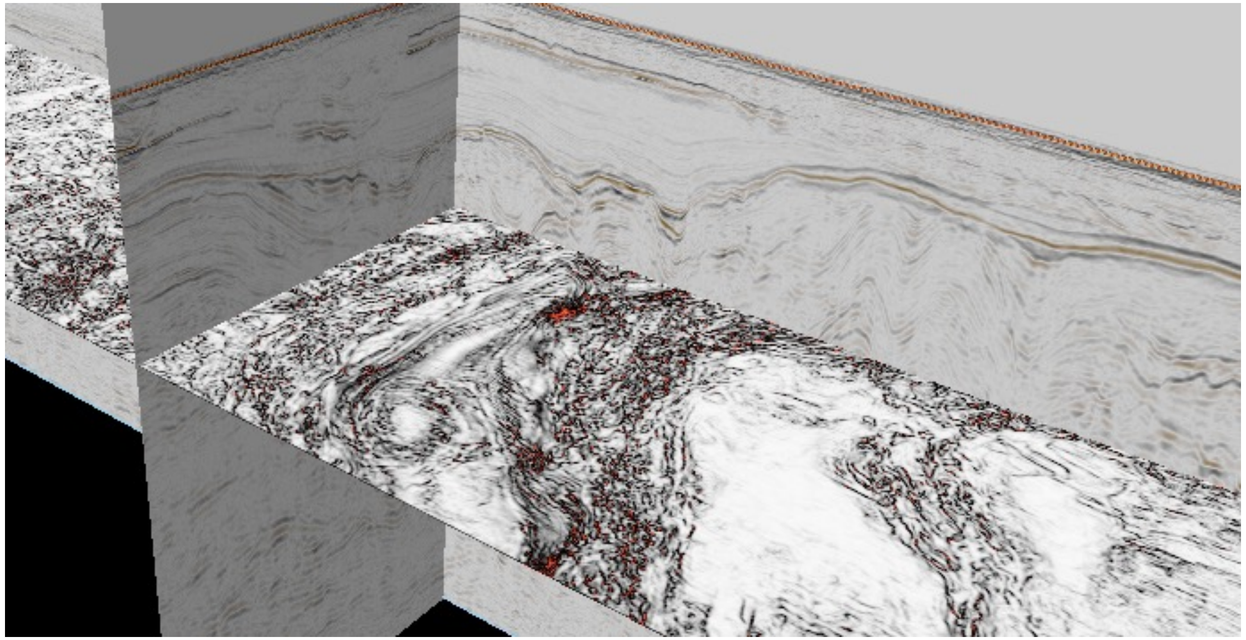
Seismic imaging
Lecture 4



Dr Rebecca (Becky) Bell

2nd floor, office 2.37a

rebecca.bell@imperial.ac.uk

Imperial College London

Module outline

Day 2 morning- Visualising and interpreting seismic data

Lecture 4: The SEGY data convention and where to find data

Lecture 5: How to interpret seismic data

Lecture 6: Seismic attribute analysis and machine learning (future outlook)

Day 2 afternoon- Exercises

Exercise 3: Reading SEGY data and visualizing 3D data

Exercise 4: Calculating seismic attributes

Lecture outline

Objective of lecture 4:

Understand how to view seismic data and the SEGY file format

1) How seismic data is recorded and visualised

2) The SEGY file format

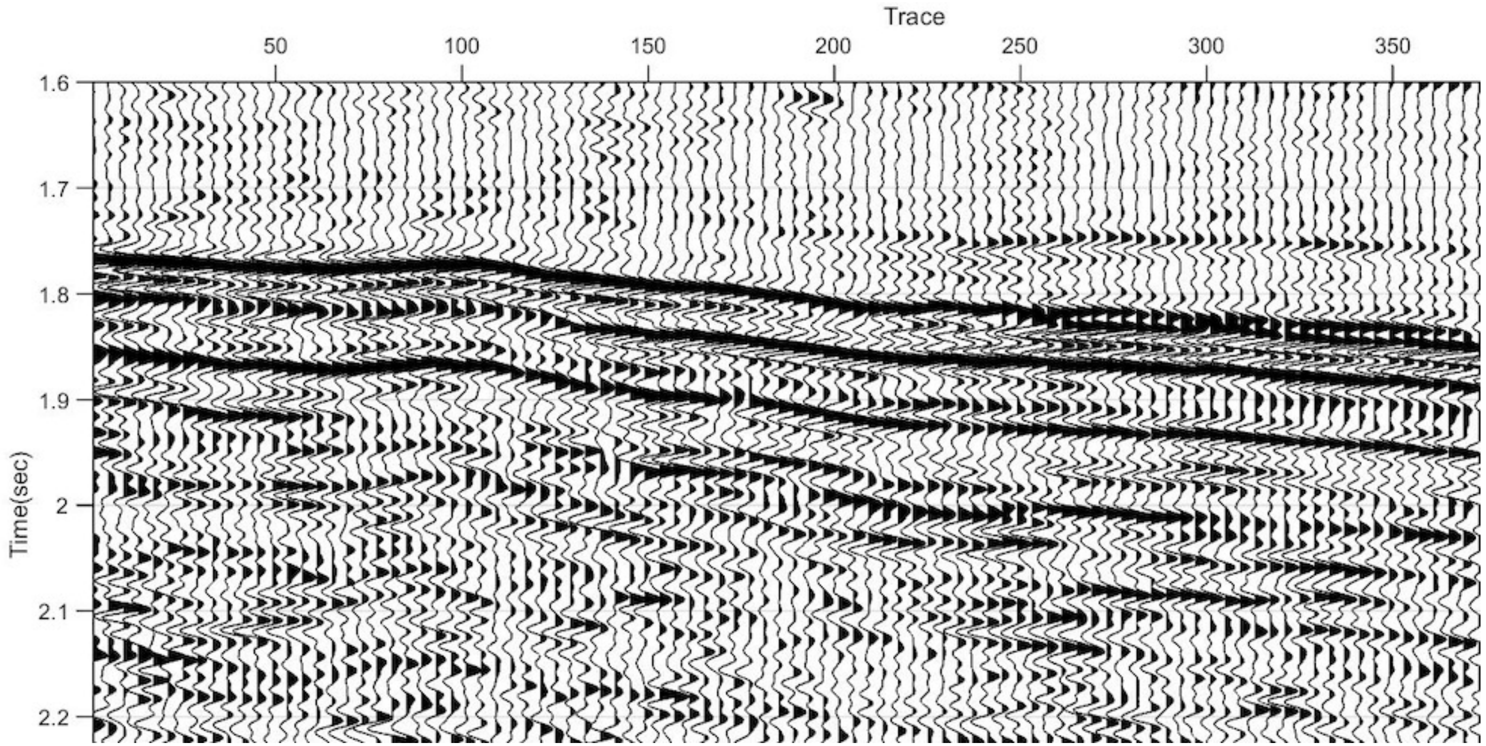
3) Open source SEGY data

This lecture directly links to Exercise 3:
Visualising 3D seismic data with Python

What is post-stack seismic data?

