Real-Time Geological Inversion for Subsurface Decision-Making

Sergey Alayev saly@norceresearch.no

Outline

- Lecture 1: Geological inversion for geosteering
    - Overview of geosteering
    - Different types of data and inverse problems
    - Reference problem of sequential inversion
    - Simple inversion
    - Regularization
- Exercise
- Lecture 3: Filtering and multi-modal inversion
    - Particle filter
    - Direct multimodal inversion with machine learning

Acknowledgments

DigiWells

NORCE
Universitetet i Stavanger
NTNU
Universitas Bergensis
AkerBP
ConocoPhillips
equinor
Harbour Energy
BR PETROBRAS
TotalEnergies
vår energi
KONGSBERG
ODFJELL DRILLING
Sekal
Leland Stanford Junior University
The University of Texas at Austin
The Research Council of Norway
sfi Centre for Research-based Innovation
The Research Council of Norway

Distinguish
AkerBP
equinor
The Research Council of Norway

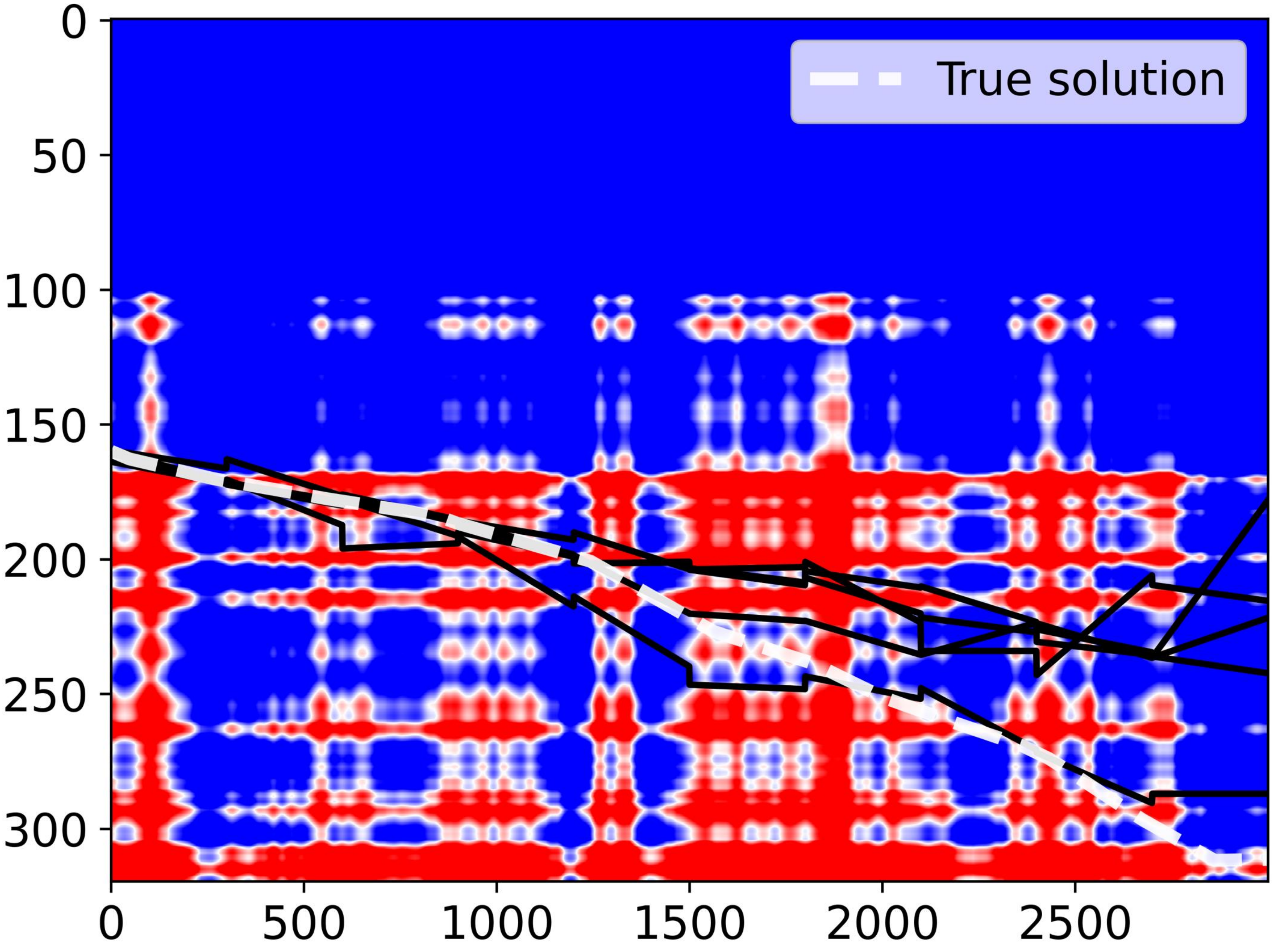
Thanks

- Organizers – for invitation
- Anna Kvashchuk – taking the kids and feedback on content
- Kristian Fossum – feedback on content
- Nazanin Jahani – slides
- Ressi Muhammad – slides
- Yasaman Cheraghi – slides
- ROGII Inc. – for Geosteering World Cup Data and animations

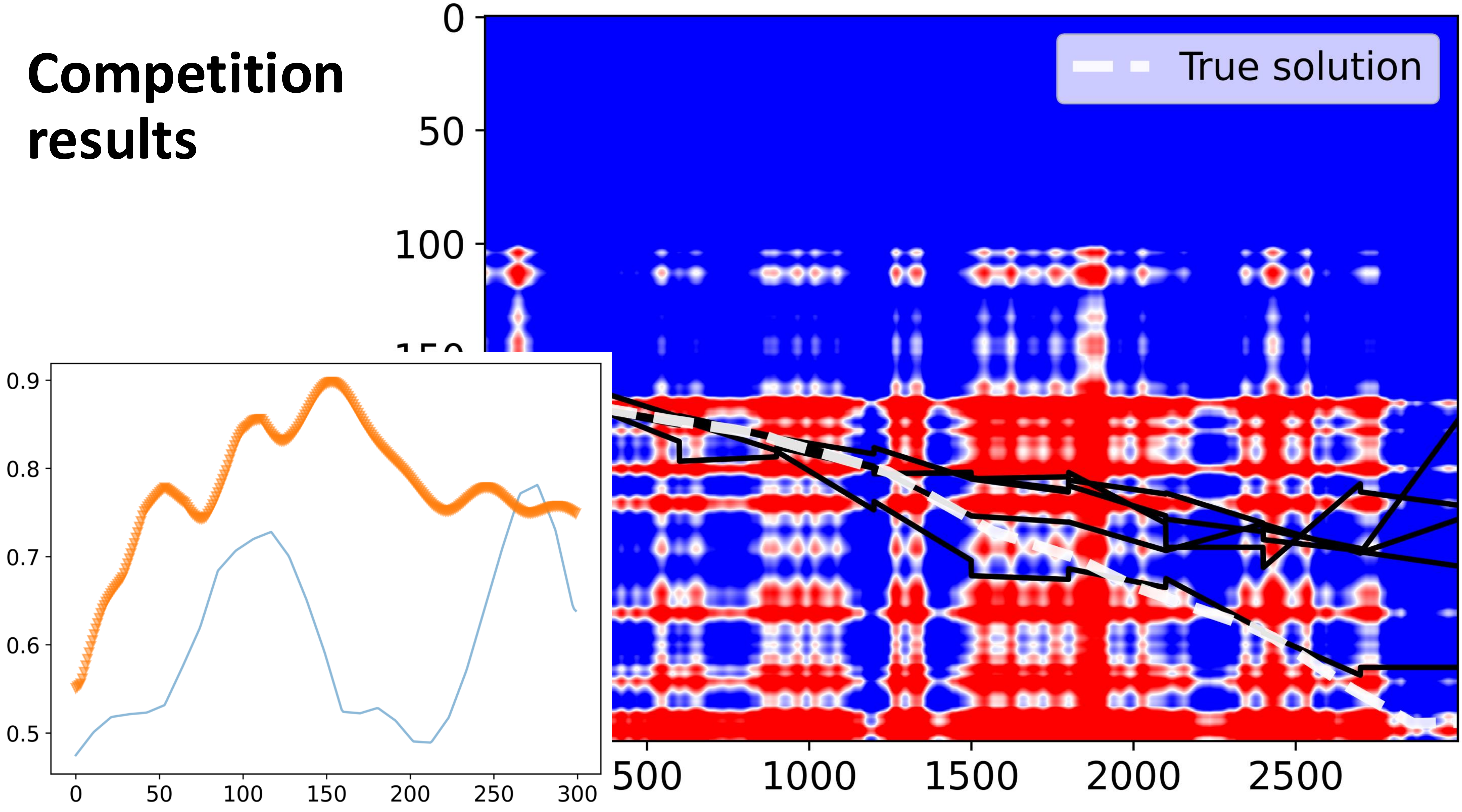
Exercise 1: results

Stratigraphic inversion

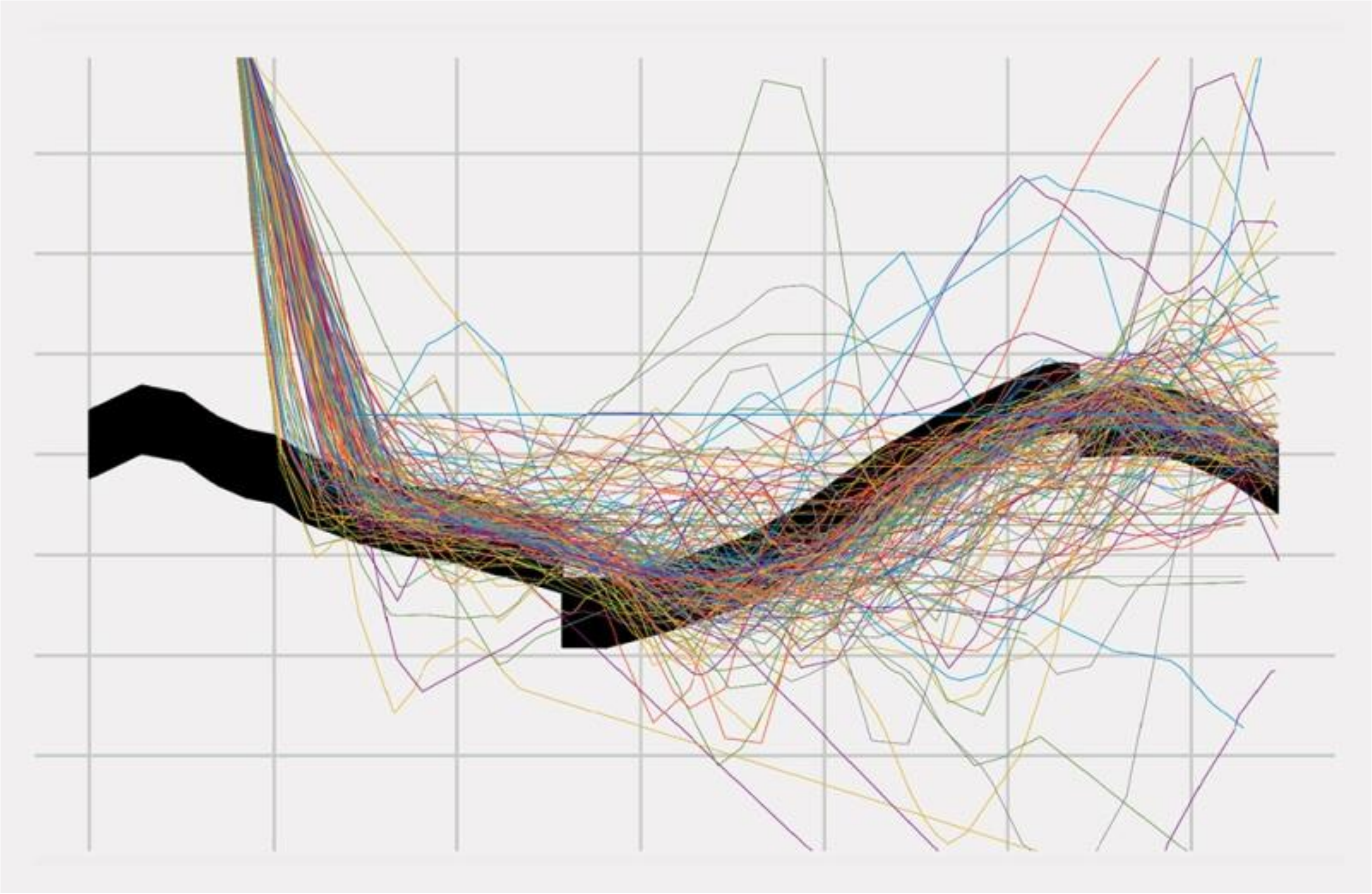
Competition results



Competition results



Geosteering decisions are challanging

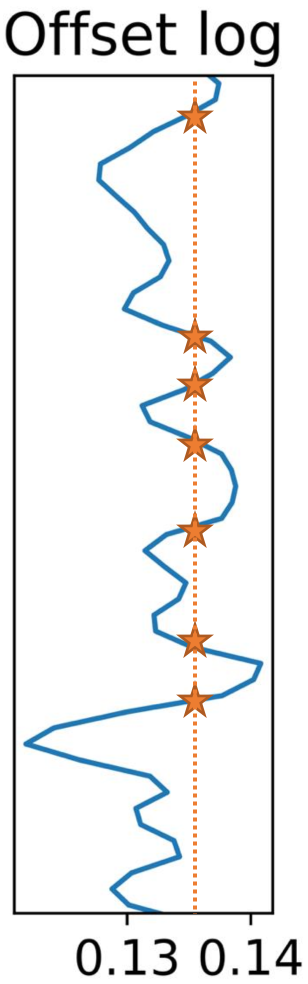


[What Can We Learn After 10,000 Geosteering Decisions? | SPE/AAPG/SEG Unconventional Resources Technology Conference | OnePetro]

Lecture 3.1: Filtering problem

Difficulty with low-data geosteering

Interpretations
of a single data point



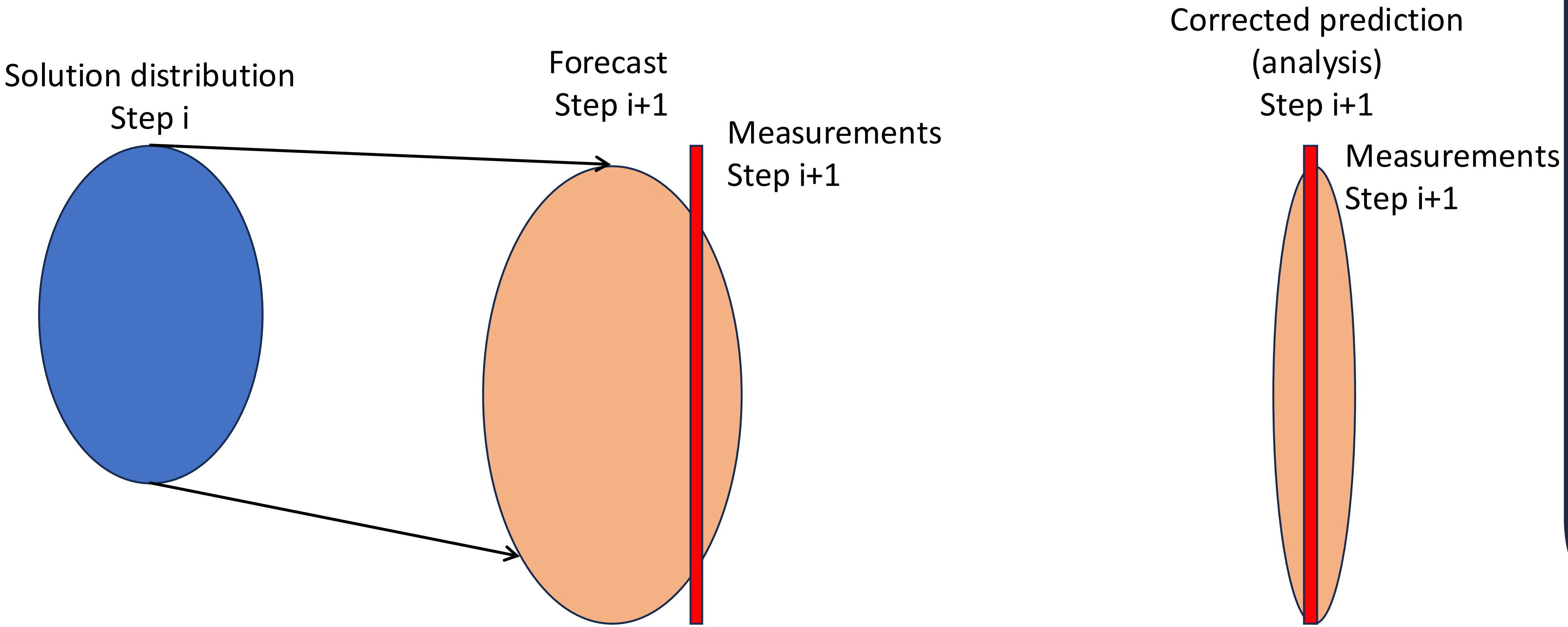
- Log data has multiple depth interpretations
- Picking one interpretation or averaging leads to errors

Multi-modal probabilistic inversion
can improve

- Geological understanding
- Geosteering decisions

Filtering problem

determining the state of a system from an incomplete and potentially noisy set of observations



Filtering problem for boundary function

[with some abuse of notation]

- $b^f_{\mathrm{k,i+1}} \equiv b^f_k(x_{i+1}) = B\left(x_i, w_{i,k}\right)$
  - b_k form a distribution of solutions
  - $B\left(x_i, w_{i,k}\right)$ - probabilistic “forward” model of geology
- $f_k(x_{i+1}) = f\left(b_k(x_{i+1})\right) + v_{i+1}$
  - f_k form a distribution of measurement prediction

Filtering problem for boundary function

[with some abuse of notation]

- $b_k(x_{i+1}) = B\left(x_i, w_{i,k}\right)$
  - b_k form a distribution of solutions
  - $B\left(x_i, w_{i,k}\right)$ - probabilistic “forward” model of geology
- $f_k(x_{i+1}) = f\left(b_k(x_{i+1})\right) + v_{i+1}$
  - f_k form a distribution of measurement predictions
  - v_{i+1} - added measurement errors
- $g_{i+1} \equiv g(x_{i+1})$ - observations

Filtering task is to find

- $\widetilde{b_{k,i+1}} \equiv b_k\widetilde{(x_{i+1})} \sim \rho\left(\widetilde{b_{\mathrm{i+1}}} \middle| \, g_{0:\mathrm{i+1}}\right) \propto \rho\left(g_{\mathrm{i+1}} \middle| \tilde{b}^f_{\mathrm{i+1}}\right) \cdot \rho\left(\tilde{b}^f_{\mathrm{i+1}}\right)$

Filtering problem for boundary function

[with some abuse of notation]

Filtering task is to find

- $\widetilde{b_{i+1}} \quad \sim \rho\left(\widetilde{b_{\mathrm{i+1}}} \mid g_{0:\mathrm{i+1}}\right) \propto \rho\left(g_{\mathrm{i+1}} \middle| \tilde{b}^{f}_{\mathrm{i+1}}\right) \cdot \rho\left(\tilde{b}^{f}_{\mathrm{i+1}}\right)$

Density of New distribution of predictions corrected to data

Likelihood of data *g* given the *forecast*

Density of the forecast distribution

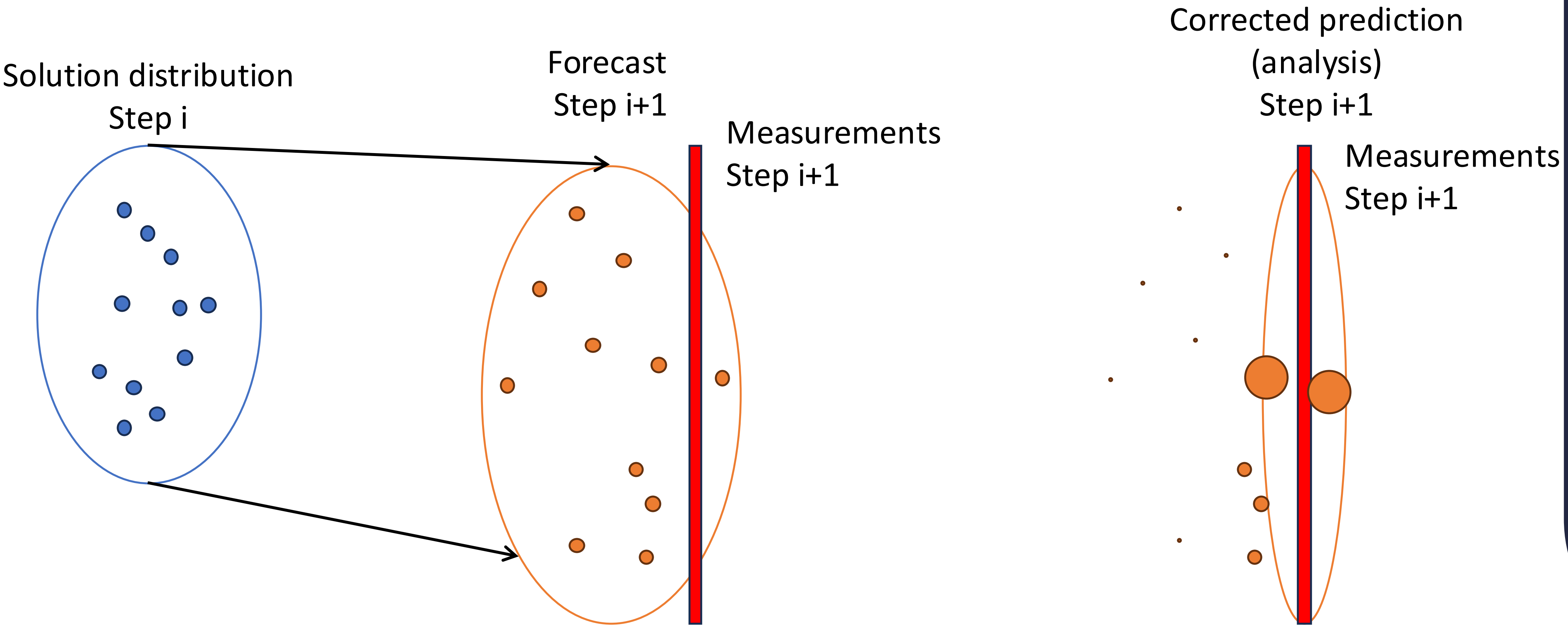
Bayes' theorem:

$$\boxed{P(A|B) = \frac{P(B|A)P(A)}{P(B)}}$$

Particle filter

Move particles individually

Update probabilities



Particle filter

Propagation

- $b^{f}_{\mathrm{k,i+1}} \equiv b^{f}_{k}(x_{i+1}) = B\left(x_i, w_{i,k}\right)$

Measurement estimation

- $f_k(x_{i+1}) = f\left(b_k(x_{i+1})\right) + v_{i+1}$

Update of weights

- $w_{k,i+1} \propto w_{k,i+1} C \exp\left(-\frac{\left(f_{k,i+1} - g_{i+1}\right)^2}{2\sigma^2}\right)$

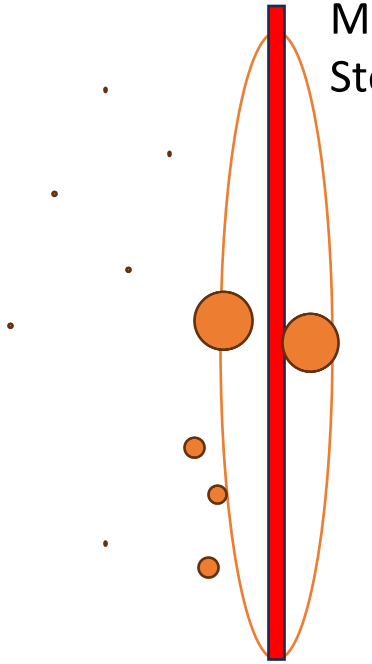
Gaussian distance to measurements

There is a problem with the new distribution

Update probabilities

Corrected prediction
(analysis)
Step i+1

Measurements
Step i+1

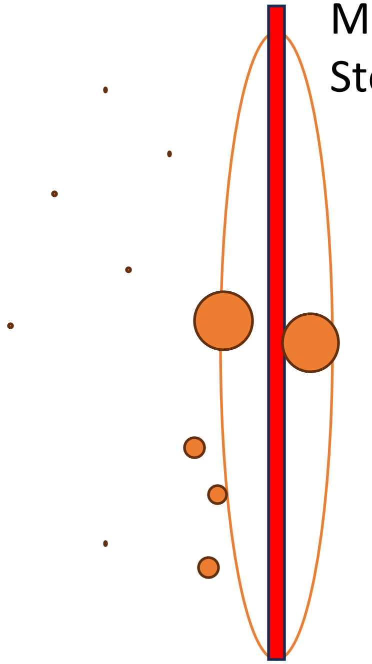


There is a problem with the new distribution

Update probabilities

Corrected prediction
(analysis)
Step i+1

Measurements
Step i+1

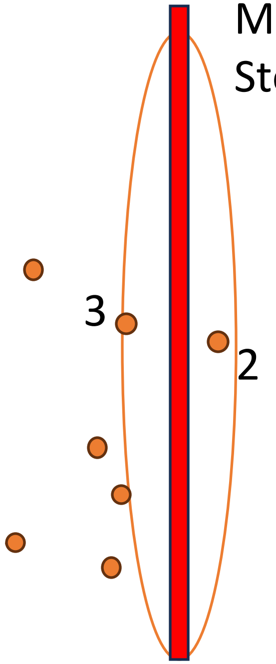


Resampling

Update probabilities

Corrected prediction
(analysis)
Step i+1

Measurements
Step i+1



Resampling

- Pick particles proportional to their weight
- The duplicates will naturally drift apart during next step

