DeepTrace: A.I. for First-Break Picking

Branton Demoss
Front Range Geosciences
Boulder, CO

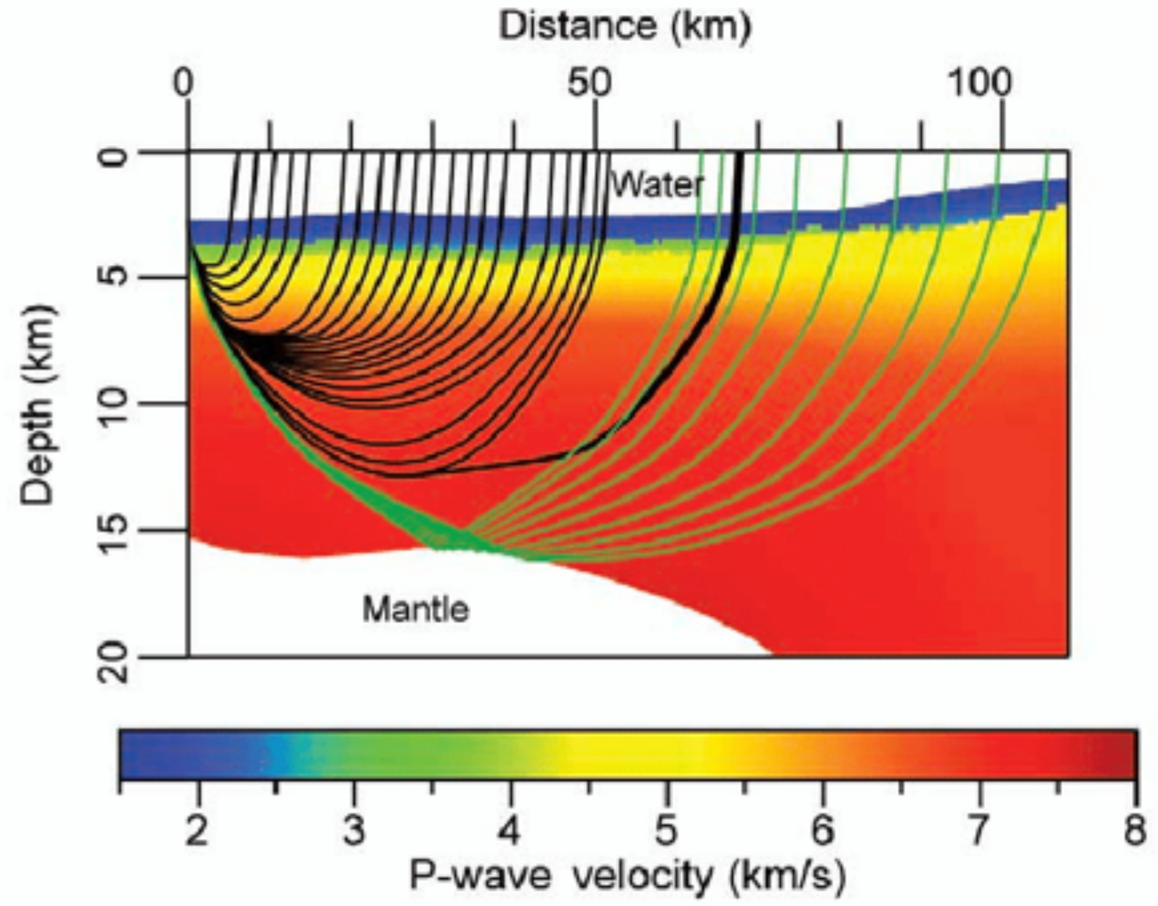
Roadmap

- Motivation and Problem
- History of Automatic First-Break Picking
- History of Machine Learning
- Hardware
- DeepTrace Results
- Phoenix
- Quantifying Confidence and Reliability

Background

why do we need first-break picks?

Near-Surface Modeling

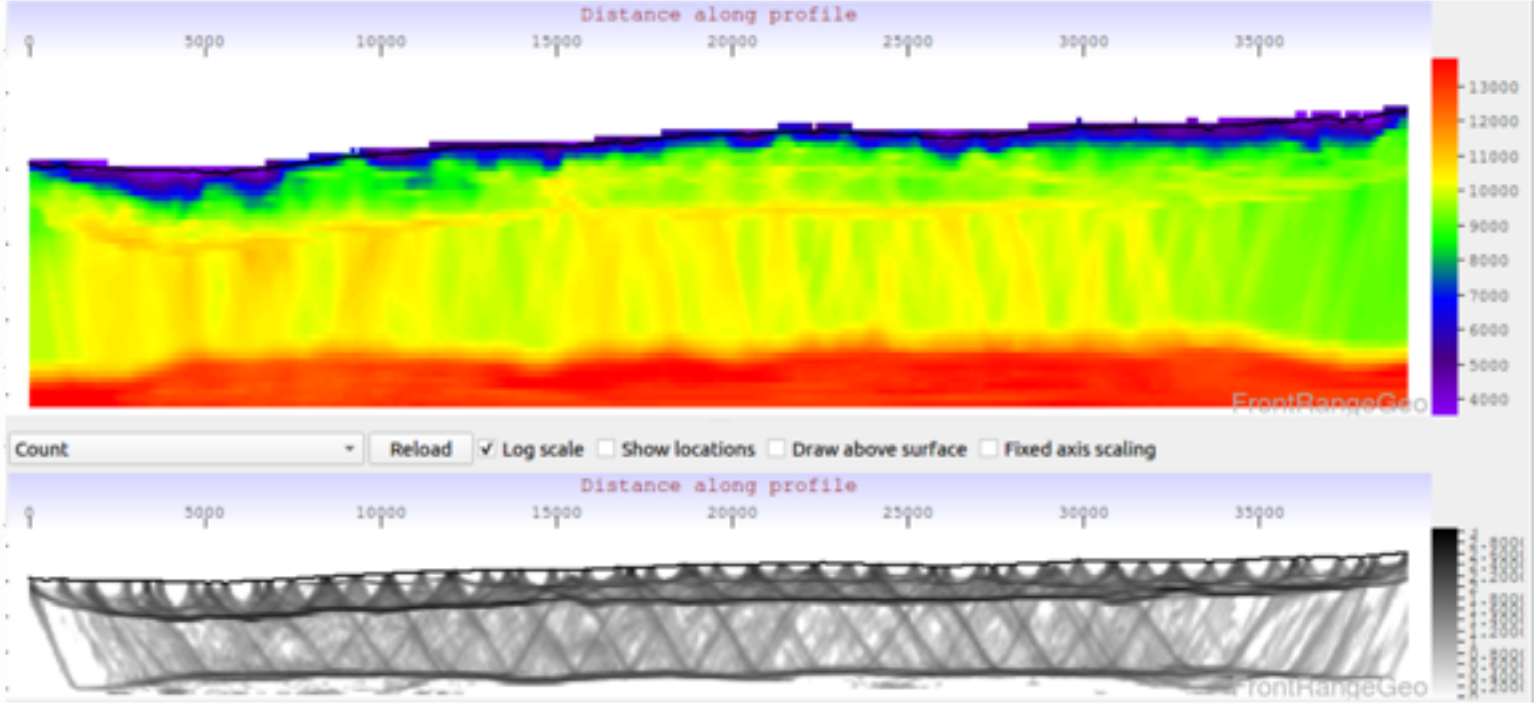


Seismic waves propagate through the earth, producing a seismic record (trace) at each receiver.

Pick the first arrival in seismic record, then invert along ray path to find a velocity model of the earth.

Lau et. al. *Low-frequency source for long-offset, sub-basalt and deep crustal penetration*. The Leading Edge. 2007

Near-Surface Modeling



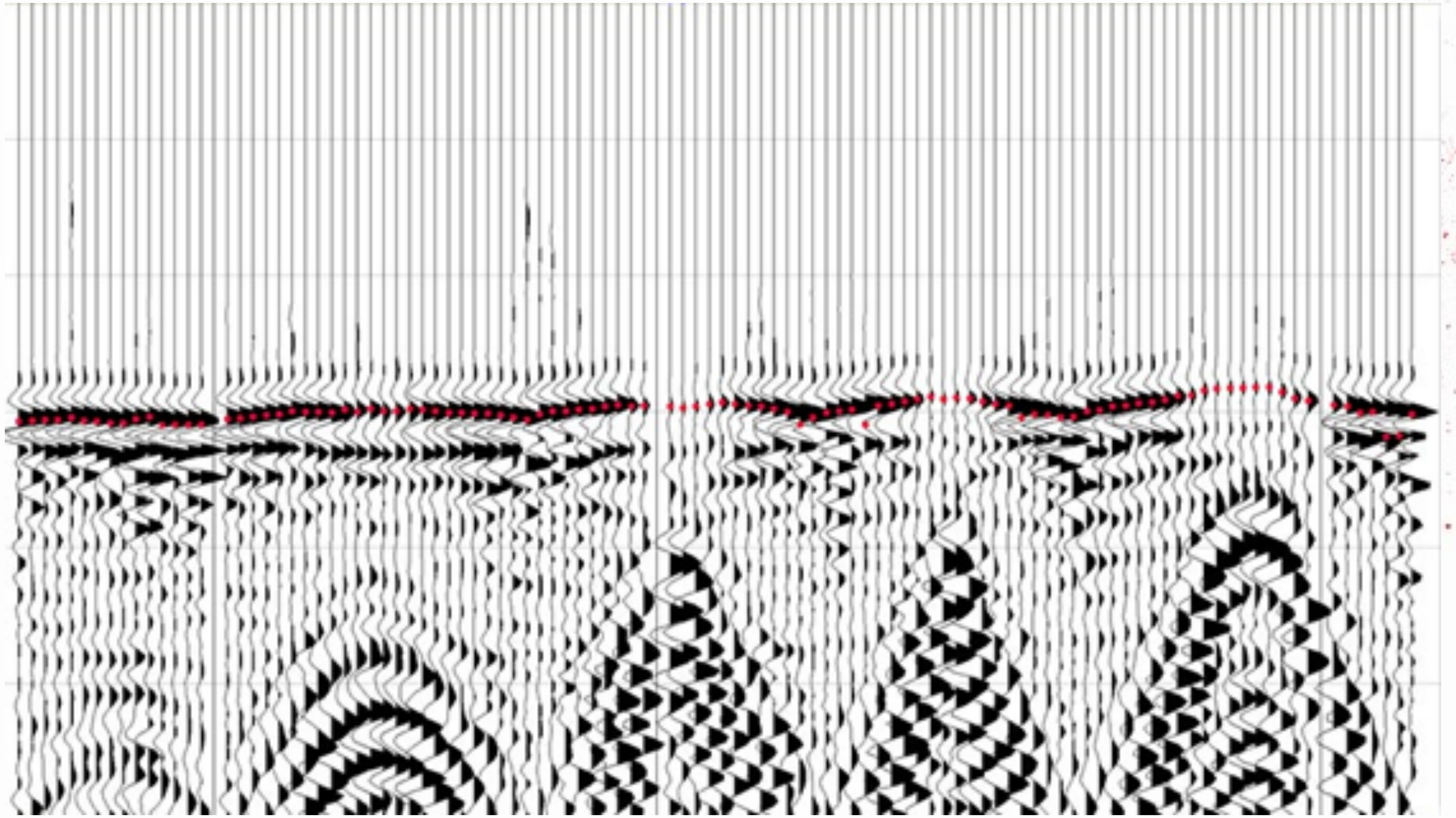
By comparing simulated travel times with picked arrival times, we can iteratively update a tomographic model.

Upper panel: profile of computed velocity field. Lower panel: simulated node hit counts (Teapot Dome Survey)

Static Correction

Final near-surface model used to correct for weathering layer.

Without static shift from near-surface, deeper imaging will have velocity anomalies.



Shot record with tomographic solution applied. Image courtesy XtremeGeo.

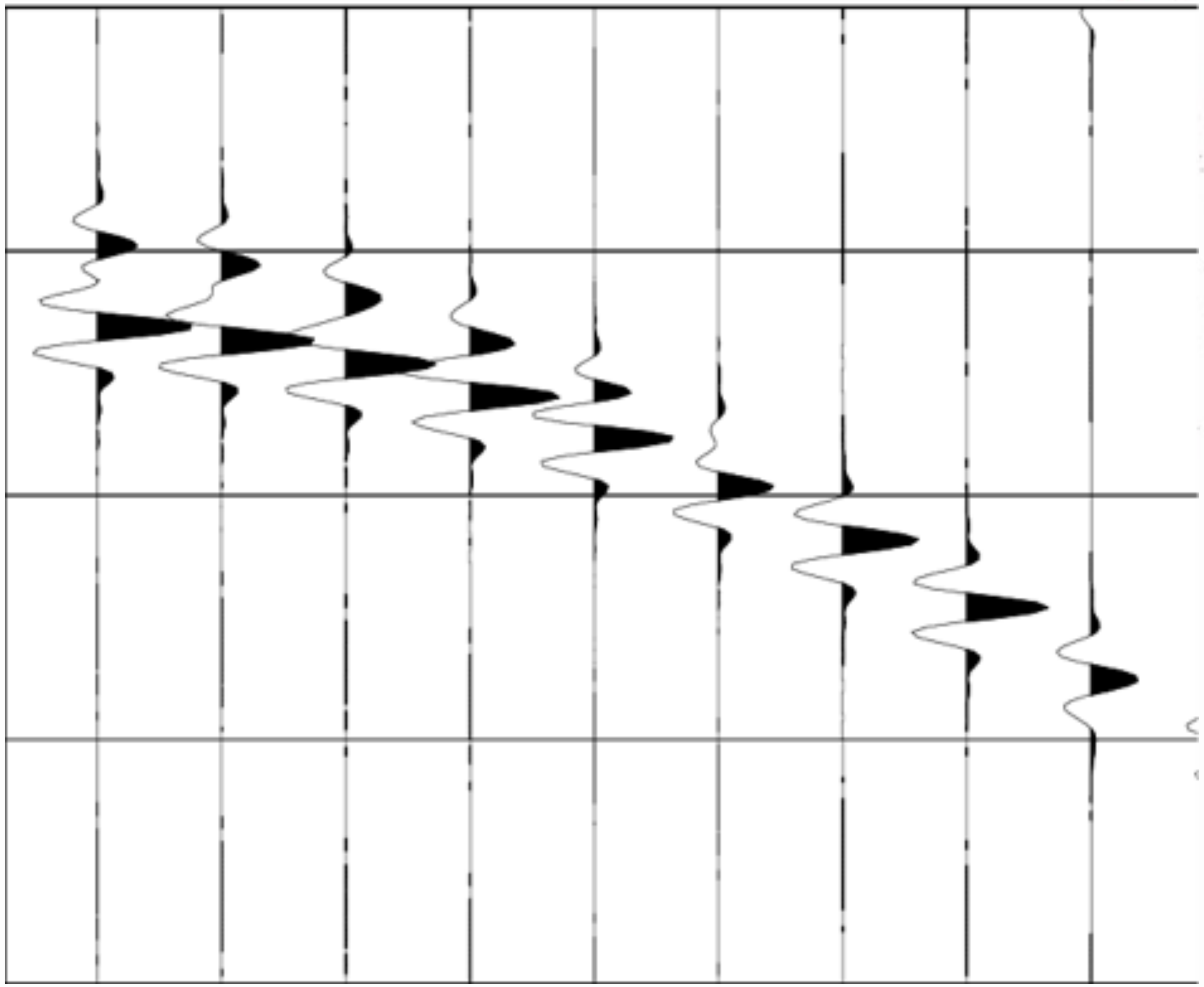
Seismic Traces

Clean traces have easy-to-spot arrival times.

Or do they? What is the arrival time of a wavelet? Is such a notion well-defined?

Peak, trough, zero crossing?

Consistency matters more than any specific event.

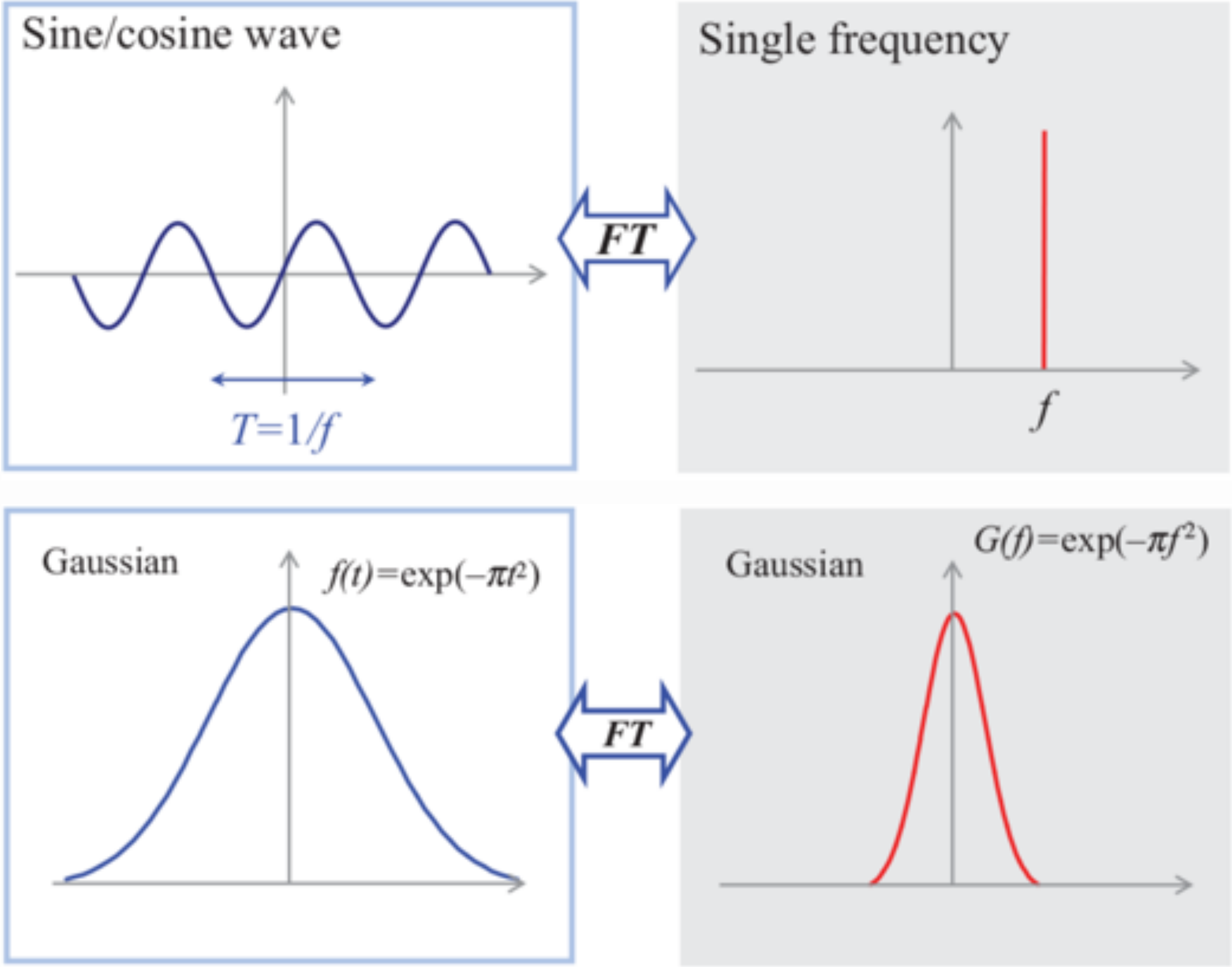


Waves

Seismic waves are spread out in time and frequency domains.

More correct model of arrival time is a *distribution.*

Different component frequencies of initial wavelet are dampened differently.



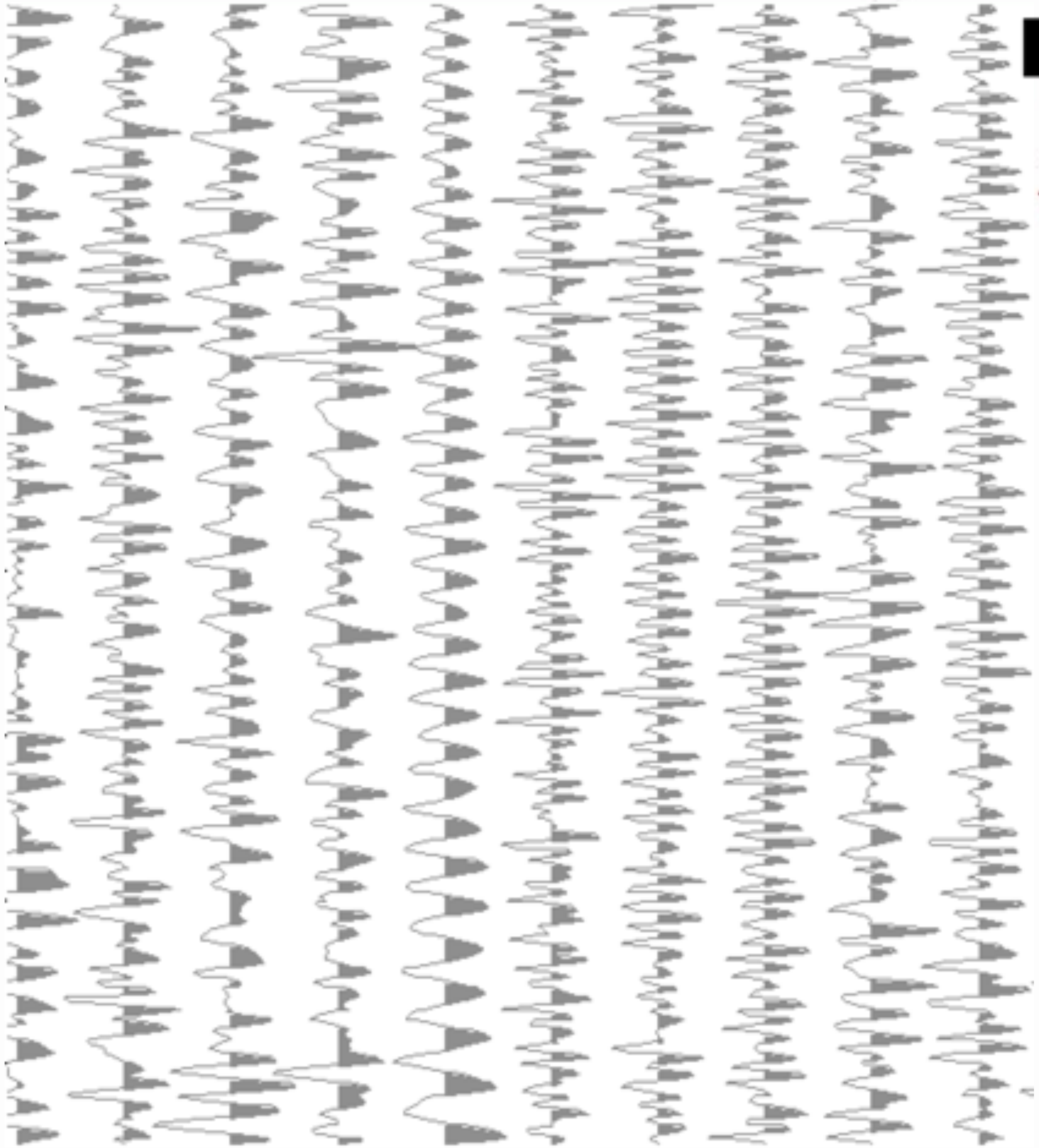
Noisy Seismic

The problem is even worse in the presence of noise.

Not only are wavelet arrival times inherently ambiguous, noise can completely obscure signal.

