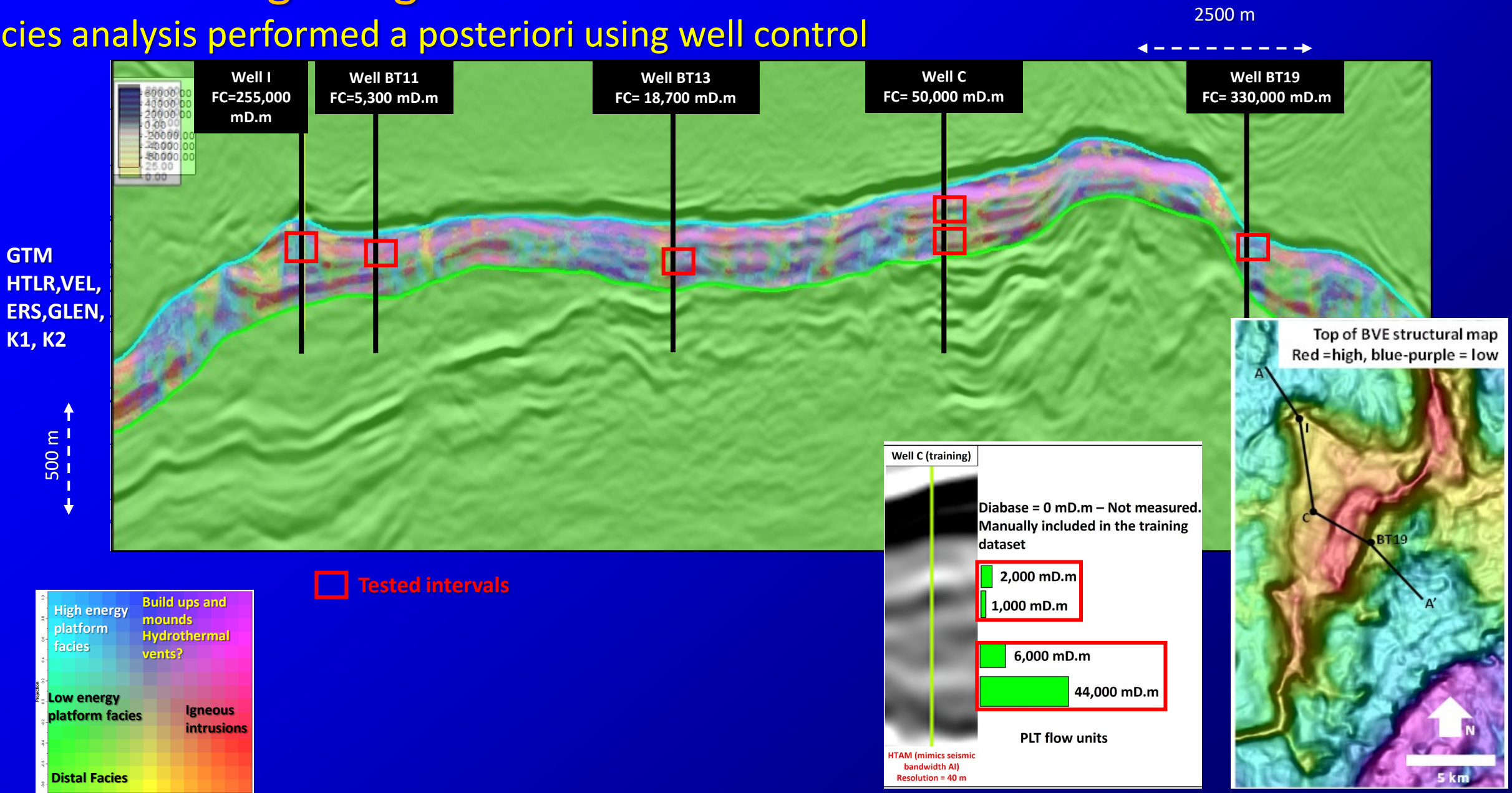
Generating Training Data from Wells for use in Supervised and Semi-Supervised Machine Learning

Kurt J. Marfurt
The University of Oklahoma

SOM clustering using six attributes.

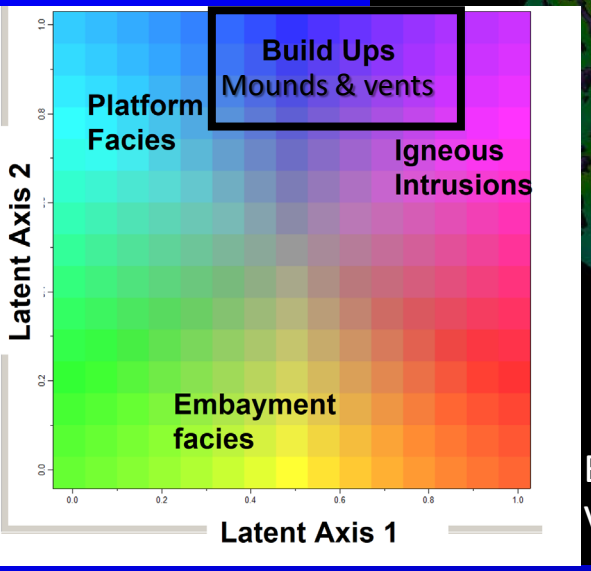
Facies analysis performed a posteriori using well control

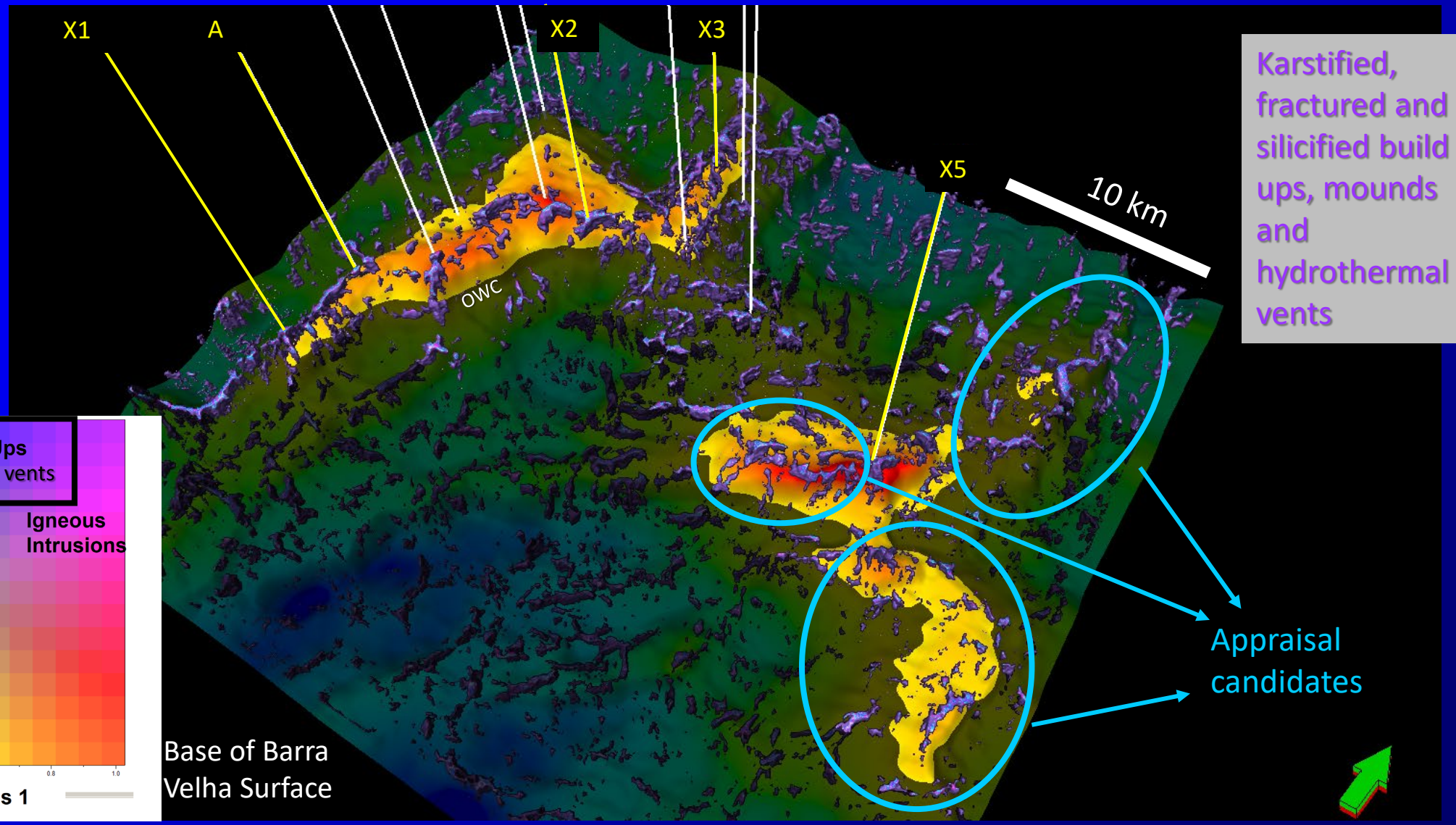


(Maas, 2025)

Final SOM model showing high potential buildups, mounds, and vents

3D view of SOM clusters

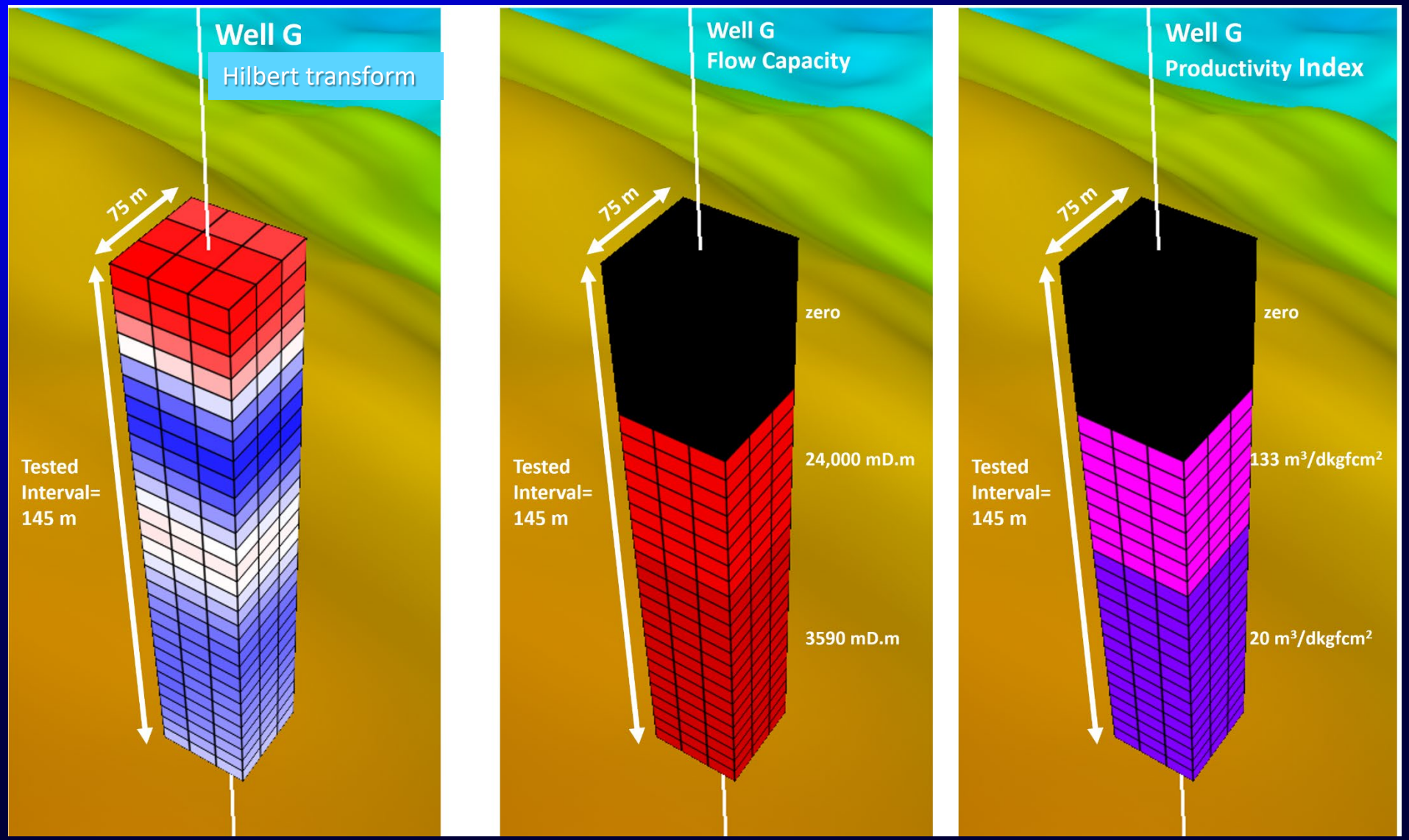




Karstified, fractured and silicified build ups, mounds and hydrothermal vents

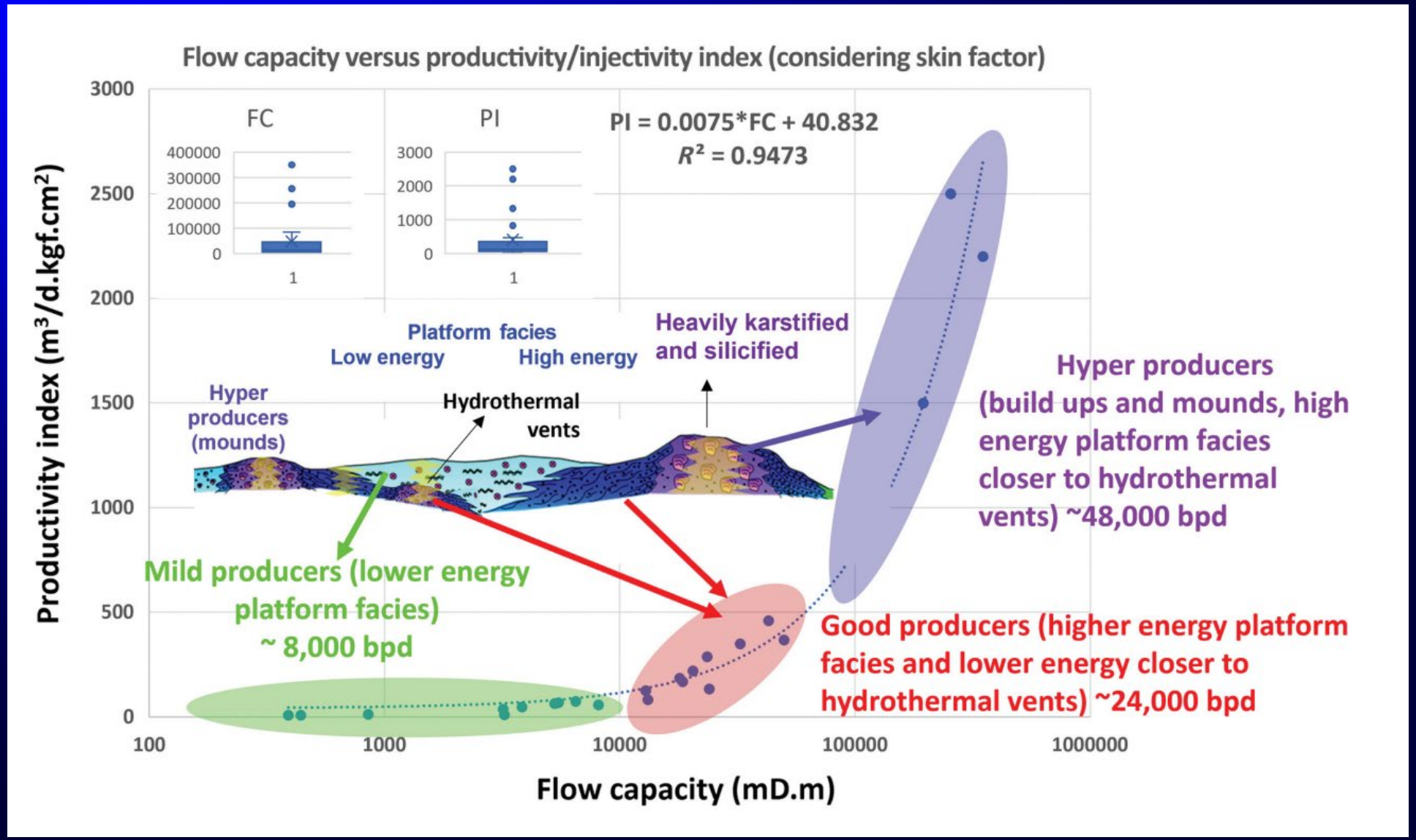
(Maas, 2025)

Training dataset building- 10 wells – 10 attributes – 4968 datapoints



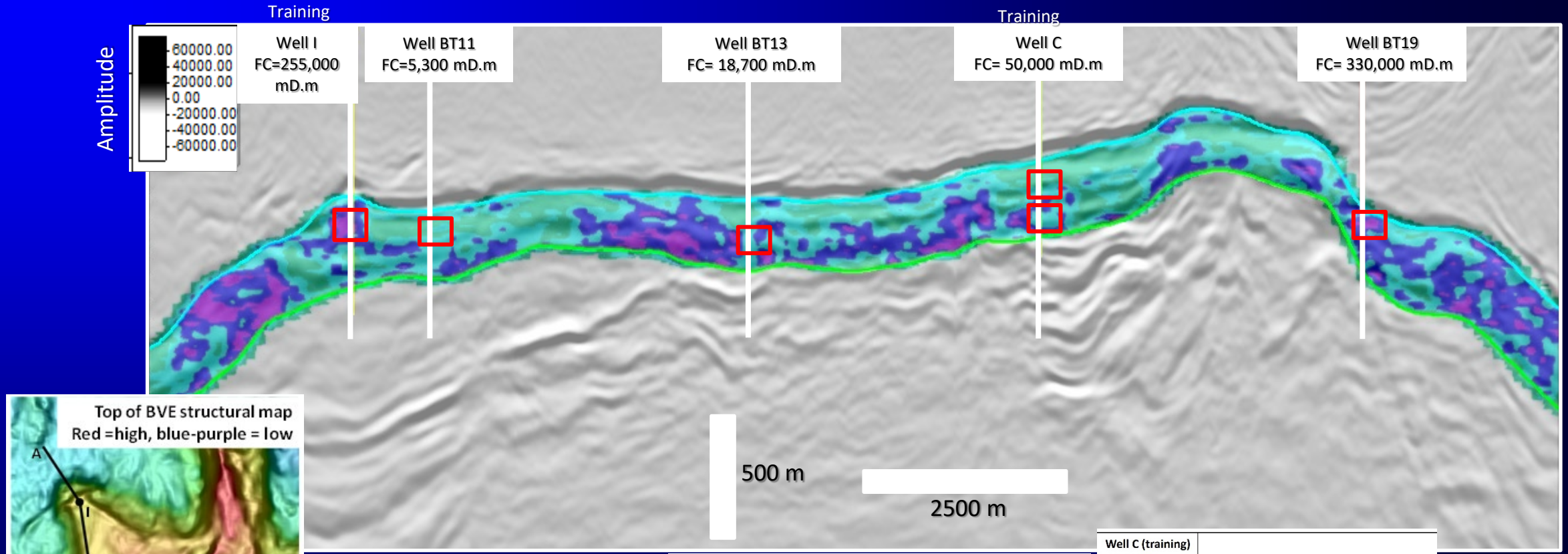
(Maas et al., 2024)