- Once we have stacked CMP gathers (lecture 3) the stacked traces are positioned at their CMP location to produce the stacked seismic section (this is called ‘post-stack’ data. The CMP gathers are ‘pre-stack’ data)

- For 2D data the traces will be spaced at the CMP interval

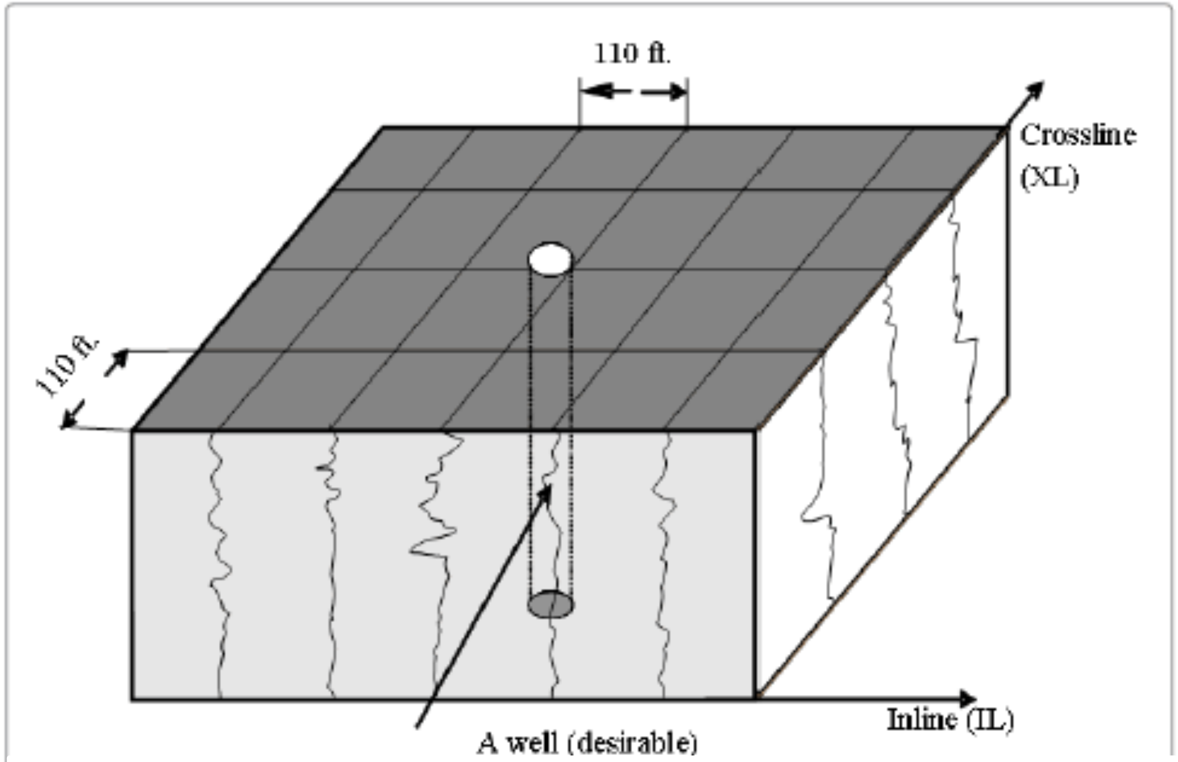


Stacked seismic sections provide an approximate image of the subsurface

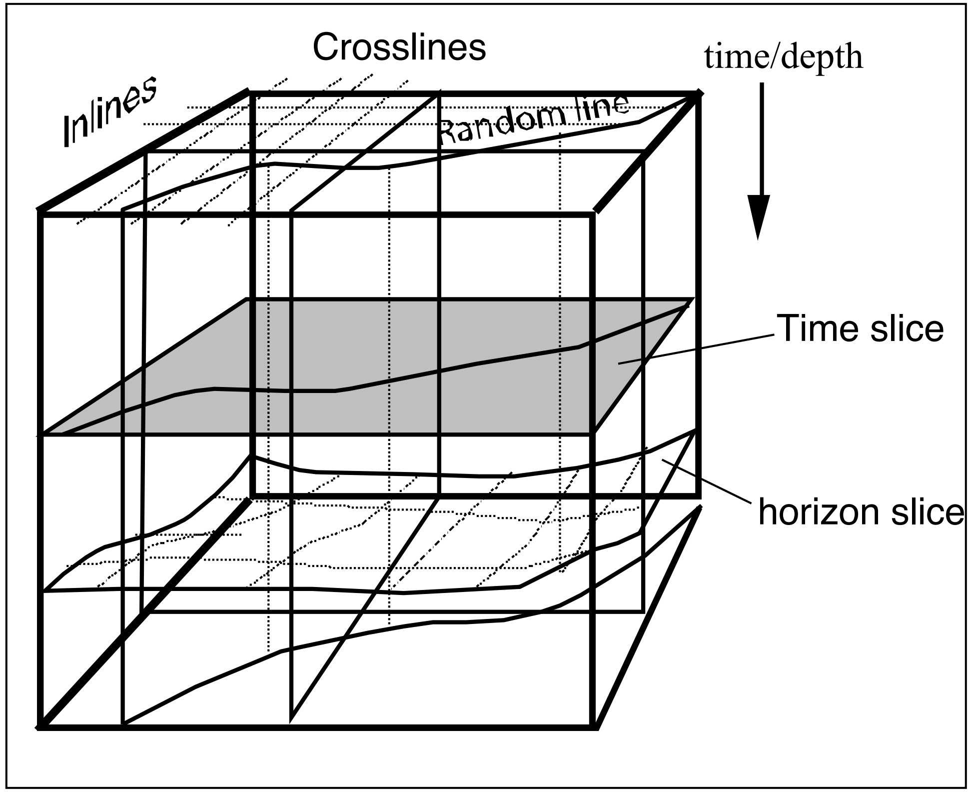
https://www.mathworks.com/matlabcentral/fileexchange/51122-cqi_plotmatrix-data-varargin

What is post-stack seismic data?

- 3D data is commonly viewed as inlines (vertical profiles in the direction the data was recorded), xlines, time/depth slices (horizontal slices) or arbitrary lines
- So a 3D seismic volume is effectively a grid of traces, with a trace at the intersection of every inline and xline.



