Complex fault system revealed from 3-D seismic reflection data with deep learning and fault network analysis

Thilo Wrona[1,2*], Indranil Pan[3,4,5], Rebecca E. Bell[6], Christopher A-L. Jackson[7], Robert L. Gawthorpe[1], Haakon Fossen[8], Edoseghe E. Osagiede[1], and Sascha Brune[2,9]

[1]Department of Earth Science, University of Bergen, Allégaten 41, N-5007 Bergen, Norway.
[2]GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany.
[3]Centre for Process Systems Engineering & Centre for Environmental Policy, Imperial College London, UK.
[4]The Alan Turing Institute, British Library, London, UK.
[5]School of Mathematics, Statistics & Physics, Newcastle University, UK.
[6]Basins Research Group (BRG), Department of Earth Science and Engineering, Imperial College, Prince Consort Road, London, SW7 2BP, UK.
[7]Department of Earth and Environmental Sciences, University of Manchester, Manchester, UK
[8]Museum of Natural History, University of Bergen, Allégaten 41, N-5007 Bergen, Norway.
[9]Institute of Geosciences, University of Potsdam, Potsdam-Golm, Germany.

*wrona@gfz-potsdam.de

ABSTRACT

Understanding where normal faults are is critical to an accurate assessment of seismic hazard, the successful exploration for and production of natural (including low-carbon) resources, and for the safe subsurface storage of CO_2. Our current knowledge of normal fault systems is largely derived from seismic reflection data imaging intra-continental rifts and continental margins. However, exploitation of these data is limited by interpretation biases, data coverage and resolution, restricting our understanding of fault systems. Applying supervised deep learning to one of the largest offshore 3-D seismic reflection data sets from the northern North Sea allows us to image the complexity of the rift-related fault system. The derived fault score volume allows us to extract almost 8000 individual normal faults of different geometries, which together form an intricate network characterised by a multitude of splays, junctions and intersections. Combining tools from deep learning, computer vision and network analysis allows us to map and analyse the fault system in great detail and a fraction of the time required by conventional interpretation methods. As such, this study shows how we can efficiently identify and analyse fault systems in increasingly large 3-D seismic data sets.

36

1. INTRODUCTION

Understanding the geometry and growth of normal fault systems is critical when assessing

seismic hazard, when identifying suitable sites for subsurface CO_2 storage and when exploring for

natural resources (traditional and low-carbon). For example, whereas probabilistic seismic hazard

analyses based on seismic event catalogues are extremely useful when trying to forecast earthquake

likelihood and location, high-resolution fault mapping, preferably in 3-D, can help us constrain the

slip tendency of faults, where seismic catalogues are discontinuous and/or incomplete (e.g. Morris et

al., 1996; Moeck et al., 2009; Yukutake et al., 2015). Moreover, faults can facilitate (or impede) fluid

and gas migration to the Earth's surface, thus determining their geometry and connectivity, as well as

their hydraulic properties is key for assessing their role in the long-term subsurface storage of CO_2

(Bissell et al., 2011; Kampman et al., 2014). In both of these examples, we need robust predictions of

3-D fault geometry over large areas and across a wide range of scales (100s m to 100 km).

Accurately mapping fault systems in 2-D and 3-D seismic reflection data typically requires

expertise and time (e.g. Bond, 2015). While we can map fault systems in great detail over small areas

using 3-D seismic reflection data (e.g. Lohr et al., 2008; Wrona et al., 2017; Claringbould et al.,

2020), we lack an understanding of the character of 3-D fault populations at the scale of entire rift

systems, as regional studies are often limited to only sparse, 2-D seismic sections (e.g. Clerc et al.,

2015; Fazlikhani et al., 2017; Phillips et al., 2019). 3-D numerical models are now capable of

simulating fault networks at the rift scale; however, there are few observational data sets of the same

scale to test the predictions of these models and, therefore, help refine them (e.g. Naliboff et al., 2020;

Pan et al., 2021).

Supervised deep learning allows us to map faults in seismic reflection data (e.g. Wu et al.,

2019; Mosser et al., 2020; Wrona et al., 2021a), but up until now many of these studies have laid the

foundation by focusing on detecting faults rather than studying the geometry of these faults. In this

study, by applying supervised deep learning to newly-acquired broadband 3-D seismic reflection data

imaging much of the northern North Sea rift (161 km wide in E-W, 266 km long area in N-S, 0-20 km

deep), we map the fault network associated with a continental rift basin at an unprecedented level of

detail. Using manually labelled data (<0.1% of data volume), we train a deep convolutional neural

network (U-Net) to predict faults in our data set. The predicted score ranges from 0 (no fault) to 1

(fault). Based on this score across the entire 3-D seismic volume we employ a second workflow to

extract the normal fault system as a network (a set of nodes and edges) allowing us to investigate the

architecture and growth of this extremely complex system consisting of thousands of intersecting

faults.

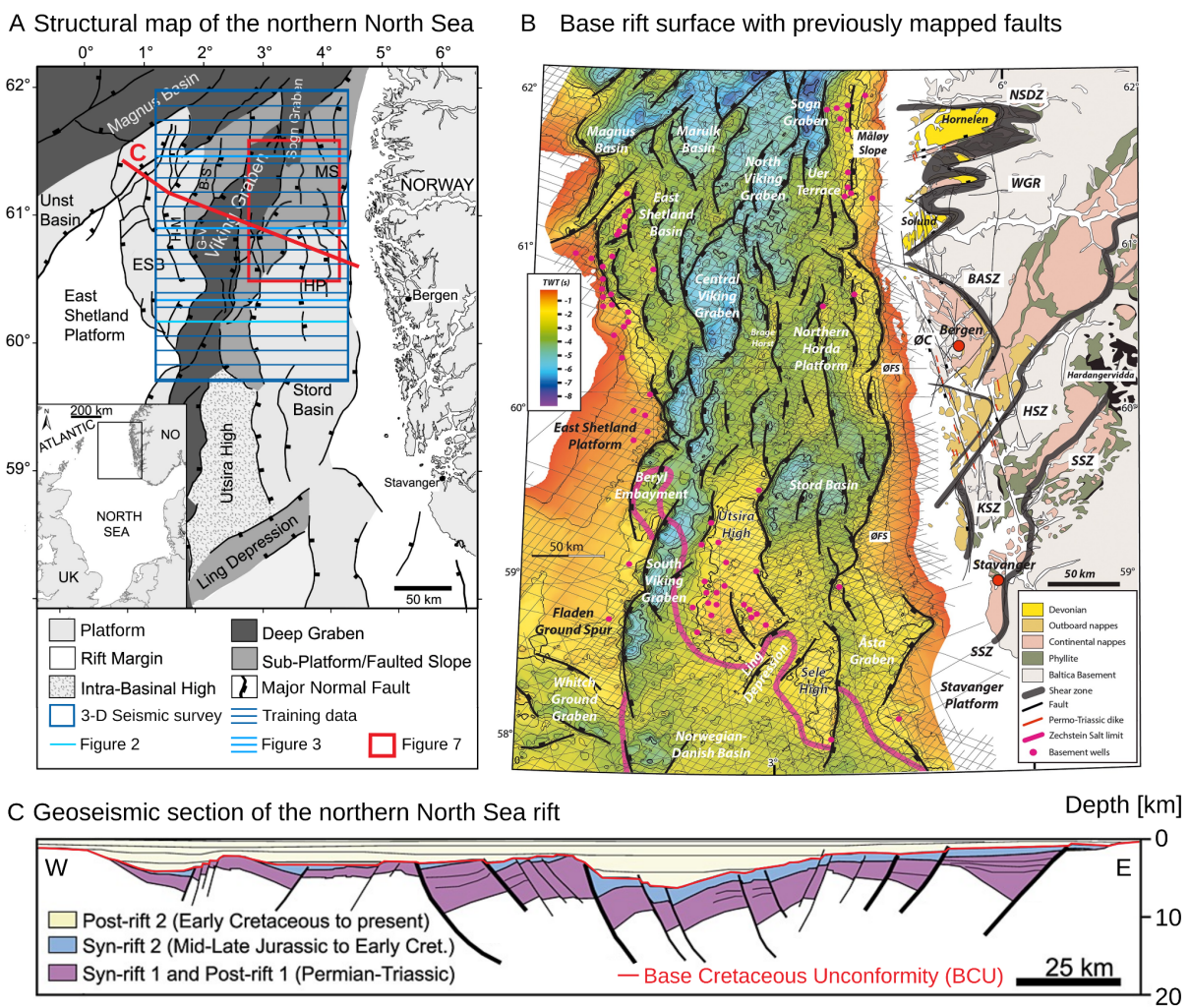



Figure 1: A Structural overview map of the northern North Sea basin system (from Tillmans et al., 2021 after Færseth, 1996). Blue rectangle marks the outline of the seismic survey in this study. ESB = East Shetland Basin, B-S = Brent-Statfjord Fault, G-V = Gullfaks-Visund Fault, MS = Måløy Slope, HP = Horda Platform. **B** The base rift surface (base Permo-Triassic rifting) time-structure map in the northern North Sea rift (from Fazlikhani et al., 2017) and the geology of southwestern Norway, showing the general onshore and offshore structural configuration in the study area. Bold black lines

highlight major rift-related normal faults displacing the base rift surface where all units older than Upper Permian are considered basement. Black lines in the background show some of the 2-D seismic reflection surveys used by Fazlikhani et al. (2017). NSDZ, Nordfjord-Sogn Detachment Zone; BASZ, Bergen Arc Shear Zone; WGR, Western Gneiss Region; ØC, Øygarden Complex (gneiss); ØFS, Øygarden Fault System; HSZ, KSZ, and; SSZ: Hardangerfjord, Karmøy, and Stavanger shear zones, respectively. **C** Regional interpretation of the structure of the northern North Sea after Færseth (1996)**.**