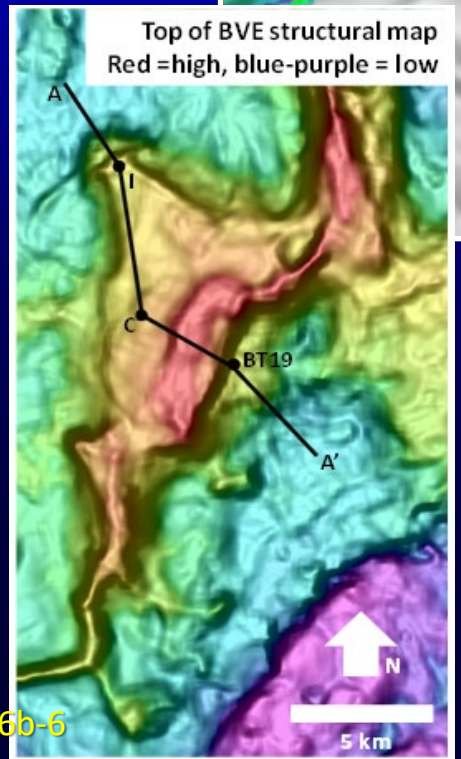
Attribute prediction of flow capacity – Mero Field offshore Brazil



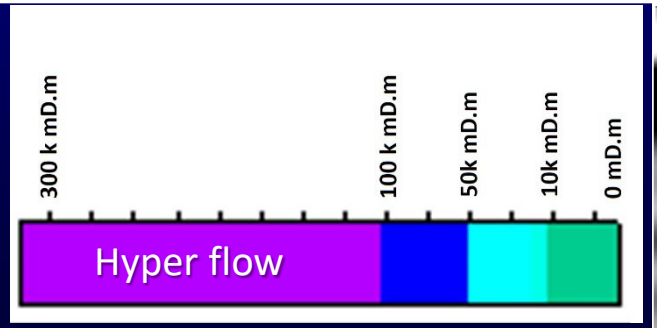
(Maas et al., 2024)

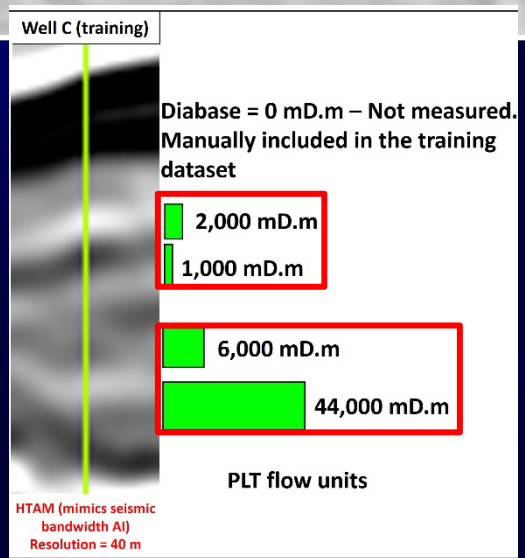
Random Forest– Flow Capacity Regression – Six attributes





Tested intervals

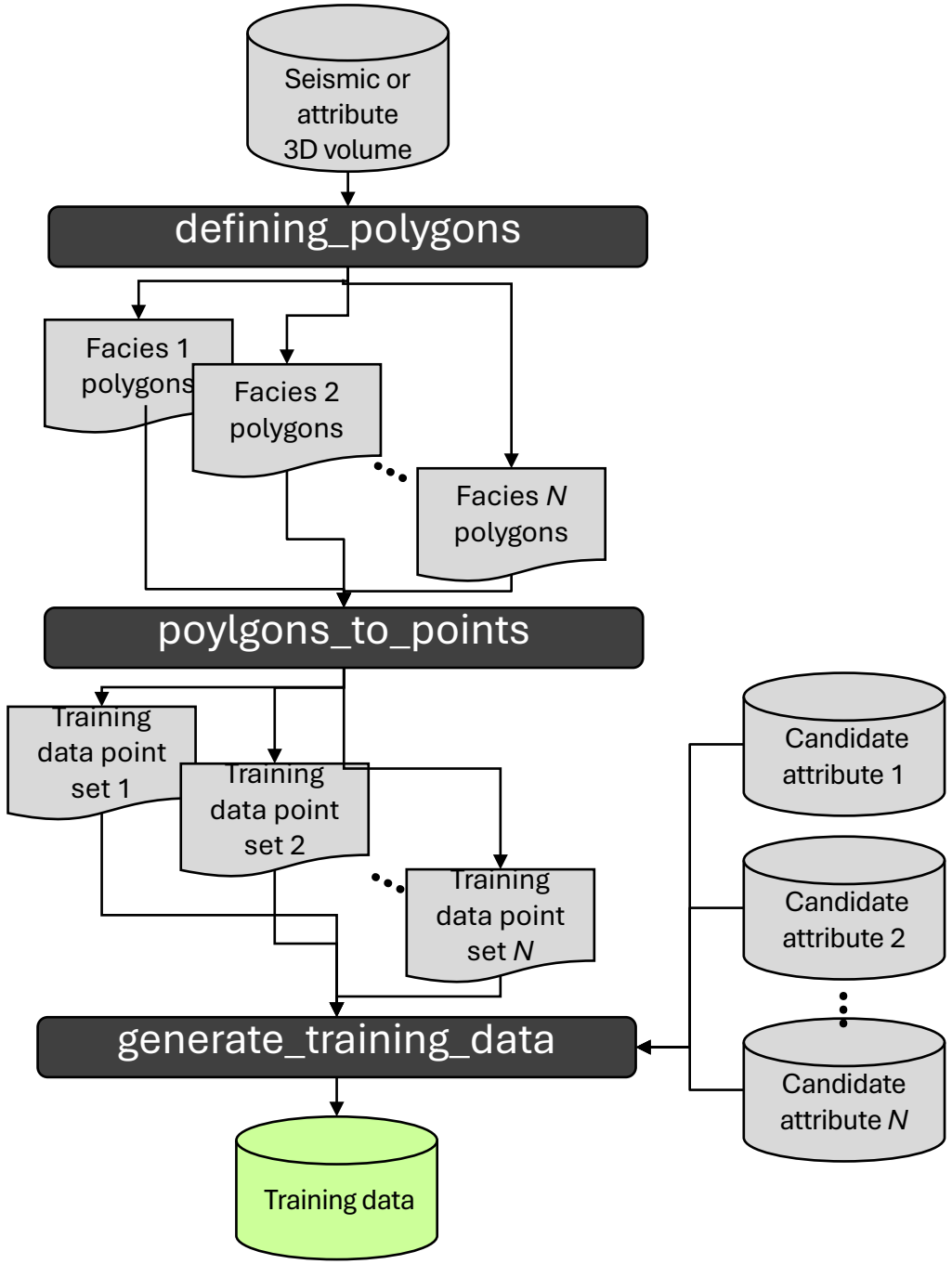




(Maas, 2025)

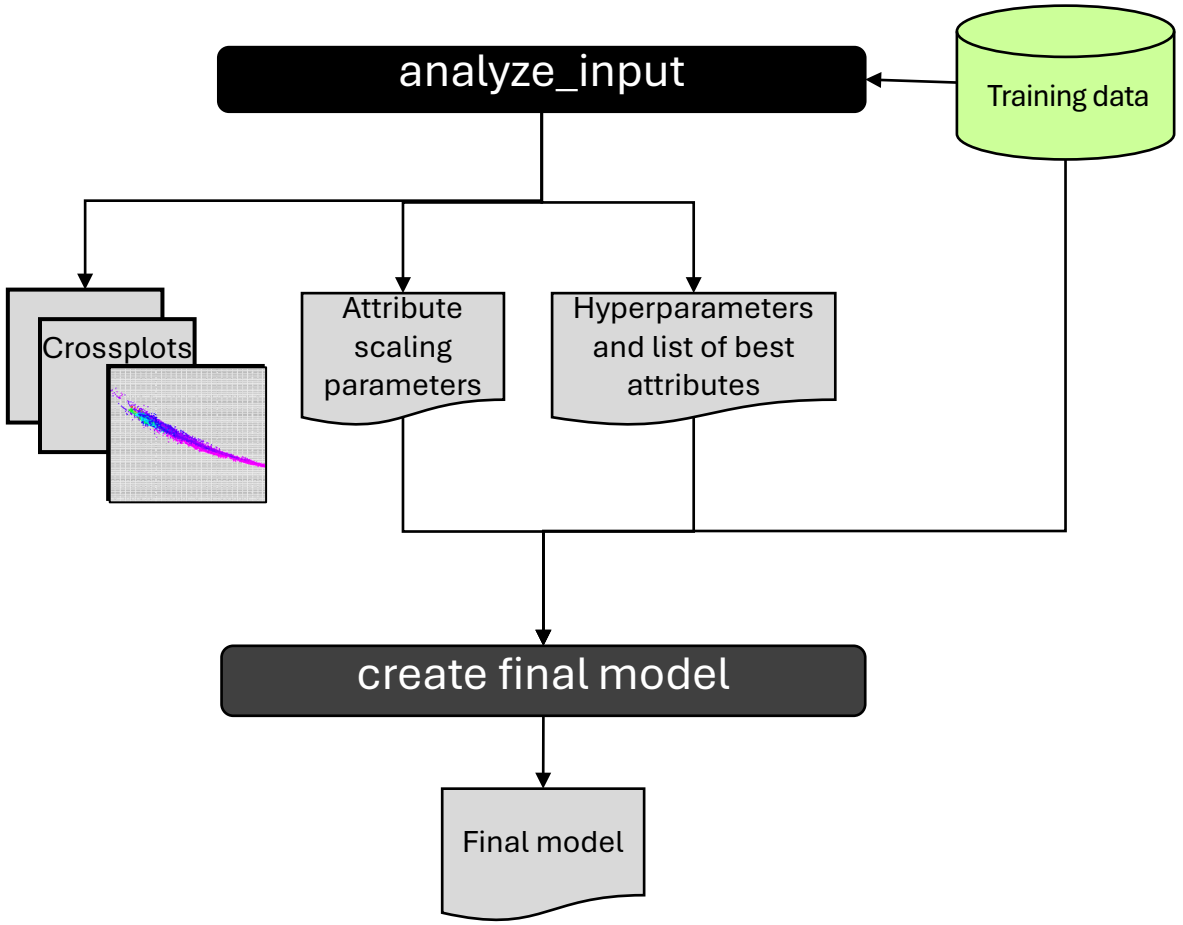
A quick review of the AASPI Multiattribute ML Workflow

Generate Training Data



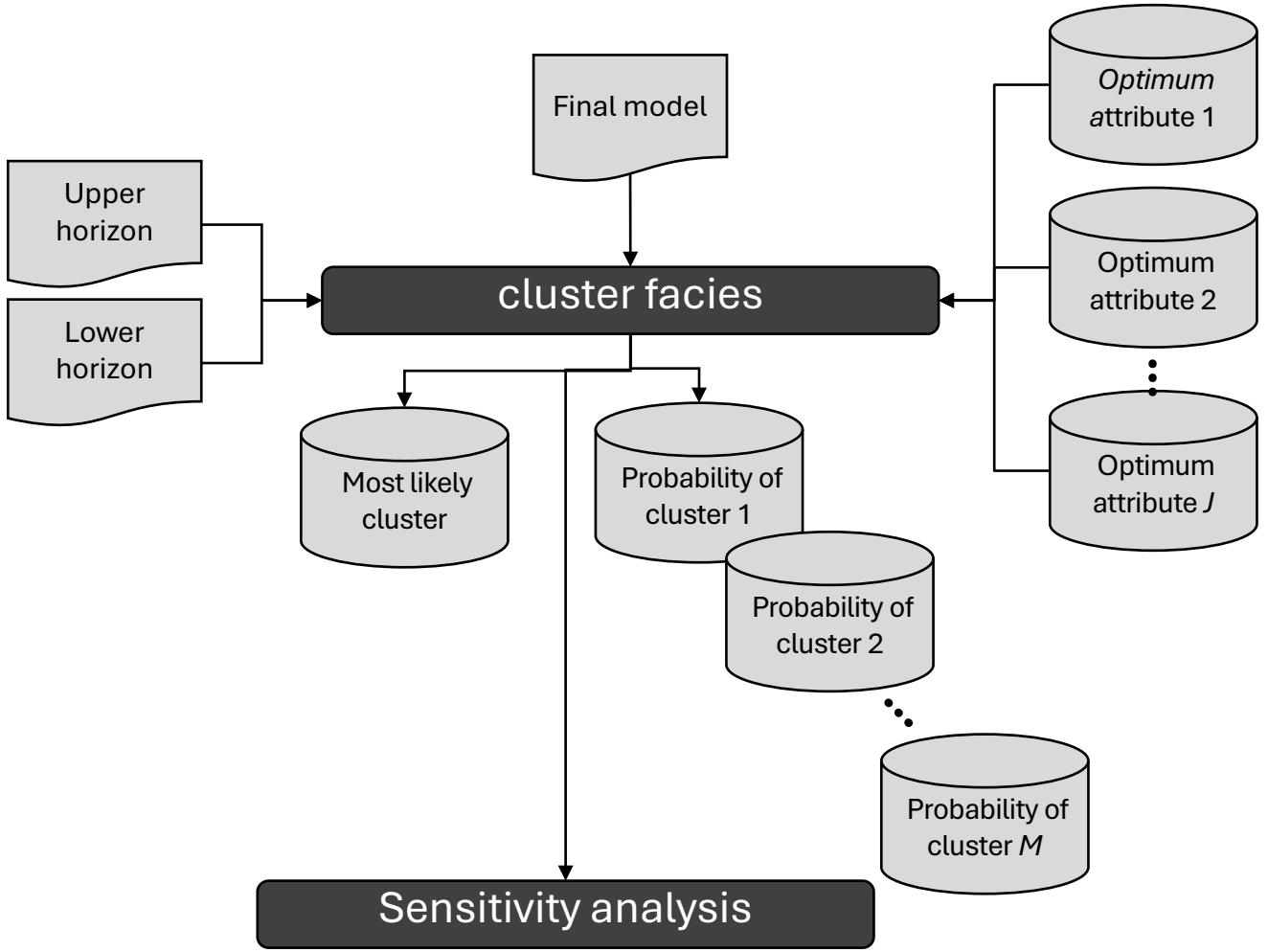
A quick review of the AASPI Multiattribute ML Workflow

- *Analyze the training data*
  - *Determine attribute scaling*
  - *Determine "best" attribute combination*
- *Generate the final model*



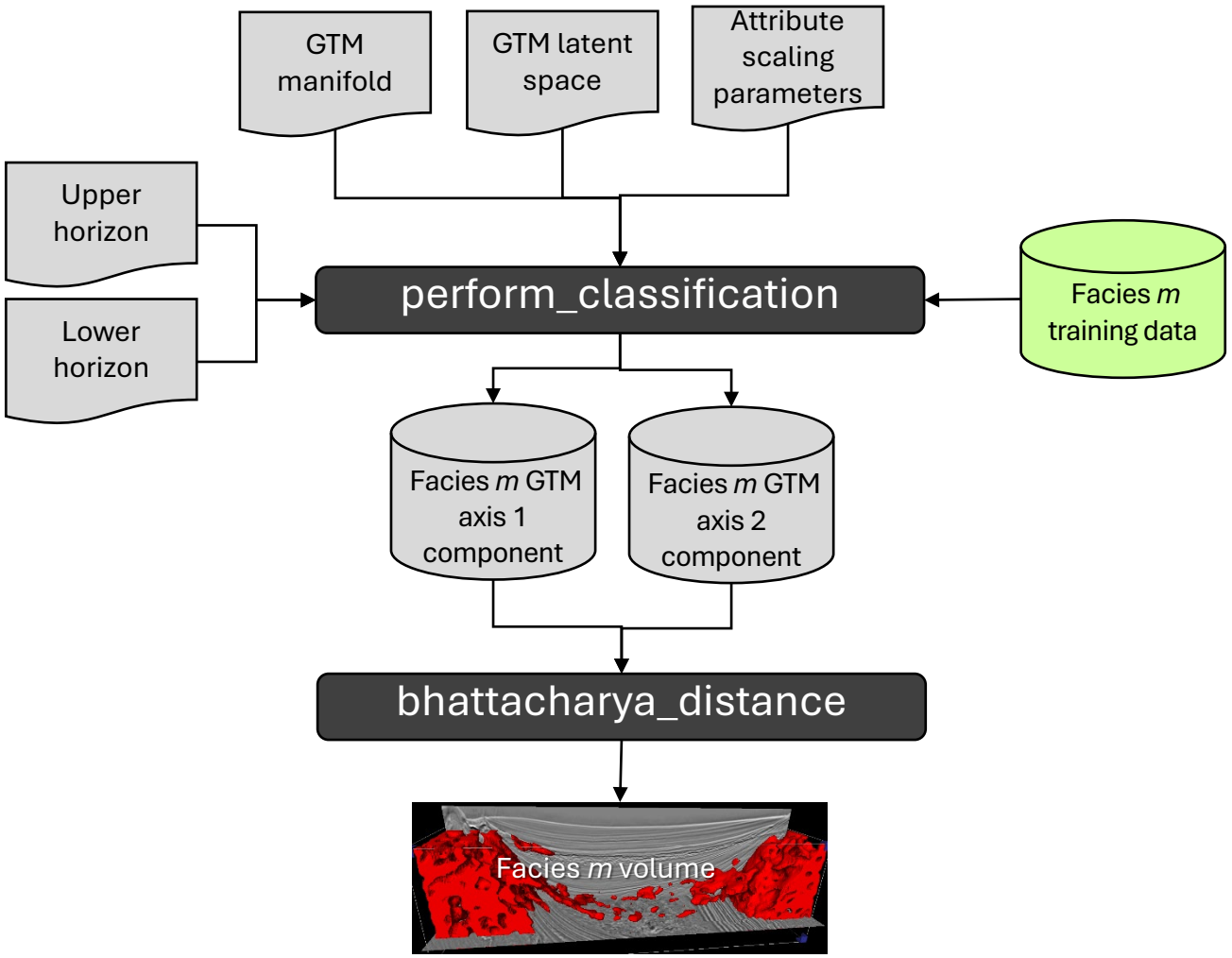
A quick review of the AASPI Multiattribute ML Workflow