Chesnokov et al. 2017



Sampling (Digitization)

- The seismic traces are not recorded continuously. The continuous (analogue) signal is recorded discretely (sampled) at a constant time interval

- The time interval (Δt) must be chosen to allow the analogue signal to be reconstructed from the sampled values

- The maximum frequency (f_{max}) to be preserved must satisfy the Nyquist criterion:

$$f_{Max} \leq \frac{1}{2\Delta t}$$

This is the minimum sampling (maximum time interval) to preserve information in a signal

Temporal aliasing (time domain)

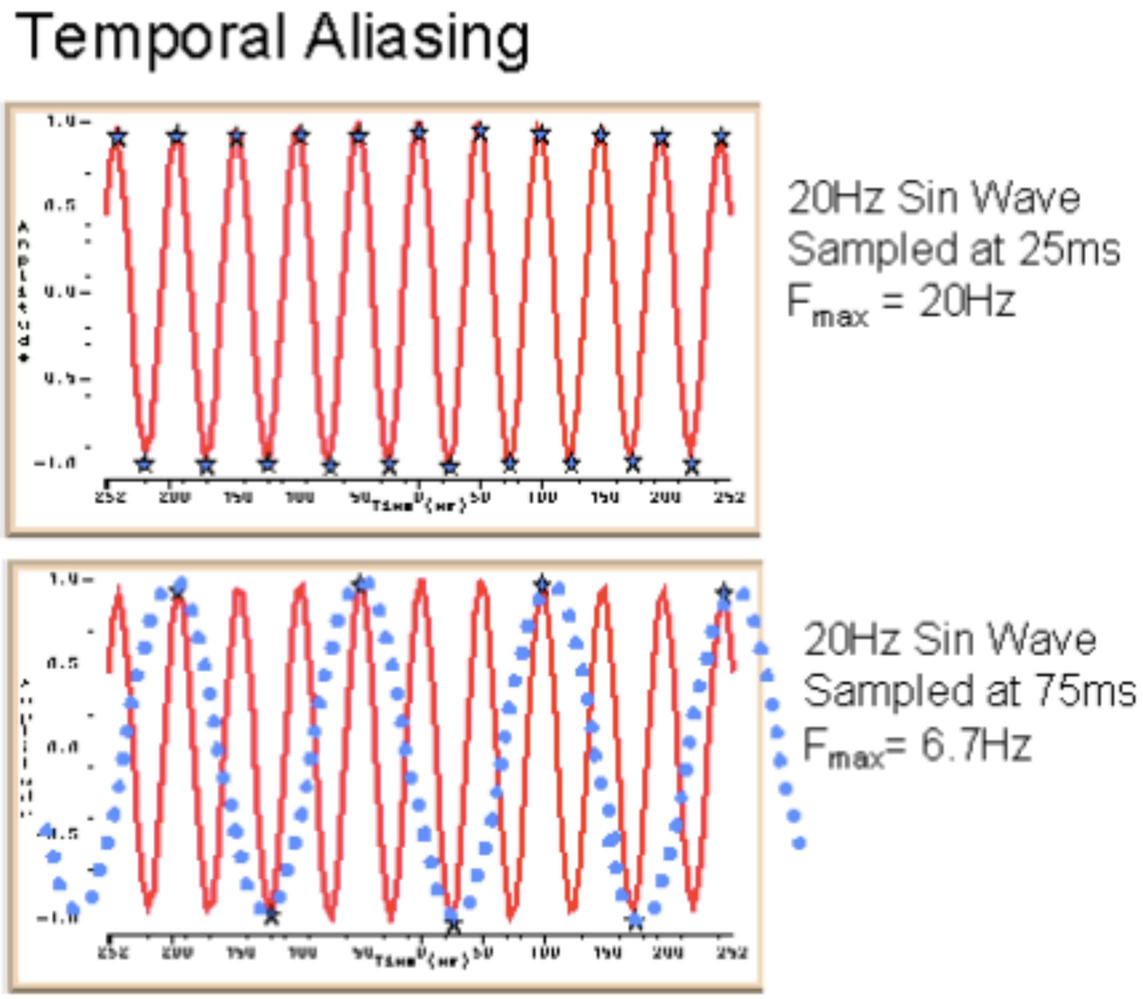
The top slide shows that a good representation of the 20Hz signal can be made by samples taken every 25ms (marked by the blue stars).

F_{max} = 1/2d =1/ 2*0.025 = 20 Hz

In the bottom slide samples are taken every 75ms. An insufficient number of samples are taken and the higher frequency information is "lost" or *aliased*. The original 20Hz red curve appears as a 6.7Hz blue dotted curve.

F_{max} = 1/2d =1/ 2*0.075 = 6.7 Hz

Before the data are sampled the higher frequencies which would be aliased by the chosen sampling interval must be removed by an analogue filter in the recording system.

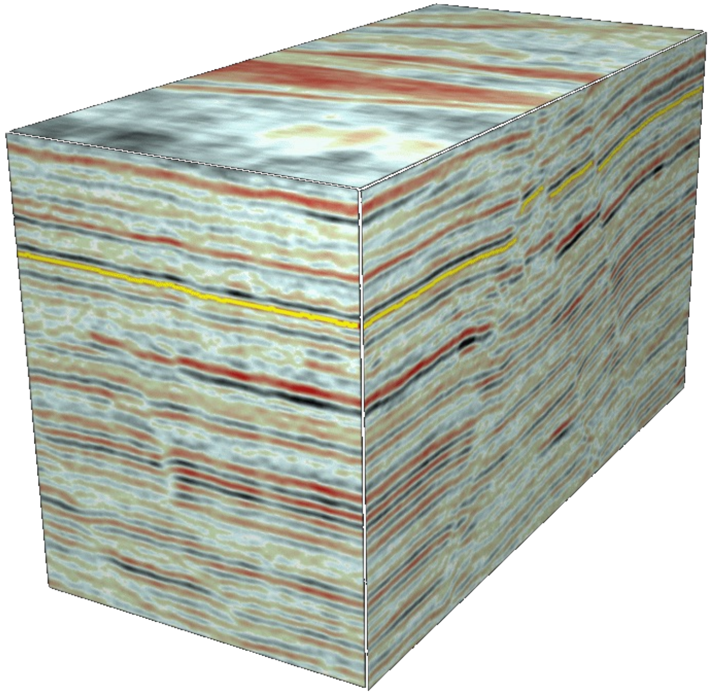


http://www.xsgeo.com/course/acq.htm

Quiz question

- You would like to collect high-frequency seismic data up to 100Hz to identify glacial channels ahead of wind farm development. What is your maximum sample rate?