A better more detailed tutorial on particle filters:

- https://ieeexplore.ieee.org/document/978374

A filtering library from colleagues in Bergen:

- https://github.com/nansencenter/DAPPER

Filtering for the geosteering problem

- Forecast
  - Predictive model of geology
  - Measurement model
- Analysis
- Resampling

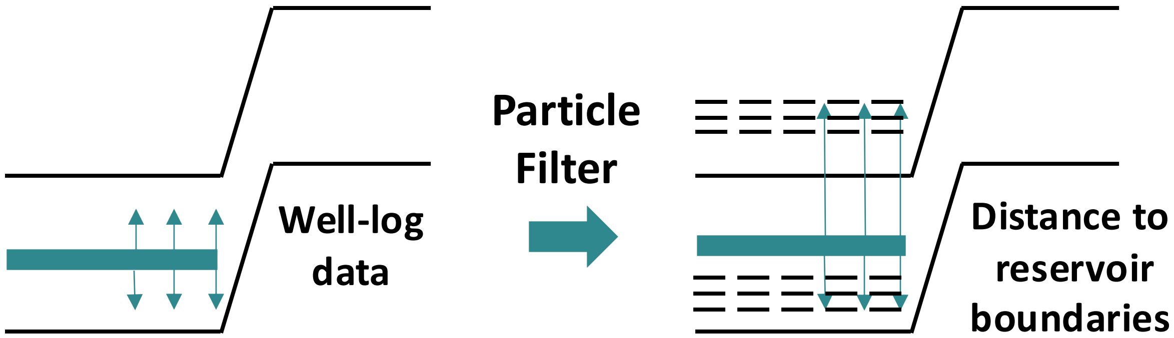
- **Tricky**
  - **Not available**
  - **Simple for low-data**
- **Standard**
- **Standard**

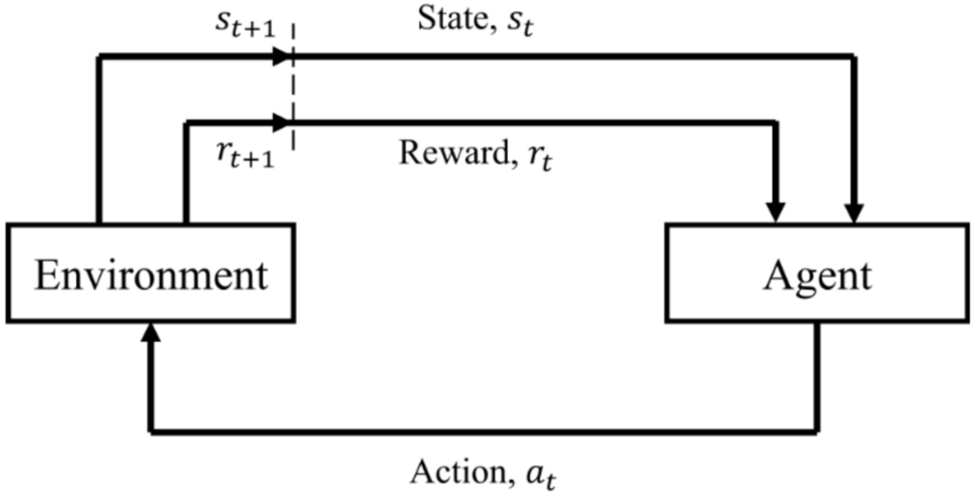
PF-based Pluralistic Geosteering Robot

- From papers
  - by Ressi B. Muhammad; Yasaman Cheraghi; Sergey Alyaev; Apoorv Srivastava; Reidar B. Bratvold,
  - SPE J. (2025) https://doi.org/10.2118/218444-PA
  - Conference (2024) https://dx.doi.org/10.2118/218444-MS

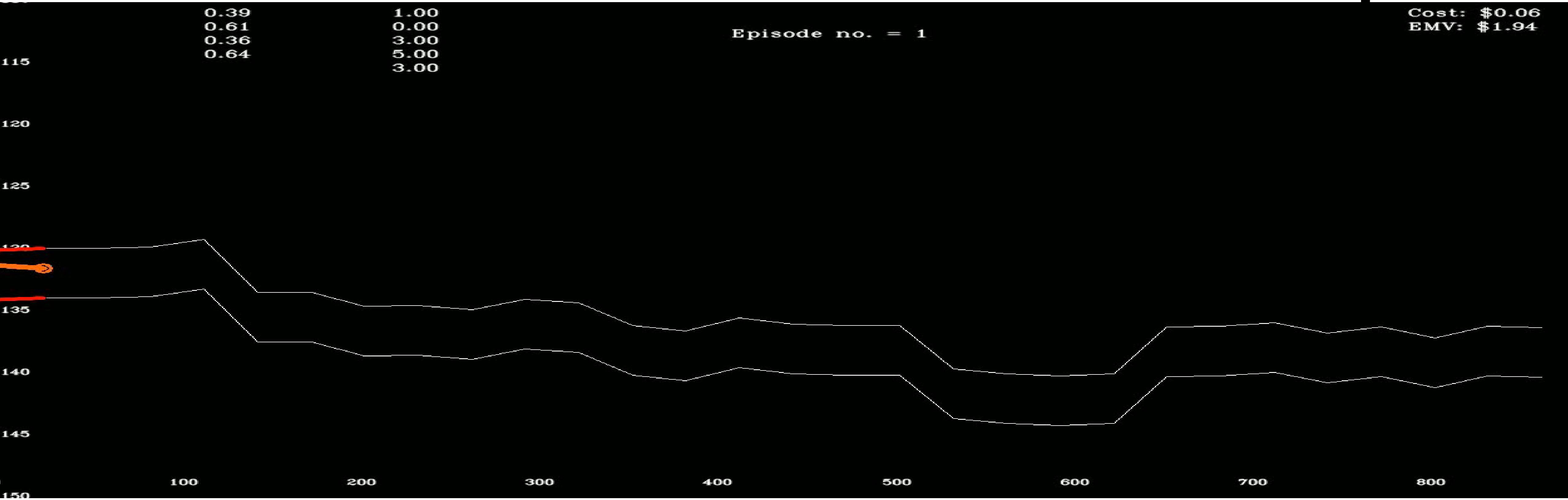
Pluralistic Geosteering Robot

- Pluralistic robot parts
- **PF** processes real-time well-log data **to make inferences about stratigraphic layers** around the bit
- **RL** uses multiple estimates **to make decisions**





RL learning process on a simple example



[2310.04772] Optimal Sequential Decision-Making in Geosteering: A Reinforcement Learning Approach (arxiv.org)

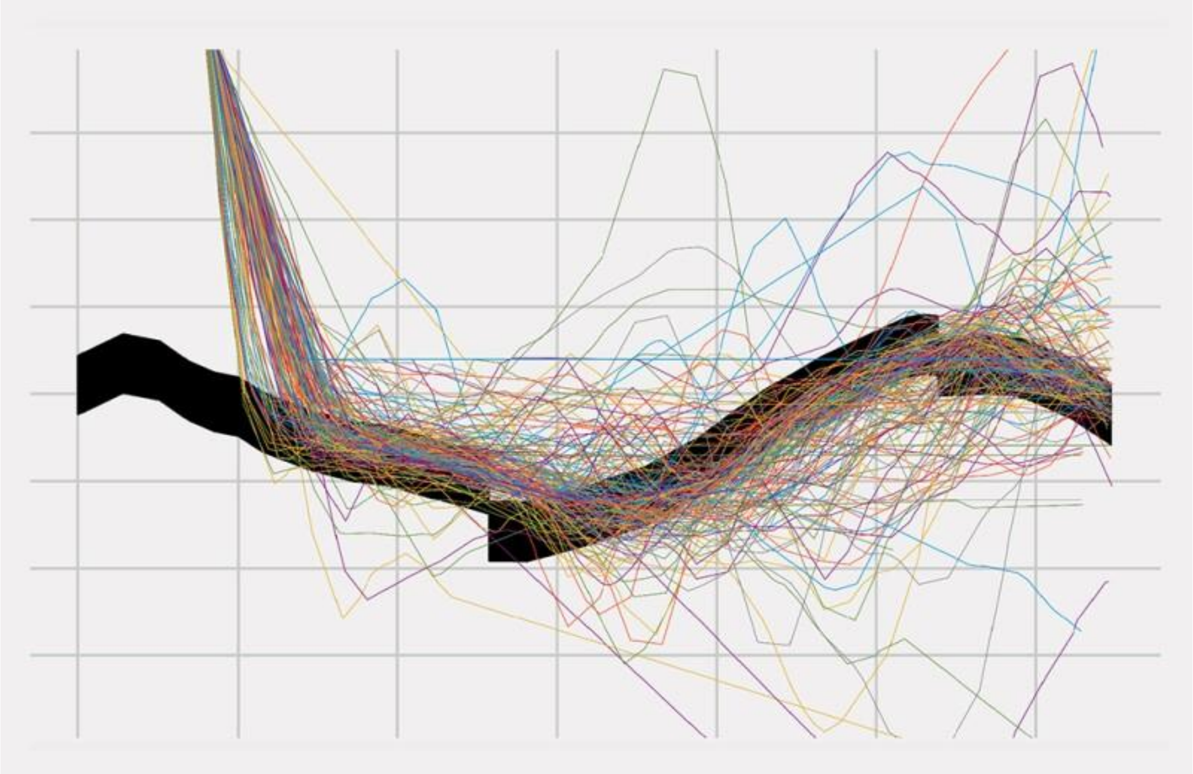
PF + RL beats
PF + Rule-based methdod
Pure RL-based AI solution

Methods	Input	Rewards	Reservoir contact (%)	MAE of Input
Rule-based[13]	1 Look-ahead	-101.36	54.75	4.27
Rule-based[13]	5 Look-ahead	-105.95	52.70	4.27
RL-Log	Gamma-ray	-74.27	66.84	-
RL-Estimation	1 PF	-25.36	88.67	1.29
RL-Estimation	5 PF	-24.99	88.91	1.31

[2402.06377] High-Precision Geosteering via Reinforcement Learning and Particle Filters (arxiv.org)

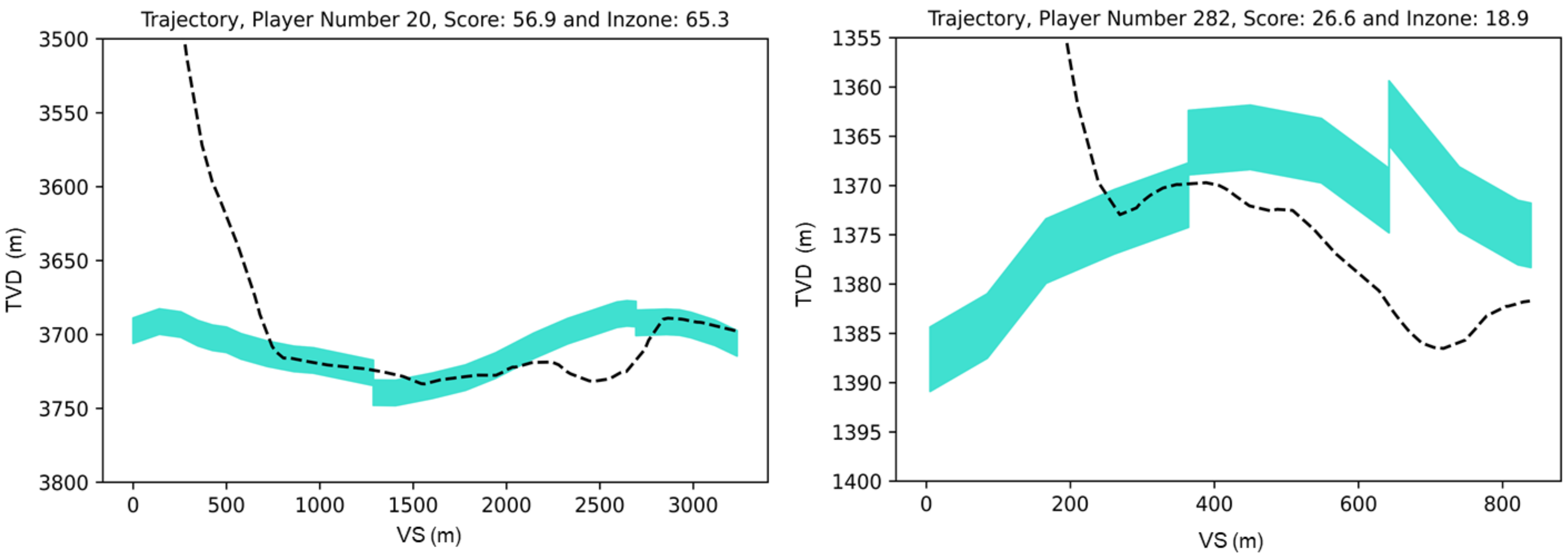
Comparison Against Human Experts

- Challenge: Surpass the performance of **the top quartile** of competitors **in the Geosteering World Cup (GWC) 2021**.



[What Can We Learn After 10,000 Geosteering Decisions? | SPE/AAPG/SEG Unconventional Resources Technology Conference | OnePetro]

Training data from geosteering experts



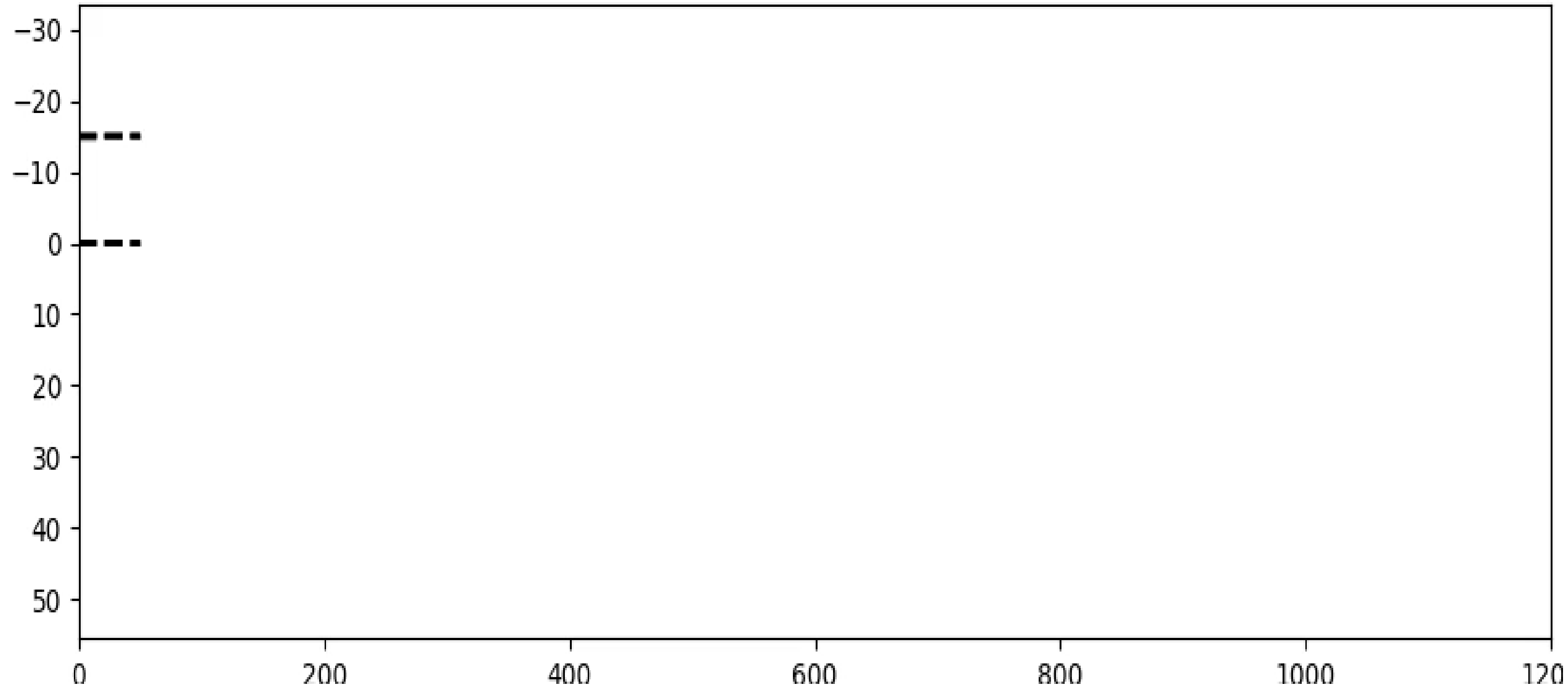
(a) (b)

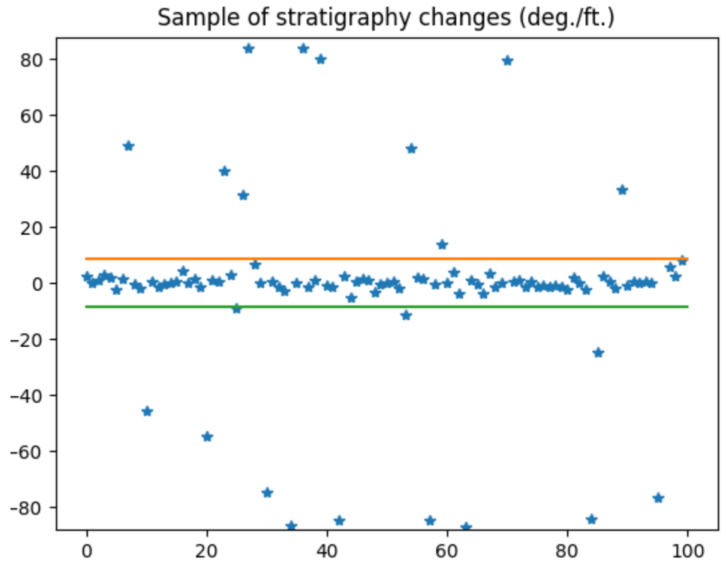
- Over 100 participants
- Over 50 decisions to make
- Almost 10000 near-well interpretations

[What Can We Learn After 10,000 Geosteering Decisions? | SPE/AAPG/SEG Unconventional Resources Technology Conference | OnePetro]

Stochastic Model for Geological Changes

- Kernel Density Estimation (KDE) based on geological interpretations made by experts during GWC 2021.





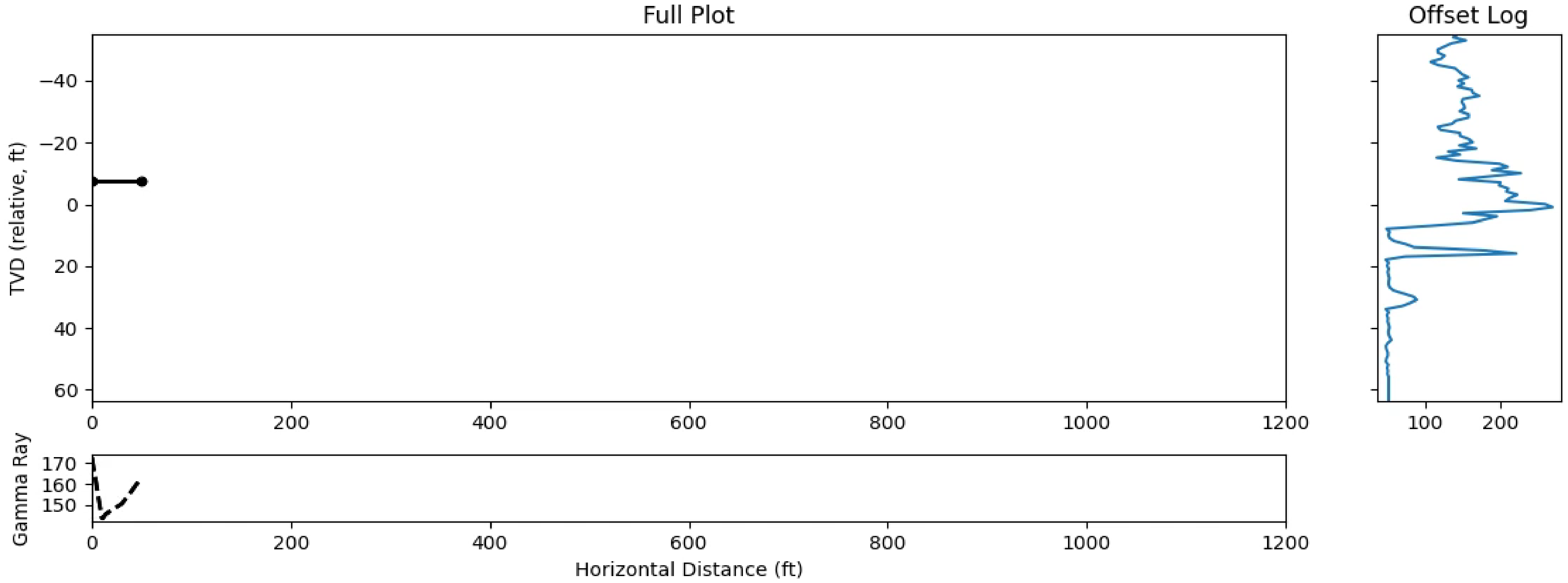
Can you do better?

- As a mathematician I would like to know that too…
- Data competition for generating 1D geology on Kaggle:
  - Small monetary prizes
  - Possibility to publish your work in a benchmark journal article
  - Submission deadline in the end of Spring Semester 2025

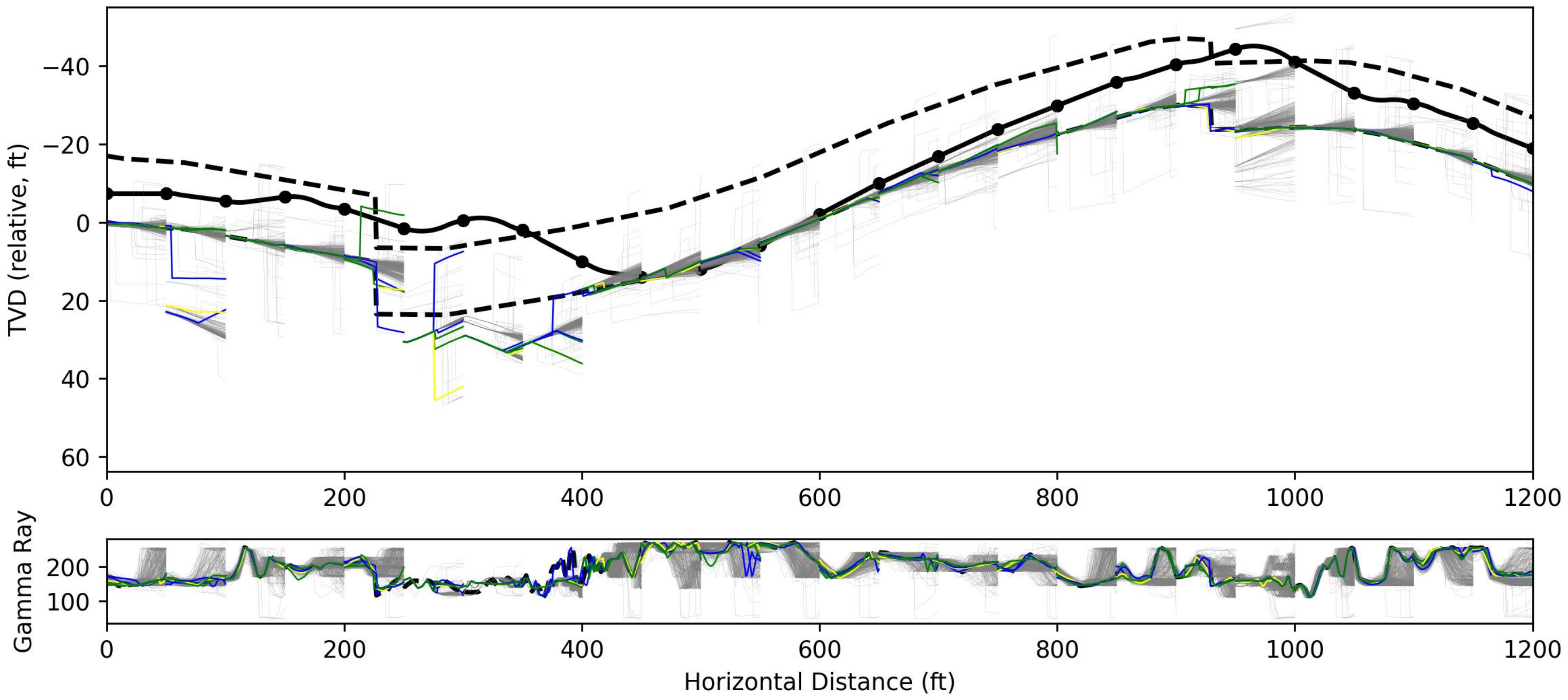
https://www.kaggle.com/competitions/geology-forecast-challenge/overview

Results when starting from within the layer

Illustration of “Pluralistic” Robot



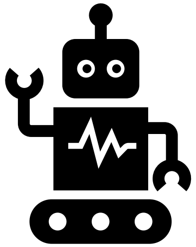
Results when starting from within the layer

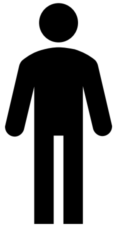


Stability of the robot

- In mathematics, **stability** refers to a property of a system to **remain unchanged or return to equilibrium** after experiencing small perturbations or disturbances.

