse multidimensional representation using anisotropic dilation and shear operators. In G. Chen and M.-J. Lai, editors, *Wavelets and Splines: Athens 2005, International Conference on "Interactions between Wavelets and Splines", 16-19 May 2005, Athens, Georgia*, pages 189–201. Nashboro Press, 2006, 2005.

L. Guo, J.-P. Latham, and J. Xiang. A numerical study of fracture spacing and through-going fracture formation in layered rocks. *International Journal of Solids and Structures*, 110-111:44 – 57, 2017. ISSN 0020-7683. doi: 10.1016/j.ijsolstr.2017.02.004.

A. A. Hagberg, D. A. Schult, and P. J. Swart. Exploring network structure, dynamics, and function using networkx. In G. Varoquaux, T. Vaught, and J. Millman, editors, *Proceedings of the 7th Python in Science Conference*, pages 11 – 15, Pasadena, CA USA, 2008.

P. Hancock and T. Engelder. Neotectonic joints. *GSA Bulletin*, 101(10):1197–1208, 10 1989. ISSN 0016-7606. doi: 10.1130/0016-7606(1989)101<1197:NJ>2.3.CO;2.

J. R. Hanke, M. P. Fischer, and R. M. Pollyea. Directional semivariogram analysis to identify and rank controls on the spatial variability of fracture networks. *Journal of Structural Geology*, 108:34 – 51, 2018. ISSN 0191-8141. doi: 10.1016/j.jsg.2017.11.012. Spatial arrangement of fractures and faults.

H. Hartle, B. Klein, S. McCabe, A. Daniels, G. St-Onge, C. Murphy, and L. Hébert-Dufresne. Network comparison and the within-ensemble graph distance. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 476(2243): 20190744, 2020. doi: 10.1098/rspa.2019.0744.

S. Harwin and A. Lucieer. Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from unmanned aerial vehicle (uav) imagery. *Remote Sensing*, 4(6):1573–1599, 2012. ISSN 2072-4292. doi: 10.3390/rs4061573.

M. Hashim, S. Ahmad, M. A. M. Johari, and A. B. Pour. Automatic lineament extraction in a heavily vegetated region using landsat enhanced thematic mapper (etm+) imagery. *Advances in Space Research*, 51(5):874–890, Mar. 2013. ISSN 0273-1177. doi: 10.1016/j.asr.2012.10.004.

K. He, G. Gkioxari, P. Dollár, and R. Girshick. Mask r-cnn. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2980–2988, Oct 2017. doi: 10.1109/ICCV.2017.322.

D. Healy, R. E. Rizzo, D. G. Cornwell, N. J. Farrell, H. Watkins, N. E. Timms, E. Gomez-Rivas, and M. Smith. Fracpaq: A matlab™ toolbox for the quantification of fracture patterns. *Journal of Structural Geology*, 95:1 – 16, 2017. ISSN 0191-8141. doi: 10.1016/j.jsg.2016.12.003.

C. Heil, D. F. Walnut, and I. Daubechies. *Fundamental Papers in Wavelet Theory, Princeton University Press, Princeton, New Jersey, USA*. Princeton University Press, 2006. ISBN 9781400827268. URL http://www.jstor.org/stable/j.ctt7rt5v.

C. Hennig, M. Meila, F. Murtagh, and R. Rocci. *Handbook of Cluster Analysis*. Handbooks of Modern Statistical Methods. Chapman and Hall, CRC Press, 1 edition, 2016. ISBN 978-1-4665-5189-3.

J. Hillier, M. Smith, R. Armugam, I. Barr, C. Boston, C. Clark, J. Ely, A. Frankl, S. Greenwood, L. Gosselin, C. Hättestrand, K. Hogan, A. Hughes, S. Livingstone, H. Lovell, M. McHenry, Y. Munoz, X. Pellicer, R. Pellitero, C. Robb, S. Roberson, D. Ruther, M. Spagnolo, M. Standell, C. Stokes, R. Storrar, N. Tate, and K. Wooldridge. Manual mapping of drumlins in synthetic landscapes to assess operator effectiveness. *Journal of Maps*, 11(5):719–729, 2015. doi: 10.1080/17445647.2014.957251.

D. Hodgetts. Laser scanning and digital outcrop geology in the petroleum industry: A review. *Marine and Petroleum Geology*, 46:335–354, Sept. 2013. ISSN 0264-8172. doi: 10.1016/j.marpetgeo.2013.02.014.

J. N. Hooker and R. F. Katz. Vein spacing in extending, layered rock: The effect of synkinematic cementation. *American Journal of Science*, 315(6):557–588, 2015. ISSN 0002-9599. doi: 10.2475/06.2015.03. URL https://www.ajsonline.org/content/315/6/557.

P. Huggenberger and W. Wildi. La tectonique du massif des bornes (chaînes subalpines, haute-savoie, france). *Eclogae Geologicae Helvetiae*, 84(1):125–149, 1991. doi: 10.5169/seals-166766.

P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5967–5976, 2017. doi: 10.1109/CVPR.2017.632.

L. Jacques, A. Coron, P. Vandergheynst, and A. Rivoldini. The YAWTb toolbox : Yet Another Wavelet Toolbox, 2011. http://sites.uclouvain.be/ispgroup/yawtb.

H. A. Jelsma, M. J. de Wit, C. Thiart, P. H. G. M. Dirks, G. Viola, I. J. Basson, and E. Anckar. Preferential distribution along transcontinental corridors of kimberlites and related rocks of southern africa. *South African Journal of Geology*, 107(1-2):301–324, 2004. ISSN 1012-0750. doi: 10.2113/107.1-2.301.

E. Jonckheere, M. Lou, F. Bonahon, and Y. Baryshnikov. Euclidean versus hyperbolic congestion in idealized versus experimental networks. *Internet Mathematics*, 7(1):1–27, 2011. doi: 10.1080/15427951.2010.554320.

P. Kamerling. The geology and hydrocarbon habitat of the bristol channel basin. *Journal of Petroleum Geology*, 2(1):75–93, 1979. doi: 10.1111/j.1747-5457.1979.tb00693.x.

H. Karbalaali, A. Javaherian, S. Dahlke, R. Reisenhofer, and S. Torabi. Seismic channel edge detection using 3d shearlets—a study on synthetic and real channelised 3d seismic data. *Geophysical Prospecting*, 66(7):1272–1289, 2018. doi: 10.1111/1365-2478.12629.

L. Kaufman. *Finding groups in data : an introduction to cluster analysis*. Wiley Series in Probability and Statistics. Wiley, New York, 1990 edition, 1990. ISBN 9780471878766. doi: 10.1002/9780470316801.