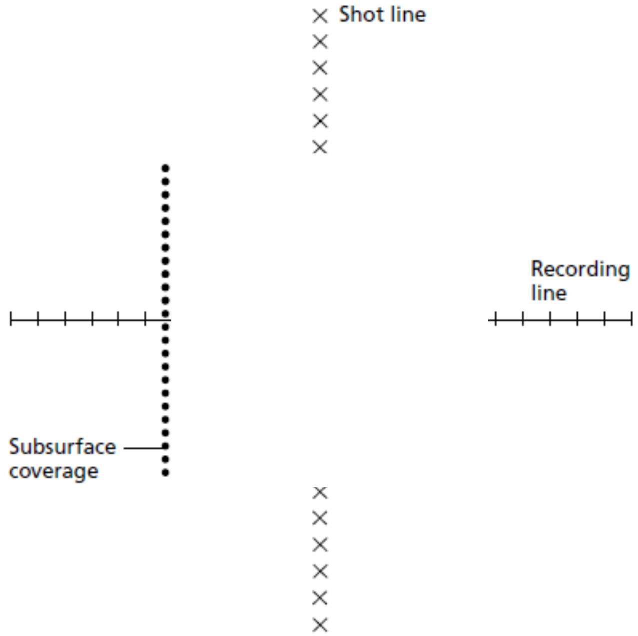
3D seismic: How big is it?



A 10 km x 10 km (100 km^2) 3D survey at 25 m CDP interval contains 400 x 400 traces which equals 160,000 traces…

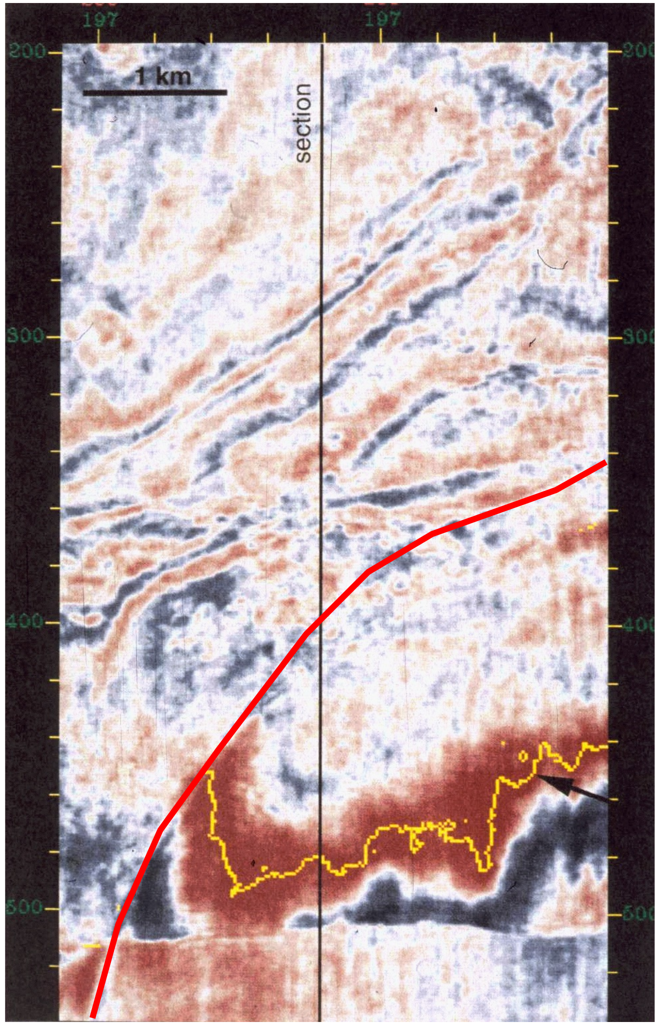
…the majority of 3D surveys now have CDP intervals of 12.5 m resulting in 640,000 traces per 100 km2 dataset!

Time samples are recorded at every 4 ms for crustal-scale surveys (not near-surface surveys), so a single trace covering 4 secs in time = 1000 samples resulting in 6.4×10^8 samples for our 100 km^2 3D seismic survey.

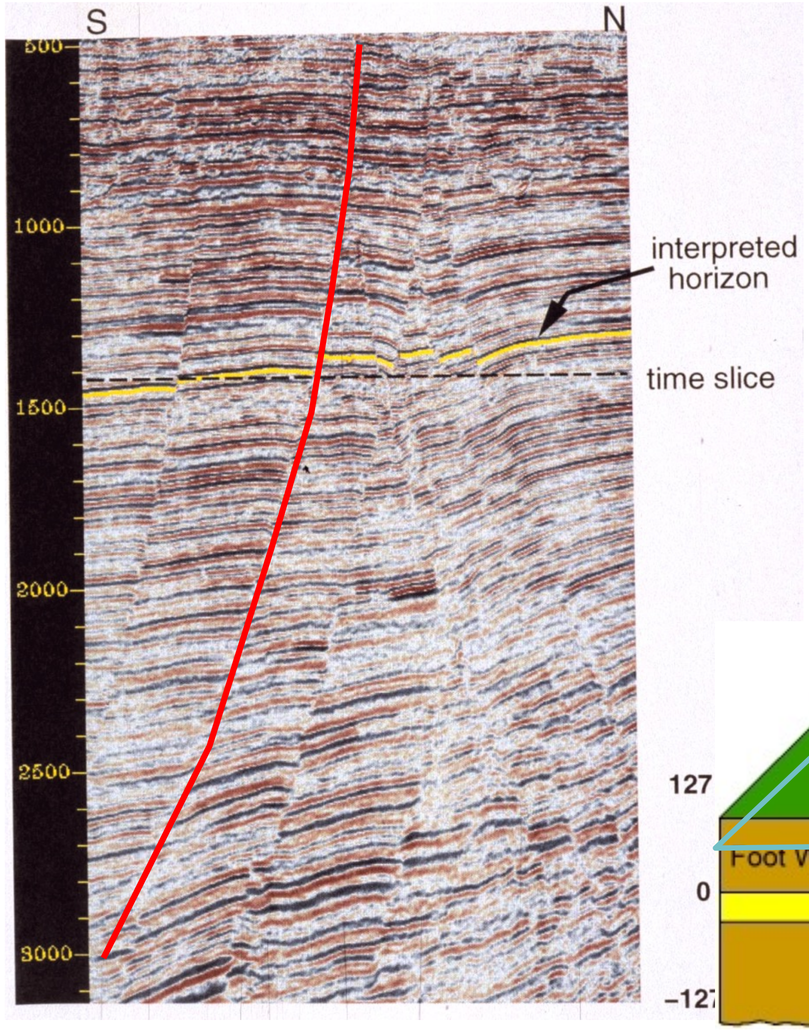


3D seismic volumes can range from a few 100 Mb to 10's or even 100's Gb

3D seismic: How do we view it?