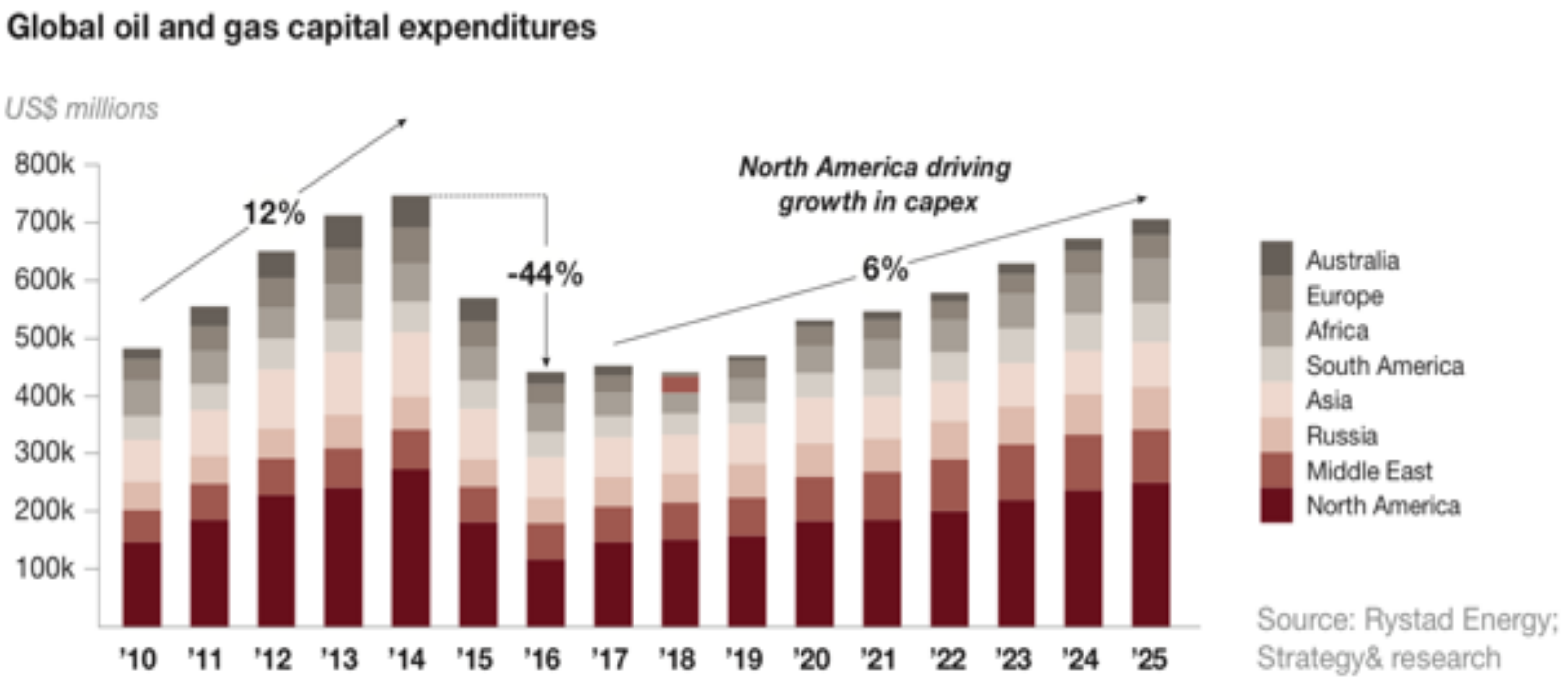
In difficult conditions, humans must label arrival times ‘by hand’.

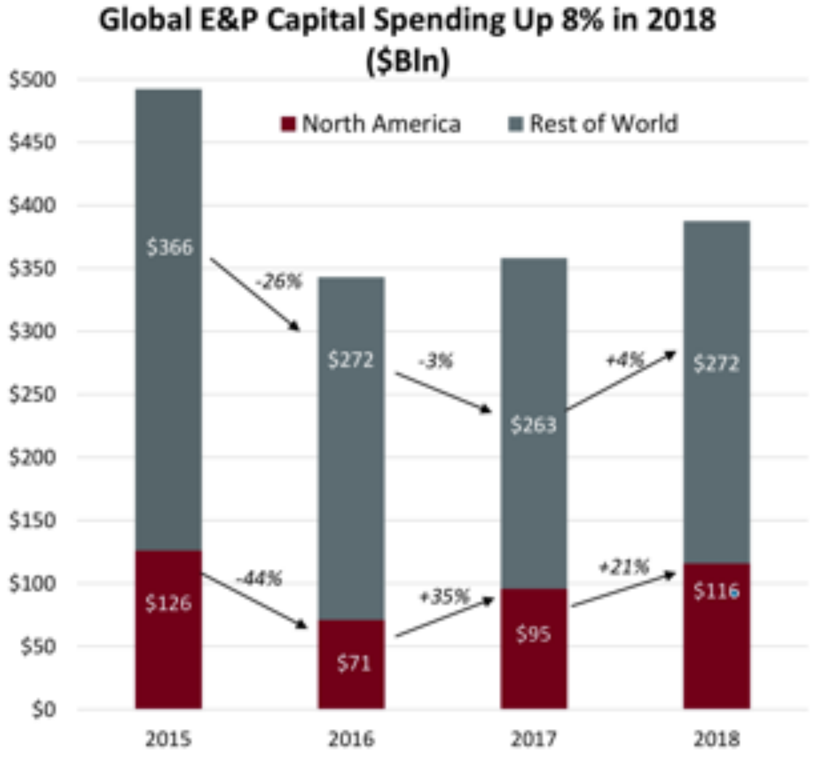
This is incredibly time consuming.



Unedited real seismic data

Seismic survey market will grow from $7.54B in 2017 to $9.28B by 2022, driven by investment in North America





1) Strategy& Research. "Oil and Gas Trends 2018-19: Strategy Shaped by Volatility." (2018).
"Seismic Survey Market by Service (…), and Region - Global Forecast to 2022." (2017). MarketsandMarkets™;
2) "Barclays Global 2018 E&P Spending Outlook." (2018). Currie.
"Impact of Technology and Innovation on Global Oil & Gas Markets." (2018).

Automatic First-Break Picking

a brief history

Historical Perspective

“Experience has shown that manual processing of […] refraction records takes a disproportionate length of time in comparison with the surveys themselves, and this is incompatible with the requirements for choosing the site of an exploration well. It thus became necessary to find an ‘industrial approach’ to the solution of this processing problem.”

- Peraldi and Clement:
Digital Processing of Refraction Data - Study of First Arrivals.
Geophysical Prospecting. ***1972***

DIGITAL PROCESSING OF REFRACTION DATA
STUDY OF FIRST ARRIVALS *

BY

R. PERALDI ** AND A. CLEMENT ***

1972 - Cross Correlations

Geophysical Prospecting 33, 1212-1231, 1985.

FIRST ARRIVAL PICKING ON COMMON-OFFSET TRACE COLLECTIONS FOR AUTOMATIC ESTIMATION OF STATIC CORRECTIONS*

F. COPPENS**

1985 - Energy Ratio

A fractal-based algorithm for detecting first arrivals on seismic traces

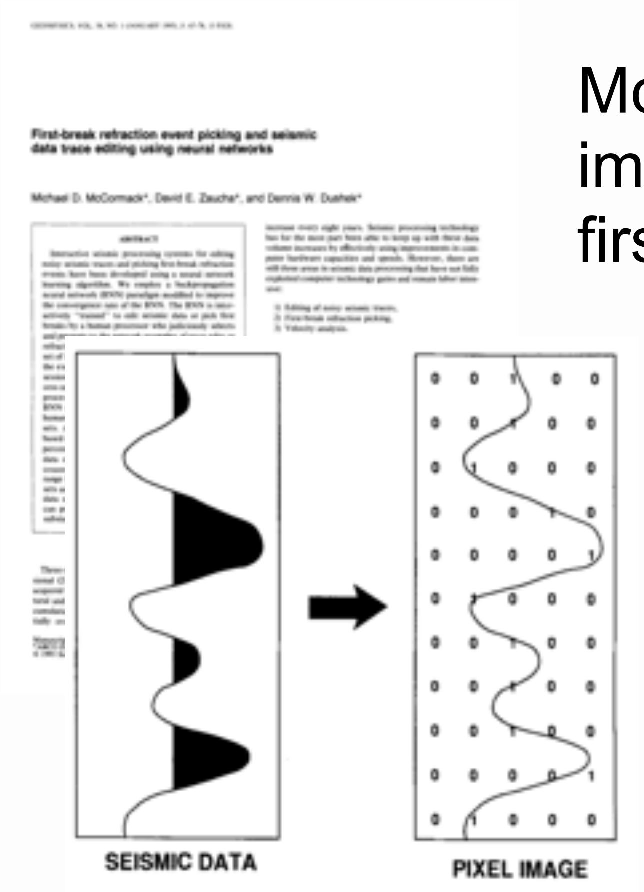
Fabio Boschetti*, Mike D. Dentith‡, and Ron D. List**

1996 - Fractal Dimension

...

- A progression of more and more sophisticated methods.
- All claim method works as well as humans in noisy areas.
- Many note that their method ‘naturally mimics the human eye’.

Neural Networks Appear



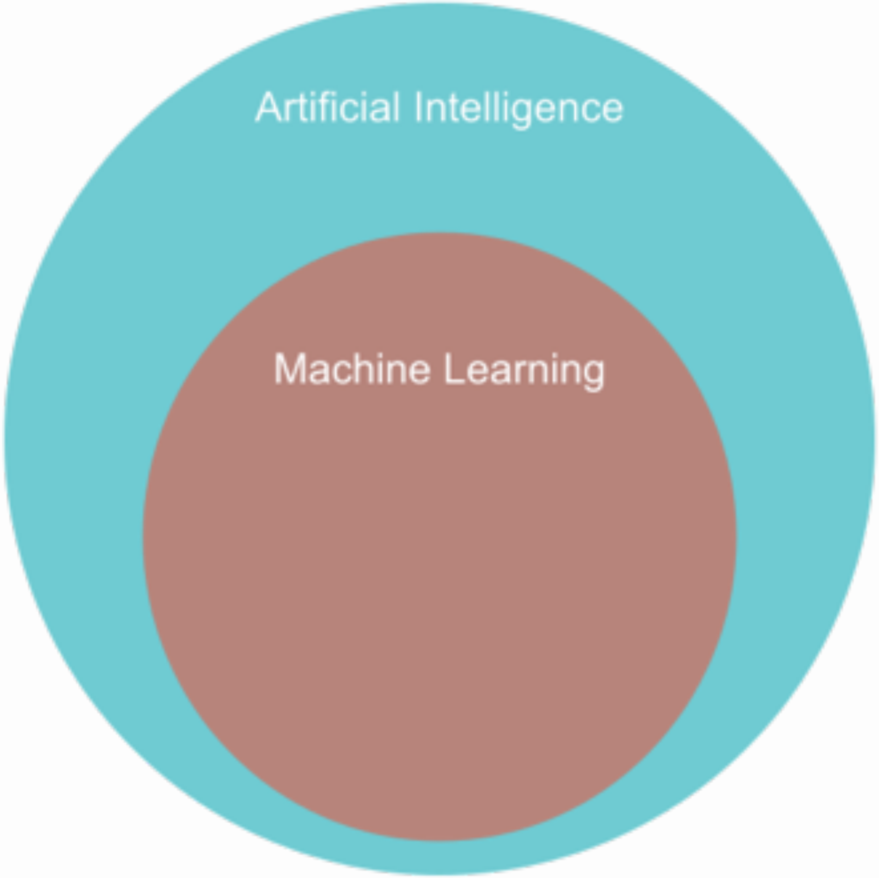
McCormack et al. (1993) describe the first implementation for a neural network to automate first-break picking.

Subsequent papers focus on the 'feature engineering' approach to neural network classification.

Machine Learning

a primer

A.I. vs Machine Learning

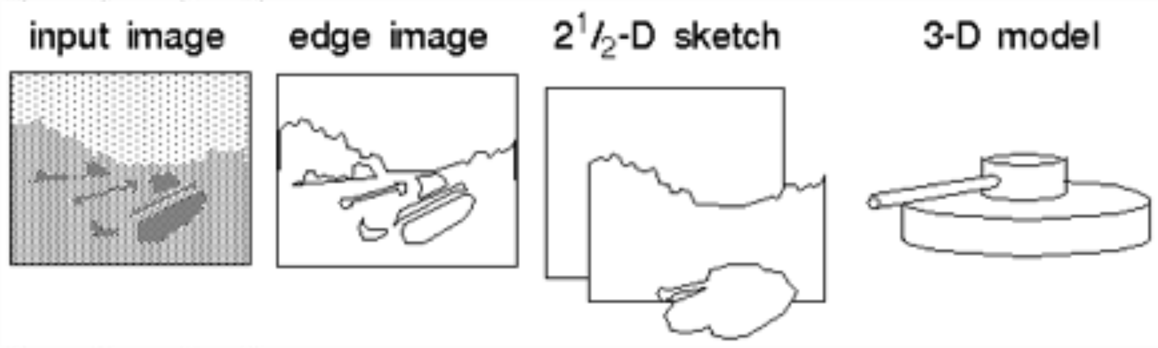


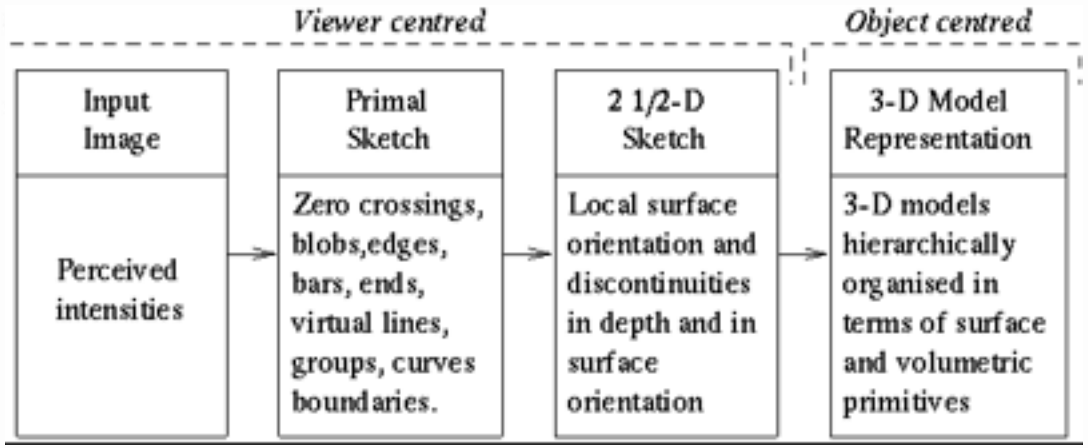
Working definitions:

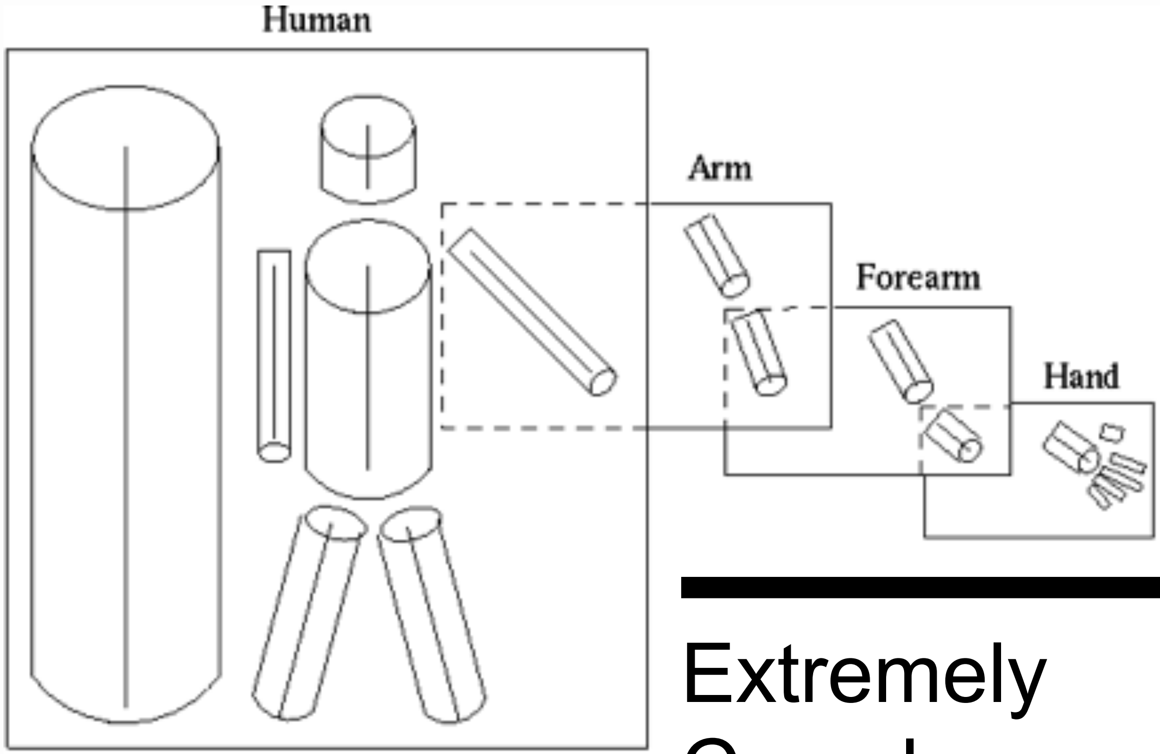
A.I. - Making computers perform tasks that humans traditionally perform well.

Machine Learning - An approach to A.I. where machines are given data, and come up with a mathematical model to fit the data “on their own”.

Old A.I.:
Classical Visual Recognition Stack



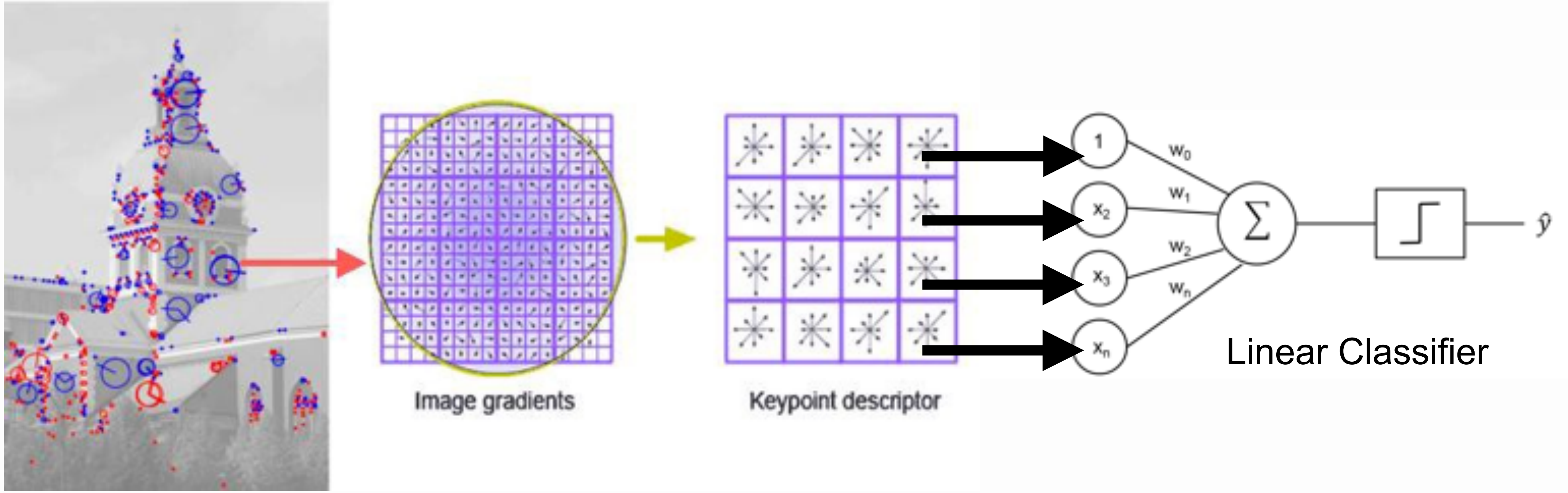




Images from David Marr's *Vision*. 1982

Extremely Complex.

Old A.I.: Feature Extraction

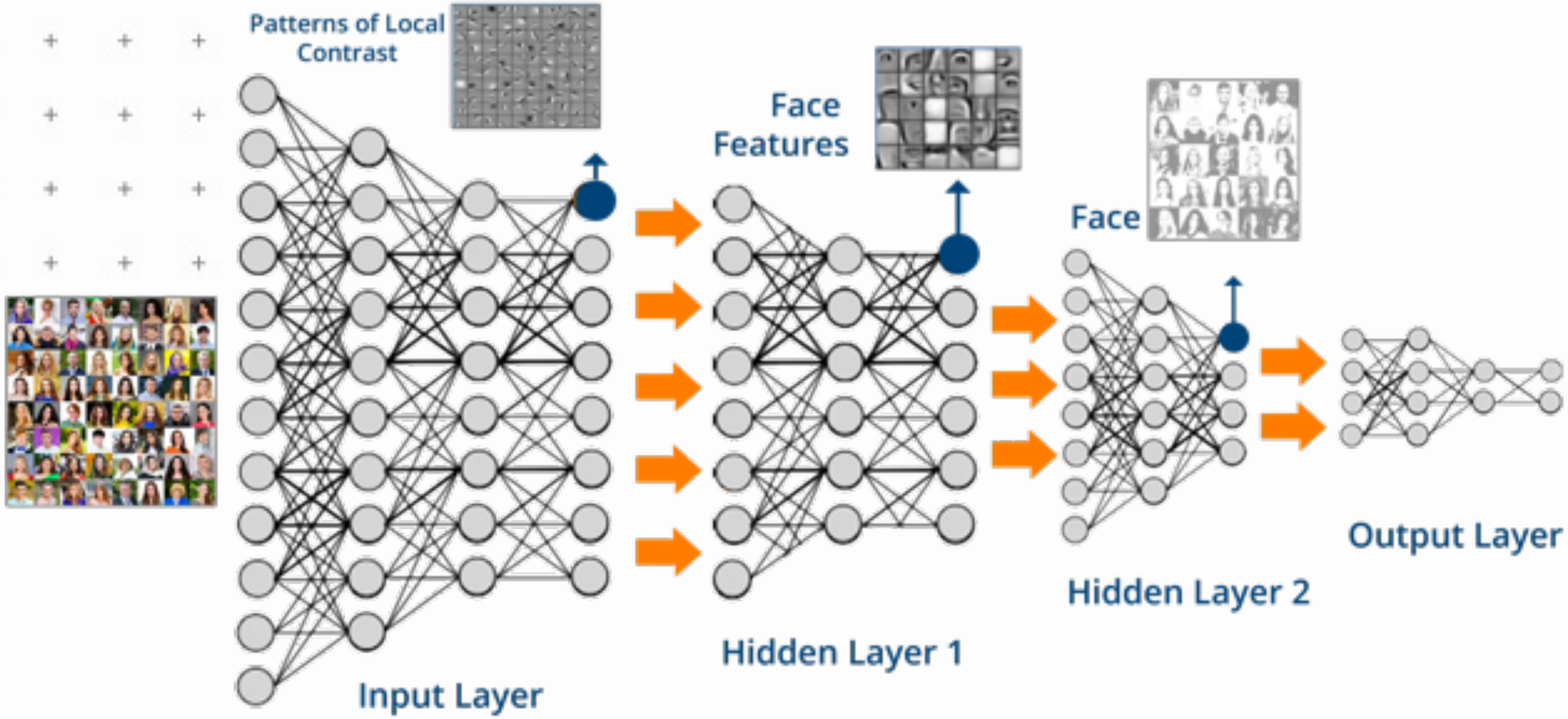


Instead of developing entire recognition stack, develop a few key features, and do statistics with them. Sometimes *many* features, extremely large state vectors.

Some recent first break picking papers focus on this (Hollander et.al 2018).

I tried this approach myself.

Modern A.I.: Deep Learning



Forgo features entirely, let A.I. learn its own representation.

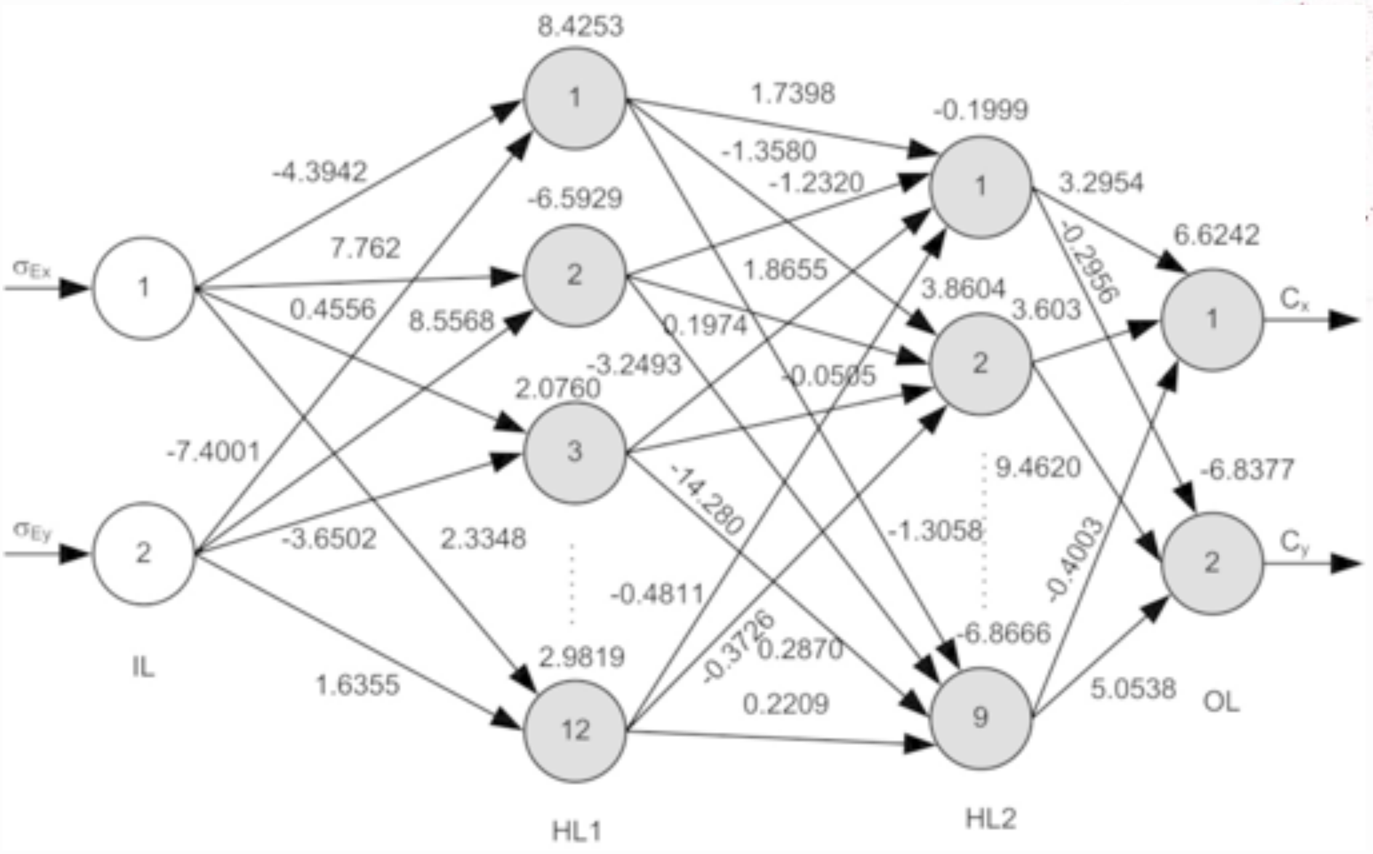
Deeper layers tend to correspond to more abstract features.

