A workflow for the characterization of fracture networks from 3D reprocessed seismic data using non local means filter

Lonneke van Bijsterveldt

July 28, 2017

Abstract

The Dinantian carbonate platforms in Northern Netherlands are marked as potential reservoirs for Ultra Deep Geothermal Energy (UDG). Permeability and therefore flow in these reservoirs is thought to be primarily defined by natural fracture networks. For the purpose of UDG, an analysis is made of the fracture network characteristics of the Dinantian Friesland platform. Due to recent developments in reprocessing of 3D seismic data and fault extraction algorithms, a first attempt is made to determine fracture network characteristics from 3D reprocessed seismic data. First, the 3D seismic data is filtered by the non local means (NLM) algorithm, using different filter parameter combinations. By doing so we aim to determine the influence of the filter settings on the characteristics of the fracture networks, extracted from the data. We are especially interested to see if by filtering the data, small fractures are removed in addition to noise. Next, the Petrel ant track workflow is used as sampling method for automatic detection and extraction of fractures, which are exported as fault stick files. A program is built to calculate fracture attributes, such as barycenter coordinates, strike, azimuth, dip, length, offset and aspect ratios. These fracture attributes are then used for fracture network characterization, computing fracture length distributions, clustering and orientation distributions. To quantify the scale invariance of the fracture network, we aim to find power law behavior in the fracture network attributes of length and spatial distribution, such that it fits Davy et al.'s (2010) double power law: $n(l, L) = \alpha L^D l^{-a}$. Length exponent a is determined via computing the density-length distribution and the fractal dimension D using the pair-correlation function. Power law behavior is found for the fracture length distribution below the truncation cut-off, with varying length exponents a depending on the NLM parameters used: $3.59 \leq a \leq 4.57$. However, the fracture length distribution is more sensitive to the Petrel attribute settings than to the NLM filter parameters. This implies a large uncertainty on the results. Furthermore, the sampling method (Petrel ant track workflow) is found to be unsuitable for detection of fractal behavior (D), due to the limits of seismic resolution.

Contents

1 Introduction **5**
1.1 Fracture network characterization 6

2 Related work **7**
2.1 Fracture length distribution of 2D sampled fracture systems 7
2.1.1 Correction methods for sampling effects 8
2.2 Spatial distribution of fracture networks sampled in 1D and 2D 9
2.3 Translating between fracture networks sampled in 1D, 2D or 3D 9
2.4 Non local means filter 10

3 Methods **11**
3.1 Non local means filter 12
3.2 Petrel ant track workflow 13
3.2.1 Variance 13
3.2.2 Ant tracking 14
3.2.3 Automatic fault extraction 15
3.3 Database design 17
3.4 Fracture network characterization 19
3.4.1 Fracture length distribution 20
3.4.2 Fractal Dimension 21
3.4.3 Orientation distribution 21

4 Results **21**

5 Discussion **27**
5.1 Petrel ant track workflow 27
5.2 Seismic resolution 30

6 Conclusions and recommendations **30**

7 Notations **31**

8 Acknowledgements **31**

9 Appendix A: Influence of NLM on the seismic volume **32**

10 Appendix B: Influence of NLM on the variance volume **33**

11 Appendix C: Influence of NLM on the ant track volume **34**

12 Appendix D: Influence of NLM on the fault extraction **35**

List of Tables

1 Ant track parameters . . . 15
2 Automatic fault extraction parameters . . . 16
3 Fault attribute calculations . . . 18

List of Figures

1 Paleogeographic map of the Netherlands during Early Carboniferous . . . 5
2 Fractal dimension and length exponent . . . 6
3 Density length distribution from Bour et al. (2002) . . . 8
4 workflow . . . 10
5 sample window . . . 11
6 Influence of NLM filter parameters . . . 12
7 Variance parameters . . . 14
8 Petrel fault patch to fault stick file . . . 17
9 Fault attributes . . . 19
10 Density length distributions as a function of NLM parameters . . . 22
11 Effect of NLM filter on fault extraction . . . 23
12 Spatial distributions as a function of NLM parameters . . . 24
13 Orientation distributions as a function of NLM parameters . . . 25
14 Dip distributions as a function of NLM parameters . . . 26
15 fracture length distribution of NLM filtered data as a function of fault extraction parameters . . . 28
16 fracture length distribution as a function of fault extraction parameters . . . 29
17 Fault extraction for different ant track settings . . . 30
18 Influence of NLM on seismic volume . . . 32
19 Influence of NLM on the variance volume . . . 33
20 Influence of NLM on the ant track volume . . . 34
21 Influence of NLM on the automatic fault extraction . . . 35

1 Introduction

As proven by several successful projects, the subsurface in the Netherlands (1000-4000 m) is suitable for deep geothermal energy exploration. Water extracted from these depths has a temperature of about 40-130°C and is mainly used for heating in green houses. However, 31% of the industrial heat demand in the Netherlands concerns temperatures of 100-200°C [EZ, 2016]. Heat to cover this demand could be obtained by Ultra Deep Geothermal energy (UDG). Sources for ultra-deep geothermal energy are characterized by temperatures exceeding 120°C and depths between 4-8 km [Boxem et al., 2016]. To date no ultra-deep doublets have been drilled in the Netherlands, due to the lack of data and knowledge of the subsurface below 4 km depth.

TNO has done research to potential UDG reservoirs in the Netherlands. By combining a 3D temperature model of the Netherlands [Bonté et al., 2012] with the Digital Geological Model (TNO, 2013), it shows which geological units are present at depths of certain isotherms. All UDG plays were ranked by parameters such as; geothermal potential, rock properties, size of the play, permeability and match with heat demand in the area. The best ranked UGD reservoirs in the Netherlands are the carbonate platforms of the Dinantian [Boxem et al., 2016]. Figure 1 shows these early carboniferous (Dinantian) carbonate platforms in blue. For this study we will focus on the platforms in northern Netherlands, since there is much more data available of this region, as a result of a rich history in gas exploration.

Due to deep burial, reservoir rocks for UDG are characterized by low porosities (primary permeability). In addition, platform carbonate rocks are already marked by low porosities of themselves [Collins et al., 2014]. However, permeability of these reservoirs can increase by the presence of karst or natural fracture systems [Goldscheider et al., 2010]. In the Dinantian platforms in the Netherlands permeability seems to be defined by natural fracture networks [Lipsey et al., 2016]. If the permeability in these fracture networks is still not sufficient to obtain economical flow rates, hydraulic stimulation can be applied. This concept is referred to as an Enhanced Geothermal System (EGS). Due to shearing of the rough fracture surfaces dilation occurs, increasing the permeability of the fracture.

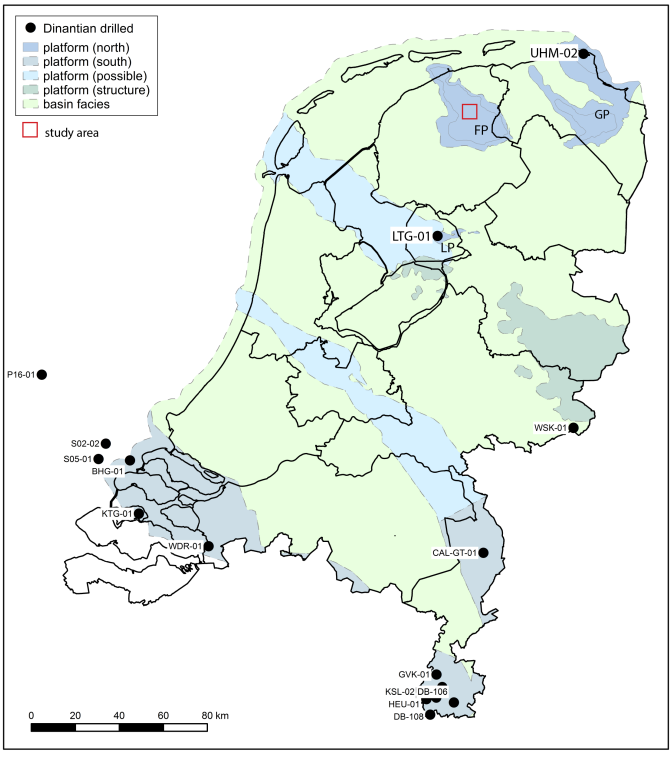


Figure 1: Paleogeograhpic map of the Netherlands during Dinantian by TNO [Boxem et al., 2016]. Red square indicates the study area. Black dots indicate wells drilled to Dinantian rocks. FP=Friesland Platform, GP=Groningen Platfrom and LP=Luttelgeest Platform.

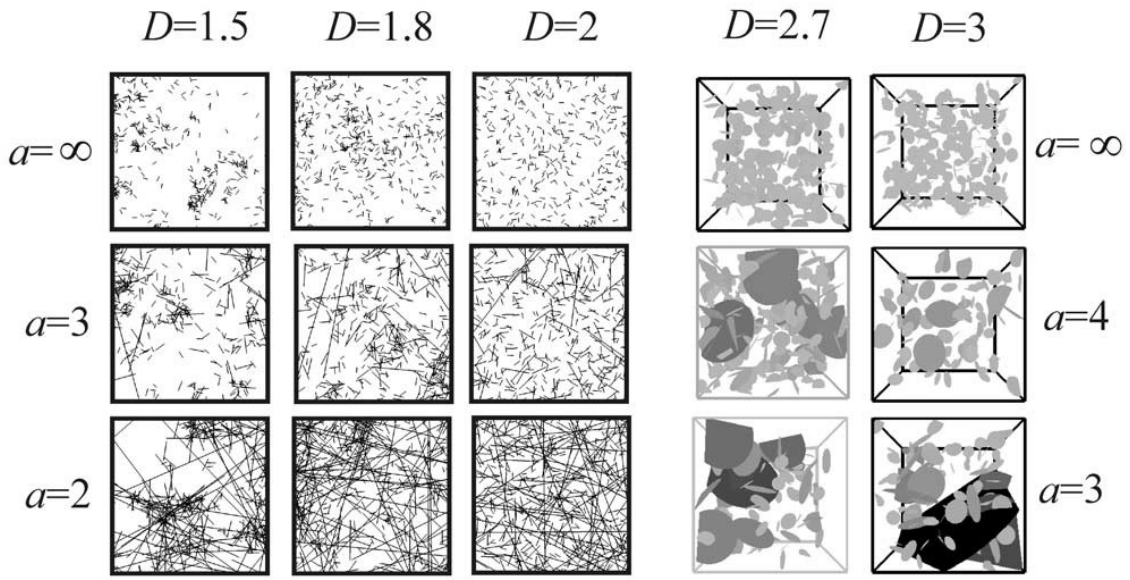


Figure 2: A visualization of parameters of Davy et al.'s (1990) double power law for Euclidean dimensions 2 and 3. A fractal dimension equal to the Euclidean dimension shows a homogeneous distribution of fractures, while a decreasing D implies an increase in clustering. A decreasing length exponent a implies a relative increase in larger fractures. Figure from [Darcel et al., 2003a]

The possible side effect of hydraulic stimulation is micro-seismicity or occasional larger earthquakes [Häring et al., 2008]. To reduce seismic risk in the case of EGS and for optimal doublet location and orientation for both EGS and UDG, it is necessary to make an in-depth analysis of the natural fracture networks in these reservoirs.

1.1 Fracture network characterization

A widely used tool to characterize fracture network attributes, such as length and spatial distribution of fractures, is Power law scaling and fractal geometry [Bonnet et al., 2001]. The advantage of power law distribution and fractal geometry is their scale invariance, which allows extrapolation of fracture network attributes to different scales [Bour et al., 2002]. For the purpose of this study, of which the fracture network data is obtained from 3D seismic data, this relationship can help us to get an idea on the abundancy and spatial distribution of small scale faults which are beyond the resolution of the seismic data.

[Davy et al., 1990] proposed a model which describes both the length and spatial distribution of fracture networks and follows a double power law:

$$n(l, L) = \alpha L^D l^{-a} \tag{1}$$

where $n(l, L)$ is $N(l, L)/dl$, the number of fractures whose length is in the range $[l, l+dl]$ divided by the bin size (dl) and whose barycenter is within the volume of size L^3, l is the fracture length and L is the typical system size, α is the fracture density term, D the fractal dimension of the fracture barycenters and a is the exponent of the density distribution of fracture lengths (see Figure 2). A fracture network is commonly called fractal when any kind of fracture characteristic (length, aperture, displacement etc.) shows a power law distribution. However the term fractal should only be used for spatial distributions [Mandelbrot, 1983]. A fractal system implies a spatial correlation (clustering) between objects which is described by the fractal dimension [Bonnet et al., 2001]. In this paper we shall stick to the original definition of the term fractal and only use it to describe spatial distributions.

Another advantage of this model is the stereological relationship. These relationships make it possible to extend spatial distributions from 1D bore hole data to 3D seismic data and vice versa [Darcel et al., 2003a, Darcel et al., 2003b] (section 2.3). There are many examples in liter-

ature of applying this model to 2D fracture trace maps [Lei, 2016, Davy, 1993, Bour et al., 2002, Davy et al., 2010, Bonnet et al., 2001] and to 1D scanlines [Valley et al., 2016, Boadu & Long, 1994]. In the case of the 1D scanline only the spatial distribution and thus the fractal dimension D can be obtained, since the length of a fractures intersecting a bore hole can not be determined. To this end, the model has not been applied to fracture networks extracted from 3D seismic data [Bonnet et al., 2001]. Note that in this paper the term fault and fracture are used interchangeably. The correct terminology defines faults to be a type of fracture which fails due to shear stress, whereas a fracture could also be caused by tensile stress and not show any displacement. However, since this study focuses on fractures extracted from seismic data, all fractures detected are per definition offset and therefore faults. In this study we, therefore, make no distinction between fractures or faults.

The aim of this study is to apply Davy et al.'s double power law model to identify the fracture network characteristics of the Dinantian carbonate platforms, based on reprocessed vintage 3D seismic data. A work flow is designed to compute fracture network characteristics from faults extracted from 3D seismic data. The faults are automatically extracted using the ant track work flow of Petrel, which allows for fast computation and minimizes human interaction. The reprocessing of the seismic data is a necessary step because the quality of the seismic data at depth of the Dinantian carbonates is relatively poor. Filtering of data is required to avoid interpretation of noise as faults. The post stack reprocessing is done by the non local means (NLM) filter [Carpentier & Steeghs, 2016], which will be discussed in more detail in section 2.4. The amount of filtering done by the NLM filter can be controlled by filter parameters. We expect, that depending on the NLM filter parameters, the reprocessing will remove small fault indications in addition to noise, which might negatively influence our fracture network characteristics. Therefore, this study also focuses on determining the influence of the NLM filter parameters on the fracture networks characteristics extracted from the seismic data.

2 Related work

2.1 Fracture length distribution of 2D sampled fracture systems

There are several ways to describe the length distribution of a fracture network: frequency (2), density (3) and the cumulative frequency (4) distribution [Bonnet et al., 2001].

$$N(l) = \alpha l^{-a} dl \tag{2}$$

$$n(l) = \alpha l^{-a} \tag{3}$$

$$C(l) = \int_{l}^{l_{max}} n(l) dl \tag{4}$$

where $N(l)$ the number of fractures whose length is in the range $[l, l + dl]$, l is the fracture length, l_{max} is the largest fracture found in the system, dl is the bin size, a is the powerlaw exponent and α is the density constant. If the length distribution of a fracture network fits a powerlaw, the length exponent can be derived from all three distributions. However, the length exponent value will differ depending on the distribution and the bin type used. When assuming the length exponent of the density distribution powerlaw fit is a, the value of the length exponent of a cumulative distribution will be $a - 1$, as will be the frequency distribution when binned logarithmically [Bonnet et al., 2001]. It is important to know the relation of the length exponent between these distributions and bin types in order to compare length distributions from different studies. In this work we will focus on the density distribution (equation 3), since this is the distribution used in the Davy's double powerlaw model. However, the frequency distribution is needed to compute the density distribution. The density distribution $n(l, L)$ is the frequency distribution $N(l, L)$ which is the number of faults with a length in the range $[l, l + dl]$, divided by the bin size dl. The cumulative distribution can be computed separately, without the need to bin the data and can thus be used as an extra check for the estimated length exponent a. The frequencies per fracture lengths are summed without binning, resulting in a smoother curve than the frequency and density distribution.