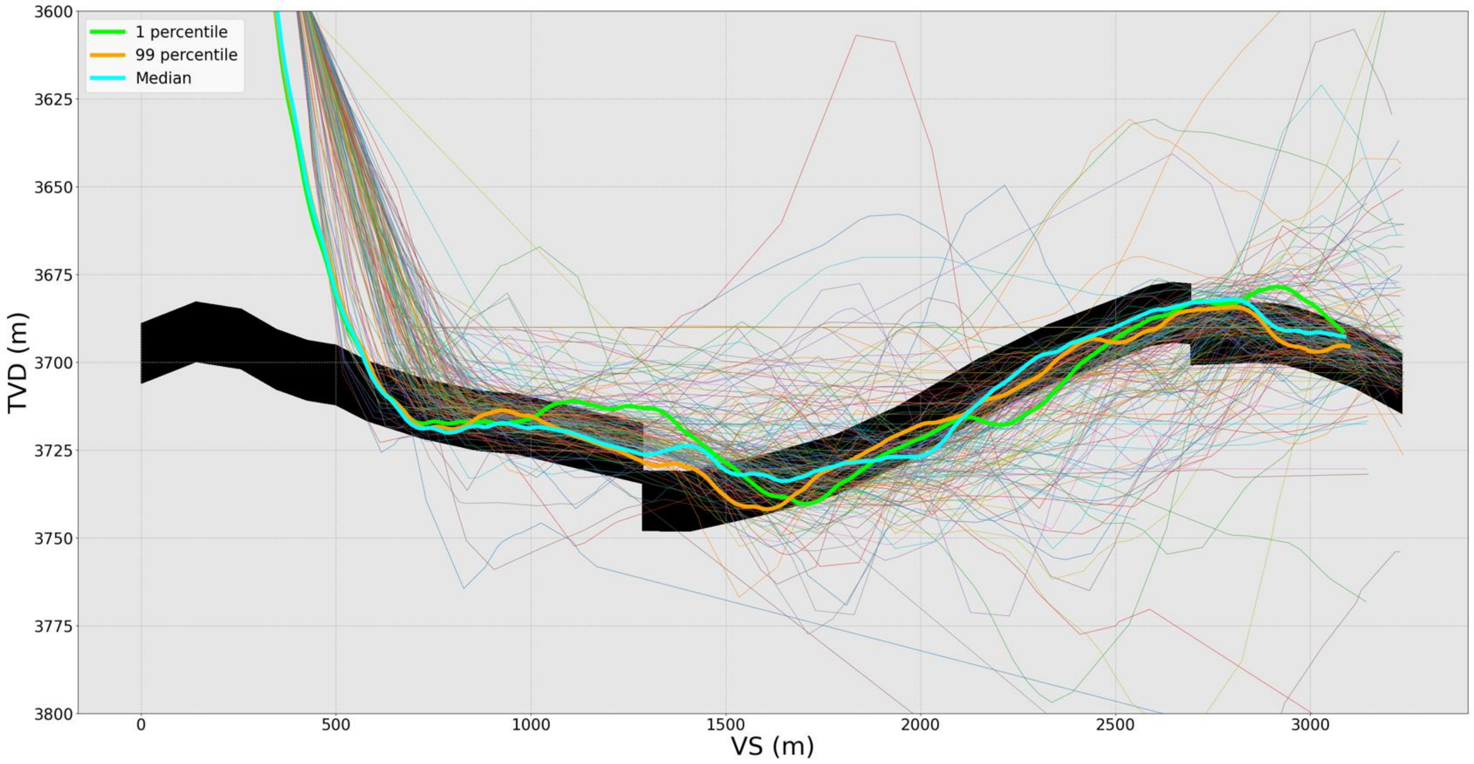
Does the robot give **consistent and stable** results?





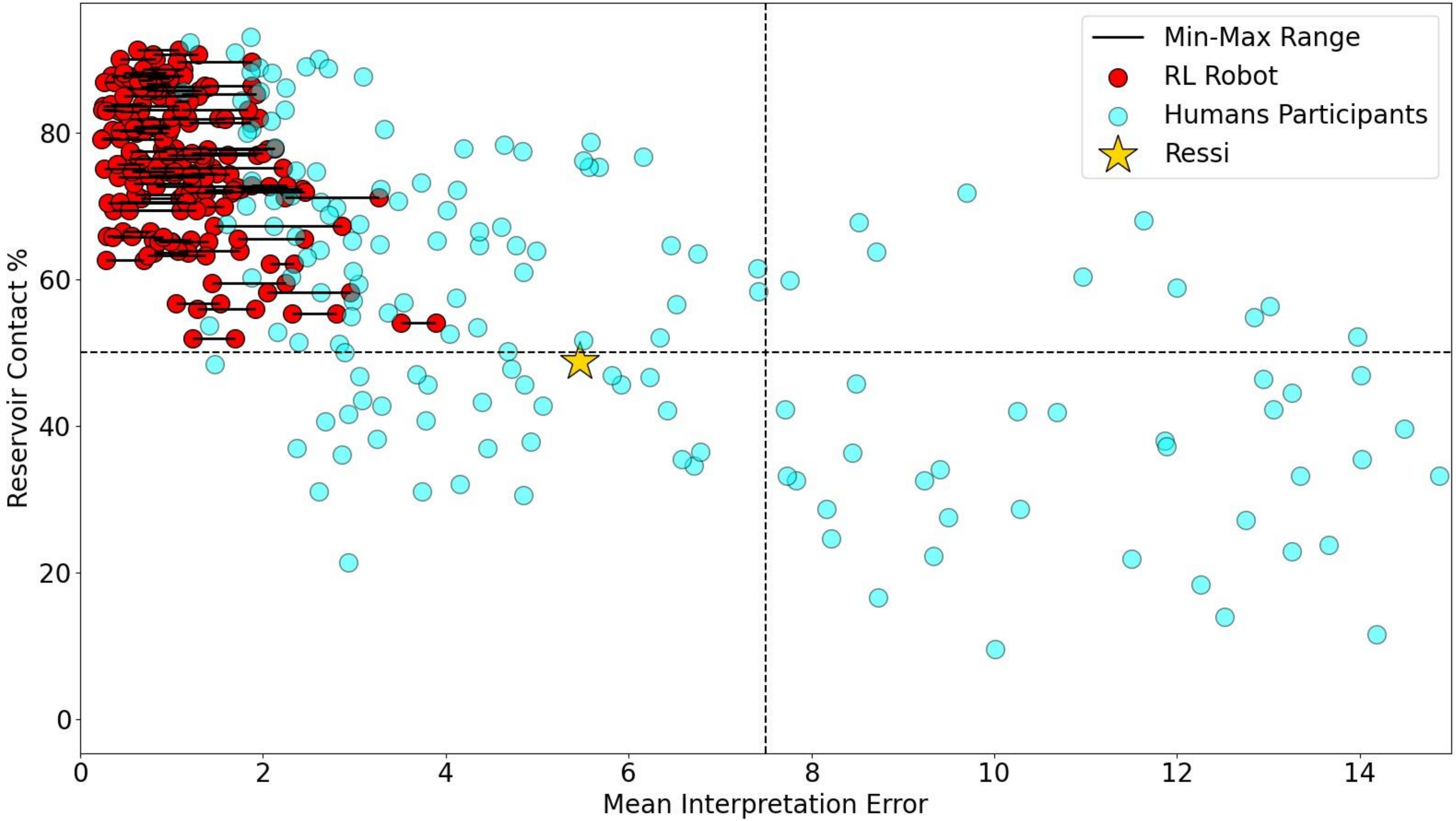
(humans do not)

3 Trajectories Comparison



1st percentile fails at all sections (first fault, incline, second fault) due to PF misinterpretations.

99th percentile manages to steer inside the layer except for the first fault



Min-Max Range
RL Robot
Humans Participants
Ressi
Reservoir Contact %
Mean Interpretation Error
0
20
40
60
80
0
2
4
6
8
10
12
14

Conclusions

- Pluralistic robot
  - **can outperform geosteering experts (with some cheating)**
  - **does outperform them most of the time**
- In low-data geosteering contexts **luck influences robots** as well as humans in terms of decision outcomes

- Ongoing work:
  - Make the robot work in the actual World Cup environment using API

Lecture 3.2: Direct Multi-modal inversion with machine learning

Direct Multi-modal inversion with machine learning for geosteering

Our MDN paper 2022

“Direct multi-modal inversion of geophysical logs using deep learning”

S Alyaev, AH Elsheikh

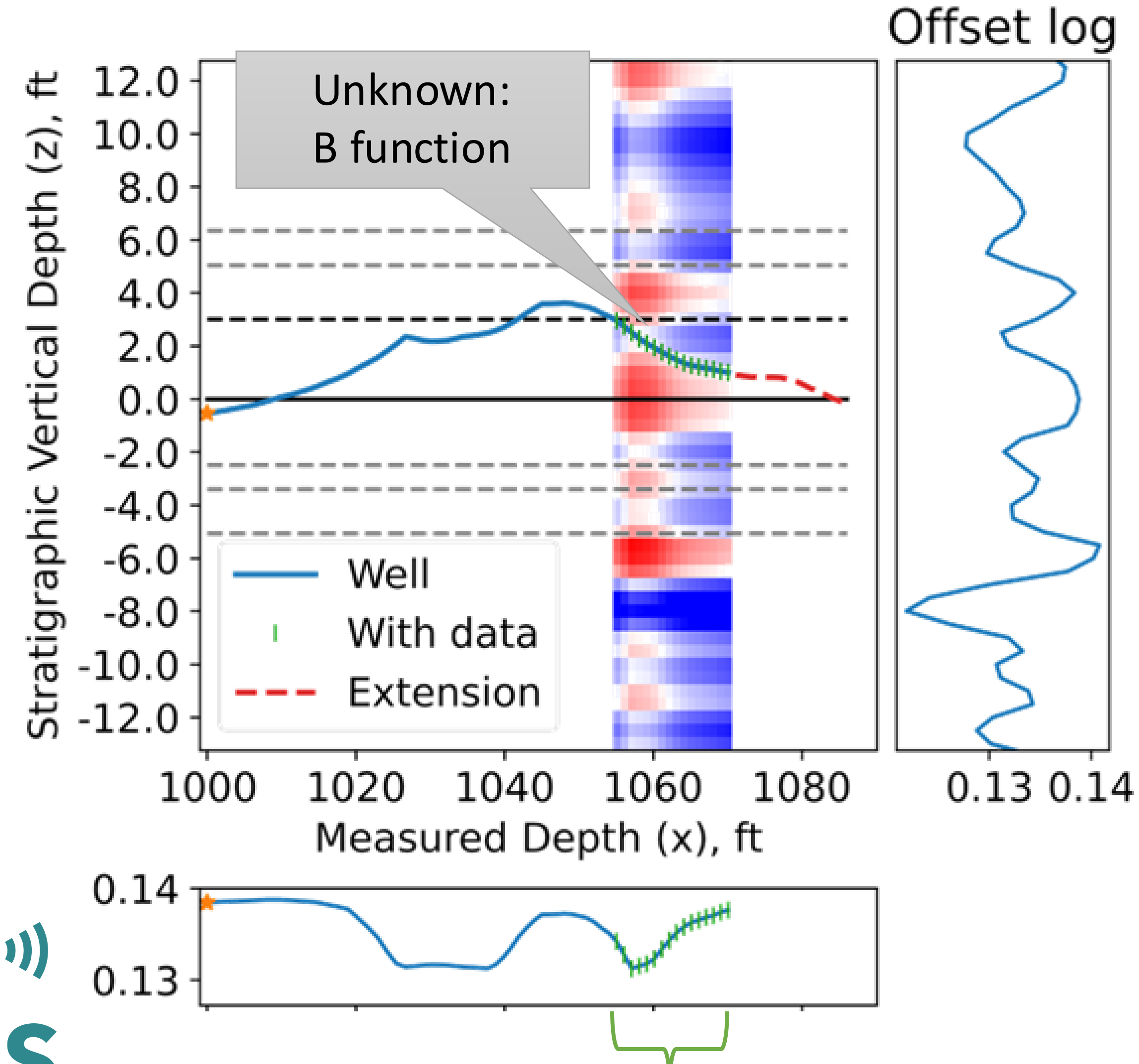
[Earth and Space Science, 2022]

Our sequencial paper SPWLA-2022-0112

“Sequential Multi-Realization…”

S Alyaev; A Ambrus; N Jahani; AH Elsheikh

Inputs and Outputs for Direct Multi-modal DNN



Find the well trajectory **B function** relative geology

Outputs of the Mixed Density Network

- *N* likely B functions
  - Along data: INVERSION
  - Ahead of data: PREDICTION
- *N* probability values

Inputs

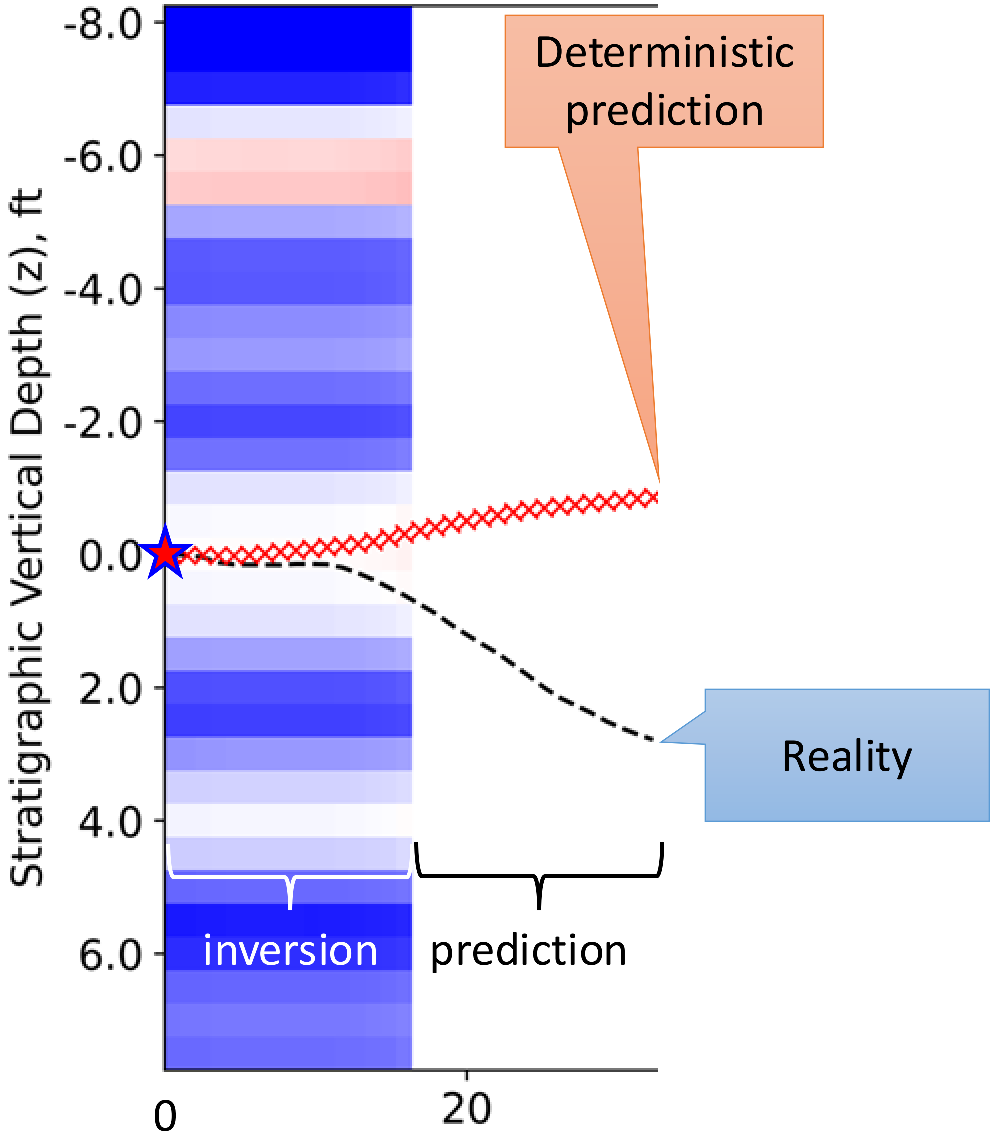
Offset well log

Current well log

→ **Heatmap image**

Examples of deterministic prediction

Geology



Weather

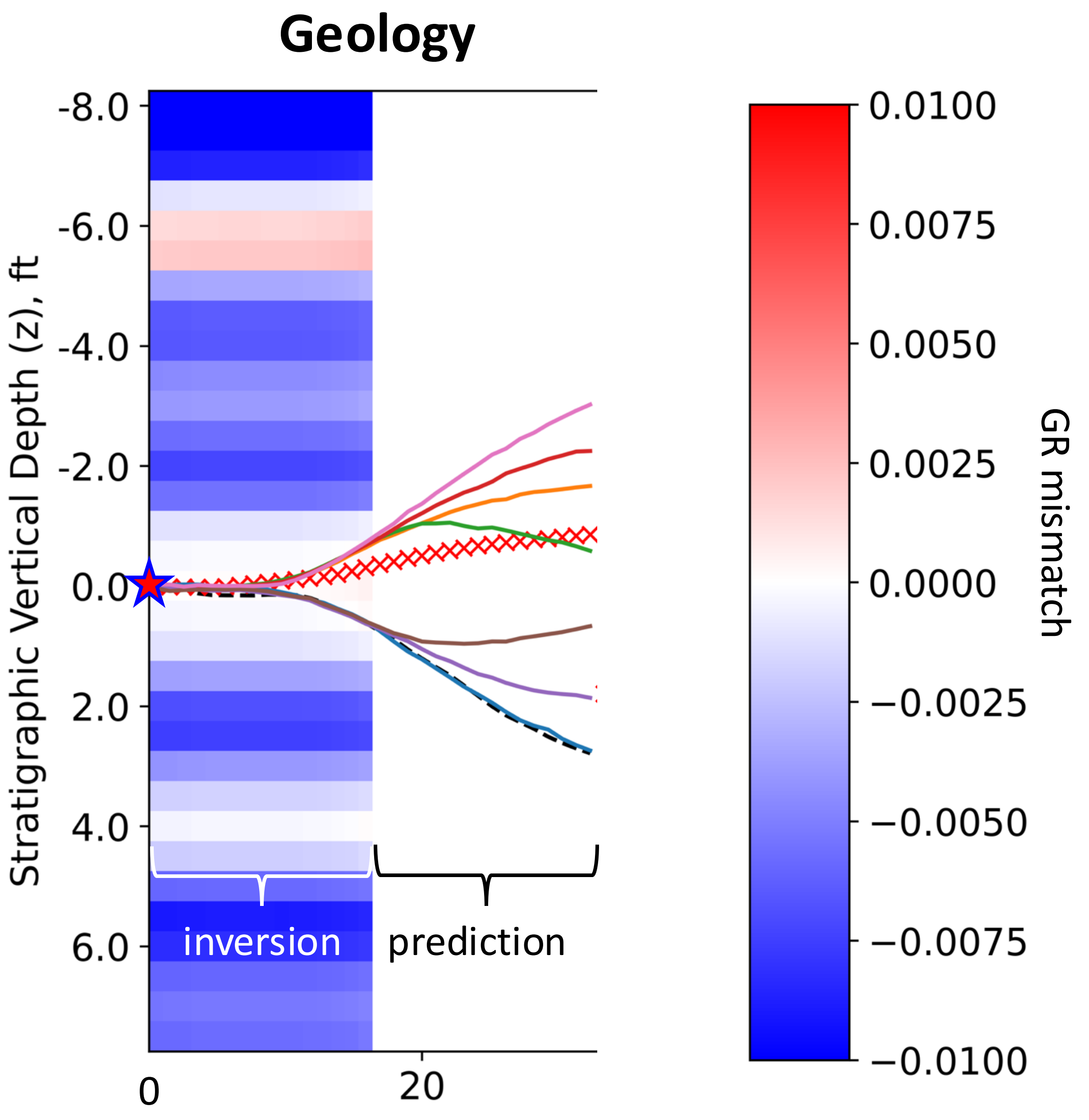
Deterministic prediction

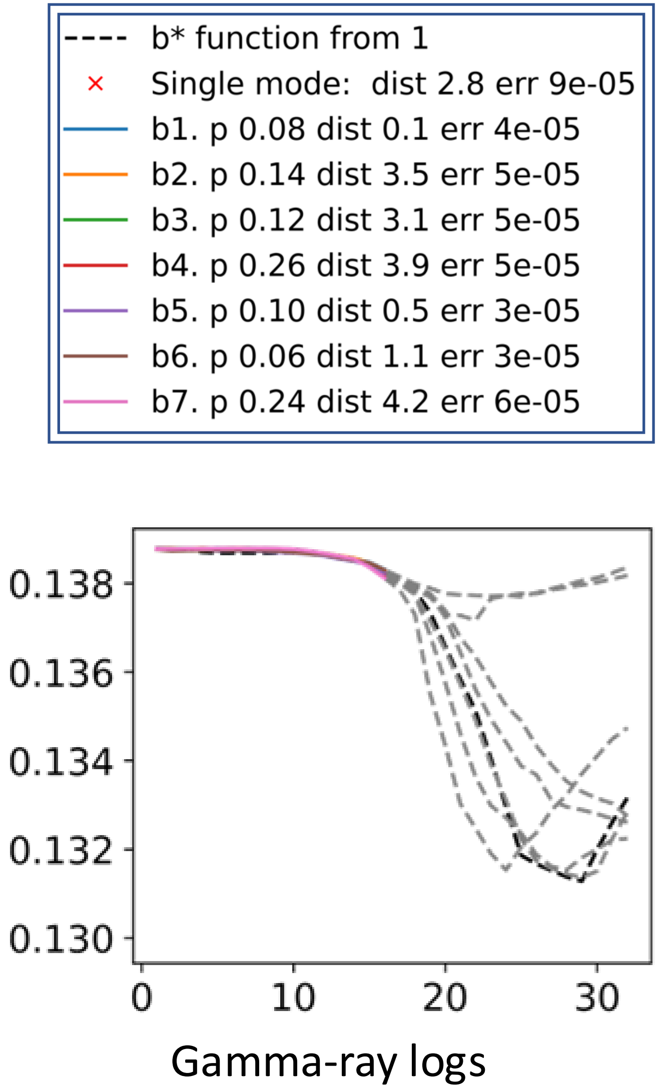
14	-5°	0.2–0.8	2 (4)
15	-5°	0.3–0.9	2 (4)
16	-5°	0.4–0.9	1 (4)
17	-6°	0.4–0.9	1 (3)