Implemented using a neural network.

Neural Networks

Neural Networks are:

- An example of a machine learning algorithm.
- Nonlinear functions (given an input, produce unique output).
- Randomly initialized, so start off knowing absolutely nothing.
- Adjusted via back-propagation and stochastic gradient descent (SGD).

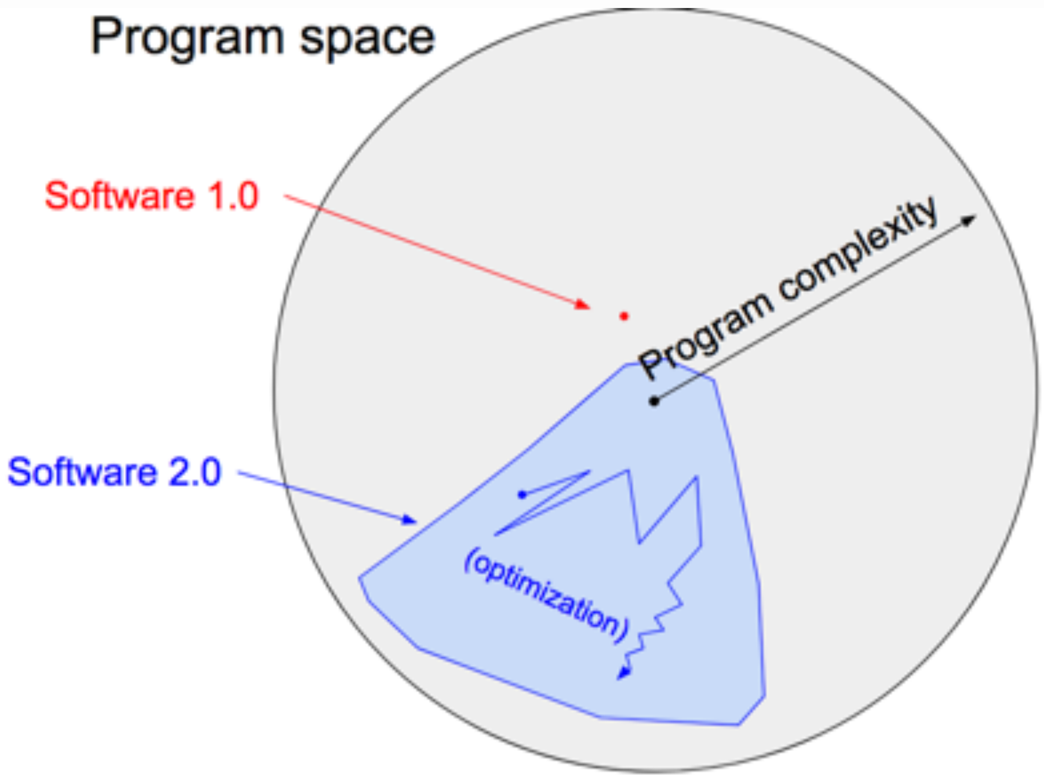


Andrej Karpathy. *Software 2.0.* Medium. 2017

Programming 2.0

In Programming 1.0 we identify a point in program space with desirable behavior by writing lines of code.

In 2.0 we specify the search space and let the optimizer find the best program (as represented by a neural network).



Andrej Karpathy. *Software 2.0*. Medium. 2017

Hardware

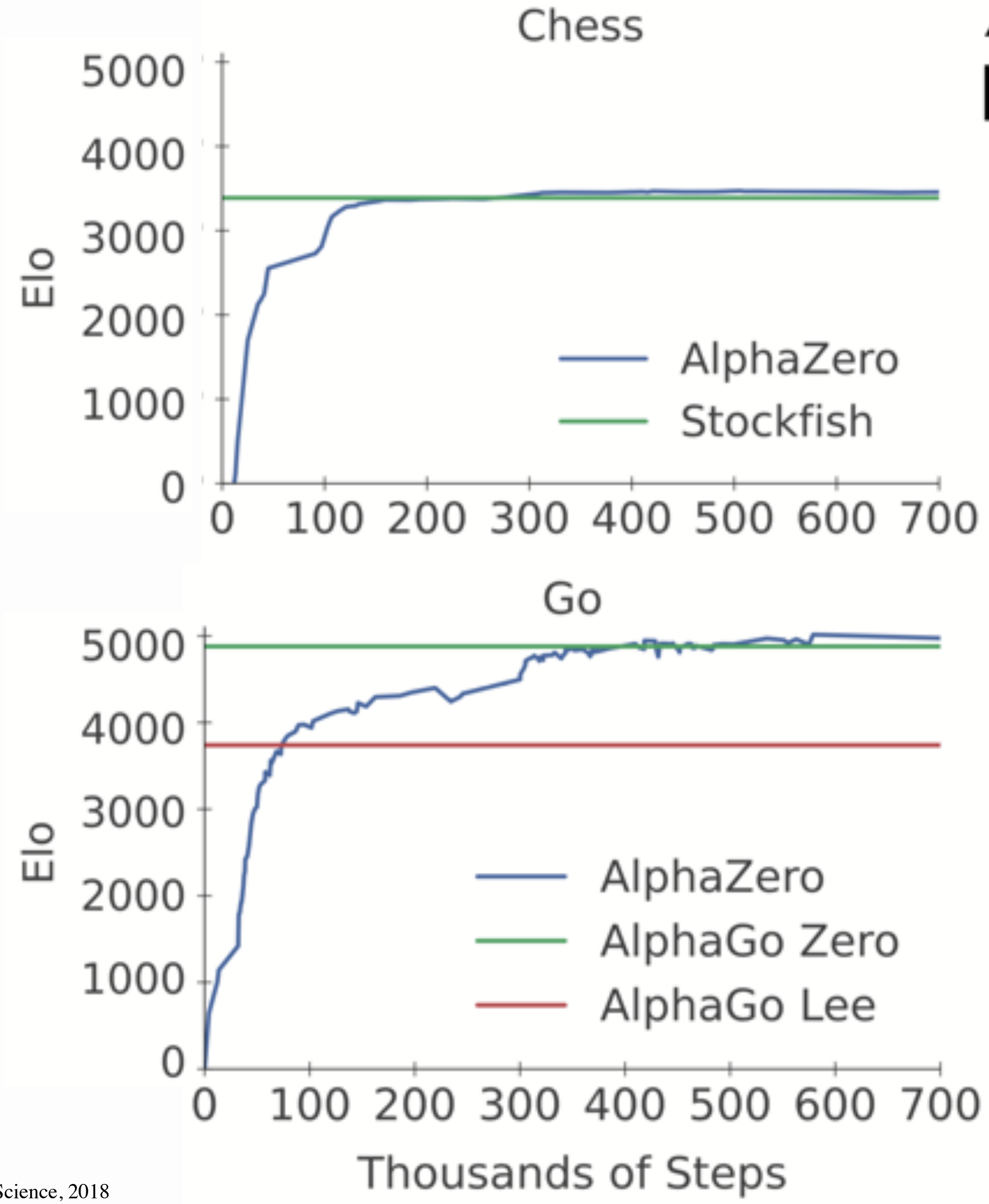
the fuel for modern machine learning

Performance Scales with Data

Generic deep learning result:

More data = higher accuracy and better generalization.

ImageNet, AlphaGo.

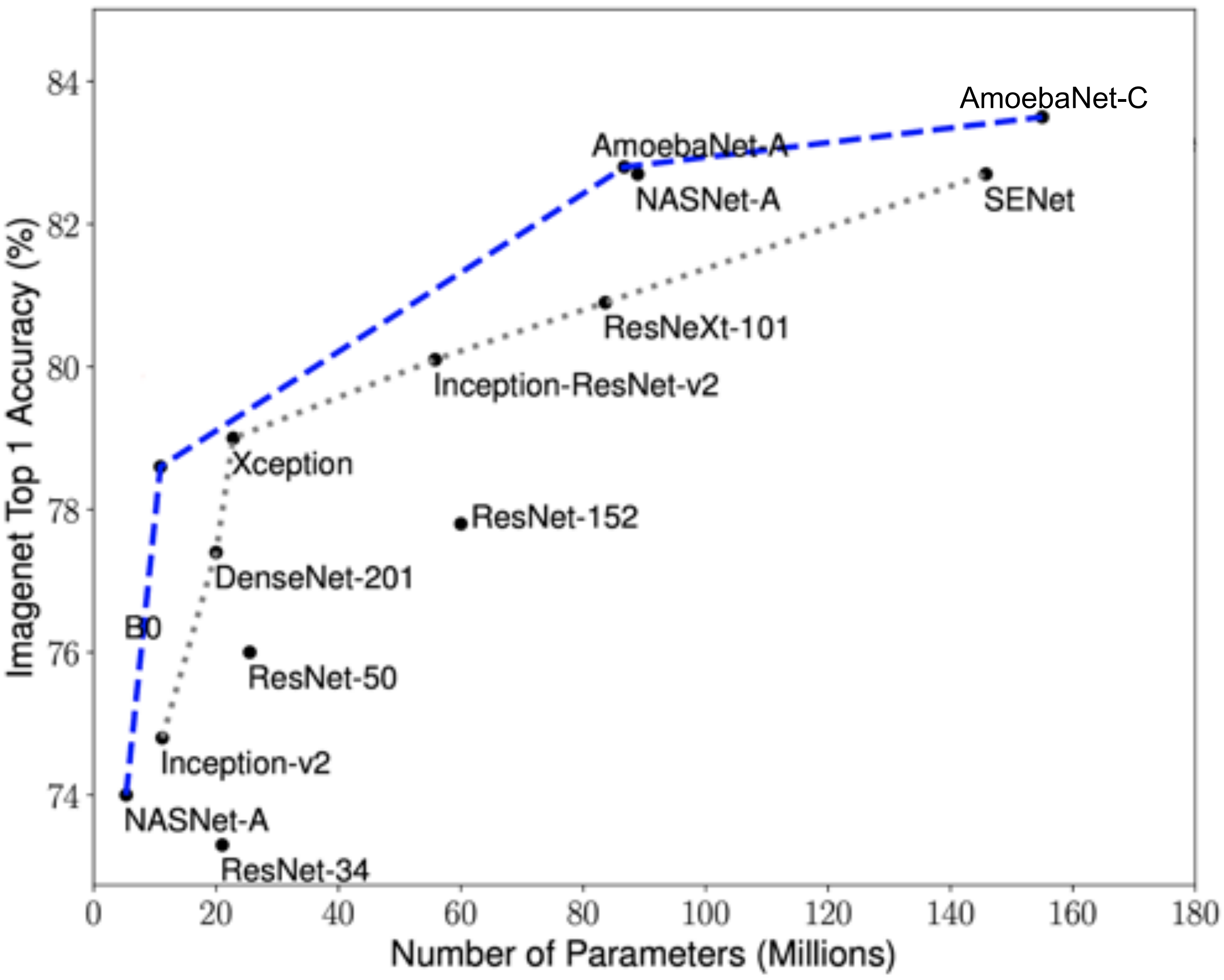


Silver et. al. *A general reinforcement learning algorithm that masters chess, shogi and Go through self-play*. Science, 2018

Performance Scales with Model Size

Neural nets aren't new, we just haven't had the computing and data scales to make them useful until recently.

The network size needed to get good image recognition results is around 10^7 neurons.

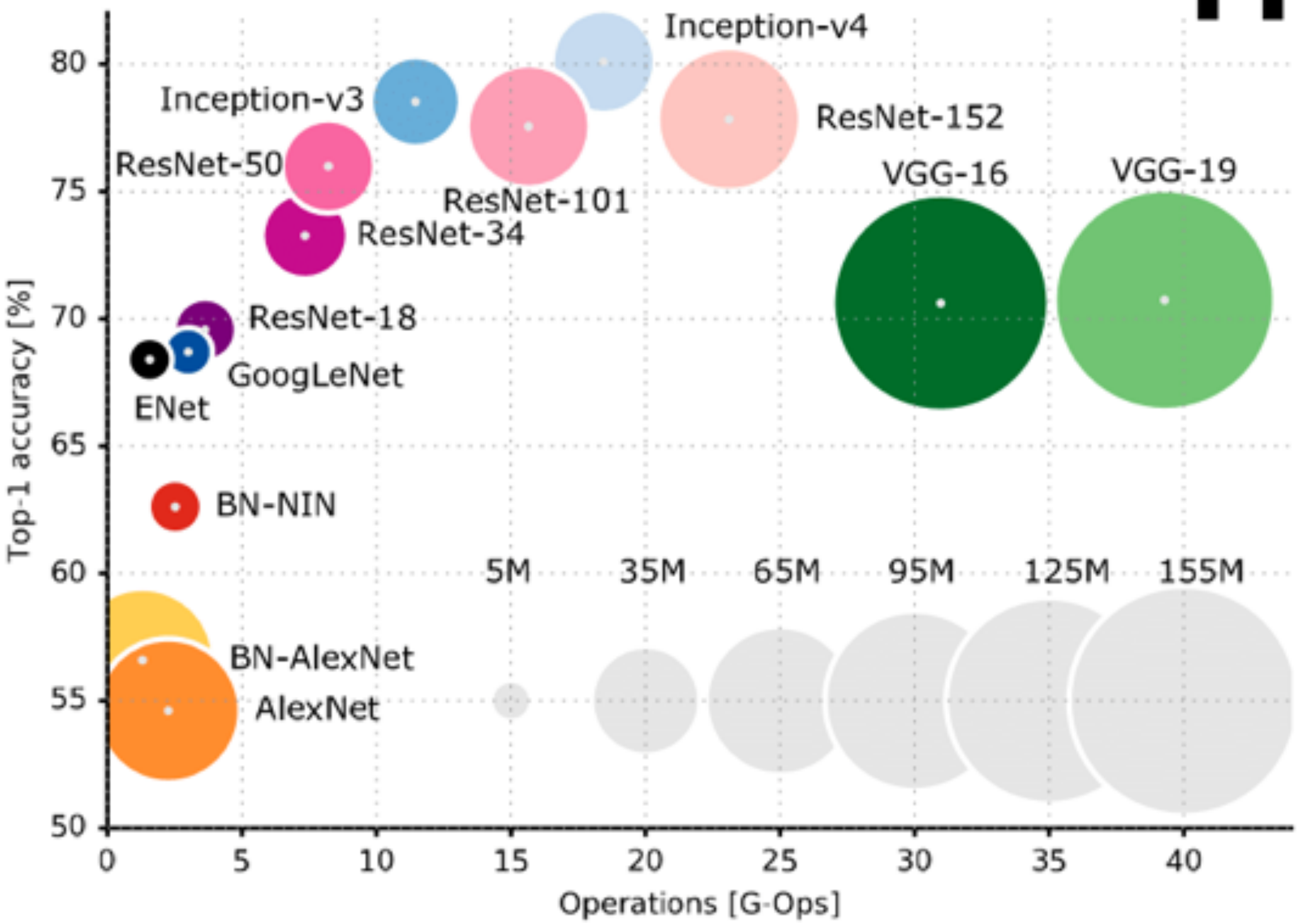


Canziani et. al. *An Analysis of Deep Neural Network Models for Practical Applications.* 2016

GPU Compute

GPUs can do many more operations in parallel than comparably priced CPUs.

Modern GPUs enable the FLOPS necessary to predict on modern large surveys in reasonable time, using deep neural networks.

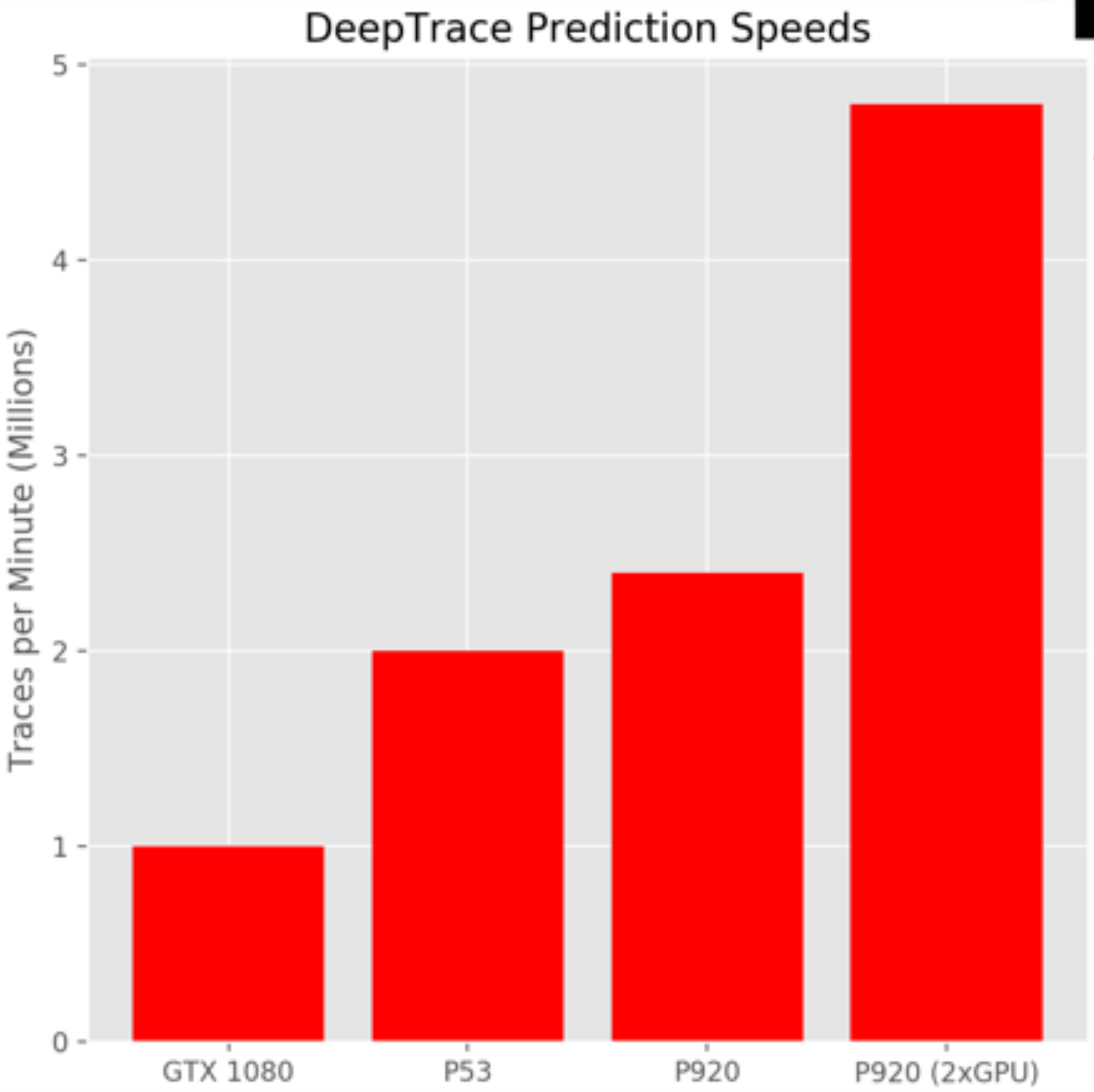


Canziani et. al. *An Analysis of Deep Neural Network Models for Practical Applications.* 2016

Lenovo Benchmarks

Running the fastest DeepTrace models, we can predict on ~2.5 million traces/minute.

With a modest 6 GPU cluster, we can get human-level predictions at nearly 1 billion traces per hour.