2. GEOLOGICAL SETTING

The study area is located in the northern North Sea (Fig. 1), where continental crust consists of 10–30-km-thick crystalline basement overlain by as much as 12 km of sedimentary strata deposited during, after, and possibly even before periods of rifting in the late Permian–Early Triassic (rift phase 1) and Middle Jurassic–Early Cretaceous (rift phase 2) (e.g. Whipp et al., 2014; Bell et al., 2014; Maystrenko et al., 2017). The extension direction of these two phases has long been debated. Whereas most studies agree that the late Permian–Early Triassic rifting was driven by E-W extension (e.g. Færseth et al., 1997; Torsvik et al., 1997), Middle Jurassic–Early Cretaceous rifting has been associated with both E-W (e.g. Bartholomew et al., 1993; Brun and Tron, 1993) and NW-SE extension (e.g. Færseth, 1996; Doré et al., 1997; Færseth et al., 1997) (Fig. 1B). This debate is further complicated by the fact that some of the largest normal faults on the Horda Platform developed during rift phase 1, but were subsequently reactivated during rift phase 2 (e.g. Whipp et al., 2014; Bell et al., 2014). The crystalline basement underlying the sedimentary strata formed by terrane accretion during the Sveconorwegian (1140–900 Ma) and Caledonian (460–400 Ma) orogenies (Bingen et al., 2008). Several studies argue that this structural template, in particular the ductile shear zones, controlled the location, strike, and overall pattern of rift-related faulting in the overlying sedimentary successions being reactivated as normal faults, or by limiting the along-strike propagation of faults (e.g. Fazlikhani et al., 2017; Phillips et al., 2019; Osagiede et al., 2020; Wiest et al., 2020).

3. DATA & METHODS

3.1. 3-D seismic reflection data

In this study, we use one of the largest offshore 3-D seismic data sets ever acquired, which images a large part of the northern North Sea rift across an area of 35,410 km^2, and with excellent depth-imaging down to 22 km (i.e., the middle-to-lower crust) (Figs. 1, 2A, 3). The data set was

106 acquired using eight, up to 8-km-long streamers that were towed ~40 m below the water surface. The

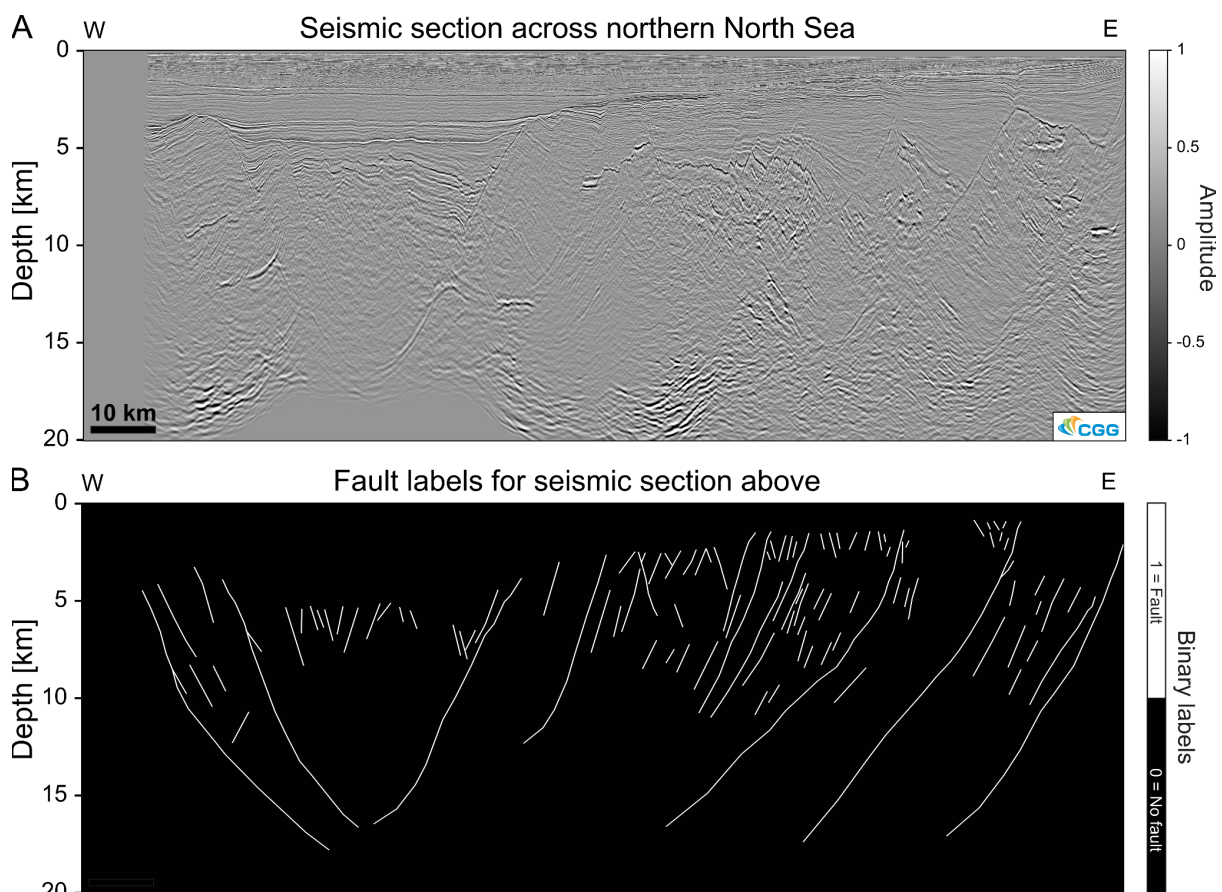
107 broadseis technology used for recording covers a wide range of frequencies (2.5-155 Hz), providing

108 high-resolution depth imaging. The data were binned at 12.5 × 18.75 m, with a vertical sample rate of

109 4 ms. The data was 3-D true amplitude prestack depth-migrated. The seismic volume was zero-phase

110 processed with SEG normal polarity; i.e., a positive reflection (white) corresponds to an acoustic

111 impedance (density × velocity) increase with depth. More details on data acquisition and pre-

112 processing steps are provided by Wrona et al., (2019, 2021a).

113
114 **Figure 2: A** Example seismic section across the northern North Sea. Amplitudes are scaled for
115 machine learning **B** Example of fault interpretation of the section used to train a deep convolutional
116 neural network for fault prediction.

117 3.2. Deep learning

118   Deep learning describes a set of algorithms and models, which learn to perform a specific task

119 (e.g. fault interpretation) on a given data set without explicit feature engineering (e.g. the calculation

120 and calibration of seismic attributes, such as coherence or variance). Deep learning allows the

121 derivation of a fault score volume that highlights normal faults within a 3-D seismic volume. This

122 approach requires that a large number of examples of faults and unfaulted strata are labelled in the

123 training seismic data. We extract 80,000 such examples (2-D squares of 128×128 pixels) from 22

124 interpreted seismic sections oriented perpendicular to the N-trending rift (Figs. 1A, 2). Note that these

125 seismic sections only constitute <0.1% of the entire 3-D seismic volume. Next, we split these

126 examples into three groups; one set for training (80%), one for validation (10%), and one for testing

127 (10%). We use the first of these to train a deep convolutional neural network (U-Net) designed to

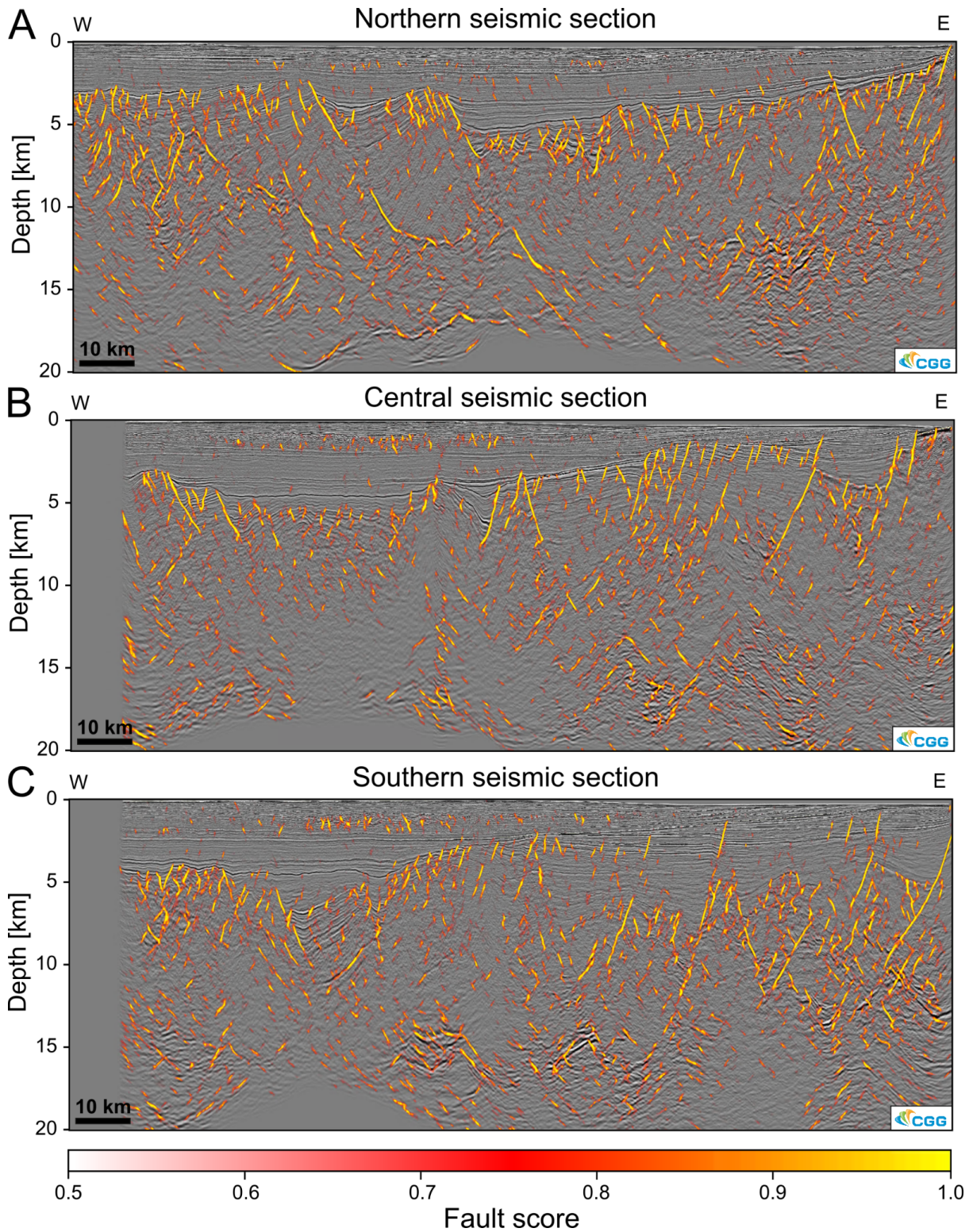
128 perform image segmentation tasks (Ronneberger et al., 2015). Using the validation set, we track the