Reality

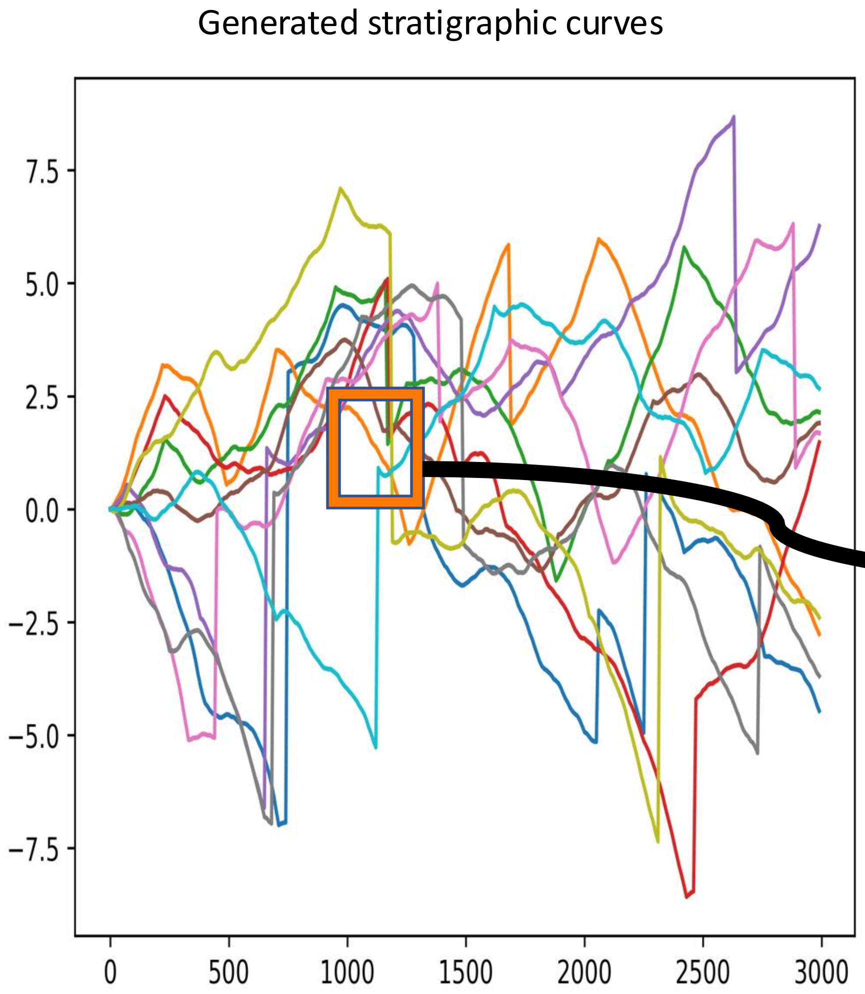
Go out and check during a break

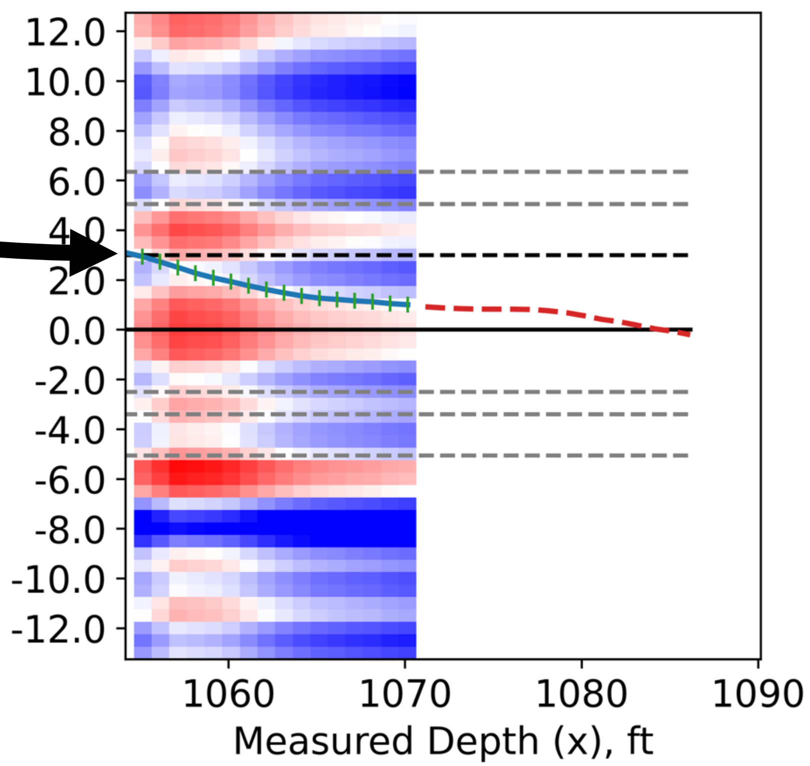
Seven-mode inversion of a segment with MDN

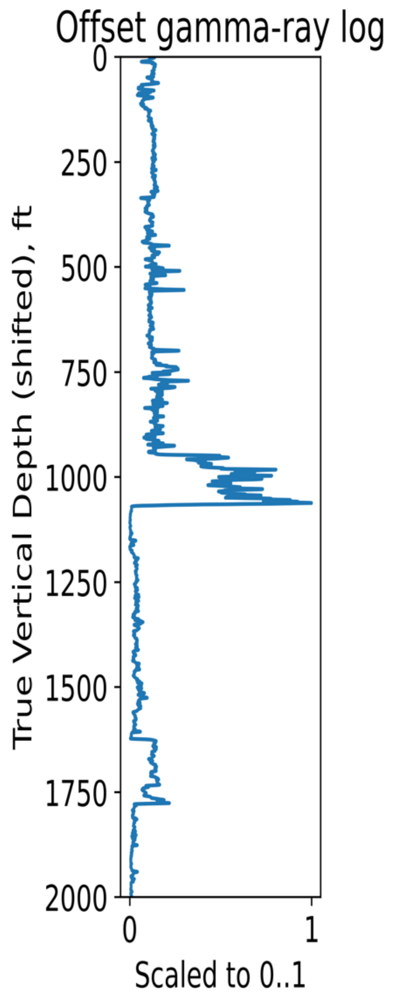




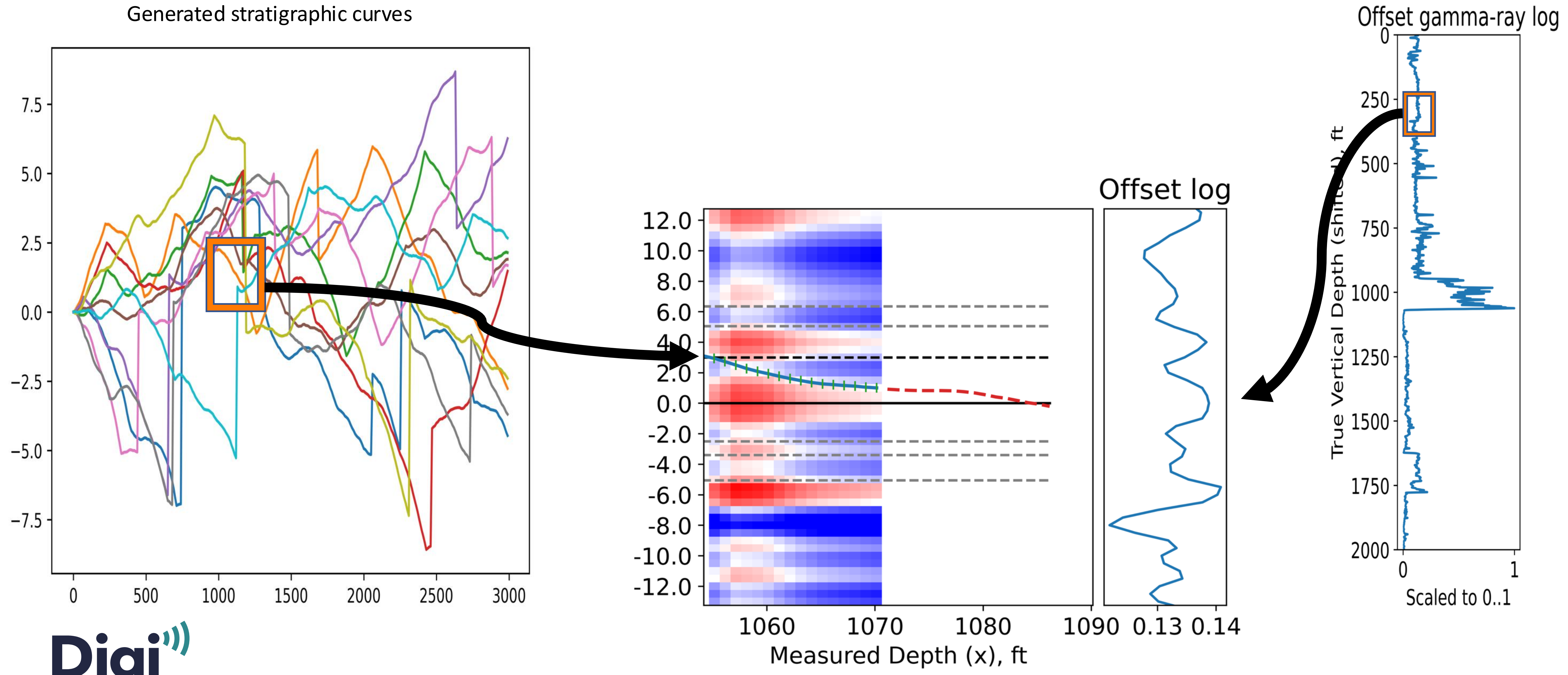
How did the model learn? Training data





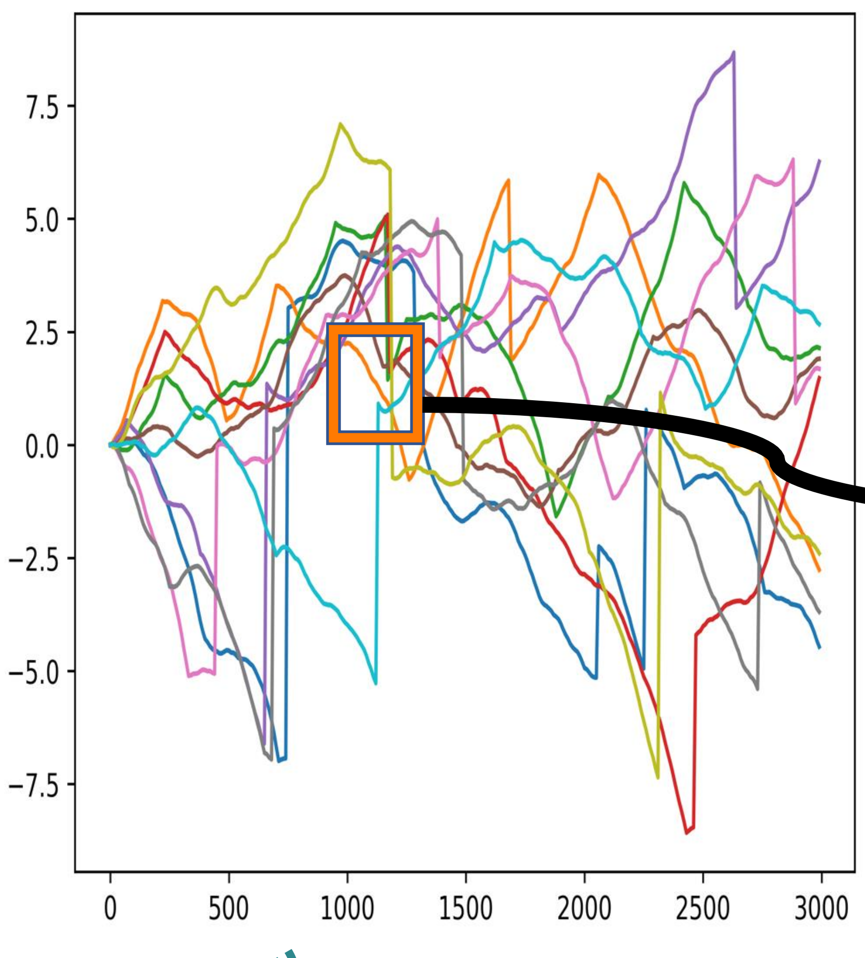


How did the model learn? Training data

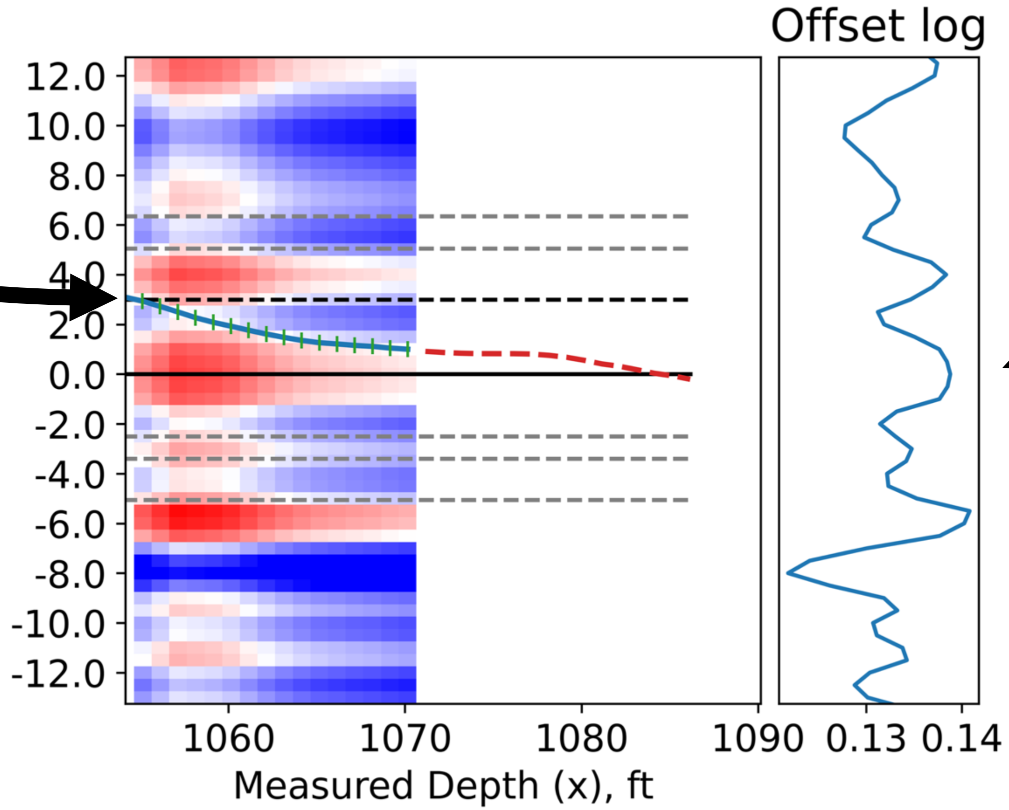


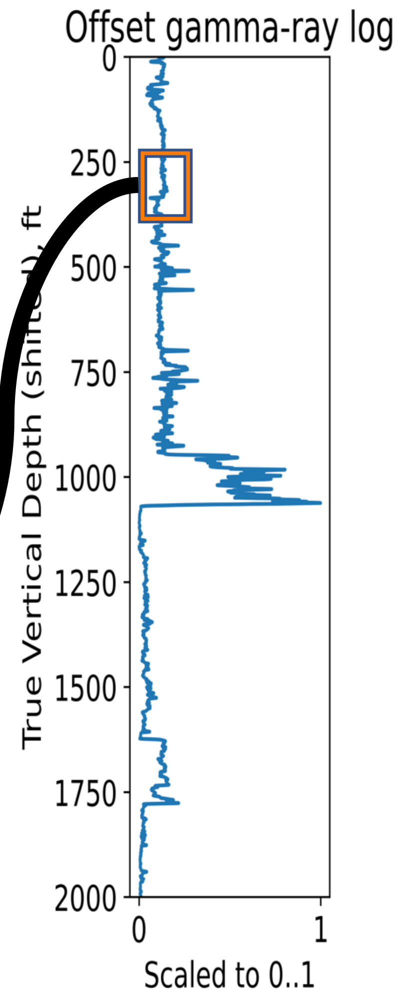
How did the model learn? Training data

Generated stratigraphic curves

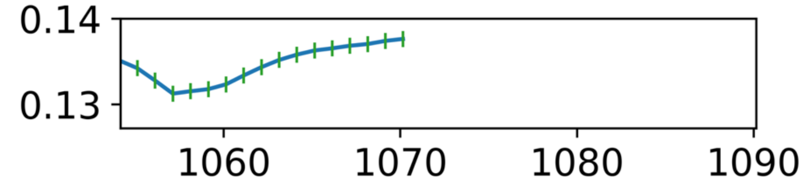


Training data:
28M log-log-curve samples



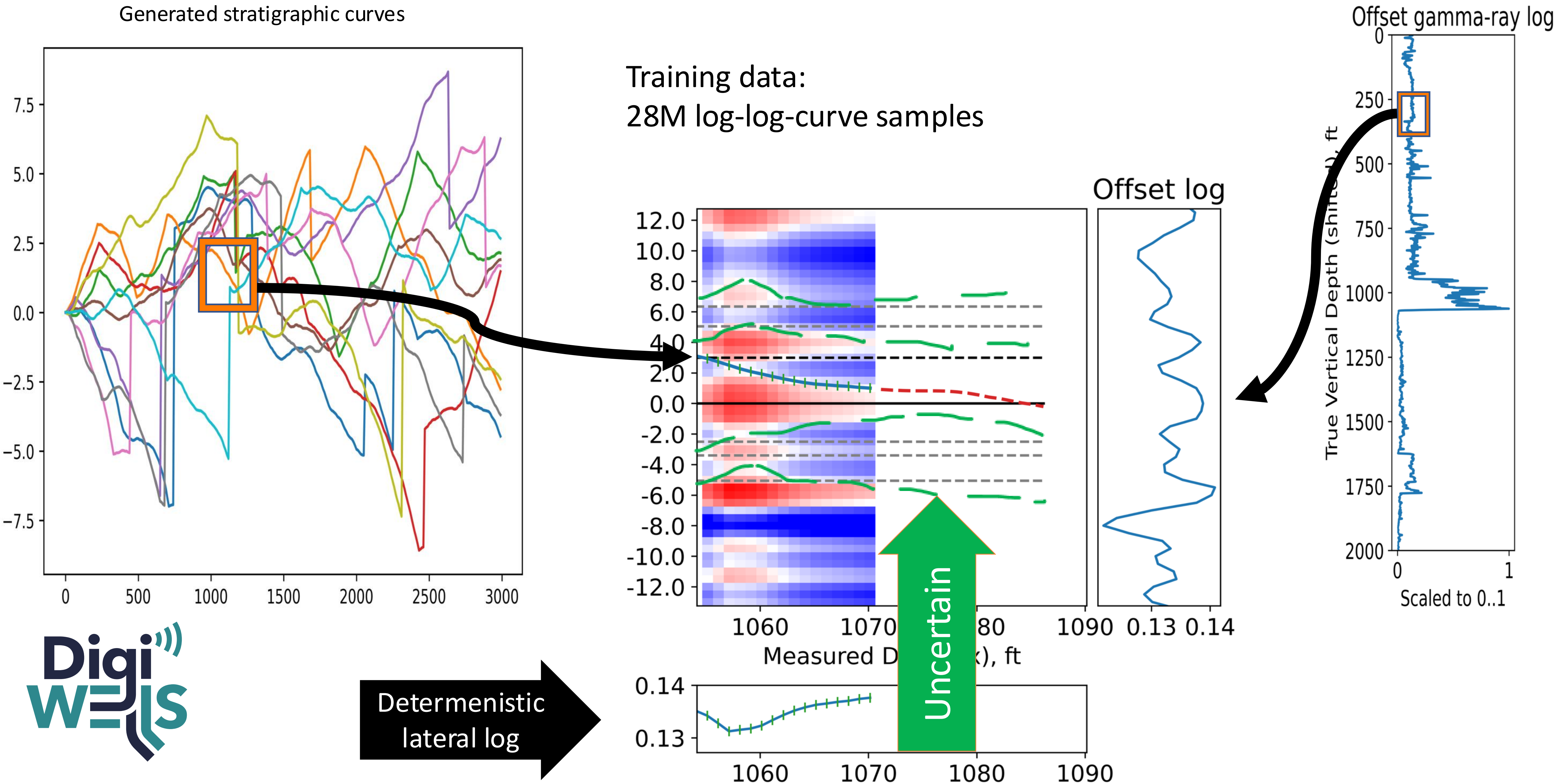


Determenistic lateral log



DigiWells

How did the model learn? Training data



Training
Multiple-Trajectory-Prediction loss

$$loss = \alpha_{probability} I_{probability} + I_{prediction}$$

$$-log\left(\frac{exp(x_{m^*})}{\sum_m exp(x_m)}\right) \qquad \frac{\left\|b^* - b^{m^*}\right\|}{l^+}$$

Probability misfit and Prediction quality of the "best mode"

minimized simultaneously

Best mode: $$m^* = \underset{1 \leq m \leq m_+}{\arg\min} \left\|b^* - b^m\right\|$$

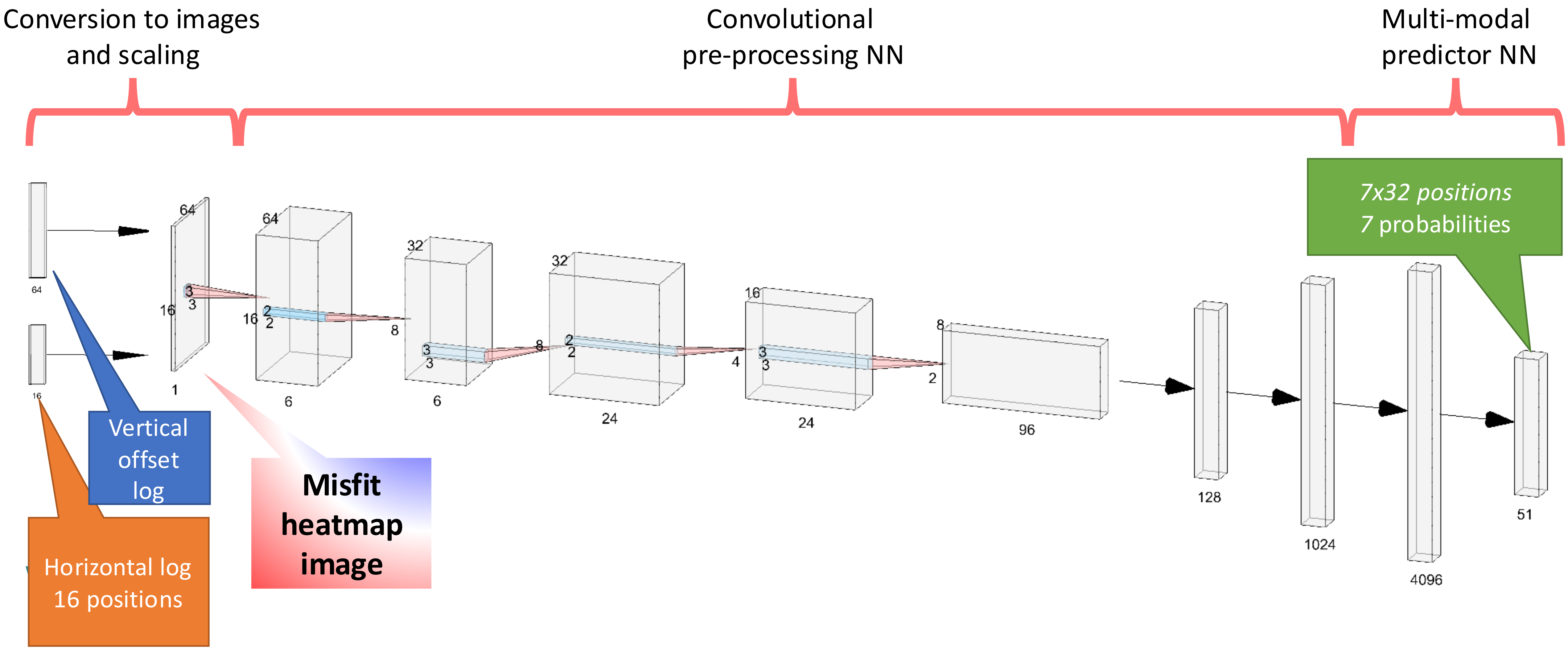
MTP loss is suggested by Cui, Henggang, et al. "Multimodal trajectory predictions..." 2019 IEER ICRA