As mentioned before, the convenience of a fracture length distribution that fits a powerlaw, is the scale invariance of the system. There is no characteristic length scale to which

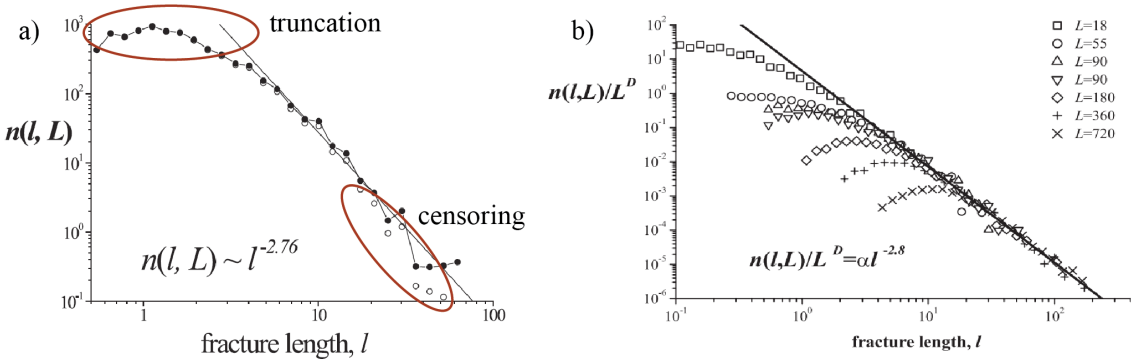


Figure 3: Figures from Bour et al.'s (2002) paper on a statistical scaling model for fracture network geometry. (a) density fracture length distribution of a single scale fracture trace map showing truncation and censoring effects. A powerlaw is fit to the non-truncated part and after correcting for censoring effects. (b) Density length distributions for fracture trace maps of different scales combined. Single distributions shows a log-normal distribution similar to (a), but when considering truncation effects the parental distribution is a powerlaw.

the system is restricted. When a characteristic length scale is present, for example lithological layering, one of the following distributions might show a good fit: log-normal, exponential or gamma law distributions [Zeeb et al., 2013, Bonnet et al., 2001]. However, when a system at multi scale is characterized by a powerlaw length distribution, the length distribution at single characteristic scale lengths often still represents a log-normal distributions, which is a result of the incomplete sampling of small fractures due to resolution effects, called truncation (figure 3) [Einstein & Baecher, 1983, Odling, 1997, Bonnet et al., 2001]. The continuity of a powerlaw length distribution is limited by two types of cut-offs, sampling effects (including truncation and censoring) and the physical upper and lower bounds of the system. Truncation is a sampling effect caused by the limit of resolution. Small fractures are incompletely sampled, resulting in a shallowing of the slope of the length distribution at the lower end of the scale range. Truncation effects are best observed in density length distributions. Censoring is the phenomenon of large fractures which are incompletely sampled because they partially fall outside of the sample (observation) window. This phenomenon causes the slope of the length distribution to steepen at the upper end of the scale range [Bonnet et al., 2001]. Both sampling effects are clearly depicted in figure 3. The second limitation to the extent of the powerlaw distribution are the physical upper and lower bounds to the system. The question of a physical lower bound to the system is not often addressed in literature, because the resolution limitations are usually encountered before the physical lower bound is reached. Bonnet et al. (2001) suggests the lower limit to be defined by atomic bonds or grain sizes, where as Odling (1997) suggests the physical lower cut off to be around 1m for joints (a fracture without displacement) in sandstones. For the purpose of our study, which is based on 3D seismic data, we do not have to worry about the physical lower cut off, since the seismic resolution limit will always be reached beforehand. The upper physical limit for powerlaw length distributions mentioned in literature are finite thickness of sedimentary beds and the brittle crust [Davy, 1993, Bonnet et al., 2001]. Apart from these limitations the cumulative distribution suffers from an additional effect that causes steepening of the distribution at larger scale. When l reaches l_{max}, $C(l)$ goes to zero, curving the cumulative distribution slope for an l range were the density distribution might still represent a powerlaw fit. This effect decreases the range to estimate length exponent a on a cumulative distribution to an even smaller part of the graph as the part which is affected by truncation and censoring.

2.1.1 Correction methods for sampling effects

To compute a powerlaw exponent from a density length distribution, the above mentioned sampling effects (truncation and censoring) should be considered and corrected for. This is to avoid under or overestimation of the ratio of small fractures with respect to large fractures. The only method mentioned in literature to correct for truncation effects is to remove the part of the distribution affected by truncation (figure 3). No quantitative method of determining the truncation threshold

is mentioned in literature. According to Bonnet et al. (2001) the truncation threshold for 2D fracture trace maps lies between 0.5-25% of the map size with an average around 5%. The map size is defined by L, the squared root of the map area. According to Odling et al. (1997) the truncation threshold lies between 5-10% of the map size. Correction methods for censoring effects are discussed more thoroughly in literature on 2D samples fracture systems. Three methods are mentioned: the method of [Laslett, 1982], which is based on the probability that a fracture of length l will be observed completely with in the system of size L. For this method fractures intersecting the system boundaries and whose real length is not known, should be excluded from the length distribution before the correction is applied [Bour et al., 2002]. The Kaplan-Meier method [Odling, 1997] is based on the number of fractures intersecting the system boundaries. The last method suggested by [Pickering et al., 1995] is to increase the frequency of large fractures until the best fit with the powerlaw is obtained.

2.2 Spatial distribution of fracture networks sampled in 1D and 2D

Similar to the studies on length distributions of fracture networks, most information on spatial distribution of fractures comes from 2D outcrops converted to trace maps. Unlike the attribute length, spatial clustering of fractures can be measured in 1D borehole data. This can be done by using the Fullbore Formation MicroImager (FMI) tool in a bore hole, as sampling method. The FMI tool creates an electrical image of the bore hole wall from micro-resistivity measurements. Micro-resistivity changes on the image can be interpreted as fractures with a resolution as high as 5 mm [Schlumberger, 2002].

Methods to determine the fractal dimension of a fracture network have thus been developed for both 1D and 2D data. Examples of these methods are the box-counting method [Mandelbrot, 1983], the Cumulative distribution method [Valley et al., 2016] and the density-density or pair-correlation method [Hentschel & Procaccia, 1983]. According to [Vicsek, 1992] the pair-correlation method is the most efficient method to determine the fractal dimension of a natural system. In addition, Bour et al. (2002) and Valley et al. (2016) have concluded that for both 2D and 1D systems, the pair-correlation function characterizes the clustering degree of fracture centers the best. The pair-correlation method is computed accordingly:

$$C_2(r) = \frac{2N_p(r)}{N(N-1)} \tag{5}$$

where N_p are the number of pairs of points whose Euclidean distance is less than r and N is the total number of points. The method in 2D is applied on fracture barycenters and in 1D on the intersection of the fracture with the borehole. Similarly, r in 2D is the radius around a barycenter, in 1D it is a distance along the borehole. The pair-correlation function $C_2(r)$ is then expected to scale with r as follows: $C_2(r) \sim r^{D_c}$. If $C_2(r)$ fits a powerlaw, then the correlation dimension D_c is equal to the fractal dimension D and is the local slope of $C_2(r)$, which is computed as follows:

$$D_c = \frac{d \log C_2(r)}{d \log r}. \tag{6}$$

The local slope is a good measure to determine the uncertainty of the computed correlation dimension. Only if the graph of the local slope shows a significant plateau, indicating the scale range over which the slope of the the pair-correlation function is constant, a line can be fit to $C_2(r)$ by linear regression, giving a meaningful determination of the fractal dimension D. If no such plateau is visible, no fractal dimension can be determined [Bonnet et al., 2001].

2.3 Translating between fracture networks sampled in 1D, 2D or 3D

Bonnet et al.(2001) suggests that characterization of fracture networks from 3D data, as seismic data sets, is not yet possible because 3D data sets seldom contain sufficient data to achieve robust statistical analysis. This is due to the lack of resolution, the presence of noise in the data and the computationally demanding process of interpreting and extracting fault surfaces from seismic data. However, with recent developments in the reprocessing of seismic data and the algorithms available to automatically extract fault surfaces from seismic data, these obstacles can be overcome. We have found no literature on attempts to analyze fracture network characteristics of 3D systems. However, there are studies that have found relations between 1D, 2D and 3D systems, which we

can use to verify our results with the numerous studies on 2D systems. Another advantage of these relations is that it allows to use sampling techniques of different dimension to determine fracture network characteristics. This can be an advantage, because when a sampling method in one dimension does not provide sufficient data due to resolution, it can be replenished by a sampling technique at a different dimension. [Darcel et al., 2003a] has found stereological rules which describe the relation between the characteristics of fracture systems which are sampled in different dimensions. These rules are

$$D_{2D} = D_{3D} - 1 \qquad a_{3D} > 2 \tag{7}$$

$$D_{2D} = D_{3D} - a_{3D} + 1 \qquad D_{3D} - 1 \leq a_{3D} \leq 2 \tag{8}$$

$$a_{3D} = a_{2D} + 1 \qquad a_{3D} > 2 \tag{9}$$

where the subscript 2D and 3D refer to two- and three-dimensional (Euclidean) systems respectively [Darcel et al., 2003b]. These equations show that simple stereological rules apply as long as $a_{3D} > 2$. [Bonnet et al., 2001] has analysed many 2D fractures systems and concludes that the length exponent a_{2D} ranges between 1.7 and 2.75, suggesting that a_{3D} should range between 2.7 and 3.75. The fractal dimension D_{2D} ranges between 1.5 and 2, which results according to the stereological rules in a D_{3D} range from 2.5 to 3. These values can provide a check on the validity of our own results from the analysis of a 3D fracture network.

2.4 Non local means filter

The attenuation or removal of random noise from seismic data is an important seismic processing step. Several random noise attenuation methods have been introduced in the past. Examples are fx-deconvolution [Canales, 1984] and the Karhunen-Loéve transform [Jones & Levy, 1987]. The disadvantage of these methods is the lack of edge preservation. While removing random noise and enhancing continuity, energy gets smeared across sharp discontinuities [Carpentier & Steeghs, 2016]. Since fracture detection is the purpose of this study, it is essential to use a noise attenuation filter that preserves fault indications. We therefore use a new method which outperforms other methods on its edge preservation ability: the non local means (NLM) random noise attenuation filter. The method was originally developed by [Buades et al., 2005] for image processing and is adapted by [Bonar & Sacchi, 2012] for 2D seismic data denoising. The algorithms used in this study are built by TNO and are suitable for filtering of both 2D and 3D seismic data. The NLM algorithm is based on the assumption of a high level of redundancy in an (seismic) image. For a more detailed description on the algorithm used, see [Carpentier & Steeghs, 2016, Buades et al., 2011, Bonar & Sacchi, 2012]. The filter parameters to be changed in the algorithm are the search window (S), the neighbourhood window (N) and the filter parameter (h).

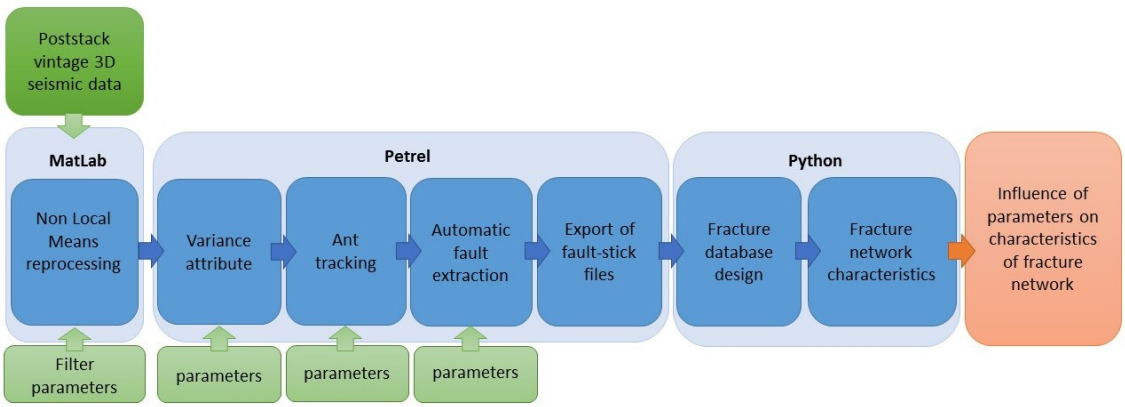


Figure 4: The workflow for the characterization of fracture networks used in this study. Green boxes are input values or data, blue boxes are processing steps and the orange box are the results of this study.

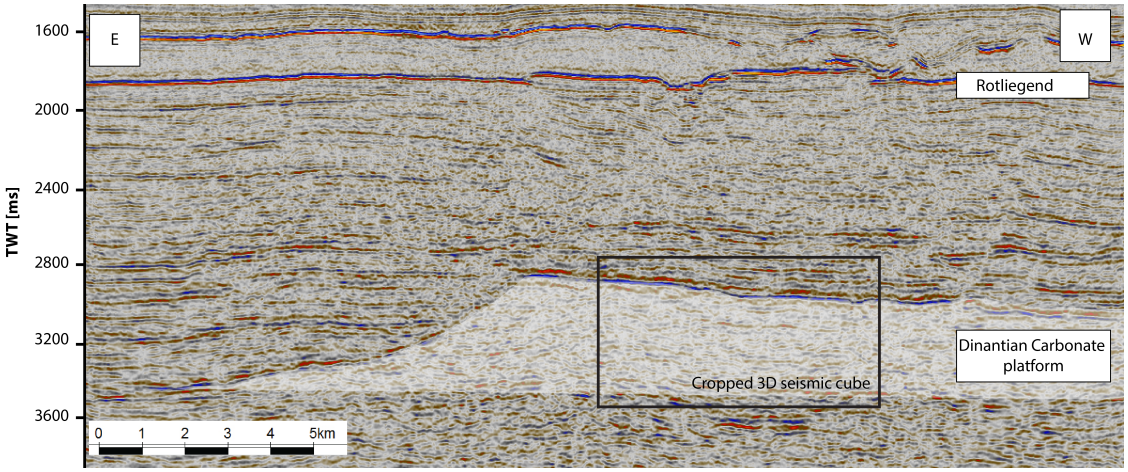


Figure 5: Crossline 1746 of L3NAM1987F Tietjerkstradeel West. Dinantian Friesland platform is marked by transparent white and the cropped minicube used in this study is indicated with a black box.

3 Methods

Before we determine the characteristics of fracture networks of the Dinantian carbonate platforms, the fractures are extracted from the 3D seismic data and translated into a data type suitable for further calculations. The total work flow to be followed in this study is visualized in figure 4).

To decrease computation time, a small (mini)cube is cropped from the original 3D seismic survey (L3NAM1987F, Tietjerkstradeel West) positioned in the Friesland Dinantian carbonate platform (figure 5). The size reduction of the data is necessary because the runtime for the 3D NLM filter increases drastically with number of samples and the size of the parameters. The size of our 3D cube after cropping is 231 inlines by 231 crosslines, where each trace contains 191 samples. A total of 231x231x191 voxels. Inline and crossline spacing is 25m, resulting in a surface span of about 33 km^2. The location and size of the cropped volume are chosen, so that the largest possible and most isotropic cube encapsulates the Friesland platform. A large study area is preferred to extract a larger amount of faults, which will increase the robustness of the statistics and an isotropic shape of the area increases the robustness especially of the spatial distribution. A polygon was drawn around a timeslice of the cube and exported as a shapefile, which was used to identify the sample boundaries and size of our sample window.

This new seismic cube was reprocessed using different NLM filter parameters. After reprocessing the data is reloaded in Petrel and several edge detection and enhancement attributes, known as the ant tracking work flow [Schlumberger, 2016], are applied to the data. The last step of the ant tracking work flow is the automatic fault extraction. The fault sticks constructing the fault patches are exported as text files. These files contain the x,y,z-coordinates of the points, constructing the fault sticks of the fault patches. Fault sticks are roughly vertical (curved) lines, representing the slope of the fault. x and y coordinates of these points are defined, using the Netherlands RD new coordinate system. This coordinate system uses meter as unit of measurement and the y-axis defines North. The z coordinate is still defined in two-way-travel time (ms) and needs to be converted to depth in order to calculate fault attributes as surface dip angle. For time depth conversion a combination of seismic data, clearly defining the top and bottom of the reservoir and a suitable well are used. There are only a few suitable wells in the Netherlands that are drilled deeper than 5 km and reach the Dinantian carbonates (figure 1). One of these wells, the Luttelgeest well LTG-01 is drilled through the Luttelgeest Dinantian carbonate platform and, therefore, was used for time-depth conversion. An average velocity was computed for the platforms overburden, by comparing the top of the platform from the 2D seismic line intersection the platform and bore hole (in TWT) and the top of the platform (in meters) from the end-slip of the bore hole. Using the same technique an interval velocity for the Dinantian carbonate platform was also computed. The average overburden velocity was calculated to be 3226 m/s and the interval velocity for the Dinantian carbonates 6039 m/s. These velocities were used for time-depth conversion of the z

coordinates of the fault sticks extracted from the Friesland platform as well as to define the depth of the top of the reservoir (4.4 km). In the following sub sections all workflow steps are discussed in more detail.

3.1 Non local means filter

Since applying the non local means filter to 3D seismic data is computationally demanding, the effects and suitable quantities of varying filter parameters are first tested on a 2D seismic line (L2NAM1981M_815015). This line intersects the Luttelgeest Dinantian carbonate platform. Figure 6a and b show the effect of the NLM filter with varying parameters h and N on the 2D data set, respectively. The effect is quantified as the normalized sum of absolute difference (L_1) between the original data and the filtered data. It is normalized by the total amount of data samples and displayed as the percentage of the amplitude range (difference between minimum and maximum amplitude in the data). These normalization steps are taken to enable the comparison between different data sets. In figure 6b and d, the NLM parameter N (neighbourhood window) is normalized by the number of samples in the dominant wavelet of the data. The dominant wavelet is determined by taking the Fourier transform of all traces and computing an average power spectrum of the whole dataset, from which the dominant frequency is extracted. The dominant frequency is then translated to a dominant period, from which the number of samples of the dominant wavelet is computed, using the sampling rate ($dt = 0.004s$). The dominant frequency of the 2D dataset is 35 Hz, resulting in a dominant wavelet of on average 7 samples.