P. Kelly, D. Peacock, D. Sanderson, and A. McGurk. Selective reverse-reactivation of normal faults, and deformation around reverse-reactivated faults in the mesozoic of the somerset coast. *Journal of Structural Geology*, 21(5):493 – 509, 1999. ISSN 0191-8141. doi: 10.1016/S0191-8141(99)00041-3.

E. J. King, R. Reisenhofer, J. Kiefer, W.-Q. Lim, Z. Li, and G. Heygster. Shearlet-based edge detection: flame fronts and tidal flats. In A. G. Tescher, editor, *Applications of Digital Image Processing XXXVIII, SPIE Optical Engineering + Applications, 2015, San Diego, California, United States*, number 959905 in Proc. SPIE. 9599, Sept. 2015. doi: 10.1117/12.2188652.

A. Klimchouk. Morphogenesis of hypogenic caves. *Geomorphology*, 106(1):100 – 117, 2009. ISSN 0169-555X. doi: 10.1016/j.geomorph.2008.09.013. Recent developments in surface and subsurface karst geomorphology.

A. Klimchouk. *Hypogene speleogenesis: Hydrogeological and Morphogenetic Perspective*. National Cave and Karst Research Institute „ 2 edition, 2011. ISBN 0-9795422-0-0. URL https://digital.lib.usf.edu/?k26.1547.

A. Klimchouk, A. N. Palmer, J. D. Waele, A. S. Auler, and P. Audra. *Hypogene Karst Regions and Caves of the World*. Springer, Cham, 1 edition, 2017. ISBN 978-3-319-53347-6. doi: 10.1007/978-3-319-53348-3.

P. Kovesi. Image features from phase congruency. *Videre: Journal of computer vision research*, 1:1–26, 1999. URL http://ftp.cs.rochester.edu/~brown/Videre/001/articles/v1n3001.pdf.

P. Kovesi. Phase congruency: A low-level image invariant. *Psychological Research*, 64 (2) (2):136–148, Dec 2000. ISSN 1430-2772. doi: 10.1007/s004260000024.

D. Krupnik and S. Khan. Close-range, ground-based hyperspectral imaging for mining applications at various scales: Review and case studies. *Earth-Science Reviews*, 198: 102952, 2019. ISSN 0012-8252. doi: 10.1016/j.earscirev.2019.102952.

G. Kutyniok and D. Labate, editors. *Shearlets Multiscale Analysis for Multivariate Data, Applied and Numerical Harmonic Analysis Book Series, Birkhäuser, Basel*. Applied and Numerical Harmonic Analysis. Birkhäuser Basel, 2012 edition, 2012. ISBN 978-0-8176-8316-0. doi: 10.1007/978-0-8176-8316-0.

G. Kutyniok, W.-Q. Lim, and R. Reisenhofer. Shearlab 3d: Faithful digital shearlet transforms based on compactly supported shearlets. *ACM Trans. Math. Softw.*, 42(1):5:1–5:42, Jan. 2016. ISSN 0098-3500. doi: 10.1145/2740960.

D. Labate, W.-Q. Lim, G. Kutyniok, and G. Weiss. Sparse multidimensional representation using shearlets. In *Wavelets XI, Optics and Photonics 2005, 31 July - 4 August 2005, San Diego, California, United States*, number 59140U in Wavelets, Aug. 2005. doi: 10.1117/12.613494.

J. Lamarche, J. Borgomano, B. Caline, F. Gisquet, S. Rigaud, S. Schröder, S. Viseur, O. J. Martinsen, A. J. Pulham, P. D. W. Haughton, and M. D. Sullivan. Characterization of fault-related dolomite bodies in carbonate reservoirs using lidar scanning, Nov. 2011. ISSN 9781565762930. URL https://doi.org/10.2110/sepmcsp.10.175.

J. Lamarche, A. P. Lavenu, B. D. Gauthier, Y. Guglielmi, and O. Jayet. Relationships between fracture patterns, geodynamics and mechanical stratigraphy in carbonates (south-east basin, france). *Tectonophysics*, 581:231 – 245, 2012. ISSN 0040-1951. doi: 10.1016/j.tecto.2012.06.042. Crustal Stresses, Fractures, and Fault Zones: The Legacy of Jacques Angelier.

S. E. Laubach, J. E. Olson, and M. R. Gross. Mechanical and fracture stratigraphy. *AAPG Bulletin*, 93(11):1413–1426, Sept. 2009. ISSN 0149-1423. doi: 10.1306/07270909094.

S. E. Laubach, J. Lamarche, B. D. Gauthier, and W. M. Dunne. Editorial: Spatial arrangement of faults and opening-mode fractures. *Journal of Structural Geology*, 108: 1, 2018a. ISSN 0191-8141. doi: 10.1016/j.jsg.2018.01.008. Spatial arrangement of fractures and faults.

S. E. Laubach, J. Lamarche, B. D. M. Gauthier, W. M. Dunne, and D. J. Sanderson. Spatial arrangement of faults and opening-mode fractures. *Journal of Structural Geology*, 108: 2–15, Mar. 2018b. ISSN 0191-8141. doi: 10.1016/j.jsg.2018.01.008.

S. E. Laubach, R. H. Lander, L. J. Criscenti, L. M. Anovitz, J. L. Urai, R. M. Pollyea, J. N. Hooker, W. Narr, M. A. Evans, S. N. Kerisit, J. E. Olson, T. Dewers, D. Fisher, R. Bodnar, B. Evans, P. Dove, L. M. Bonnell, M. P. Marder, and L. Pyrak-Nolte. The role of chemistry in fracture pattern development and opportunities to advance interpretations of geological materials. *Reviews of Geophysics*, 57(3):1065–1111, 2019. doi: 10.1029/2019RG000671.

E. Lavoine, P. Davy, C. Darcel, and R. Munier. A discrete fracture network model with stress-driven nucleation: Impact on clustering, connectivity, and topology. *Frontiers in Physics*, 8:9, 2020. ISSN 2296-424X. doi: 10.3389/fphy.2020.00009.

E. Le Pennec and S. Mallat. Sparse geometric image representations with bandelets. *IEEE Transactions on Image Processing*, 14(4):423–438, Apr. 2005. ISSN 1057-7149. doi: 10.1109/TIP.2005.843753.

Q. Lei, J.-P. Latham, and C.-F. Tsang. The use of discrete fracture networks for modelling coupled geomechanical and hydrological behaviour of fractured rocks. *Computers and Geotechnics*, 85:151–176, Jan. 2017. doi: 10.1016/j.compgeo.2016.12.024.