Training
Multiple-Trajectory-Prediction loss

$$loss = \alpha_{probability} I_{probability} + I_{prediction}$$

Note: No explicit fitting to measurements

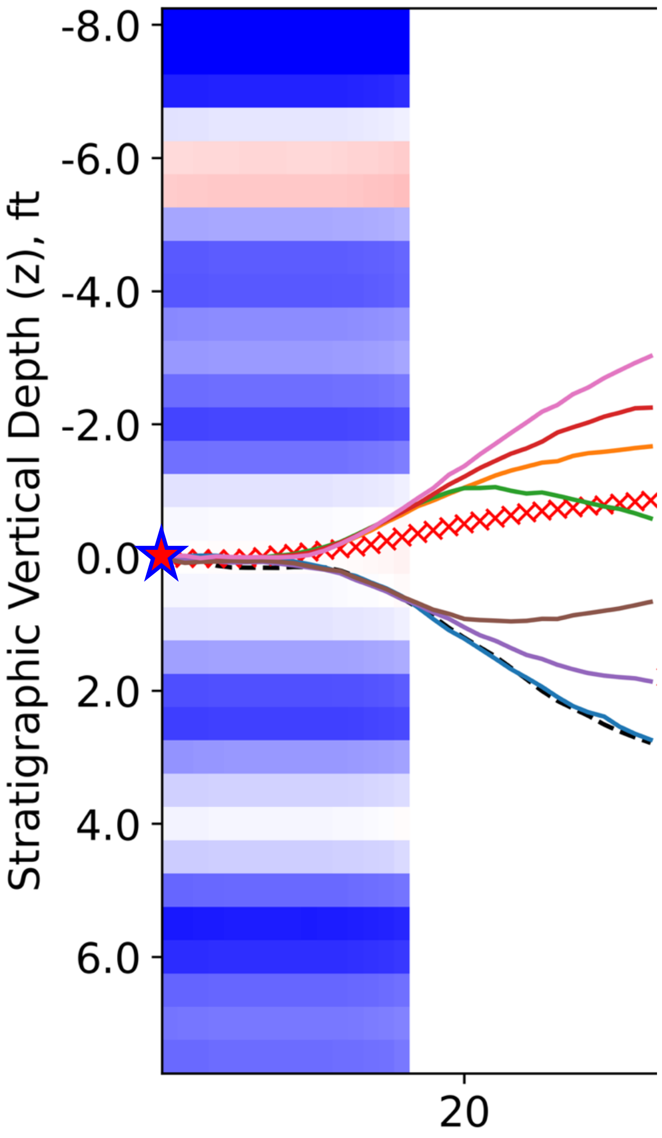
CNN Mixture Density Network **(MDN)** Architecture



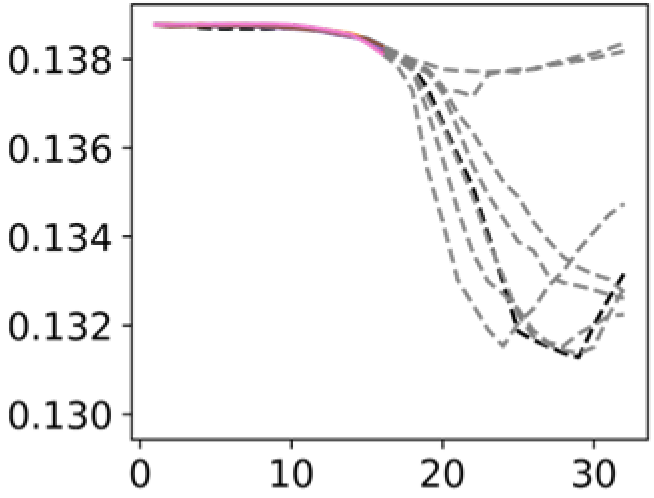
Mixture Density Network (MDN)

- A neural network that tries to reconstruct distribution from the data
- The distribution is represented as a sum of kernels
  - (similar to particles),
  - each having its own parameters

The results
Easy: prediction around 0

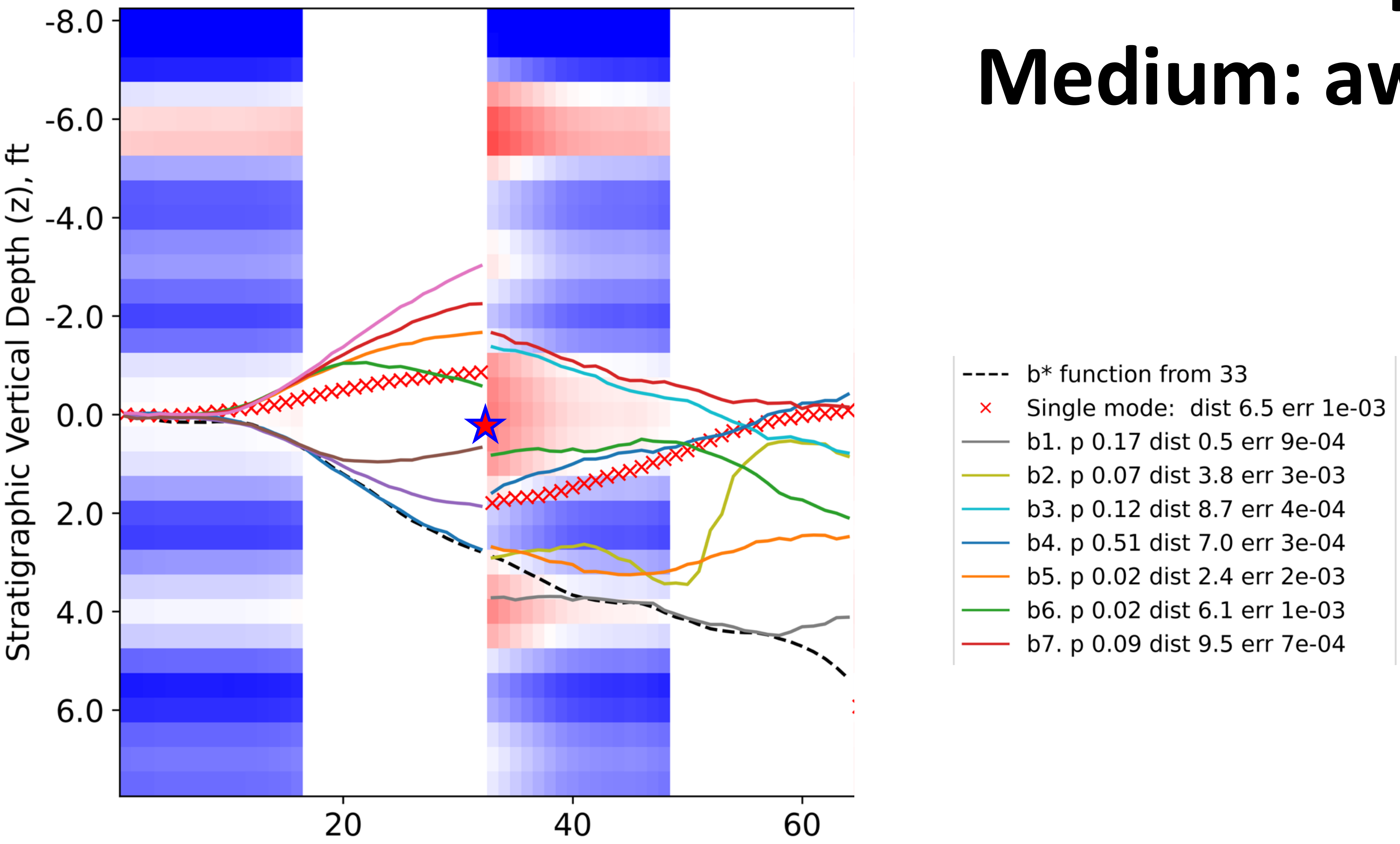


- - - b* function from 1
× Single mode: dist 2.8 err 9e-05
b1. p 0.08 dist 0.1 err 4e-05
b2. p 0.14 dist 3.5 err 5e-05
b3. p 0.12 dist 3.1 err 5e-05
b4. p 0.26 dist 3.9 err 5e-05
b5. p 0.10 dist 0.5 err 3e-05
b6. p 0.06 dist 1.1 err 3e-05
b7. p 0.24 dist 4.2 err 6e-05

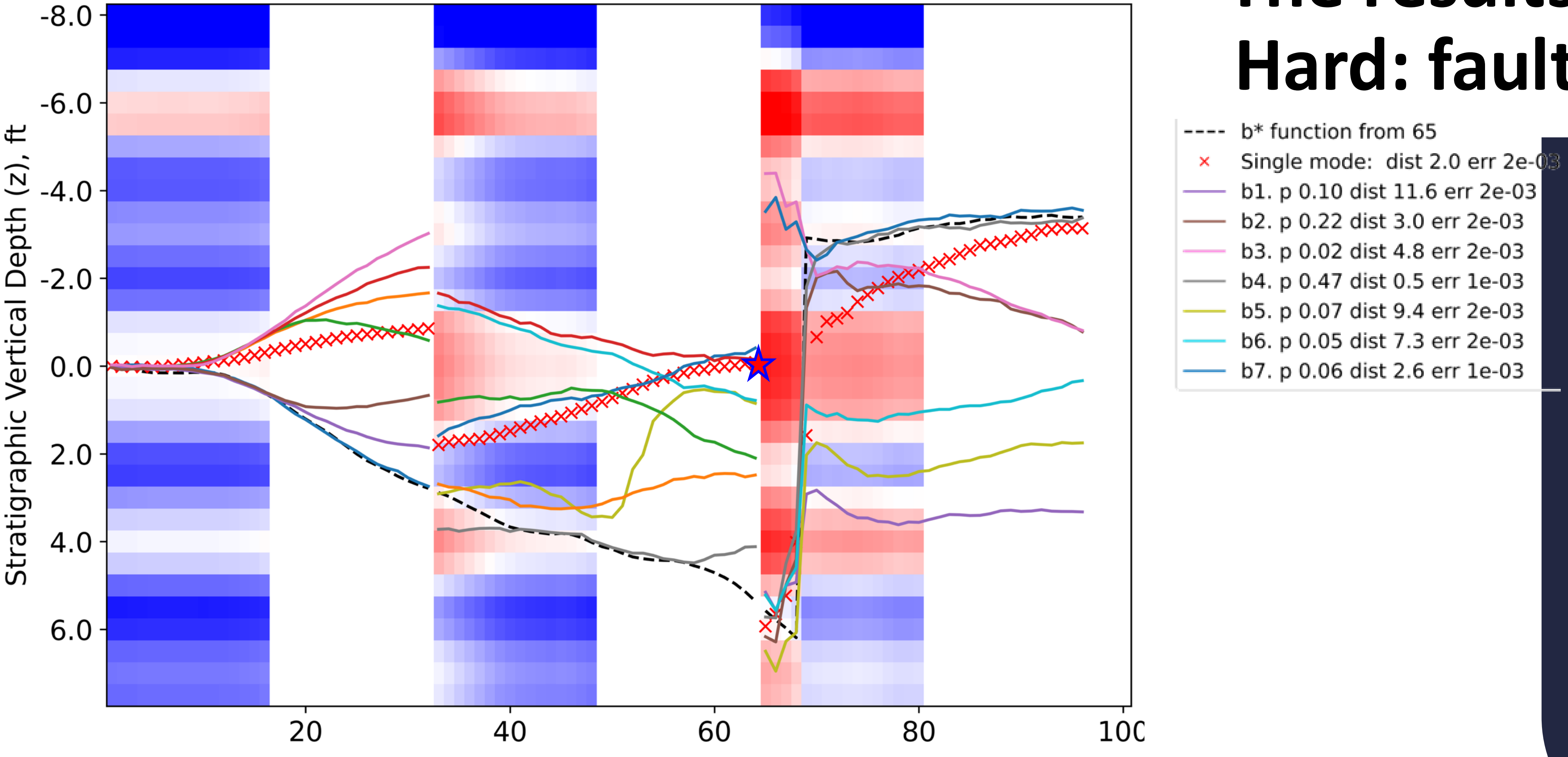


Gamma-ray logs

The results
Medium: away from 0



The results
Hard: fault



To invert a full well…

We need to apply MDN iteratively

- Needs to be robust outside training dataset
- Requires a heuristic algorithm for tracking previous inversions’ end points and their probabilities

Challenge:

Possible curse of dimensionality:

- 1 point ➔ 7 predictions ➔ 49 prediction ➔ 343 predictions
- Can loose computational efficiency

Sequential predictor (aka filter)

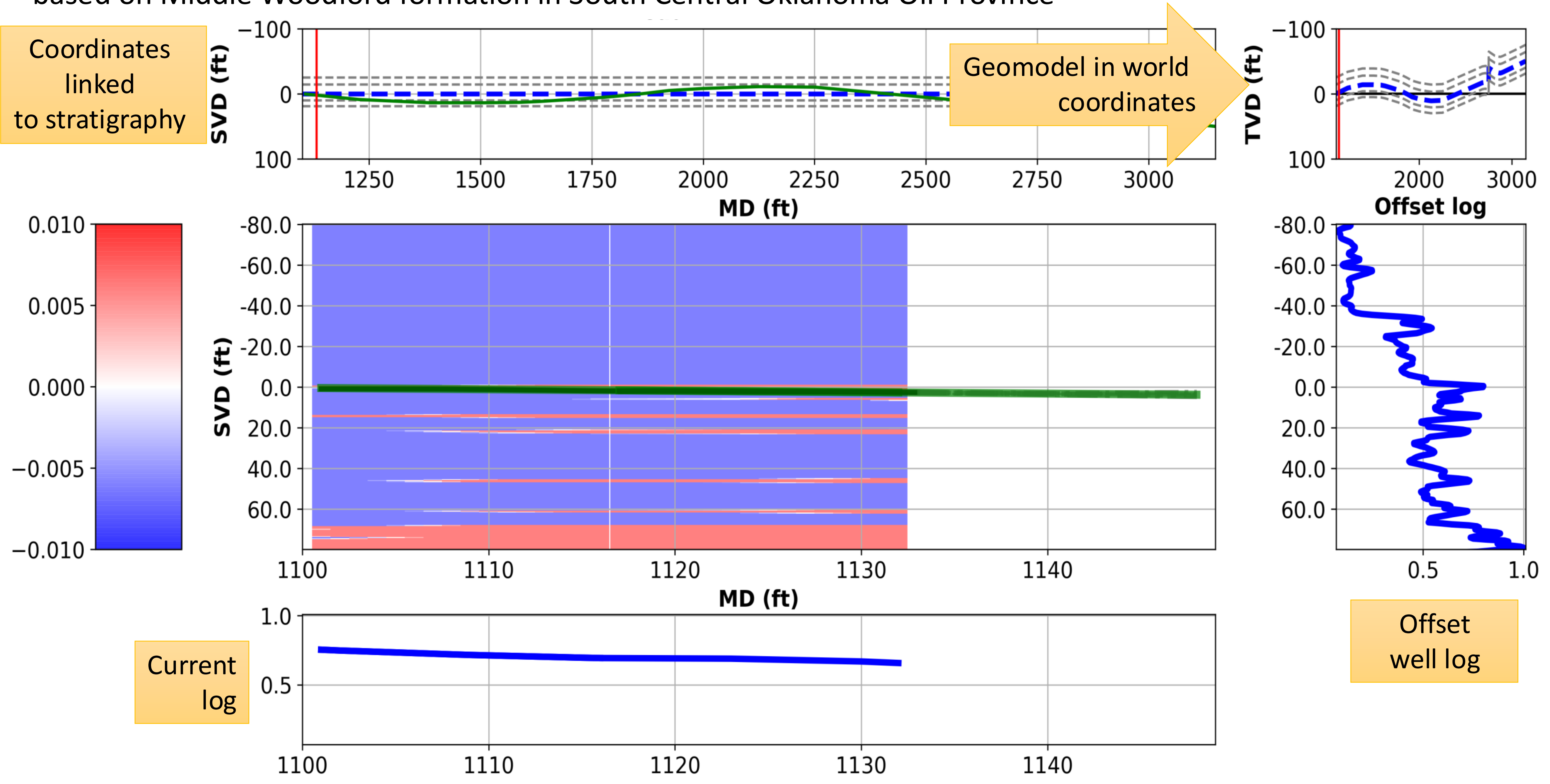
Inputs

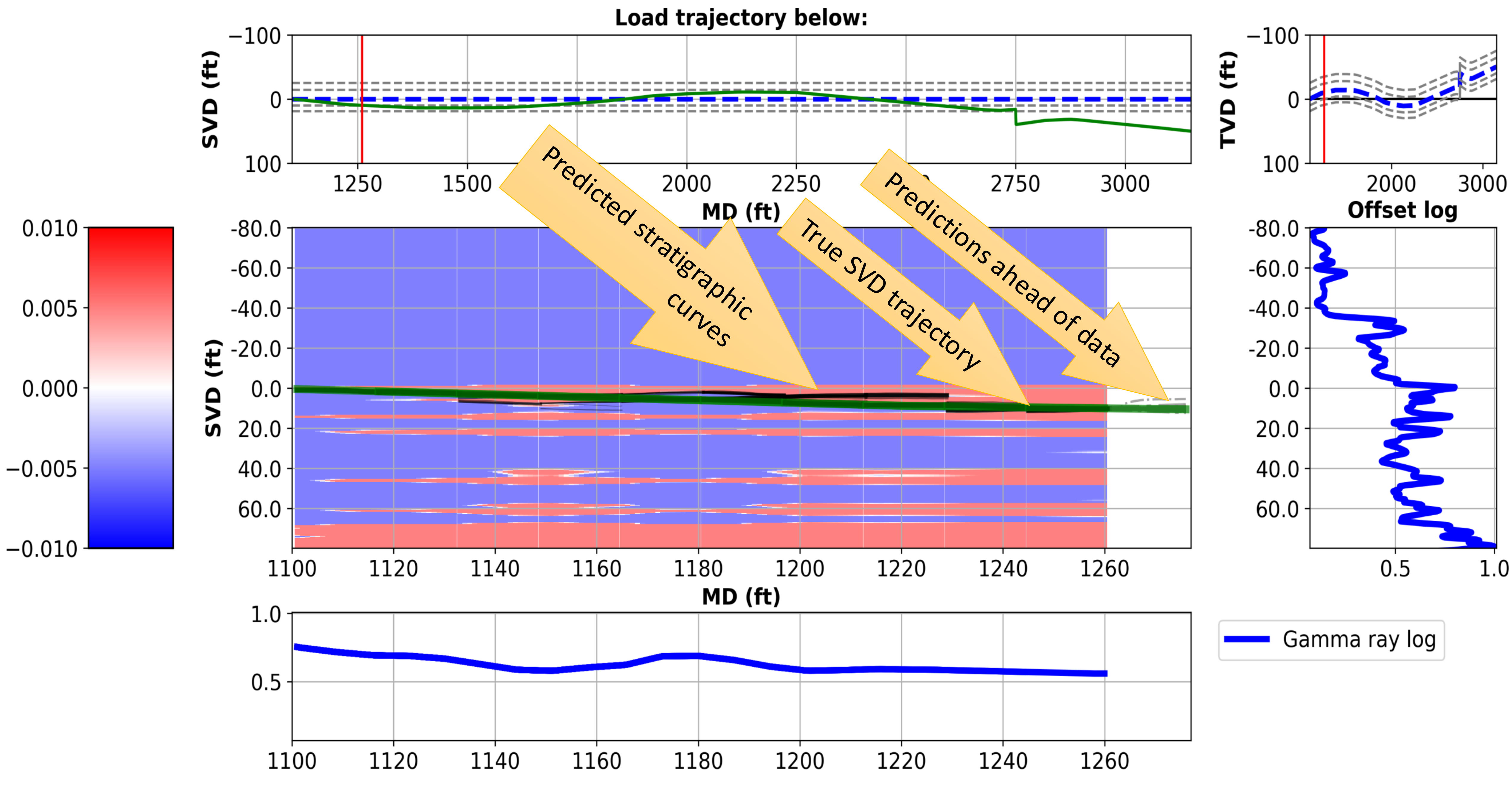
- Offset well log (any size)
- Horizontal well log
- Starting points and probabilities (any number *M*)

Outputs

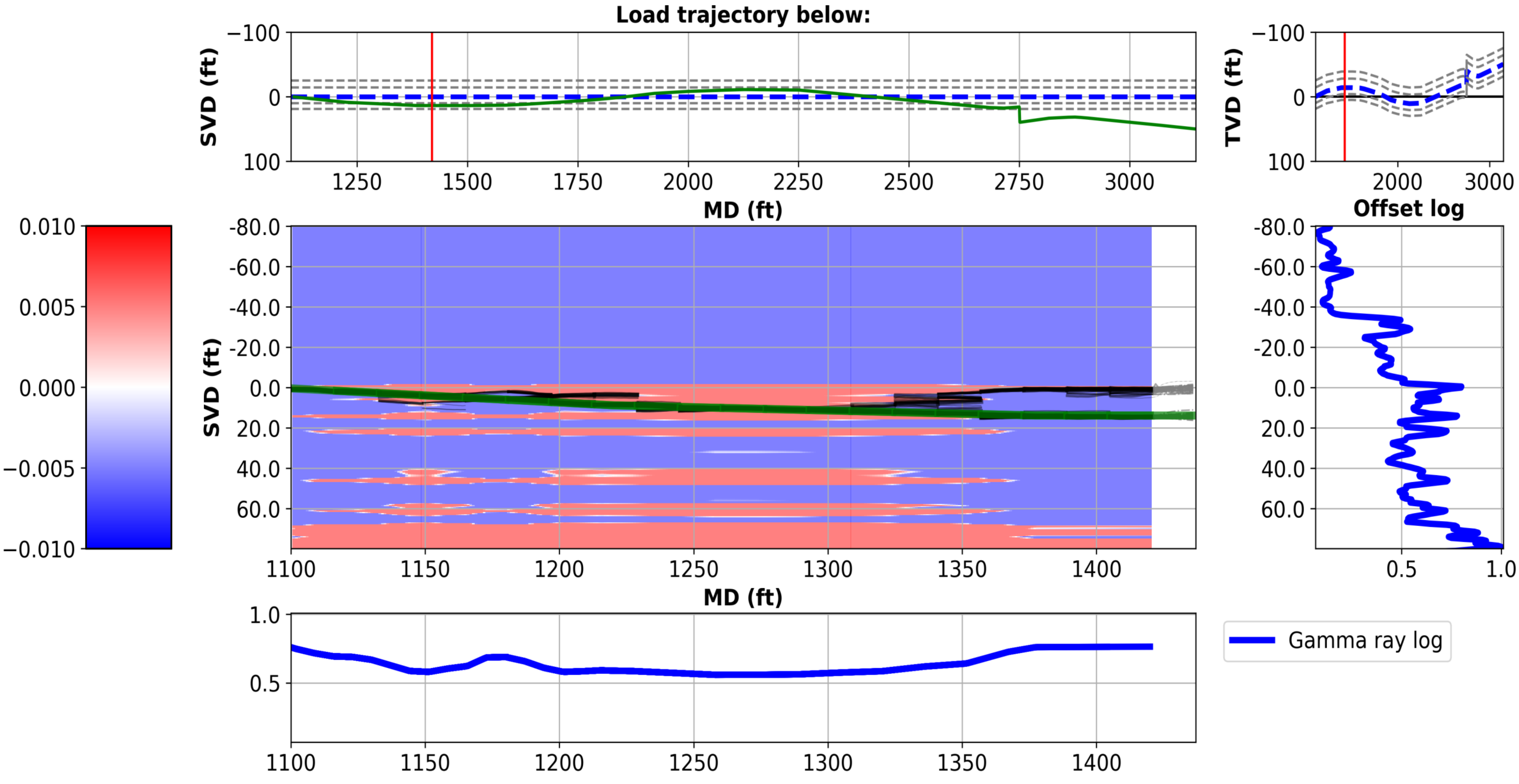
- *N*M* likely stratigraphic curves
  - Along well: INVERSION
  - Ahead of well: PREDICTION
- *N*M* probability values

Geological ‘truth’ and logs from GWC 2020
based on Middle Woodford formation in South Central Oklahoma Oil Province