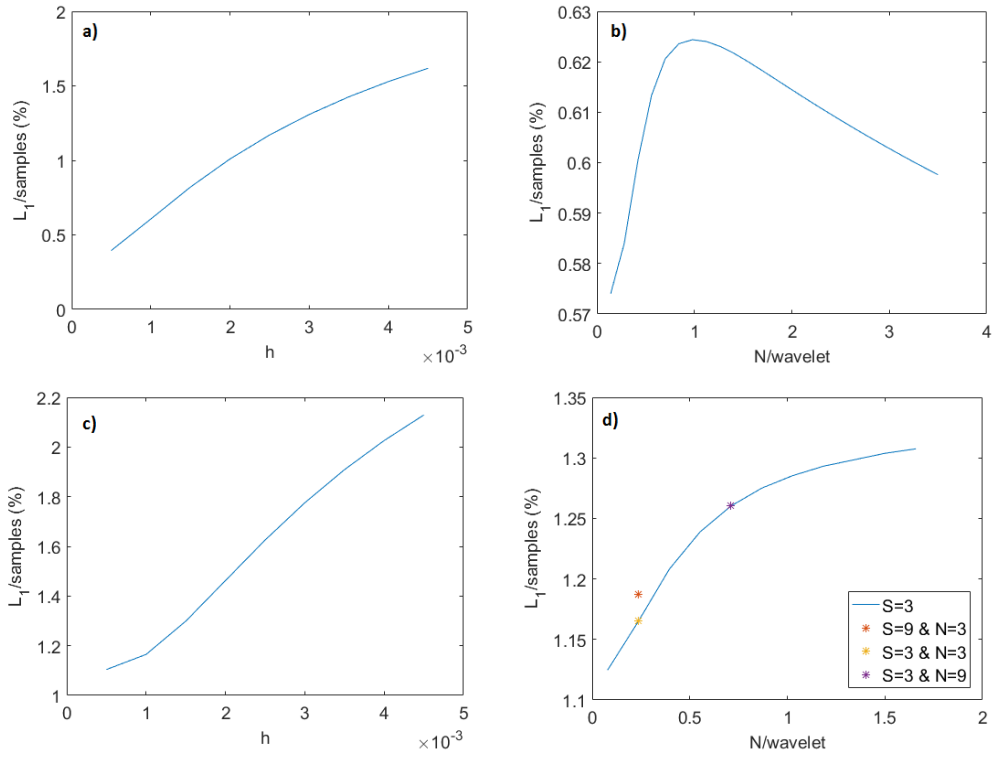


Figure 6: Normalized sum of absolute difference (L_1) between original (non-filtered) data and filtered data with different NLM parameters for a 2D seismic line (L2NAM1981M_815015, a and b) and a 3D seismic cube (L3NAM1987F minicube, c and d). NLM parameters are (a) S=9, N=19 and h is varying between 0.0005 and 0.0045, (b) S=9, N is varying between 1 and 30 and h=0.001, (c) S=3, N=3 and h is varying between 0.0005 and 0.0045 and (d) the blue line is S=3, N is varying from 1 to 21 and h=0.001, the orange star is L_1 for S=9, N=3 and h=0.001, the yellow star S=3, N=3, h=0.001 and the purple star is S=3, N=9 and h=0.001.

Figure 6b clearly shows that in a 2D seismic image the amount of filtering increases until the neighbourhood window (radius) is roughly the same size as the dominant wavelet. When the size of the neighbourhood window exceeds the dominant wavelet size, the algorithm becomes more selective and the amount of filtering decreases. Since the neighborhood window parameter is defined as the radius of the window, the dominant wavelet actually fits twice in the diameter of the window at the point were filtering becomes more selective. However, the power spectrum of the 2D dataset shows a wide range of frequencies, since frequencies change with depth. Therefore many wavelets will be larger than the dominant one and the decrease in NLM filtering is therefore also at a higher number of samples. This is in agreement with [Bonar & Sacchi, 2012], who suggests that the the neighbourhood size must be chosen such that it is large enough to encapsulate the structures of interest within the data. Figure 6c and d show the effect of varying the filter parameters h and N on the 3D minicube of the Friesland platform. The dominant frequency of the 3D dataset is 20 Hz, resulting in a dominant wavelet of on average 13 samples. The power spectrum of frequencies of the 3D data set is much more defined as the one of the 2D set, since we have cropped the 3D dataset at a small depth interval, which excludes all the high frequencies present at shallow depths. Striking in figure 6c and d is the absence of the kink in graph d at N = dominant wavelet (N/wavelet=1). We suggest this effect is somewhat eased by the addition of an extra dimension perpendicular to the time (trace) direction. Similarity of the neighbourhood windows is calculated in 3 dimensions, of which 2 are horizontal and only one parallel to the traces depicting the wavelets. The influence of the wavelet size on the similarity calculations has therefore decreased, compared to the 2D case. Due to long computation times of filtering 3D data, we were unable to denoise the 3D data set with an neighbourhood window larger than 21 samples, to test if the sum of absolute difference decreases with larger N values. Figure 6d shows the difference in L_1 for increasing the neighbourhood window N and keeping the search window S constant (yellow and purple stars) and the difference in L_1 when keeping the neighbourhood window constant and increasing the search window S (yellow and orange stars) with the same amount of samples. Figure 6 does not show the extend of h and N toward zero. Note that setting h and N to zero, results in zero filtering of the data and therefore also an L_1 value of zero.

Since changing the search window S does not have a great influence on the amount of filtering and is computationally very demanding, we have chosen to not vary this parameter, but keep it constant. The search window is chosen as large as the computation time allows for, which is $S = 3$ in the 3D case. For filtering the 3D cube and the rest of the study, filter parameter h is varied between 0.001 and 0.005 and neighbourhood window radius N between 3 and 15.

3.2 Petrel ant track workflow

3.2.1 Variance

The trace-to-trace variance (Edge method) volume attribute in Petrel is a semblance based coherency analysis [Petrel, 2015]. Variance is an edge detection method and is often used as a tool for better identification of faults in seismic data and is assumed to be a suitable input for the ant tracking algorithm [Silva et al., 2005]. There are several ways to quantify similarity of seismic data. According to [Petrel, 2015] Petrel software uses a semblance algorithm to compute variance (1-semblance). Semblance is calculated between a target trace and its surrounding traces. The amount of surrounding traces is the horizontal analysis window (inline and crossline range). Within a chosen time window (vertical smoothing) the energy of the input traces is calculated and an average is computed. The semblance is the ratio of the energy of the average trace to the energy of the input trace [Liu & He, 2012, Taner et al., 1979, Marfurt et al., 1998]. Semblance should always be calculated parallel to the orientation of the reflector, to avoid the computation of high values of variance due to the dipping reflectors rather than actual structural discontinuities. Figure 7 shows the effect of the size of the horizontal and vertical analysis windows on the computed variance of our 3D seismic cube. In general, keeping the windows small, increases resolution and will result in the detection of smaller faults. However, calculating variance with small windows is more sensitive to noise. Figure 7a to c show the effect of the vertical analysis window. Making the vertical window (time interval) smaller than the dominant wavelet (figure 7a) results in a very wormy pattern. When looking at an inline of this variance cube, it shows that this pattern is the result to high variance values along the reflectors, rather than discontinuities in the horizontal continuity of amplitude. Even applying a dip correction to the attribute does not entirely fix this response. In order for the variance attribute to focus on the vertical discontinuities like faults, the

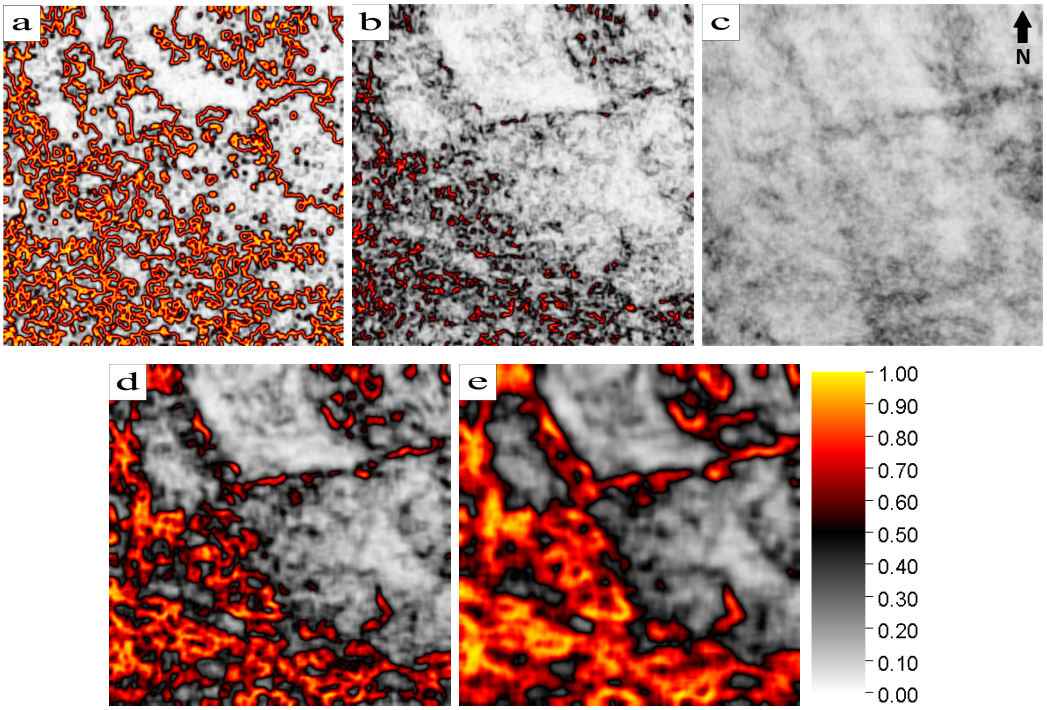


Figure 7: Variance of time slice at 2968 ms of original seismic data, using window sizes: inlines x crosslines x time samples (a) 3x3x6, (b) 3x3x15, (c) 3x3x191, (d) 6x6x15 and (e) 11x11x15.

lower bound for the vertical windows is the number of samples of the dominant wavelet. Figure 7 b, d and e show the effect of varying the horizontal analysis window. Enlarging this window increases the averaging and between traces and reduces the resolution. A larger horizontal window is less sensitive to noise and when looking at a larger area it is suitable for detecting large scale faults. However, since our cube only covers a small area and we want to find both small and large faults, a small analysis window was chosen (3x3x15). The choice of variance parameters for the purpose of fault detection is thus dependent on the size of your study area, the size of faults you want to detect, the amount of noise in your data and the length of the dominant wavelet.

3.2.2 Ant tracking

The ant tracking algorithm in Petrel is based on the behavior of ant colonies in nature. Ants in a colony mark their paths using pheromones for other ants to follow in their search for the shortest way towards food [Silva et al., 2005]. This principle is used in the algorithm. Virtual ants (agents) are placed on a seismic discontinuity volume (variance cube) and move along planar discontinuity structures while emitting pheromone. Surfaces which are likely to be faults will be traced by many agents deployed at different positions in the volume and will thus be strongly marked by pheromones, while unstructured surfaces (like noise) will be tracked by fewer ants and therefore weakly marked by pheromones [Pedersen et al., 2005]. Ants are only allowed a 15° deviation from their original direction when tracking discontinuities [Schlumberger, 2016]. This constrained is to enlarge the likelihood of ants tracking faults (planar structures) rather than random noise [Fang et al., 2017, Zhao & Sun, 2013]. The resulting ant track discontinuity volume is strongly affected by several parameters: Initial ant boundary, ant track deviation, ant step size, illegal step allowed, legal step required and the stop criteria. Table 1 shows the default settings (Passive and Aggressive) and the range for all ant track parameters. The definitions of these parameters and their effect on the resulting fault interpretation of the variance cube, are discussed below.

	Initial ant boundary	Ant track deviation	Ant step size	Illegal step allowed	Legal step required	stop criteria
Passive	7	2	3	1	3	5
Aggressive	5	2	3	2	2	10
Minimum	1	0	2	0	0	0
Maximum	30	3	10	3	3	50

Table 1: Ant track parameters for predefined settings: passive and aggressive and the range per parameter.

The first step in the ant track algorithm is defining the initial distribution of agents (ants). This is done by choosing the initial ant boundary, which is the radius (in voxels) around each ant. No ant is placed within the radius of another ant. A smaller boundary means more ants are initially placed in the volume and so the seed points will be denser. Each ant makes an initial estimate of orientation for the local maximum of variance within its boundary. This estimation defines the tracking direction for that particular ant, from which it can not deviate more than 15°. A larger initial ant boundary might help in a better estimation of the orientation of faults. However, it has a much larger effect on the number of faults that are extracted from the volume. It also results in less defined faults. Since ants are located relatively far from each other, they are less likely to follow the same tracks. The user has no influence on the location where ants are placed during this initial distribution, nor what minimum variance value to define as an edge value (local maximum). Ants are not allowed to track outside of the 15° range from the initial orientation, to maintain planar structures. Voxels within this 15° range are defined as legal positions. A local maximum just outside this range can be taking into account however, using the ant track deviation parameter. When the ant track deviation parameter is set to 1 the ant is allowed to deviate by one voxel in both directions from the legal position to search for a local maximum. If it finds one, the legal position closed to the maximum will be tracked (marked as a legal step). The ant step size parameter defines the number of voxels an ant advances in one step and can search in for a local maximum. A higher value will increase the chance of finding a local maximum parallel to the track orientation per step, but will decrease the resolution of the tracks. The illegal steps allowed constrains how many steps (amount of voxels depending on ant step size) an ant track can continue without finding any local maximum. If it exceeds the set value for this parameter, the ant track is terminated. This parameter has to be considered in combination with the legal step required parameter. This parameter defines the number of continues steps that must contain a valid edge value (local maximum). It is a constraint on how connected a track has to be, to be determined as a fault. The last ant track parameter is also related to the illegal step parameter. The stop criteria is the percentage of illegal steps allowed throughout an ants entire search space. A track might satisfy the conditions of the illegal steps allowed and legal steps required, but after continuation the percentage of total illegal steps might become so large, the likelihood of the track being a fault becomes questionable. When this percentage is reached the ant track in this direction is terminated. Petrel offers two different default settings of ant track parameters, passive and aggressive (table 1). The passive mode contains parameters which results in detection of only major regional faults zones. The ants require a strong and continuous signal in order to proceed. The aggressive mode consists of parameters which allow the ants more freedom to detect also subtle connections. This mode is designed to extract both major and subtle faults zones [Schlumberger, 2016]. Since we want to determine the widest possible range of fracture sizes, the aggressive mode was chosen for fault detection.

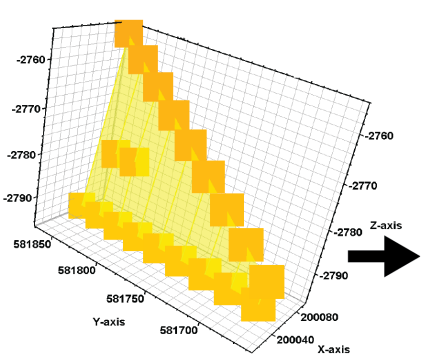
3.2.3 Automatic fault extraction

Automatic fault extraction is the next step in the Petrel ant tracking workflow. It generates fault patches from the ant track volume. The automatic fault extraction is used rather than the classic human interpretation to decrease computation time of fracture interpretation and to increase objectiveness. Faults are interpreted with minimum human intervention. This step in the total workflow can therefore be seen as a constant, allowing us to compare differences of fracture network characteristics based on differences in the seismic input rather than human subjectiveness. However, the result of the automatic fault extraction process is still influenced by parameters set by the user (see table 2). These parameters influence the accuracy of the fault picking and directs the process to search for small or large faults [Schlumberger, 2016]. There is no standard set of

Parameter	Minimum	Maximum	Chosen value
Extraction sampling distance	3	30	6
Extraction sampling threshold	Top 10%	All	Top 30%
Extraction background threshold	Top 10%	All	Top 60%
Deviation from plane	2	21	5
Connectivity constraint	1	3	2
Minimum patch size (points)	3	100000	3
Patch down sampling (voxels)	1	50	4

Table 2: Range and chosen user parameters for the automatic fault extraction algorithm.

parameters suitable for each dataset. Parameters should always be chosen based on the quality, size and expected fault patterns in the seismic data. The parameters and our argumentation on the chosen value are described below. The first two parameters are the Extraction sampling distance and Extraction sampling threshold. The first defines the distance between seed points, which controls the minimum distance between extracted fault patches. The distance is a radius and is quantified in voxels. The algorithm starts searching with steps defined by the extraction sampling distance for the highest values in the volume (top 10%), when such a value is found a seed point is created. The algorithm will continue searching and creating seed points of lower values (top 20 %, 30 % etc.) until the extraction sampling threshold is reached. The extraction sampling threshold is defined as a percentage of the ant track volume signal range (figure 17c). It is important to realize that the extraction sampling threshold percentages are based on the maximum and minimum ant track values [-1,1] and not on the maximum and minimum ant track values found in the specific input ant track volume, which might have a smaller range (e.g. [-1,0.2]). Once all the seed points in the volume are picked, the algorithm will search around every seed point for values within the value range defined by the extraction background threshold, the value range for this parameter is the same as for the extraction sampling threshold. The search around a seed point is restricted spatially by the Deviation from a plane parameter. A initial plane orientation is defined around the seed point and surrounding points will only be included in the fault patch if they have a value higher than the extraction background threshold and their deviation from the initial orientation (in voxels) is less than the value set by the Deviation from a plane parameter. The deviation from a plane parameter should always have a smaller value than the extraction sampling distance. This to prevent the tracking of multiple events in one fault patch, which is likely to happen if two near parallel faults are close together. The connectivity constraint defines how well connected a point has to be to be included in the fault patch. A voxel is defined as box with faces around the sample (point). The value of the connectivity constrained is the number voxel faces that has to be connected to another voxel which also satisfied all previous parameters. Whenever there is a lot of noise in your data, setting a high connectivity constraint will help in excluding noise from the fault patch. Points that satisfy all previous parameter thresholds are points that together form a possible fault patch. We shall call these *validpoints*. The last two parameters, minimum patch size and patch down sampling together control the minimum size of fault patches extracted and thereby the total number of fault patches extracted. The value for the minimum patch size parameter defines the minimum number of points a fault patch has to contain to be added to the fault patch set. Its minimum value is 3, since that is the minimum number of points needed to create a surface. Whenever the patch down sampling parameter is set to 1 voxel, the value for the minimum patch size defines the actual minimum number of valid points a fault patch should contain to be extracted. The Patch down sampling parameter controls the density of points of the extracted fault patch, a value of 1 is the highest possible density (every voxel is sampled). Lower densities decrease the computation time and smooths the faults. However, increasing the sampling distance (Patch down sampling) excludes fault patches with a small number valid points. By multiplying the patch down sampling value with the minimum fault patch size value we get the actual minimum size in valid points (voxels) a fault patch needs to contain, to be extracted. The choice in filter parameter for this study (table 2) are based on finding the widest fracture size range and to determine their spatial distribution within the cube. The constraints on both requirements are vertical and lateral resolution. Faults with a throw (or offset) smaller than the vertical resolution ($\lambda/4$) [Yilmaz, 2001] should not be visible in the seismic data and can thus not be tracked as faults. Note that our study area is a cropped volume between the two-way-travel time interval 2744 and 3504 ms. The power spectrum of the frequencies in the Fourier domain




```
200018.69696  581663.25271  2788.80000 L3NAM1987F_minicube_NLM_3_3_0_001_4 1
200022.03044  581663.25271  2784.44444 L3NAM1987F_minicube_NLM_3_3_0_001_4 1
200023.69719  581688.25090  2789.60000 L3NAM1987F_minicube_NLM_3_3_0_001_4 2
200030.36417  581688.25090  2780.88889 L3NAM1987F_minicube_NLM_3_3_0_001_4 2
200028.69742  581713.24910  2790.40000 L3NAM1987F_minicube_NLM_3_3_0_001_4 3
200038.69789  581713.24910  2777.33333 L3NAM1987F_minicube_NLM_3_3_0_001_4 3
200033.69766  581738.24729  2791.20000 L3NAM1987F_minicube_NLM_3_3_0_001_4 4
200047.03162  581738.24729  2773.77778 L3NAM1987F_minicube_NLM_3_3_0_001_4 4
200038.69789  581763.24549  2792.00000 L3NAM1987F_minicube_NLM_3_3_0_001_4 5
200055.36534  581763.24549  2770.22222 L3NAM1987F_minicube_NLM_3_3_0_001_4 5
200043.69813  581788.24368  2792.80000 L3NAM1987F_minicube_NLM_3_3_0_001_4 6
200063.69906  581788.24368  2766.66667 L3NAM1987F_minicube_NLM_3_3_0_001_4 6
200048.69836  581813.24188  2793.60000 L3NAM1987F_minicube_NLM_3_3_0_001_4 7
200057.44877  581813.24188  2779.00000 L3NAM1987F_minicube_NLM_3_3_0_001_4 7
200072.03279  581813.24188  2763.11111 L3NAM1987F_minicube_NLM_3_3_0_001_4 7
200053.69859  581838.24007  2794.40000 L3NAM1987F_minicube_NLM_3_3_0_001_4 8
200063.69906  581838.24007  2780.00000 L3NAM1987F_minicube_NLM_3_3_0_001_4 8
200080.36651  581838.24007  2759.55556 L3NAM1987F_minicube_NLM_3_3_0_001_4 8
200058.69883  581863.23827  2795.20000 L3NAM1987F_minicube_NLM_3_3_0_001_4 9
200088.70023  581863.23827  2756.00000 L3NAM1987F_minicube_NLM_3_3_0_001_4 9
```

Figure 8: Export of a fault patch in Petrel consisting of 9 fault sticks to an ASCII text file. Column (1) x-coordinates, (2) y-coordinates, (3) TWT [ms], (4) name of fault and (5) fault stick number.