- *Cluster analysis (e.g. k-means++, GMM)*

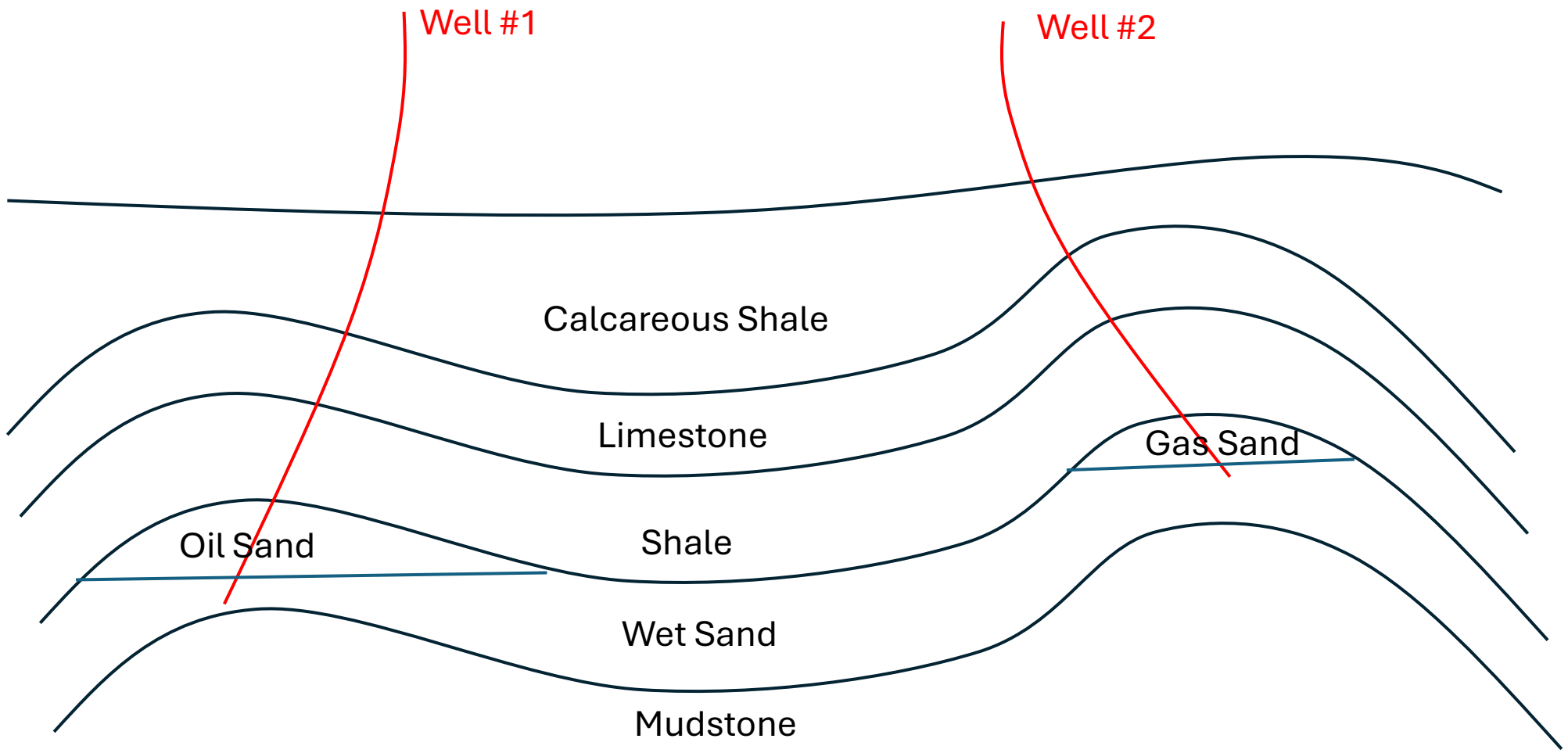


A quick review of the AASPI Multiattribute ML Workflow

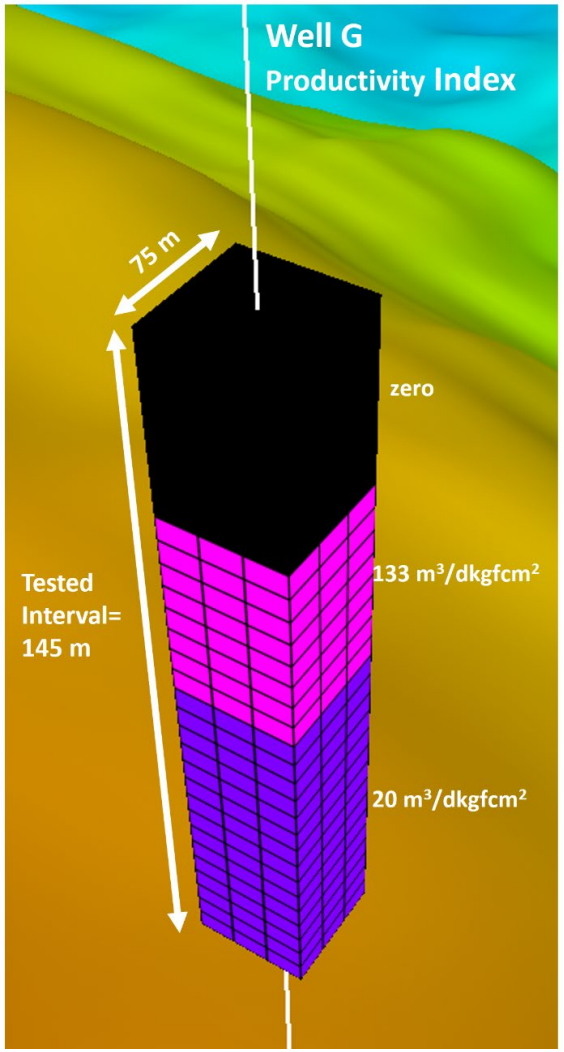
- *Projection techniques (e.g. GTM, SOM)*



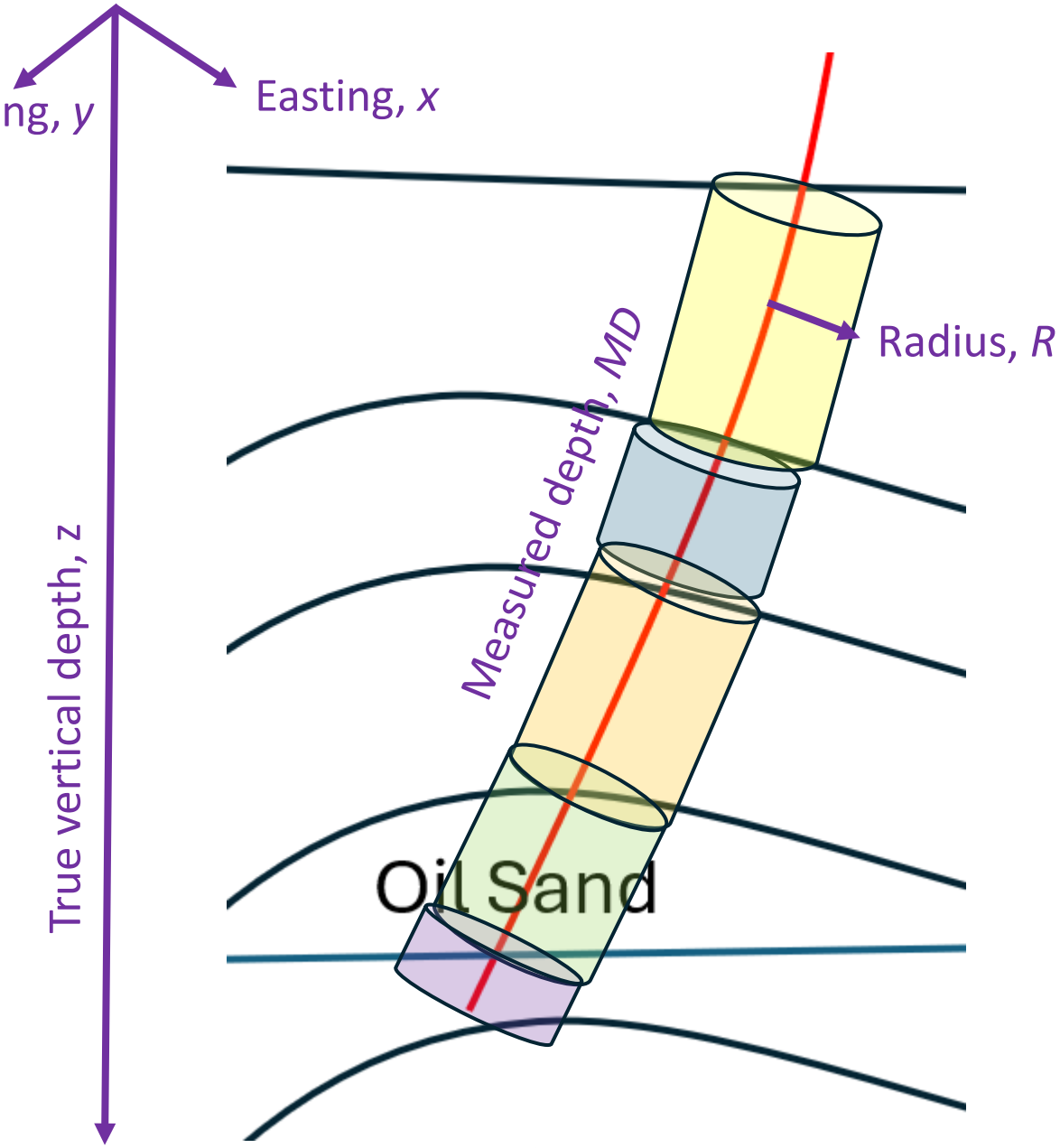
Today’s objective: Generate point sets about wells



Generating point sets about wells



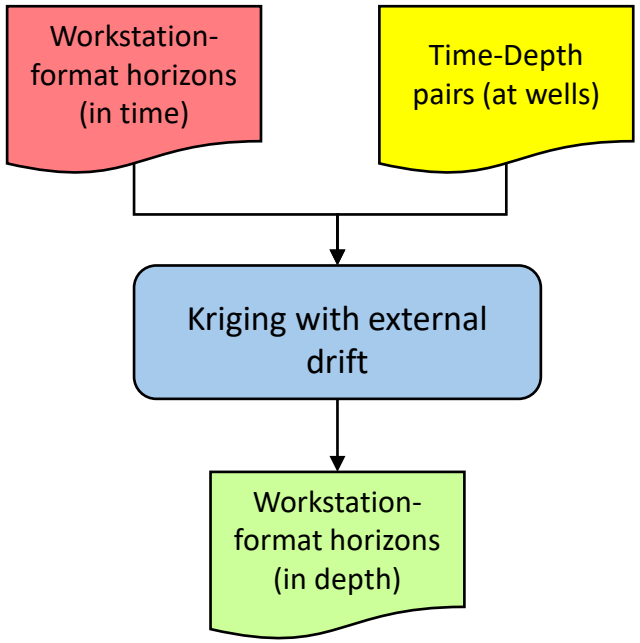
(Maas et al., 2024)



AASPI program
define_point_sets_about_wells

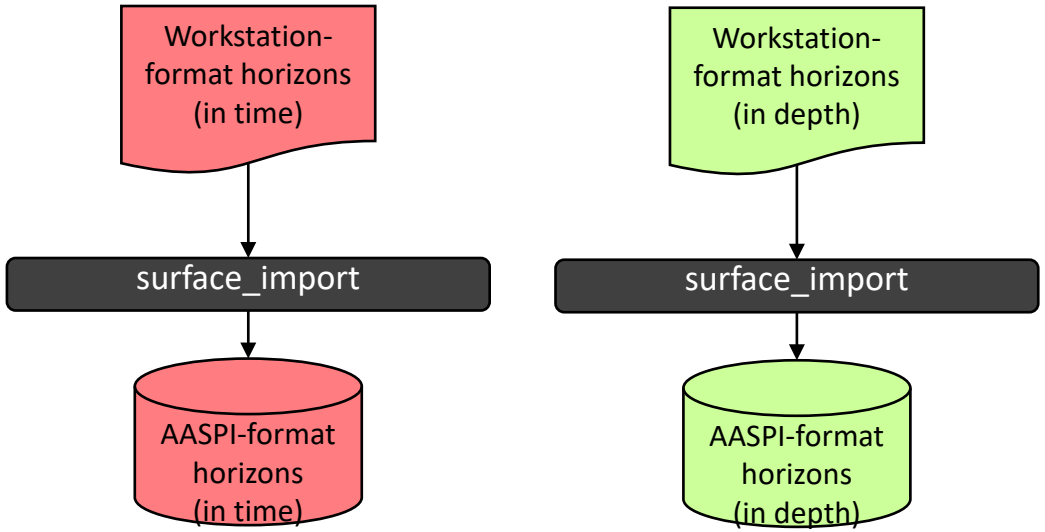
Generating point sets about wells

Step 0: Use your time-depth pairs and workstation software to convert time horizons to depth horizons