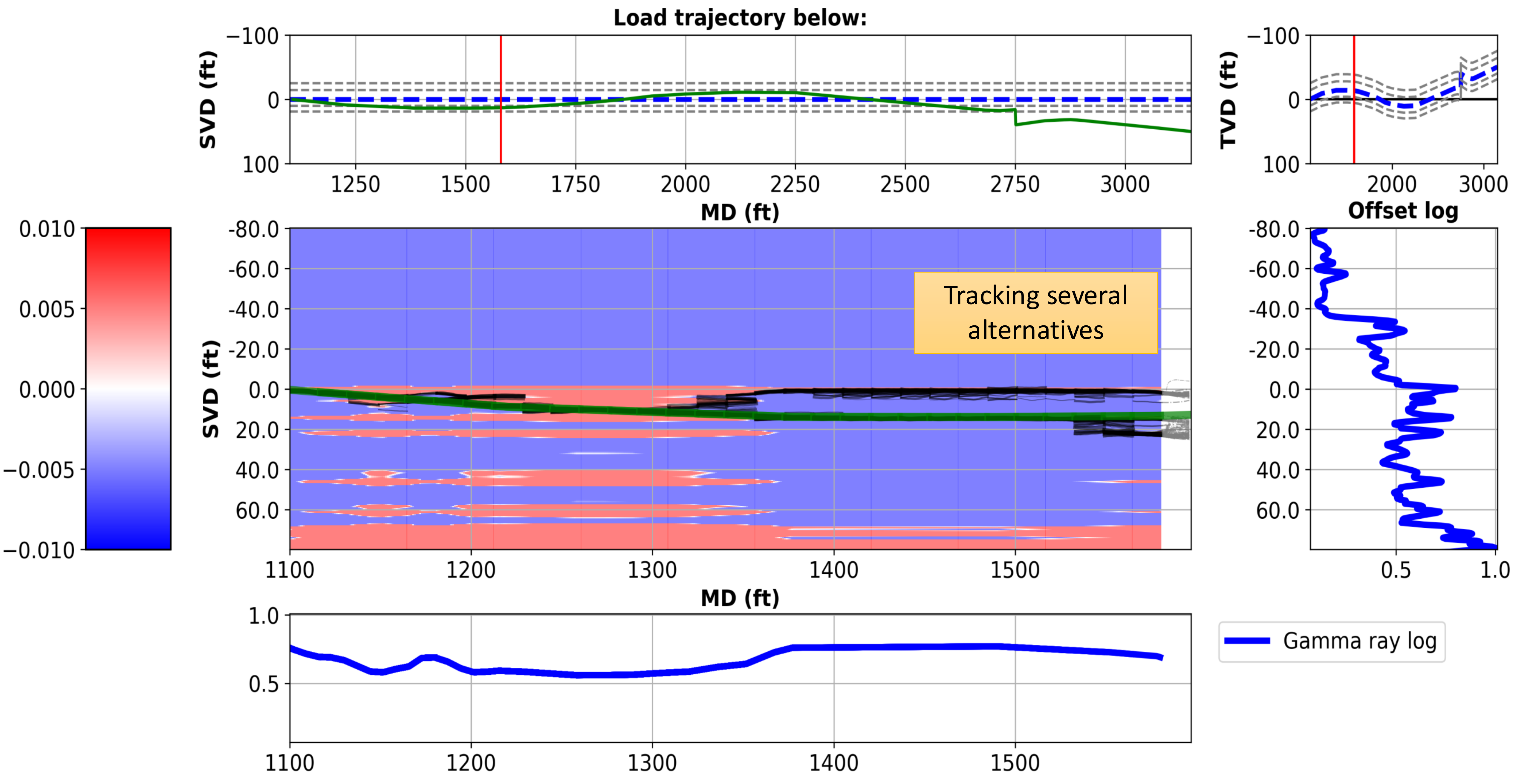




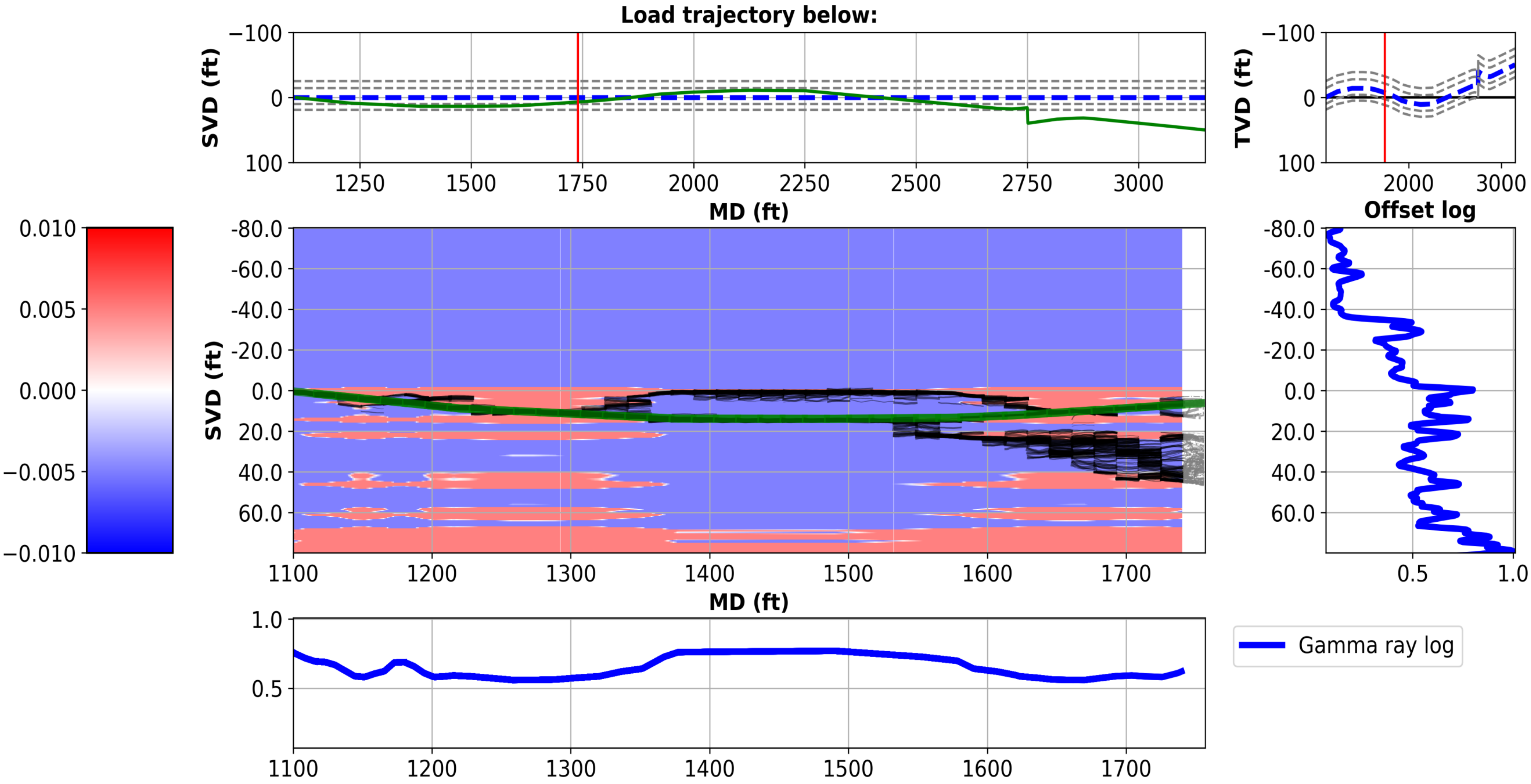
Load trajectory below:
SVD (ft)
MD (ft)
TVD (ft)
Offset log
Predicted stratigraphic curves
True SVD trajectory
Predictions ahead of data
Gamma ray log



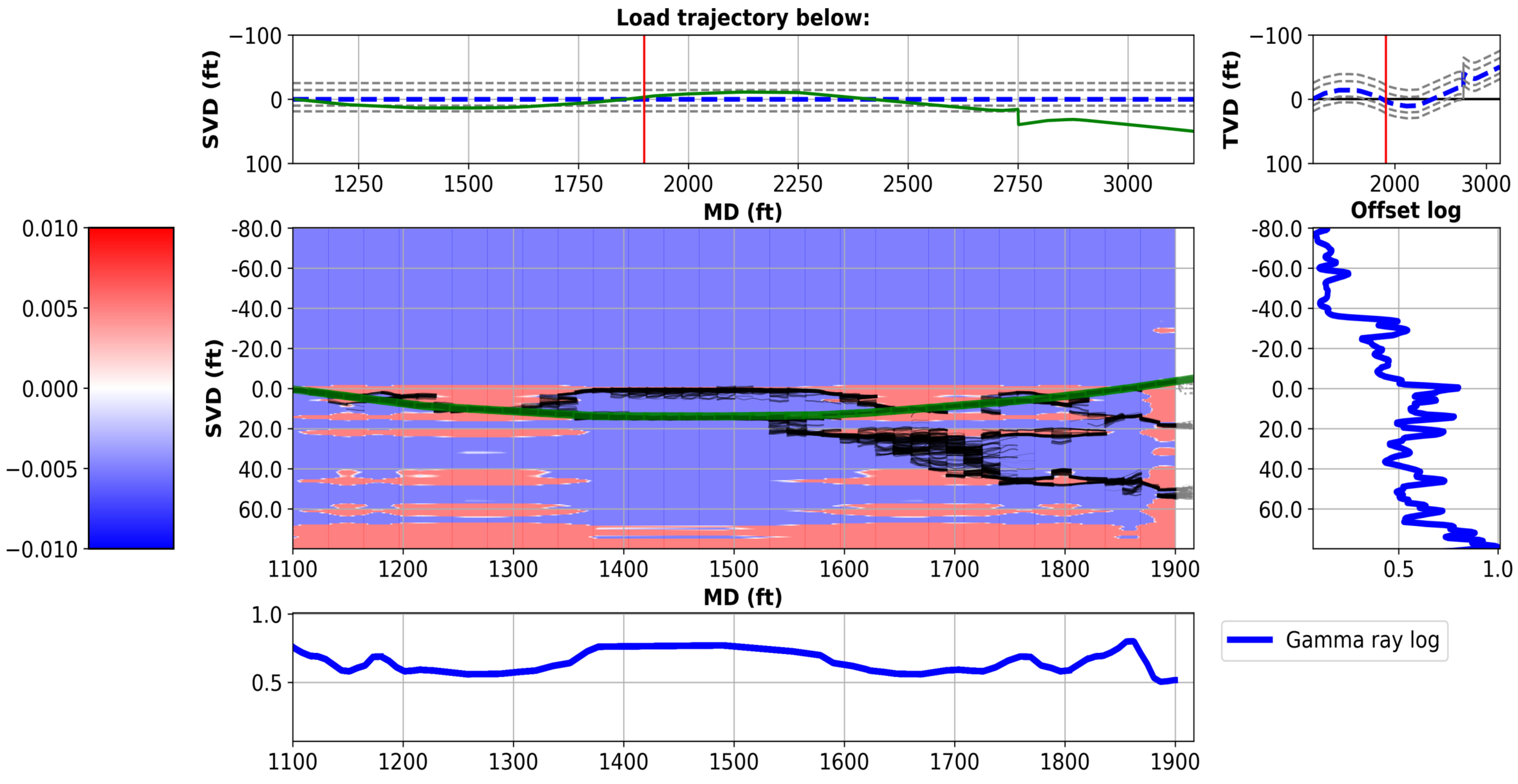
Load trajectory below:
SVD (ft)
MD (ft)
TVD (ft)
Offset log
Gamma ray log
0.010
0.005
0.000
−0.005
−0.010



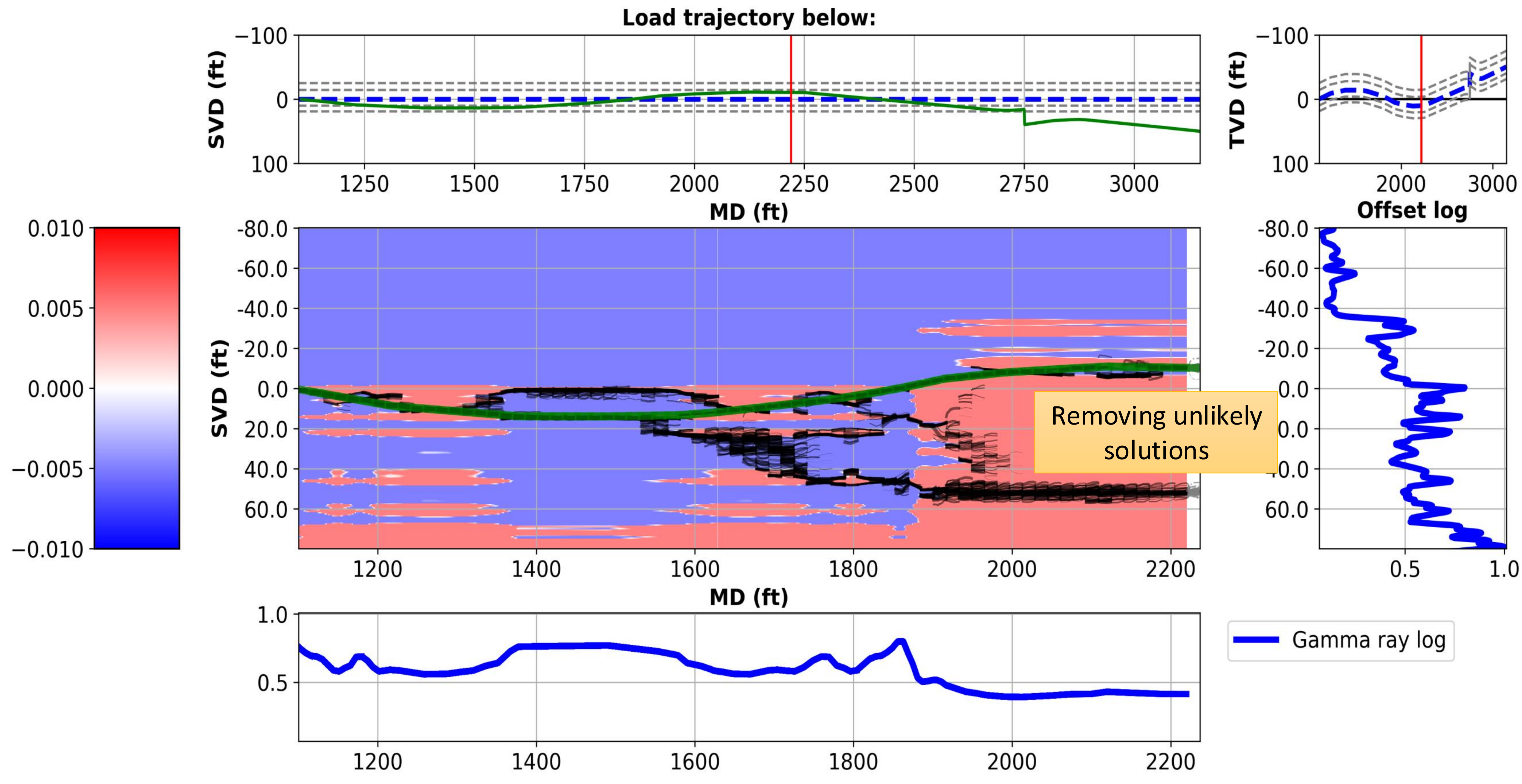
Load trajectory below:
SVD (ft)
MD (ft)
TVD (ft)
Offset log
Tracking several alternatives
Gamma ray log



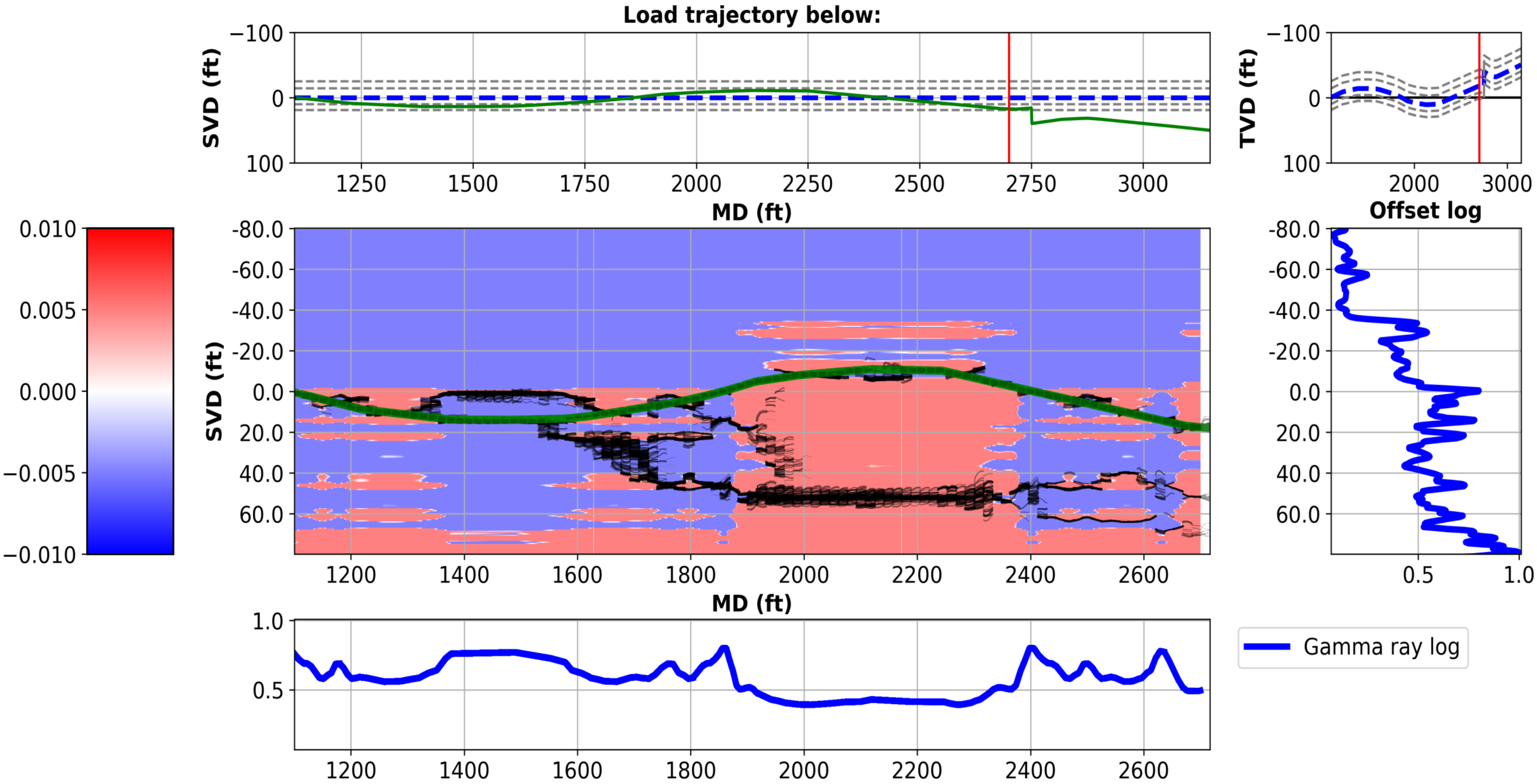
Load trajectory below:
SVD (ft)
MD (ft)
TVD (ft)
Offset log
Gamma ray log



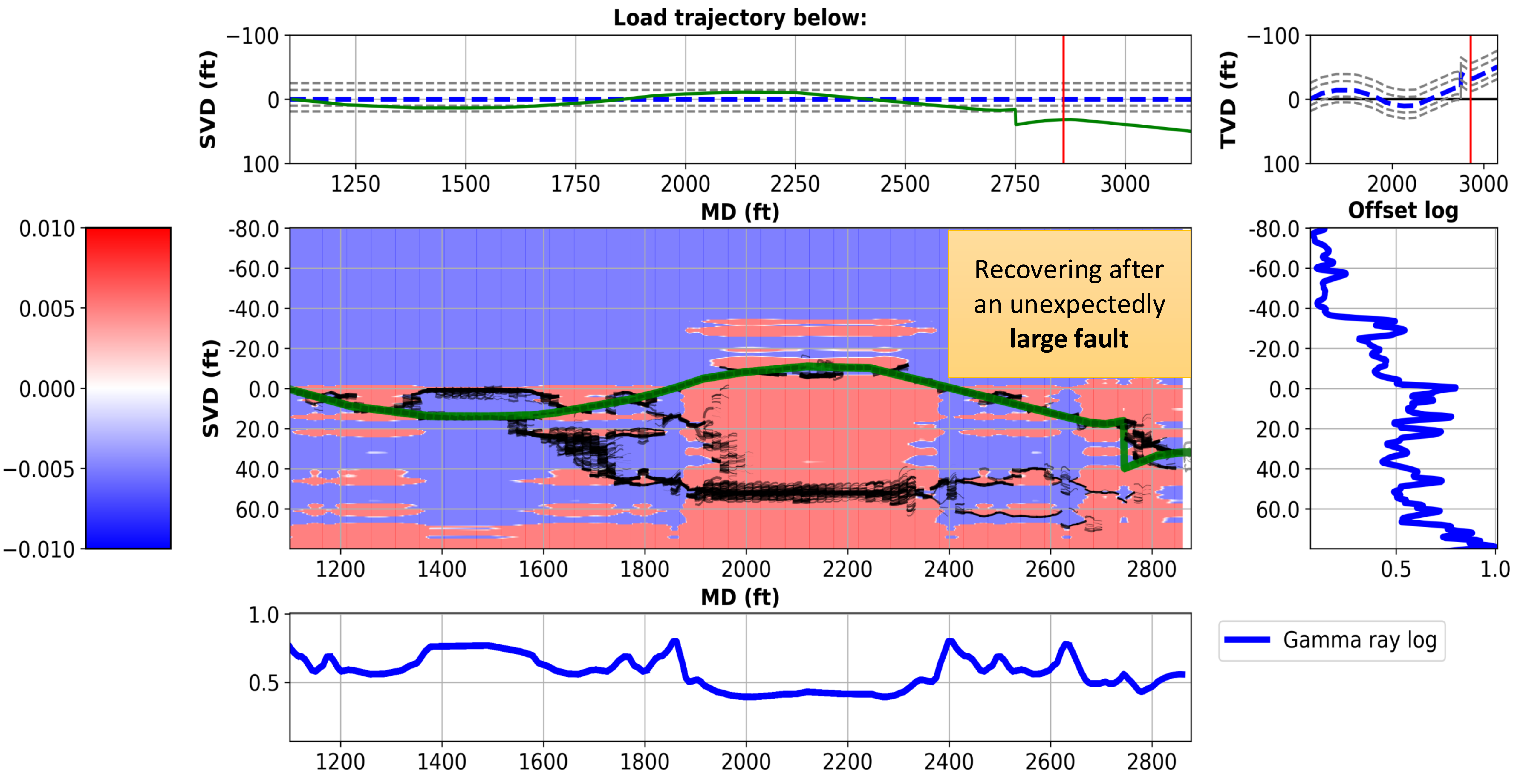
Load trajectory below:
SVD (ft)
MD (ft)
TVD (ft)
Offset log
Gamma ray log



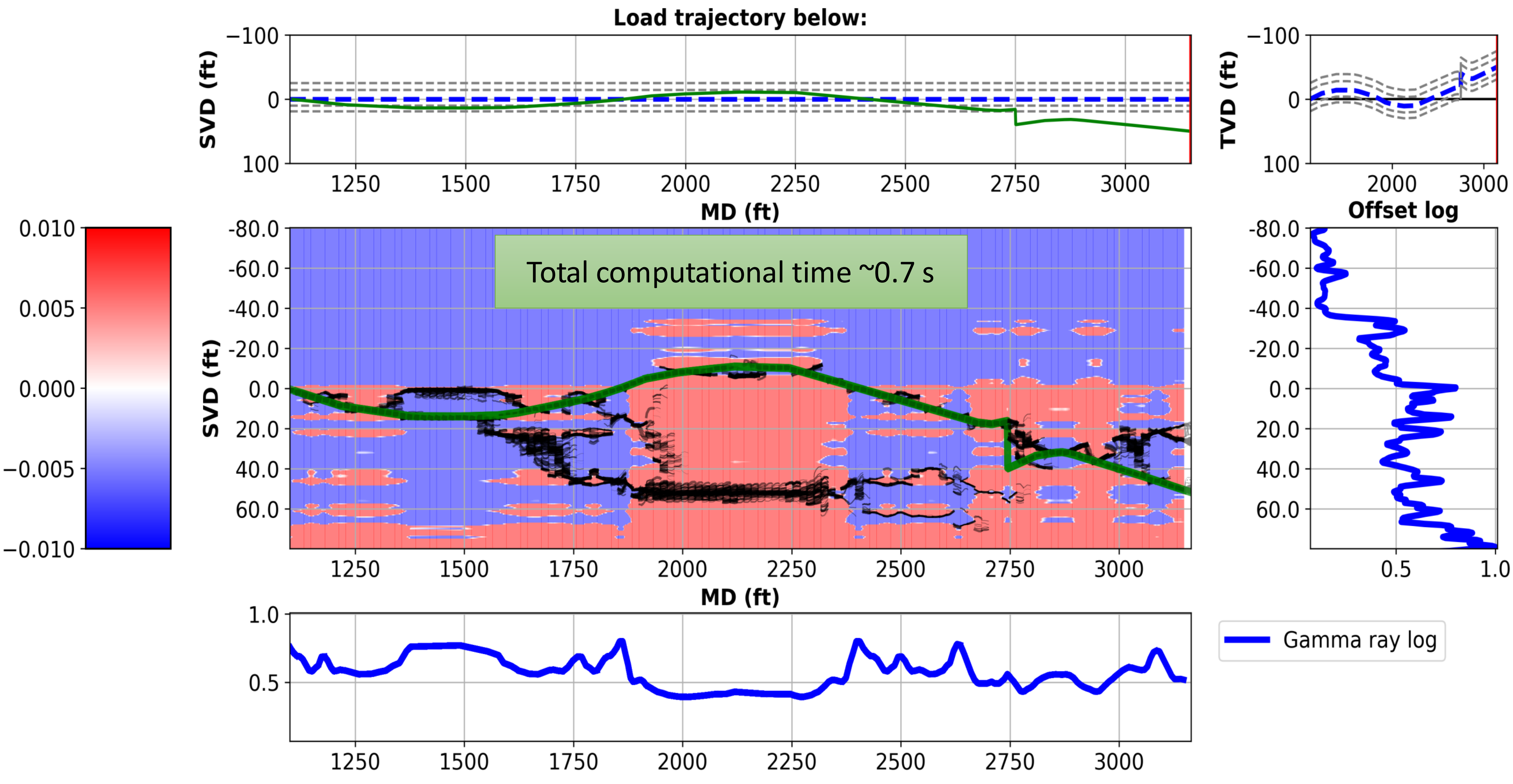
Load trajectory below:
SVD (ft)
MD (ft)
TVD (ft)
Offset log
Removing unlikely solutions
Gamma ray log