(see section 3.1) is narrow and shows a clear peak around 20 Hz. We therefore assume a constant frequency and wavelength with depth within our study area. Whenever discontinuities of a size smaller than the vertical resolution are determined, it is most likely to be noise. The minimal vertical component of the fault throw for our data should be 3 voxels. The number of samples per wavelet is computed as the period of the dominant wavelet (∼50 ms in TWT) divided by sampling rate (4 ms), see section 3.1 . A quarter of this wavelet is the vertical resolution, which is ∼3 voxels. This is the minimal throw for a vertical fault. For a fracture with a dip of less than 90° the minimal throw has to be larger to overcome the vertical seismic resolution threshold. We have therefore picked the patch down sampling parameter a bit higher than 3. No constrained was set on the maximum fault length (along strike) to be extracted. The maximum fault size is already restricted by the sampling volume. The lateral resolution was used as a lower bound for the extraction sampling distance. Assuming, migration was carried out successfully, the Fresnel zone of lateral resolution has collapsed to λ [Stolt & Benson, 1986], which translates to 12 voxels (300m/trace spacing). The extraction sampling distance parameter was set to 6 voxels ($\lambda/2$) however, for more detailed sampling. The choice of values for the parameters Extraction sampling threshold and Extraction background threshold are based on the one of the predefined settings for normal confidence fault extraction. All chosen fault extraction parameters are listed in table 2.

After automatic fault extraction was carried out, a fault patch set is created in Petrel. To export these faults from Petrel, the fault patches first need to be converted to fault interpretations. For this step it is important to set the fault stick interval (under the Settings tab in the Automatic fault extraction window) to 1 trace. This should be done before the automatic fault extraction is carried out. By doing so, detail will be preserved when transforming (discretize) fault surfaces to fault sticks. These faults, consisting of fault sticks can then be exported as ASCII text files (figure 8), which are the input data for the python codes described in Database design.

3.3 Database design

The next step in the overall fracture network characterization workflow is to convert the fault stick ASCII files exported from Petrel into individual fracture attributes. Per fault stick file, which describes one fault, the following attributes are computed: fracture barycenter, strike length, dip length (offset), strike, azimuth, dip and aspect ratio. The database design is constructed in python and consists of 3 python codes: *userinput.py*, $Fault\_data\_3D.py$ and $faultcalculations\_3D.py$. In *userinput.py* the user defines the following parameters: number of faults extracted, fault stick interval in meters, average overburden velocity, top of reservoir in two-way-travel time (TWT), bottom of reservoir in TWT, the path of the folder containing all the fault stick files, the path where to save the faultcalculations file and the path where to find the polygon shape file determining the systems edges. This data is then imported into the $Fault\_data\_3D.py$ code. This code loops over the total number of faults and during every iteration it reads the fault stick file under

Fault attribute	Input	Equation
Barycenter	End points of faultsticks $i = [1, N]$	
X_0	$(x_1, y_1, z_1)_i$ and $(x_2, y_2, z_2)_i$	$X_0 = \frac{1}{2N} \sum_{i=1}^{N} [x_1 + x_2, y_1 + y_2, z_1 + z_2]$
Strike length	Point at same z per faultstick	$l_i = \sqrt{(x_{z,i+1} - x_{z,i})^2 + (y_{z,i+1} - y_{z,i})^2}$
l	$(x_z, y_z)_i$	$l = \sum_{i=1}^{N-1} l_i$
Strike	$l_i, (x_z, y_z)_i$	$0° \leq \psi \leq 180°$
ψ	and $(x_z, y_z)_{i+1}$	$\psi = \frac{1}{l} \sum_{i=1}^{N-1} \tan^{-1} \left(\frac{x_{z,i+1} - x_{z,i}}{y_{z,i+1} - y_{z,i}} l_i \right)$
Offset	$(x_1, y_1, z_1)_i$ and $(x_2, y_2, z_2)_i$	$O_i = \sqrt{(x_1 - x_2)_i^2 + (y_1 - y_2)_i^2 + (z_1 - z_2)_i^2}$
O		$O = max\,[O_i]_{i=1}^{N}$
Dip	$O_{avg} = \frac{1}{N} \sum_{i=1}^{N} O_i$	
θ	$dz_{avg} = \frac{1}{N} \sum_{i=1}^{N} \lvert z_1 - z_2 \rvert_i$	$\theta = sin^{-1}\,(dz_{avg}/O_{avg})$
Aspect ratio	ψ and O	
ψ/O		$\psi/O = \frac{\psi}{O}$
Azimuth	ψ, $(x_1, y_1, z_1)_i$	
ϕ	and $(x_2, y_2, z_2)_i$	$0° < \phi \leq 360°$
		$\phi = \begin{cases} \psi + 90° & \psi \leq 45° \text{ and } x_1 \geq x_2 \\ \psi + 270° & \psi \leq 45° \text{ and } x_1 < x_2 \\ \psi + 270° & 45° < \psi \leq 135° \text{ and } y_1 \geq y_2 \\ \psi + 90° & 45° < \psi \leq 135° \text{ and } y_1 < y_2 \\ \psi - 90° & \psi > 135° \text{ and } x_1 \geq x_2 \\ \psi + 90° & \psi > 135° \text{ and } x_1 < x_2 \end{cases}$

Table 3: Fault attribute calculations carried out in $faultcalculations\_3D.py$. Note that $(x_1, y_1)_i$ is the first point of every fault stick of which the z-value is the highest (deepest) and $(x_2, y_2)_i$ is the last point of every fault stick. The range of fault sticks per fault is denoted as $i = [1, N]$.

consideration, applies a time-depth conversion to the Z-coordinates of the fault sticks and calls the functions which calculate fracture length, azimuth, dip, aspect ratio and barycenter.

These attributes are then stored and the loop moves on to the next fault stick file. After all iterations are completed, all attributes of all faults are stored in in a text file and saved (figure 9). Some exceptions are built into the code. Faults that intersect or overlap with the system boundary are excluded. In general, the ant track workflow of Petrel does not allow faults to touch or intersect with the edge boundaries. However, if your seismic volume has indentations of no data, the intersection between the no-data zone and the seismic volume are marked as high variance zone and the ant track workflow extracts faults along these boundaries. Two functions are implemented to correct for this misinterpretation. The first excludes faults which intersect or overlap with the polygon surrounding the system and the second excludes faults which have a strike of exactly 0 or 90 degrees, which is the orientation of our cube. The purpose of this second function is also to exclude faults which are based on migration artifacts, which are parallel to the inline and crossline and are not filtered out by the NLM filter. The reason for assuming this exclusion to be valid, will be clear when we explain the calculation of the strike attribute. Note that the code thus assumes the study area to be orientated parallel to the inlines and crosslines and that the inlines and crosslines are assumed to be north-south and east-west orientated. Another exception build into the code is that faults that only consist of one fault stick are not taken into account. These faults do not contribute to the statistics since no strike length or strike can be computed.

For the calculations of the fault attributes a few assumptions are made. The code is built for the Netherlands RD new coordinate system, which uses meter as unit of measure and for which

	A	B	C	D	E	F	G	H	I	J
1	# faultID	barycenterx	y	z	strike	azimuth	dip	maxdiplength	length	aspectratio
2	0	198013.603	580015.2948	4577.511538	90.52203271	0.522032715	61.4110431	163.1889476	300.3728727	1.885523554
3	2	200169.0165	582125.7193	4601.957875	12.45829285	282.4582929	70.79336092	146.1461022	487.7173783	3.351138256
4	3	200138.7026	583408.96	4564.969	80.6280453	170.6280453	32.78989629	89.20895624	73.2375393	0.70060506
5	4	200047.0316	581763.2455	4538.129	12.65401417	282.6540142	75.77697982	122.107392	204.9639963	1.670226383
6	5	198963.6475	578442.5294	4670.473882	111.539225	201.539225	54.57371077	130.9449585	433.2746848	3.210150009
7	6	198593.8386	578725.9648	4595.66725	138.9223058	228.9223058	78.49265954	48.312	165.8801463	3.367894769
8	7	199776.1856	582002.2907	4516.657	107.7366303	17.73663033	77.08509931	88.16425647	78.78739894	0.862456612
9	8	200701.2289	582254.7279	4663.534107	51.36752329	321.3675233	68.77112933	184.0642261	420.1801609	2.421425031
10	9	200338.7119	582029.4763	4636.2292	126.8675642	36.86756423	63.93217552	108.7926302	125.1287083	1.148979485
11	10	198176.1107	580035.2453	4589.125	79.39912846	349.3991285	75.33785809	77.99041966	127.2138899	1.666959612

Figure 9: Output text file containing all fault attributes, of *Fault_data_3D.py* and *faultcalculations_3D.py*. Barycenters x,y,z, maxdiplength (offset) and length are rendered in meters. The unit for strike, azimuth and dip is degrees (°).

the north is equal to the y-axis. Furthermore, fault stick files are assumed to have the standard structure of fault stick files from automatically extracted faults. The features of this structure are: per fault stick the first point is always the deepest (largest z-value, figure 8), the counting of fault sticks is ordered from left to right or right to left, but does not start in the middle and there is no limit to the number of fault sticks per fault or the number of points per fault stick.

Table 3 shows all the calculations carried out in *faultcalculations_3D.py*. Some additional remarks concerning these calculations are discussed below. For most calculations the end points of all fault sticks are needed, these are denoted by $(x_1, y_1, z_1)_i$ and $(x_2, y_2, z_2)_i$. Where i is the fault stick number within the fault, subscript 1 indicates the first point of fault stick i and subscript 2 the last point of fault stick i. The barycenter of each fault is computed by taking the mean of all the end points within the fault patch. To compute the strike and the strike length some linear algebra is used to find a point $(x_z, y_z)_{i+1}$ on the line between the two end-points of a fault stick at the same height as the middle point of the neighboring fault stick, as the strike along a fault plane should always be measured horizontally. The strike is then computed as a weighted average of all the separately computed strikes between fault sticks. The weights are defined as the length between the succeeding fault sticks l_i. The total strike length is the sum of l_i. The strike is computed with an implementation making sure the strike cannot exceed 180°. Due to linear algebra and the weighted average applied, it is very unlikely that a fault has an exact strike of 0 or 90 degrees. For this reason we conclude that our implementation of excluding these faults, mentioned previously, is valid. The offset (dip length) is defined as the maximum fault stick length in the fault. Whereas the dip is computed by trigonometry using the average fault stick length and dz. This averaging is applied to minimize the effect of an odd fault stick. For the computation of the azimuth of the fault patch under consideration, the already computed strike is used as input variable, as well as the end points of all the fault sticks. Based on the relative difference in x and y coordinates of the end points of the fault stick, we can determine whether 90° (or 270° to ensure $0° < \phi \leq 360°$) has to be added or subtracted from the strike to compute the azimuth. This is computed for every fault stick in the fault, after which the most numerous occurring fault stick azimuth is adapted as azimuth for the whole fault. This final implementation is included for faults whose dip is close to 90° and of which some fault sticks might be dipping in the opposite direction as the majority.

After all calculations are carried out, all fault attributes along with its fault ID (number) are written to a text file (figure 9) which is used as input for the *Fault_statistics.py* code.

3.4 Fracture network characterization

Now that all fault attributes are determined, fracture network characteristics can be analysed. To realize this we use 2 codes: *Fault_statistics.py*, which uses *userinput.py* and the text file containing all fault attributes, as input. The second code is *PairCorrelationFunction_3D.py*, which computes $C_2(r)$ and D_c from equation 5 and 6. In the *Fault_statistics.py* code, the distributions of the following network characteristics are computed: fracture (strike) length distribution, spatial distribution, orientation (azimuth) distribution and dip distribution. The computation and analysis of these distributions will be discussed below.

3.4.1 Fracture length distribution

The most commonly used method to obtain the length exponent (or scaling parameter) a of a data set of which the length distribution is expected to follow a powerlaw (equation 1, 2 or 3), is to calculate the slope of a histogram of double logarithmic axis, by least-squares linear regression. However, when using this method the choice of bin size dl is critical because it defines the smoothness of the histogram [Davy, 1993, Bonnet et al., 2001]. A larger bin size will result in a smoother curve, but the smoothness of the curve might influence the slope as well and thus the length exponent a, associated with the distribution. To estimate a for a powerlaw fit to our length distribution, we therefore used a method proposed by [Clauset et al., 2009], called the method of maximum likelihood. A python package has been provided by [Alstott et al., 2014], which computes a between user defined l_{min} and l_{max}. This method is based on the probability that the data were drawn from a powerlaw model with scaling parameter a. The derivative of the logarithmic likelihood is taken with respect to a and set to zero to get the maximum of this function and thus the maximum likelihood that the data fits the powerlaw with parameter a. The result is the following function, to estimate a without the need to bin our data:

$$a = 1 + n \left[\sum_{i=1}^{n} ln \frac{l_i}{l_{min}} \right] \tag{10}$$

where n is the number of observations of the data, in our case the number of fractures of which the length has been determined and l_{min} is the lower bound for which the powerlaw behavior holds. The determination of l_{min} is discussed below. For visualization of the length distribution, fractures were binned using logarithmic binning. The same binrange (25 to L) and number of bins was used (40) for all data sets, for good comparison between distributions of the same data set but which were filtered with different NLM parameters. Logarithmic binning was used rather than linear binning to avoid empty bins at larger fracture lengths, since large faults are expected to be less abundant than small faults.

The type of distribution used, is the density length distribution (equation 3), since it is less affected by finite size effects as the cumulative distribution (equation 4) and the truncation effects are more clearly visible. This makes estimation of the lower-cut off of the powerlaw fit, due to resolution, easier. Furthermore, Davy's double powerlaw model (equation 1), which we hope to find for our data, is based on the density distribution.

Fitting of the powerlaw line with exponent a was done by finding the minimum first norm between our binned data and a powerlaw fit line with constant slope a (determined by equation 10) but varying intersection points (b) with the y-axis ($y^{fit} = bl^{-a}$). In order to let all data points weigh equally on the log-log scale, L^1 was computed according to the exponents on the log scale, to simulate linear regression on a log-log scale. The minimum found first norm was then normalized by the number of data points (filled bins, I), which was used as a measure for the misfit of the powerlaw (equation 11).

$$\hat{L}^1 = \frac{min \left[\sum_{i=1}^{I} | \log y_{i,b}^{fit} - \log y_i^{data} | \right]_b}{I} \tag{11}$$

From studies on fracture length distributions of 2D trace maps the lower bound of the powerlaw, due to resolution (truncation) effects, as explained previously, was found to be between 5-10% of the domain size L ($L = \sqrt{area}$), as explained previously [Odling, 1997, Lei et al., 2015]. The lower cut-off for the fracture length distribution from our 3D fracture network also seems to fall in this range (where $L = \sqrt[3]{volume}$). Since all length distributions tested in this study are from the same original 3D seismic data set the lower cut-off for the length distribution, associated with resolution, is assumed to be the same. The lower cut-off was determined within the range mentioned above, by finding the minimum of the sum of the normalized first norms ($\hat{L}^1$) of the length distributions from the same data set filtered by different NLM parameters. For determination we have used 9 different combinations of NLM filter parameters ($N = [3, 9, 15]$ and $h = [0.001, 0.002, 0.003]$) and fracture network from the non-filtered data ($NLM = [0, 9]$).

$$l_{min} = min \left[\sum_{NLM=0}^{9} \hat{L}_{NLM}^{1} \right]_{0.05L}^{0.1L} \tag{12}$$

For the fracture networks extracted from the L3NAM1987F minicube the trunctaion cut-off was estimated to be at 7.5% of L, using the above method. By estimating and encountering the truncation cut-off we have corrected for one of the two sampling effect, truncation. Correction for the other sampling effect, censoring, was not applied. Since the ant track workflow of Petrel does not allow fault patches to extend and intersect with the system boundaries, no estimation can be made on the amount of faults affected by censoring.