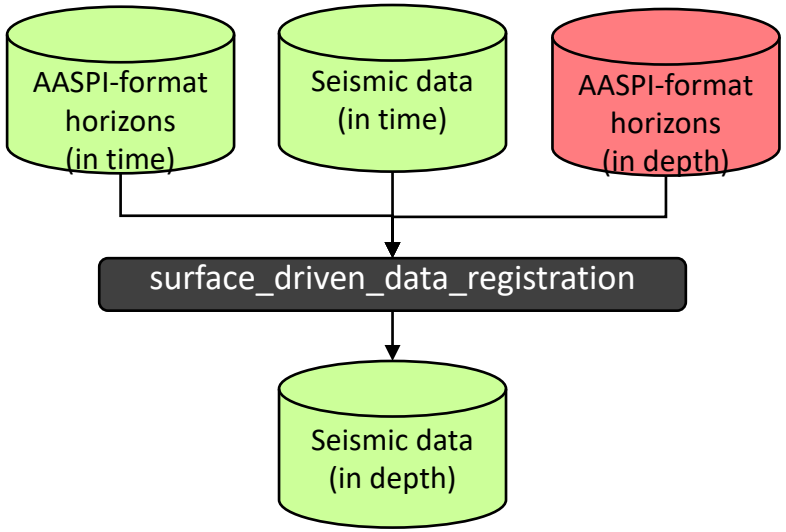


Generating point sets about wells

Step 1: Convert workstation-format horizons to AASPI format

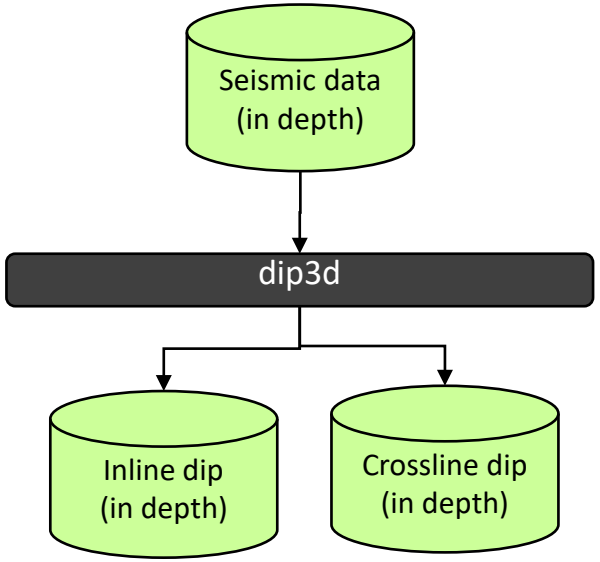


Step 2: Convert time-domain seismic data to depth

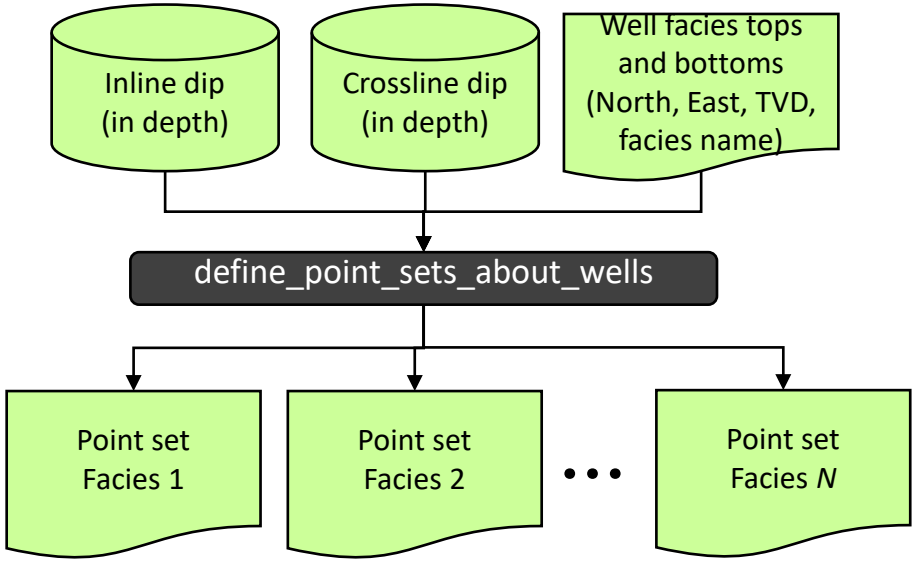


Generating point sets about wells

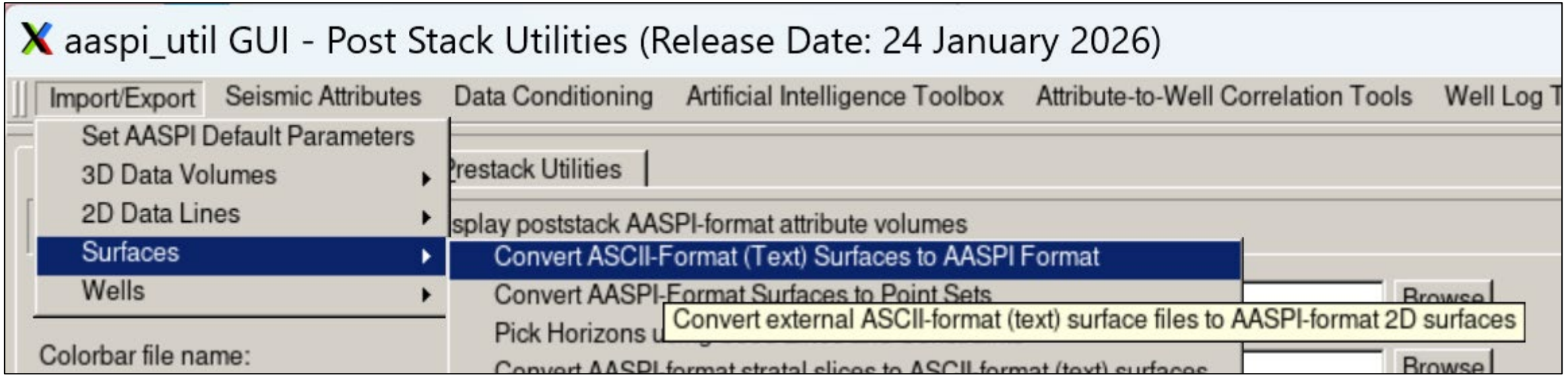
Step 3: Compute structural dip



Step 4: Define the point sets about the wells

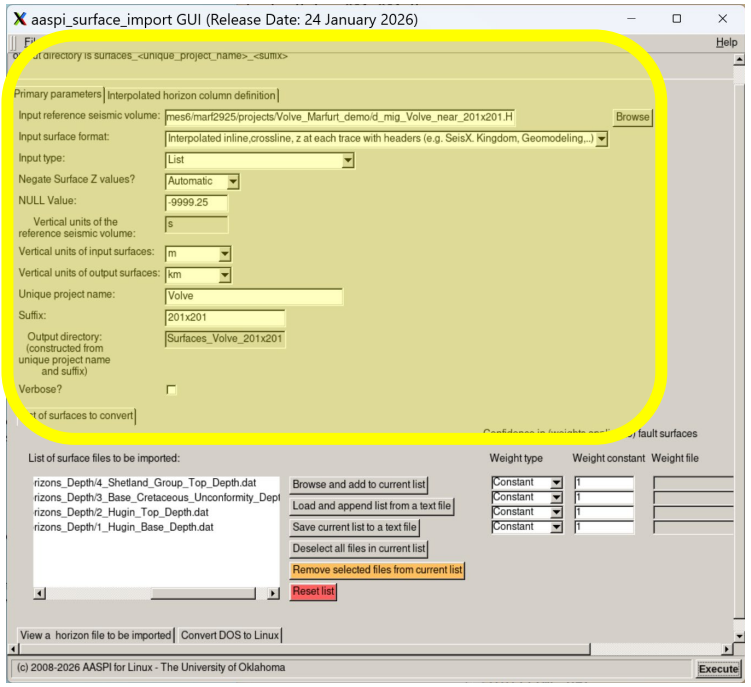


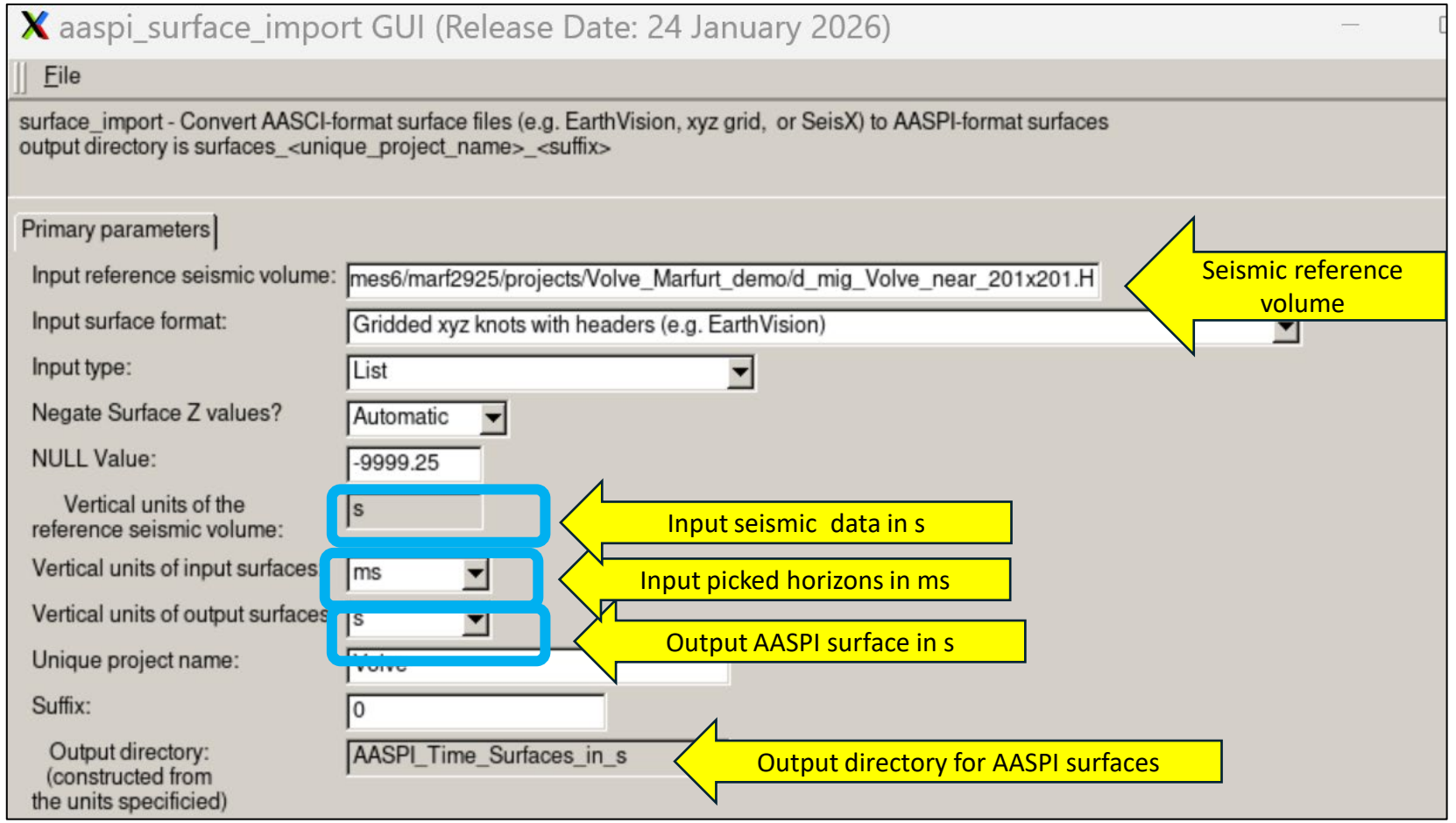
Step 1: Convert workstation surfaces to AASPI format



Step 1a: Convert workstation time surfaces to AASPI format

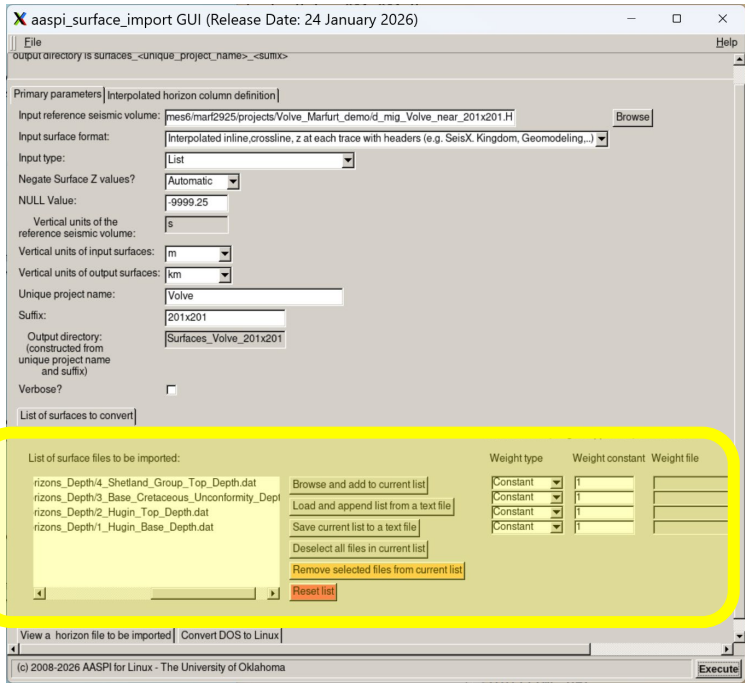
- *First, read in the reference seismic data volume and define vertical units*

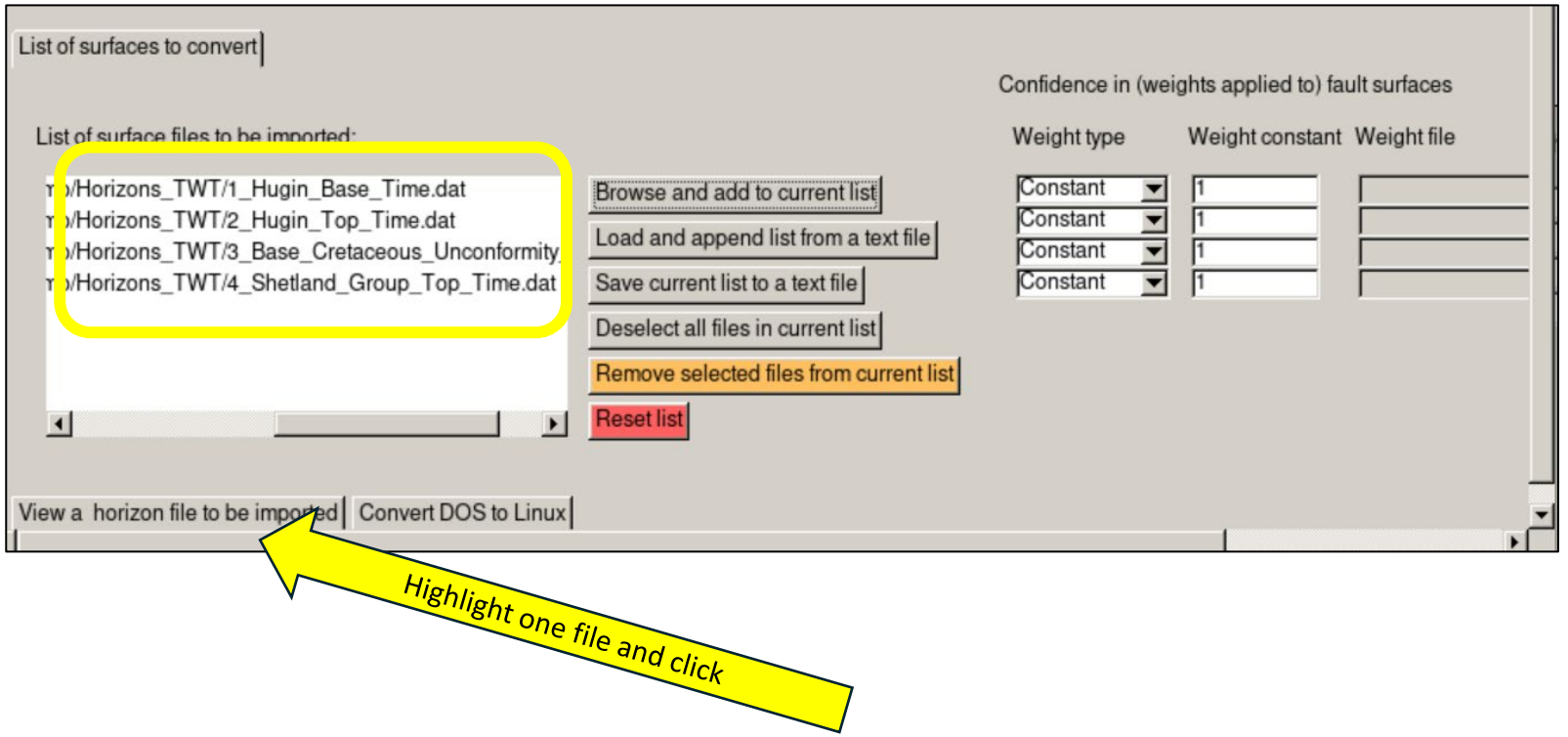




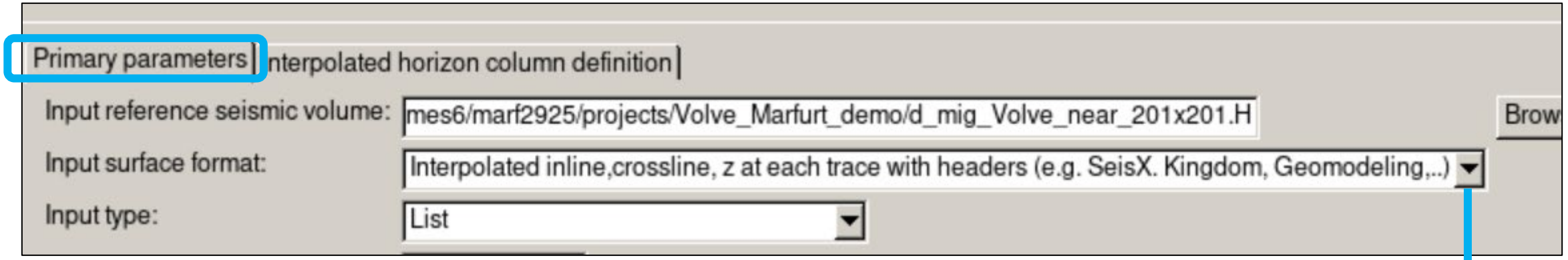
Step 1a: Convert workstation time surfaces to AASPI format

- *Second, select the workstation **Time** horizons (in the Horizons_TWT folder)*





- *Then, third, define the format of the workstation-generated TIME horizon files*

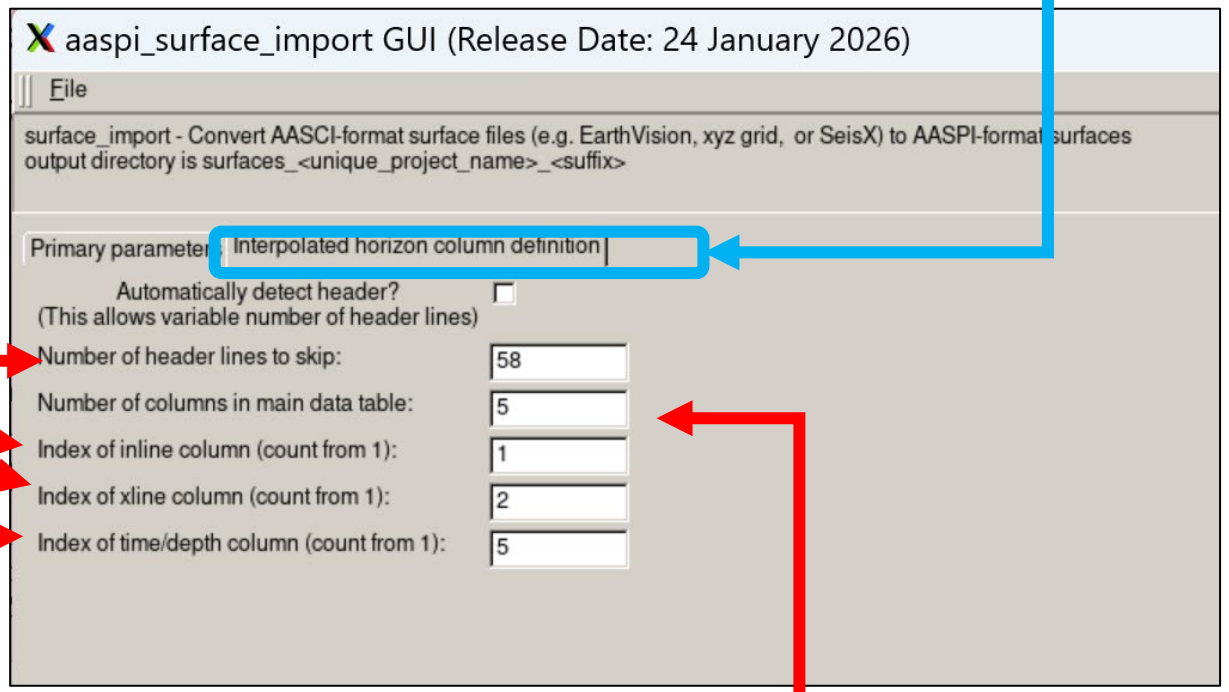


Input file 1_Hugin_Base.dat

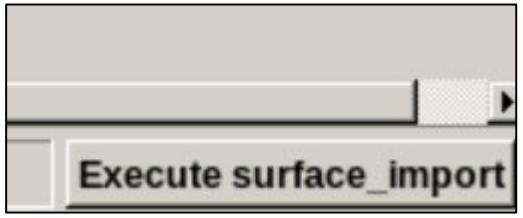
Input file 1 format="SeisX"

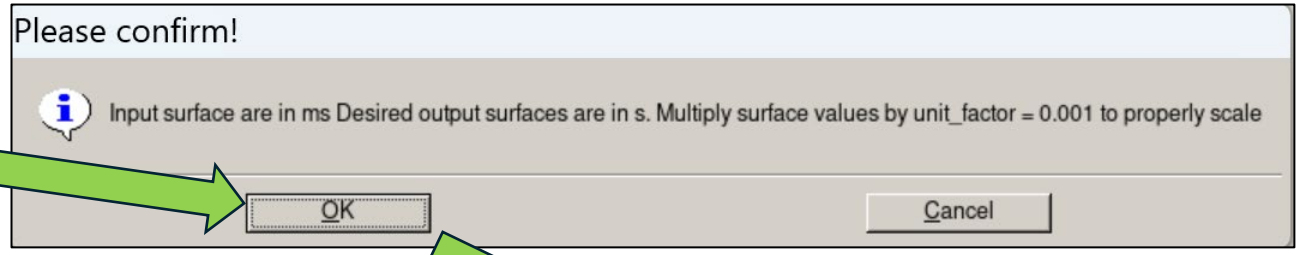
kmarfurt@xphi1:~/projects/Volve_Marfurt_demo/Horizons_TWT