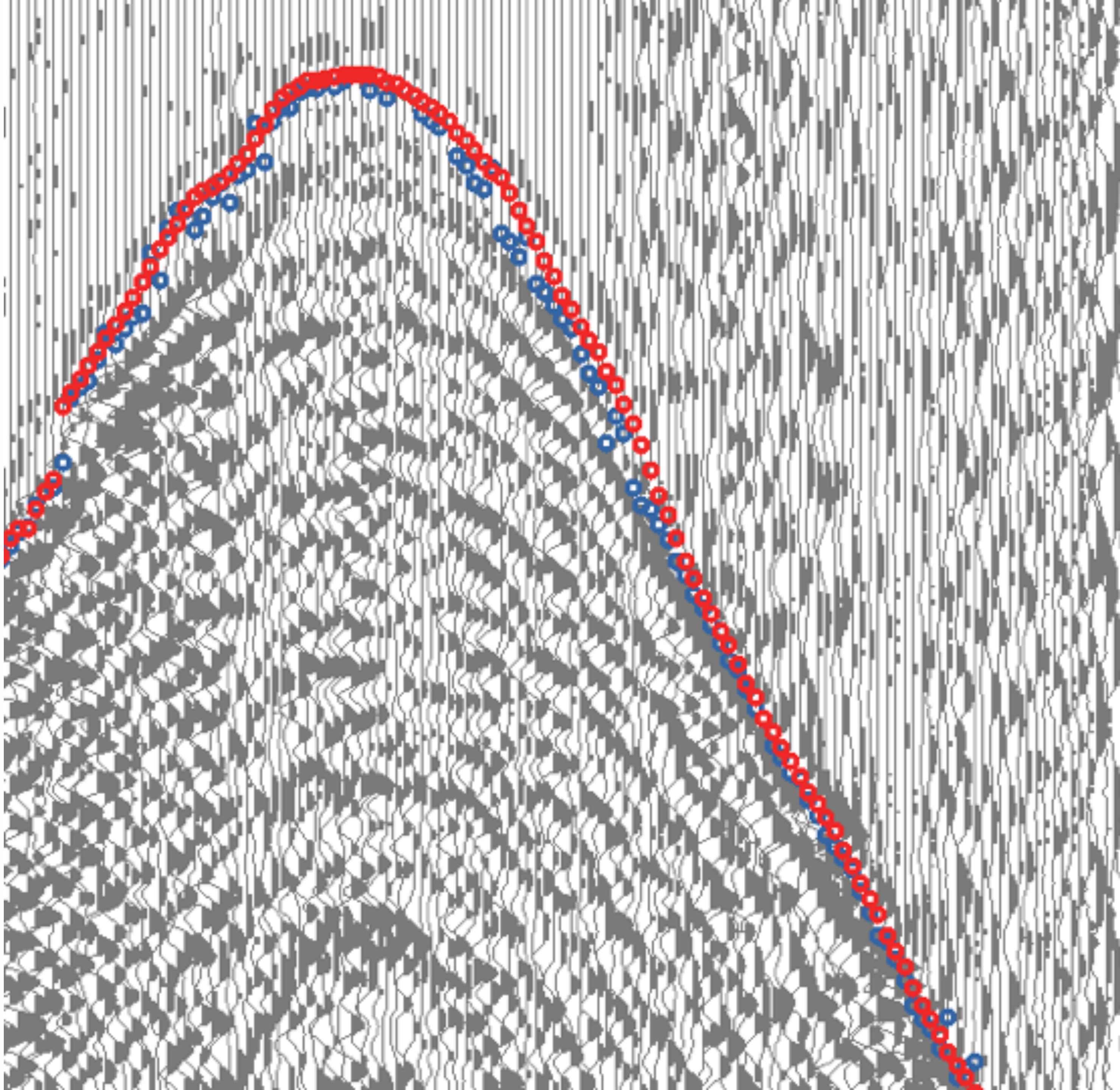


DeepTrace Results

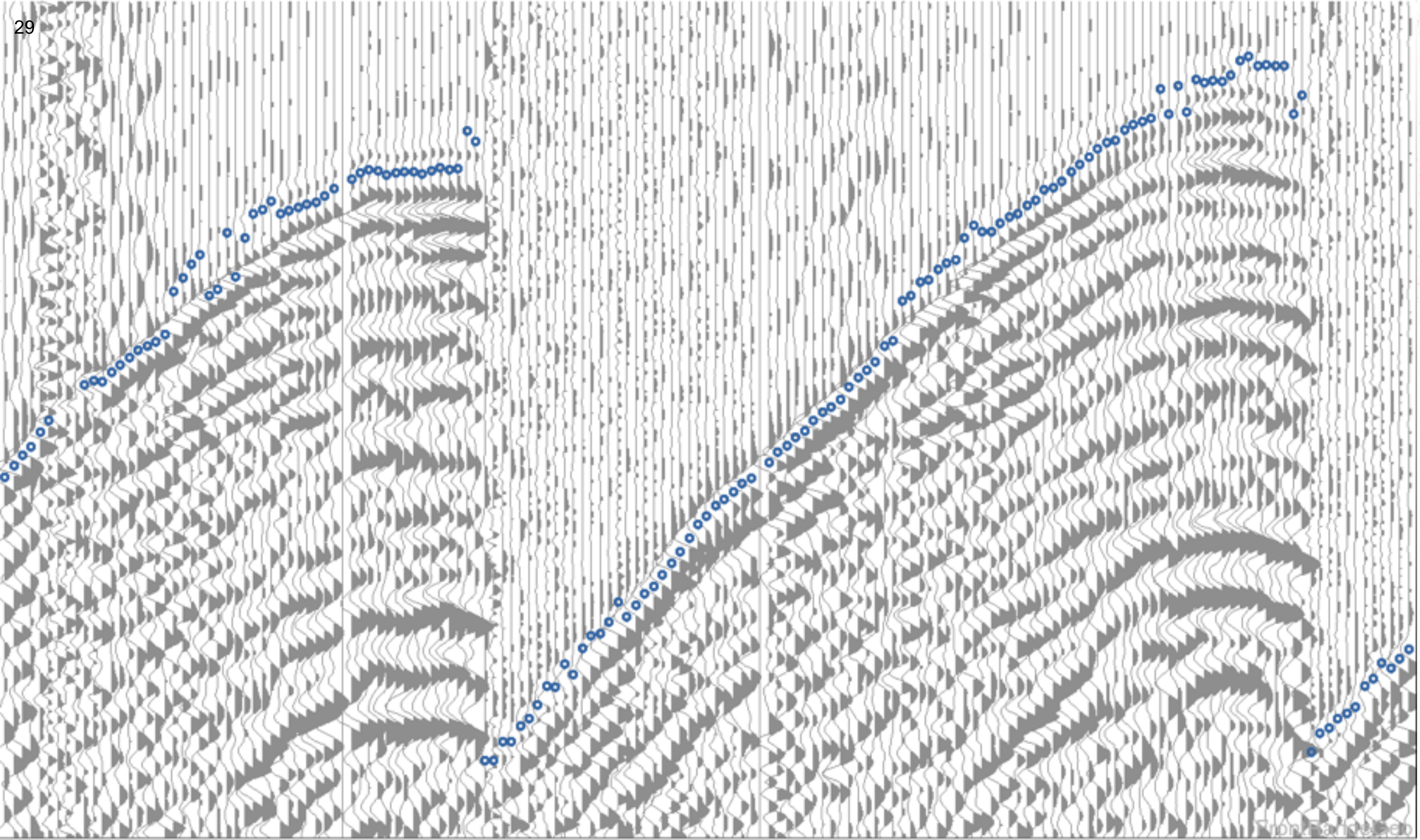
in collaboration with XtremeGeo

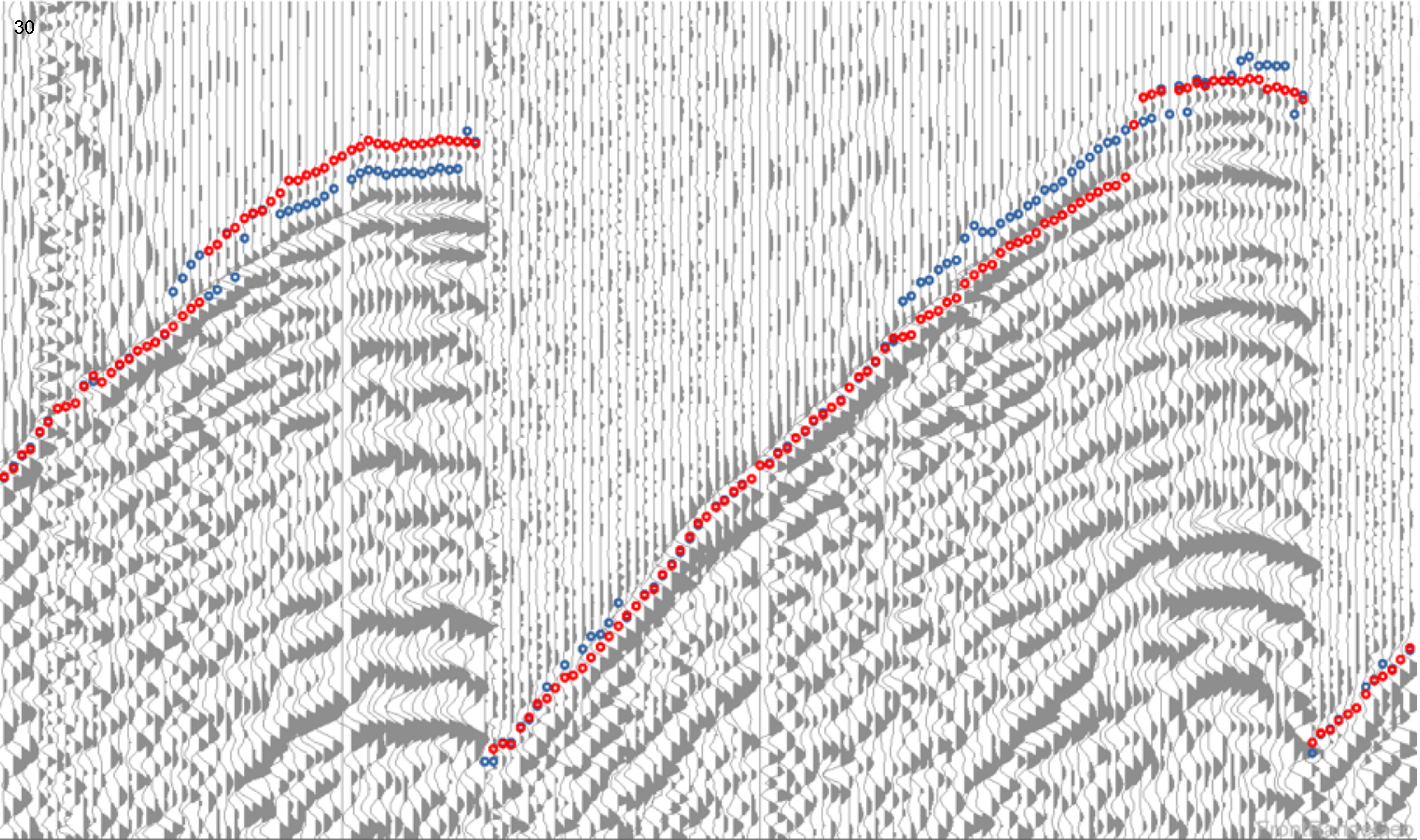
Noisy Seismic Data

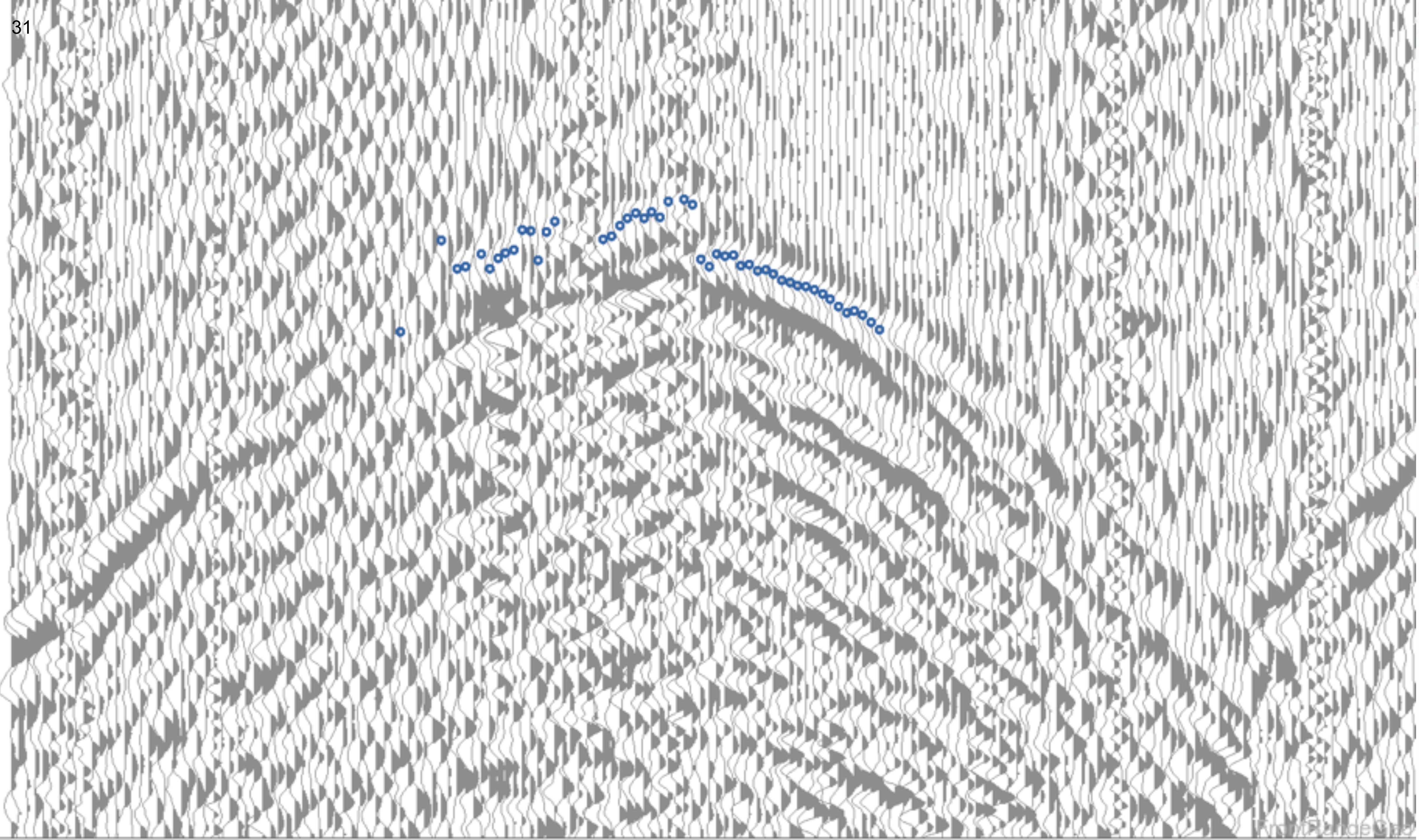
- Human
- DeepTrace

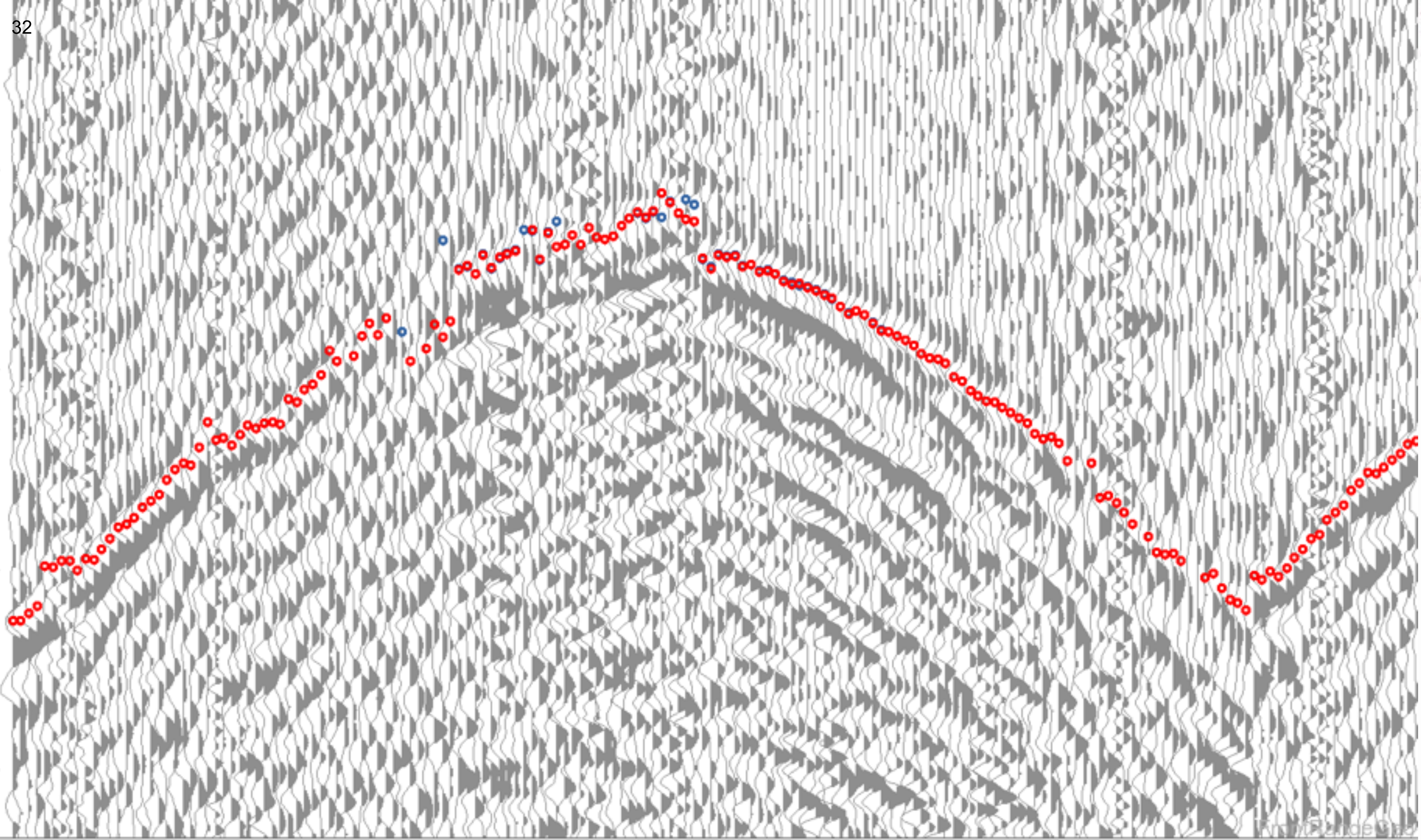


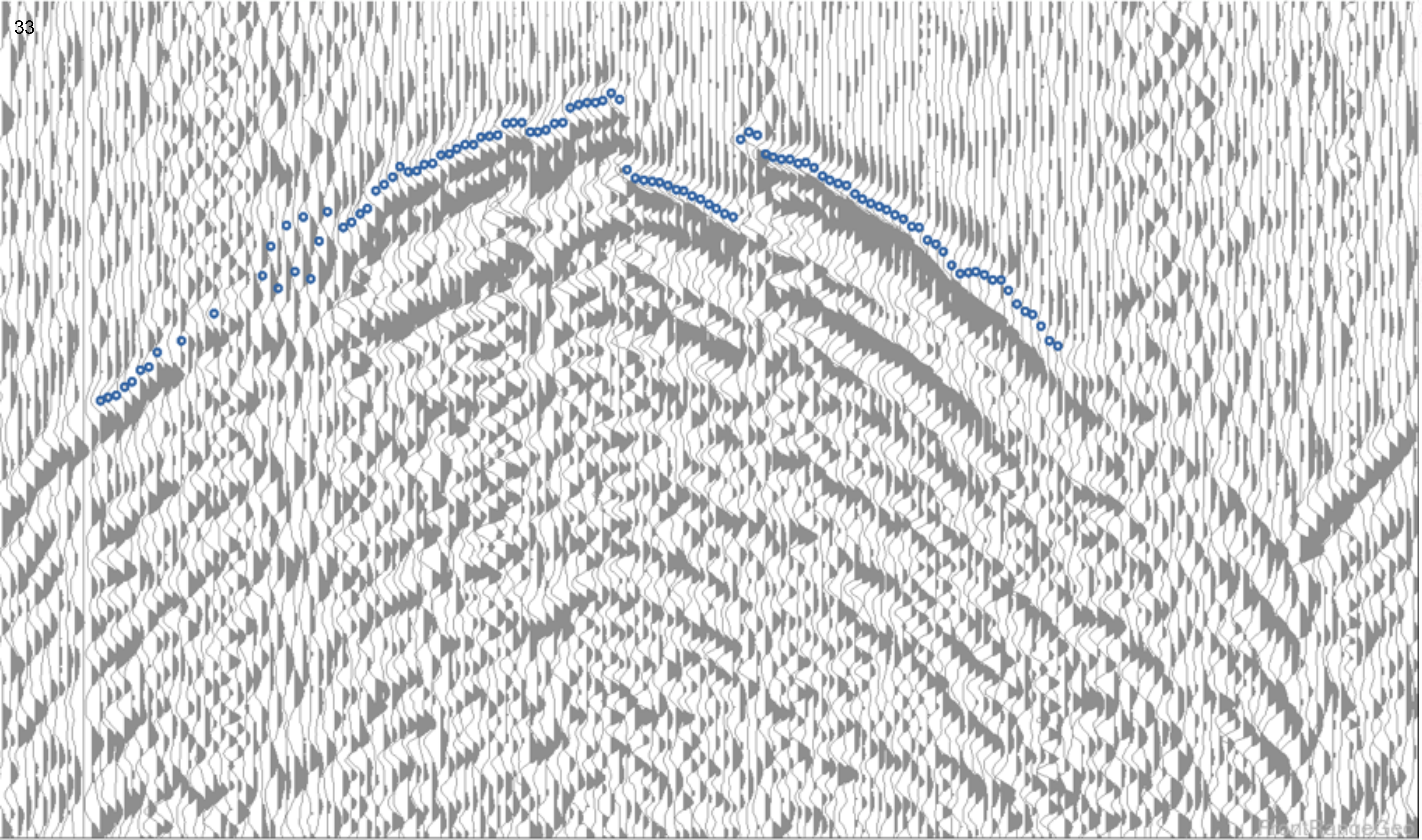
Data from the Permian Basin

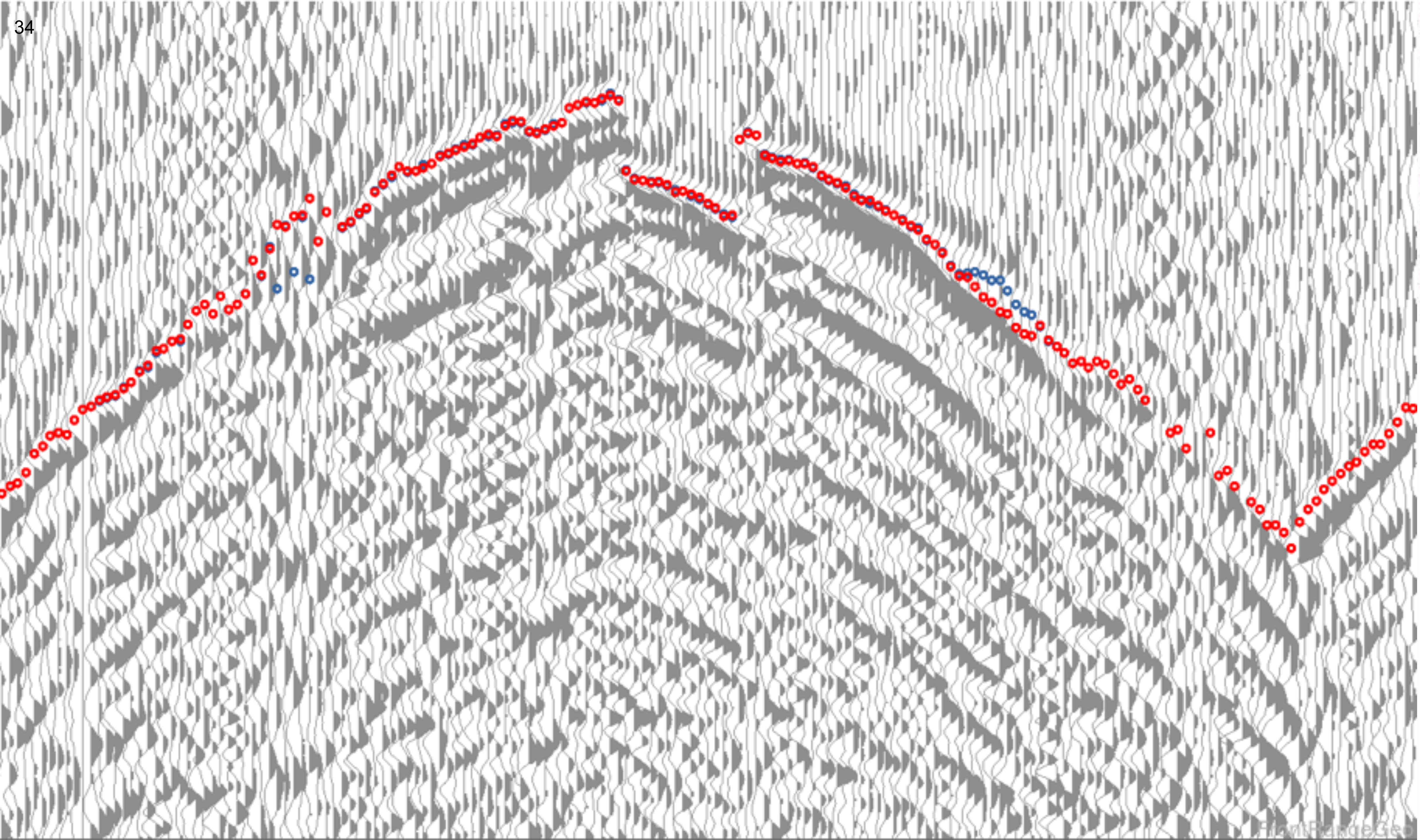


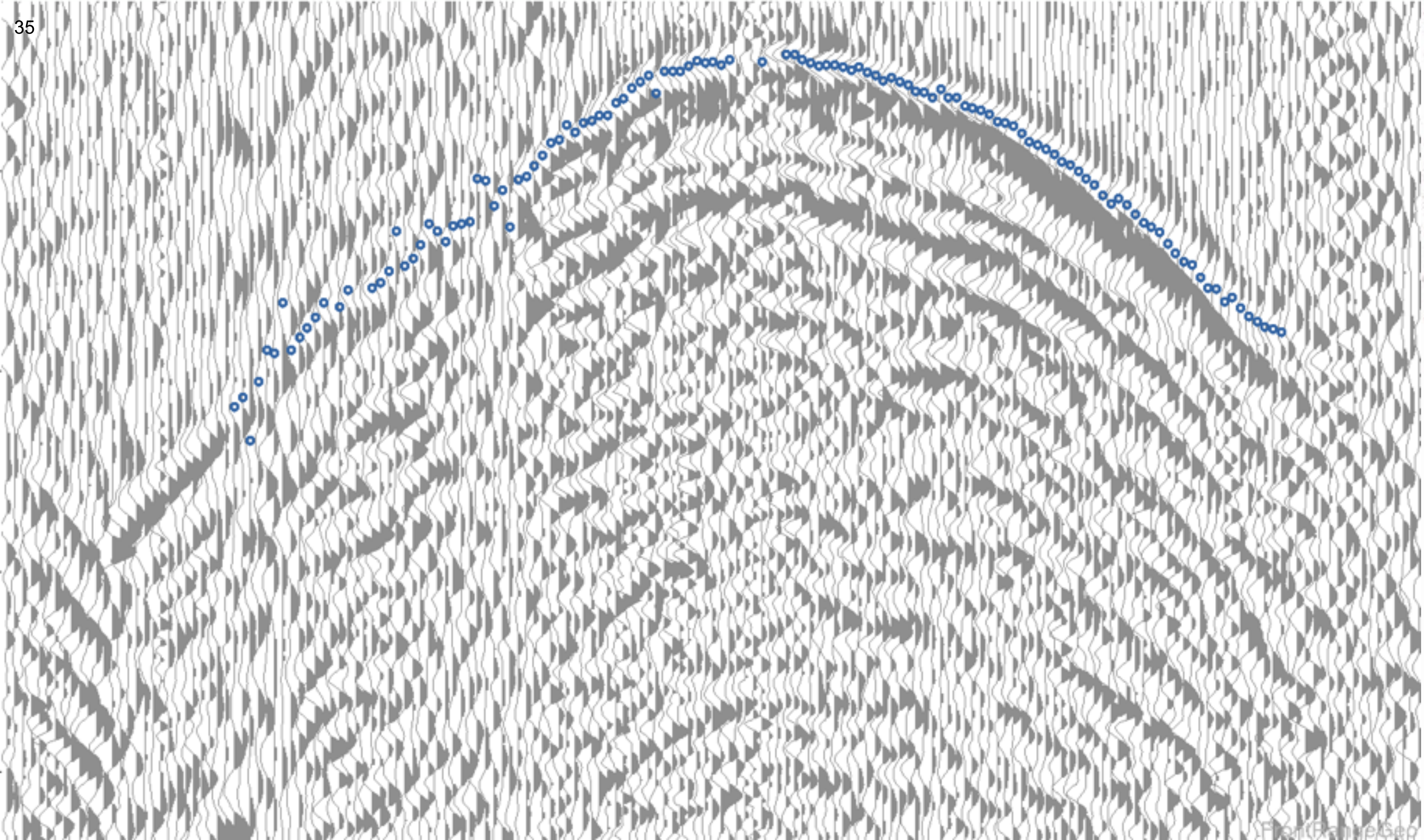


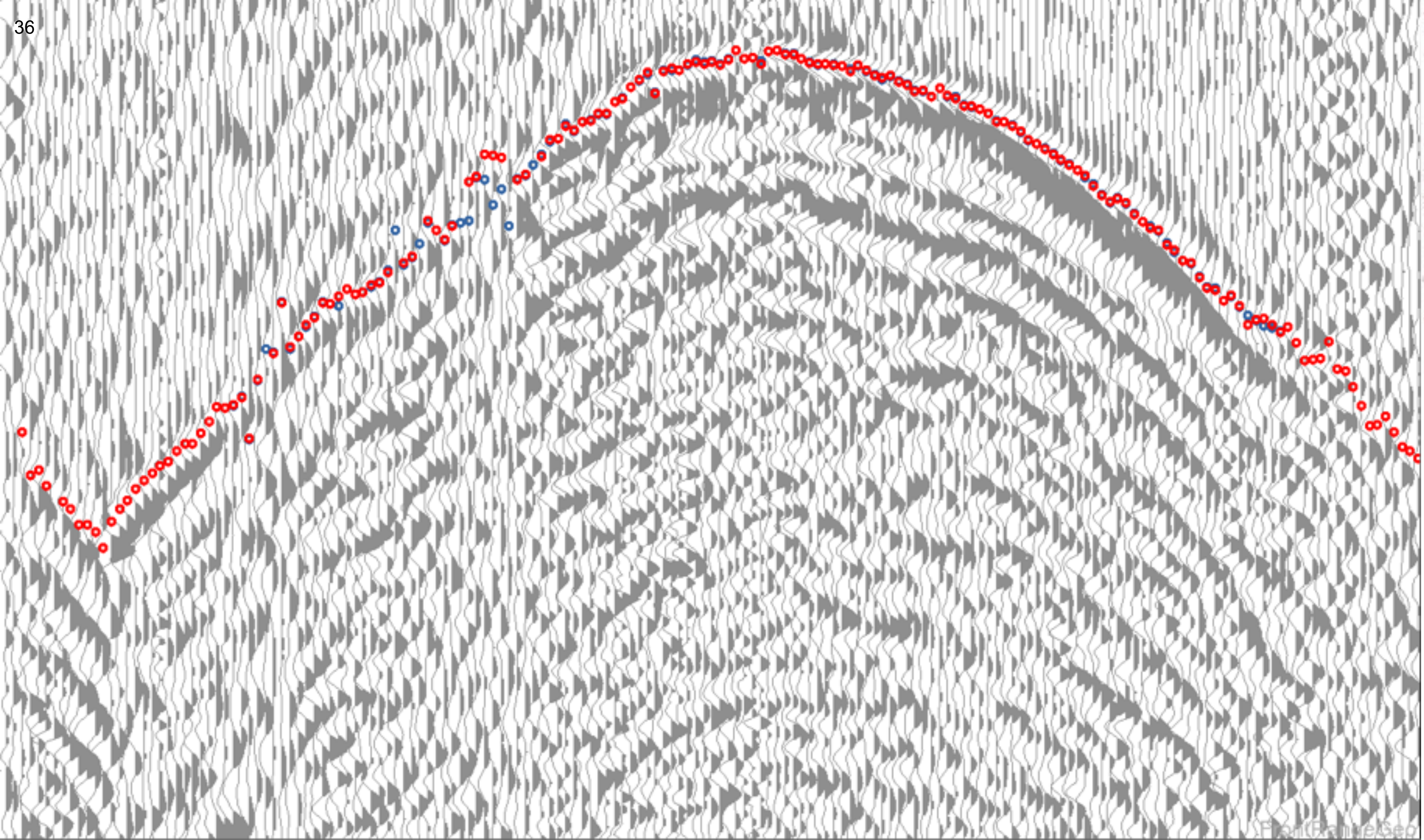


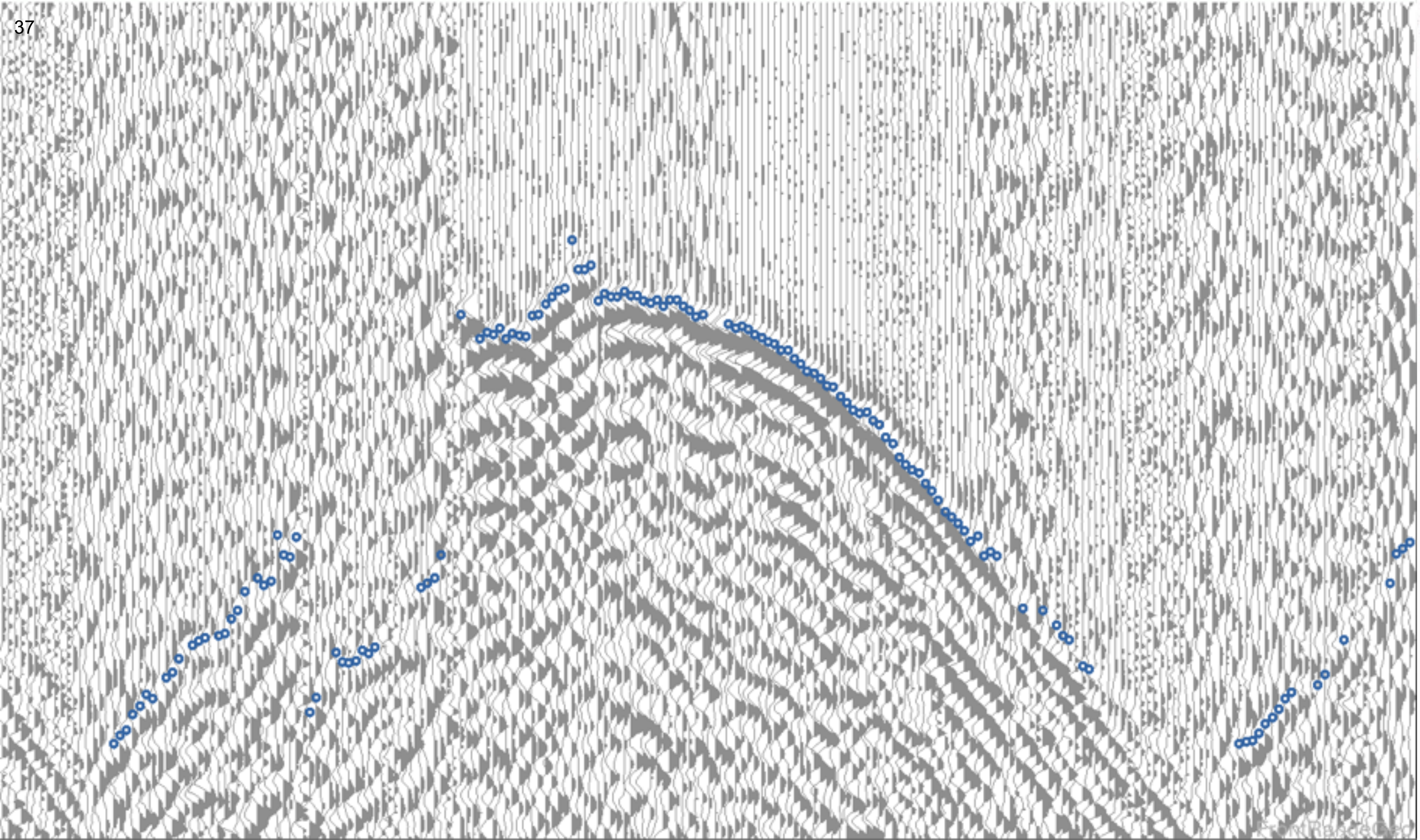


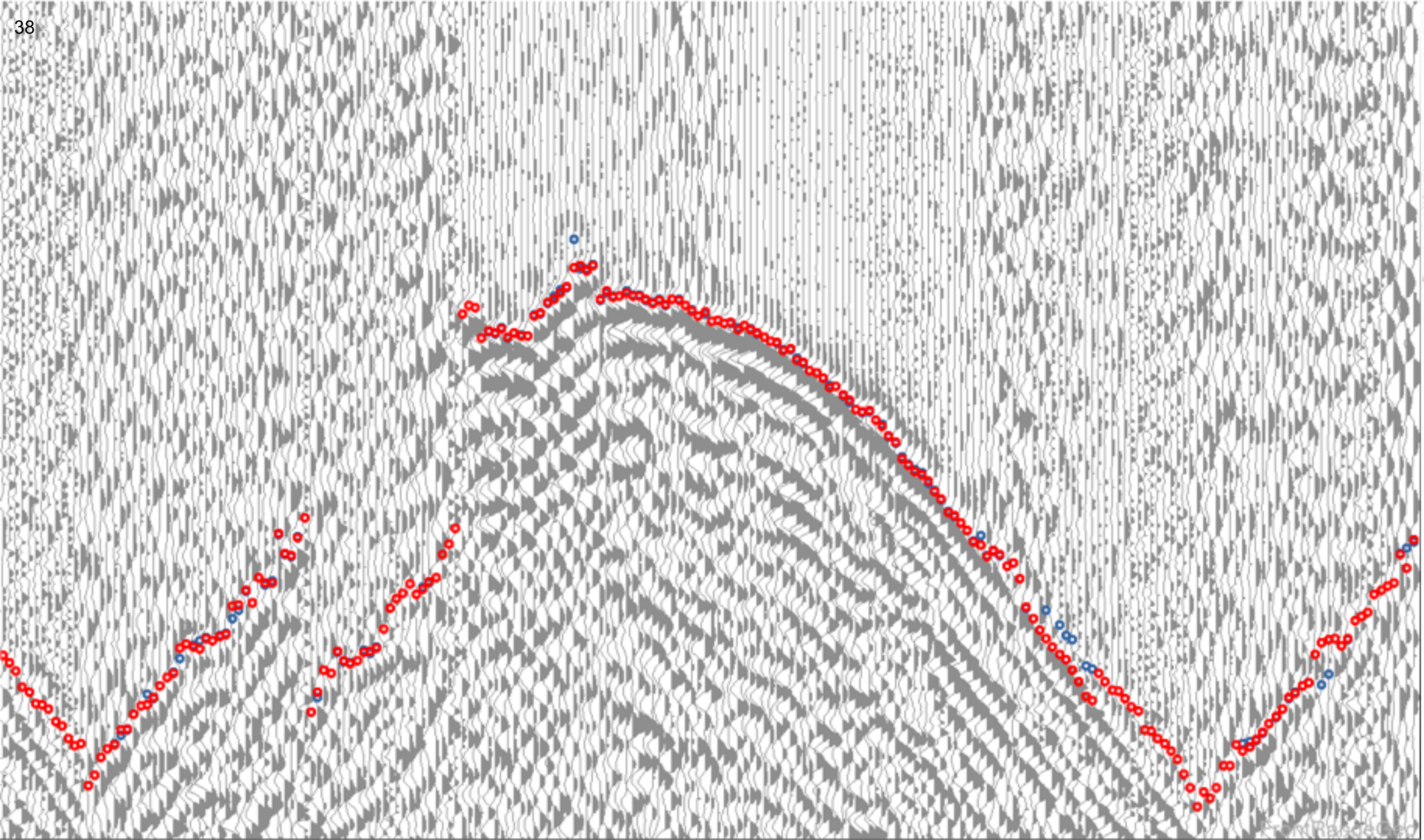


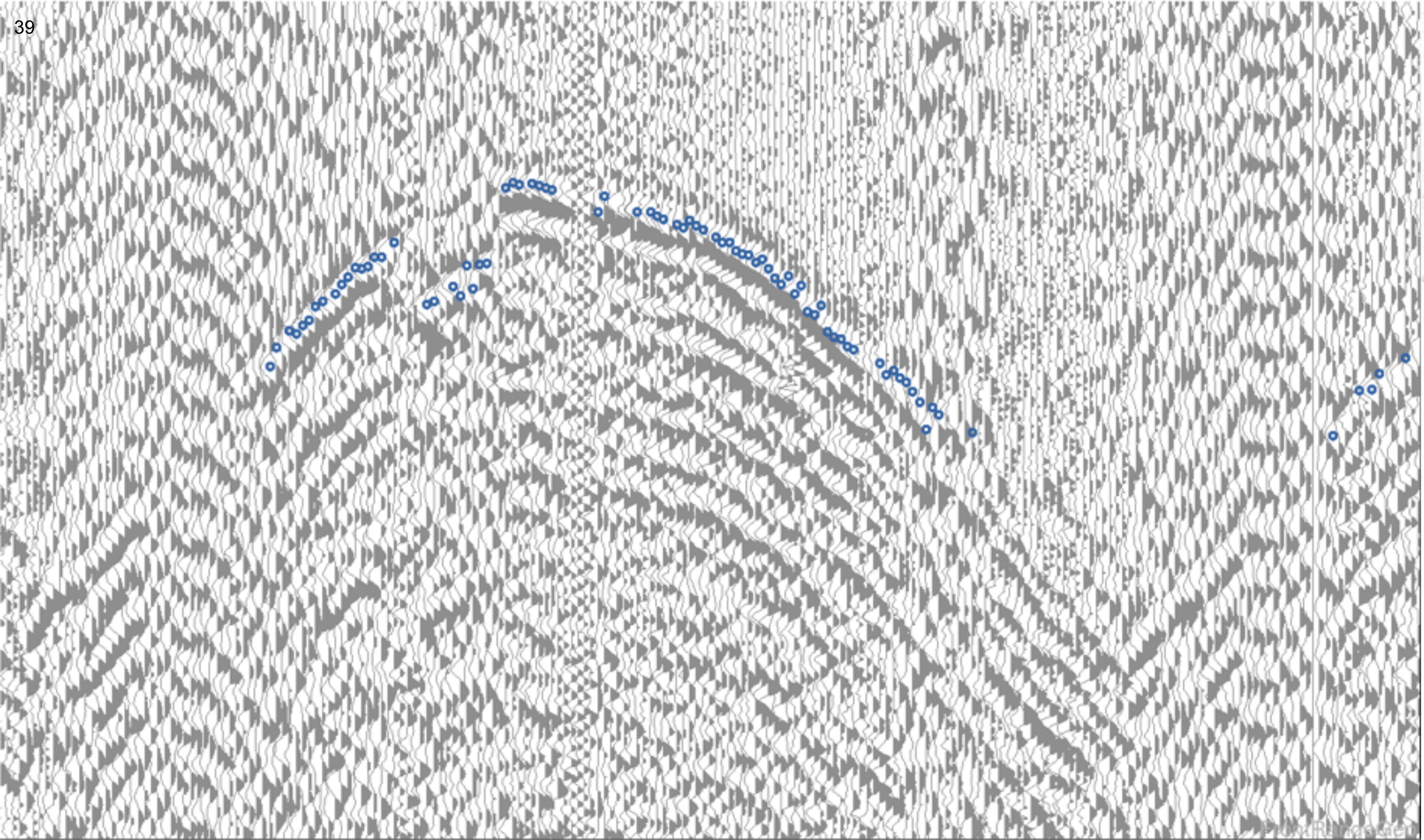