3.4.2 Fractal Dimension

The spatial distribution of fractures were determined by computing the pair-correlation function as in equation 5 and the local slope associated with it, equation 6. We have adapted this method to our 3D situation and applied it to the fracture barycenters. $Fault\_statistics.py$ calls the pair-correlation function in $PairCorrelationFunction\_3D.py$ and gives the the coordinates of the fracture barycenters and the typical system size L as input parameters. Typical system size L is computed as follows $L = \sqrt[3]{polygonarea * depthinterval}$. The pair correlation function computes $C_2(r)$ as follows. The code loops over all points (barycenters) in the system and computes for every points its distance to all the other points in the system, using Pythagoras for 3D. From this computation the minimum distance between two points is used to define the lower bound r_{min} for the bin range in which the number of pairs of points will be counted. The maximum value for the bin range is the typical system size L and the bins are spaced in logarithmic order. The algorithm than counts how many pairs of points have a distance between r_{min} and the edge of the bin r, resulting in an array containing N_p values per bin which upper edge is the corresponding r. $C_2(r)$ is than computed according to equation 5 and D_c as in equation 6. Our is system is not an isotropic volume (cube which sides all have the same length), L_x and L_y have the same length, but L_z is smaller. This will affect the extend for which the pair-correlation function can follow a powerlaw. Because L is larger than L_z, the number of pairs of points for $r > L_z$ will go down as an effect of our volume being anisotropic. Whenever a system is analysed which is even more anisotropic as our system ($L_x \neq L_y \neq L_z$) it is possibly better to compute the pair correlation function in 2D in the orientation for which the plane shows the highest isotropy and convert the 2D fractal dimension to a 3D fractal dimension using the stereological rules of [Darcel et al., 2003a]. An implementation to compute the 2D pair correlation function for different depth intervals is also built into the code.

3.4.3 Orientation distribution

The last network characteristic distributions to be computed, are related to orientation. Strike and azimuth distributions are visualized by rose diagrams, whereas the dip distribution is represented by a normal histogram. Data is binned in linear order. Bin sizes for the dip distribution are 5° and for strike and azimuth 10°. The bars indicate the percentage of the total number of faults $N(L)$ in the system, having an orientation in the range $[\psi, \psi + 10]$ or $[\phi, \phi + 10]$.

4 Results

We have determined the fracture network characteristics of automatically extracted fracture patterns from a cropped 3D seismic cube, enclosing part of the Friesland Dinantian carbonate platform. Several fracture patterns were extracted from the same 3D seismic data but filtered using different NLM parameter combinations. Our results show the effect of the use of different NLM filter parameters on the fracture network characteristics as length, spatial and orientation distributions. Data was filtered using a wide range of filter parameters. For visualization of the effects we have chosen the following range, $N = [3, 9, 15]$ and $h = [0.001, 0.002, 0.003]$. As mentioned before, the settings of Petrel attributes (Variance, ant tracking and automatic fault extraction) were kept constant (see tabel 1 and 2).

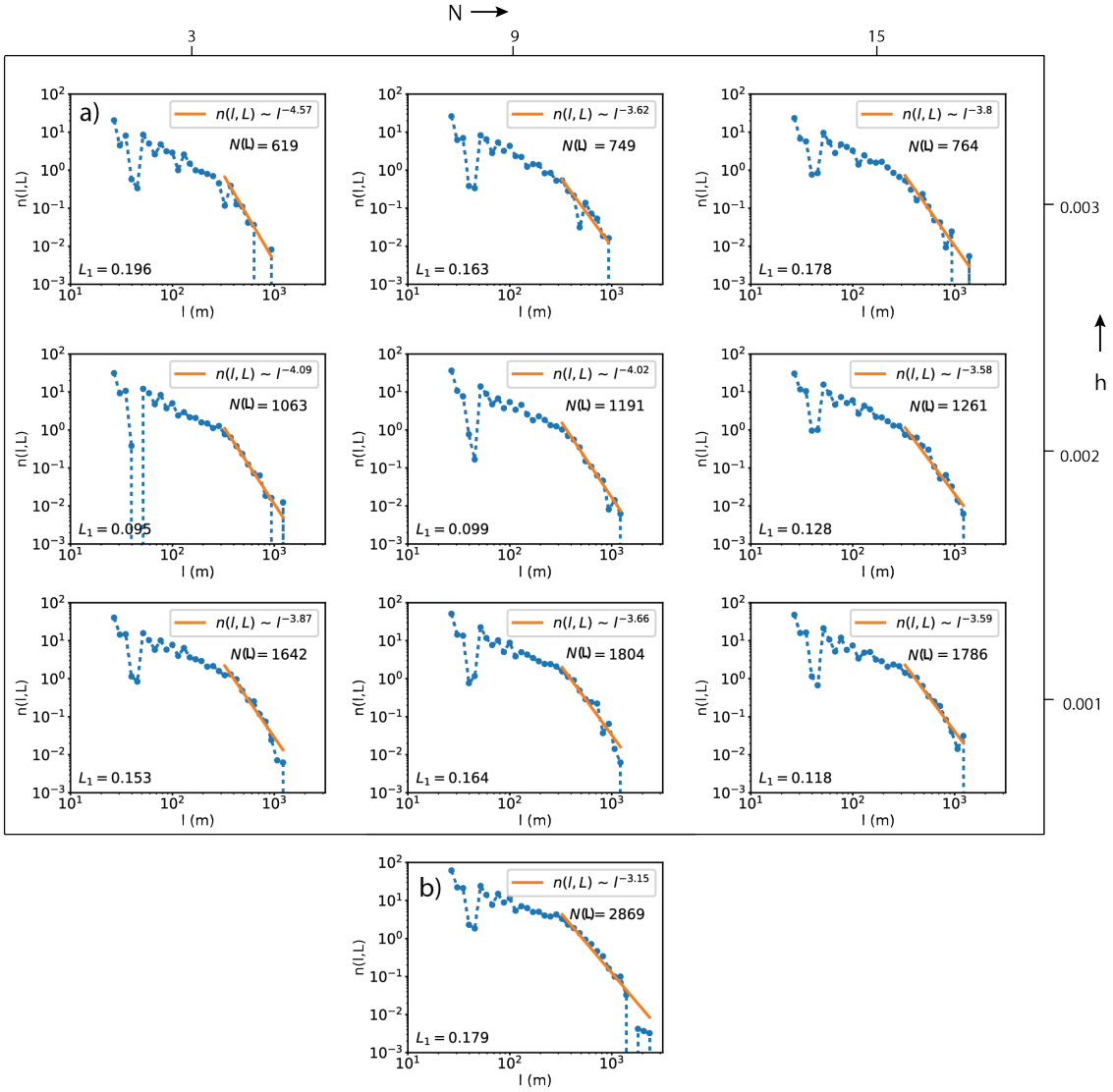


Figure 10: Density length distributions of fracture networks extracted from L3NAM1987F Friesland platform minicube (a) filtered using different NLM parameters and (b) non-filtered. NLM parameters, N and h, are varied as indicated along the outer axes. The search window parameter (S) was kept constant at 3 for all filtered data. Note that the N indicated on the top-outer axis is the NLM parameter: neighbourhood window and the N(L) inside the individual graphs is the number of fractures.

Figure 10a shows the effect of the NLM filter on the length distribution of the fracture network with respect to length distribution of the non-filtered data (figure 10b). The length distribution is represented by a density distribution. The blue dots represent the middle points of the logarithmically spaced bins. $N(L)$ indicates the number of faults which were extracted, from which the length has been determined. This means faults which only consist of one fault stick were not included in this distribution. The orange line is the powerlaw fit to the data, bounded by the truncation cut-off on the lower end, at $l_{min} = 7.5\%L$ and at the upper bound by l_{max}, the largest fault found in the system. Since all distributions are derived from the same system L, the lower bound (7.5% of L) is equal for all distributions. L_1 is the measure of misfit of the powerlaw to the data, the normalized sum of absolute difference (equation 11).The smaller this value the better the fit of the powerlaw to the data. The length exponents of the powerlaw fits computed, range between 3.15 for the non-filtered data, to max 4.57 for NLM filtered data. Recalling from section 2.3, the expected value of a powerlaw length exponent for a 3D system is between 2.7 and 3.75. Possible reasons for slightly higher results are discussed in 5. A striking feature in all graphs is the dip in the data between $l = 25$ and $l = 50$. This dip is the cause of a sampling effect, it is the

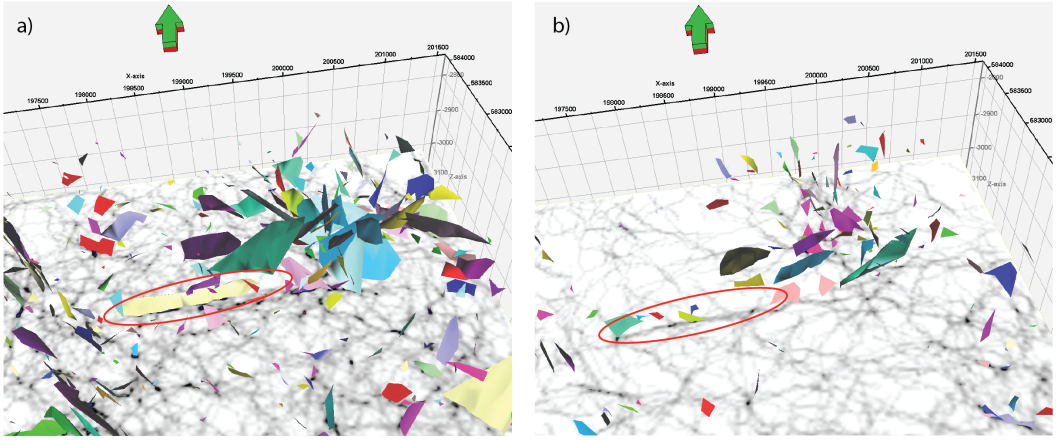


Figure 11: Fault extraction and ant track time slice (2988 ms) of (a) non-filtered data set and (b) of NLM filtered (S=3, N=15, h=0.002) data set. Red ellipse encircles large fault (a) in non-filtered data which is extracted as (b) several small faults in filtered data.

difference between faults consisting of 2 ($l \sim 25m$) and faults consisting of 3 fault sticks ($l \sim 50m$).

When comparing the length exponent a of the powerlaw fit for the non-filtered data (figure 10b) with the filtered data (figure 10a), we observe an increase in the powerlaw exponent a for filtered data. This suggests that the relative abundance of small fractures with respect to large ones, has increased after filtering the data. At first sight this is in contradiction with our hypothesis that filtering of the data will remove indications of small fractures along with noise and therefore reduce the number of small fractures. However, we should take into account that our powerlaw fit and thus the purpose of a does not extend to the distribution of the very small fractures, due to truncation. When merely looking at the data point for the smallest detected faults ($l \sim 25m$), we do observe a clear decrease in the abundance of small fractures for the filtered data, in agreement with our hypothesis. To explain the increase in length exponent a for the powerlaw range of the filtered data with respect to the non-filtered data, we have looked at the fault patterns extracted from Petrel in more detail. Figure 11a shows how a large fault (marked by the red ellipse) is extracted from the ant track volume. Whereas in figure 11b we see that at that same location this large fault is extracted as several small faults. We suggest this effect is the cause of the change in the slope of the powerlaw between filtered and non filtered data. In non-filtered data, large faults are extracted by linking small fault patches with noise to create large surfaces. When NLM filtering is applied, these large fault patches are aligned as a series of smaller faults. Possibly, this indicates the phenomenon of large fault zones being constructed of many aligned small faults [Walsh et al., 2003]. This effect is mainly a result of the filter parameter h, as will be discussed below. When looking at the trade-off between the NLM filter parameters (figure 10a) we observe that in general the increase of filter parameter h causes the length exponent a to increase. which results in the steepening of the slope. In contrast, increasing the neighbourhood window N, results in a decrease of a and thus a shallowing of the slope of the length distribution. When looking at the number of faults extracted $N(L)$, this decreases drastically with increasing h value but increases with increasing N. We suggest that the decrease in the number of faults $N(L)$ extracted with increasing filter parameter h, is a result of the decrease in the amplitude range of the seismic data. Since the filter parameter h defines how much filtering is applied, a larger value of h results in the overall weakening of the seismic amplitudes. The seismic amplitude range decreases approximately with 1-3% per increase of 0.001 h. This results in weaker defined discontinuities, therefore lower variance values and lower ant track values. Because the fault extraction threshold is kept constant, most ant track discontinuities will not be registered by the fault extraction algorithm, resulting in a smaller number of faults extracted. See figure 19 and 20 in the appendix to see the decrease in amplitude in the variance and ant track cube, with increasing h value. To explain the increase in $N(L)$ and decrease in a with increasing neighborhood window N we looked at the values of the ant track cube in more detail. This analysis shows that with increasing neighbourhood window N, the percentage of high ant track values (0.4-0.6) increases with a few percent. This means that that the algorithm is more robust with regard to the validity of the faults it has tracked. A larger

number of values above the extraction threshold not only results in more faults to be extracted but also in existing faults to be tracked over longer distances.

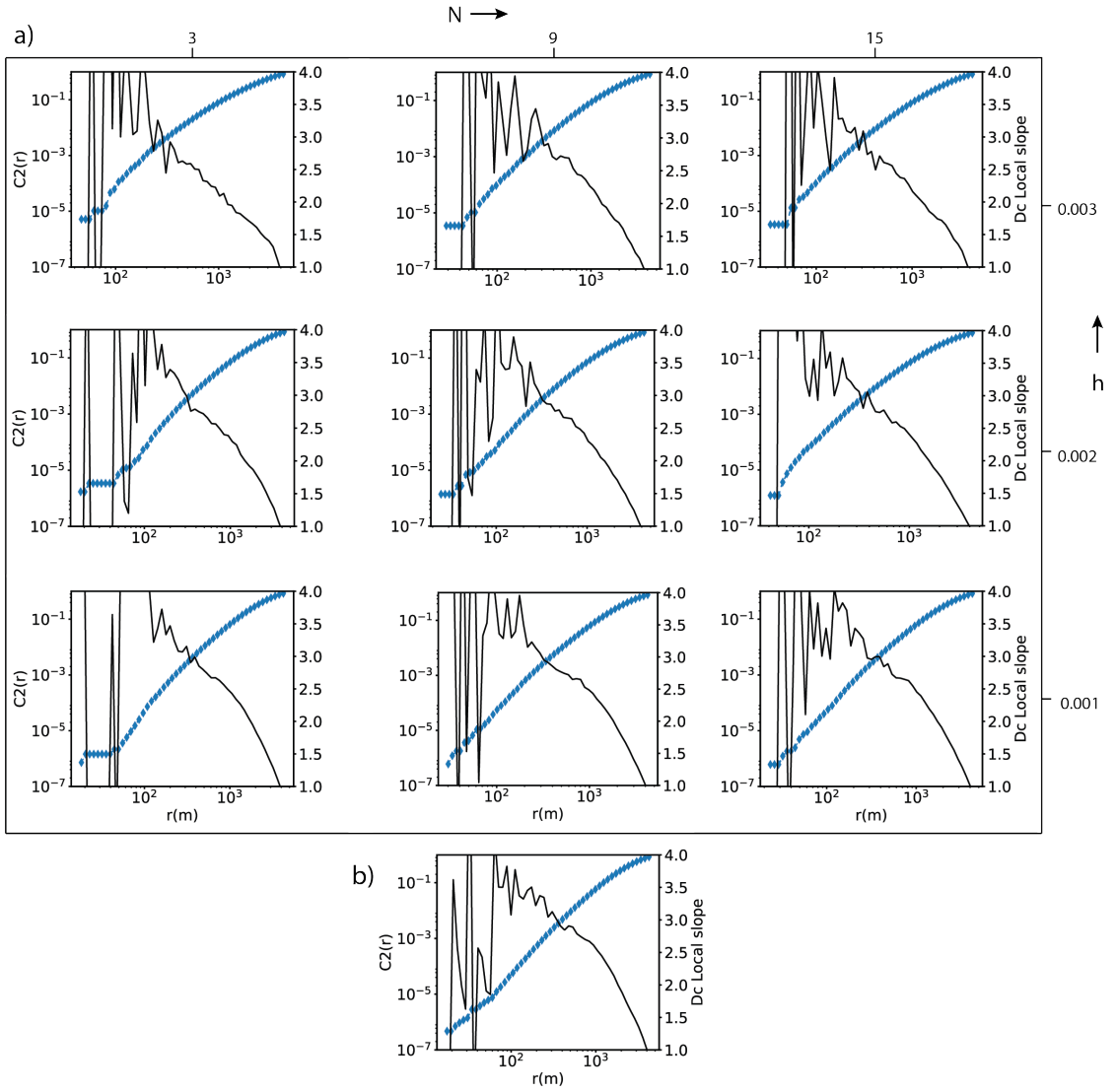


Figure 12: The spatial distribution of fracture barycenters of fracture networks extracted from the 3D seismic dataset: L3NAM1987F Friesland platform minicube, both (a) filtered using different NLM parameters and (b) non-filtered. Petrel attribute parameters used can be found in table 1 and 2. C_2 (blue curve) was computed by the pair correlation function. The right axis and black line represent the local slope $(-d(log(C_2(r)))/d(log(r)))$ of the C_2 curve. No plateau in the local slope is visible in any of the graphs.

Figure 12 shows the the pair-correlation function and associated local slope of fractures extracted from non-filtered (figure 12b) and filtered data (figure 12a). As explained before, a fractal dimension can only be extracted from the local slope of the pair-correlation function when a clearly defined plateau is visible in the graph of the local slope. A plateau in the local slope indicates if a sufficiently linear section exists to determine if the spatial distribution is fractal. Unfortunately, no significant plateau is visible in any of the graphs of figure 12. We suggest the absence of a sufficiently linear section of the pair correlation function is caused by two effects. The first is the insufficient spatial resolution of the seismic data, mainly influencing the lower part of the the pair correlation graph. The second causes the shallowing of the slope at the upper end of the graph and is the result of the anisotropic shape ($L_z < L_{x,y}$, see section 3.4.2) of our system. Less pairs of barycenters are detected at distances $r > L_z$, since at distances larger than L_z the algorithm starts searching partially outside the study area, where no faults are extracted. The range of these

two obstructions seems to be so large, that their upper and lower bound meet, leaving no range to determine a fractal dimension.