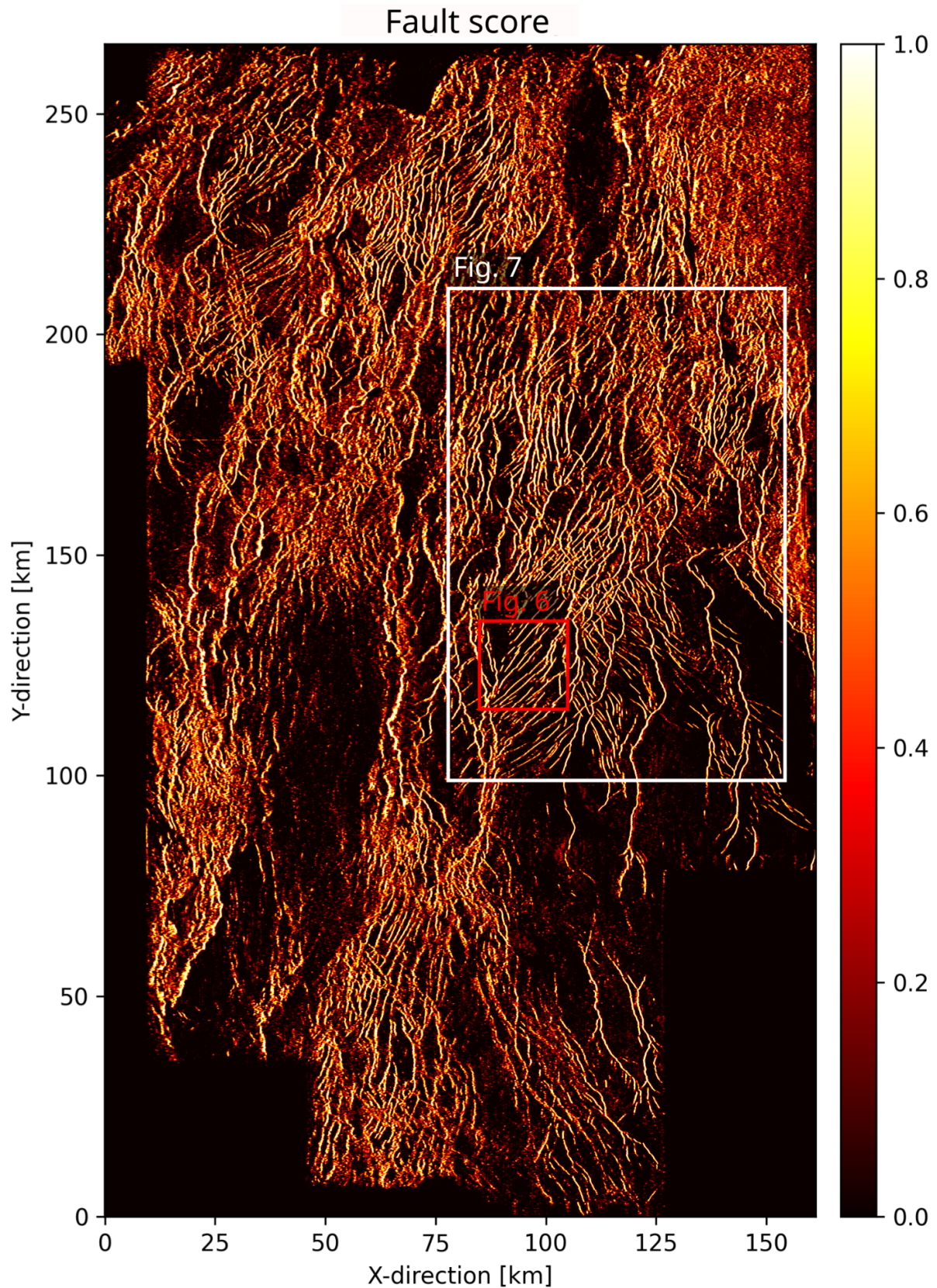
129 accuracy and loss of the model during training and stop once the validation loss does not decrease

130 further, resulting in a final binary accuracy of 0.83 and F1-score of 0.76 (see Wrona et al., 2021a).

131 Finally, we apply the model to the entire 3-D seismic volume to derive a fault score volume (Figs. 3,

132 4), an attribute, which ranges from 0 (no fault) to 1 (fault). All details of the workflow and the code

133 are provided by Wrona et al. (2021b, 2021a).

134
Figure 3: Examples of seismic sections extracted from fault score volume of the 3-D seismic data set.
135
Note that these sections were not part of the training data, but are actually 6.25 km away from the
136
closest interpreted seismic section (see Fig. 1A). To show the correspondence between seismic data
137
and fault score, we needed to define a cutoff value (0.5) below which the fault score becomes
138
transparent and the seismic data becomes visible.
139

140

141
142 **Figure 4:** Surface capturing tectonic faults extracted from fault likelihood volume. The surface was
143 extracted 500 m below the Base Cretaceous Unconformity, where we observe a large number of
144 faults, which were either formed or reactivated in the second rift phase. White rectangle shows the

area used for validation (Fig. 8) and the red rectangle indicates the area where we demonstrate our fault network extraction workflow (Fig. 6). Note that this figure shows to whole range of values of the fault score [0,1].

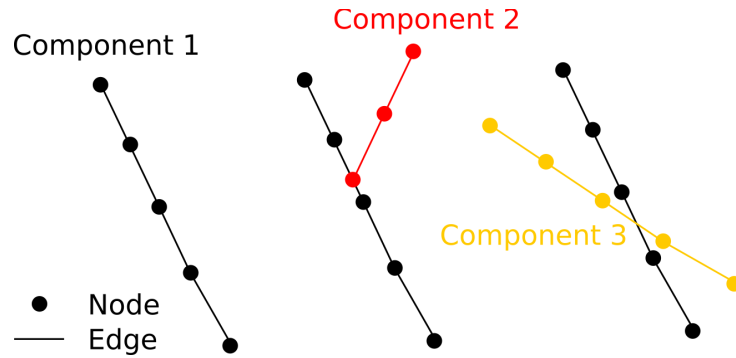
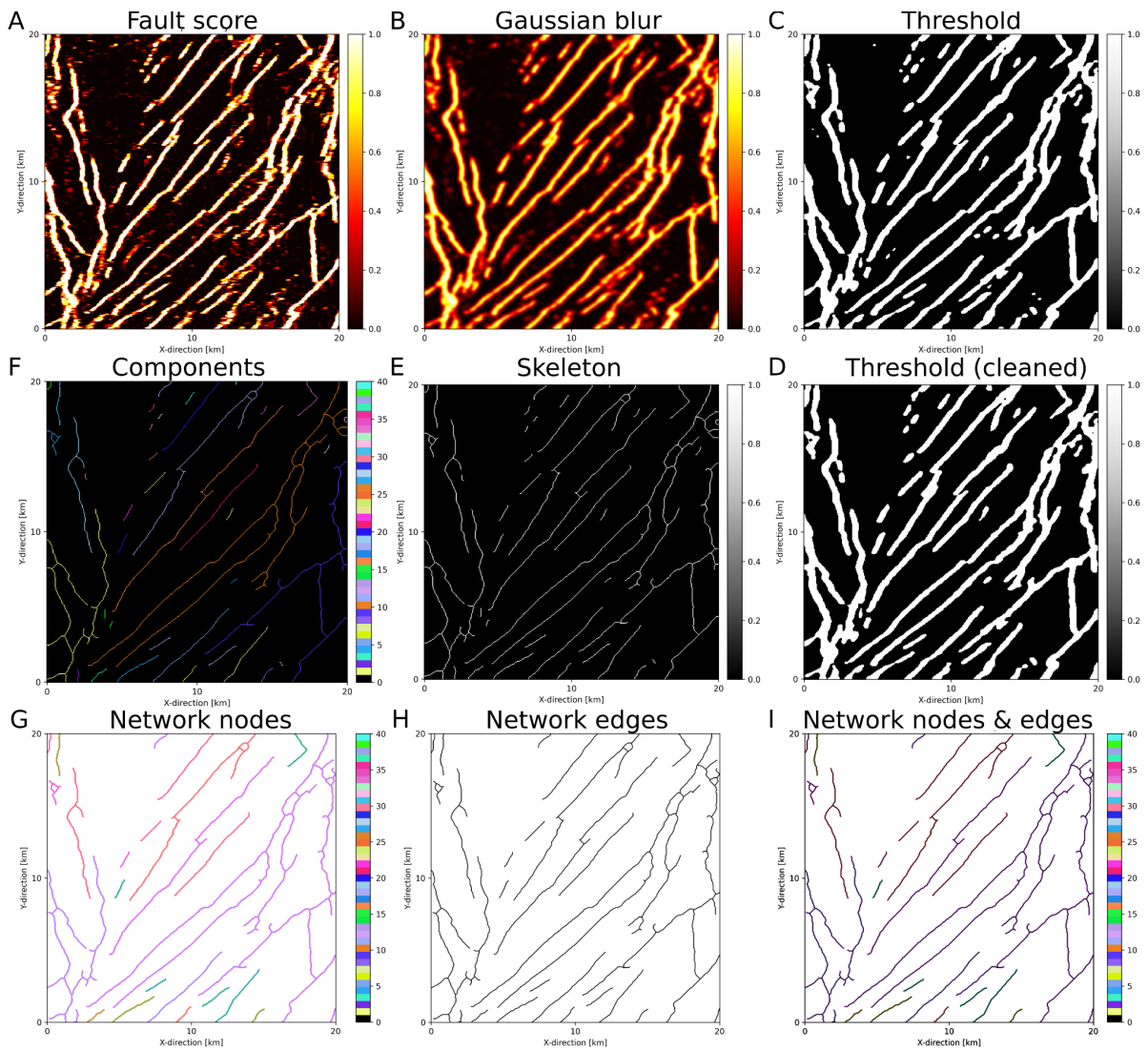