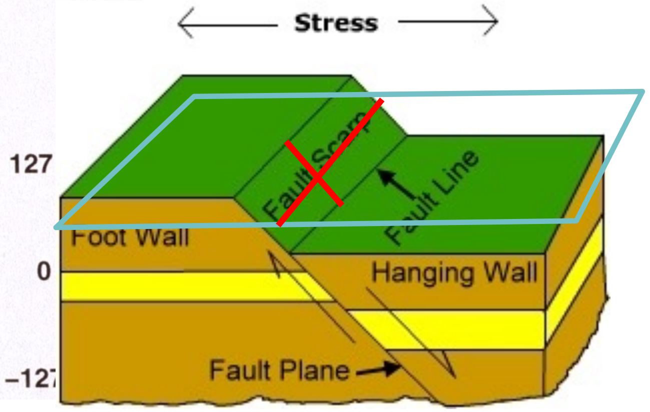


Time slice at 1420 ms



Vertical seismic section

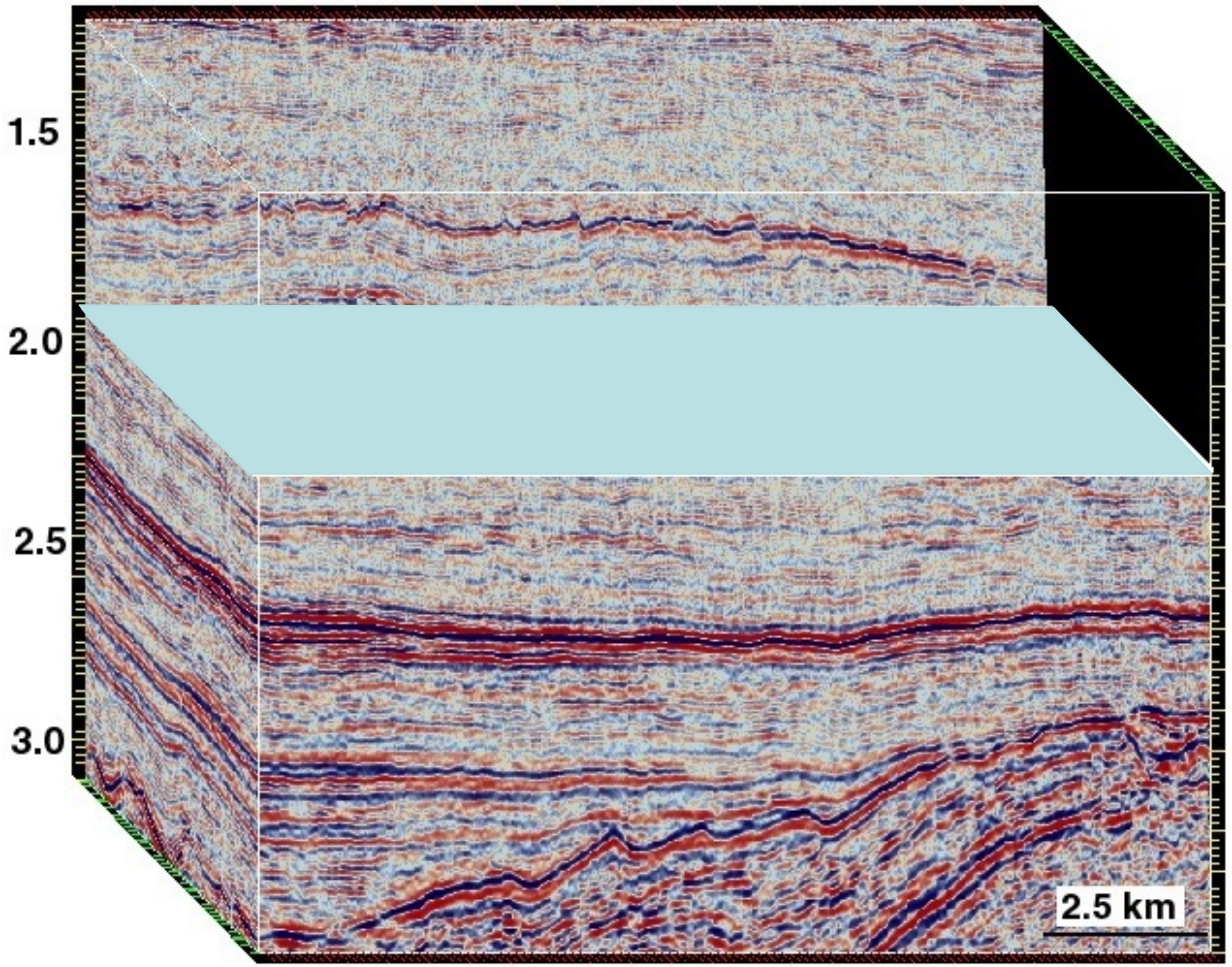
This timeslice (left) and vertical seismic section (right) come from the same dataset in the Gulf of Mexico. Note the development of normal faults and mapped seismic horizon (yellow)



3D seismic: How do we view it?

Understanding of the relationship between features observed in vertical and map-view (timeslice) seismic sections can be aided by using 3D visualisation packages. **We are going to use a package called *Mayavi in Ex 3 and 4.***