Load trajectory below:
SVD (ft)
MD (ft)
TVD (ft)
Offset log
Gamma ray log



Load trajectory below:
SVD (ft)
MD (ft)
TVD (ft)
Offset log
Recovering after an unexpectedly **large fault**
Gamma ray log

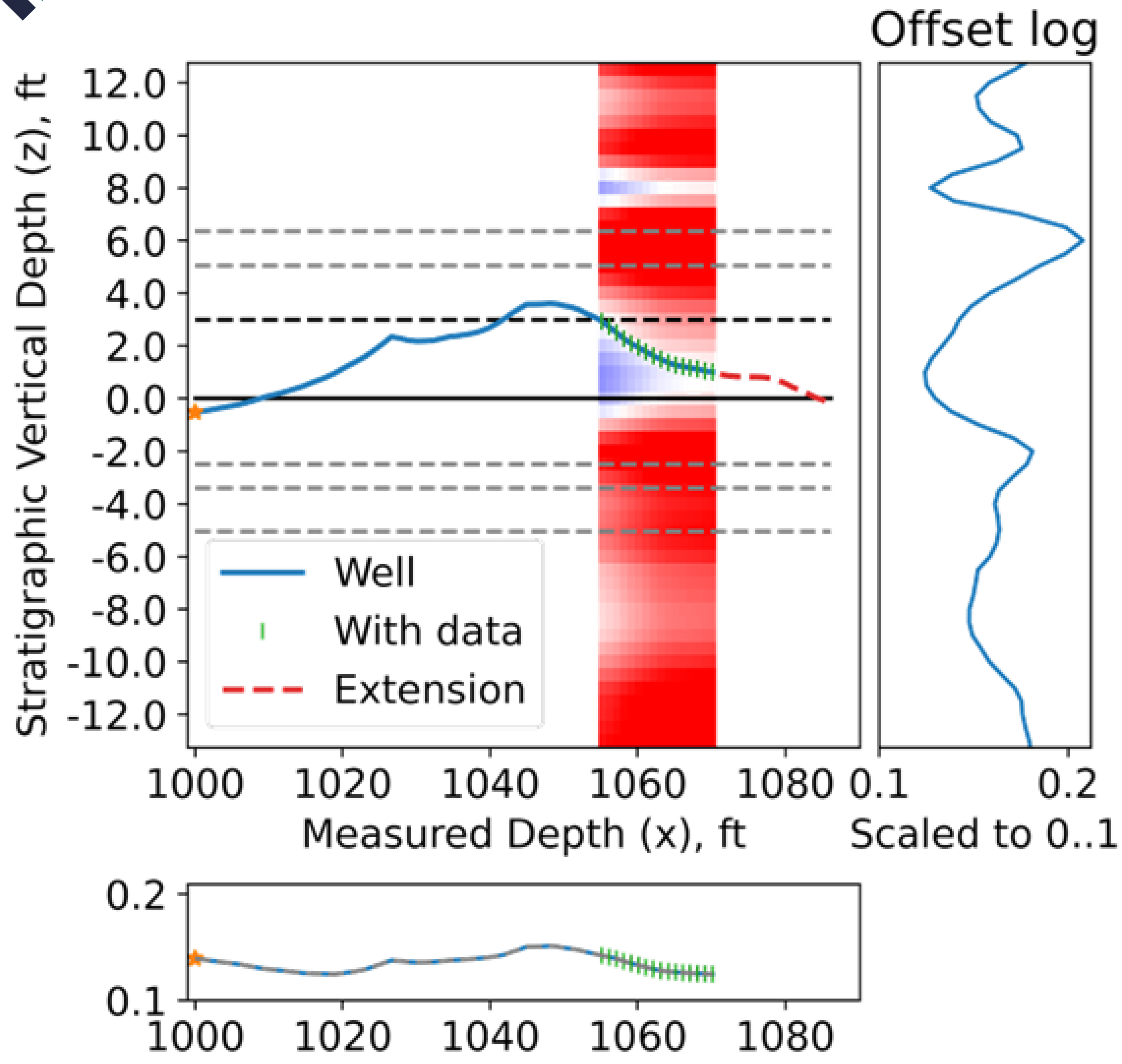

Load trajectory below:
SVD (ft)
MD (ft)
TVD (ft)
Offset log
Total computational time ~0.7 s
Gamma ray log

The MDN deep-learning method

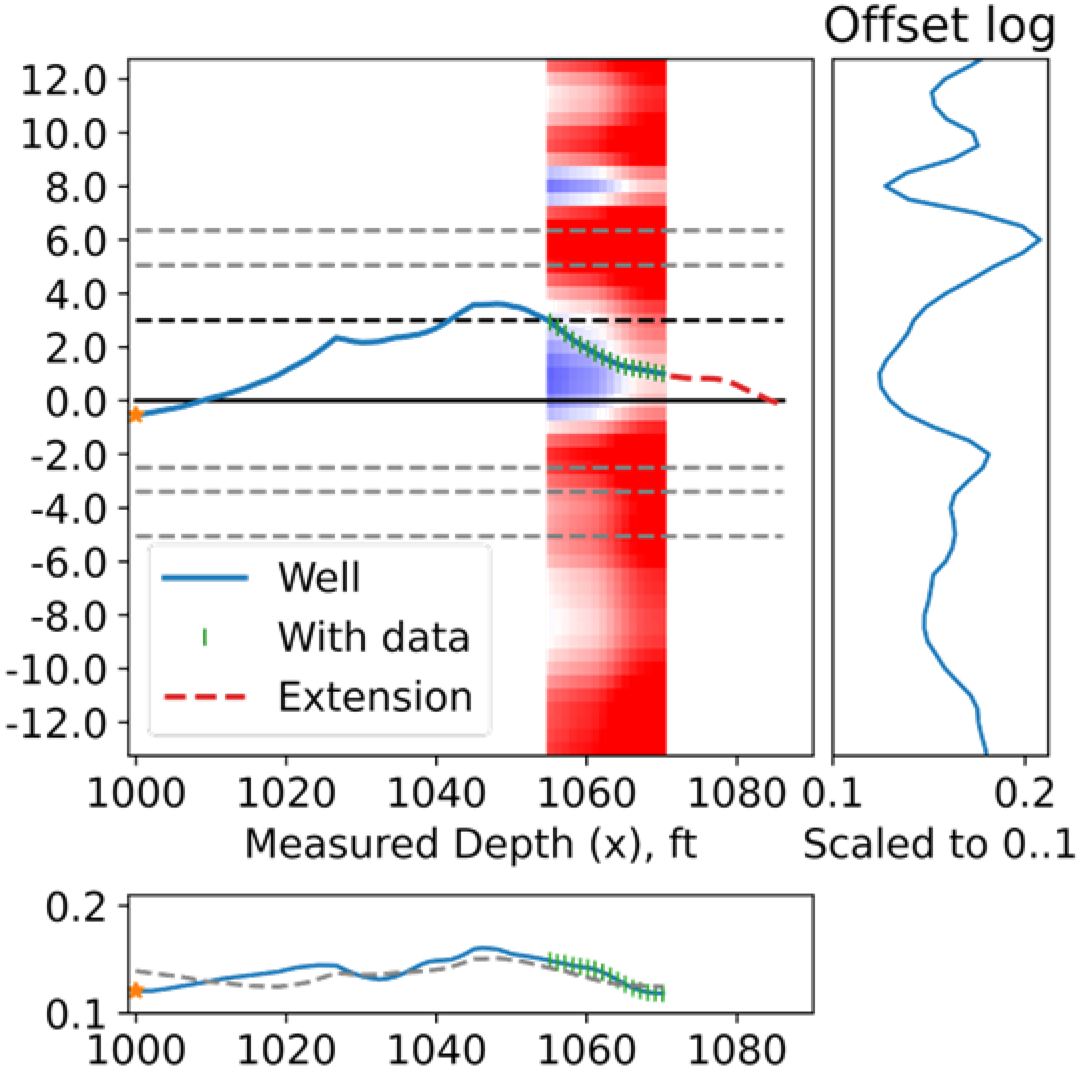
- Two orders of magnitude faster than other current methods
- Can predict and track several realizations in ‘low-data’ enviornment