B. Lepillier, K. Yoshioka, F. Parisio, R. Bakker, and D. Bruhn. Variational phase-field modeling of hydraulic fracture interaction with natural fractures and application to enhanced geothermal systems. *Journal of Geophysical Research: Solid Earth*, 125(7): e2020JB019856, 2020. doi: 10.1029/2020JB019856.

J. J. Long, R. R. Jones, and S. E. Daniels. Reducing uncertainty in fracture modelling: assessing the sensitivity of inputs from outcrop analogues. In *The Geology of Fractured Reservoirs, 24-25 October 2018, The Geological Society of London*. The Geological Society of London, Oct. 2018. Oral Presentation.

R. J. H. Loosveld and R. C. M. W. Franssen. Extensional vs. shear fractures: Implications for reservoir characterisation, 1992. ISSN 978-1-55563-498-8.

R. Louf and M. Barthelemy. A typology of street patterns. *Journal of The Royal Society Interface*, 11(101):20140924, 2014. doi: 10.1098/rsif.2014.0924.

S. B. Mabee, K. C. Hardcastle, and D. U. Wise. A method of collecting and analyzing lineaments for regional-scale fractured-bedrock aquifer studies. *Groundwater*, 32(6): 884–894, Nov. 1994. ISSN 0017-467X. doi: 10.1111/j.1745-6584.1994.tb00928.x.

L. Maerten, P. Gillespie, and J.-M. Daniel. Three-dimensional geomechanical modeling for constraint of subseismic fault simulation. *Aapg Bulletin*, 90(9):1337–1358, Sept. 2006. doi: 10.1306/03130605148.

S. Mallat and W. L. Hwang. Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, 38(2):617–643, 1992. ISSN 0018-9448. doi: 10.1109/18.119727.

T. Manzocchi. The connectivity of two-dimensional networks of spatially correlated fractures. *Water Resources Research*, 38(9):1–1–1–20, 2002. doi: 10.1029/2000WR000180.

R. Marrett, J. F. Gale, L. A. Gómez, and S. E. Laubach. Correlation analysis of fracture arrangement in space. *Journal of Structural Geology*, 108:16 – 33, 2018. ISSN 0191-8141. doi: 10.1016/j.jsg.2017.06.012. Spatial arrangement of fractures and faults.

A. Masoud and K. Koike. Applicability of computer-aided comprehensive tool (linda: Lineament detection and analysis) and shaded digital elevation model for characterizing and interpreting morphotectonic features from lineaments. *Computers & Geosciences*, 106:89–100, Sept. 2017. ISSN 0098-3004. doi: 10.1016/j.cageo.2017.06.006.

M. Mauldon and J. G. Mauldon. Fracture sampling on a cylinder: From scanlines to boreholes and tunnels. *Rock Mechanics and Rock Engineering*, 30(3):129–144, July 1997. ISSN 1434-453X. doi: 10.1007/BF01047389.

M. Mauldon, W. Dunne, and M. Rohrbaugh. Circular scanlines and circular windows: new tools for characterizing the geometry of fracture traces. *Journal of Structural Geology*, 23(2):247 – 258, 2001. ISSN 0191-8141. doi: 10.1016/S0191-8141(00)00094-8.

Q. Meng, J. Hooker, and J. Cartwright. Progressive accretion of antitaxial crystal fibres: Implications for the kinematics and dynamics of vein dilation. *Journal of Structural Geology*, 126:25–36, 2019. ISSN 0191-8141. doi: 10.1016/j.jsg.2019.05.006.

J. Molron, N. Linde, L. Baron, J.-O. Selroos, C. Darcel, and P. Davy. Which fractures are imaged with ground penetrating radar? results from an experiment in the Äspö hardrock laboratory, sweden. *Engineering Geology*, 273:105674, 2020. ISSN 0013-7952. doi: 10.1016/j.enggeo.2020.105674.

National Research Council. *Rock Fractures and Fluid Flow: Contemporary Understanding and Applications, Washington, DC*. The National Academies Press, Washington, DC, 1996 edition, 1996. ISBN 978-0-309-10371-8. doi: 10.17226/2309.

R. Nelson. *Geologic Analysis of Naturally Fractured Reservoirs*. Gulf Professional Publishing, 2 edition, 2001. ISBN 978-0-88415-317-7.

M. Nemčok, R. Gayer, and M. Miliorizos. Structural analysis of the inverted bristol channel basin: implications for the geometry and timing of fracture porosity. *Geological Society, London, Special Publications*, 88(1):355, Jan. 1995. doi: 10.1144/GSL.SP.1995.088.01.20.

M. E. J. Newman. Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Phys. Rev. E*, 64:016132, Jun 2001. doi: 10.1103/PhysRevE.64.016132.

M. E. J. Newman. Assortative mixing in networks. *Phys. Rev. Lett.*, 89:208701, Oct 2002. doi: 10.1103/PhysRevLett.89.208701.

M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69:026113, Feb 2004. doi: 10.1103/PhysRevE.69.026113.

T. Nishizeki and M. Rahman. *Planar Graph Drawing*. World Scientific Publishing, 2004. ISBN 981-256-033-5. doi: 10.1142/5648. URL https://www.worldscientific.com/doi/abs/10.1142/5648.

B. Nyberg, C. W. Nixon, and D. J. Sanderson. Networkgt: A gis tool for geometric and topological analysis of two-dimensional fracture networks. *Geosphere*, 14(4):1618–1634, Sept. 2018. ISSN 1553-040X. doi: 10.1130/GES01595.1.

N. E. Odling. Scaling and connectivity of joint systems in sandstones from western norway. *Journal of Structural Geology*, 19(10):1257–1271, 1997. ISSN 0191-8141. doi: 10.1016/S0191-8141(97)00041-2.

J. E. Olson, S. E. Laubach, and R. H. Lander. Natural fracture characterization in tight gas sandstones: Integrating mechanics and diagenesis. *AAPG Bulletin*, 93(11):1535–1549, Nov. 2009. ISSN 0149-1423. doi: 10.1306/08110909100.

N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979. ISSN 0018-9472. doi: 10.1109/TSMC.1979.4310076.

R. Palamakumbura, M. Krabbendam, K. Whitbread, and C. Arnhardt. Data acquisition by digitizing 2-d fracture networks and topographic lineaments in geographic information systems: further development and applications. *Solid Earth*, 11(5):1731–1746, 2020. doi: 10.5194/se-11-1731-2020.