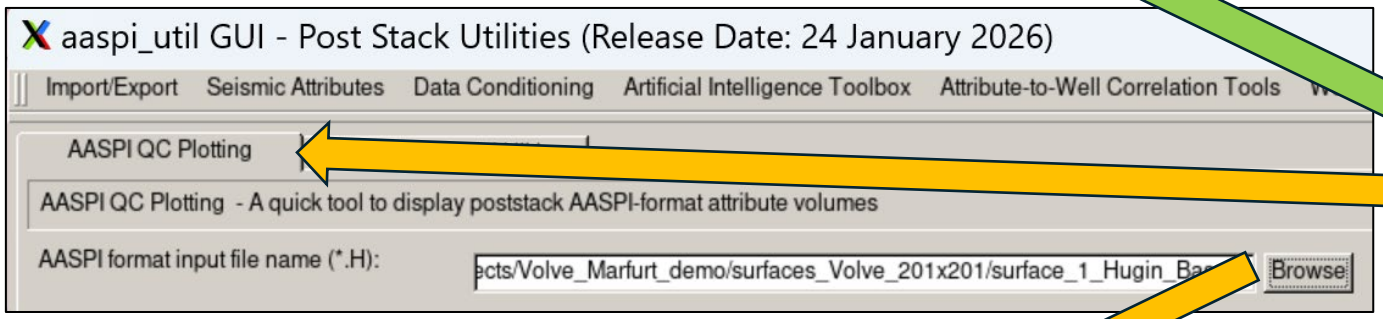
```
32 #       value:null
33 #       trigger:false
34 #
35 #     Delimited:
36 #       repeating:false
37 #       delimiter:,
38 #       whitespace:false
39 #
40 #     local delimiter:","
41 #
42 #     Item:surveyName               line number: 3 delimiter field number: 1
43 #     Item:horizonName              line number: 4 delimiter field number: 1
44 #     Item:dataSource               line number: 5 delimiter field number: 1
45 #     Item:horizonAttribute         line number: 6 delimiter field number: 1
46 #     Item:remark                   line number: 7 delimiter field number: 1
47 #     Item:line                     line number: 8 delimiter field number: 1
48 #     Item:trace                    line number: 8 delimiter field number: 2
49 #     Item:x                        line number: 8 delimiter field number: 3
50 #     Item:y                        line number: 8 delimiter field number: 4
51 #     Item:z                        line number: 8 delimiter field number: 5
52 #     Item:surveyType               constant:3D Survey
53 #
54 ST10010
55 Hugin_Fm_Base+ST10010ZC11_Near_190314
56 STAT
57 TTME
58 ,
59 10031.0000000,2470.0000000,432340.9250911,6477699.2535549,2767.3488770
60 10032.0000000,2470.0000000,432343.9494647,6477711.3821630,2767.9340820
61 10033.0000000,2470.0000000,432346.9738383,6477723.5107711,2771.2666016
62 10034.0000000,2470.0000000,432349.9982119,6477735.6393792,2771.6528320
63 10035.0000000,2470.0000000,432353.0225855,6477747.7679873,2772.0661621
64 10005.0000000,2469.0000000,432274.4199852,6477380.8853657,2759.6750488
```

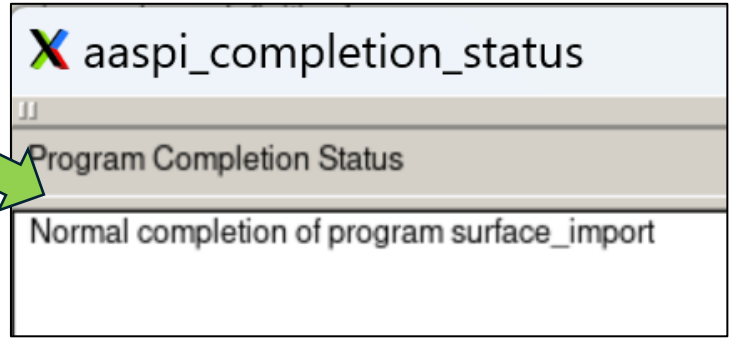


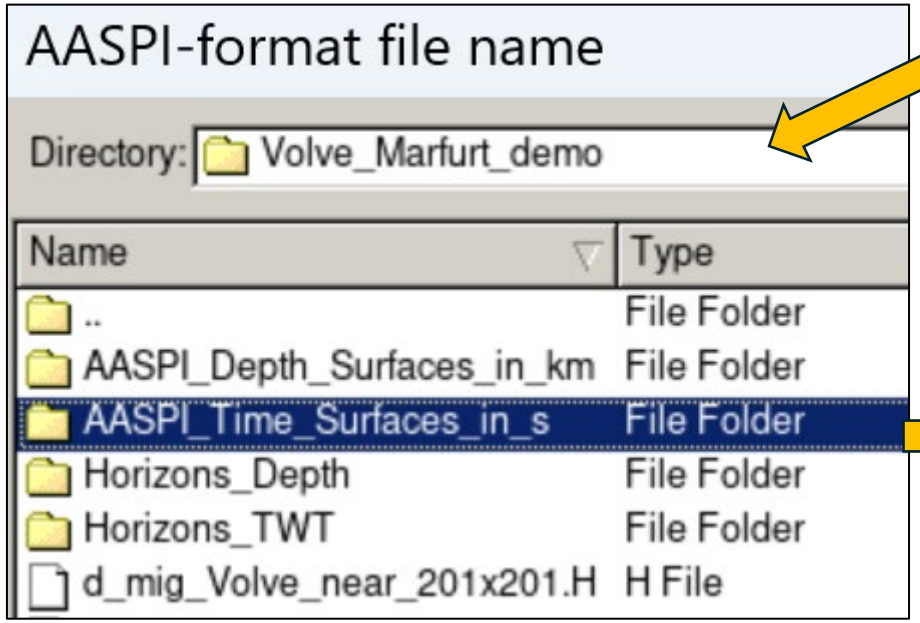
- *Fourth, execute the program and plot the results for quality control*