- Assumption: noiseless data

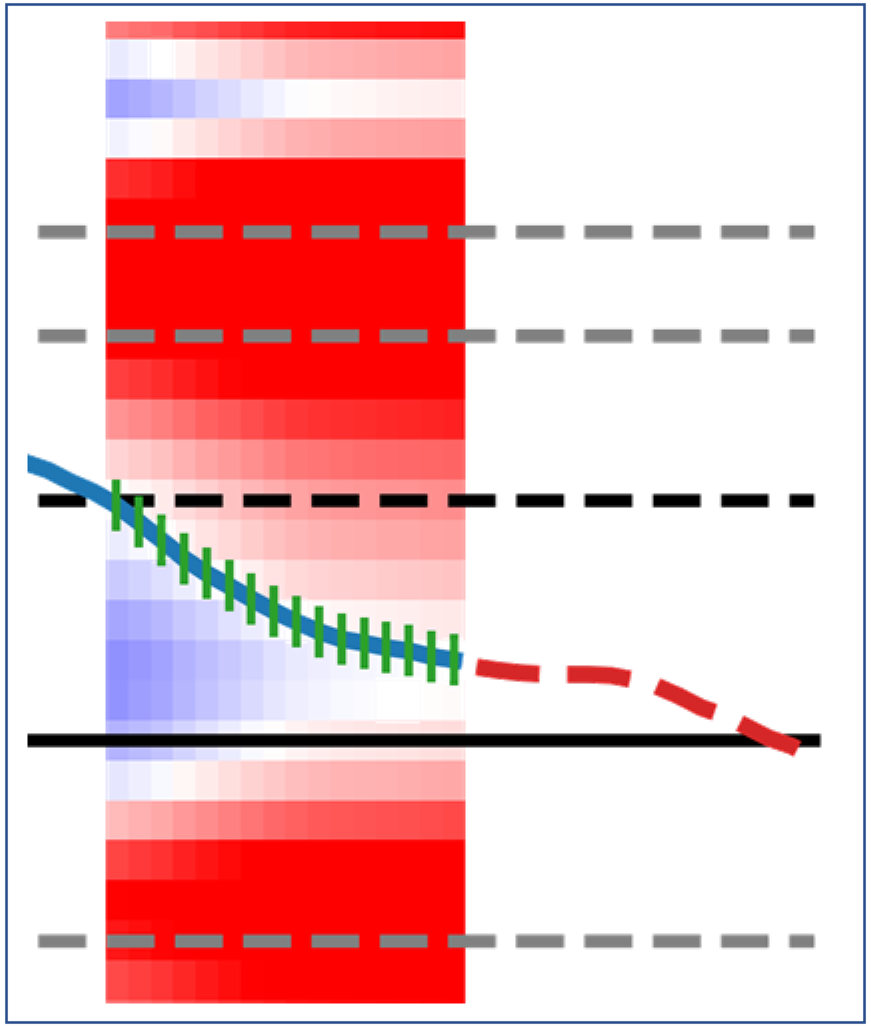
No noise



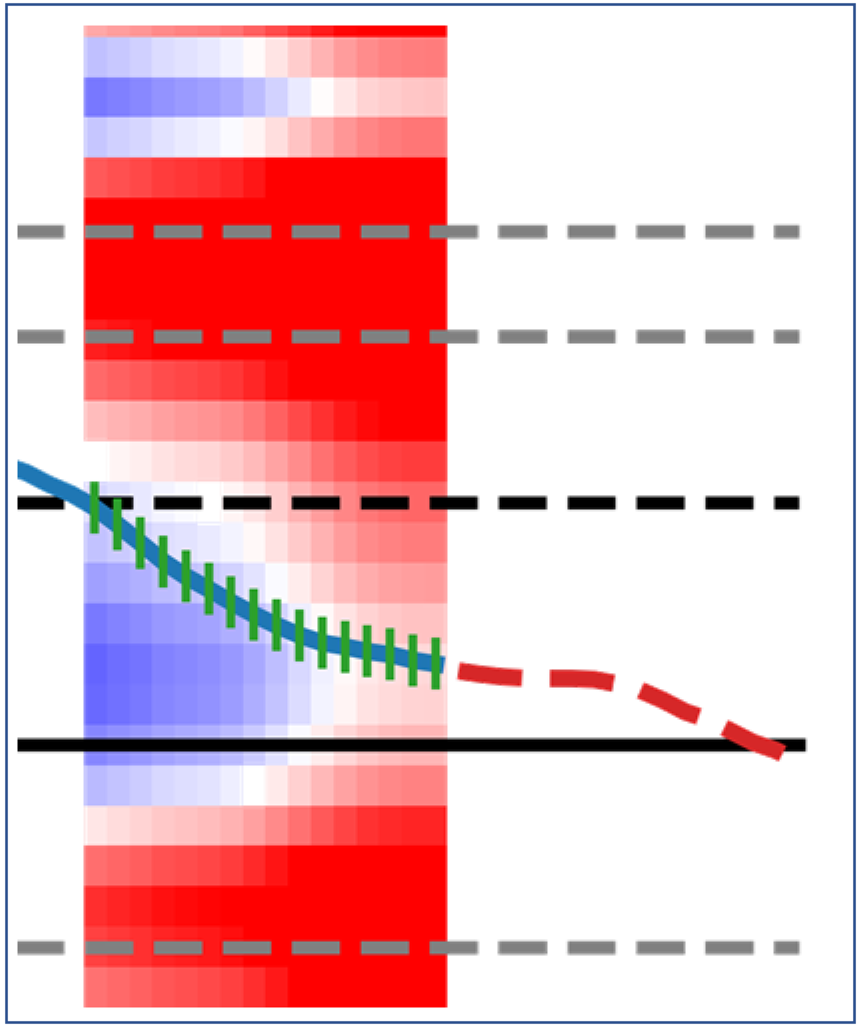
With noise



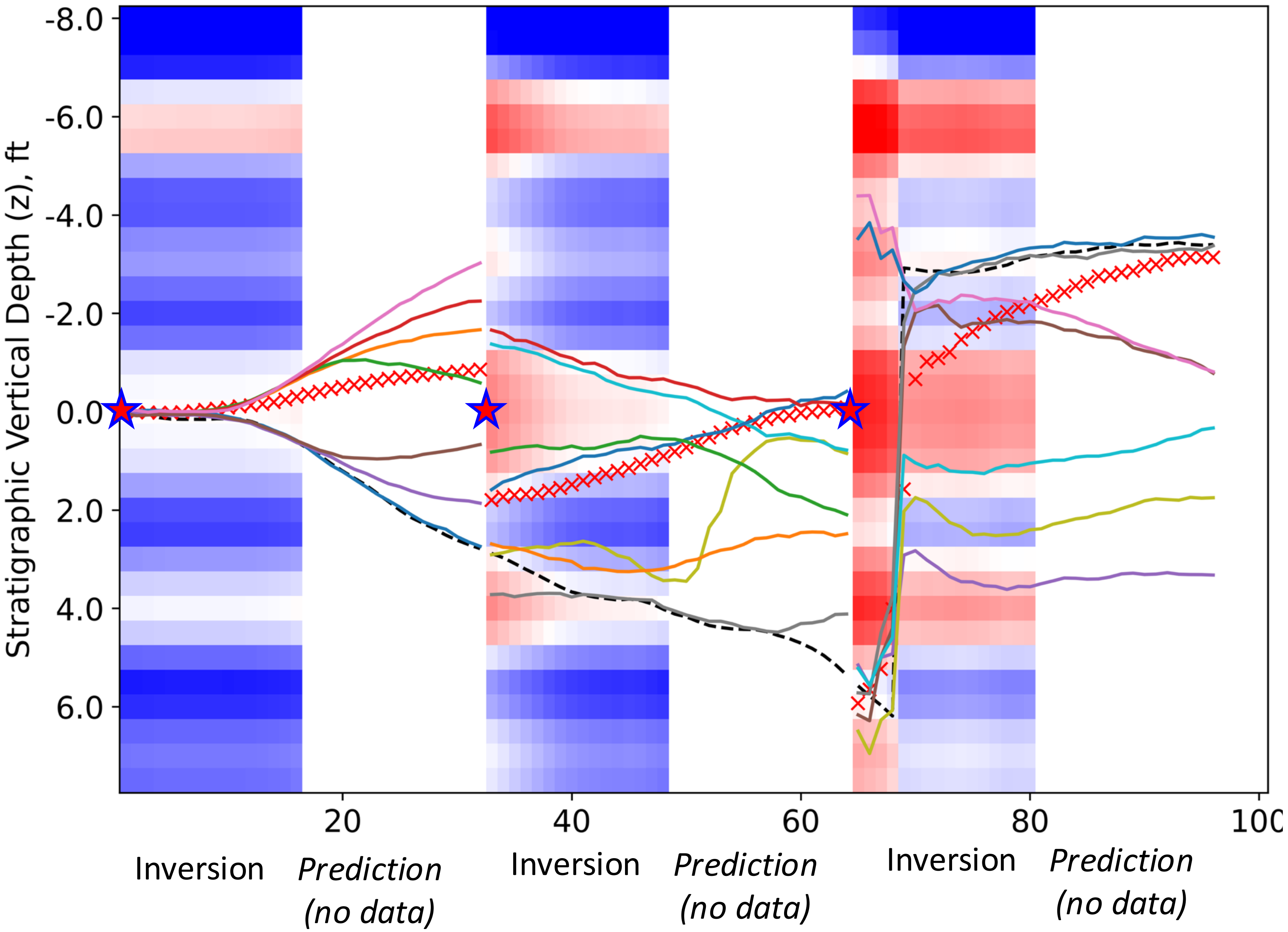
No noise



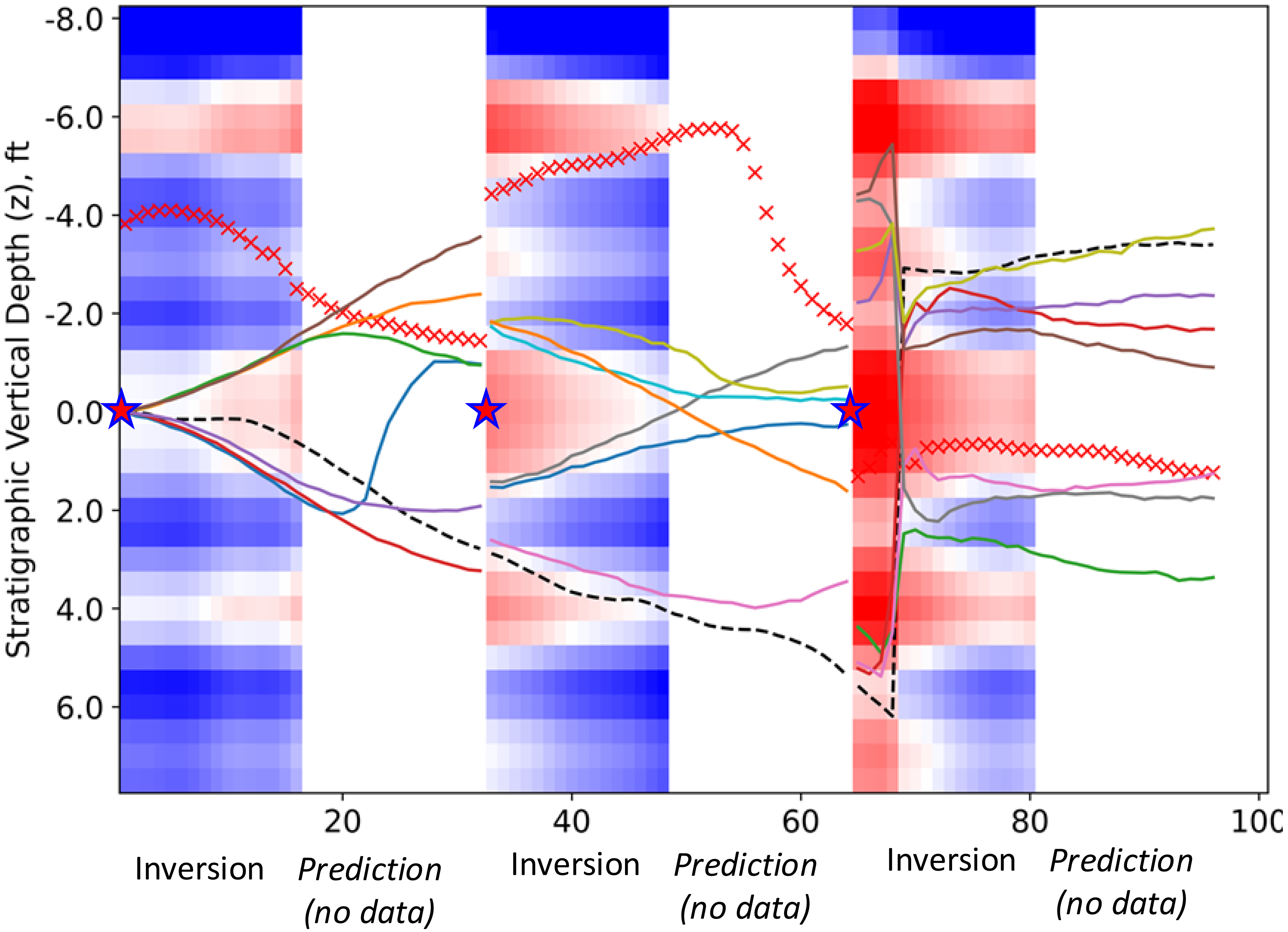
With noise



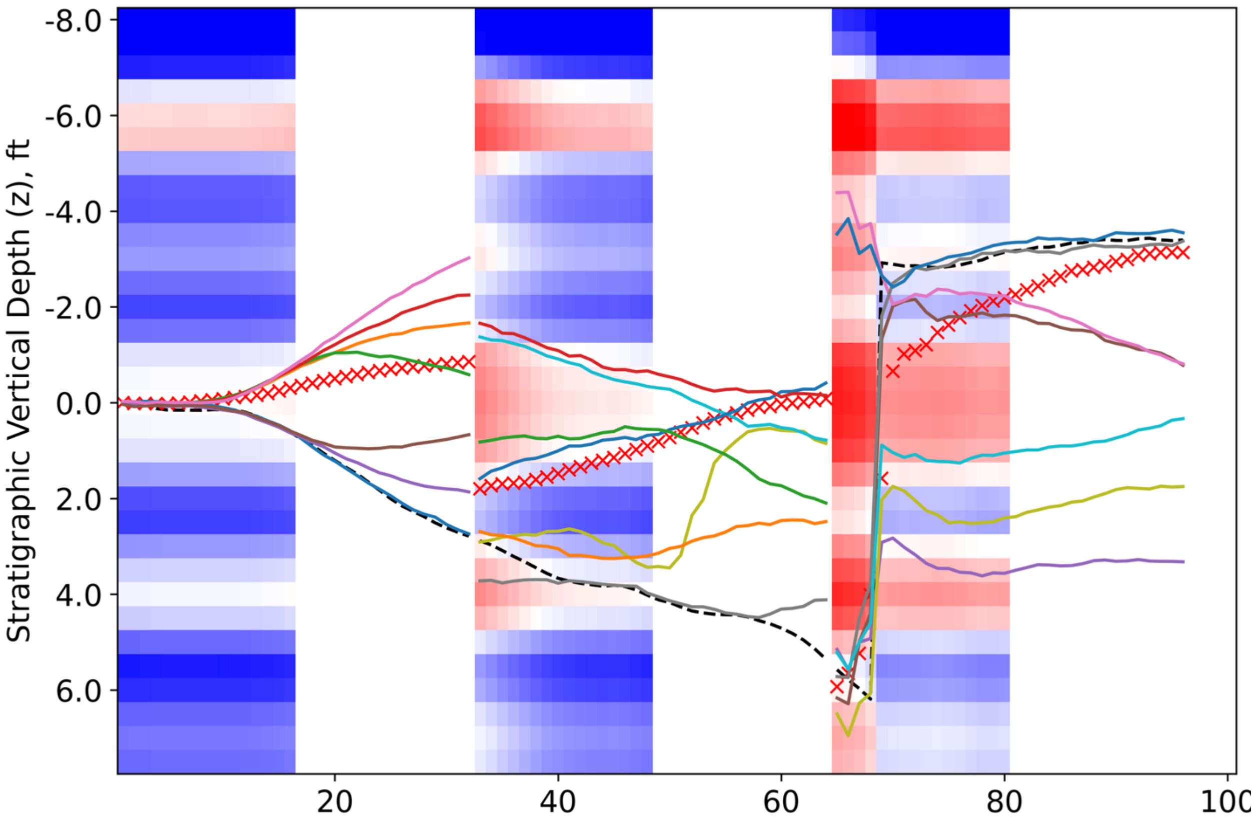
Summary
No noise

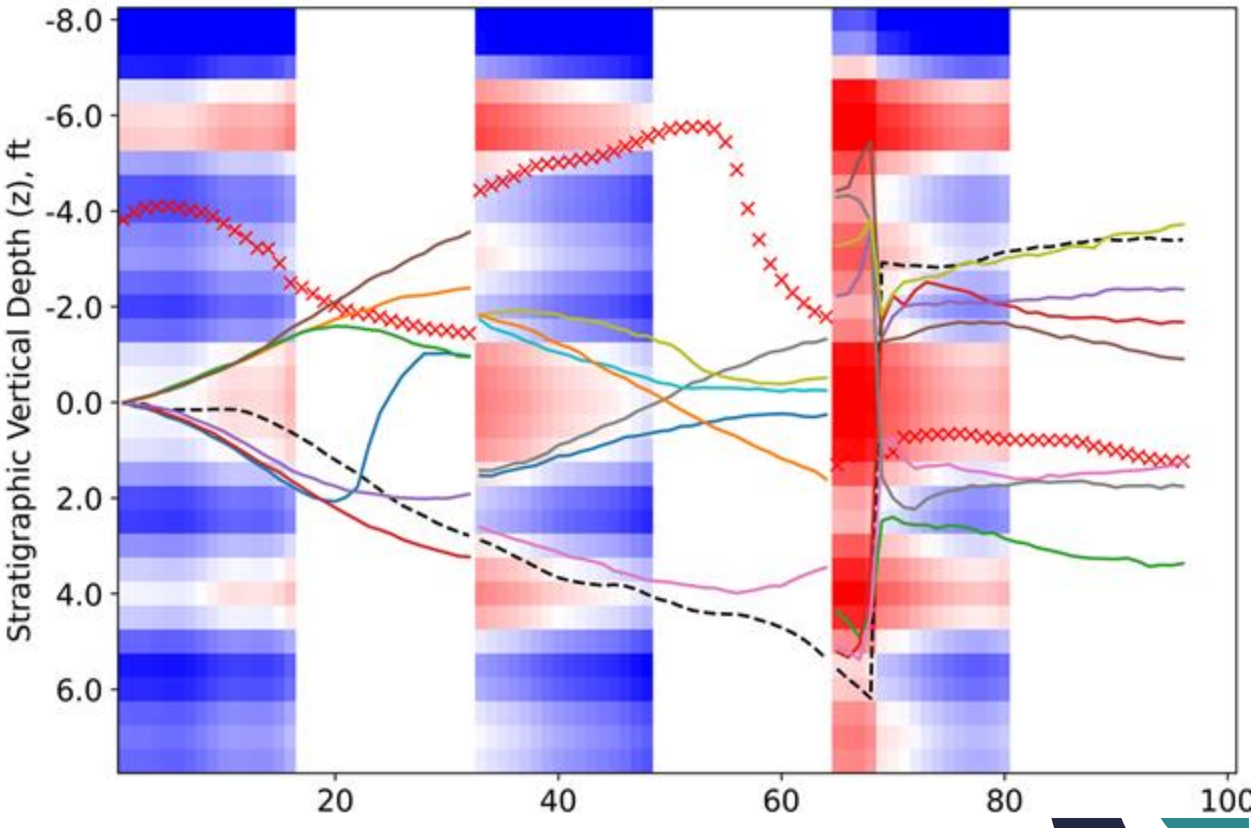


Summary 4% noise



Effect of 4% noise on prediction (right)





Measuring fit quality

$$I_{NLL} = -\log \sum_m \underbrace{p_m}_{\text{Probability}} \underbrace{\exp\left(\frac{-\|b^* - b^m\|}{\sigma^*}\right)}_{\text{Mismatch for one mode scaled with kernel 'width'}}$$

Negative-Log-Likelihood misfit goes to zero if

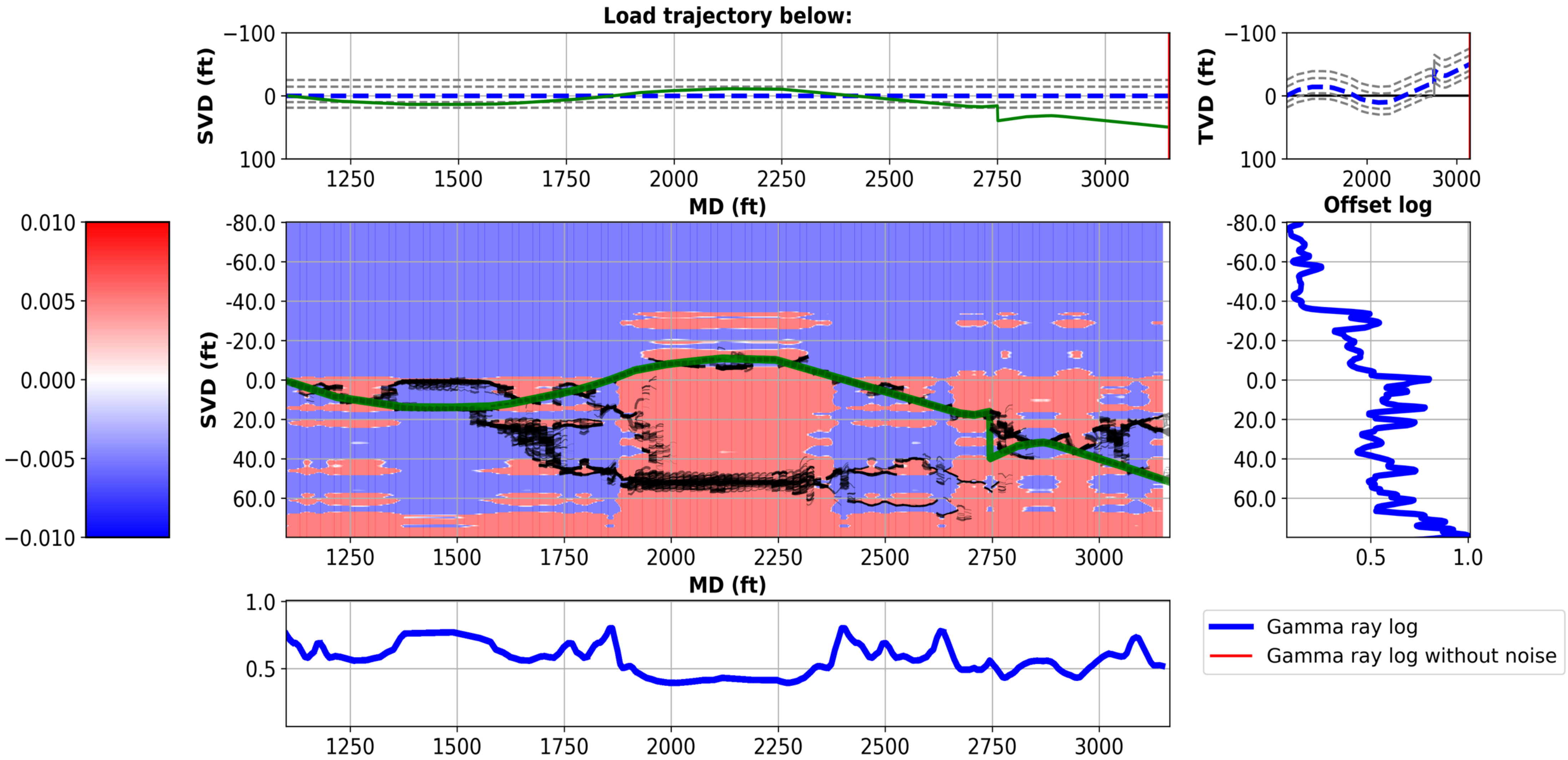
- the mode kernel coincides with the solution ***b*==b***^{m} and
- has ***p=1***

Results on noisy data revisited

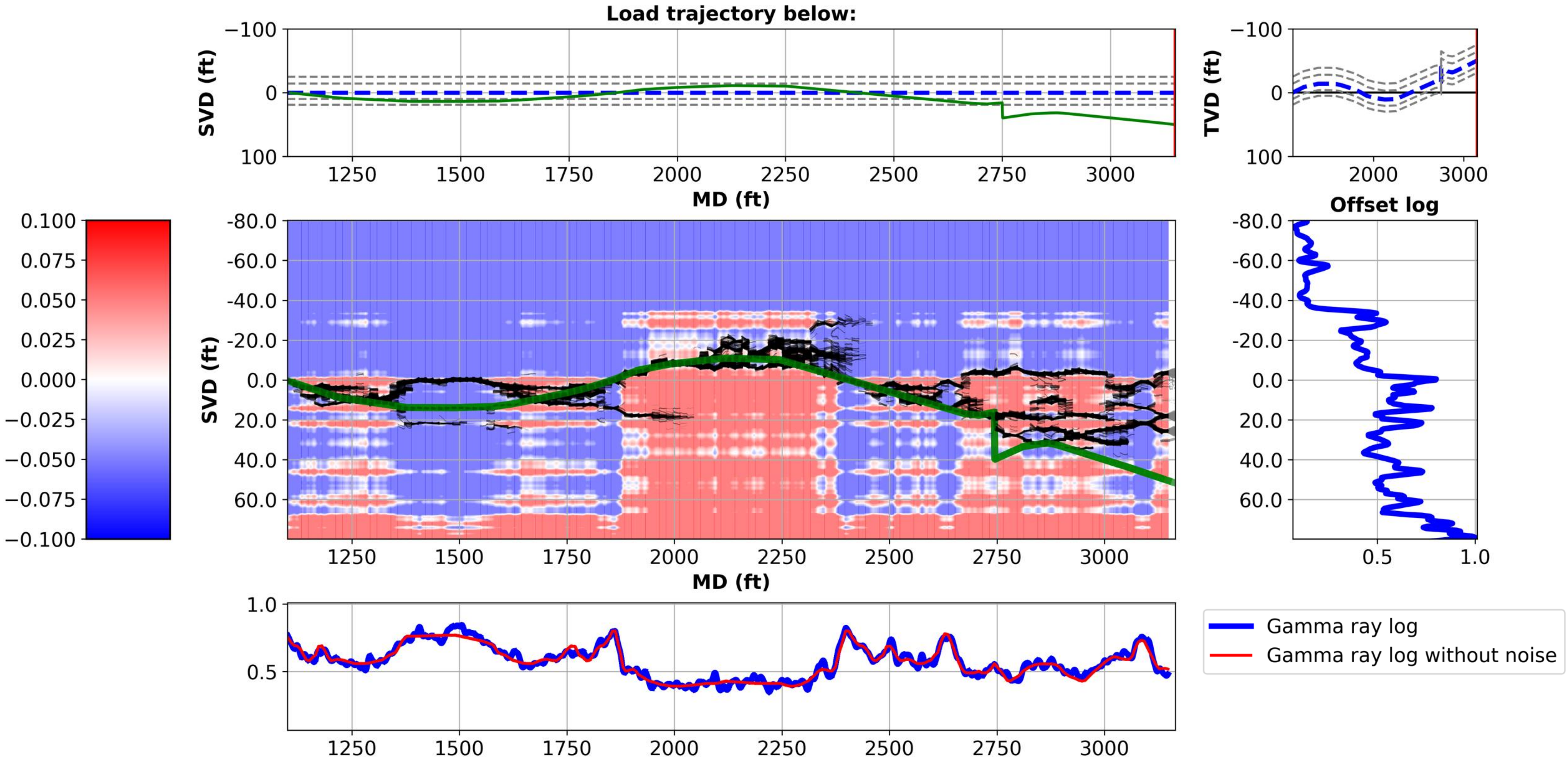
7-mode predictor	Evaluation noise			
Noise, training ↓	**1%**	**2%**	**4%**	**10%**
0-base-case	0.9444	0.9444	0.9444	0.9444
***0%**	0.7687	1.0673	1.4507	1.8966
1%	0.5609	0.6824	1.0129	1.5737
2%	0.5592	0.6008	0.7497	1.2594
4%	0.6226	0.6354	0.6809	0.9458
10%	0.7290	0.7322	0.7429	0.7996

- Best results in a column when training noise is equal to evaluation noise
- Best results in a row when noise is smallest
- Trained with 10% noise is on-par noiseless model tested with 1% noise
  - Our method can handle the log noise above the industry standard 6-7%.

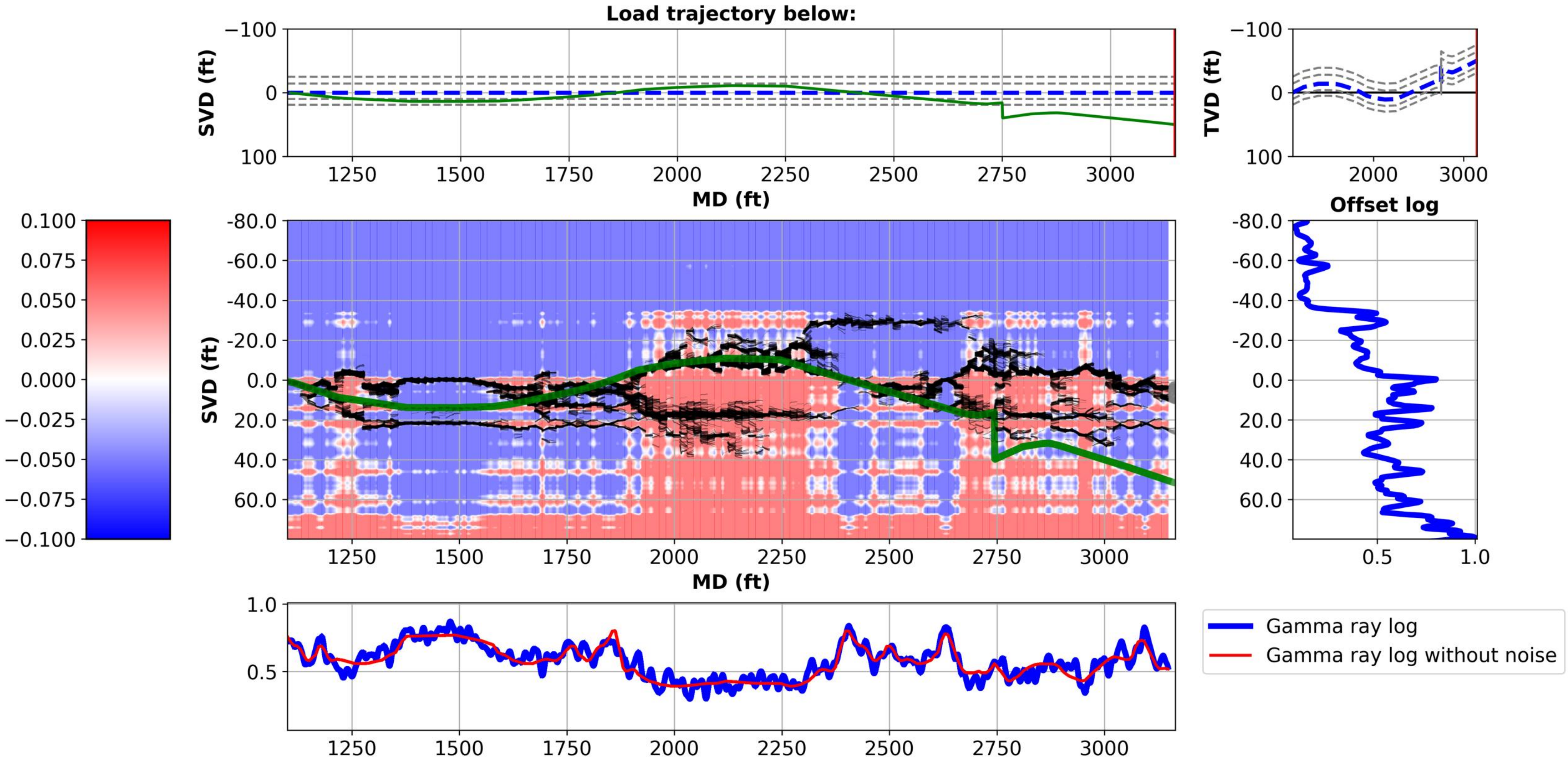
GWC dataset 0% noise



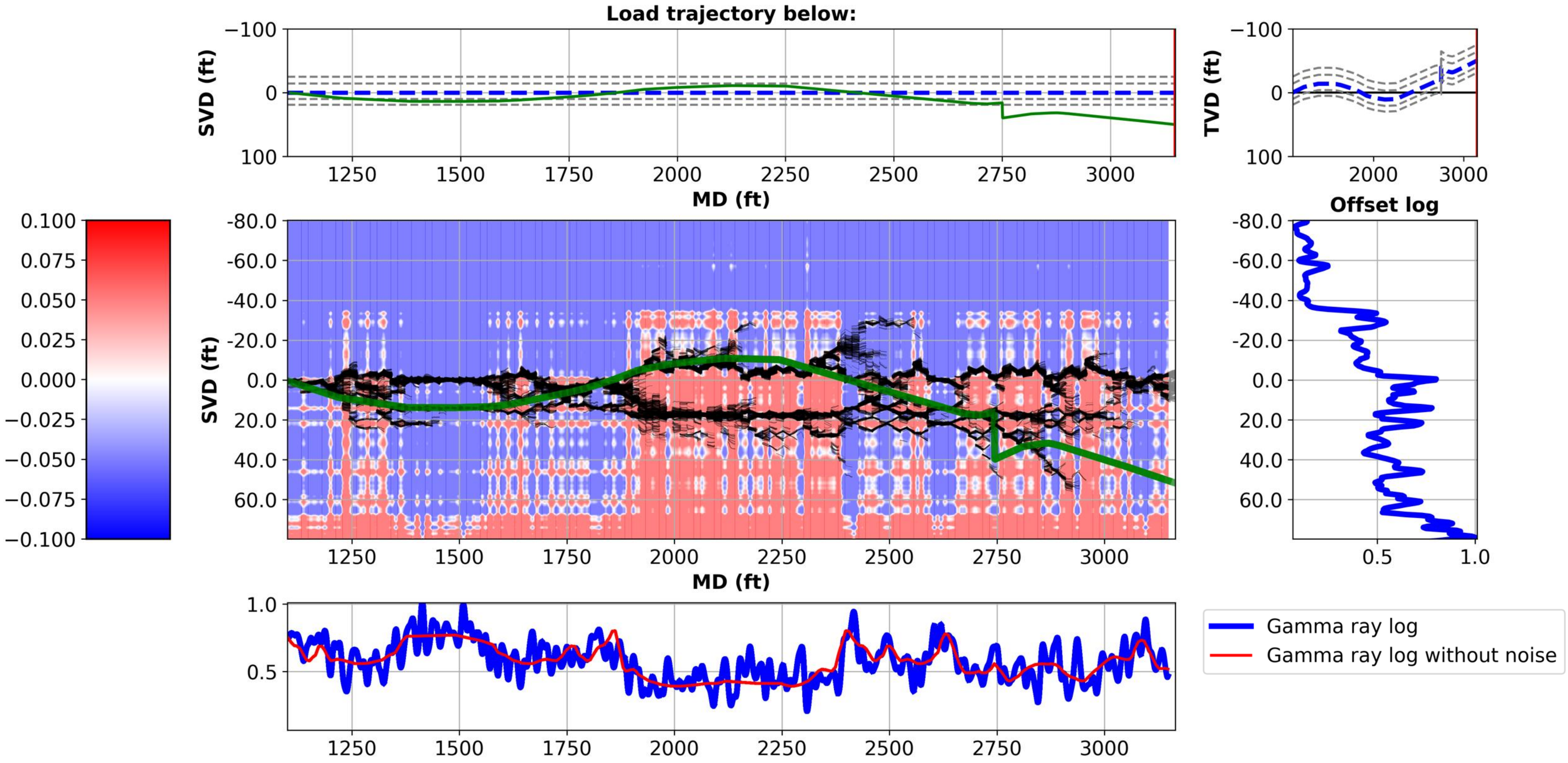
GWC dataset 1% noise



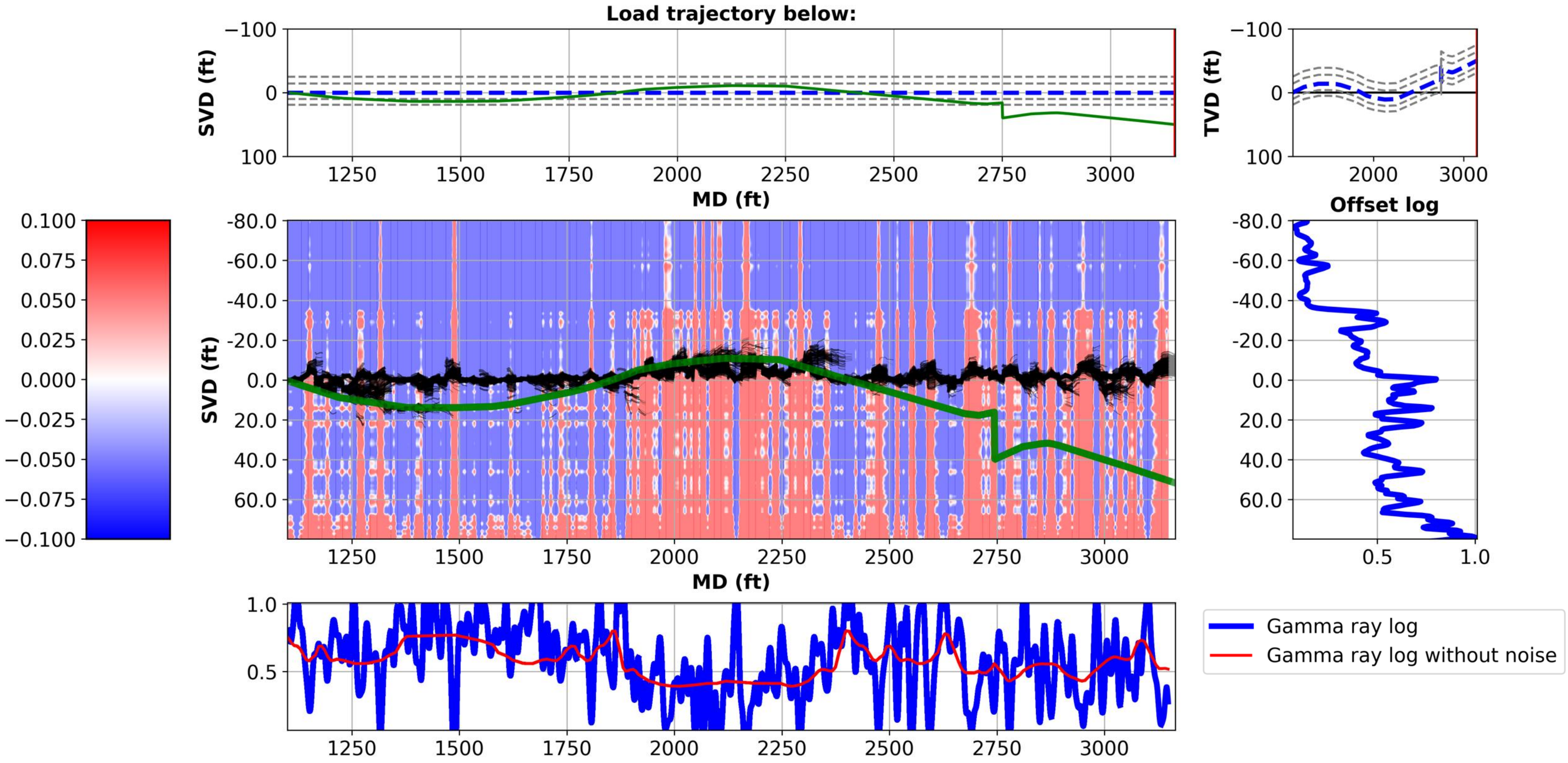
GWC dataset 2% noise



GWC dataset 4% noise



GWC dataset 10% noise



Conclusions

- New MDN learns
    1. **ability to interpret logs**
    2. **geological information to quantify likelihoods**
    3. **handles realistic 4% noise**

- Probabilistic interpretation tracks several solutions
    - Robust even in complex environment
    - At 800 m/s **100 times faster** than earlier probabilistic methods

- Current status:
    - Waiting for more funding since my parental leave ☺