FrontRangeGeo

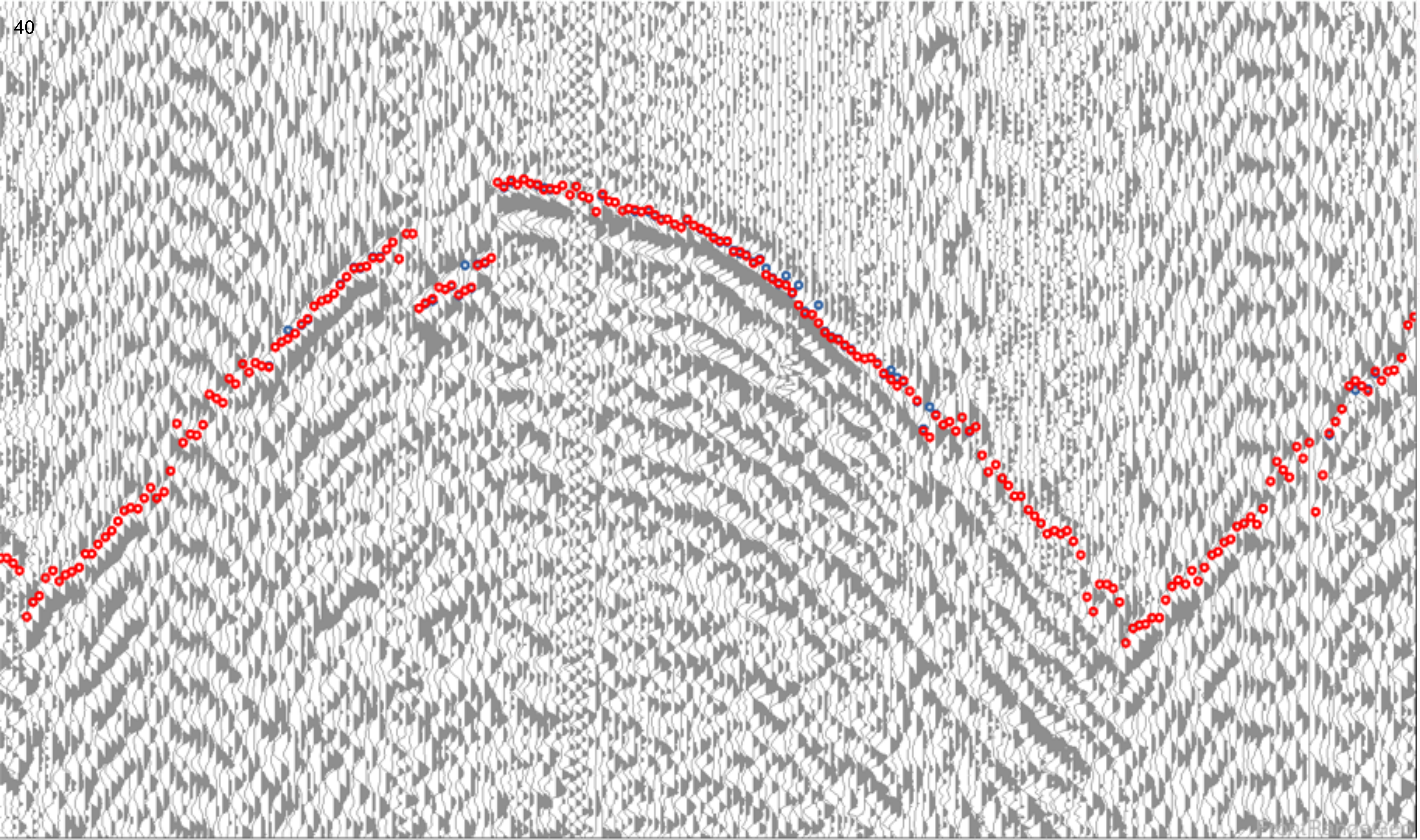


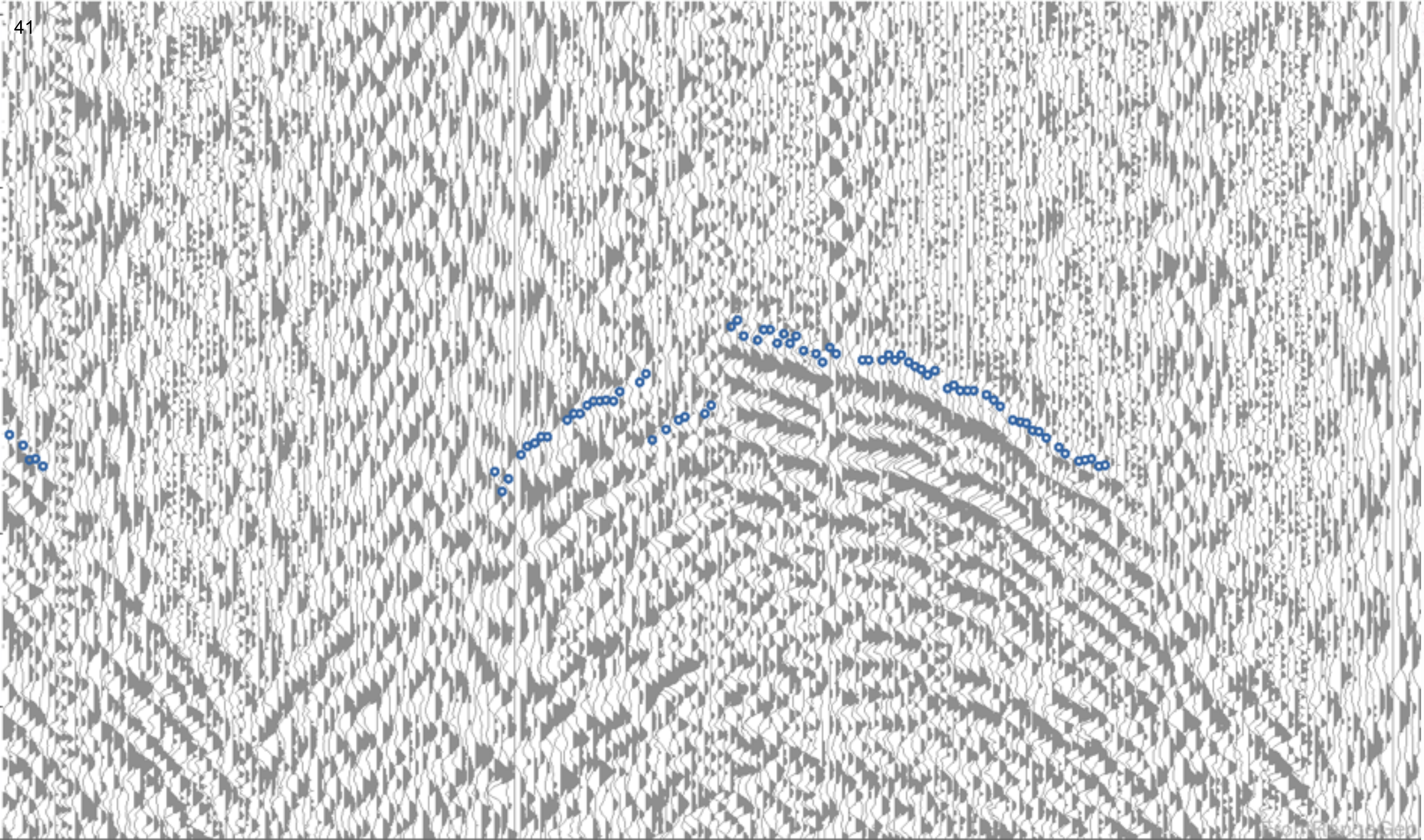


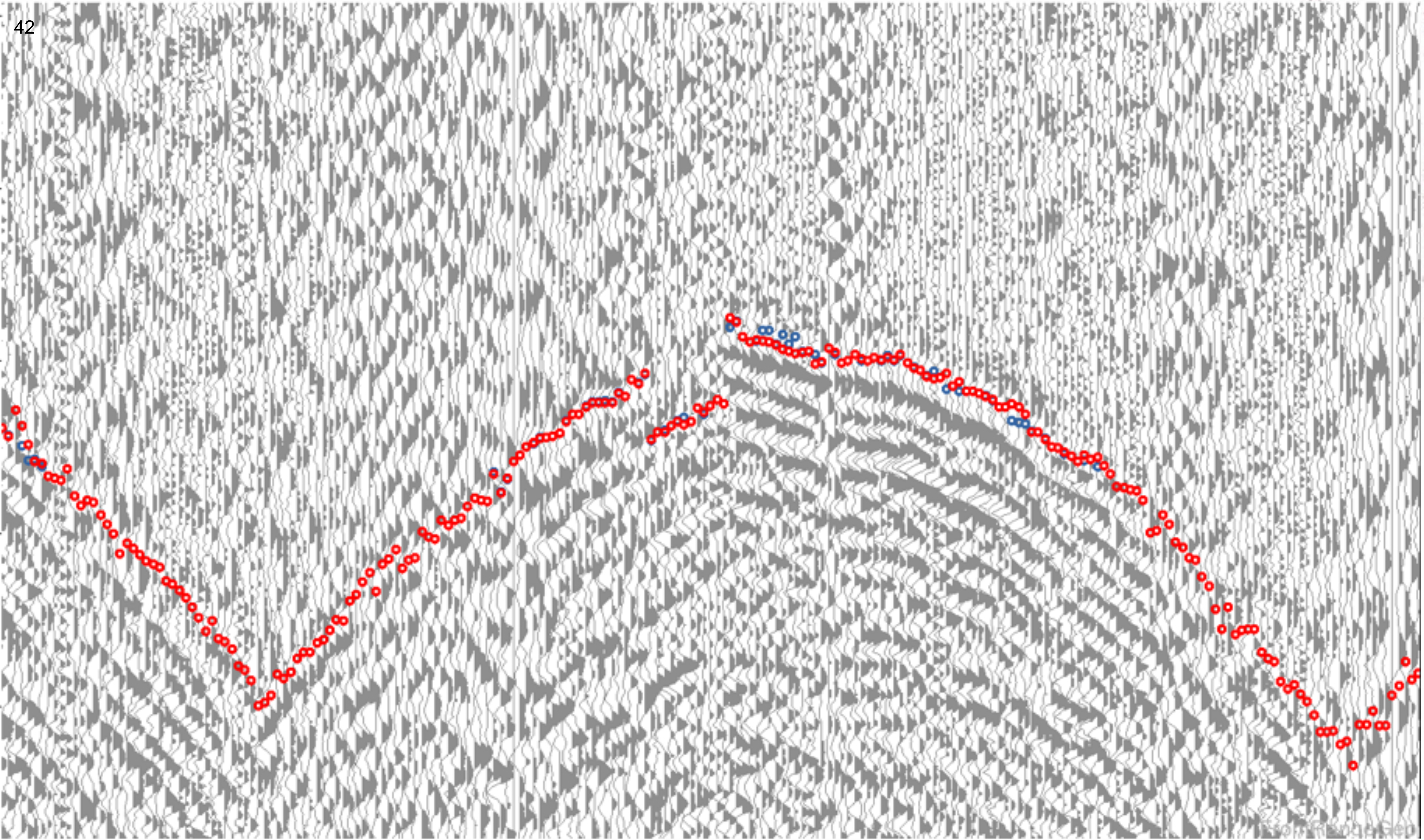


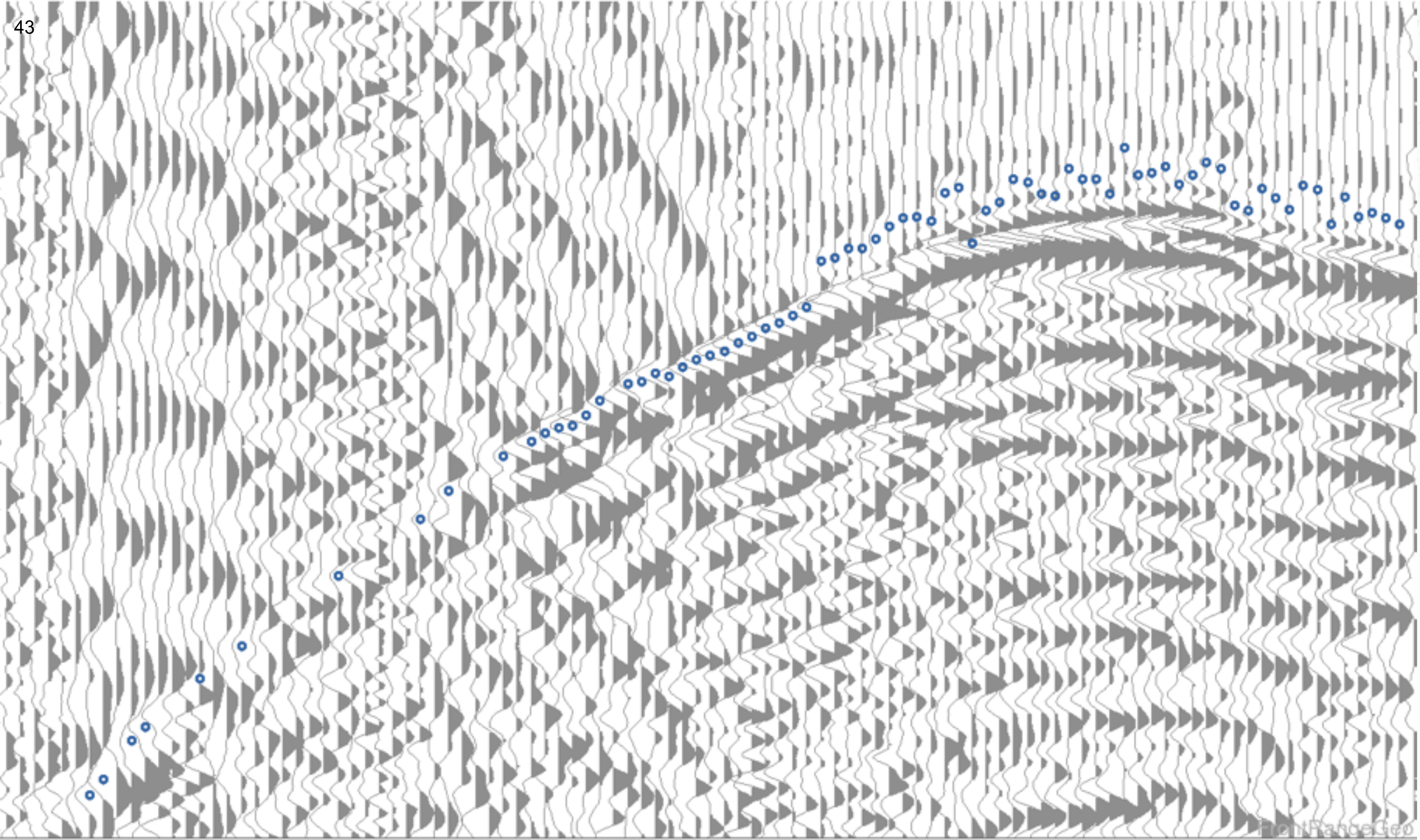


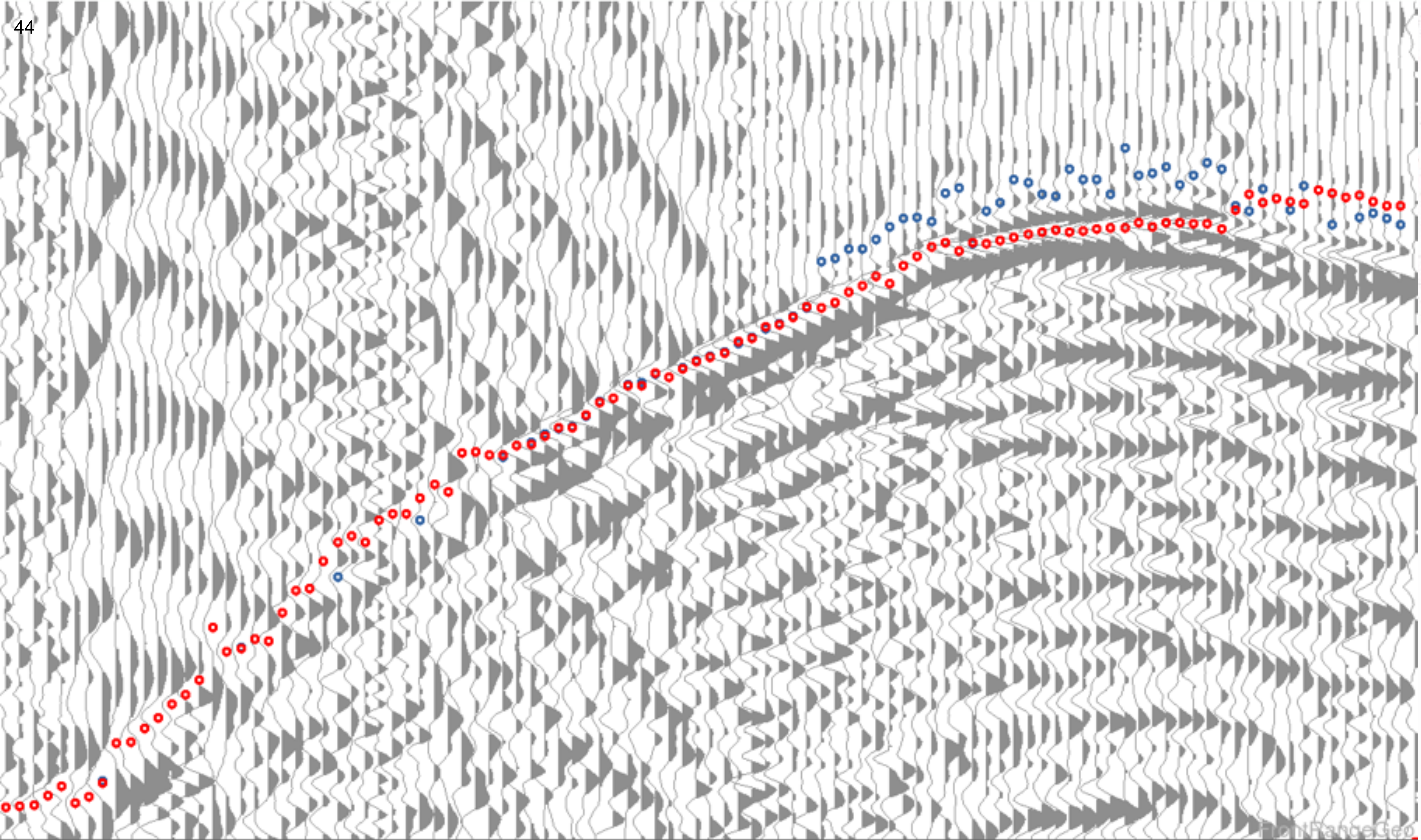


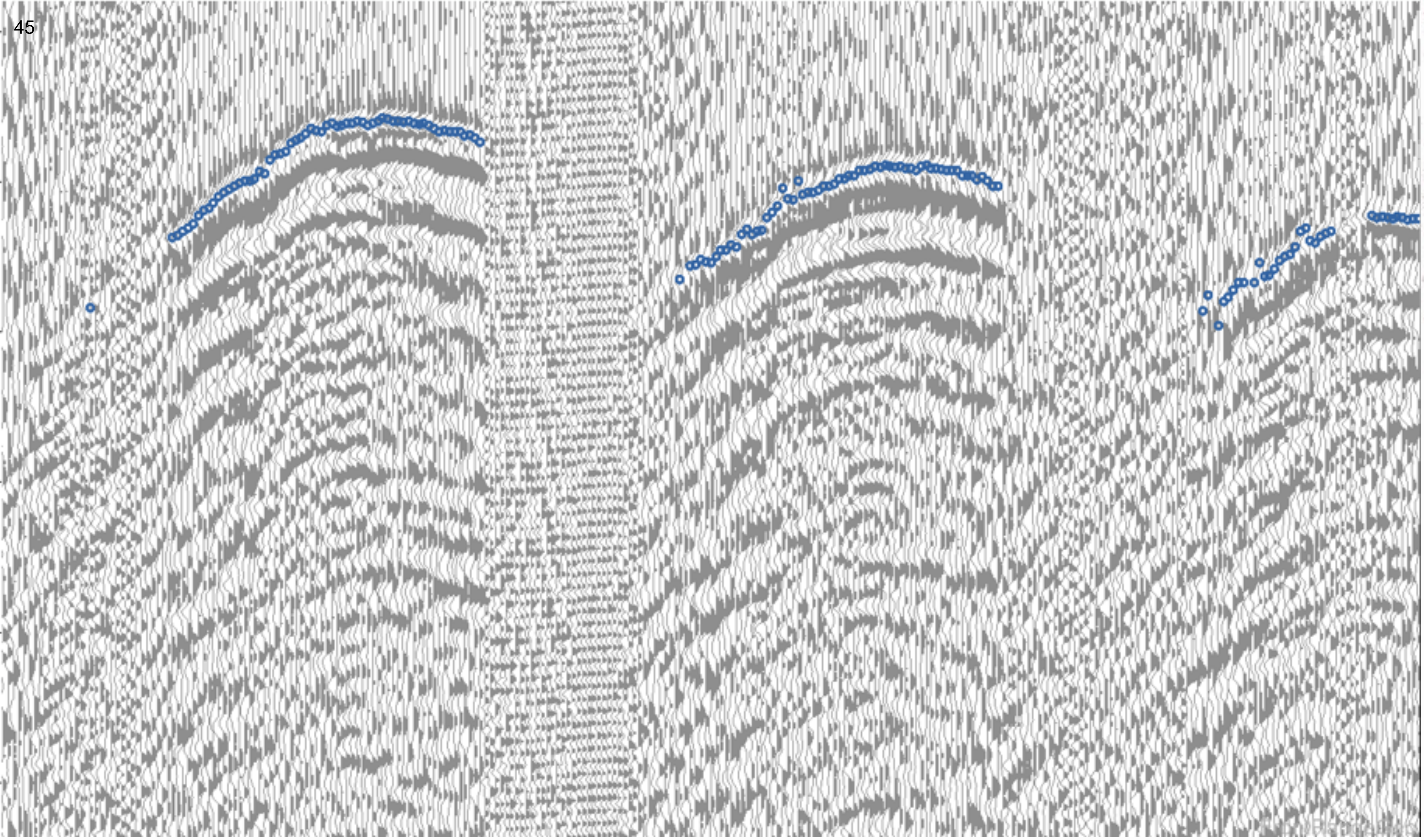


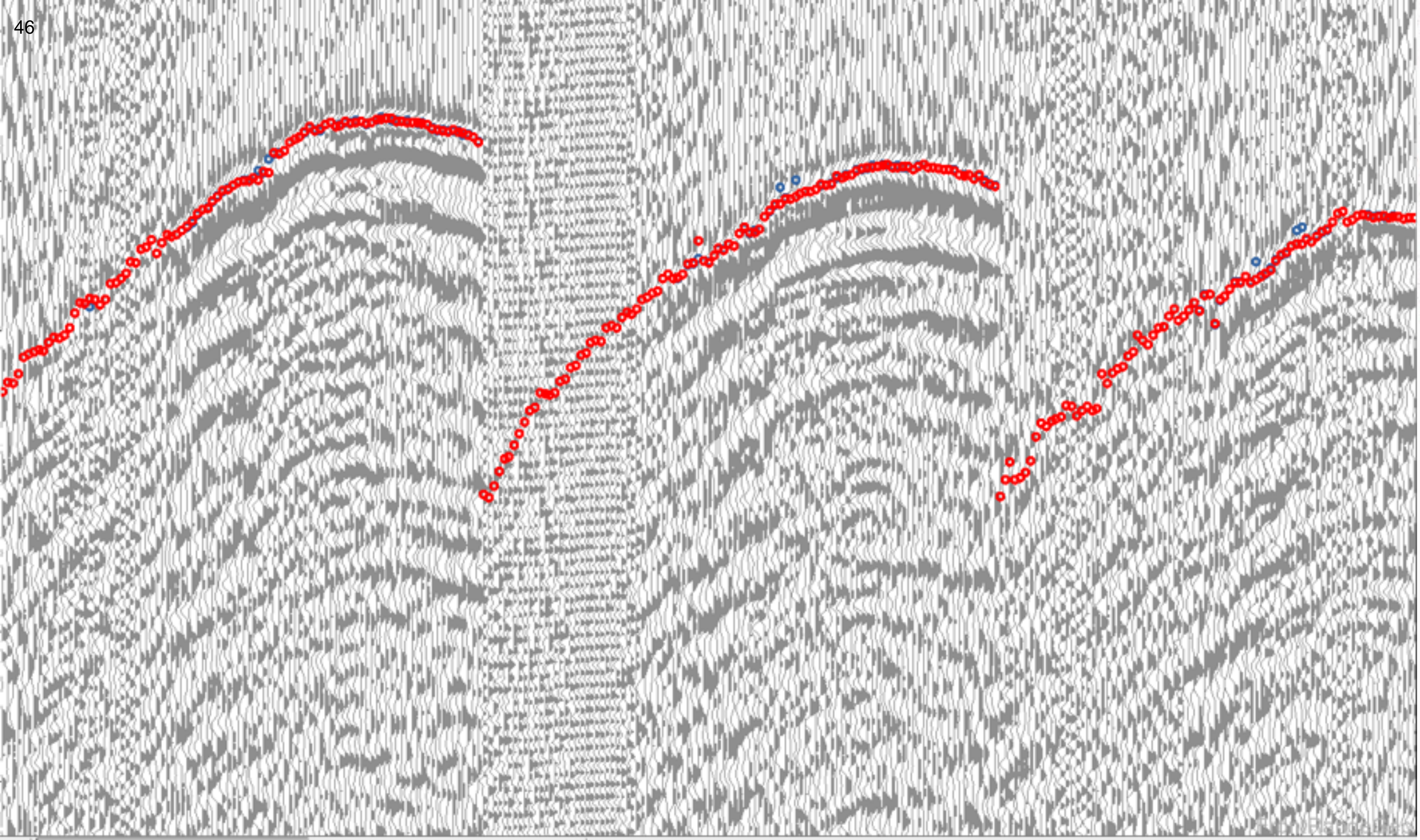


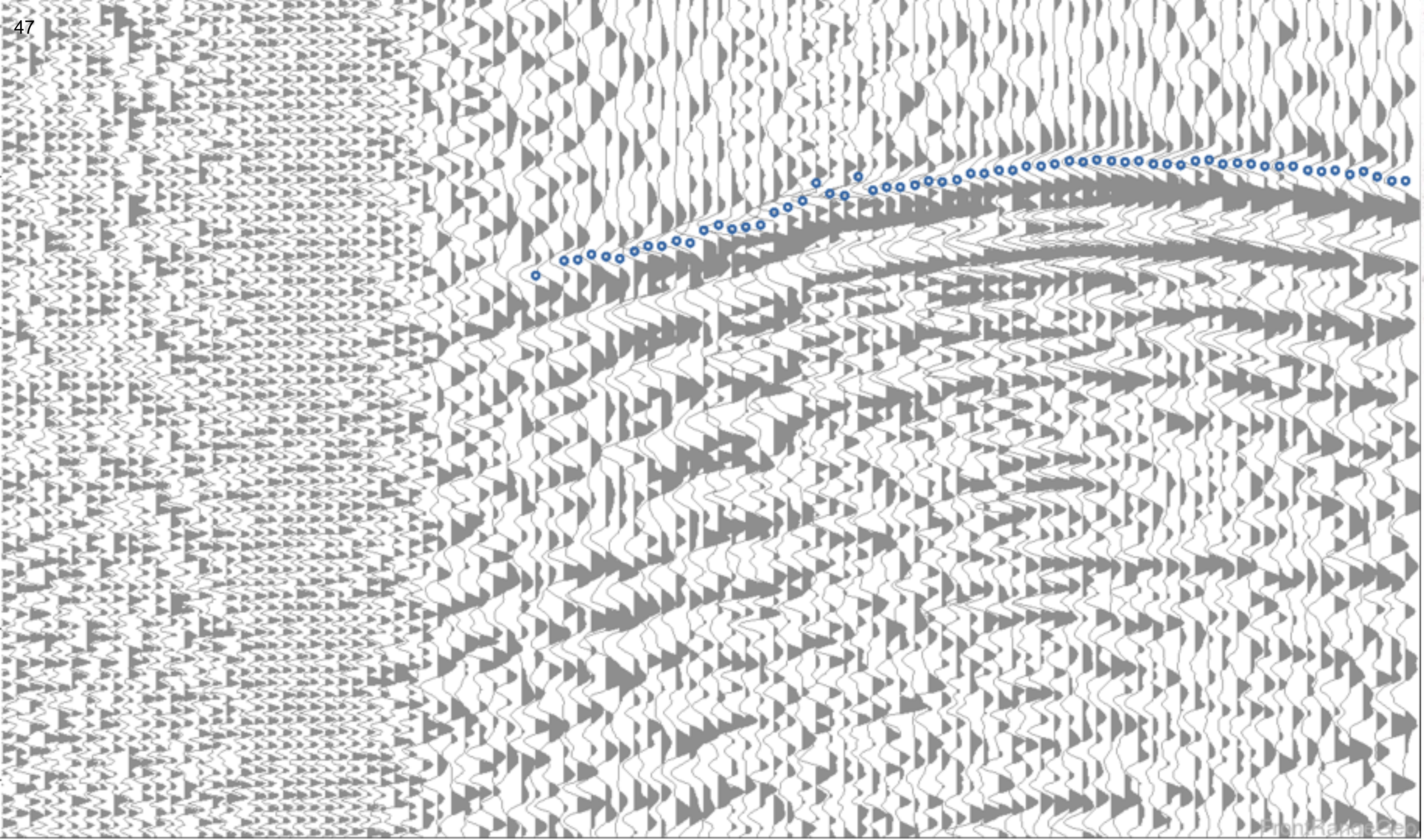


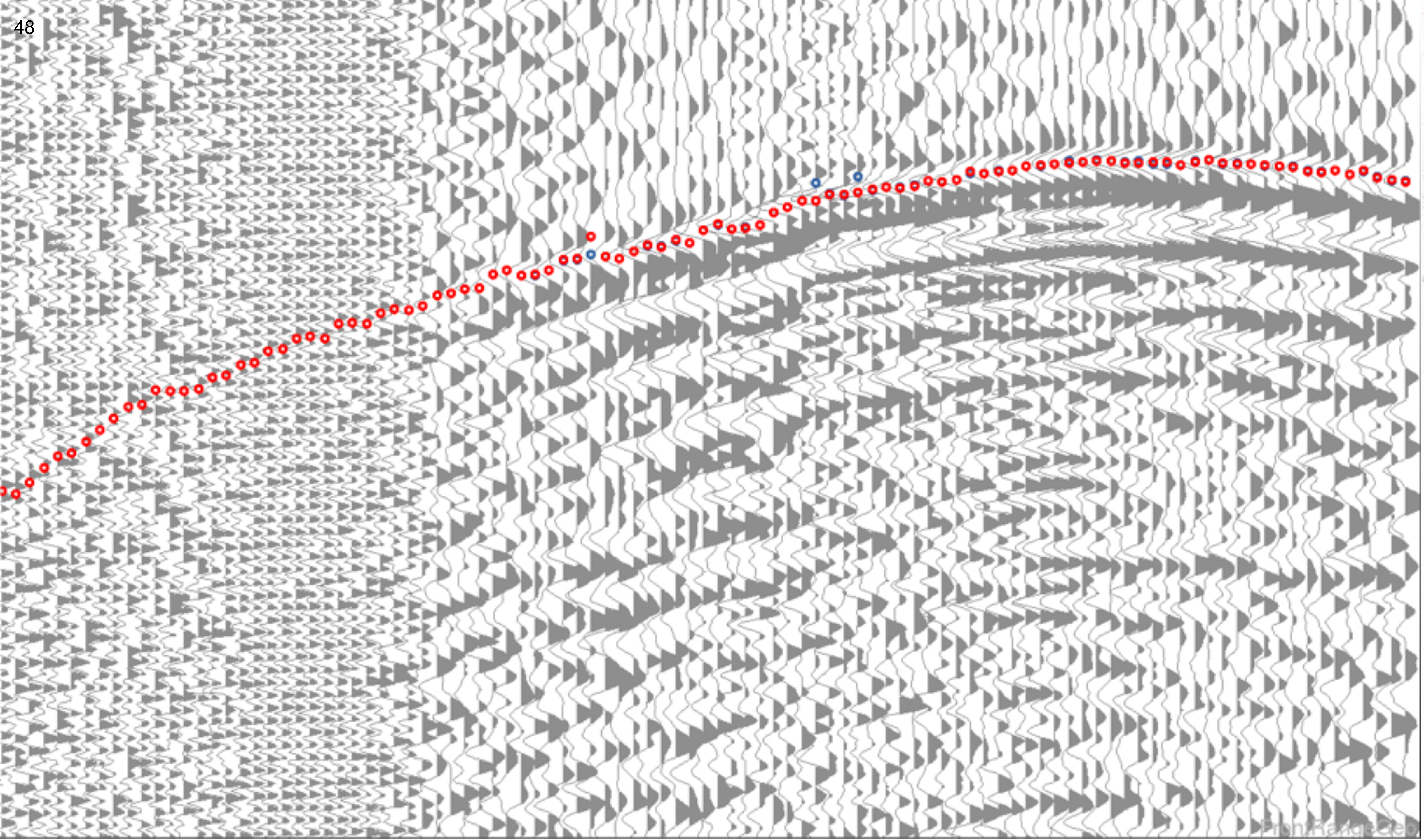

FrontRangeGeo

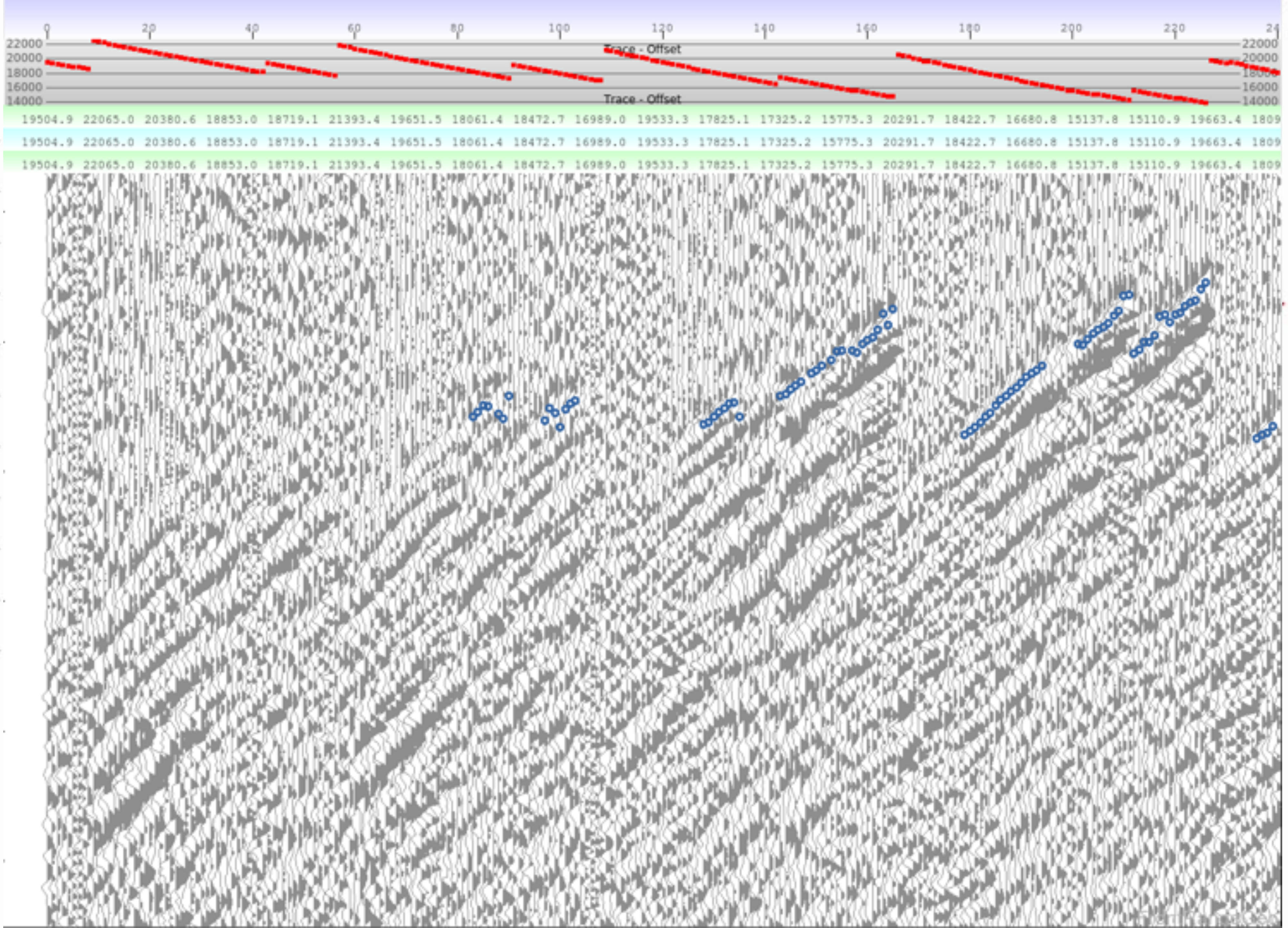


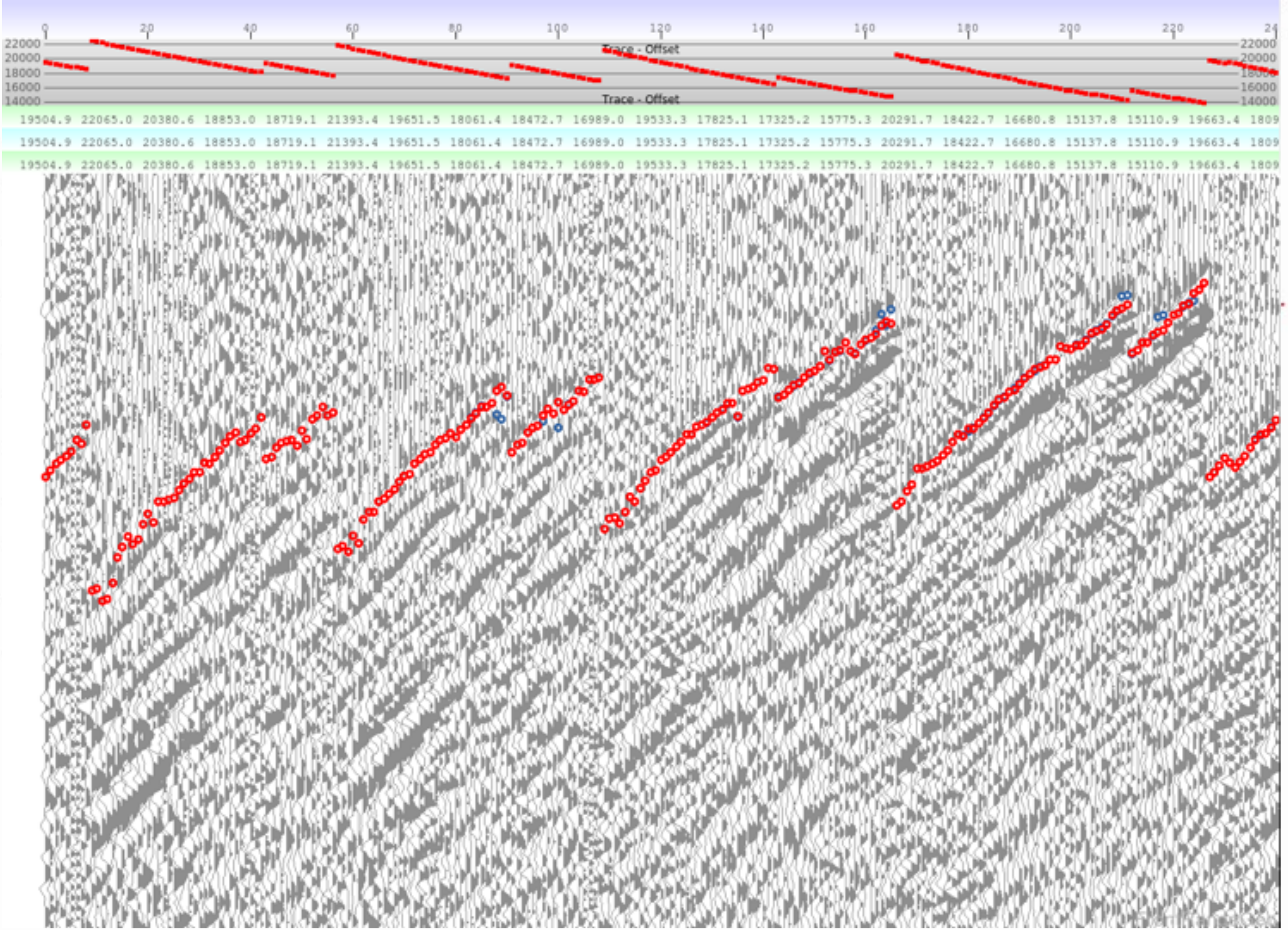


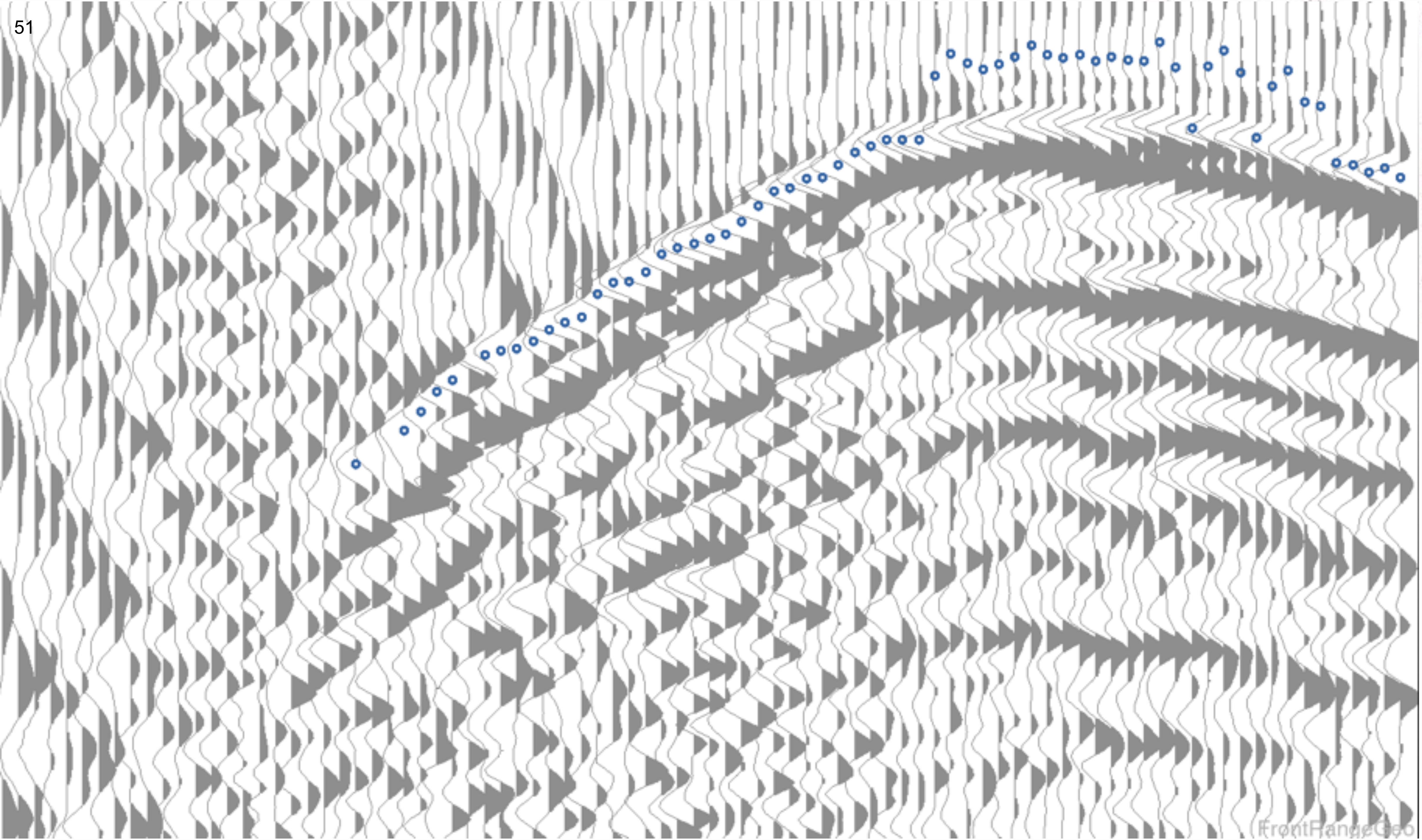


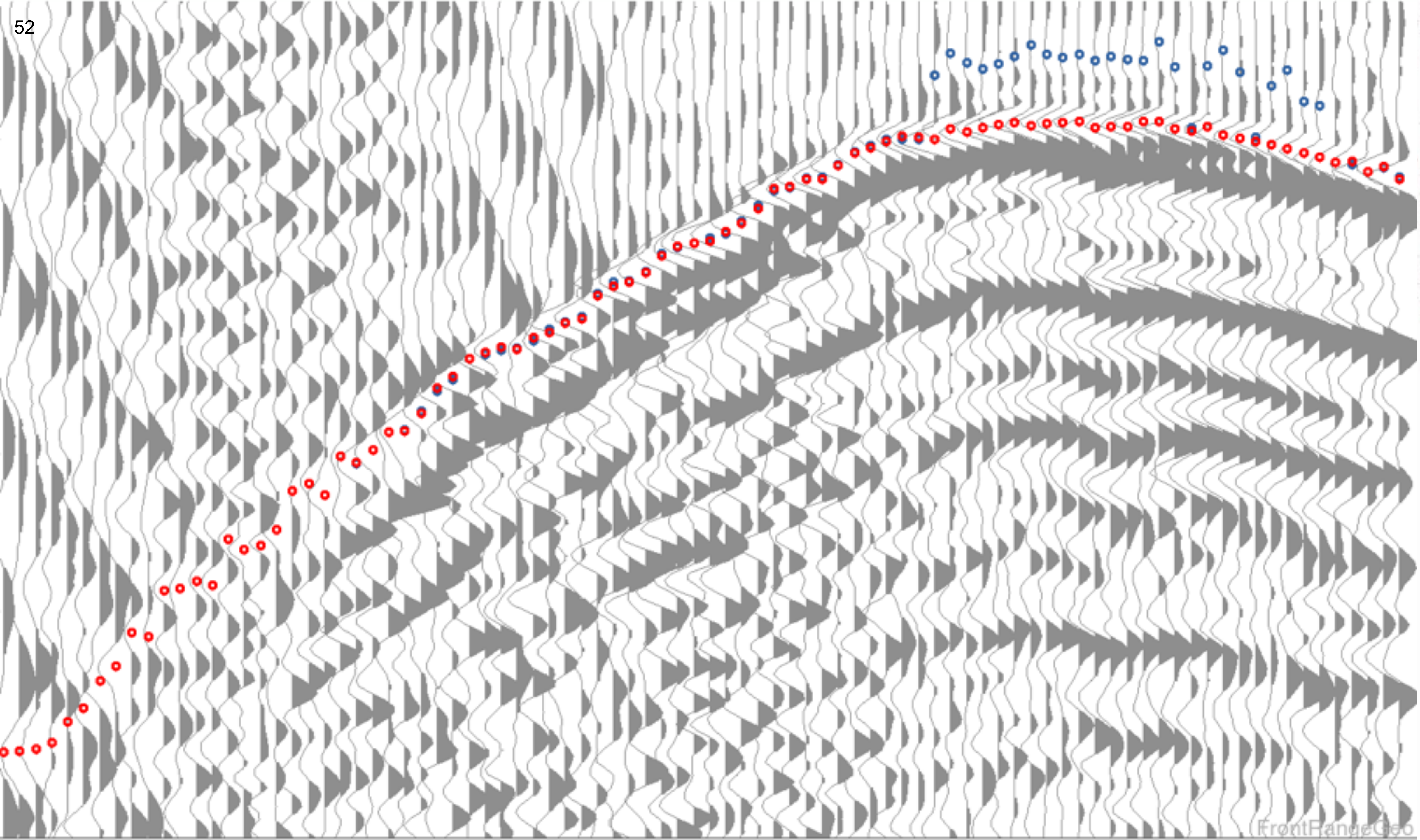