Figure 5: Schematic illustration of fault network (or graph) with nodes, edges and components. Each node marks a location along the fault. Each edge connects two nodes and each (connected) component indicates all nodes connected to one another by edges.

154 **Figure 6:** Fault network extraction workflow showing: **A** Fault score extracted along the surface (500
155 m below BCU). **B** Gaussian Blur filter (σ=2) of surface. **C** Threshold (0.35) of filter. **D** Cleaned
156 threshold where small patches are removed. **E** Skeleton of cleaned threshold. **F** Connected
157 components of skeleton. **G** Network nodes based on components. **H** Network edges based on
158 components. **I** Network nodes and edges combined. Note that colours in F, G and I indicate connected
159 components (i.e. individual faults), before splitting (see Fig. 6).

Fault network editing

A Extracted fault network   **B** Cleaned fault network   **C** Split fault network

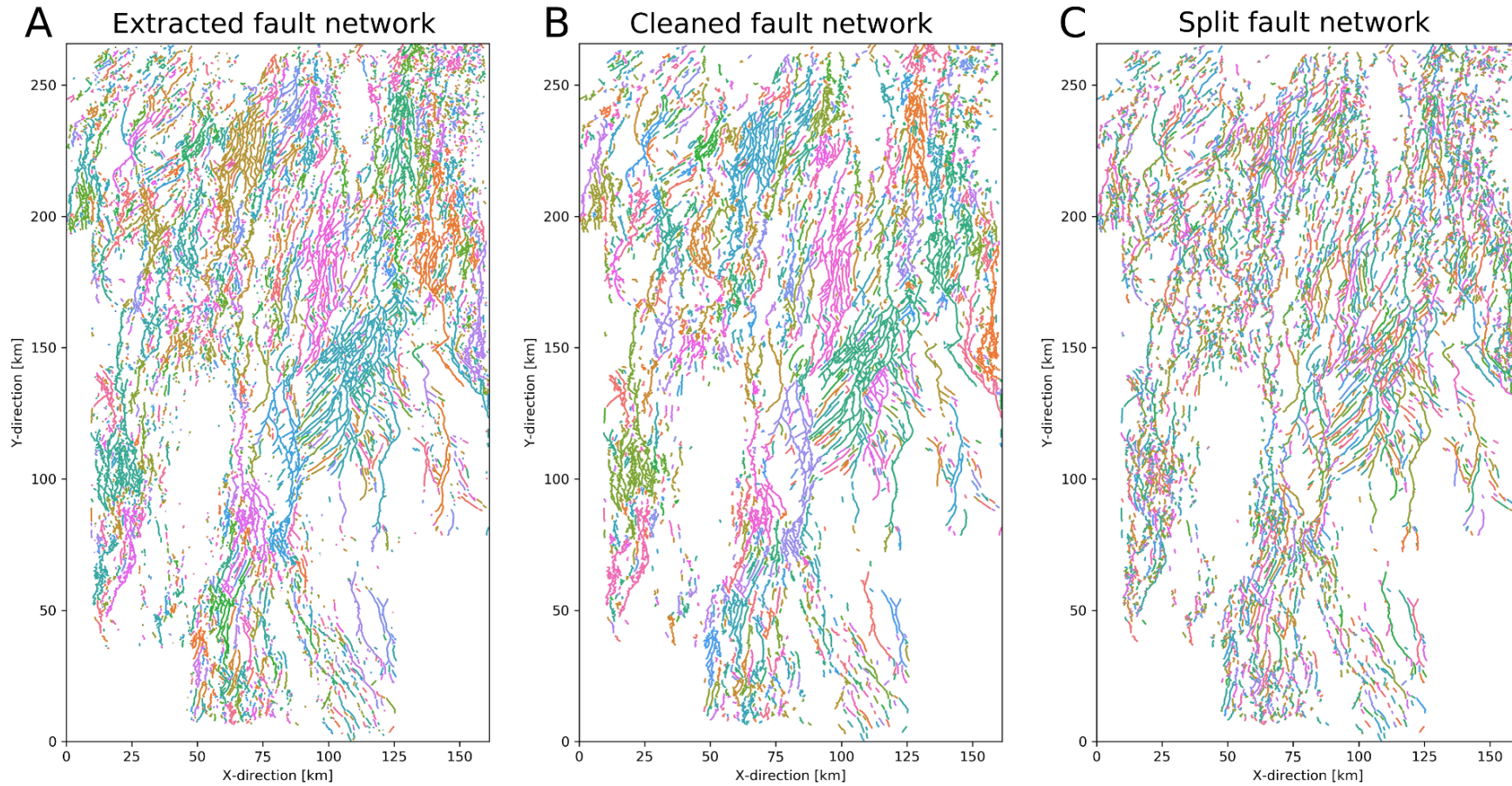



Figure 7: A Fault network extracted from BCU (Fig. 4D). Note the large areas with the same colours resulting from multiple faults being grouped into one connected component **B** Fault network after removal of noise (i.e. small components). **C** Fault network after splitting junctions previously connecting splaying and intersecting faults. Note that large connected components are split up and individual faults are highlighted by different colours.

3.3. Automated fault network extraction and analysis

Extracting a fault network from the 3-D volume allows us to perform a comprehensive geometric analysis of the fault system using our fault analysis toolbox - fatbox (Wrona et al., 2022). The basic idea is to describe a fault system in 2-D as a network (or graph), i.e. sets of nodes and edges (Fig. 5). Each node marks a location along the fault and each edge connects two nodes. All nodes connected to one another by edges are labelled as a (connected) component.

Our fault extraction workflow consists of these eight steps: (1) extract horizon, (2) Gaussian blur filter, (3) thresholding, (4) cleaning, (5) skeletonization, (6) connect components, (7) add nodes to graph, (8) add edges to graph and (9) split junctions. While applying it to our North Sea target region, we first attempt to capture as many faults as possible by extracting the fault score along a horizon 500 m below Base Cretaceous Unconformity (BCU) (Fig. 1C). Here we observe a large number of faults, which were either formed in the second rift phase, or formed in the first rift phase and reactivated in the second rift phase (Figs. 4, 6A). Second, we apply a Gaussian blur filter to increase lateral fault continuity (Fig. 6B), which allows us to extract long, geologically plausible faults. Using a small filter ($\sigma=2$) results in local smoothing without connecting distant faults with one another. Third, we apply a threshold of 0.35 to separate the faults from the background in the fault likelihood (Fig. 6C). This threshold is a tradeoff, which balances capturing as many faults as possible (lower values) and identifying clearly resolvable faults (high values). Four, we further restrict this threshold and essentially filter these points by removing areas smaller than 25 pixels (Fig. 6D). Five, we collapse the faults to one-pixel wide lines using skeletonization (Guo and Hall, 1992) (Fig. 6E). Six, we label adjacent pixels in the image as connected components (Wu et al., 2009) (Fig. 6F). Each component consists of pixels which are connected to each other. These components represent the faults in the network. At this point, we can build our graph using these connected components of the image (Fig. 6F). Each pixel belonging to a component becomes a node whereas edges are created between neighbouring nodes (Fig 6G-I). This process results in a number of faults with splays, junctions or intersections being grouped into one connected component (Fig. 7A). To correct this, we split up junctions (nodes with three edges) based on the similarity of strike, i.e. aligned branches

192   remain connected (Fig. 7B,C). This final network is compared to the Base Late Jurassic horizon

193   mapped by Tillmans et al., (2021) (Fig. 8). Additionally, we perform the exact same workflow on ten

194   slices through the fault score volume (1-10 km depth) to capture 3-D fault geometries with depth (Fig.

195   9).

196   After extracting the fault system, we calculate a series of typical fault properties using our fault

197   analysis toolbox - fatbox (Wrona et al., 2022) (Fig. 10). First, we calculate the fault length as the sum

198   of the edge lengths of each component (Fig. 10B). Second, we calculate the strike along the fault from

199   neighbouring nodes (Fig. 10C). If we were to calculate the overall fault strike, we would overlook

200   along-strike variations in strike. If we were to calculate the strike as the orientation of each edge, we

201   would only obtain values of 0, 45 or 90°, because the nodes are closely spaced. Instead, we calculate

202   the strike from the 3rd degree neighbouring nodes (i.e. neighbours of neighbours of neighbours). This

203   assures a robust, high resolution fault strike calculation. Combining the fault length and strike, we can

204   generate a length-weighted Rose diagram (Fig. 10C). Finally, we calculate the fault density as the

205   fault length per area (Fig. 10D).