M. Passchier, C. W. Passchier, C. Weismüller, and J. L. Urai. The joint sets on the lilstock benches, uk. observations based on mapping a full resolution uav-based image. *Journal of Structural Geology*, 147:104332, 2021. ISSN 0191-8141. doi: 10.1016/j.jsg.2021.104332.

D. Peacock and D. Sanderson. Strike-slip relay ramps. *Journal of Structural Geology*, 17 (10):1351 – 1360, 1995. ISSN 0191-8141. doi: 10.1016/0191-8141(95)97303-W.

D. Peacock, D. Sanderson, and A. Rotevatn. Relationships between fractures. *Journal of Structural Geology*, 106:41 – 53, 2018. ISSN 0191-8141. doi: 10.1016/j.jsg.2017.11.010.

D. Peacock, D. Sanderson, E. Bastesen, A. Rotevatn, and T. Storstein. Causes of bias and uncertainty in fracture network analysis. *Norwegian Journal of Geology*, 99(1)(1), 2019. doi: 10.17850/njg99-1-06.

D. C. P. Peacock. *Differences between veins and joints using the example of the Jurassic limestones of Somerset*, volume 231, pages 209–221. Geological Society of London, Special Publications, 2004. doi: 10.1144/GSL.SP.2004.231.01.12.

Z. G. Philip, J. W. Jennings, J. E. Olson, S. E. Laubach, and J. Holder. Modeling coupled fracture-matrix fluid flow in geomechanically simulated fracture networks. *SPE Res Eval & Eng*, 8(04):300–309, Oct. 2005. ISSN 1094-6470. doi: 10.2118/77340-PA.

G. J. Potts and S. M. Reddy. Application of younging tables to the construction of relative deformation histories—1: Fracture systems. *Journal of Structural Geology*, 22(10): 1473–1490, 2000. ISSN 0191-8141. doi: 10.1016/S0191-8141(00)00044-4.

R. Prabhakaran. rahulprabhakaran/fracture-graphs (supplement to jsg manuscript 10.1016/j.jsg.2021.104405), 3 2021a.

R. Prabhakaran. rahulprabhakaran/Fracture-Fingerprint), Apr. 2021b.

R. Prabhakaran, P.-O. Bruna, G. Bertotti, and D. Smeulders. An automated fracture trace detection technique using the complex shearlet transform. *Solid Earth*, 10 (6):2137–2166, 2019a. doi: 10.5194/se-10-2137-2019.

R. Prabhakaran, P.-O. Bruna, G. Bertotti, D. Smeulders, and M. Meda. Fracture network patterns from the parmelan anticline, france. 4TU Centre for Research Data. Dataset, 2019b.

R. Prabhakaran, G. Bertotti, J. Urai, and D. Smeulders. Investigating spatial heterogeneity within fracture networks using hierarchical clustering and graph distance metrics. *Solid Earth Discussions*, 2021a. doi: 10.5194/se-2021-45.

R. Prabhakaran, J. Urai, G. Bertotti, C. Weismüller, and D. Smeulders. Large-scale natural fracture network patterns: Insights from automated mapping in the lilstock (bristol channel) limestone outcrops. *Journal of Structural Geology*, 150:104405, 2021b. ISSN 0191-8141. doi: 10.1016/j.jsg.2021.104405.

S. Priest and J. Hudson. Discontinuity spacings in rock. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 13(5):135 – 148, 1976. ISSN 0148-9062. doi: 10.1016/0148-9062(76)90818-4.

S. Priest and J. Hudson. Estimation of discontinuity spacing and trace length using scanline surveys. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 18(3):183–197, 1981. ISSN 0148-9062. doi: 10.1016/0148-9062(81)90973-6.

A. Procter and D. J. Sanderson. Spatial and layer-controlled variability in fracture networks. *Journal of Structural Geology*, 108:52 – 65, 2018. ISSN 0191-8141. doi: 10.1016/j.jsg.2017.07.008. Spatial arrangement of fractures and faults.

K. Rawnsley, D. Peacock, T. Rives, and J.-P. Petit. Joints in the mesozoic sediments around the bristol channel basin. *Journal of Structural Geology*, 20(12):1641 – 1661, 1998. ISSN 0191-8141. doi: 10.1016/S0191-8141(98)00070-4.

R. Reisenhofer. The complex shearlet transform and applications to image quality assessment. Master's thesis, Technical University of Berlin, 2014.

R. Reisenhofer and E. J. King. Edge, ridge, and blob detection with symmetric molecules. *SIAM Journal on Imaging Sciences*, 12(4):1585–1626, 2019. doi: 10.1137/19M1240861.

R. Reisenhofer, J. Kiefer, and E. J. King. Shearlet-based detection of flame fronts. *Experiments in Fluids*, 57:41, Feb. 2016. ISSN 1432-1114. doi: 10.1007/s00348-016-2128-6.

E. W. Remij, J. J. C. Remmers, F. Pizzocolo, D. M. J. Smeulders, and J. M. Huyghe. A partition of unity-based model for crack nucleation and propagation in porous media, including orthotropic materials. *Transport in Porous Media*, 106(3):505–522, 2015. ISSN 1573-1634. doi: 10.1007/s11242-014-0399-z.

Ren and Malik. Learning a classification model for segmentation. In *Proceedings Ninth IEEE International Conference on Computer Vision*, pages 10–17 vol.1, 2003. doi: 10.1109/ICCV.2003.1238308.

J. Rohrbaugh, M. B., W. M. Dunne, and M. Mauldon. Estimating fracture trace intensity, density, and mean length using circular scan lines and windows. *AAPG Bulletin*, 86(12):2089–2104, Sept. 2002. ISSN 0149-1423. doi: 10.1306/61EEDE0E-173E-11D7-8645000102C1865D.

P. Sander, T. B. Minor, and M. M. Chesley. Ground-water exploration based on lineament analysis and reproducibility tests. *Groundwater*, 35(5):888–894, Sept. 1997. ISSN 0017-467X. doi: 10.1111/j.1745-6584.1997.tb00157.x.

D. J. Sanderson and C. W. Nixon. The use of topology in fracture network characterization. *Journal of Structural Geology*, 72:55 – 66, 2015. ISSN 0191-8141. doi: 10.1016/j.jsg.2015.01.005.

D. J. Sanderson, D. C. Peacock, C. W. Nixon, and A. Rotevatn. Graph theory and the analysis of fracture networks. *Journal of Structural Geology*, 125:155–165, 2019. ISSN 0191-8141. doi: 10.1016/j.jsg.2018.04.011. Back to the future.

