Soil Heterogeneity Statistics Overview

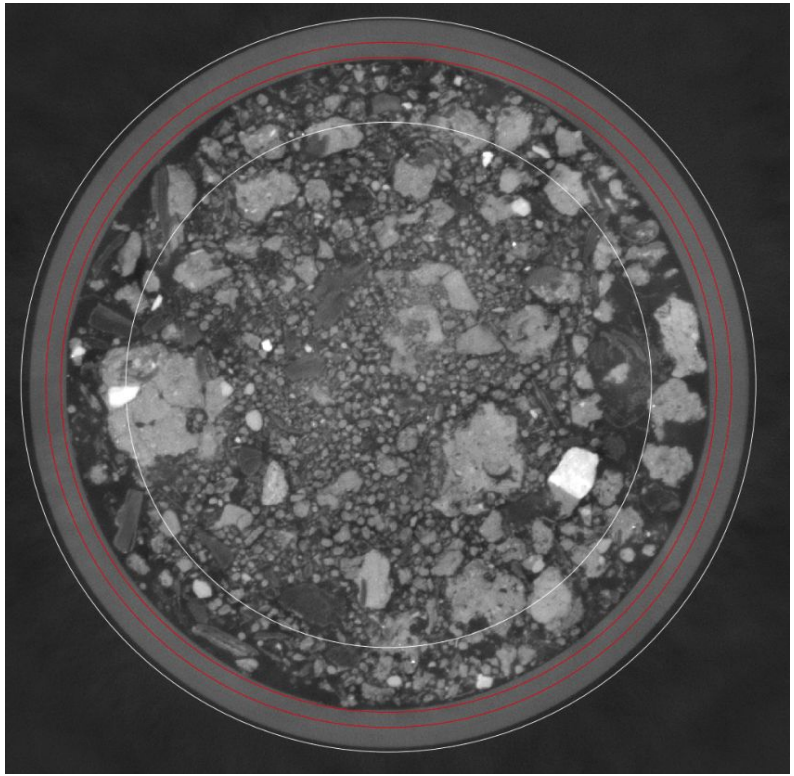
8/7/23

Overview

- Goals
  - Explore local soil density
  - Identify metrics that capture heterogeneity of local soil density