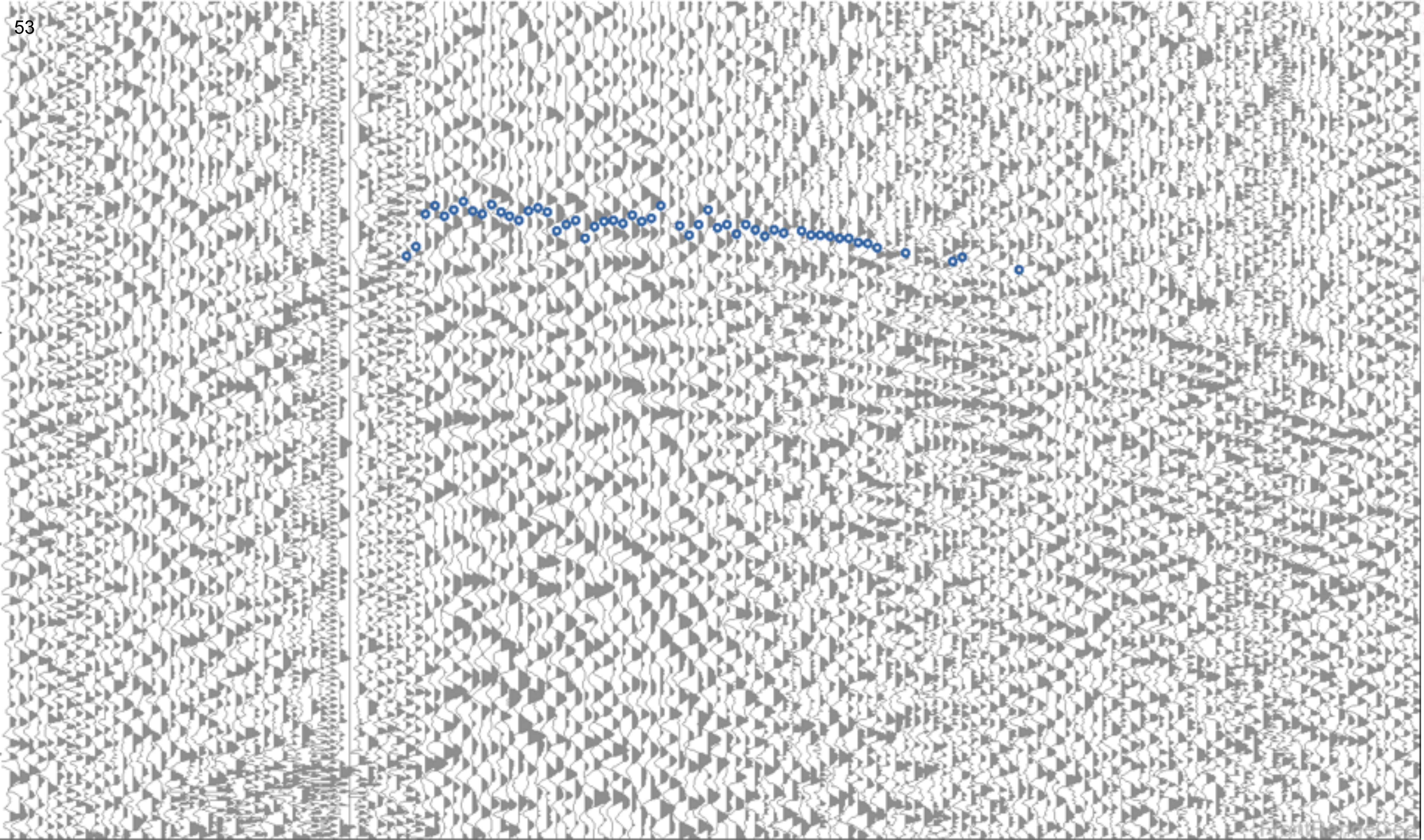


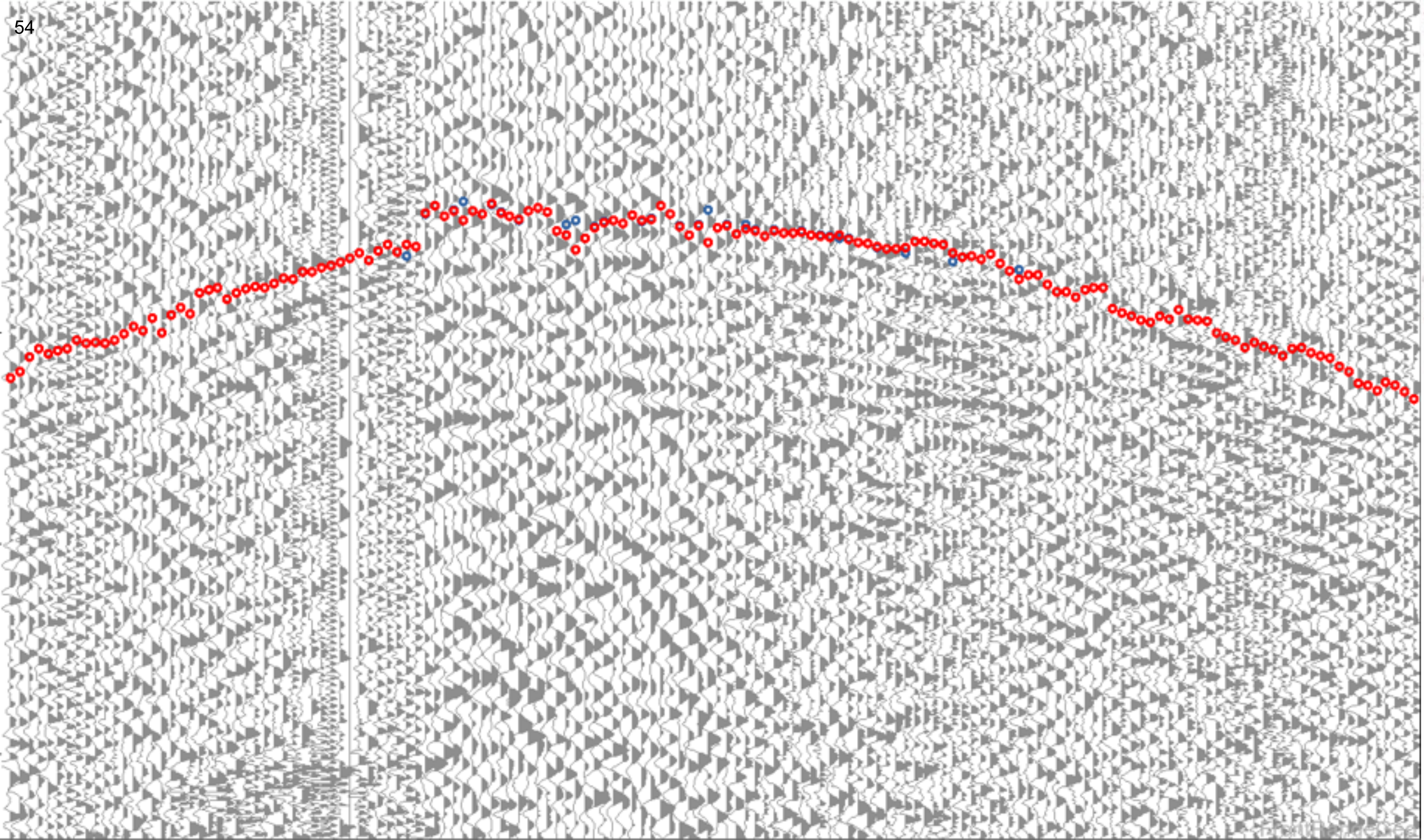

0 20 40 60 80 100 120 140 160 180 200 220 24
22000
20000
18000
16000
14000
Trace - Offset
Trace - Offset
19504.9 22065.0 20380.6 18853.0 18719.1 21393.4 19651.5 18061.4 18472.7 16989.0 19533.3 17825.1 17325.2 15775.3 20291.7 18422.7 16680.8 15137.8 15110.9 19663.4 1809
19504.9 22065.0 20380.6 18853.0 18719.1 21393.4 19651.5 18061.4 18472.7 16989.0 19533.3 17825.1 17325.2 15775.3 20291.7 18422.7 16680.8 15137.8 15110.9 19663.4 1809
19504.9 22065.0 20380.6 18853.0 18719.1 21393.4 19651.5 18061.4 18472.7 16989.0 19533.3 17825.1 17325.2 15775.3 20291.7 18422.7 16680.8 15137.8 15110.9 19663.4 1809