E. Santiago, J. X. Velasco-Hernández, and M. Romero-Salcedo. A descriptive study of fracture networks in rocks using complex network metrics. *Computers and Geosciences*, 88:97 – 114, 2016. ISSN 0098-3004. doi: 10.1016/j.cageo.2015.12.021.

T. Scheiber, O. Fredin, G. Viola, A. Jarna, D. Gasser, and R. Łapińska Viola. Manual extraction of bedrock lineaments from high-resolution lidar data: methodological bias and human perception. *Geologiska Föreningen i Stockholm (GFF)*, 137(4):362–372, Oct. 2015. ISSN 1103-5897. doi: 10.1080/11035897.2015.1085434.

T. A. Schieber, L. Carpi, A. Díaz-Guilera, P. M. Pardalos, C. Masoller, and M. G. Ravetti. Quantification of network structural dissimilarities. *Nature Communications*, 8(1): 13928, 2017. ISSN 2041-1723. doi: 10.1038/ncomms13928.

D. N. Scott and E. E. Wohl. Bedrock fracture influences on geomorphic process and form across process domains and scales. *Earth Surface Processes and Landforms*, 44(1):27–45, 2019. doi: 10.1002/esp.4473.

J. Shackleton, M. Cooke, and A. Sussman. Evidence for temporally changing mechanical stratigraphy and effects on joint-network architecture. *Geology*, 33(2):101–104, Feb. 2005. doi: 10.1130/G20930.1.

M. E. Snyder and J. W. Waldron. Fracture overprinting history using markov chain analysis: Windsor-kennetcook subbasin, maritimes basin, canada. *Journal of Structural Geology*, 108:80 – 93, 2018. ISSN 0191-8141. doi: 10.1016/j.jsg.2017.07.009. Spatial arrangement of fractures and faults.

D. A. Spielman and S.-H. Teng. Spectral partitioning works: Planar graphs and finite element meshes. *Linear Algebra and its Applications*, 421(2):284 – 305, 2007. ISSN 0024-3795. doi: 10.1016/j.laa.2006.07.020. Special Issue in honor of Miroslav Fiedler.

L. Spruženiece, M. Späth, J. L. Urai, E. Ukar, M. Selzer, B. Nestler, and A. Schwedt. Formation of wide-blocky calcite veins by extreme growth competition. *Journal of the Geological Society*, 178, 2021. ISSN 0016-7649. doi: 10.1144/jgs2020-104.

S. Srinivasan, J. Hyman, S. Karra, D. O'Malley, H. Viswanathan, and G. Srinivasan. Robust system size reduction of discrete fracture networks: a multi-fidelity method that preserves transport characteristics. *Computational Geosciences*, 22:1515—-1526, Dec. 2018. doi: 10.1007/s10596-018-9770-4.

S. Srinivasan, S. Karra, J. Hyman, H. Viswanathan, and G. Srinivasan. Model reduction for fractured porous media: a machine learning approach for identifying main flow

pathways. *Computational Geosciences*, 23(3):617–629, 2019. ISSN 1573-1499. doi: 10.1007/s10596-019-9811-7.

M. Tantardini, F. Ieva, L. Tajoli, and C. Piccardi. Comparing methods for comparing networks. *Scientific Reports*, 9(1):17557, 2019. ISSN 2045-2322. doi: 10.1038/s41598-019-53708-y.

S. T. Thiele, L. Grose, A. Samsu, S. Micklethwaite, S. A. Vollgger, and A. R. Cruden. Rapid, semi-automatic fracture and contact mapping for point clouds, images and geophysical data. *Solid Earth*, 8(6):1241–1253, 2017. doi: 10.5194/se-8-1241-2017.

R. N. Thomas, A. Paluszny, and R. W. Zimmerman. Effect of fracture growth velocity exponent on fluid flow through geomechanically-grown 3d fracture networks. In *2nd International Discrete Fracture Network Engineering Conference, 20-22 June 2018, Seattle, Washington, USA*, Seattle, Washington, USA, Nov. 2018. ARMA-DFNE-18-0239. URL https://www.onepetro.org/conference-paper/ARMA-DFNE-18-0239.

J.-F. Thovert, V. Mourzenko, and P. Adler. Percolation in three-dimensional fracture networks for arbitrary size and shape distributions. *Physical Review E*, 95(4):042112, Apr. 2017. doi: 10.1103/PhysRevE.95.042112.

V. A. Traag, L. Waltman, and N. J. van Eck. From louvain to leiden: guaranteeing well-connected communities. *Scientific Reports*, 9(1):5233, 2019. ISSN 2045-2322. doi: 10.1038/s41598-019-41695-z.

A. Tsitsulin, D. Mottin, P. Karras, A. Bronstein, and E. Müller. Netlsd: Hearing the shape of a graph. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '18, 2018.

A. Tsitsulin, J. Palowitch, B. Perozzi, and E. Müller. Graph clustering with graph neural networks, 2020. URL https://arxiv.org/abs/2006.16904.

C.-L. Tu, W.-L. Hwang, and J. Ho. Analysis of singularities from modulus maxima of complex wavelets. *IEEE Transactions on Information Theory*, 51(3):1049–1062, 2005. ISSN 0018-9448. doi: 10.1109/TIT.2004.842706.

D. Turner, A. Lucieer, and C. Watson. An automated technique for generating georectified mosaics from ultra-high resolution unmanned aerial vehicle (uav) imagery, based on structure from motion (sfm) point clouds. *Remote Sensing*, 4(5), 2012. ISSN 2072-4292. doi: 10.3390/rs4051392.

E. Ukar, S. E. Laubach, and J. N. Hooker. Outcrops as guides to subsurface natural fractures: Example from the nikanassin formation tight-gas sandstone, grande cache, alberta foothills, canada. *Marine and Petroleum Geology*, 103:255–275, May 2019. ISSN 0264-8172. doi: 10.1016/j.marpetgeo.2019.01.039.

L. Valentini, D. Perugini, and G. Poli. The 'small-world' nature of fracture/conduit networks: Possible implications for disequilibrium transport of magmas beneath mid-ocean ridges. *Journal of Volcanology and Geothermal Research*, 159(4):355–365, 2007a. ISSN 0377-0273. doi: 10.1016/j.jvolgeores.2006.08.002.

L. Valentini, D. Perugini, and G. Poli. The “small-world” topology of rock fracture networks. *Physica A: Statistical Mechanics and its Applications*, 377(1):323–328, 2007b. ISSN 0378-4371. doi: 10.1016/j.physa.2006.11.025.