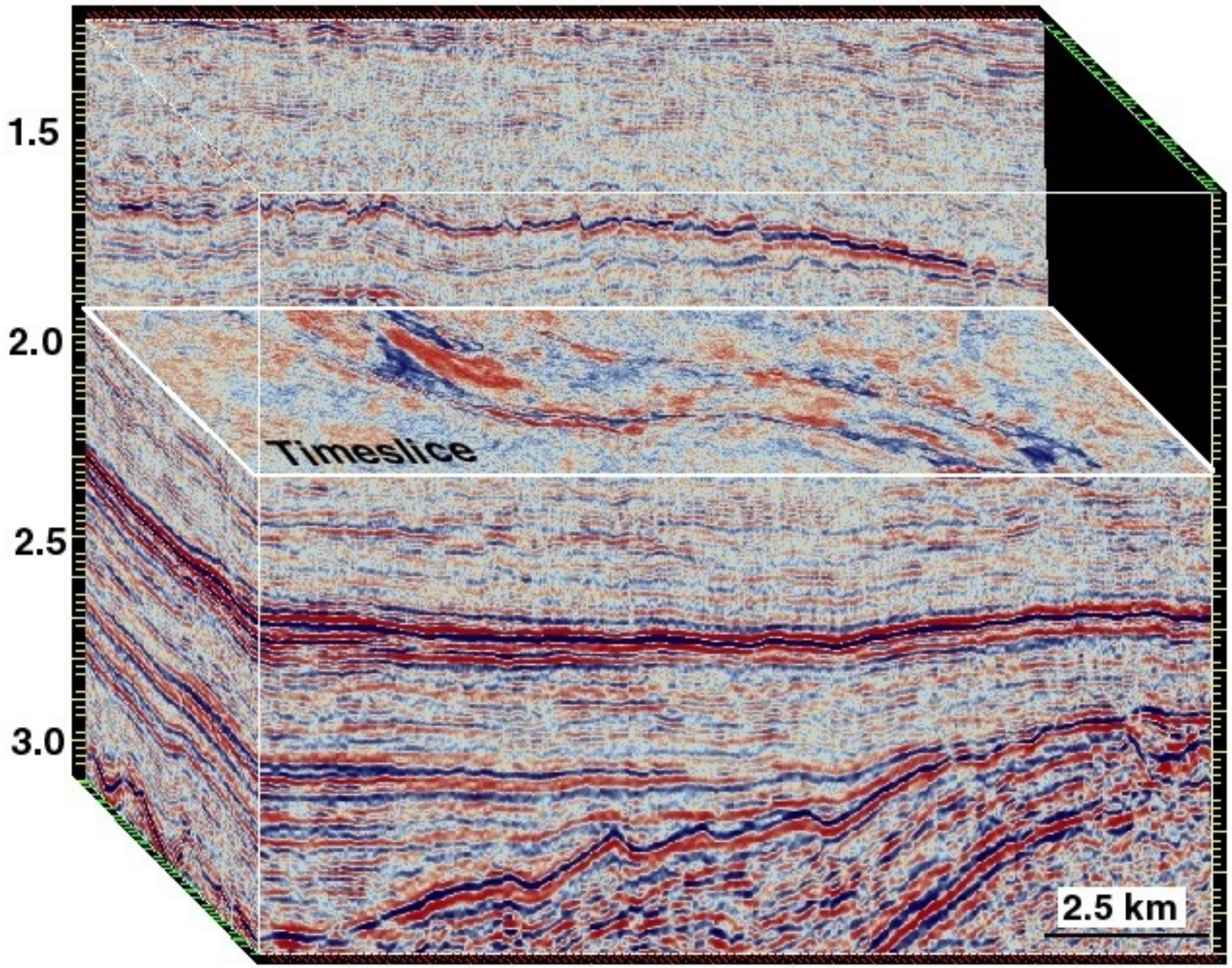
In the example to the right from the Northern North Sea it is difficult to identify channels in the vertical section, but they are clear in the horizontal time slice.



3D seismic: How do we view it?

Understanding of the relationship between features observed in vertical and map-view (timeslice) seismic sections can be aided by using 3D visualisation packages. **We are going to use a package called *Mayavi in Ex 3 and 4.***

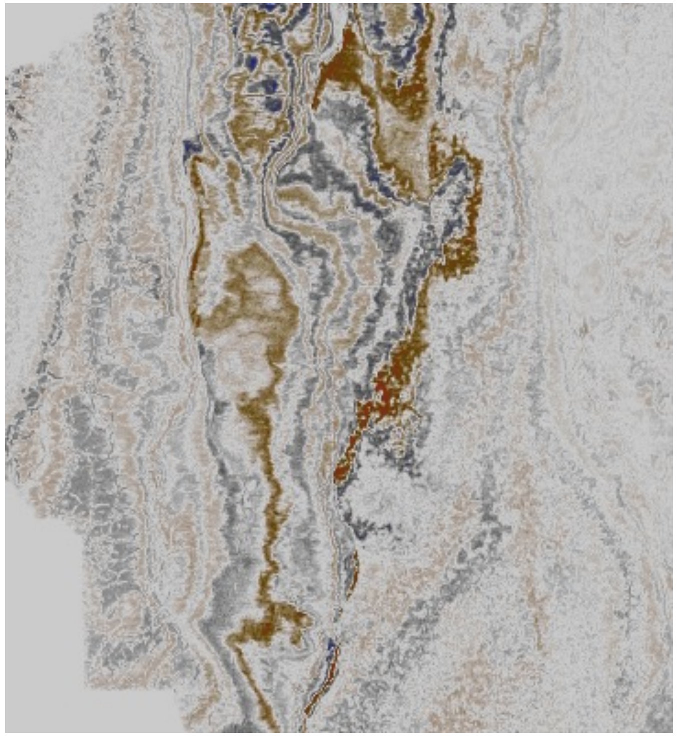
In the example to the right from the Northern North Sea it is difficult to identify channels in the vertical section, but they are clear in the horizontal time slice.



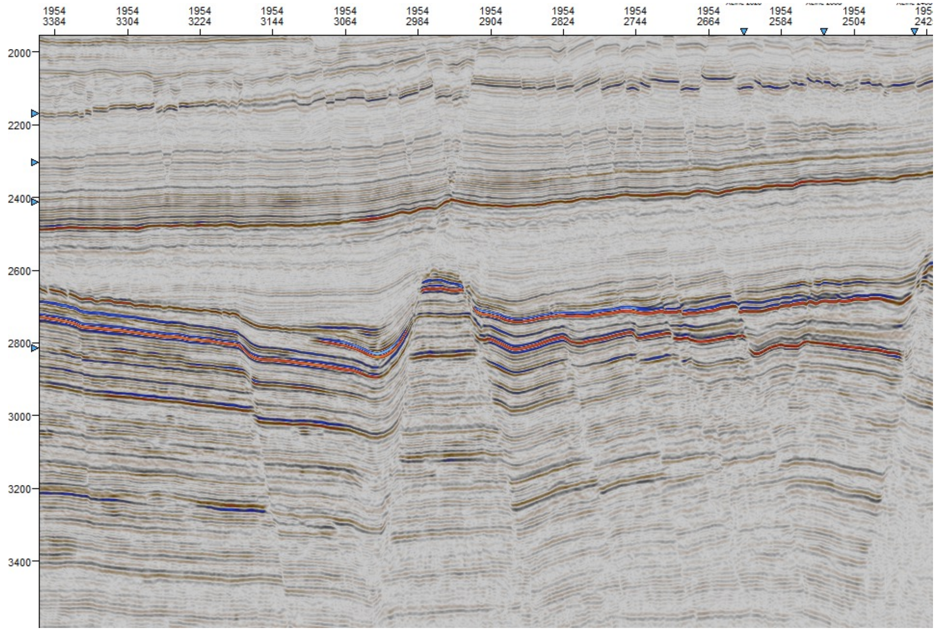
Quiz question

Which of these images shows an inline?

A



B



3D seismic: How do we view it?