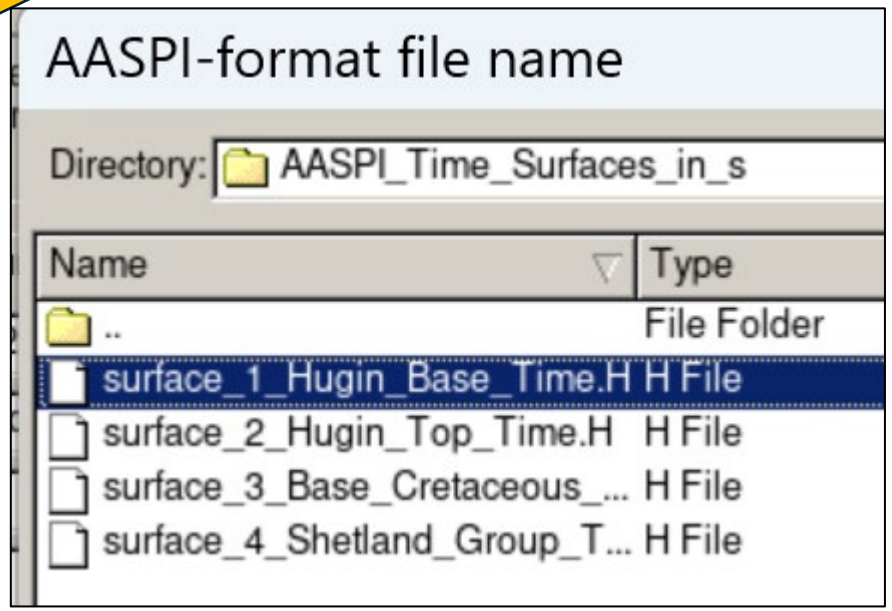


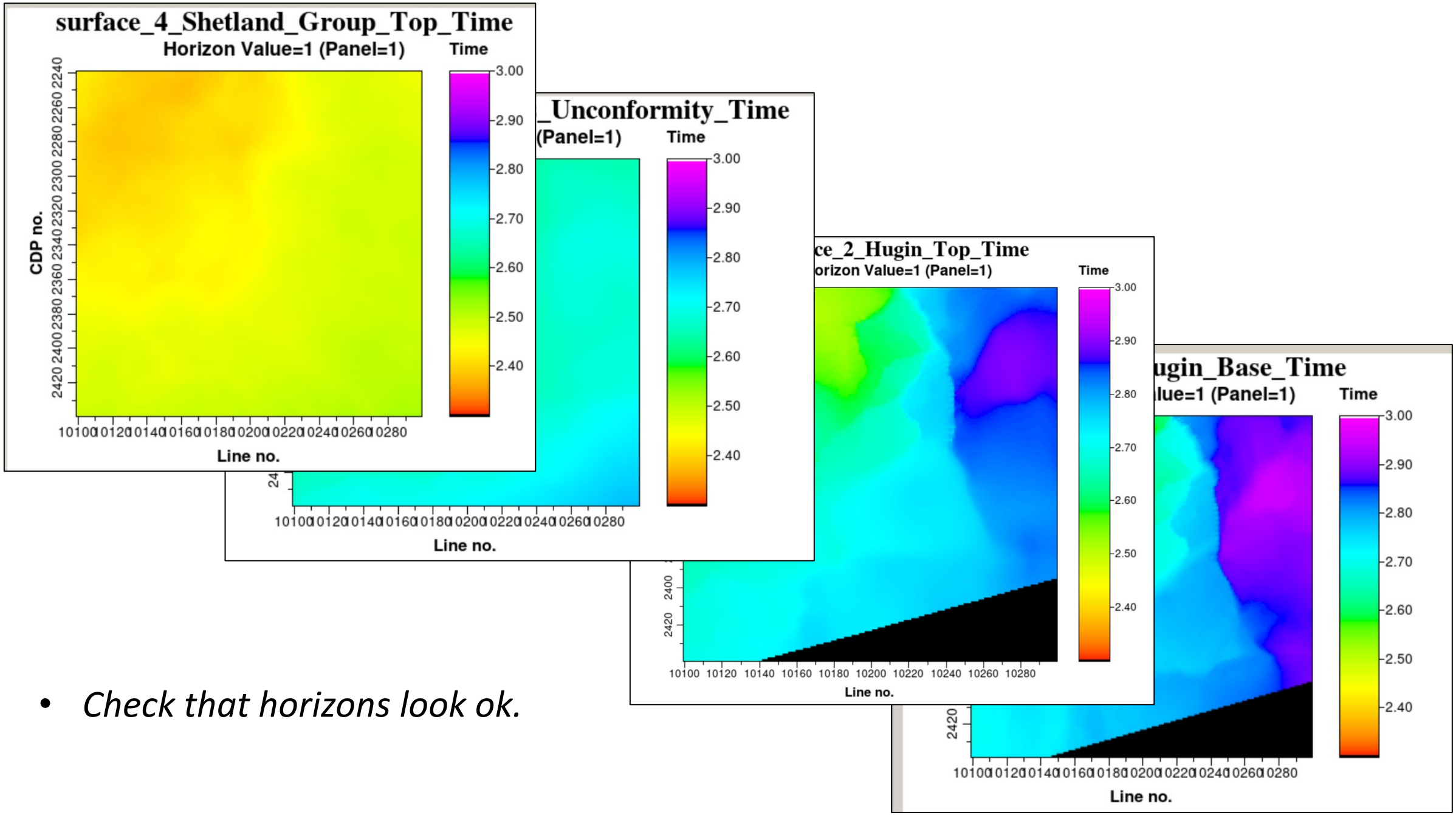








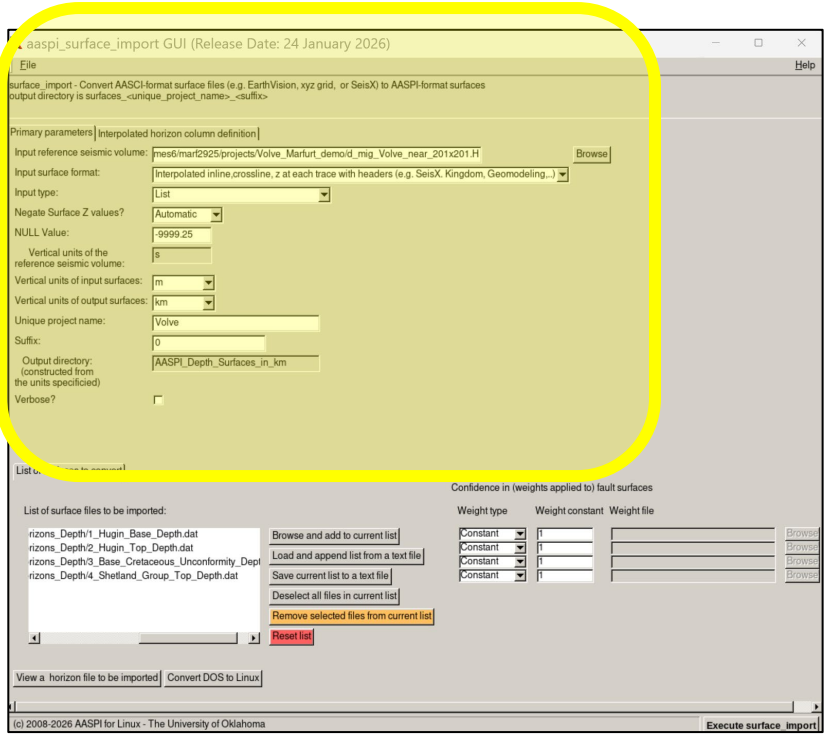


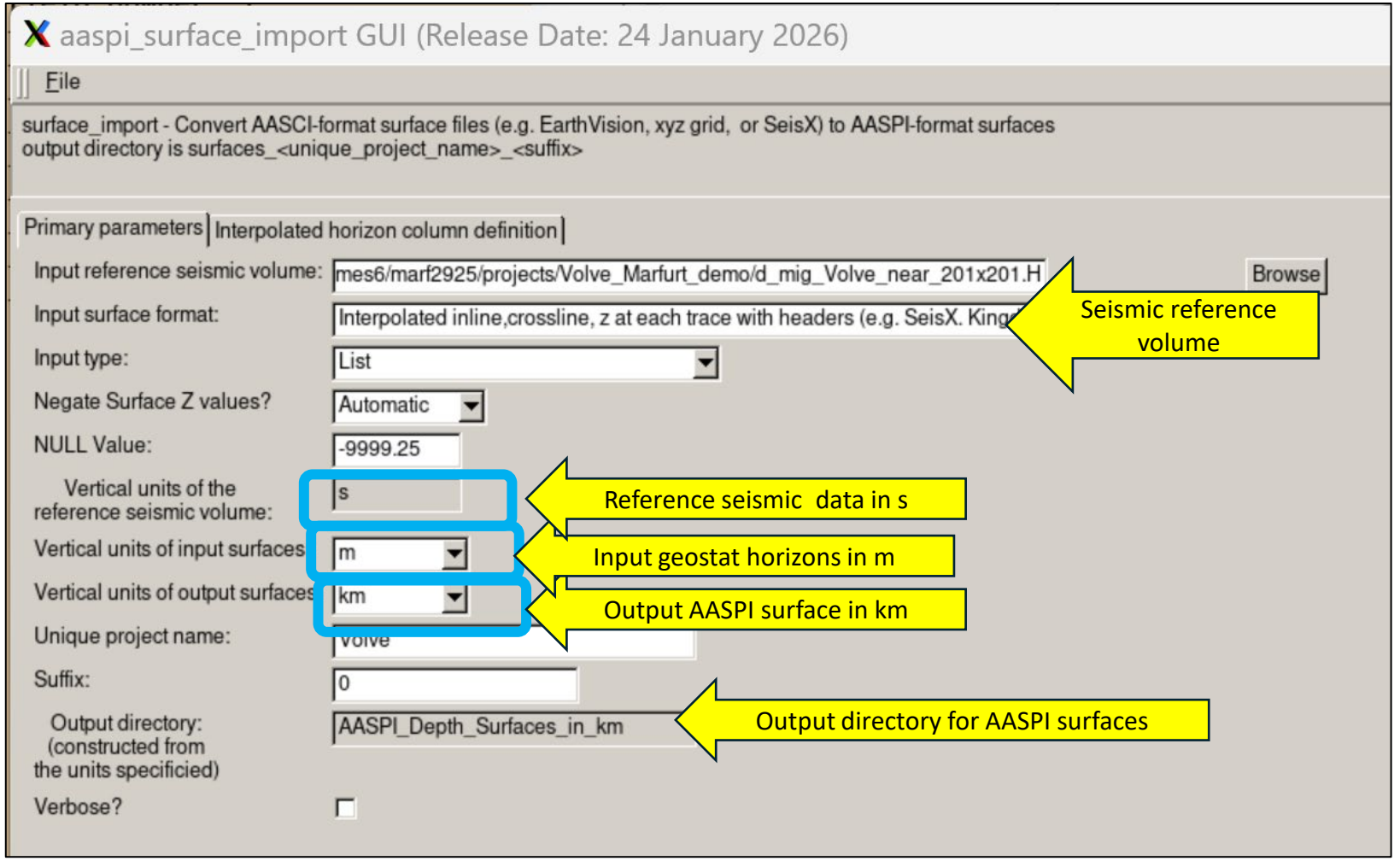


- *Check that horizons look ok.*

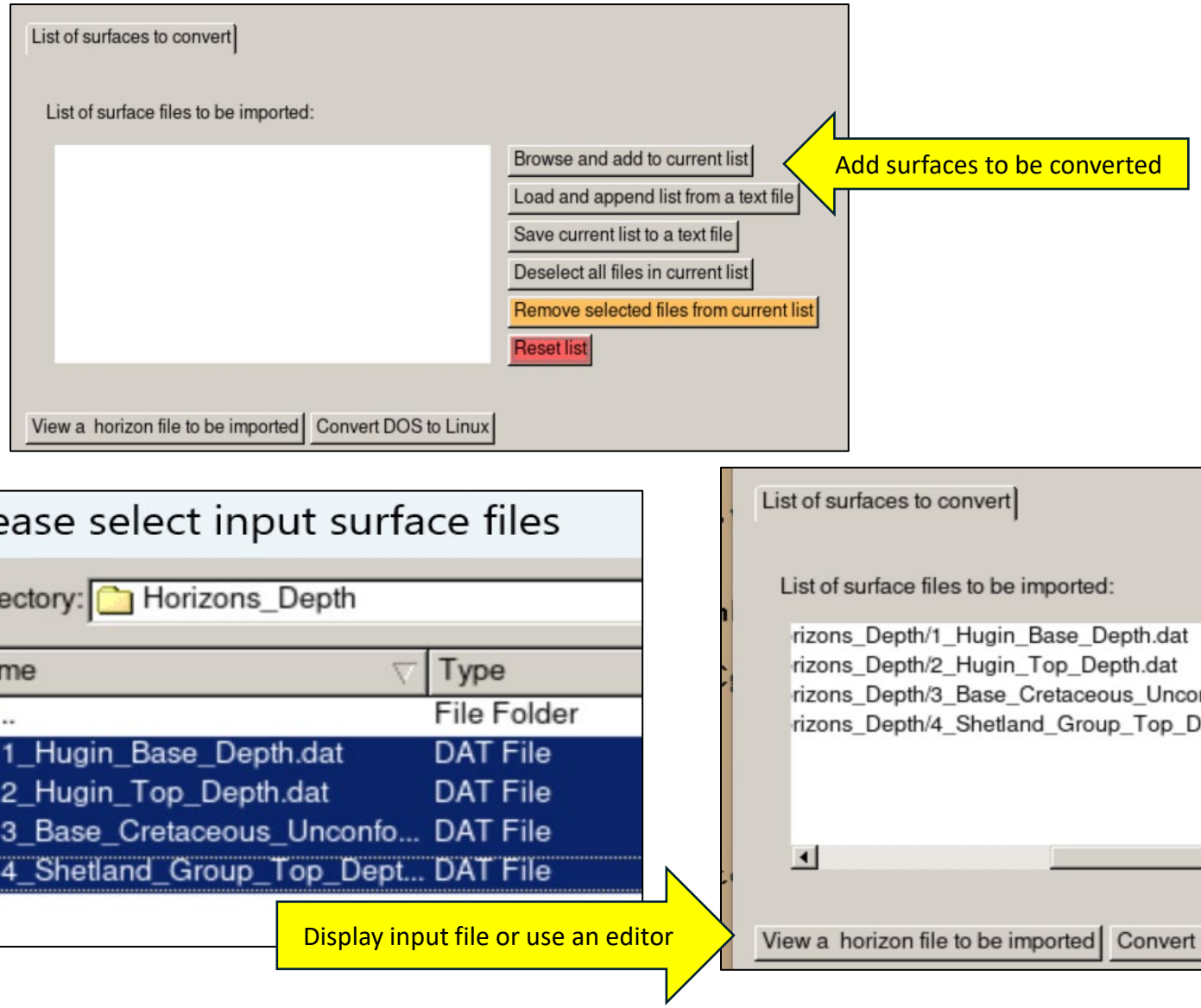
Step 1b: Repeat the process and convert workstation depth surfaces:

- *First, read in the reference seismic data volume and define vertical units*

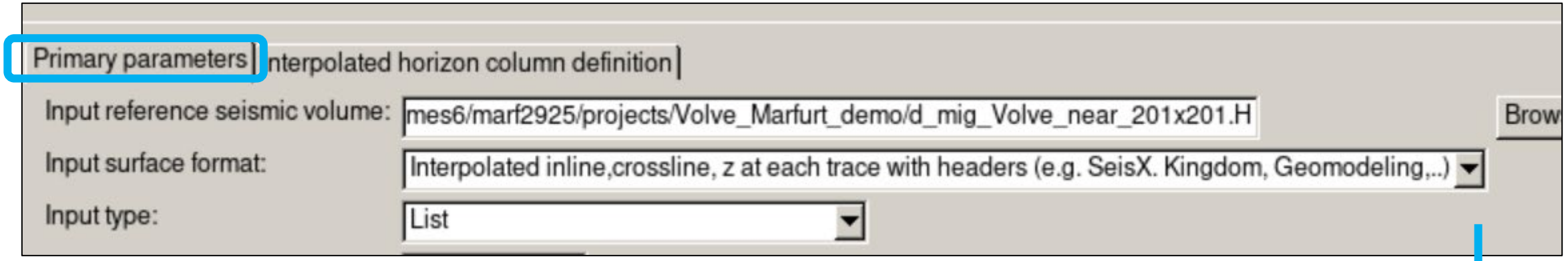




- *Second, select the workstation **DEPTH** horizons*

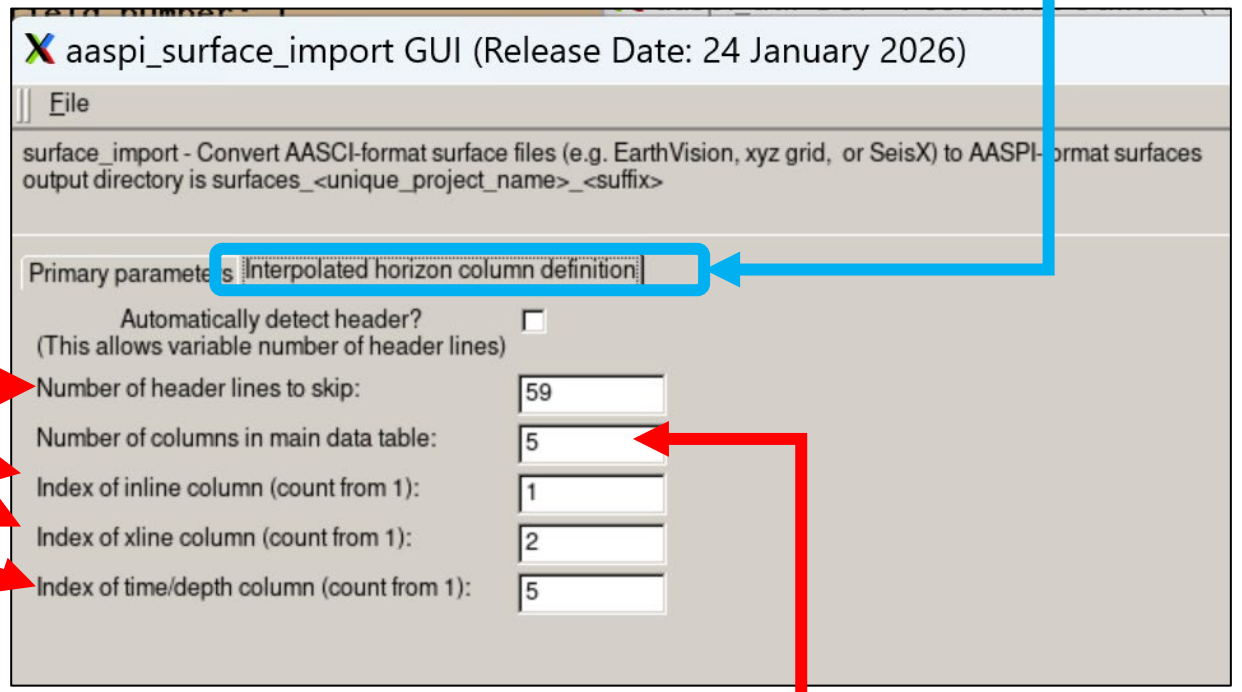


- *Third, determine the workstation-format of the **DEPTH** horizons*



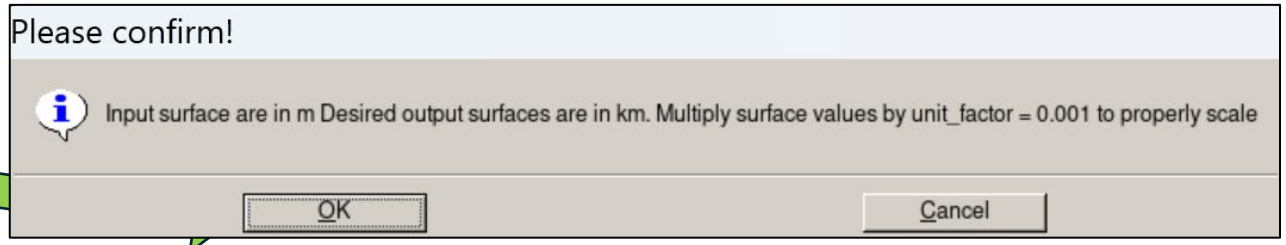
Input file
1_Hugin_Base_Depth.dat

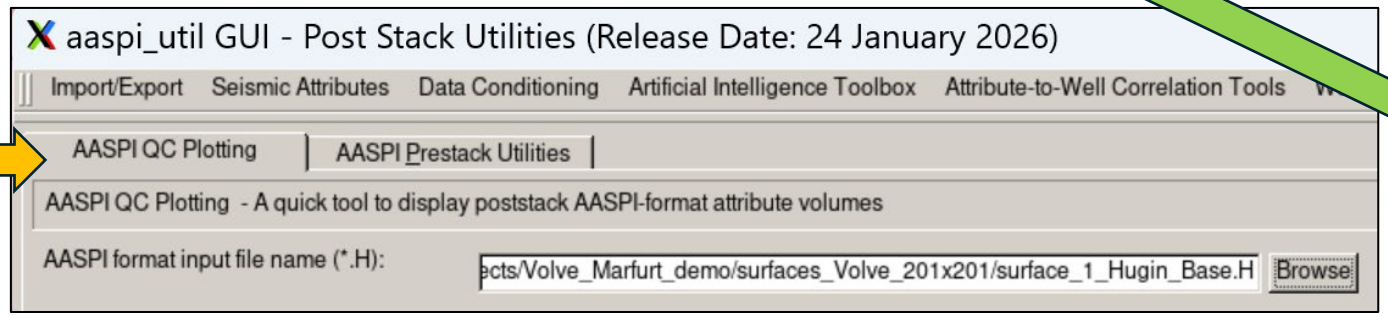
Input file 1 format="SeisX"

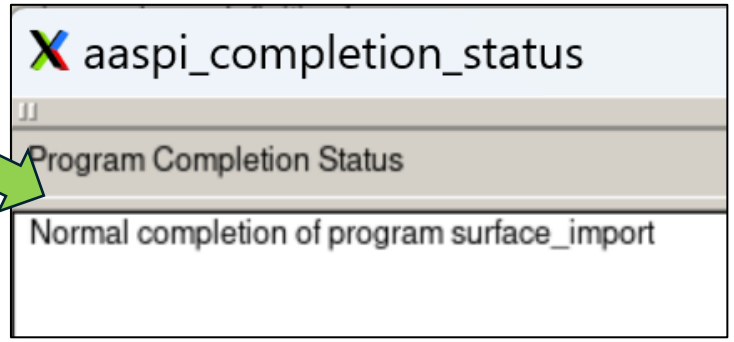


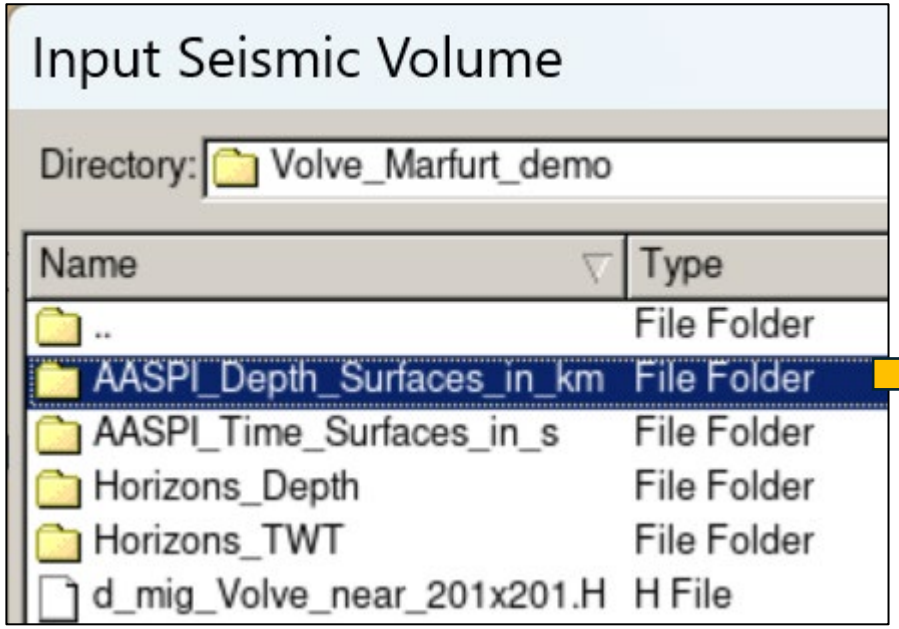

```
42 #
43 #     Item:surveyName              line number: 3 delimiter field number: 1
44 #     Item:horizonName             line number: 4 delimiter field number: 1
45 #     Item:dataSource              line number: 5 delimiter field number: 1
46 #     Item:horizonAttribute        line number: 6 delimiter field number: 1
47 #     Item:remark                  line number: 7 delimiter field number: 1
48 #     Item:line                    line number: 8 delimiter field number: 1
49 #     Item:trace                   line number: 8 delimiter field number: 2
50 #     Item:x                       line number: 8 delimiter field number: 3
51 #     Item:y                       line number: 8 delimiter field number: 4
52 #     Item:z                       line number: 8 delimiter field number: 5
53 #     Item:surveyType               constant:3D Survey
54 #
55 ST10010
56 Hugin_Fm_Base+ST10010ZC11_Near_190314_adj_2990_EasyDC
57 STAT
58 DEPTH
59 ,
60 10031.0000000,2477.0000000,432256.0248349,6477720.4242047,3172.8281250
61 10032.0000000,2477.0000000,432259.0492085,6477732.5528128,3175.6452637
62 10033.0000000,2477.0000000,432262.0735821,6477744.6814209,3177.7917480
63 10034.0000000,2477.0000000,432265.0979557,6477756.8100290,3180.6823730
```

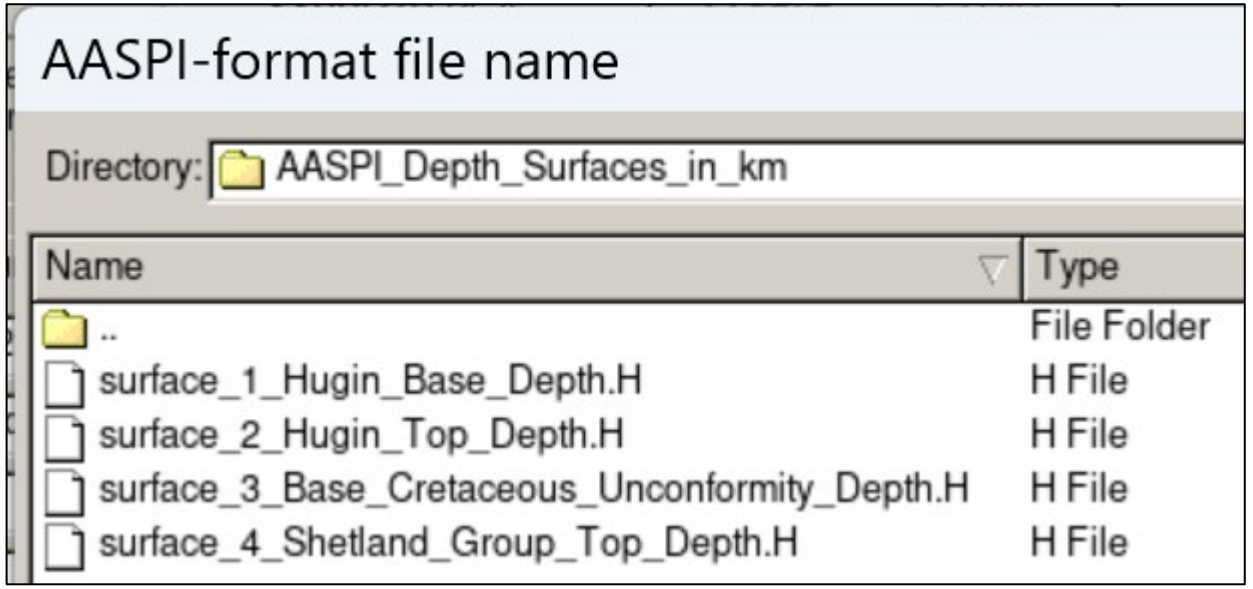
- *Fourth, execute the program and plot the results for quality control*