V. Valliappan, J. J. C. Remmers, A. Barnhoorn, and D. M. J. Smeulders. A numerical study on the effect of anisotropy on hydraulic fractures. *Rock Mechanics and Rock Engineering*, 52(2):591–609, 2019. ISSN 1434-453X. doi: 10.1007/s00603-017-1362-4.

B. Van Hoorn. The south celtic sea/bristol channel basin: origin, deformation and inversion history. *Tectonophysics*, 137(1):309–334, 1987. ISSN 0040-1951. doi: 10.1016/0040-1951(87)90325-8. Special Issue: Compressional Intra-Plate Deformations in the Alpine Foreland.

M. Van Steen. *Graph Theory and Complex Networks: An Introduction*. 1 edition, 2010. ISBN 9081540610.

Y. Vasuki, E.-J. Holden, P. Kovesi, and S. Micklethwaite. Semi-automatic mapping of geological structures using uav-based photogrammetric data: An image analysis approach. *Computers & Geosciences*, 69:22–32, Aug. 2014. ISSN 0098-3004. doi: 10.1016/j.cageo.2014.04.012.

Y. Vasuki, E.-J. Holden, P. Kovesi, and S. Micklethwaite. An interactive image segmentation method for lithological boundary detection: A rapid mapping tool for geologists. *Computers & Geosciences*, 100:27–40, Mar. 2017. ISSN 0098-3004. doi: 10.1016/j.cageo.2016.12.001.

J. N. Vevatne, E. Rimstad, S. M. Hope, R. Korsnes, and A. Hansen. Fracture networks in sea ice. *Frontiers in Physics*, 2:21, 2014. ISSN 2296-424X. doi: 10.3389/fphy.2014.00021.

M. P. Viana, E. Strano, P. Bordin, and M. Barthelemy. The simplicity of planar networks. *Scientific Reports*, 3(1):3495, Dec. 2013. ISSN 2045-2322. doi: 10.1038/srep03495.

J. Vidal, A. Genter, and F. Chopin. Permeable fracture zones in the hard rocks of the geothermal reservoir at rittershoffen, france. *Journal of Geophysical Research: Solid Earth*, 122(7):4864–4887, 2017. doi: 10.1002/2017JB014331.

S. Virgo, S. Abe, and J. L. Urai. The influence of loading conditions on fracture initiation, propagation, and interaction in rocks with veins: Results from a comparative discrete element method study. *Journal of Geophysical Research: Solid Earth*, 121(3):1730–1738, 2016. doi: 10.1002/2016JB012792.

J. S. Y. Wang and J. A. Hudson. *Fracture Flow and Underground Research Laboratories for Nuclear Waste Disposal and Physics Experiments*, chapter 2, pages 19–41. American Geophysical Union (AGU), 2015. ISBN 9781118877517. doi: 10.1002/9781118877517.ch2.

Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004. ISSN 1057-7149. doi: 10.1109/TIP.2003.819861.

D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393 (6684):440–442, 1998. ISSN 1476-4687. doi: 10.1038/30918.

C. Weismüller, M. Passchier, J. Urai, and K. Reicherter. The fracture network in the lilstock pavement, bristol channel, uk: digital elevation models and orthorectified mosaics created from unmanned aerial vehicle imagery. *RWTH Publications*, 2020. doi: 10.18154/RWTH-2020-06903.

C. Weismüller, R. Prabhakaran, M. Passchier, J. L. Urai, G. Bertotti, and K. Reicherter. Mapping the fracture network in the lilstock pavement, bristol channel, uk: manual versus automatic. *Solid Earth*, 11(5):1773–1802, 2020. doi: 10.5194/se-11-1773-2020.

M. J. Welch, M. Luthje, and A. C. Glad. Influence of fracture nucleation and propagation rates on fracture geometry: insights from geomechanical modelling. *Petroleum Geoscience*, 25(4):470–489, 2019. ISSN 1354-0793. doi: 10.1144/petgeo2018-161.

A. E. Whitaker and T. Engelder. Characterizing stress fields in the upper crust using joint orientation distributions. *Journal of Structural Geology*, 27(10):1778 – 1787, 2005. ISSN 0191-8141. doi: 10.1016/j.jsg.2005.05.016.

S. Wierzchoń and M. Kłopotek. *Modern Algorithms of Cluster Analysis*, volume 34 of *Studies in Big Data*. Springer International Publishing, 1 edition, 2018. ISBN 978-3-319-69307-1. doi: 10.1007/978-3-319-69308-8.

C. E. Wilson, A. Aydin, M. Karimi-Fard, L. J. Durlofsky, S. Amir, E. E. Brodsky, O. Kreylos, and L. H. Kellogg. From outcrop to flow simulation: Constructing discrete fracture models from a lidar survey. *AAPG Bulletin*, 95(11):1883–1905, Nov. 2011. ISSN 0149-1423. doi: 10.1306/03241108148.

P. A. Witherspoon. Flow of groundwater in fractured rocks. *Bulletin of the International Association of Engineering Geology - Bulletin de l'Association Internationale de Géologie de l'Ingénieur*, 34(1):103–115, Oct. 1986. ISSN 1435-9537. doi: 10.1007/BF02590241.

H. Wu and D. D. Pollard. Imaging 3-d fracture networks around boreholes. *AAPG Bulletin*, 86(4):593–604, Oct. 2002. ISSN 0149-1423. doi: 10.1306/61EEDB52-173E-11D7-8645000102C1865D.

P. Wüstefeld. Capturing a world-class outcrop of a quality calcite vein network on a polished limestone outcrop in the oman mountains: Creation of a high resolution panorama and microstructural vein description, 2010. URL https://core.ac.uk/download/pdf/36587925.pdf.

F. A. Wyller. Spatio-temporal development of a joint network and its properties: a case study from lilstock, uk. Msc thesis, 2019. URL https://hdl.handle.net/1956/20414.

S. Yi, D. Labate, G. R. Easley, and H. Krim. A shearlet approach to edge analysis and detection. *IEEE Transactions on Image Processing*, 18(5):929–941, 2009. ISSN 1057-7149. doi: 10.1109/TIP.2009.2013082.

K. Yoshioka and B. Bourdin. A variational hydraulic fracturing model coupled to a reservoir simulator. *International Journal of Rock Mechanics and Mining Sciences*, 88:137–150, 2016. ISSN 1365-1609. doi: 10.1016/j.ijrmms.2016.07.020.