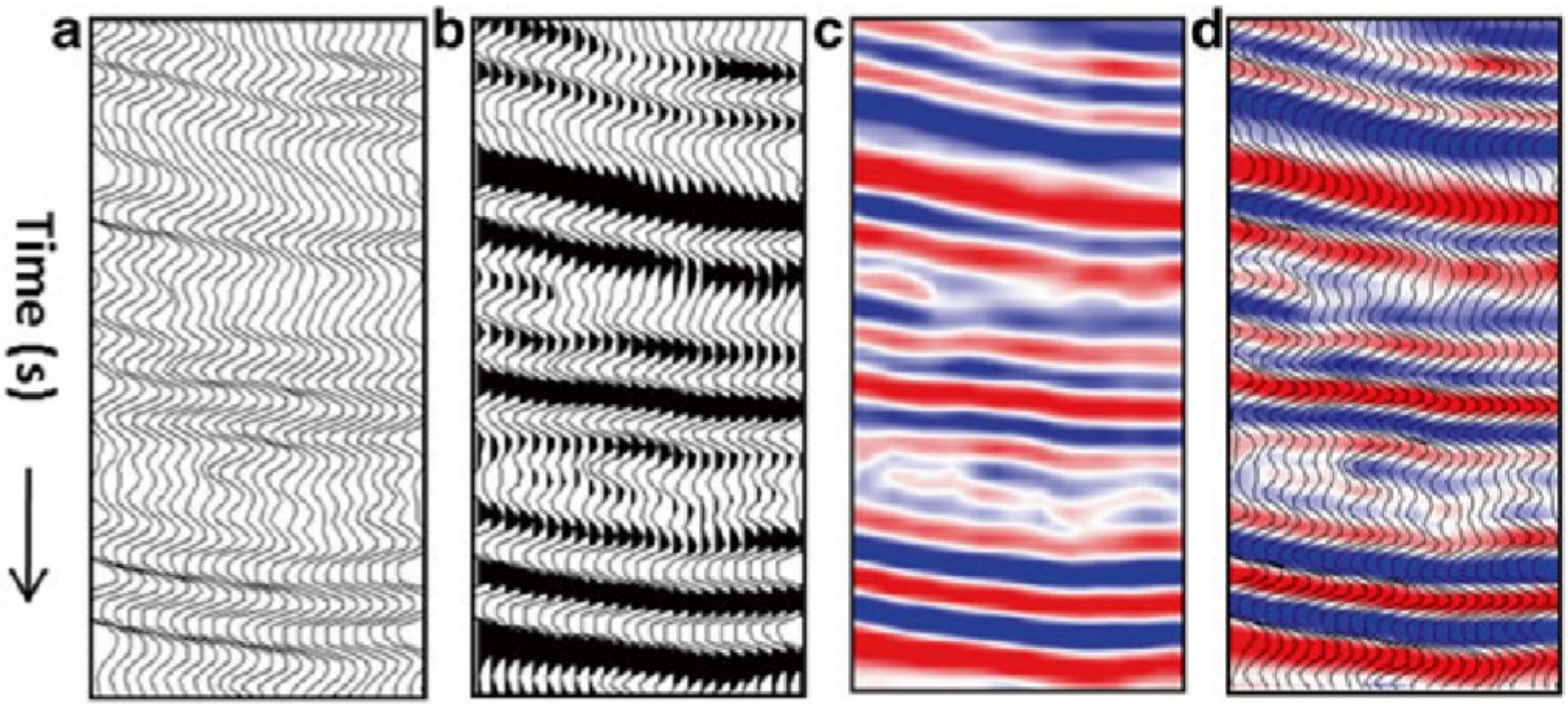
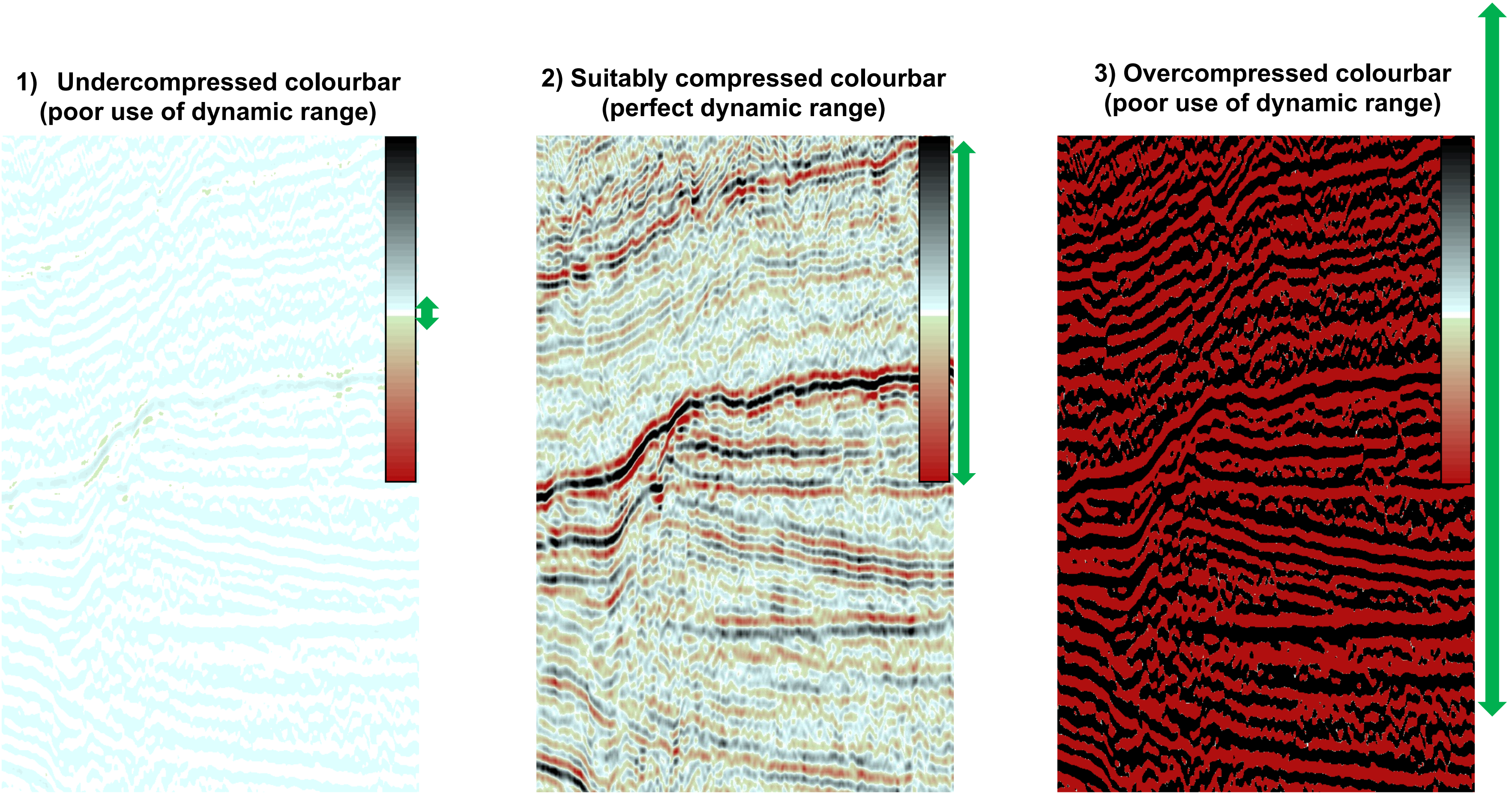


Figure 6: Types of seismic data display modes: (a) Wiggle. (b) Wiggle and variable area. (c) Variable density. (d) Combination of (a) and (c).