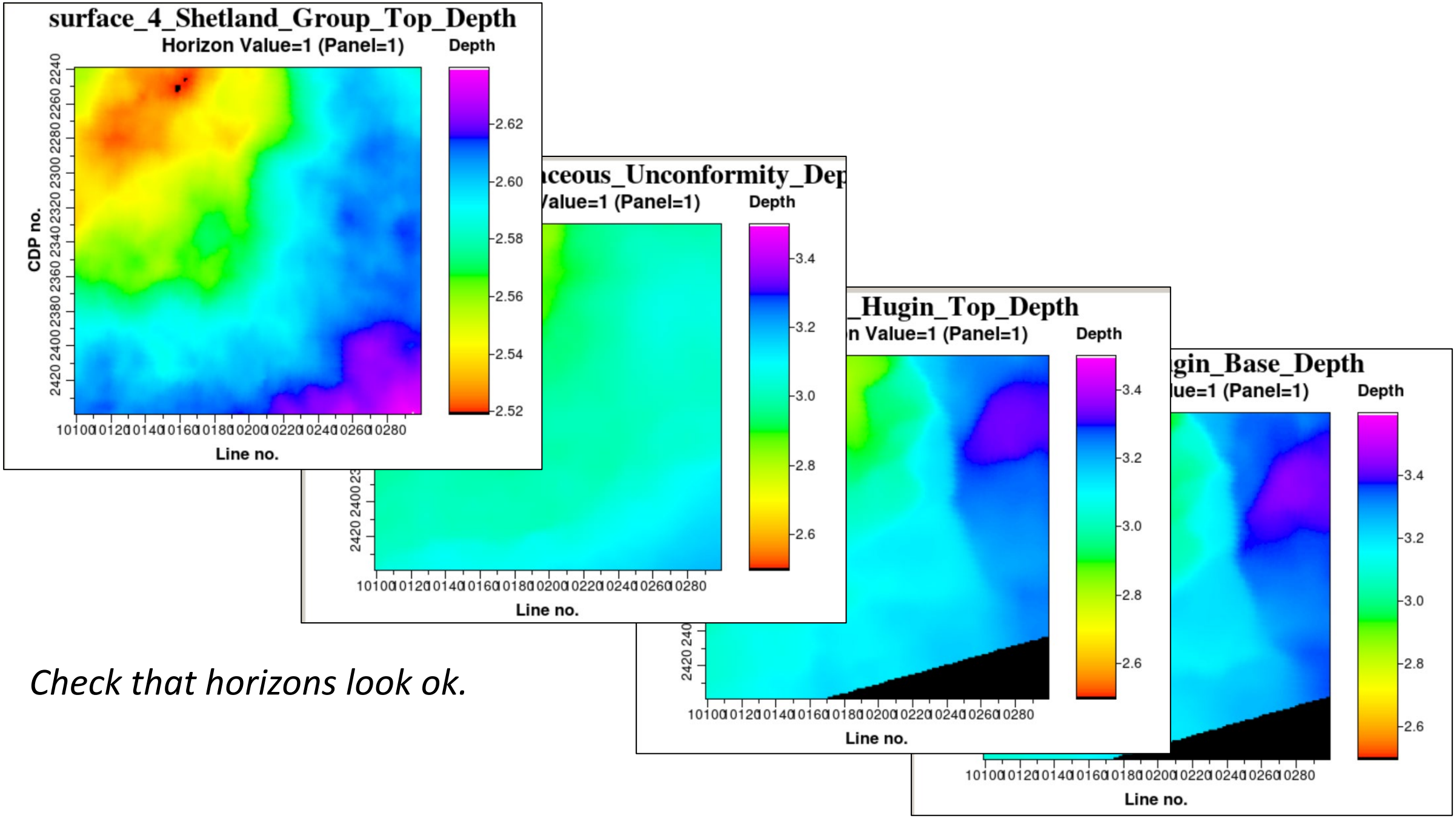






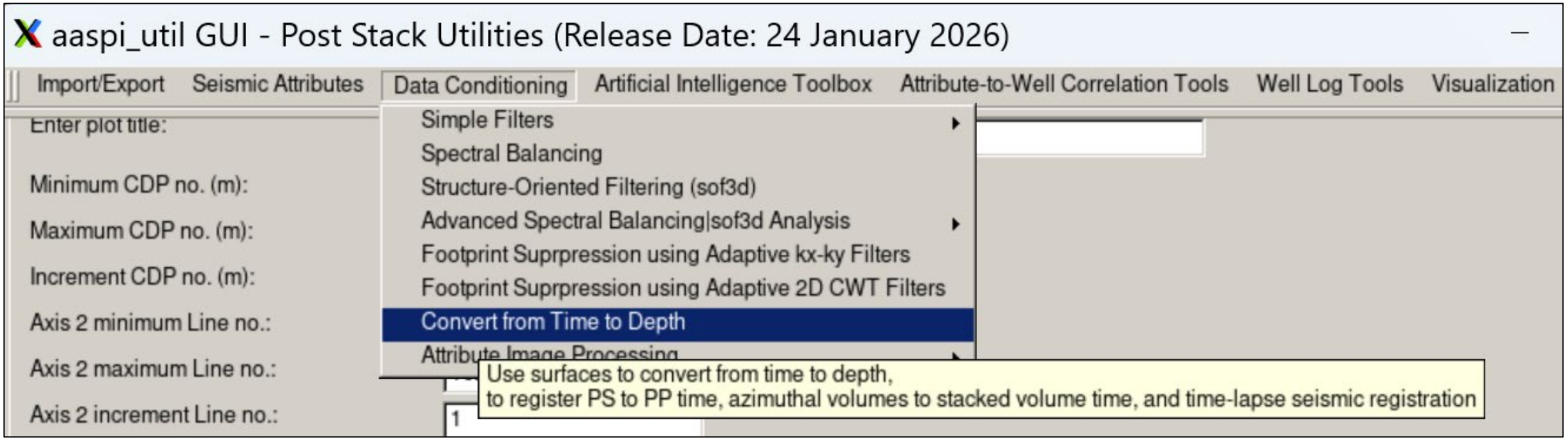




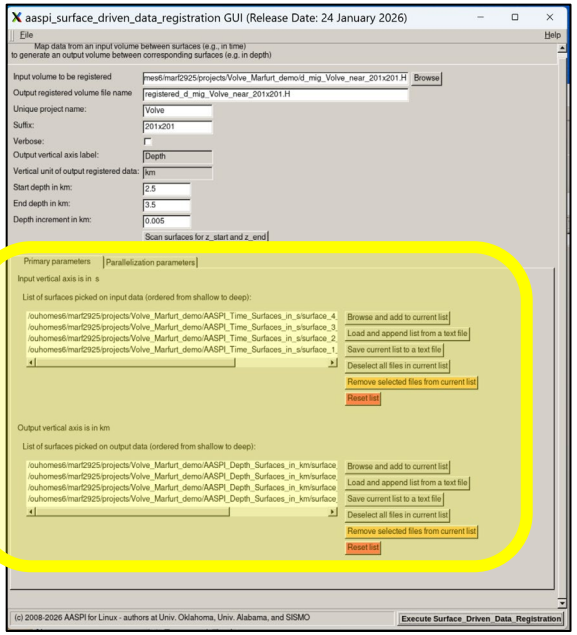


- *Check that horizons look ok.*

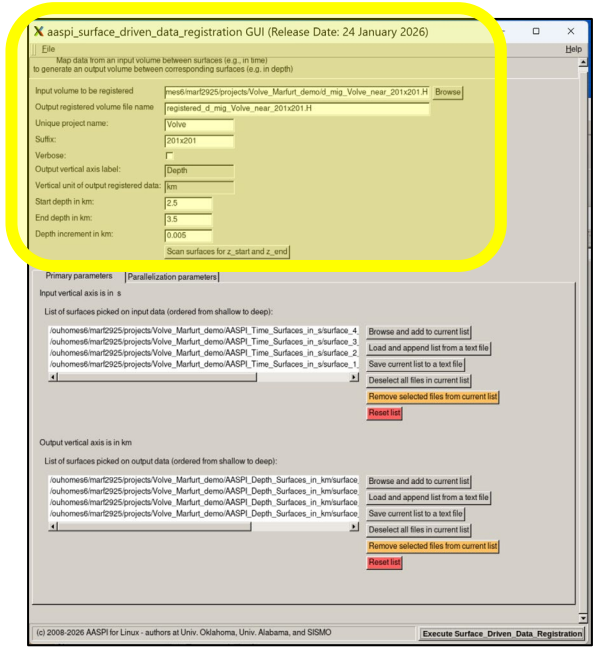
Step 2: Convert seismic volume from time to depth

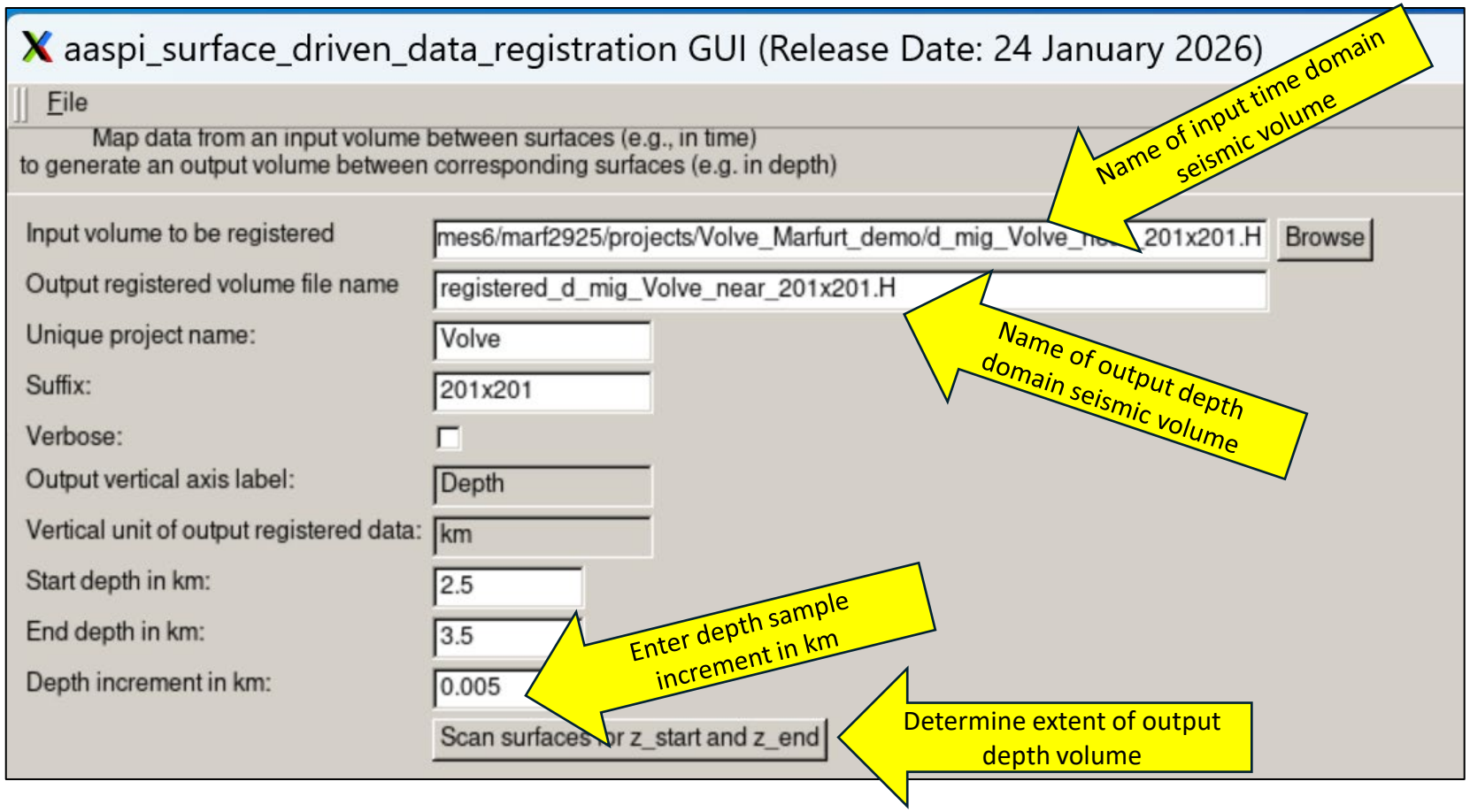


Enter the corresponding time and depth horizons IN ORDER from bottom to top



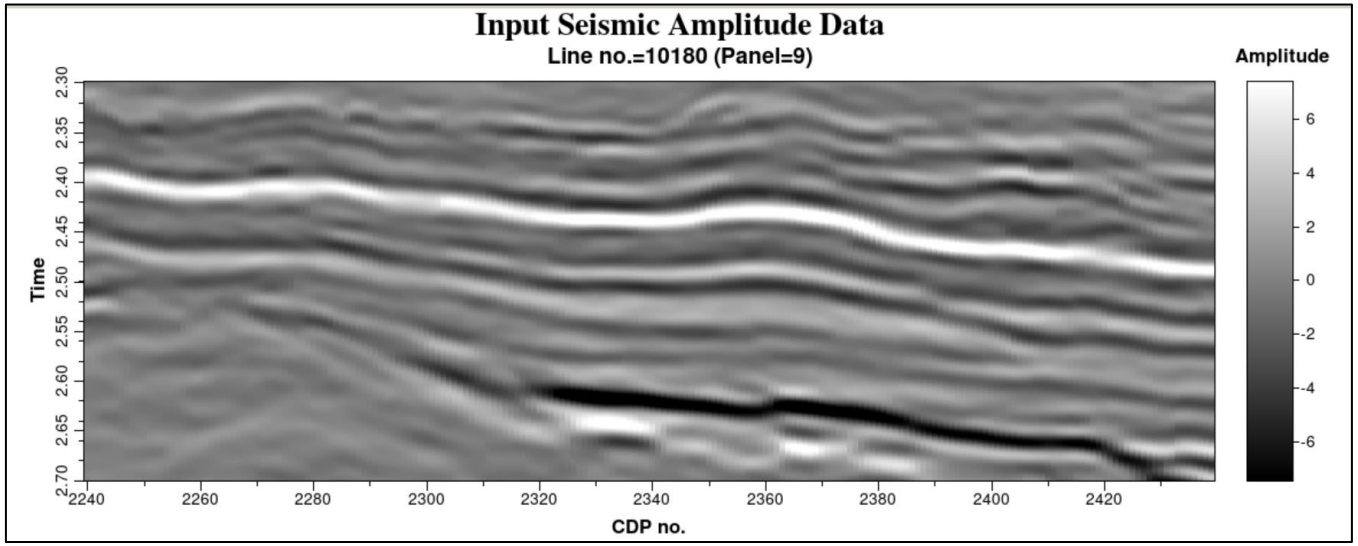
Enter the name of the time domain seismic volume
and the name and vertical range of the output depth domain volume





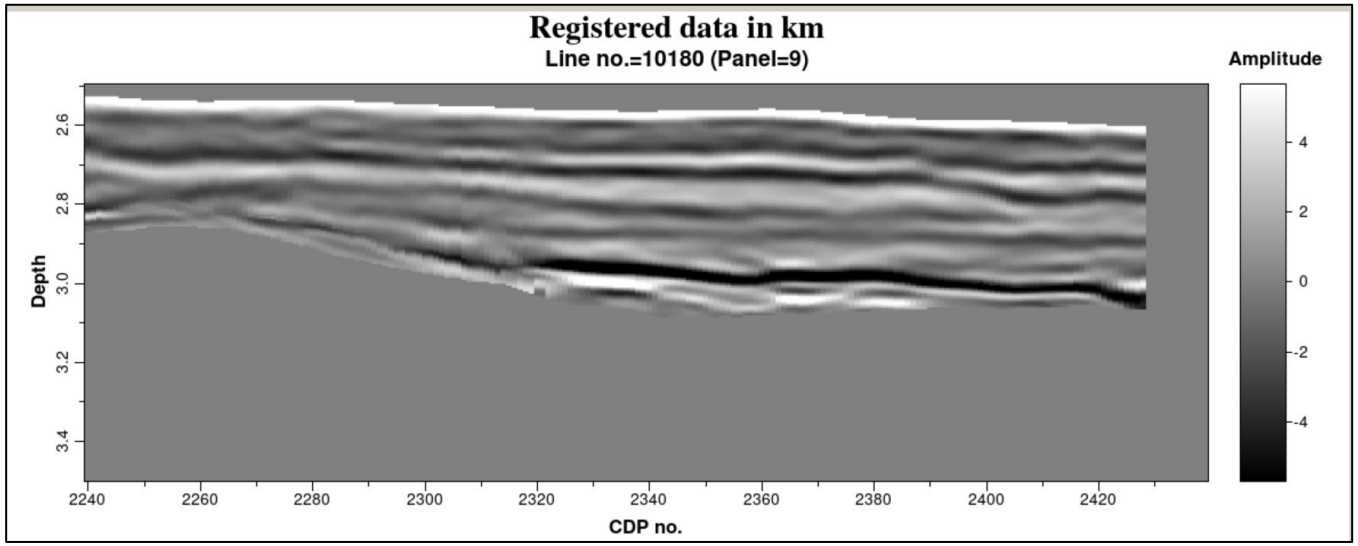
Execute the program

AASPI format input file name (*.H): mes6/marf2925/projects/Volve_Marfurt_demo/d_mig_Volve_near_201x201.H Browse

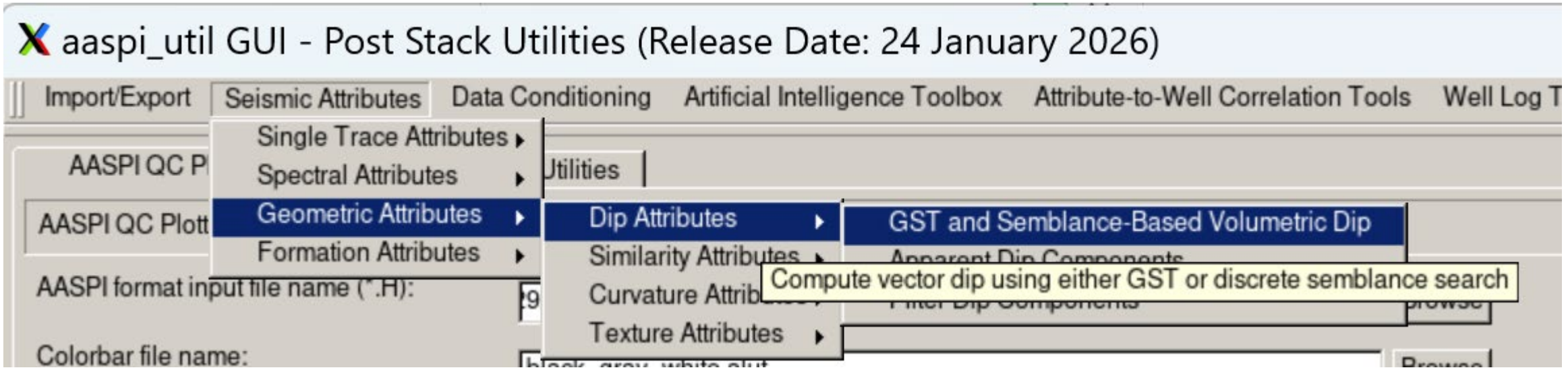


AASPI format input file name (*.H): 925/projects/Volve_Marfurt_demo/registered_d_mig_Volve_near_201x201.H Browse

Plot the results!