206   3.4. Comparison to conventional seismic interpretation

207   We can ask ourselves, "how good are our results compared to a state-of-the-art fault

208   interpretation from the same data set using conventional fault mapping techniques?" (Fig. 8). Tillmans

209   et al., (2021) map the Base Late Jurassic (base of syn-rift sediments associated with rift phase 2) on

210   the eastern flank of the North Viking Graben (see Figs. 1A, 4 for location) using a combination of

211   manual picking and auto-tracking on the same seismic dataset. This horizon is calibrated with 40

212   exploration wells, which provide direct constraints on the depth of the surface. Tillmans et al. (2021)

213   highlight the fault system by computing the variance attribute (Chopra and Marfurt, 2007) along the

214   horizon (Fig. 7A). On top of the horizon, we plot the fault network mapped from the fault score

215   extracted 500 m below the easily-mappable Base Cretaceous Unconformity (BCU) (Fig. 8B). This

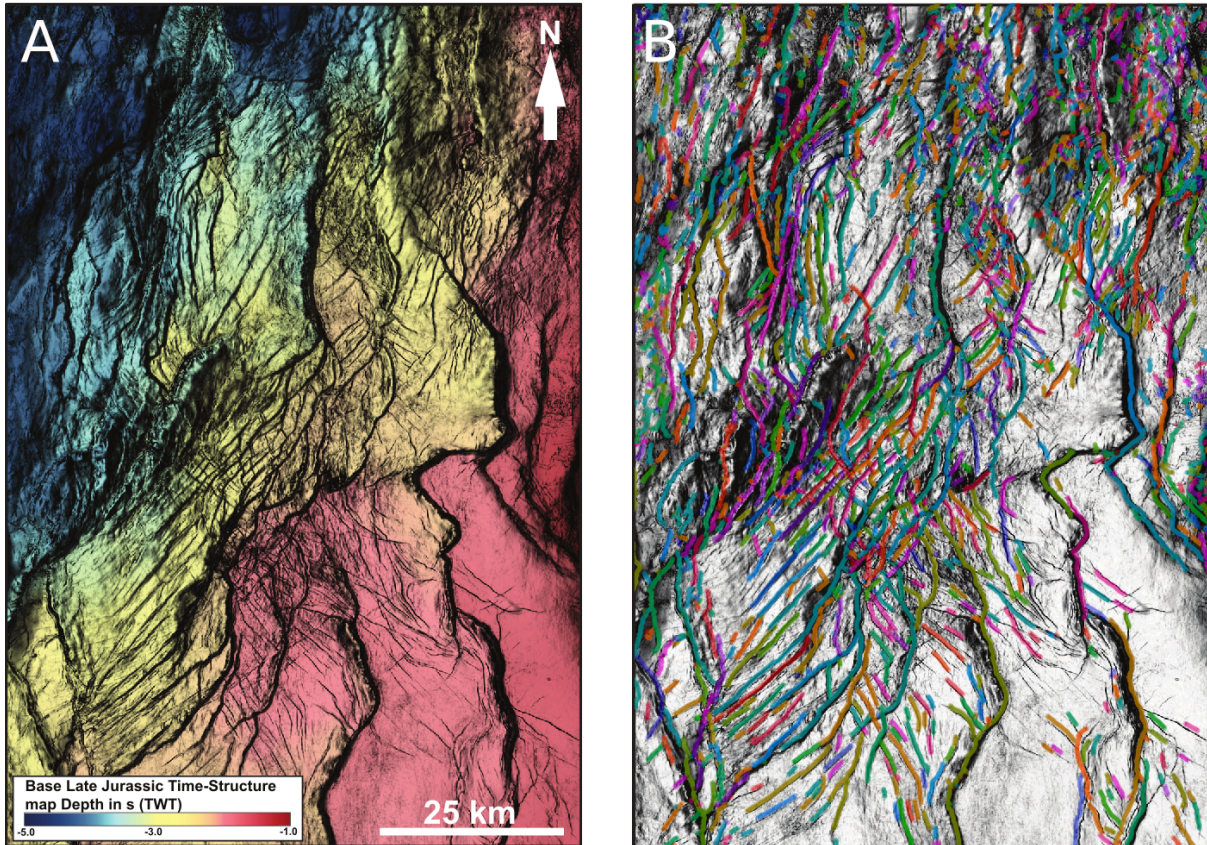
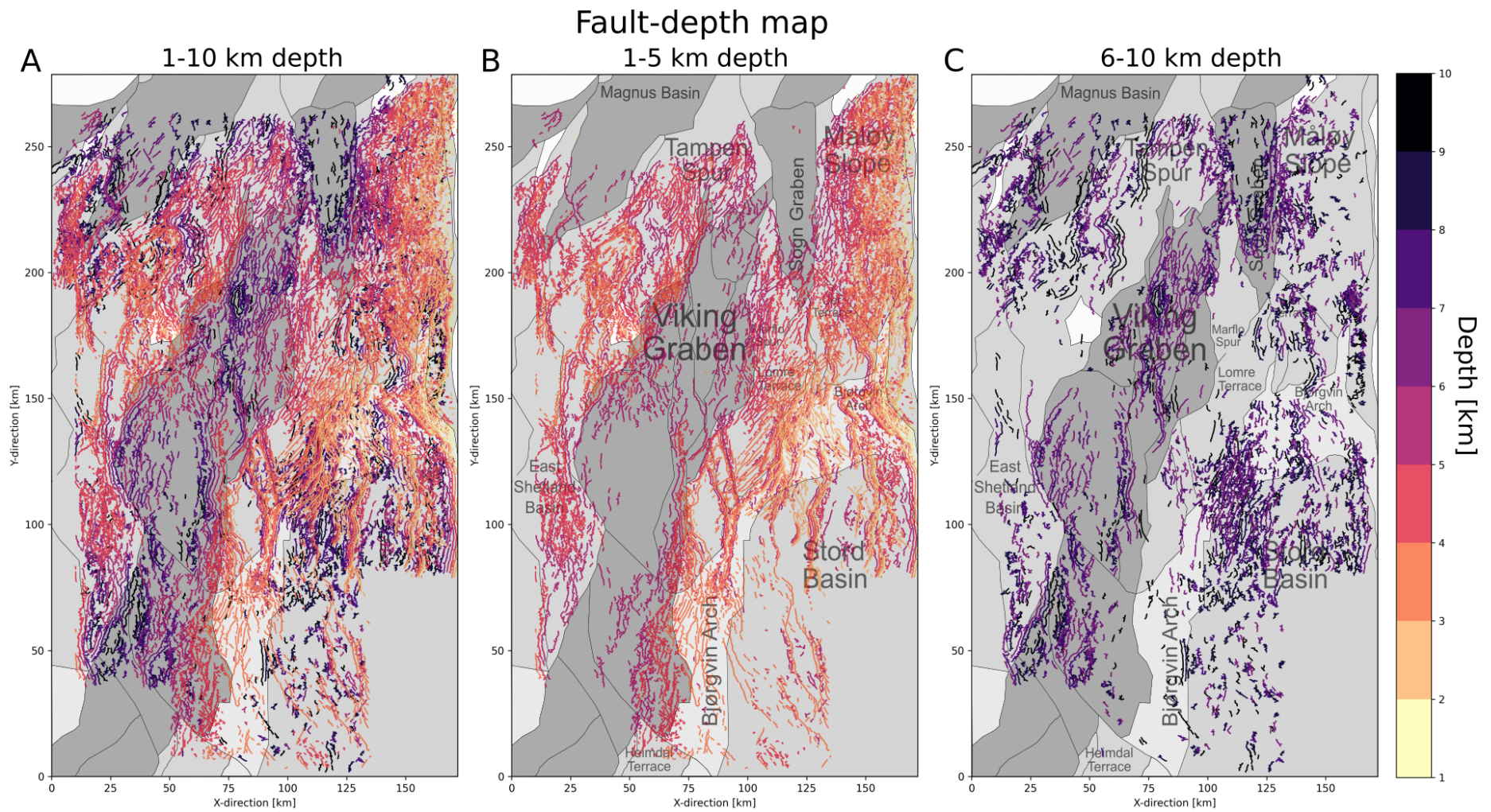
216   visual comparison shows that while we are missing a few faults in the southwest of the map, we are

217   able to identify and accurately represent most of the faults identified by Tillmans et al. (2021). The

218   missing faults are either overlooked by our model (i.e. false negatives) or result from the difference in

219 the horizons that we compare: Base Cretaceous Unconformity (our study) versus Base Late Jurassic

220 (Tillmans et al., 2021).

221

222
223 **Figure 8:** Comparison of **A** Base Late Jurassic time-structure map interpreted by Tillmans et al.,
224 (2021) and **B** Automatically-extracted fault network 500 m below Base Cretaceous Unconformity
225 using the same seismic dataset. Faults are distinguished by colour.

226
Figure 9: Fault map of the northern North Sea extracted every kilometre between 1-10 km depth (**A**), 1-5 km depth (**B**) and 6-10 km depth (**C**) with structural
228   elements from the Norwegian Petroleum Directorate or NPD (2022).