0 20 40 60 80 100 120 140 160 180 200 220 24
22000 20000 18000 16000 14000
Trace - Offset
Trace - Offset
19504.9 22065.0 20380.6 18853.0 18719.1 21393.4 19651.5 18061.4 18472.7 16989.0 19533.3 17825.1 17325.2 15775.3 20291.7 18422.7 16680.8 15137.8 15110.9 19663.4 1809


FrontRangeGeo


FrontRangeGeo





Phoenix

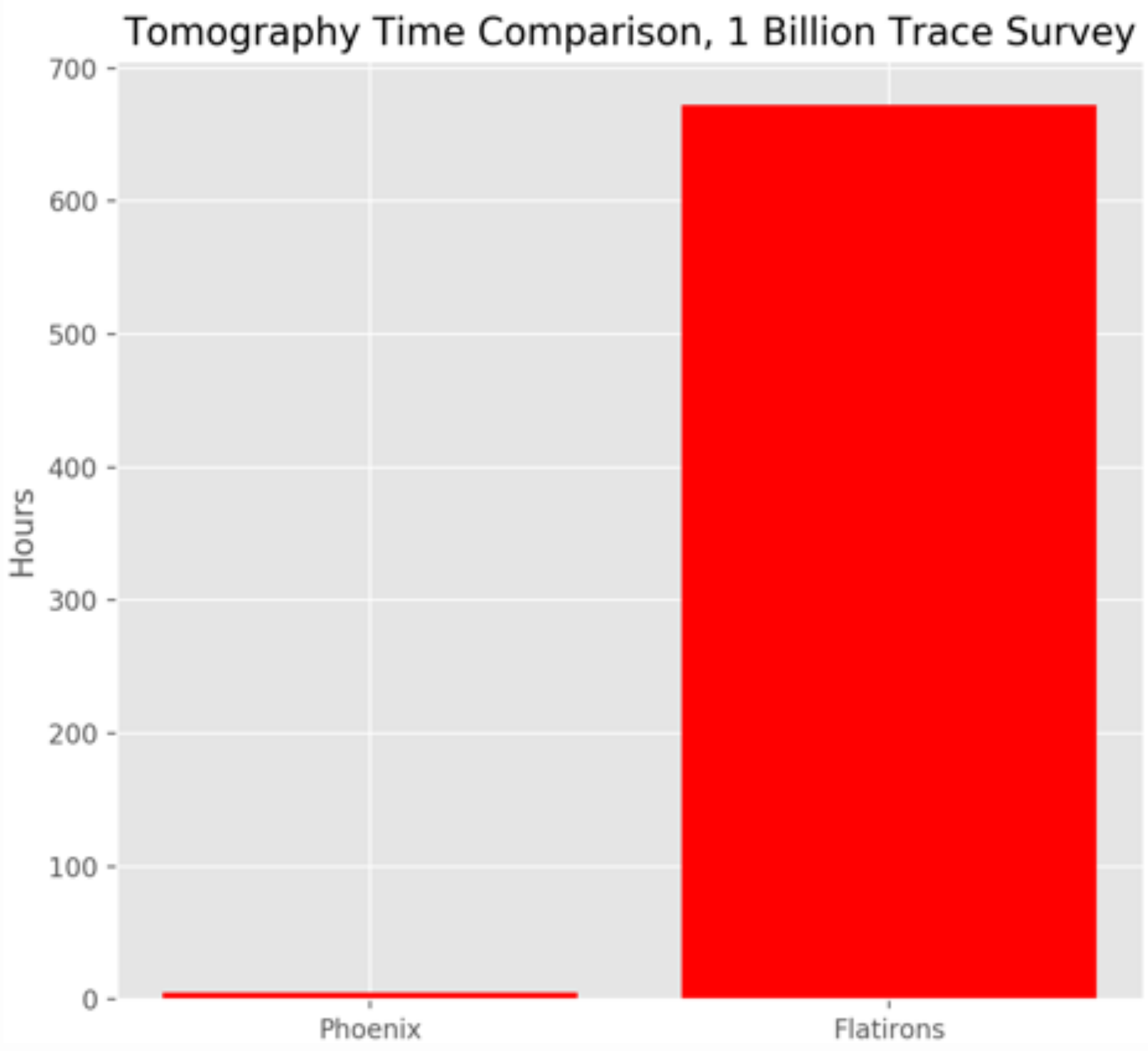
the next generation refraction statics solution

Phoenix

Phoenix is the next generation Refractions Statics solution, designed for the largest and most difficult surveys.