T. Zhang, P. Tilke, E. Dupont, L. Zhu, L. Liang, and W. Bailey. Generating geologically realistic 3d reservoir facies models using deep learning of sedimentary architecture with generative adversarial networks. In *International Petroleum Technology Conference, 26-28 March 2019, Beijing, China*, number IPTC-19454-MS in International Petroleum Technology Conference, Beijing, China, Mar. 2019. International Petroleum Technology Conference. doi: 10.2523/IPTC-19454-MS.

ACKNOWLEDGEMENTS

This PhD was partially sponsored by Eni S.p.A. for the duration of the SEFRAC consortium. I would like to thank my promotors Giovanni Bertotti at Delft and David Smeulders at Eindhoven for putting together such a unique arrangement as a joint PhD candidate within the two universities, while also being sponsored by a major upstream company. This dissertation would have been impossible if it wasn't for Giovanni Bertotti, who carved out such a unique opportunity for me. Dear Giovanni, I sincerely appreciate your patience, guidance, supervision, and for encouraging me to tackle problems at the interface of disciplines. With your active backing and the freedom that I was given to explore, I forayed beyond my background, into avenues such as image processing, graph theory, lidar, karst, machine learning, and more. Some of the explored trajectories didn't fructify into proper academic outputs, but nevertheless I am extremely thankful for your encouragement. Dear David, thanks to your clinical insight in our review meetings, planning and support, I was always able to remain on-track.

I would like to thank Marco Meda for organizing the SEFRAC meetings at ENI headquarters and allowing me to take a peek at real-world reservoir modelling, showing us the sights of beautiful Milano, and for facilitating the fieldwork at the scenic Parmelan plateau in France. I also thank Fabrizio Storti, Fabrizio Balsamo, Silvia Mittempergher, and Luigi Berio for interesting discussions on fractures, folds, and other structures during the Parmelan expedition. I am grateful to Janos Urai and Christopher Weismüller for the collaboration with RWTH Aachen that was started in the latter half of my PhD. The excellent photogrammetric dataset from the 'Mecca' outcrop of fractures, the Bristol Channel, is now the basis for four journal publications with the fifth in the pipeline. Janos, it is indeed a privilege for me to work under you.

Over the past six years, I received scientific inputs from many directions that enabled the shaping of my eventual PhD trajectory. While hunting for a master thesis project, both Hamid Nick and Nico Hardebol gave me crisp, yet exhaustive enough, coverage of the state-of-the-art in fracture network modelling. I am indebted to my master thesis supervisors Hans de Pater and Josef Shaoul at Fenix Delft, who were always willing to enlighten me on myriad aspects of fracturing, from physical experiments to numerical modelling to oilfield stories. I credit my frank discussion with Hans and Pacelli Zitha that led me to consider doing a PhD. Kevin Bisdom was instrumental for his guidance during the early months, for agreeing to participate and giving me feedback during my second year review meeting, and for inviting me and showing me around the Shell office in Amsterdam. I am thankful to Dr. Julia Gale who during her short course and feedback on my poster at the AAPG conference, gave valuable insight. Thanks are also due to Dr. Benoit Noetinger for the suggestion on using graph theory. I would like to thank Denis Voskov for his advice on DFM modelling and support with ADGPRS.

I would like to thank my independent doctoral examination committee members, Dr. David Bruhn, Dr. Bill Rossen, Dr. Auke Barnhoorn, Dr. Casey Nixon, and Dr. Ulrich Kelka

for accepting to review my work. I thank Dr. Bertrand Gauthier, Dr. Juliette Lamarche, Dr. Steven Laubach, Dr. David Sanderson, and Dr. Billy Andrews for critically reviewing my journal papers and improving the quality of my doctoral dissertation.

I would like to thank Quinten for being a valued 'fracture colleague', co-author, and for driving me around in Utah after our conference at Salt Lake City. It was a pleasure to work with you during the years we jointly supervised the Field Development Project master course. Dear Pierre-Olivier, you were not just a timely supervisor and co-author, but also a discussion sink for all things related to fractures. Thanks for the guidance during Parmelan fieldwork and other continuous feedbacks during the entire period of my PhD. Thanks Stephan for the cooperation and collaboration on the cave data acquisition, and various discussions on reservoir simulation, meshing, and fractures.

The Molengraaff fund is thanked for funding the cave expedition to Bahia, Brazil. The fieldwork would not have happened without the groundwork laid done by Hilario Bezerra under the POROCARSTE project. I would like to thank Hilario for his generous hospitality during our time in Brazil and his tireless enthusiasm during the fieldwork. In Brazil, I benefited greatly from the knowledge and speleological expertise of Augusto Auler and Phillippe Audra. I would like to thank Cayo and Rebeca for their company at Natal, in the caves, showing us around picturesque Lençois, bringing me to a live capoeira performance, and for helping me get the 'best' coffee in Brazil.

Thanks to my fellow colleagues from geology, geophysics, and petroleum engineering: Remi, Koen, Max, Helena, Lisanne, Siddharth, Runhai, Swej, Myrna, Mark, Navid, Boris, Thais, Quinten, Baptiste, Joeri, Youwei, Reuben, Jianhuan, Santhosh, Lele, Karlien, Atsushi, Alexandros, Ranjani, Aulia, Maria, Anne, Hiranya, Stephan, Kai, Tim, Akeel, Martha, Florencia, Diego, Nicolas, Sixue, Musab, Emilio, Andrea, for a great 4.5 years at TU Delft! Special thanks to Remi, Tim, and Emilio for organizing the Think and Drink sessions !!

Since my PhD was a joint project between TU Delft and TU Eindhoven, I would like to take a moment to thank the support staff at both universities: Linda, Angie, Mariska, Hetty, Kelsi (TU Eindhoven) and Marlijn, Lydia, Margot, Ralf, Marja, Marijke, Anke (TU Delft), for all assistances that made my time in the office comfortable. Thanks also to Joost and Guus for the software support and licensing.

Thanks to Petrobowl and facilitation by SPE Student Chapter Delft, I found an excuse to escape temporarily from my work and wander the streets of Krakow twice. Thank you Kai, Swej, Luiza, Rhadityo, Sean, and Ahmad. It was really fun, although Covid spoiled another potential trip to Zagreb.

A shoutout to my Delft housemates, Karthik, Hamid, Nitheesh, and Naveen! If it wasn't for Karun Nooney who was my junior at ONGC and who suggested TU Delft and the Netherlands as a possible study destination, I would have probably continued my rig career, boring kilometres and kilometres of holes into the earth. Thanks Karun for the timely perturbation that eventually changed my life for the better!!