229

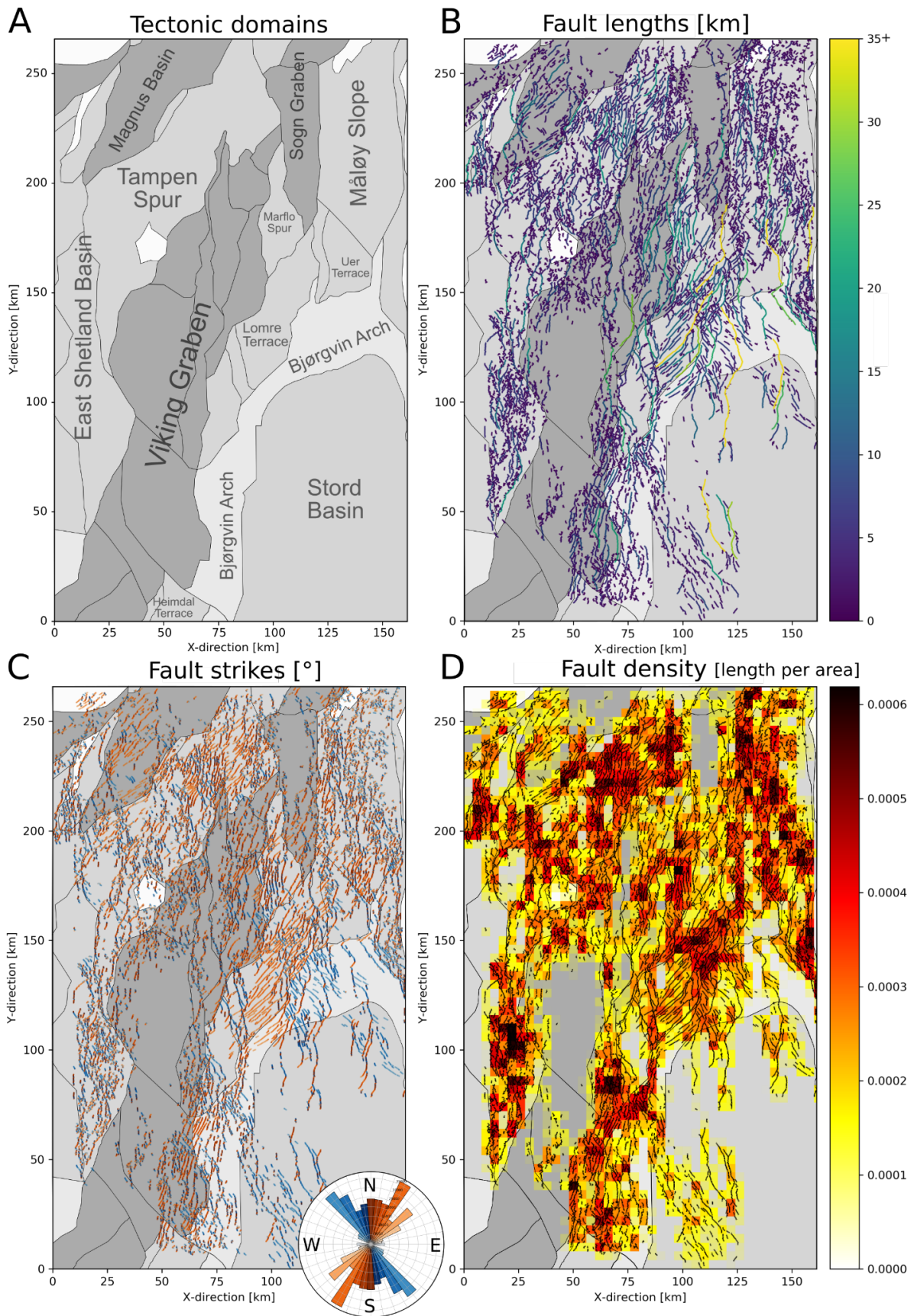



Figure 10: **A** Structural elements of the northern North Sea Rift (NPD, 2022) **B** Fault lengths (500 m below BCU) on top of structural elements. **C** Fault strikes (500 m below BCU) on top of structural elements with length-weighted Rose diagram. **D** Fault density on top of structural elements. Note that

233   fault density was measured as fault length per square area. These squares have an edge length of 3.6
234   km; a value chosen for visual purposes.

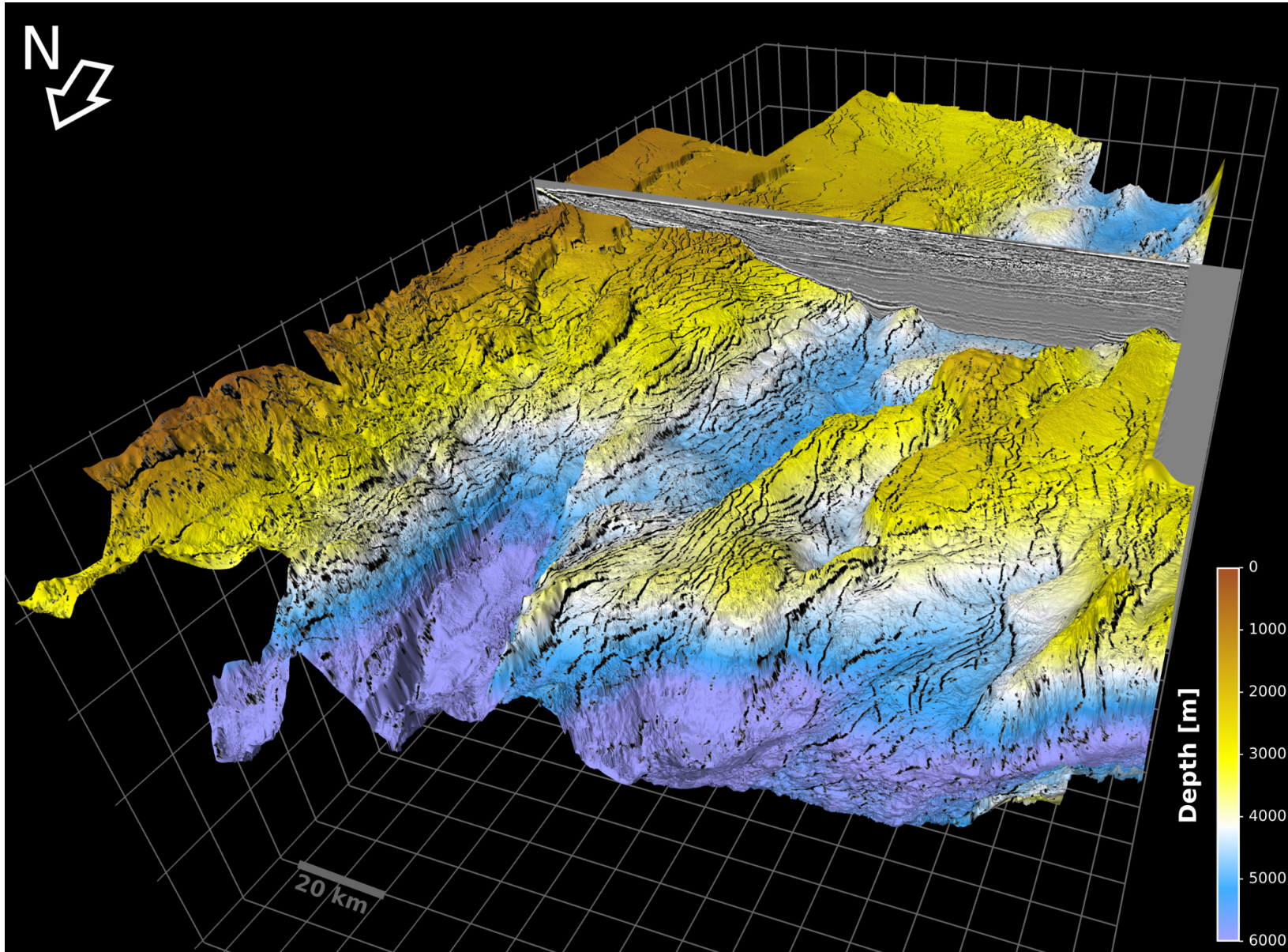



Figure 11: 3-D perspective of the northern North Sea rift showing the Base Cretaceous Unconformity overlain with faults (black) extracted from 3-D seismic reflection data with deep learning. Vertical exaggeration of 5.

4. OBSERVATIONS

Our fault extraction allows us to map a complex network consisting of 7983 individual faults across an approximately 161 km-wide and 266 km-long area, covering 35,410 km^2 of the northern North Sea rift (Fig. 7C).

4.1. Fault length

Faults vary in length by 3 orders of magnitude - from 50 m to 75.9 km, with some of the longest faults (>30 km) extending from the Stord Basin and Bjørgvin Arch in the south to the Uer and Lomre Terrace in the north (Fig. 10B). In cross-section, these faults have up to several kilometres of displacement and bound rotated half-graben (e.g. Whipp et al., 2014; Bell et al., 2014) (Fig. 3B,C). While we observe some long (up to 20 km) faults in the Viking Graben and Tampen Spur, most faults (>90%) are closely spaced (< 5 km) and relatively short (<10 km long) (Fig. 10B).

4.2. Fault strikes

In map view, we observe a complex network consisting of a large number of variably trending faults that display a broad range of intersection styles (e.g., oblique, perpendicular). These faults show a large range of strikes, varying from NW-SE to NE-SW (Figs. 9, 10C). The length-weighted rose plot shows that most faults strike NW-SE (light blue) or NNE-SSW (light orange), with a large number showing intervening strike directions (Fig. 10C). This general divide occurs between predominantly NW-SE-striking faults along the eastern part of the rift and NE-SW-striking faults in the central and northwestern part of the rift. This divide becomes most evident when comparing faults on the Lomre Terrace (NE-SW) to the adjacent Bjørgvin Arch (NW-SE), at least at the structural level of the Base Cretaceous Unconformity (Fig. 10C).

4.3. Fault density

In map view, we observe large variations in fault density 500 below the BCU (Fig. 10D). While dense networks of intersecting faults result in high density areas (e.g. Lomre Terrace, Bjørgvin Arch) we observe low densities in the Viking and Sogn Graben, where faults occur at greater depths (e.g. Fig. 9C).