Phoenix achieves unheard-of speeds by massively parallelizing and threading most processes.

New physics and inversion constraint methods increase accuracy and quality of solution.

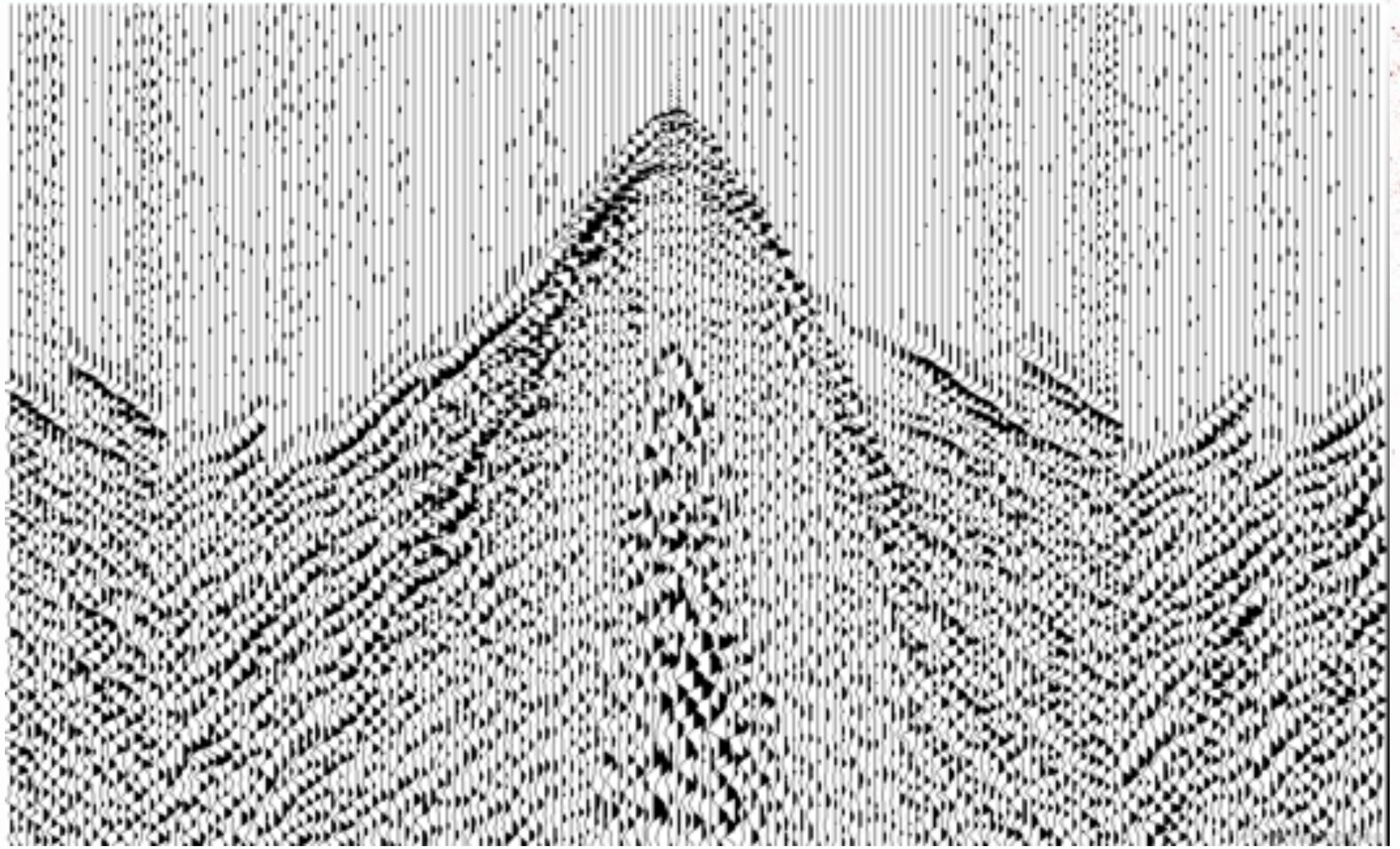


DeepTrace + Phoenix

By iteratively updating the physics model with new picks, and adjusting the network's view of the picks based on the model, we can quickly refine both the picks and model.

Phoenix and DeepTrace models *work together*, while avoiding circularity by taking different approaches (computer vision vs physics).

Phoenix and DeepTrace work together seamlessly to automate the entire near-surface processing workflow, requiring minimal human oversight.

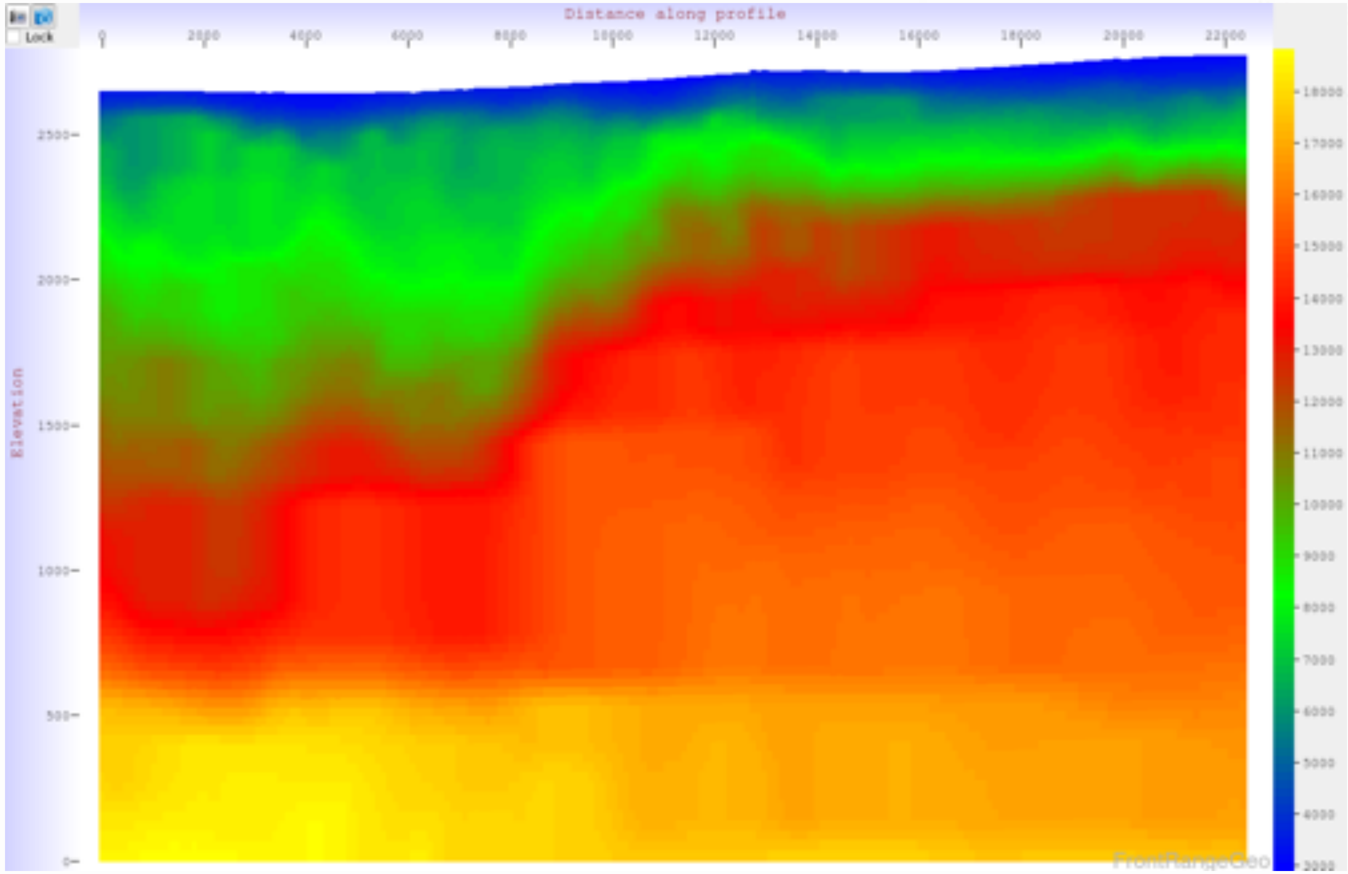


Animation: Shot line gather before and after applying tomographic solution.

Model Comparison

First: Velocity profile from human picks.

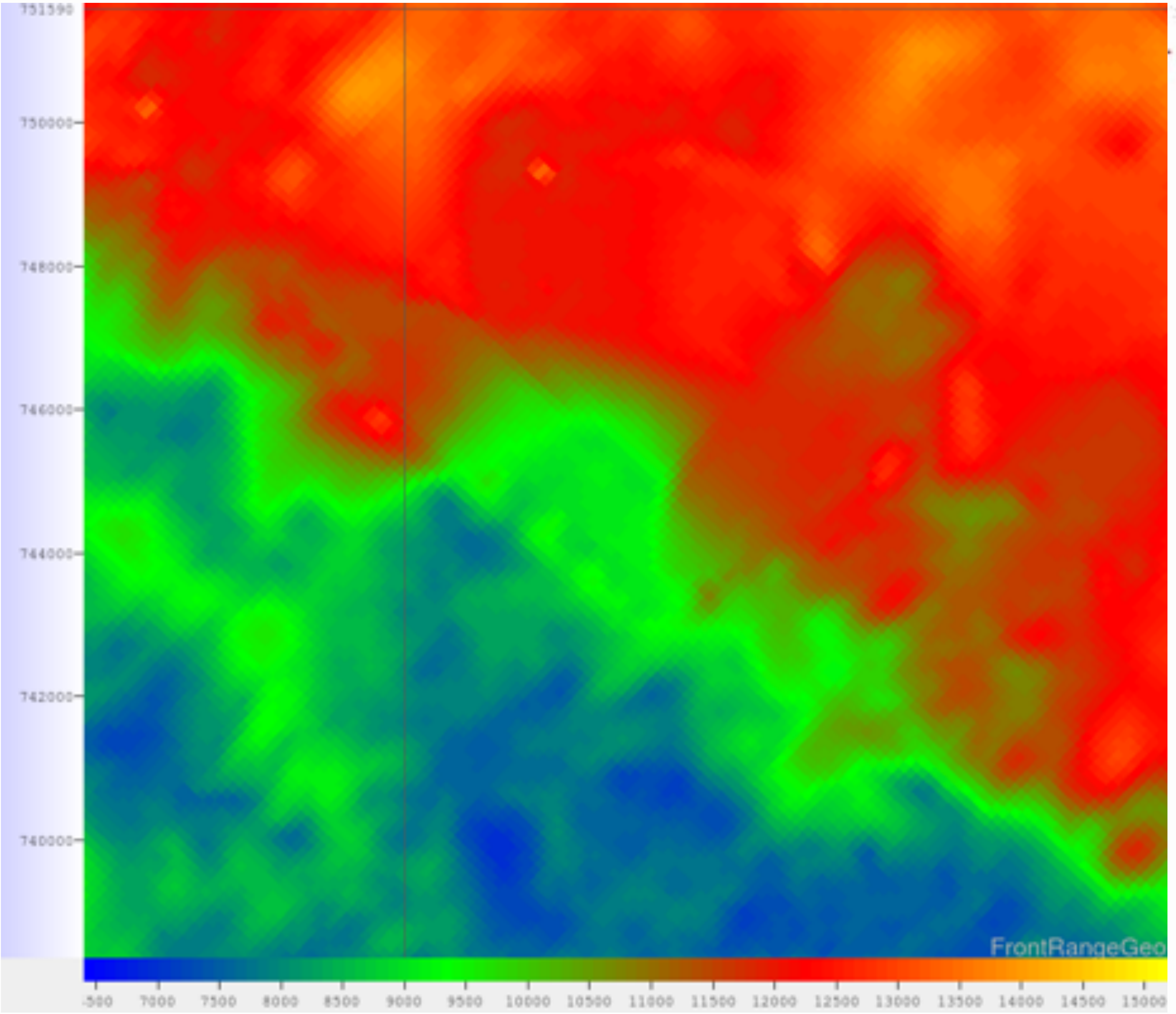
Second: Velocity profile from DeepTrace picks. No post-processing.



Model Comparison

First: Velocity depth slice from human picks.

Second: Velocity depth slice from DeepTrace picks. No post-processing.



Reliability

automating quality control

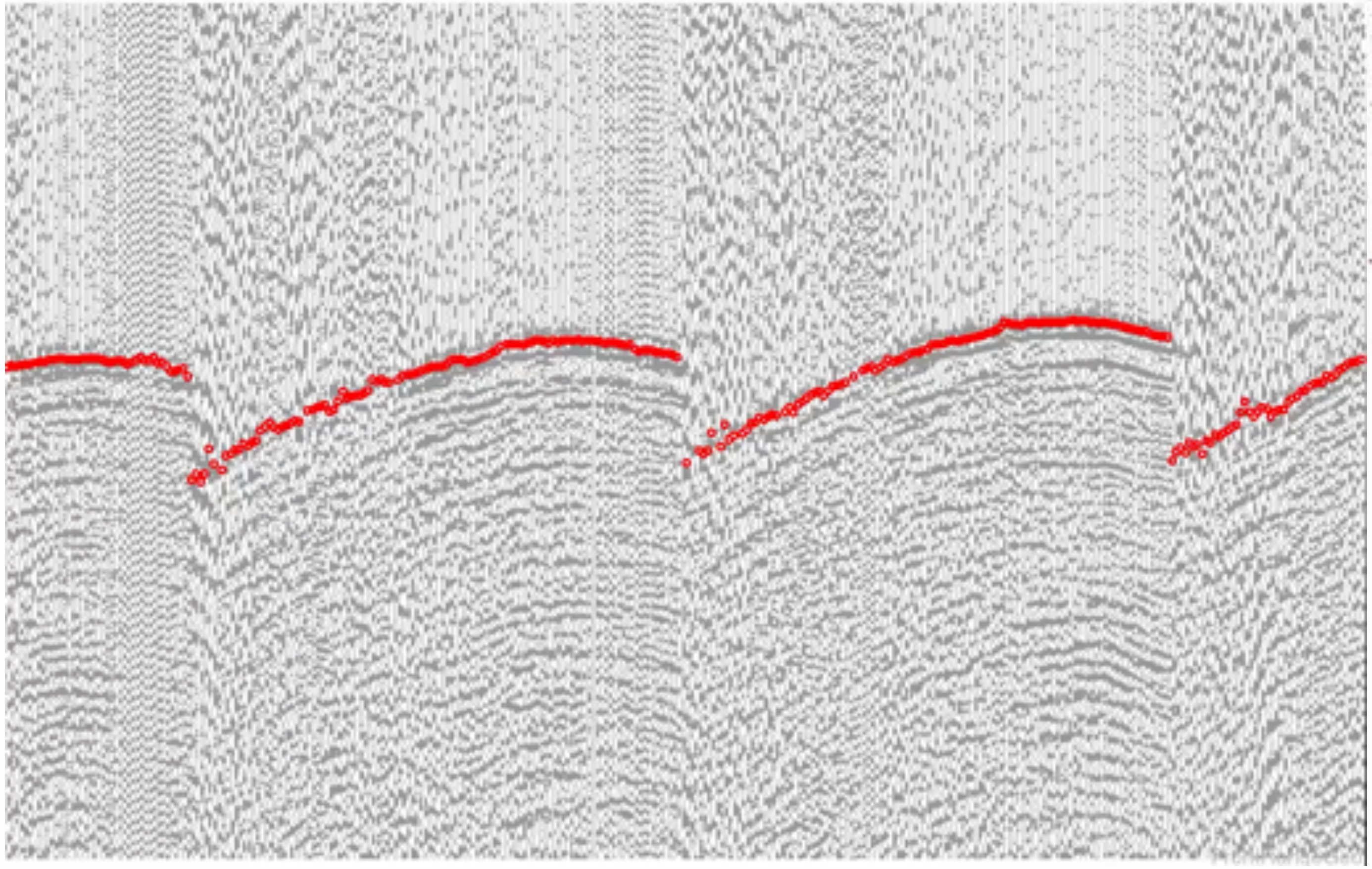
STA/LTA Thresholding

Wong et al. first proposed the STA/LTA method to find first breaks.

Their method involves calculating many STA/LTA windows per trace to find a first break.

We calculate the STA/LTA only once around the DeepTrace pick to use as a quantitative measure of confidence.

Users can specify an STA/LTA confidence threshold that the DeepTrace pick must satisfy to be considered valid.



Video: Killing picks based on varying STA/LTA threshold.

Based on the methods of:
Wong et al. *Automatic time-picking of first arrivals on noisy microseismic data*. 2009

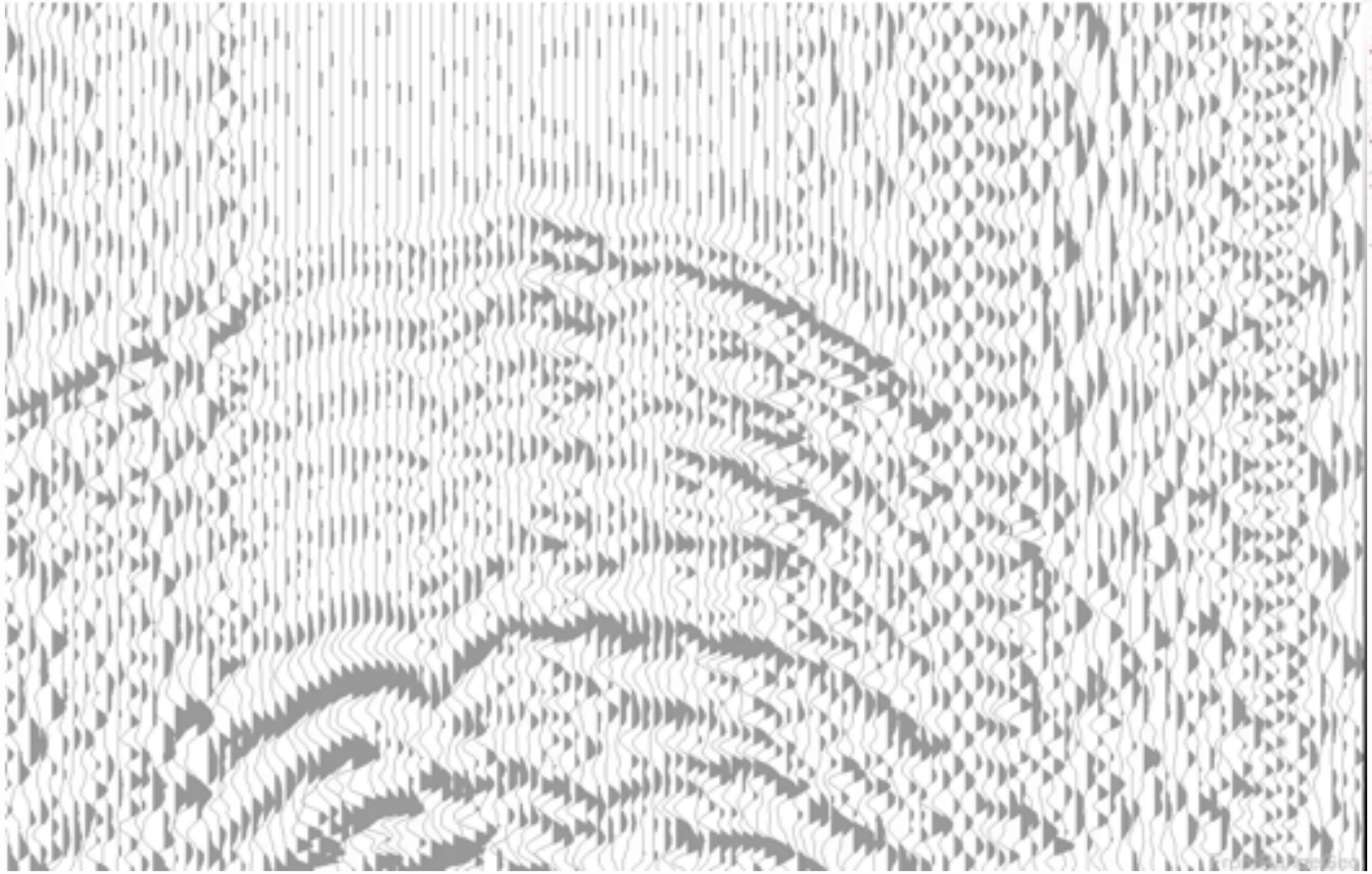
Model Ensembling

Accuracy increases by averaging over multiple models (generic deep learning result).

Variance of independent model picks gives another measure of confidence.

This distribution also conveniently solves our issue of assigning a single scalar value to a wavelet arrival.

Model ensembling is computationally expensive.



DeepTrace Model #1
DeepTrace Model #2
DeepTrace Model #3
DeepTrace Average

Summary

- Engineering a solution to first-break picking is nearly impossible.
- Modern compute enables training deep neural networks to learn the best approach to the problem.
- The machine learning paradigm of computing is here to stay. Large labeled data sets are the key to training A.I.
- Physics modeling and A.I. work together to give the best first-break picking results.

Thank you

Q&A?

We thank Lenovo for the opportunity to present at SEG 2019.

Contact: demoss@frontrangegeosciences.com