I would like to thank my parents, brother, and sister for their complete support, love and faith in me. This work is dedicated to you all! Last but not the least, I thank my wife Anjana. It is not easy to get married to and put-up with a PhD student husband in his mid-30s. I look forward to the new future that awaits us in Erlangen.

CURRICULUM VITÆ

Rahul PRABHAKARAN

1984/10/10	Born in Manama, Kingdom of Bahrain

Education

2000 - 2002	All India Senior School Certificate Examination Sabarigiri Central School, Kerala, India
2002 - 2006	B.Tech in Mechanical Engineering University of Kerala, India
2014 - 2016	M.Sc in Applied Earth Sciences Specialization: Petroleum Engineering and Geosciences TU Delft, the Netherlands
2017 - 2021	Ph.D in Applied Geology TU Delft / TU Eindhoven , the Netherlands

Work Experience

2006 - 2008	Engineer - Research and Development Elgi Equipments, Coimbatore, India
2010 - 2014	Assistant Executive Engineer - Drilling Oil and Natural Gas Corporation, India
2015 - 2016	Research Internship Fenix Consulting Delft, the Netherlands

List of Publications

Journal Publications

8. **R. Prabhakaran**, G. Bertotti, J.L. Urai, D. Smeulders, *Investigating spatial heterogeneity within fracture networks using hierarchical clustering and graph distance metrics*, Solid Earth, In Press, (2021)

7. **R. Prabhakaran**, J.L. Urai, G. Bertotti, C. Weismüller, D.M.J Smeulders, *Large-scale natural fracture network patterns: Insights from automated mapping in the Lilstock (Bristol Channel) limestone outcrops*, Journal of Structural Geology, 150, 104405, (2021)

6. C. Weismüller, **R. Prabhakaran**, M. Passchier, J.L. Urai, G. Bertotti, K. Reicherter, *Mapping the fracture network in the Lilstock pavement, Bristol Channel, UK: manual versus automatic*, Solid Earth, 11 (5) , 1773-1802, (2020)

5. G. Bertotti, P. Audra, F.H. Bezerra, S. de Hoop, C. Pontes, **R. Prabhakaran**, Rebeca Lima, *The Morro Vermelho hypogenic karst system (Brazil): Stratigraphy, fractures and flow in a carbonate strike-slip fault zone with implications for carbonate reservoirs*, AAPG Bulletin, 104 (10), 2029-2050, (2020)

4. **R. Prabhakaran**, P-O. Bruna, G. Bertotti, D. Smeulders, *An automated fracture trace detection technique using the complex shearlet transform*, Solid Earth, 10 (6), 2137-2166, (2019).

3. P-O. Bruna, J. Straubhaar, **R. Prabhakaran**, G. Bertotti, K. Bisdom, G. Mariethoz, M. Meda, *A new methodology to train fracture network simulation using multiple-point statistics*, Solid Earth 10 (2), 537-559, (2019).

2. Q. Boersma, **R. Prabhakaran**, F.H. Bezerra, G. Bertotti, *Linking natural fractures to karst cave development: A case study combining drone imagery, a natural cave network and numerical modelling*, Petroleum Geoscience, 25 (4), 454-469, (2019)

1. **R. Prabhakaran**, H. de Pater, J. Shaoul, *Pore pressure effects on fracture net pressure and hydraulic fracture containment: insights from an empirical and simulation approach*, Journal of Petroleum Science and Engineering, 157, 724-736, (2017)

Conference Proceedings

6. Q.Boersma, **R. Prabhakaran**, F.H. Bezerra, G. Bertotti, *Linking Natural Fractures to Karst Cave Development: A Case Study on Naturally Fractured Carbonates*, 82nd EAGE Annual Conference and Exhibition, 8 - 11 December, Amsterdam, the Netherlands, (2020).

5. G. Bertotti, P. Audra, A. Auler, F.H. Bezerra, S. de Hoop, **R. Prabhakaran**, C. Pontes, R. Lima, *Characteristics and genesis of hypogenic karsts in carbonate successions: inferences for hydrocarbon and geothermal systems*, Fourth EAGE Workshop on Naturally Fractured Reservoirs, 11 - 13 February, Ras Al Khaimah, UAE, (2020).

4. P-O. Bruna, **R. Prabhakaran**, G. Bertotti, J. Straubhaar, R. Plateaux, L. Maerten, G. Marietho z, M. Meda, *The MPS-based fracture network simulation method: Application to subsurface domain*, 81st EAGE Conference and Exhibition, 3 - 6 June, London, U.K., (2019).

3. **R. Prabhakaran**, P-O. Bruna, G. Bertotti, S. Mittempergher, A. Succo, A. Bistacchi, F. Storti, *Characterizing dynamic stress sensitive fracture apertures in a DFN representation: An example from the Island of Pag (Croatia)*, AAPG Annual Convention and Exhibition, 20 - 23 May, Salt Lake City, USA, (2018).

2. P-O. Bruna, **R. Prabhakaran**, G. Bertotti, S. Mittempergher, A. Succo, A. Bistacchi, *Multiscale 3D prediction of fracture network geometry and fluid flow efficiency in folded carbonate reservoir analogues: Case study of the Island of Pag (Croatia)*, Third EAGE Workshop on Naturally Fractured Reservoirs – Calibration Challenges, 5 - 7 February, Muscat, Oman, (2018).

1. C.J. de Pater, **R. Prabhakaran**, J.R. Shaoul, *Observations and modeling of the relation between effective reservoir stress and hydraulic fracture propagation pressure*, SPE Hydraulic Fracturing Technology Conference and Exhibition, 24 - 26 January, Texas, USA, (2017).

Datasets

4. **R. Prabhakaran**, J.L. Urai, G.Bertotti, Christopher Weismüller, D.M.J Smeulders, *Fracture Patterns from the Lilstock Pavement, Bristol Channel, UK*, 4TU Centre for Research Data, (2021)

3. Q. Boersma, **R. Prabhakaran**, F.H. Bezerra, G. Bertotti, *Brejões fractured pavements and cliffs: drone data, interpretations and modelling results*, 4TU Centre for Research Data, (2020)

2. **R. Prabhakaran**, Q. Boersma, F.H. Bezerra, G. Bertotti, *Fracture network patterns from the Brejões Outcrop, Irecê Basin, Brazil*, 4TU Centre for Research Data, (2019)

1. **R. Prabhakaran**, P-O. Bruna, G. Bertotti, D. Smeulders, M. Meda, *Fracture network patterns from the Parmelan Anticline, France*, 4TU Centre for Research Data, (2019)