Niranjan 2016

Dynamic range



It is important to use the dynamic range of the data in the colour bar to display seismic data well

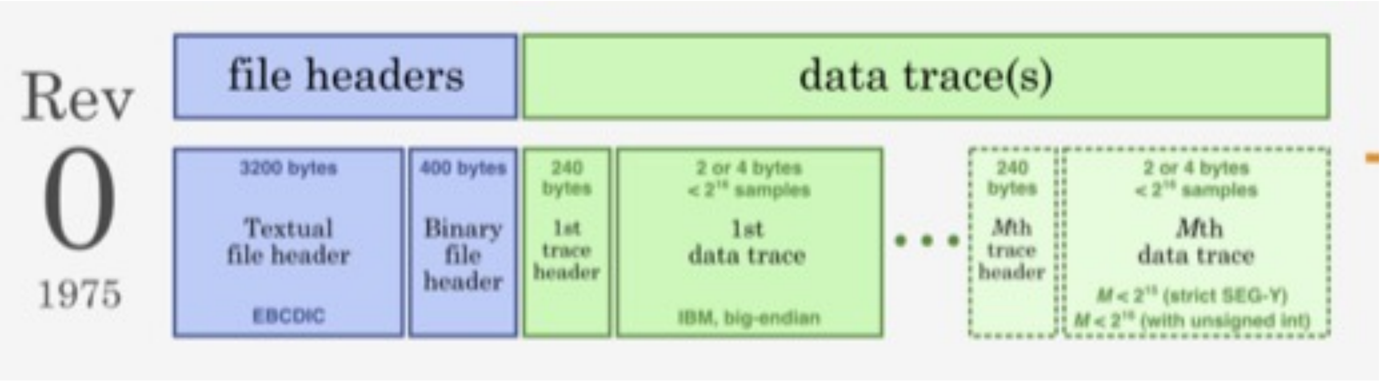
You will explore this issue in Ex 3

How are seismic data stored?

- 3D Post-stack seismic data are effectively a grid of samples. As a data scientist you may think this is quite a simple data type to store
- However, seismic data storage has a long history dating back to 1975 when the SEGY file format was established for saving single-line seismic data to magnetic tapes (needed to make sure all the metadata stored with the trace data)
- We still use the same file-format today, although it has been updated periodically
- Not the most efficient file format for modern data science
- In Exercise 3 the first thing you will do is convert SEGY data into a more efficient format (e.g. a numpy array)

The SEGY file format

- Developed by the Society of Exploration Geophysicists (SEG) released in 1975 for storing seismic data in a consistent way throughout industry
- Built for reading and writing data to tapes, not for efficiently viewing inlines, xlines and time slices
- The SEG-Y file mostly involves the trace data, but it also includes a number of headers: A text header, a binary header, a header for each trace



https://www.troika-int.com/about-us/resources/segy-seismic-data-exchange-format

The SEGY file format

Text header: known as the EBCDIC. Human-readable header describing the data (bit like a ReadMe file). Contains information on acquisition, processing and geometry.

Summary | Text Header | Bin Header | Trace Header | Trace Data

```
C 1 CLIENT                         COMPANY                        CREW NO
C 2 LINE TAN0106_LINE21_MIG_NZTM_CDP_AMPLITUDES MAP ID
C 3 REEL NO            DAY-START OF REEL     YEAR      OBSERVER
C 4 SHOTPOINT:    5        LONG       SHOTPOINT FOR 2D AND 3D SURVEY DATA
C 5 LINE NUMBER:  17        LONG       LINE RECORD
C 6 TRACE NUMBER: 25        LONG       TRACE RECORD(CDP Number)
C 7 SHOTPOINT:   21        LONG       SHOTPOINT FOR ARBITRARY LINE
C 8 CDP_X:         73        LONG       CDPBIN NUMBER
C 9 CDP_Y:         77        LONG       CDPBIN NUMBER
C10 NSAMPLE:      115 UNSIGNED SHORT  NUMBER OF SAMPLES
C11 SRATE:        117 UNSIGNED SHORT  SAMPLE RATE
C12
C13 IHS CRS Conversion - Apply scalar:Yes - Scalar Bytes: 71-72
C14   From: NZGD2000 / New Zealand Transverse Mercator 2000
C15   To:    WGS 84 / UTM zone 60S
C16   Bytes: (73-76 -> 73-76, 77-80 -> 77-80)
C17
C18
```

Challenges for data scientists

- There are a number of tools available to convert SEGY data into a format which is more efficient for viewing inlines, xlines, time slices and for doing calculations with it. We are going to use ***segyio*** in Exercise 3

- The size of SEGY data makes it difficult to work with on an individual workstation. In Exercise 3 we will work on a 500 Mb subset of an 8Gb data volume (TNW_small2.segy)

- People are working on web API's to host data in the cloud and allow people to access parts of the data quickly and easily (e.g. oneseismic)

- Data loading problems will occur if the data is not a perfect cube and inline/xline spacing not regular. You can explore this in the bonus exercise and this may be an issue if you download data from an open-source repository

- Not all SEGY files will contain all the information you expect (the metadata quality will vary between files…)

How can we access seismic data?

- Seismic data is often proprietary when first collected
- After a moratorium period academics and industry are often required to make the data open access (some countries are great at insisting on this- e.g. Geoscience Australia)
- There is no single centralized location to access these datasets
- The SEG wiki have a good summary of many repositories for open SEGY data- https://wiki.seg.org/wiki/Open_data
- The Marine Geoscience Data System is a good option for academic datasets

Exercise 3

- Explore the SEGY file format
- You will load and visualize 3D SEGY data in terms of inlines and xlines
- You will visualise the data in 3D using Mayavi
- Fix data loading issues (in the Bonus exercise…)

Key points

- Appreciate how 3D seismic data are commonly visualized (inlines, xlines, horizontal slices, arbitrary lines)
- Understand that seismic data in stored in the SEGY file format involving various header information as well as the trace sample data
- Know some open-access repositories for downloading free SEGY data
- Appreciate the key challenges of working with SEGY data (primarily, huge file sizes)