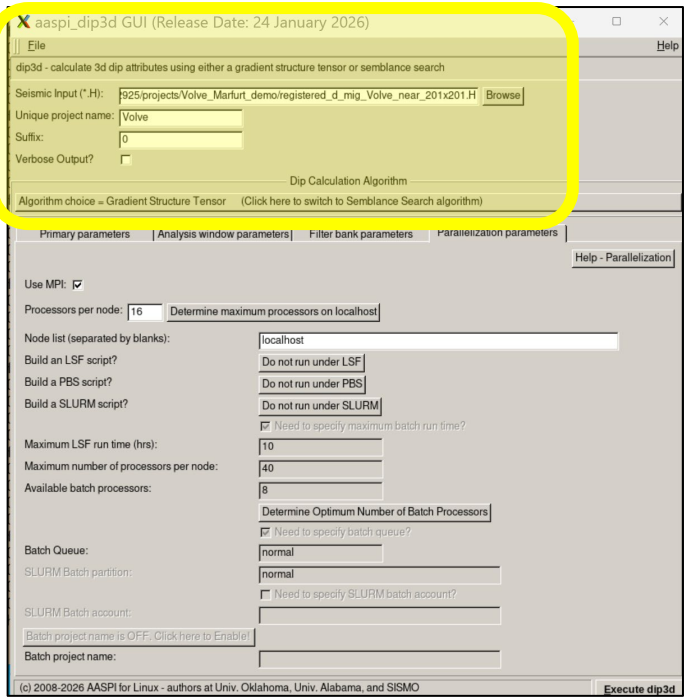


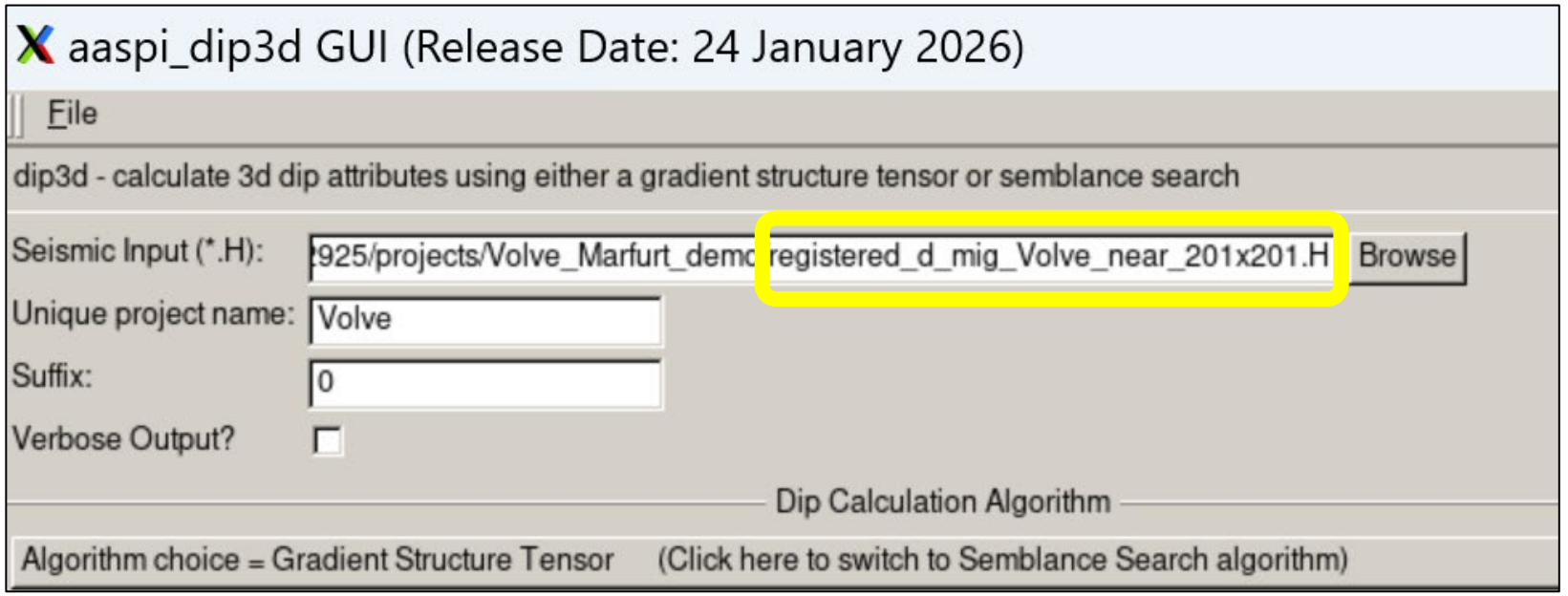
Step 3: Compute structural dip



Step 3: Compute structural dip

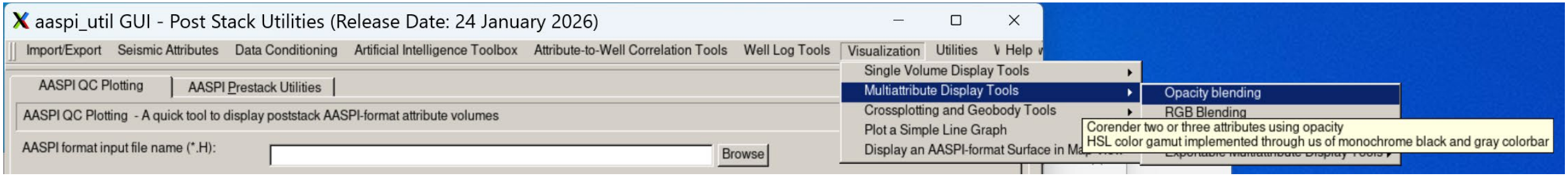
Defaults are fine.
But choose the registered DEPTH data, not the original TIME volume as input!

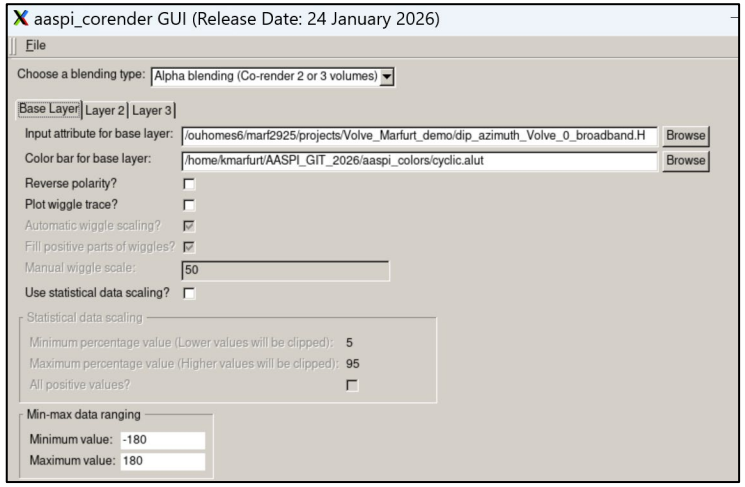


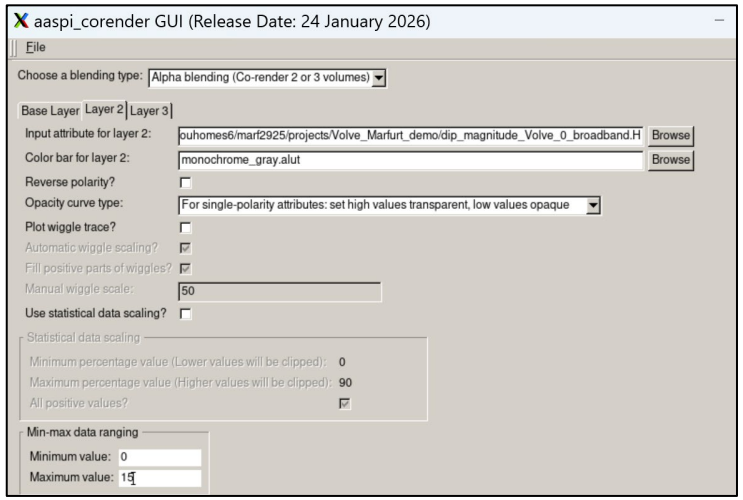


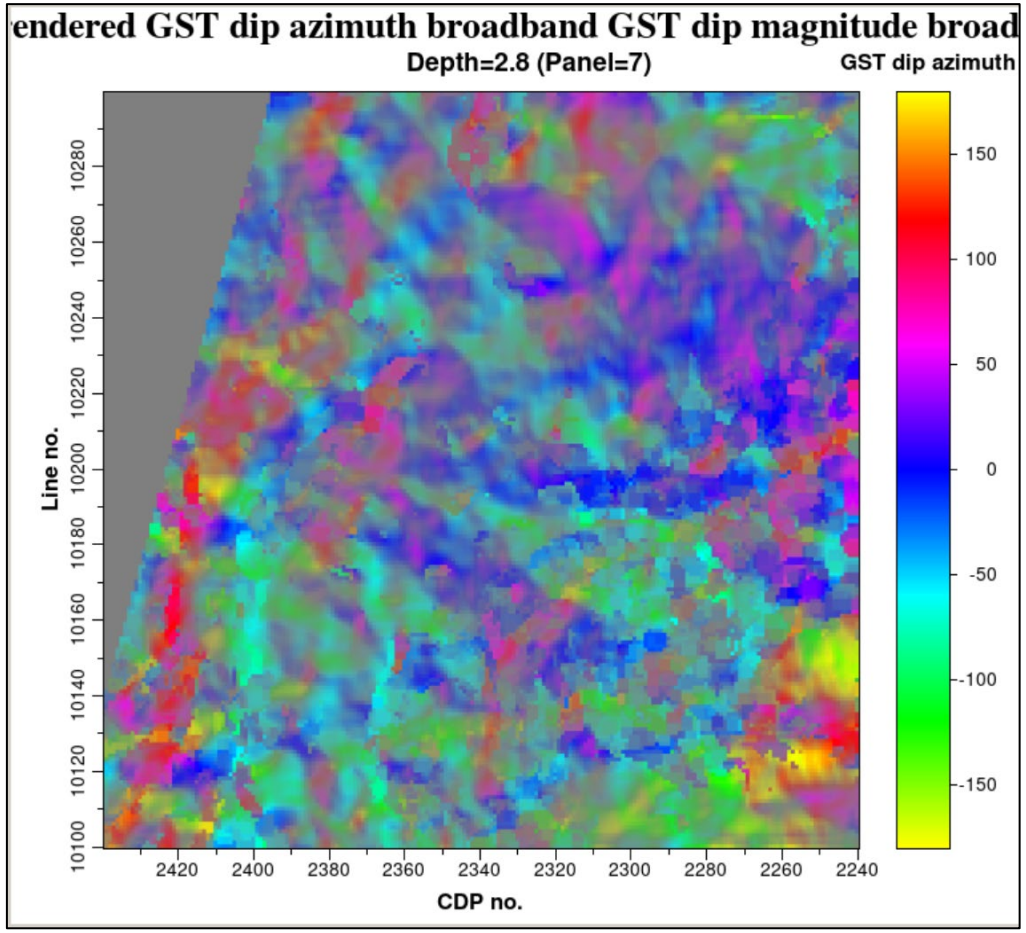
Then execute the program:

Plot the results using corendering

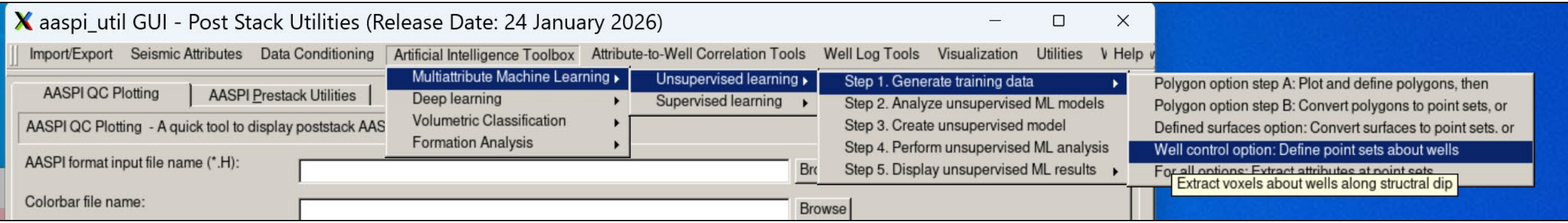




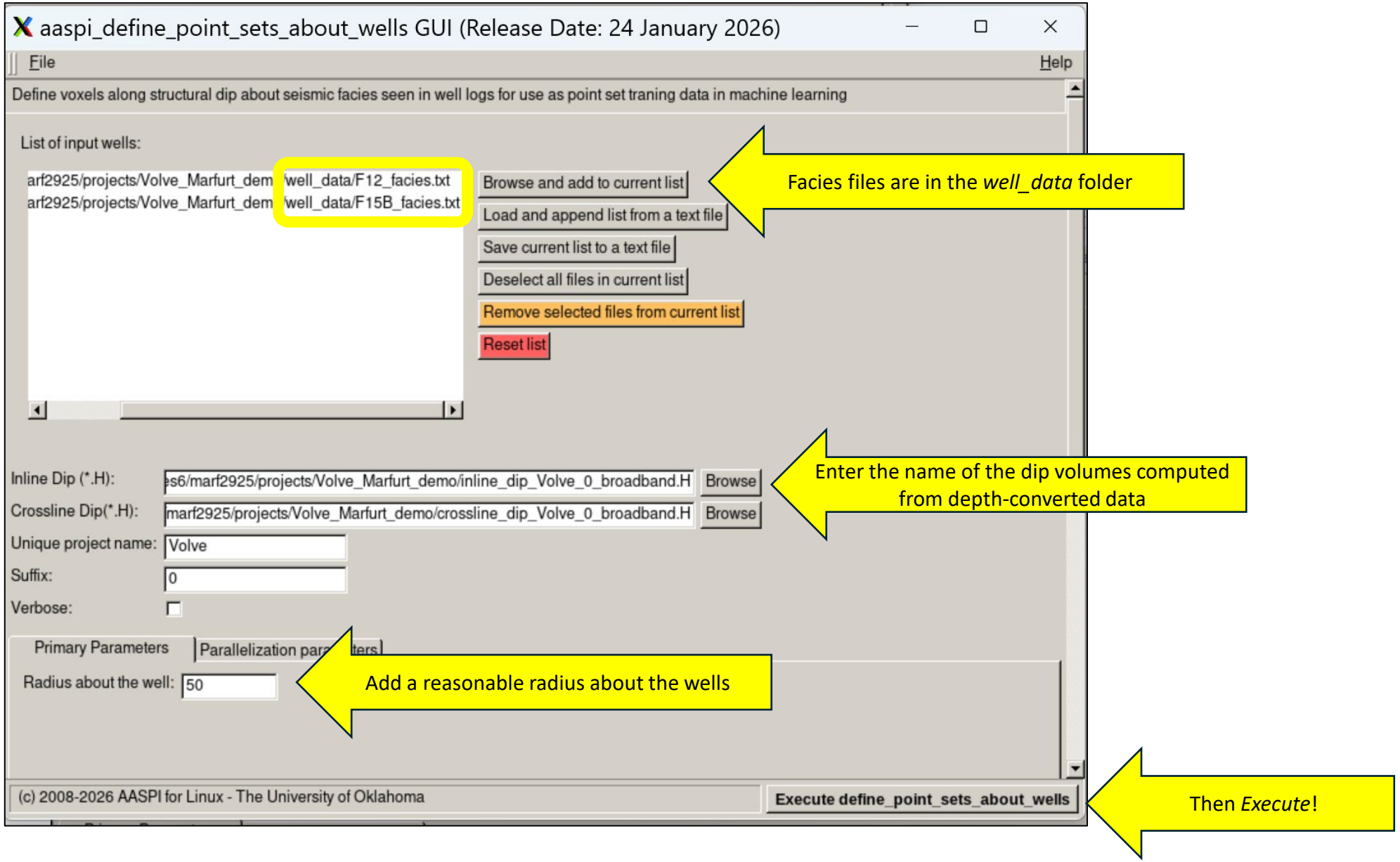




Step 4: Define the point sets about the wells



Step 4: Define the point sets about the wells



Step 4: Define the point sets about the wells

Facies file for well F12 (this might change)

Unconsolidated Quaternary deposits			
145.9	145.9	6478566.22	435050.21
893	888.13	6478596.47	435026.94
Unconsolidated to weakly consolidated sand			
903	897.8	6478596.18	435024.41
1213.00	1196.00	6478.581.86	434942.55
Silty sandstone			
1223.00	1205.39	6478580.62	434939.33
1323.00	1300.14	6478570.58	434909.07
Claystone to shale			
1333.00	1309.81	6478569.71	434906.67
2223.00	2144.62	6478461.98	434628.80
Volcaniclastic tuff and claystone			
2233.00	2154.50	6478460.59	434628.08
2283.00	2203.87	6478452.92	434626.44
Claystone			
2293.00	2213.73	6478451.23	434626.53
2610.92	2523.59	6478384.45	434689.73
Chalky limestone			
2658.30	2523.59	6478384.45	434689.73
2860.02	2715.05	6478321.65	434820.56
Sandstone Oil			
2874.45	2740.77	6478312.76	434850.43
3008.61	2804.89	6478292.97	434935.21
Sandstone Wet			
3060.63	2834.92	6478284.38	434976.79
3520.00	3108.36	6478201.68	435335.43

MD *TVD* *Northing* *Easting*

Output files containing the point sets for all wells

```
Jan 27 18:26 facies_Unconsolidated_Quaternary_deposits.dat
Jan 27 18:26 facies_Unconsolidated_to_weakly_consolidated_sand.dat
Jan 27 18:26 facies_Silty_sandstone.dat
Jan 27 18:26 facies_Claystone_to_shale.dat
Jan 27 18:26 facies_Volcaniclastic_tuff_and_claystone.dat
Jan 27 18:26 facies_Claystone.dat
Jan 27 18:26 facies_Chalky_limestone.dat
Jan 27 18:26 facies_Sandstone_Oil.dat
Jan 27 18:26 facies_Sandstone_Wet.dat
```

Step 4: Define the point sets about the wells

Top part of the "point sets" for the Sandstone_Oil well facies

```
                                                          Sandstone_Oil
Depth
CDP no.
Line no.
 j_well =           1
       2.830       2285.      10124.
       2.840       2285.      10124.
       2.850       2285.      10124.
       2.800       2286.      10124.
       2.805       2286.      10124.
       2.810       2286.      10124.
       2.815       2286.      10124.
       2.820       2286.      10124.
       2.830       2286.      10124.
       2.795       2287.      10124.
       2.805       2287.      10124.
       2.815       2287.      10124.
       2.825       2287.      10124.
       2.800       2288.      10124.
       2.810       2288.      10124.
       2.820       2288.      10124.
       2.810       2289.      10124.
       2.820       2289.      10124.
       2.805       2290.      10124.
       2.815       2290.      10124.
       2.810       2291.      10124.
       2.820       2291.      10124.
       2.755       2284.      10125.
```

Next steps:

- *Determine the simplest way to convert alternative workstation facies logs as input for define_point_sets_about_wells*
- *Rework generative topographic mapping (GTM) and Bhattacharya distance to provide semi-supervised classification*
- *Repeat the process for self-organizing mapping (SOM)*