278 ## 4.4. Vertical continuity

279   The faults extracted at different depths are variable in their vertical continuity (i.e., fault height;

280 Fig. 8). Whereas some faults, in particular in the Stord Basin, the Tampen Spur, and the Magnus

281 Basin show parallel fault traces from 1 to 10 km depth (Fig. 9A), we also observe a large number of

282 faults that occur only at shallower (1-5 km) or at greater depths (6-10 km) (Fig. 9B, C). Upon closer

283 inspection, we observe that the faults, which occur continuously between 1-10 km depth, e.g. in the

284 eastern Stord Basin and the Bjørgvin Arch, are typically large-displacement normal faults with tens of

285 kilometres spacing (e.g. Fig. 3B, C), whereas the other faults, which only occur between 6-10 km

286 depth (e.g. northwestern Stord Basin), are usually shorter and more closely spaced (a few kilometres)

287 (e.g. Fig. 9C).

288
289
290 # 5. DISCUSSION

291 ## 5.1. Advantages of deep learning based fault interpretation

292   When comparing our results to conventional interpretation methods, we can ask ourselves

293 "what value does deep learning add?". Here, we highlight the advantages of the supervised deep

294 learning-based fault interpretation workflow which we present in this study. First, we can predict

295 faults in a seismic section in a fraction of the time (5 seconds) required by expert interpreters (~10

296 minutes). These differences accumulate, in particular when interpreting such a large data set with

297 >22000 inlines. A conventional fault interpretation of such a large data set can take several months,

298 whereas a trained convolutional neural network can identify faults across the entire volume within a

299 day on a single GPU (GeForce GTX 1080 Ti). Note that this comparison does not include the time

300 required to label the training data (~2 days), train the initial model (~4 hours), fine-tune and select the

301 final model (days-months). Second, after identifying faults in seismic reflection data, they also need to

302 be mapped before we can perform the relevant fault analysis. Here we map the fault network using a

303 series of tools from computer vision and network analysis compiled in our fault analysis toolbox -

304 fatbox (Wrona et al., 2022) (Figs. 6, 7). Our automated workflow extracts the fault network in less

305 than five minutes compared to the several weeks to months that would have been required to

306 manually map the faults in this large data set. Furthermore, once extracted, we can immediately

307 conduct a number of typical fault analyses using predefined functions implemented in fatbox (Wrona

308 et al., 2022) (e.g. Fig. 10).

309 Third, conventional fault interpretations are often binary (fault vs. no fault), but deep learning delivers

310 a score ranging from 0 (no fault) to 1 (fault). Although this score is no true fault probability (see

311 discussion by Mosser and Naeini, 2022), the fault score nevertheless correlates with the visibility of

312 faults (i.e. faults, which are well-resolved by the data, are associated with higher fault scores). This

313 allows users to qualitatively select the faults that they want to analyze using a threshold (as done

314 herein).. This selection is particularly useful for assessing the sealing potential of certain layers for

315 CO_2 storage and for predicting fluid flow during geothermal exploration. Fourth, seismic interpreters

316 typically focus on the largest faults, whereas our model performs the same prediction across the entire

317 data set irrespective of the size of the faults encountered. Fifth, given the same data, labels, model and

318 training, our model and results are fully reproducible, which is not the case for conventional fault

319 interpretations, where the interpreter has to make a myriad of decisions in the process of mapping a

320 fault network.

321 ## 5.2. Complex fault system in the northern North Sea

322 Our study shows how to reveal the complex geometry of normal fault systems in 3-D seismic

323 reflection data using a combination of deep learning and automated fault extraction. We were able to

324 map an intricate network consisting of almost 8000 individual faults that cover an area approximately

325 161 km wide and 266 km long (e.g. Figs. 4, 6, 10). This fault network shows large variations in fault

326 length, strike and density, with extremely complex splays, junctions and intersections between these

327 faults (Figs. 7-11). As such, our work goes far beyond typical seismic interpretations in previous case

328 studies, which covered only a fraction of the rift (e.g. Duffy et al., 2015; Deng et al., 2017; Tillmans

329 et al., 2021), or regional studies that mapped <100 of the largest faults using primarily sparse, 2-D

330 seismic sections (e.g. Fig. 1B; Fazlikhani et al., 2017; Phillips et al., 2019).

5.3. Uncertainties during fault mapping

While there are several advantages to our approach, it is worth remembering the uncertainties associated with mapping faults in seismic reflection data. First, seismic reflection data can only image faults with displacement above the seismic resolution (and level of noise) of the data set. The seismic resolution of our data set decreases from 15 m (vertical) and 30 m (lateral) around 3 km depth down to 180 m (vertical) around 20 km depth (see Wrona et al., 2019; Tillman et al., 2021). Second, the labels we use to train our model are derived from 22 interpreted seismic sections, which, like any seismic interpretation, contains the expertise and biases of the interpreter (e.g. Bond et al., 2007, Bond 2015). Third, our current model has not been trained and is thus unable to distinguish between different fault types (normal, reverse, strike-slip). We labelled all major faults in the training data, which are predominantly normal faults (probably >99%). A handful of these normal faults may show evidence of minor inversion, but they all remain in net-extension, i.e. the hanging wall has moved down relative to the footwall. While strike-slip faults are notoriously difficult to resolve in seismic reflection data, as they show little to no vertical offset of reflectors, normal and reverse faults show differing offsets, which neural networks could learn to recognize by correlating reflectors across the fault. Machine learning models could thus be able to distinguish fault types based on their seismic signature in the future. Fourth, the convolutional neural network that we trained achieves an accuracy of 83%, implying that 17% of the data is misclassified (see Wrona et al., 2021). A closer inspection reveals that 36% are false positives (i.e. faults that were overlooked) and 5% are false negatives (i.e. faults that were misinterpreted) (see Wrona et al., 2021). Despite these limitations, the robustness of our approach is evident when considering along-strike fault continuity across a large number of different seismic lines (Fig. 10, 11).

5.4. Future research on automated fault mapping

Based on our work, we can identify three related areas for future research. First, conventional neural networks predict a fault score from 0 to 1, which seems to correspond to the visibility of the fault in the dataset. Bayesian neural networks, on the other hand, allow the prediction of true fault probabilities (e.g. Mosser et al., 2020). Predicting fault probabilities in regional seismic data sets

could significantly accelerate the screening for and risk assessment of potential CO_2 storage sites (see Wrona and Pan, 2021). Second, we currently map faults on seismic in- and crosslines, which may contain redundant information regarding the faults. In the future, it may be advantagous to maximize the diversity of the training set (i.e. different fault types or levels of noise) using uncertainty estimates and active learning. Third, in addition to predicting where faults occur, we can explore the prediction of other fault properties, such as displacement, fault zone permeability or even the time when they were active. This would significantly allow us to study the spatial and temporal evolution of fault systems in high resolution at a regional scale. ~~Third~~Fourth, while our fault extraction workflow currently focuses on mapping fault networks in a series of 2-D slices or horizons, we really need freely-available methods to generate 3-D fault surfaces, which allow for complex fault splays, junctions and intersections, as they could be applied to large 3-D seismic data sets, but also to analogue and numerical models.~~observed here.~~

6. CONCLUSIONS

This study shows that the combination of deep learning and network analysis applied to 3-D seismic reflection data allows us, for the first time, to map almost 8000 normal faults across the entire northern North Sea rift. These faults form an intricate network with complex relationships (e.g. splays, junctions, intersections) including large variations in fault length (50 m to 75.9 km) and strikes (NW-SE to NE-SW). As such, this work goes far beyond previous seismic studies by providing high resolution fault maps at a regional scale in a fraction of the time required by conventional interpretation methods.

ACKNOWLEDGEMENTS

We would like to thank The Norwegian Academy of Science and Letters (VISTA), The University of Bergen and The Initiative and Networking Fund of the Helmholtz Association through the project "Advanced Earth System Modelling Capacity (ESM), The Geo.X Network and Deutsche Forschungsgemeinschaft (Projektnummer 460760884) for supporting this research. I. Pan acknowledges the NUAcT fellowship for partially supporting the work. We are grateful to CGG, in

384 particular Stein Åsheim and Jaswinder Mann, for the permission to present this data and publish this

385 work. We thank Schlumberger for providing the software Petrel 2019© and Leo Zijerveld for IT

386 support.