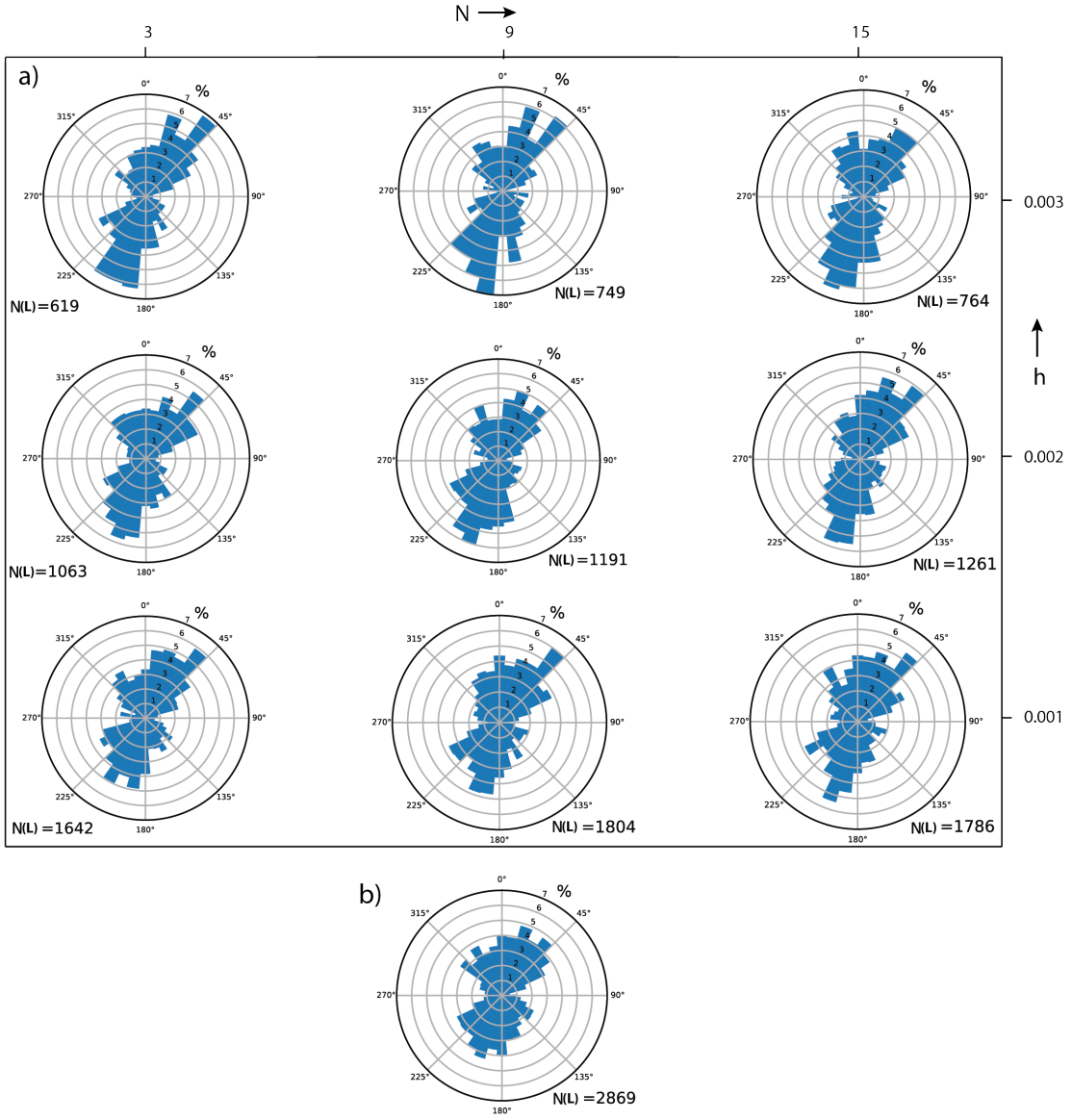


Figure 13: Rose diagrams showing the azimuth (dip direction) of fractures extracted from (a) NLM filtered data and (b) non-filtered data. The polar axis represent North (0°), East (90°), South (180°) and West (270°) and the radial axis represent the percentage of faults with a dip direction $[\phi, \phi + 10]$. Radial axis tick mark intervals are 1% and the maximum radial axis is 7%.

Figure 13 shows the orientation distribution of the fractures extracted from the filtered and the non filtered cropped 3D seismic cube. The highest frequency of fractures binned is 7% of the total number of fractures extracted, suggesting no clear preferred orientation. The rose diagram patterns do not seem to vary much between filtered and non-filtered distributions or between differently filtered data. This suggests that the NLM filter only filters the random noise and does not filter out structured patterns as faults.

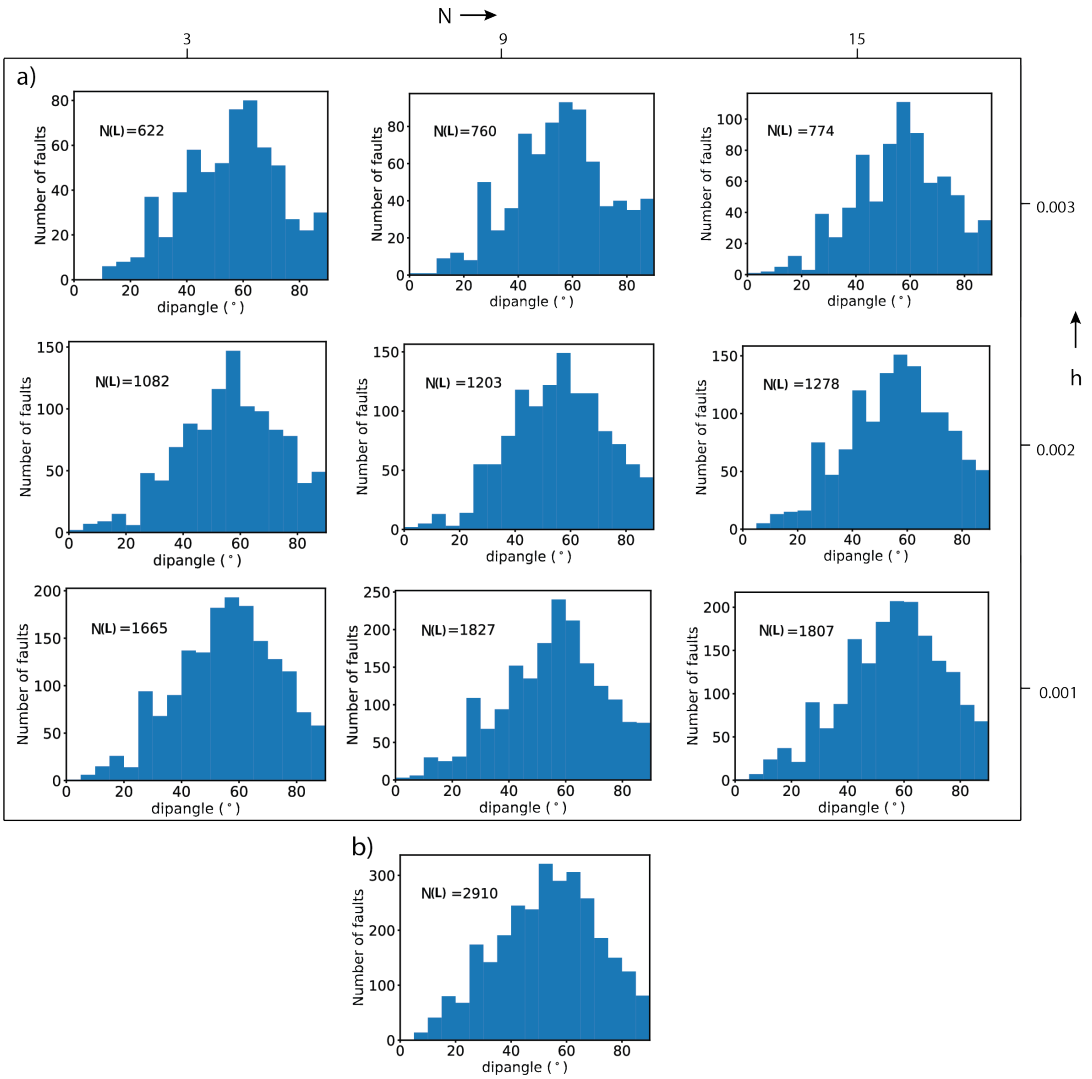


Figure 14: Histograms showing the dip distribution of fractures extracted from (a) NLM filtered data and (b) non-filtered data.

Figure 14 shows the last computed fracture network characteristic, the dip distribution. All distributions, for filtered (figure 14a) and non-filtered (figure 14b) show a peak at about 60° and all distributions show relative more faults with high dip angles than low dip angles. This is expected since fault offsets in seismic data can only be distinguished at higher dips. The histograms of NLM filtered data with filter parameter $h = 0.003$ show more fluctuation than the other filtered data. We suggest this effect is mainly the result of fewer fractures being extracted, resulting in less robust statistics.

5 Discussion

5.1 Petrel ant track workflow

All fracture characteristics discussed in section 4 show to some extent the influence of the varying NLM filter parameters. However in many cases these differences seem to be related by choices made in the different Petrel attribute settings. The threshold parameters of the Petrel fault extraction algorithm seems to have a great effect, especially on the fracture length distribution. To check the robustness of the length distributions found, we assessed their dependence on the Petrel attribute settings. We have tested the influence of the Petrel attribute settings on the non-filtered dataset as well as on a NLM filtered dataset. The NLM filtered dataset chosen is the one filtered using setting parameter settings $N = 3$ and $h = 0.002$, since the fracture length distribution of this dataset shows the best fit to a powerlaw (lowest L_1 value). The first Petrel ant track workflow attribute to be analysed is the fault extraction algorithm. The original choice for the settings of this attribute is discussed in section 3.2.3. The following automatic fault extraction parameters are varied: the connectivity constraint, the extraction sampling and background thresholds combined and the extraction sampling distance and deviation from a plane combined. The last two sets of parameters are changed simultaneously, since their influences are strongly dependent on one another (see section 3.2.3). Figure 15 shows the effect of changing these attribute parameters, on the NLM ($S = 3$, $N = 3$ and $h = 0.002$) filtered dataset. Figure 16 shows this effect on the non-filtered data set. Both figures show a wide range in the length exponent variability. The length exponent a for the non-filtered dataset varies between 2.39 and 3.58. For the NLM filtered dataset the length exponent a varies between 2.73 and 4.88. The range of length exponents of the NLM filtered dataset, affected by fault extraction parameters still shows overall higher a values as the range of length exponents of the non-filtered data set. The range of length exponents as a result of different fault extraction parameters is larger than the range of length exponents computed through varying the NLM filter parameters. This suggests that the fracture network characteristic of fracture length distribution is affected more by the settings of the fault extraction algorithm than by the NLM filter.

An attempt is made to carry out a similar analysis of the influence of the ant track parameters on the fracture length exponent a. An independent analysis appears to be impossible however, due to the non-linearity of the ant track algorithm. Applying slight chances to the input parameters of the algorithm, results in non proportional changes in the output values. This is the result of the algorithm being self enhancing. The value of a planar discontinuity structure is dependent on the amount of ants that have tracked the structure. Once one ant has found this connecting path, many ants will follow, increasing the value of the discontinuity. The certainty of a planar discontinuity being a fault is therefore not only dependent on how well connected the structure is, but also on how many ants are placed on the volume in the first place and thus how many ants can walk along the same track. Figure 17 illustrates this concept. Figure 17a shows the ant track and fault extraction for normal aggressive ant track settings (initial ant boundary of 5 voxels) and 17b shows the ant track and fault extraction for the same settings, however for an initial ant boundary value of 10 voxels. This means less ants are placed on the volume at further distance from each other, resulting in less ants walking along the same paths and therefore lower ant track values. However the certainty of the discontinuities found to be faults in figure 17b is not necessarily less than the ones found in 17a. This effect is then enhanced by the fault extraction parameters Extraction sampling threshold and extraction background threshold, which are based on the scale bar in 17c. As discussed in section 3.2.2 this scale bar is set according to the minimum and maximum values the ant track algorithm can produce, not the minimum and maximum values actually found in the volume. For this reason, it is not possible to make an independent analysis of the influence of the ant track parameters on the fracture network characteristics. To get a realistic fault extraction pattern, the fault extraction parameters have to be adapted along with the ant track volume, to avoid results as in figure 17b.

Another effect of the Petrel ant track workflow on the fracture network characteristics is related to censoring. The fault extraction algorithm of Petrel does not allow the extraction of faults touching the system boundary. Due to this implementation in the algorithm no estimation can be made on the faults affected by censoring. This results in the fact that no censoring correction method could be applied to the data. The slightly higher values estimated for a, compared to the range indicated by [Bonnet et al., 2001], might thus be explained by the absence of a censoring correction.

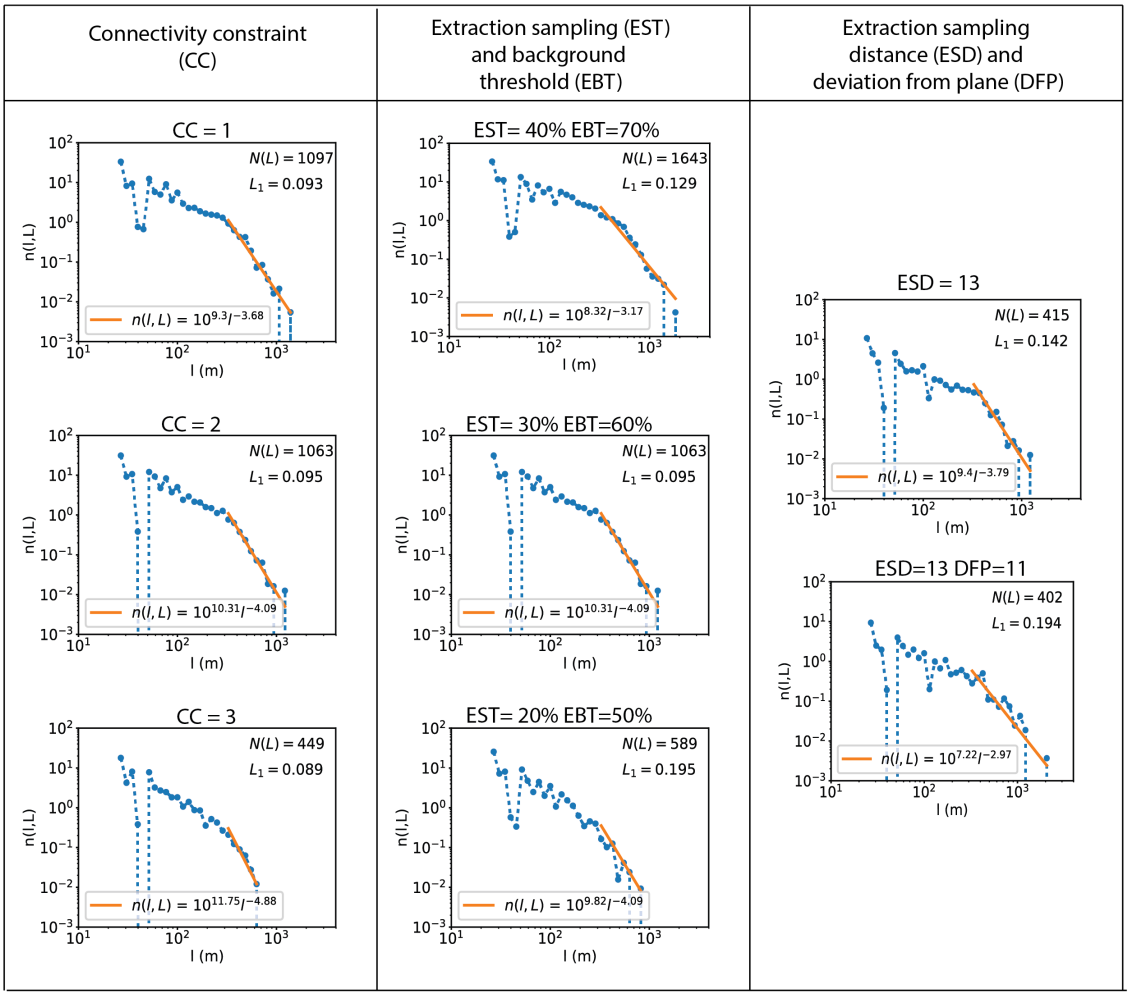


Figure 15: Varying fault extraction parameters on the cropped 3D seismic volume filtered using NLM parameters S=3, N=3 and h=0.002. Fault extraction parameters which are varied are: the connectivity constraint (cc), the extraction sampling and background threshold (EST and EBT, respectively) and the extraction sampling distance (ESD) and deviation from a plane (DFP). The original settings used are CC=2 and EST=30% and EBT=60%.

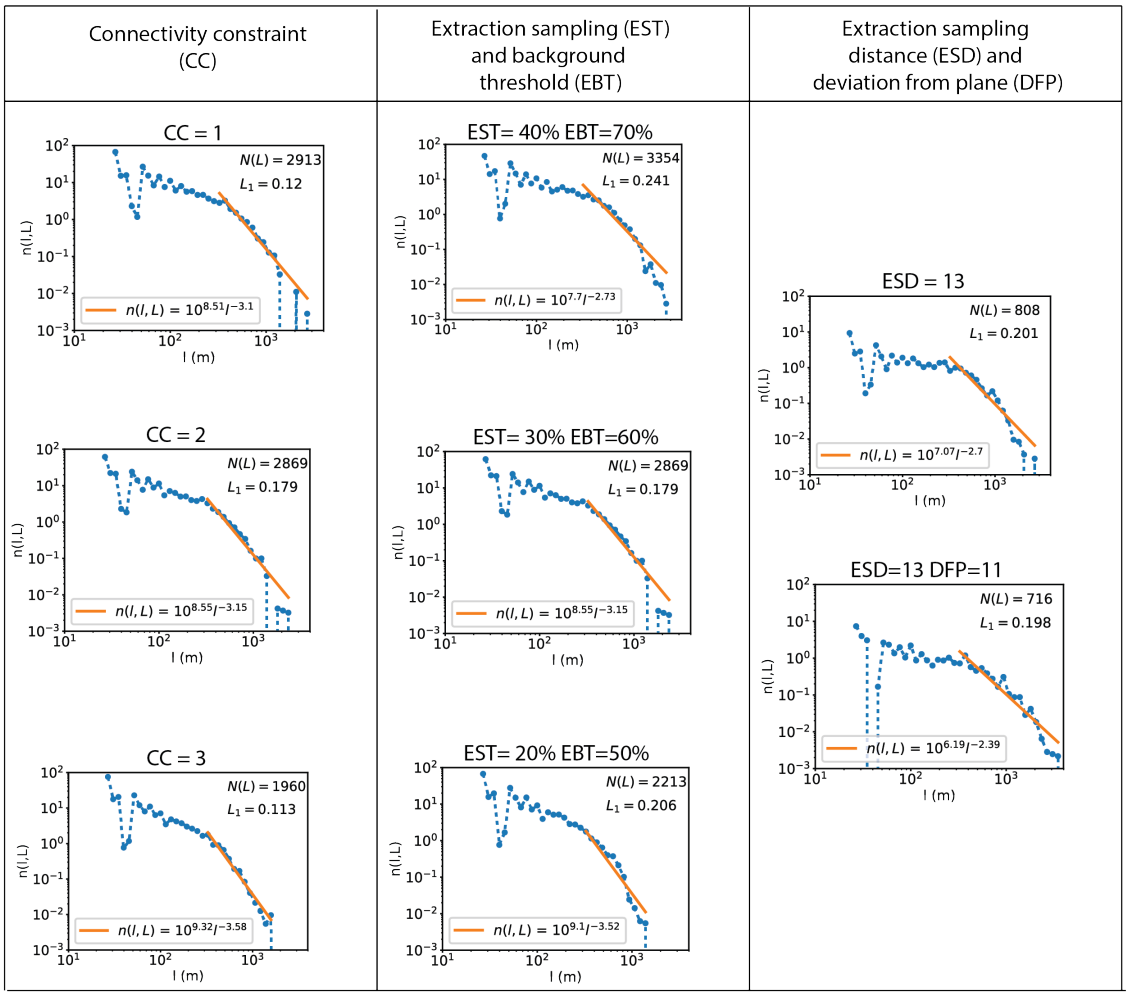


Figure 16: Varying fault extraction parameters on the non-filtered cropped 3D seismic volume. Fault extraction parameters which are varied are: the connectivity constraint (cc), the extraction sampling and background threshold (EST and EBT, respectively) and the extraction sampling distance (ESD) and deviation from a plane (DFP). The original settings used are CC=2 and EST=30% and EBT=60%.

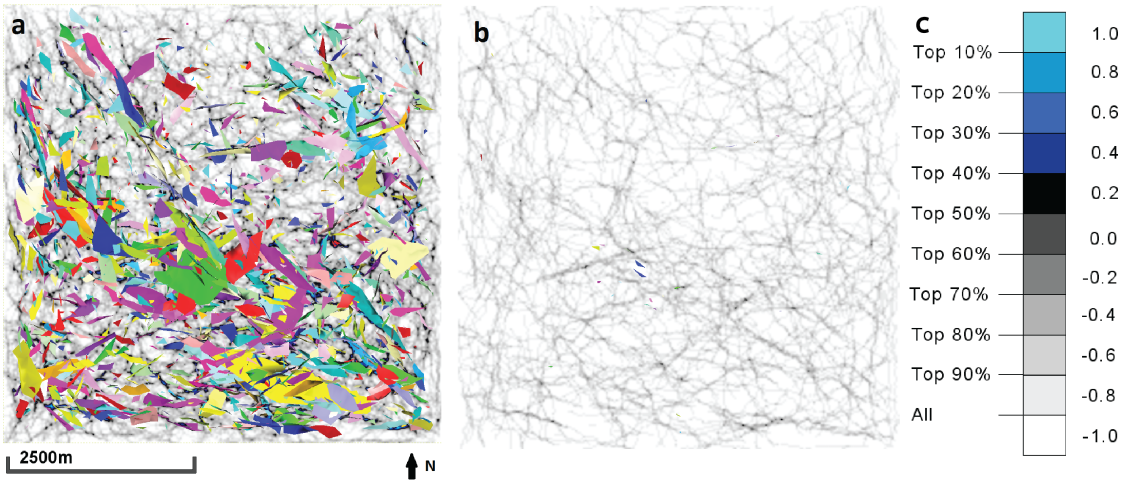


Figure 17: The difference between fault extraction of non-filtered L3NAM1987F minicube with (a) aggressive ant track parameters (initial ant boundary of 5) and (b) aggressive ant track parameters accept an initial ant boundary of 10. The same fault extraction parameters were used. The corresponding ant track time slice at 3124 ms is shown in the background.

The ant track workflow attributes in Petrel highly affect the characteristics of our fracture networks. This implies a great uncertainty on the validity of the fractures extracted from the Petrel ant track workflow. Therefore, a small study is carried out to look for alternative automatic fault detection and extraction algorithms. OpendTect is another software which uses the thin fault likelihood algorithm [Hale, 2012] to detect and extract faults. This algorithm seems to be more linear than the ant track algorithm, which could result in more robust fault patterns. However, a great disadvantage of the fault extraction algorithm used in OpendTect is that the user has to define the maximum number of faults to be extracted. The algorithm will aim to extract exactly this amount of faults, which sometimes results in duplication of already detected faults. Since the number of faults in the data is generally not known, this is a great disadvantage that will have a large impact on the statistics of the fracture network. It would be interesting to compare the patterns extracted using different detection and extraction algorithms, however, this goes beyond the scope of this study.

5.2 Seismic resolution

Apart from the limitations the Petrel ant track workflow inflict on the fracture network analysis, the largest obstacle of this study remains the seismic resolution of the data itself, as resolution decreases drastically with increasing depth. We have aimed to take these limitations into account by the choice in fault extraction settings (section 3.2.3), however, considering the quality of the dataset this may not be sufficient. A quality check of the validity of the extracted fractures is a necessary step, which could not be carried out due to the low seismic resolution of this particular dataset. Further research is needed to apply the full workflow, presented in this study, on a synthetic dataset of which the fracture characteristics are known.

6 Conclusions and recommendations

Due to recent developments in reprocessing of 3D seismic data and fault extraction algorithms, an attempt is made to determine fracture network characteristics from 3D reprocessed seismic data. The non local means (NLM) denoising algorithm is applied to a cropped 3D seismic volume encapsulating part of the Dinantian Friesland platform, a potential reservoir for ultra deep geothermal (UDG) energy. The Petrel ant track workflow is used as sample method, for automatic fault detection and extraction. A python code was build to compute the fracture network characteristics of the faults extracted from the reprocessed seismic data. From analysis of these fracture network characteristics the following conclusions were drawn:

Fracture length distributions of both filtered and non-filtered data show powerlaw behavior above the truncation cut-off. The powerlaw exponent a of the fracture length distributions from NLM filtered data varies depending on the NLM filter parameters used, but remains in the following range: $3.59 \leq a \leq 4.57$. The NLM filter removes small discontinuities that fall below the truncation threshold, but which are interpreted as small fractures in the non-filtered data. This dataset does not allow us to determine whether these small discontinuities are in fact small faults which are being removed by the filter or noise. We therefore recommend to apply this workflow on a synthetic seismic dataset to determine if the Petrel ant track workflow can distinguish effectively between noise and faults. This recommendation is supported also by the detection that the fracture length distribution is more sensitive to the Petrel attribute settings than to the NLM filter parameters.

The sampling method applied in this study is not suitable to detect fractal spatial behavior, due to the resolution limitations of 3D seismic data. However, since the requirement for stereological rules is reached ($a_{3D} \geq 2$), translations can be made between sampling methods in 1D, 2D and 3D. We therefore recommend to apply spatial distribution analysis on FMI of borehole data to determine the fractal behavior of the fractures, since the resolution of scanline data is sufficiently higher than seismic data.

7 Notations

l	fracture length
L	Typical system (sample window) size
L^2	area of sample window fracture network
L^3	volume of sample window fracture network
$n(l, L)$	density length distribution
$N(L)$	number of fractures in system of size L
$N(l, L)$	frequency length distribution
$C(l, L)$	cumulative length distribution
D	fractal dimension
a	length exponent of density length distribution
$1, 2, 3D$	euclidean dimension
N	Non Local Mean parameter: radius in samples of the Neighbourhood window
S	Non Local Mean parameter: radius in samples of the Search window
h	Non Local Mean parameter filter parameter
L_1	first norm or sum of absolute difference
$\hat{L}_1$	normalized first norm or sum of absolute difference

8 Acknowledgements

This study was carried within the framework of the IMAGE (Integrated Methods for Advanced Geothermal Exploration) project at TNO Utrecht. I would like to thank my supervisors at TNO, Jan-Diederik van Wees and Hans Veldkamp for their input and help, Stefan Carpentier for sharing his work on the non local means algorithm and Mark Vrijlandt for his help and suggestions on the coding in Python. Lastly, thanks to Fred Beekman for input and support as my supervisor from Utrecht University.

9 Appendix A: Influence of NLM on the seismic volume

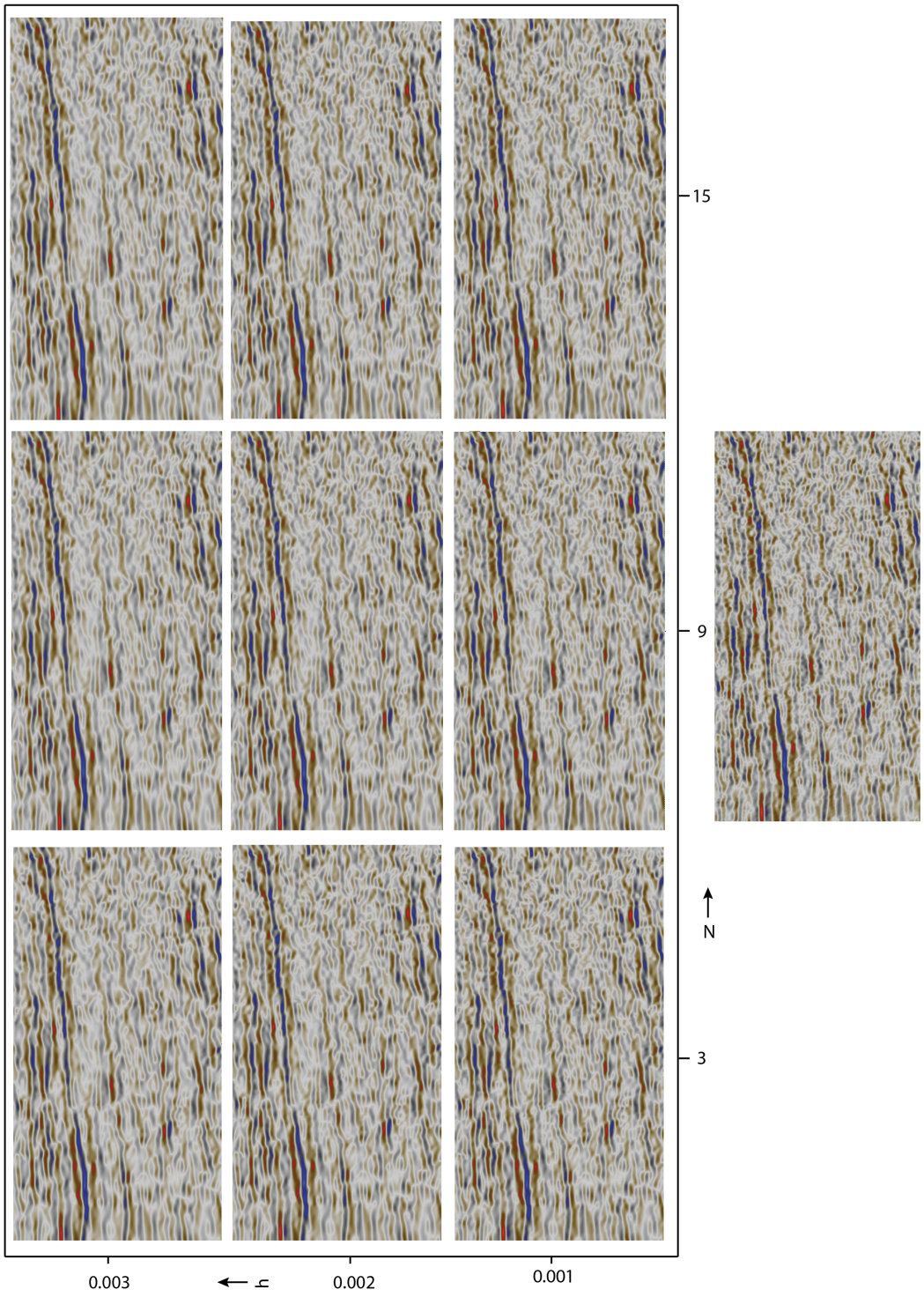


Figure 18: Influence of NLM parameters on the cropped 3D seismic cube. Inline 1536 of the minicube is displayed for the differently filtered cubes.(a) NLM filtered cube, (b) non-filtered cube.

10 Appendix B: Influence of NLM on the variance volume

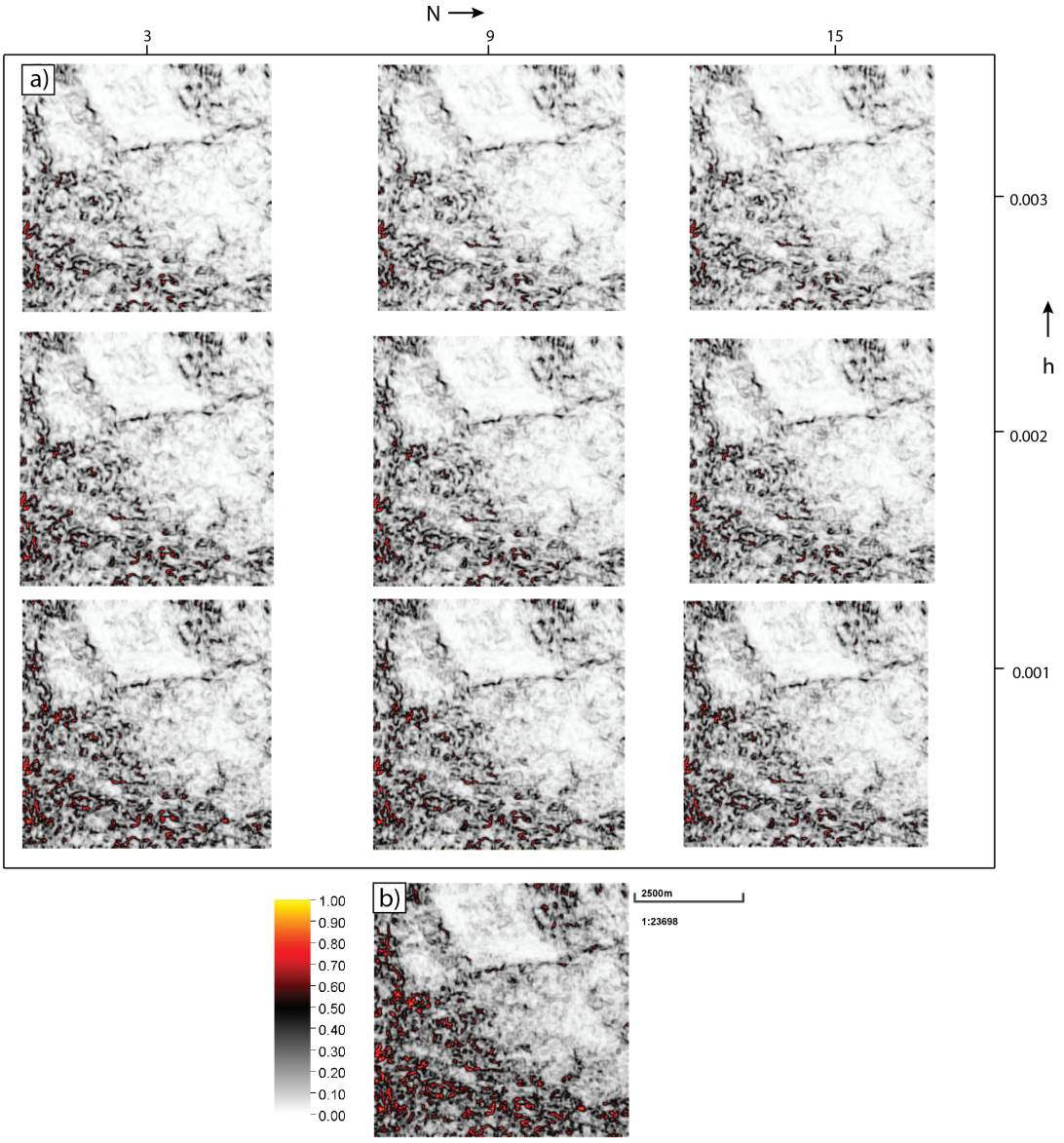


Figure 19: Influence of NLM parameters on the cropped variance cube. Timeslice at 2968 ms of the variance minicube is displayed for the differently filtered cubes. (a) NLM filtered cube, (b) non-filtered cube.