387

388 REFERENCES

389 Bartholomew, I.D., Peters, J.M., and Powell, C.M., 1993, Regional structural evolution of the North
390   Sea: Oblique slip and the reactivation of basement lineaments, *in* Petroleum Geology
391   Conference Proceedings, v. 4, p. 1109–1122, doi:10.1144/0041109.
392 Bell, R.E., Jackson, C.A.L., Whipp, P.S., and Clements, B., 2014, Strain migration during multiphase
393   extension: Observations from the northern North Sea: Tectonics, v. 33, p. 1936–1963,
394   doi:10.1002/2014TC003551.
395 Bissell, R.C., Vasco, D.W., Atbi, M., Hamdani, M., Okwelegbe, M., and Goldwater, M.H., 2011, A
396   full field simulation of the In Salah gas production and CO2 storage project using a coupled
397   geo-mechanical and thermal fluid flow simulator: Energy Procedia, v. 4, p. 3290–3297.
398 Bond, C.E., 2015, Uncertainty in structural interpretation: Lessons to be learnt: Journal of Structural
399   Geology, v. 74, p. 185–200, doi:10.1016/j.jsg.2015.03.003.
400 Brun, J.P., and Tron, V., 1993, Development of the North Viking Graben: inferences from laboratory
401   modelling: Sedimentary Geology, v. 86, p. 31–51, doi:10.1016/0037-0738(93)90132-O.
402 Chopra, S., and Marfurt, K.J., 2007, Seismic Attributes for Prospect Identification and Reservoir
403   Characterization: Society of Exploration Geophysicists and European Association of
404   Geoscientists and Engineers, doi:10.1190/1.9781560801900.
405 Claringbould, J.S., Bell, R.E., Jackson, C.A.L., Gawthorpe, R.L., and Odinsen, T., 2020, Pre-breakup
406   Extension in the Northern North Sea Defined by Complex Strain Partitioning and
407   Heterogeneous Extension Rates: Tectonics, v. 39, doi:10.1029/2019TC005924.
408 Clerc, C., Jolivet, L., and Ringenbach, J.C., 2015, Ductile extensional shear zones in the lower crust
409   of a passive margin: Earth and Planetary Science Letters, v. 431, p. 1–7,
410   doi:10.1016/j.epsl.2015.08.038.
411 Deng, C., Fossen, H., Gawthorpe, R.L., Rotevatn, A., Jackson, C.A.L., and FazliKhani, H., 2017,
412   Influence of fault reactivation during multiphase rifting: The Oseberg area, northern North
413   Sea rift: Marine and Petroleum Geology, v. 86, p. 1252–1272,
414   doi:10.1016/J.MARPETGEO.2017.07.025.
415 Doré, A.G., Lundin, E.R., Fichler, C., and Olesen, O., 1997, Patterns of basement structure and
416   reactivation along the NE Atlantic margin: Journal of the Geological Society, v. 154, p. 85–
417   92, doi:10.1144/gsjgs.154.1.0085.
418 Duffy, O.B., Bell, R.E., Jackson, C.A.L., Gawthorpe, R.L., and Whipp, P.S., 2015, Fault growth and
419   interactions in a multiphase rift fault network: Horda Platform, Norwegian North Sea: Journal
420   of Structural Geology, v. 80, p. 99–119, doi:10.1016/J.JSG.2015.08.015.
421 Færseth, R.B., 1996, Interaction of permo-triassic and jurassic extensional fault-blocks during the
422   development of the northern North Sea: Journal of the Geological Society, v. 153, p. 931–
423   944, doi:10.1144/gsjgs.153.6.0931.
424 Færseth, R.B., Knudsen, B.E., Liljedahl, T., Midbøe, P.S., and Søderstrøm, B., 1997, Oblique rifting
425   and sequential faulting in the Jurassic development of the northern North Sea: Journal of
426   Structural Geology, v. 19, p. 1285–1302, doi:10.1016/s0191-8141(97)00045-x.
427 Fazlikhani, H., Fossen, H., Gawthorpe, R.L., Faleide, J.I., and Bell, R.E., 2017, Basement structure
428   and its influence on the structural configuration of the northern North Sea rift: Tectonics, v.
429   36, p. 1151–1177, doi:10.1002/2017TC004514.
430 Guo, Z., and Hall, R.W., 1992, Fast fully parallel thinning algorithms: CVGIP: Image Understanding,
431   v. 55, p. 317–328, doi:10.1016/1049-9660(92)90029-3.
432 Kampman, N., Bickle, M., Wigley, M., and Dubacq, B., 2014, Fluid flow and CO2–fluid–mineral

433   interactions during CO2-storage in sedimentary basins: Chemical Geology, v. 369, p. 22–50.

434 Lohr, T., Krawczyk, C.M., Oncken, O., and Tanner, D.C., 2008, Evolution of a fault surface from 3D

435   attribute analysis and displacement measurements: Journal of Structural Geology, v. 30, p.

436   690–700, doi:10.1016/j.jsg.2008.02.009.

437 Maystrenko, Y.P., Olesen, O., Ebbing, J., and Nasuti, A., 2017, Deep structure of the northern north

438   sea and southwestern Norway based on 3D density and magnetic modelling: Norsk Geologisk

439   Tidsskrift, v. 97, p. 169–210, doi:10.17850/njg97-3-01.

440 Moeck, I., Kwiatek, G., and Zimmermann, G., 2009, Slip tendency analysis, fault reactivation

441   potential and induced seismicity in a deep geothermal reservoir: Journal of Structural

442   Geology, v. 31, p. 1174–1182.

443 Morris, A., Ferrill, D.A., and Henderson, D.B., 1996, Slip-tendency analysis and fault reactivation:

444   Geology, v. 24, p. 275–278.

445 Mosser, L., Purves, S., and Naeini, E.Z., 2020, Deep bayesian neural networks for fault identification

446   and uncertainty quantification: 1st EAGE Digitalization Conference and Exhibition,

447   doi:10.3997/2214-4609.202032036.

448 Mosser, L., and Naeini, E.Z., 2022, A comprehensive study of calibration and uncertainty

449   quantification for Bayesian convolutional neural networks — An application to seismic data,

450   GEOPHYSICS 87: IM157-IM176. https://doi.org/10.1190/geo2021-0318.1

451

452 Naliboff, J.B., Glerum, A., Brune, S., Péron-Pinvidic, G., and Wrona, T., 2020, Development of 3D

453   rift heterogeneity through fault network evolution: Geophysical Research Letters,

454   doi:10.1029/2019gl086611.

455 Osagiede, E.E., Rotevatn, A., Gawthorpe, R., Kristensen, T.B., Jackson, C.A., and Marsh, N., 2020,

456   Pre-existing intra-basement shear zones influence growth and geometry of non-colinear

457   normal faults, western Utsira High–Heimdal Terrace, North Sea: Journal of Structural

458   Geology, v. 130, p. 103908.

459 Pan, S., Naliboff, J., Bell, R.E., and Jackson, C., 2021, Bridging spatiotemporal scales of normal fault

460   growth using numerical models of continental extension: EarthArXiv,

461   doi:10.31223/X5DG7M.

462 Phillips, T.B., Fazlikhani, H., Gawthorpe, R.L., Fossen, H., Jackson, C.A.L., Bell, R.E., Faleide, J.I.,

463   and Rotevatn, A., 2019, The Influence of Structural Inheritance and Multiphase Extension on

464   Rift Development, the NorthernNorth Sea: Tectonics, v. 38, p. 4099–4126,

465   doi:10.1029/2019TC005756.

466 Tillmans, F., Gawthorpe, R.L., Jackson, C.A. -L., and Rotevatn, A., 2021, Syn-rift sediment gravity

467   flow deposition on a Late Jurassic fault-terraced slope, northern North Sea: Basin Research,

468   v. 33, p. 1844–1879, doi:10.1111/BRE.12538.

469 Torsvik, T.H., Andersen, T.B., Eide, E.A., and Walderhaug, H.J., 1997, The age and tectonic

470   significance of dolerite dykes in western Norway: Journal of the Geological Society, v. 154,

471   p. 961–973, doi:10.1144/gsjgs.154.6.0961.

472 Whipp, P.S., Jackson, C.A.L., Gawthorpe, R.L., Dreyer, T., and Quinn, D., 2014, Normal fault array

473   evolution above a reactivated rift fabric; a subsurface example from the northern Horda

474   Platform, Norwegian North Sea: Basin Research, v. 26, p. 523–549, doi:10.1111/bre.12050.

475 Wiest, J.D., Wrona, T., Bauck, M.S., Fossen, H., Gawthorpe, R.L., Osmundsen, P.T., and Faleide, J.I.,

476   2020, From Caledonian Collapse to North Sea Rift: The Extended History of a Metamorphic

477   Core Complex: Tectonics, v. 39, doi:10.1029/2020TC006178.

478 Wrona, T., Brune, S., Gayrin, P., and Hake, T., 2022, Fatbox - Fault Analysis Toolbox: GFZ Data

479   Services, doi:10.5880/GFZ.2.5.2022.002.

480 Wrona, T., Magee, C., Fossen, H., Gawthorpe, R.L.L., Bell, R.E.E., Jackson, C.A.-L.A.L., and

481   Faleide, J.I.I., 2019, 3-D seismic images of an extensive igneous sill in the lower crust:

482   Geology, v. 47, p. 729–733, doi:10.1130/G46150.1.

483 Wrona, T., Magee, C., Jackson, C.A.L.C.A.-L.C.A.L., Huuse, M., and Taylor, K.G.K.G., 2017,

484   Kinematics of polygonal fault systems: Observations from the northern north sea: Frontiers in

485   Earth Science, v. 5, p. 101, doi:10.3389/feart.2017.00101.

486 Wrona, T., and Pan, I., 2021, Can machine learning improve carbon storage? Synergies of deep

487   learning, uncertainty quantification and intelligent process control:

488   Wrona, T., Pan, I., Bell, R.E., Gawthorpe, R.L., Fossen, H., and Brune, S., 2021a, 3D seismic
489   interpretation with deep learning: A brief introduction: Society of Exploration Geophysicists,
490   v. 40, p. 524–532, doi:10.1190/tle40070524.1.
491   Wrona, T., Pan, I., Bell, R.E., Gawthorpe, R.L., Fossen, H., and Brune, S., 2021b, 3-D seismic
492   interpretation with deep learning: a set of Python tutorials:, doi:10.5880/GFZ.2.5.2021.001.
493   Wu, X., Liang, L., Shi, Y., and Fomel, S., 2019, FaultSeg3D: Using synthetic data sets to train an end-
494   to-end convolutional neural network for 3D seismic fault segmentation: GEOPHYSICS, v.
495   84, p. IM35–IM45, doi:10.1190/geo2018-0646.1.
496   Wu, K., Otoo, E., and Suzuki, K., 2009, Optimizing two-pass connected-component labeling
497   algorithms: Pattern Analysis and Applications, v. 12, p. 117–135, doi:10.1007/S10044-008-
498   0109-Y.
499   Yukutake, Y., Takeda, T., and Yoshida, A., 2015, The applicability of frictional reactivation theory to
500   active faults in Japan based on slip tendency analysis: Earth and Planetary Science Letters, v.
501   411, p. 188–198.