11 Appendix C: Influence of NLM on the ant track volume

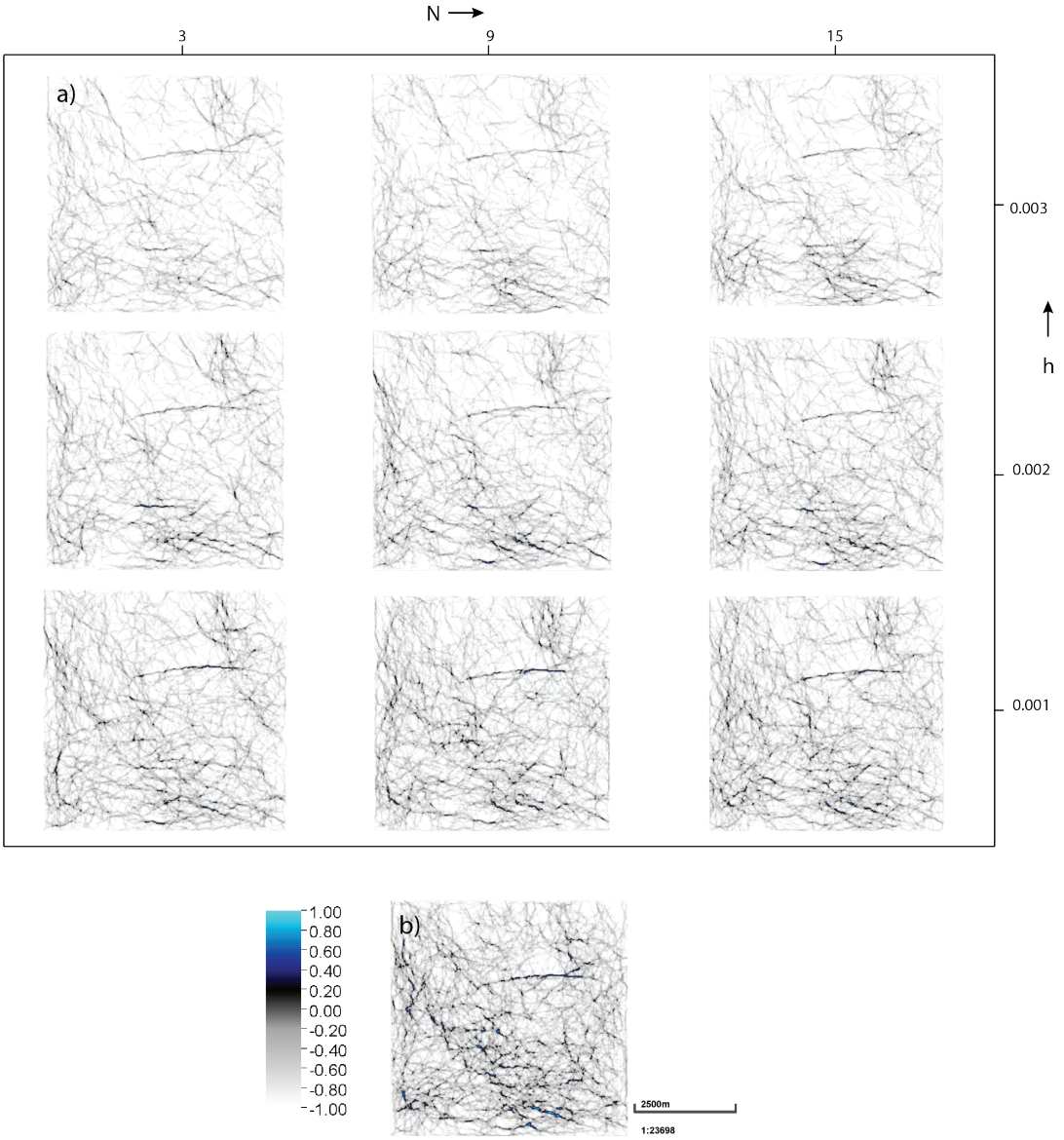


Figure 20: Influence of NLM parameters on the cropped ant track volume. Timeslice at 2968 ms of the ant track minicube is displayed for the differently filtered cubes. (a) NLM filtered cube, (b) non-filtered cube.

12 Appendix D: Influence of NLM on the fault extraction

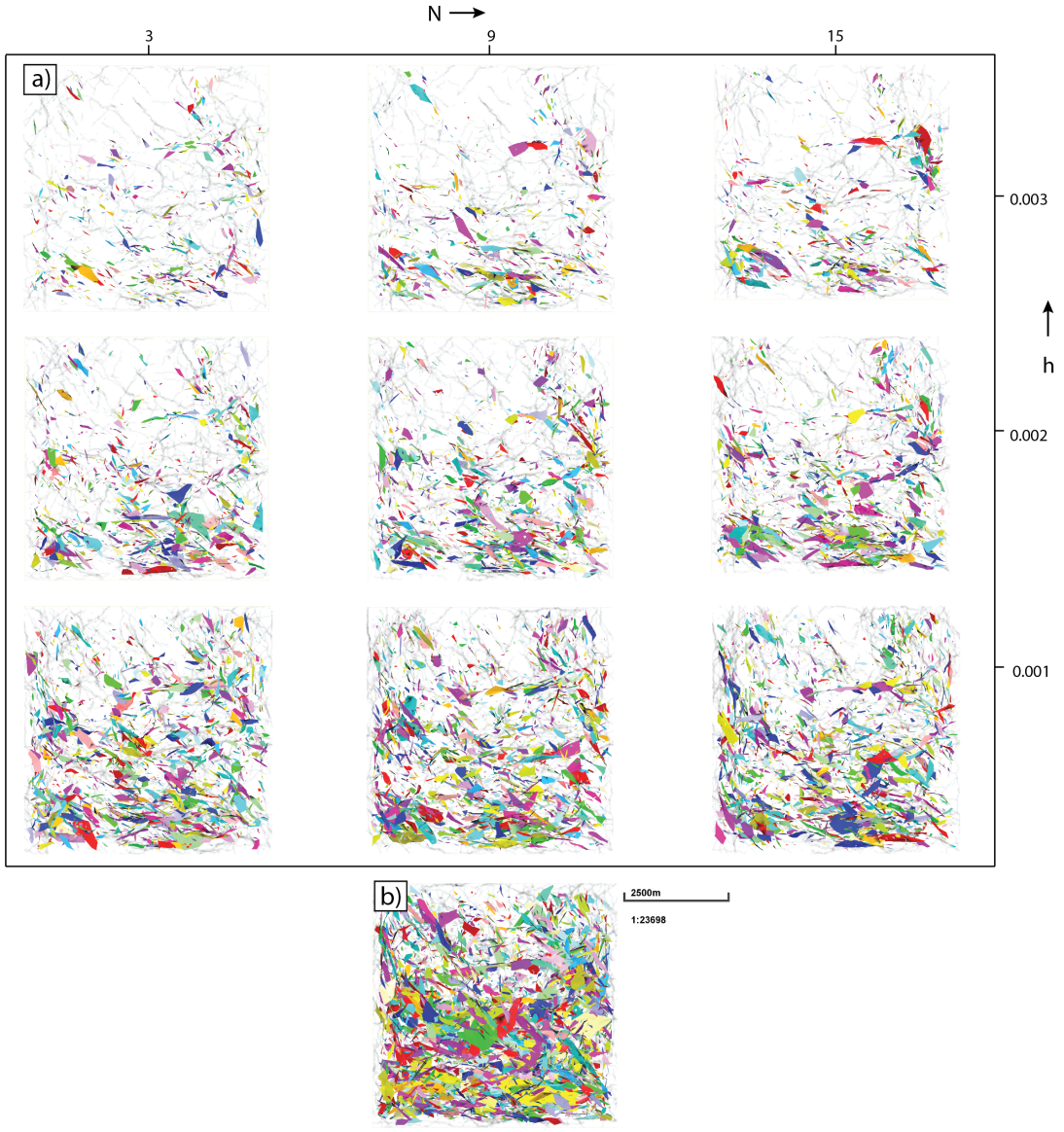


Figure 21: Influence of NLM parameters on the faults extracted by the automatic fault extraction algorithm. Timeslice at 2968 ms of the ant track minicube is displayed in the background. (a) NLM filtered cube, (b) non-filtered cube.

References

[Alstott et al., 2014] Alstott, J., Bullmore, E., & Plenz, D. (2014). Powerlaw: A python package for analysis of heavy-tailed distributions. *PLoS ONE*, 9(1), 1–18.

[Boadu & Long, 1994] Boadu, F. K. & Long, L. T. (1994). The fractal character of fracture spacing and RQD. *International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts*, 00(0), 1–8.

[Bonar & Sacchi, 2012] Bonar, D. & Sacchi, M. (2012). Denoising seismic data using the nonlocal means algorithm. *Geophysics*, 77(1), A5–A8.

[Bonnet et al., 2001] Bonnet, E., Bour, O., Odling, N. E., Davy, P., Main, I., Cowie, P., & Berkowitz, B. (2001). Scaling of fracture systems in geolocial media. *Reviews of Geophysics*, 39(3), 347–383.

[Bonté et al., 2012] Bonté, D., van Wees, J.-D., & Verweij, J. (2012). Subsurface temperature of the onshore Netherlands: new temperature dataset and modelling. *Netherlands Journal of Geosciences / Geologie en Mijnbouw*, 91(4), 491–515.

[Bour et al., 2002] Bour, O., Davy, P., & Darcel, C. (2002). A statistical scaling model for fracture network geometry, with validation on a multiscale mapping of a joint network (Hornelen Basin, Norway). *Journal of Geophysical Research*, 107(B6), 2113.

[Boxem et al., 2016] Boxem, T., Veldkamp, J., & van Wees, J. (2016). *Ultra-diepe geothermie: Overzicht, inzicht & to-do ondergrond*. Technical report.

[Buades et al., 2005] Buades, A., Coll, B., & Morel, J.-M. (2005). A review of image denoising algorithms, with a new one. *SIAM Journal on Multiscale Modeling and Simulation: A SIAM Interdisci-plinary Journal*, 4(2), 490–530.

[Buades et al., 2011] Buades, A., Coll, B., & Morel, J.-M. (2011). Non-Local Means Denoising. *Image Processing On Line*, 1, 208–212.

[Canales, 1984] Canales, L. L. (1984). Random Noise Reduction. *Extended Abstracts SEG meeting Atlanta*.

[Carpentier & Steeghs, 2016] Carpentier, S. & Steeghs, P. (2016). Final Report Work Package: New active seismic processing techniques developed. *EU Project IMAGE*, 7.1.

[Clauset et al., 2009] Clauset, A., Shalizi, C. R., & Newman, M. (2009). Power-law distributions in empirical data. *SIAM Review 51*, (pp. 661–703).

[Collins et al., 2014] Collins, J. F., Katz, D., Harris, P. M. M., & Narr, W. (2014). Burial Cementation and Dissolution in Carboniferous Slope Facies , Tengiz Field , Kazakhstan : Evidence for Hydrothermal Activity. *AAPG 2013 Annual Convention and Exhibition, Pittsburgh, Pennsylvania*.

[Darcel et al., 2003a] Darcel, C., Bour, O., & Davy, P. (2003a). Stereological analysis of fractal fracture networks. *Journal of Geophysical Research*, 108, 1–14.

[Darcel et al., 2003b] Darcel, C., Bour, O., Davy, P., & Dreuzy, J. R. D. (2003b). Connectivity properties of two-dimensional fracture networks with stochastic fractal correlation. *Water Resources Research*, 39(10), 1–13.

[Davy, 1993] Davy, P. (1993). On the Frequency-Length Distribution of the San Andreas Fault System. *Journal of Geophysical Research*, 98(l), 12,141 – 12,151.

[Davy et al., 2010] Davy, P., Le Goc, R., Darcel, C., Bour, O., de Dreuzy, J.-r., & Munier, R. (2010). A likely universal model of fracture scaling and its consequence for crustal hydromechanics. *Journal of Geophysical Research*, 115.

[Davy et al., 1990] Davy, P., Sornette, A., & Sornette, D. (1990). Some consequences of a proposed fractal nature of continental faulting. *Nature*, 348(6296), 56–58.

[Einstein & Baecher, 1983] Einstein, H. H. & Baecher, G. B. (1983). Probabilistic and statistical methods in engineering geology. *Rock Mechanics and Rock Engineering*, 16, 39–72.

[EZ, 2016] EZ (2016). *Energierapport Transitie naar duurzaam.* Technical report.

[Fang et al., 2017] Fang, J., Zhou, F., & Tang, Z. (2017). Discrete Fracture Network Modelling in a Naturally Fractured Carbonate Reservoir in the Jingbei Oilfield, China. *Energies*, 10(2), 183.

[Goldscheider et al., 2010] Goldscheider, N., Mádl-Szőnyi, J., Erőss, A., & Schill, E. (2010). Review: Thermal water resources in carbonate rock aquifers. *Hydrogeology Journal*, 18(6), 1303–1318.

[Hale, 2012] Hale, D. (2012). Fault surfaces and fault throws from 3D seismic images. *CWP-721*, (pp. 1–6).

[Häring et al., 2008] Häring, M. O., Schanz, U., Ladner, F., & Dyer, B. C. (2008). Characterisation of the Basel 1 enhanced geothermal system. *Geothermics*, 37(5), 469–495.

[Hentschel & Procaccia, 1983] Hentschel, H. G. E. & Procaccia, I. (1983). The infinite number of generalized dimensions of fractals and strange attractors. *Physica D: Nonlinear Phenomena*, 8(3), 435–444.

[Jones & Levy, 1987] Jones, I. F. & Levy, S. (1987). Signal-to-noise ratio enhancement in multichannel seismic data via the Karhunen-Loéve transform. *Geophysical Prospecting*, 35(1), 12–32.

[Laslett, 1982] Laslett, G. M. (1982). Censoring and edge effects in areal and line transect sampling of rock joint traces. *Journal of the International Association for Mathematical Geology*, 14(2), 125–140.

[Lei, 2016] Lei, Q. (2016). *Characterisation and modelling of natural fracture networks: geometry, geomechanics and fluid flow.* PhD thesis, Imperial College London.

[Lei et al., 2015] Lei, Q., Latham, J.-p., Tsang, C.-f., Xiang, J., & Lang, P. (2015). A new approach to upscaling fracture network models while preserving geostatistical and geomechanical characteristics. *Journal of Geophysical Research: Sollid Earth*, 120.

[Lipsey et al., 2016] Lipsey, L., Pluymaekers, M., Goldberg, T., van Oversteeg, K., Ghazaryan, L., Cloetingh, S., & van Wees, J. D. (2016). Numerical modelling of thermal convection in the Luttelgeest carbonate platform, the Netherlands. *Geothermics*, 64(3), 135–151.

[Liu & He, 2012] Liu, W. & He, Y. (2012). Seismic geometric attributes and information integration techniques for interpretation of small faults. *SEG Technical Program Expanded Abstracts 2012*, (pp. 1–5).

[Mandelbrot, 1983] Mandelbrot, B. B. (1983). *The Fractal Geometry of Nature.* New York: W.H. Freeman and Company.

[Marfurt et al., 1998] Marfurt, K. J., Kirlin, R. L., Farmer, S. L., & Bahorich, M. S. (1998). 3D seismic attributes using a semblance based coherency algorithm. *Geophysics*, 63(4), 1150–1165.

[Odling, 1997] Odling, N. E. (1997). Scaling and connectivity of joint systems in sandstones from western Norway. *Journal of Structural Geology*, 19(10), 1257–1271.

[Pedersen et al., 2005] Pedersen, S. I., Skov, T., Randen, T., & Sønneland, L. (2005). Automatic fault Extraction Using Artificial Ants. In A. Iske & T. Randen (Eds.), *Mathematical Methods and Modelling in Hydrocarbon Exploration and Production* chapter 3, (pp. 107–116). Springer-Verlag Berlin Heidelberg.

[Petrel, 2015] Petrel (2015). Recommended Seismic Volume Attributes.

[Pickering et al., 1995] Pickering, G., Bull, J. M., & Sanderson, D. J. (1995). Sampling power-law distributions. *Tectonophysics*, 248, 1–20.

[Schlumberger, 2002] Schlumberger (2002). Borehole geology, geomechanics and 3D reservoir modeling.

[Schlumberger, 2016] Schlumberger (2016). Petrel Help Center.

[Silva et al., 2005] Silva, C. C., Marcolino, C. S., & Lima, F. D. (2005). Automatic Fault Extraction Using Ant Tracking Algorithm in the Marlim South Field, Campos Basin. *SEG/Houston Annual Meeting*, (pp. 857–861).

[Stolt & Benson, 1986] Stolt, R. H. & Benson, A. K. (1986). Seismic migration: Theory and practice. In *Handbook of geophysical exploration*, volume 5. Geophysical Press.

[Taner et al., 1979] Taner, M. T., Koehler, F., & Sheriff, R. E. (1979). Complex seismic trace analysis. *Geophysics*, 44(6), 1041–1063.

[Valley et al., 2016] Valley, B., Moein, M., & Evans, K. F. (2016). Physical linkage between power-law scaling relations for fractures inferred for observed stress heterogeneity, b-value of induced seismicity and fracture scaling using data-constrained models. *IMAGE*, D6.03, 20–51.

[Vicsek, 1992] Vicsek, T. (1992). Fractal Growth Phenomena. In *Fractal Growth Phenomena* (pp. 558). Singapore: World Scientific.

[Walsh et al., 2003] Walsh, J. J., Bailey, W. R., Childs, C., Nicol, A., & Bonson, C. G. (2003). Formation of segmented normal faults: A 3-D perspective. *Journal of Structural Geology*, 25(8), 1251–1262.

[Yilmaz, 2001] Yilmaz, O. (2001). *Seismic Data Analysis Processing, Inversion and Interpretation of Seismic Data*, volume II. Tulsa: Society of Exploration Geophysicists, 2 edition.

[Zeeb et al., 2013] Zeeb, C., Gomez-Rivas, E., Bons, P. D., Virgo, S., & Blum, P. (2013). Fracture network evaluation program (FraNEP): A software for analyzing 2D fracture trace-line maps. *Computers and Geosciences*, 60, 11–22.

[Zhao & Sun, 2013] Zhao, J. & Sun, S. Z. (2013). Automatic fault extraction using a modified ant-colony algorithm. *Journal of Geophysics and Engineering*, 10, 1–7.