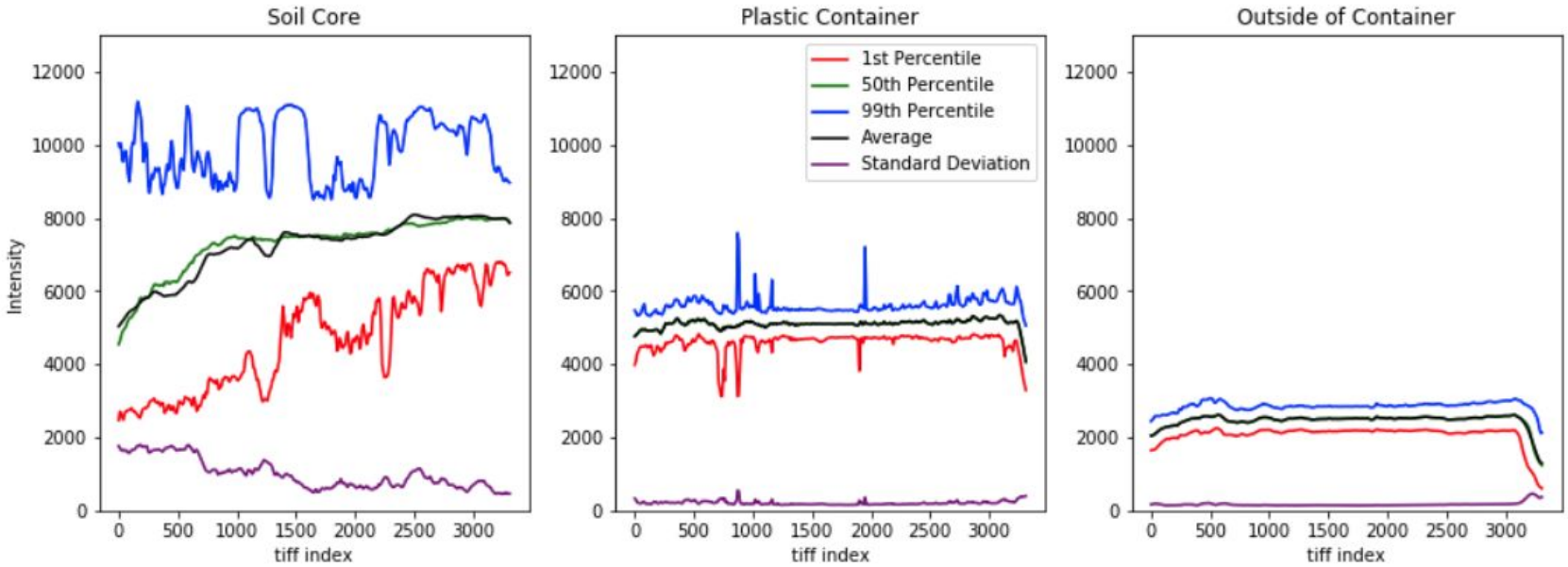
Local Soil Density

- Gather some basic statistics by depth (horizontal xray slices)
- Split into 3 different sections
  - Inner soil core
    - Ignore the edges with the container to avoid boundary effects
  - Plastic ring
    - This should be uniform density throughout
  - Outside the container
    - Should be a density of zero?



Local Soil Density

scan 14, no till, high manure

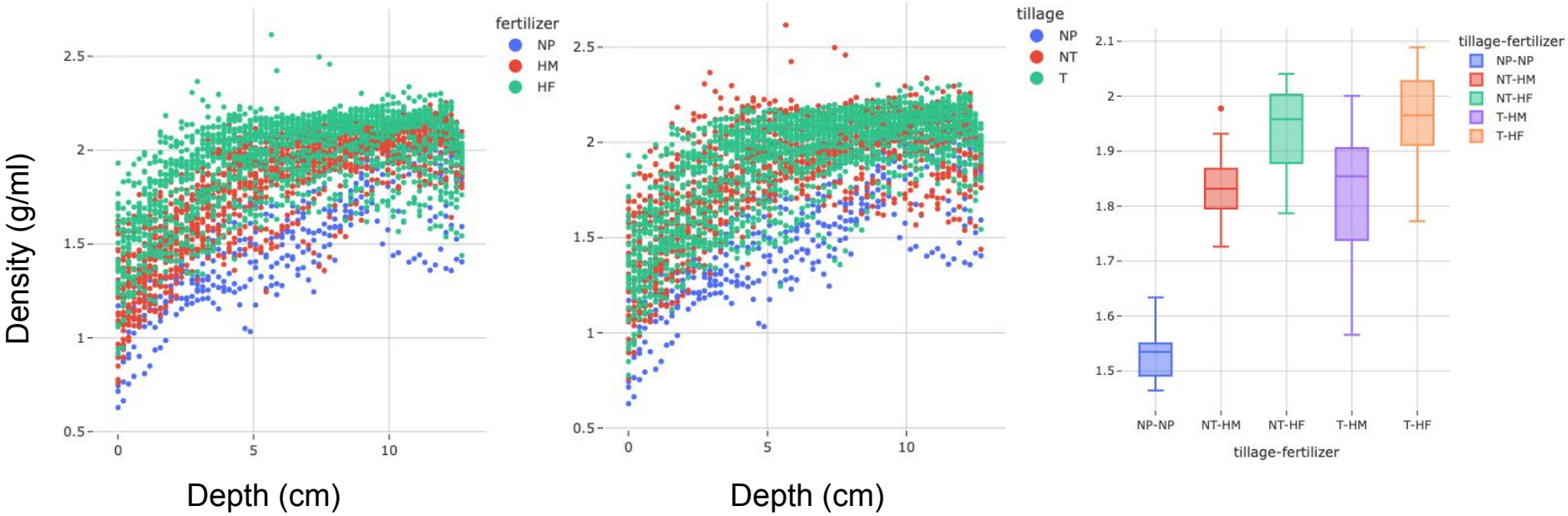


To approximate density:

density = $(PI_{soil}-PI_{outside\_of\_container})/(PI_{container}-PI_{outside\_of\_container})*1.022$, units: g/ml

Conversion from displacement measurement of PETG container: 70mL volume displacement, 71.55g

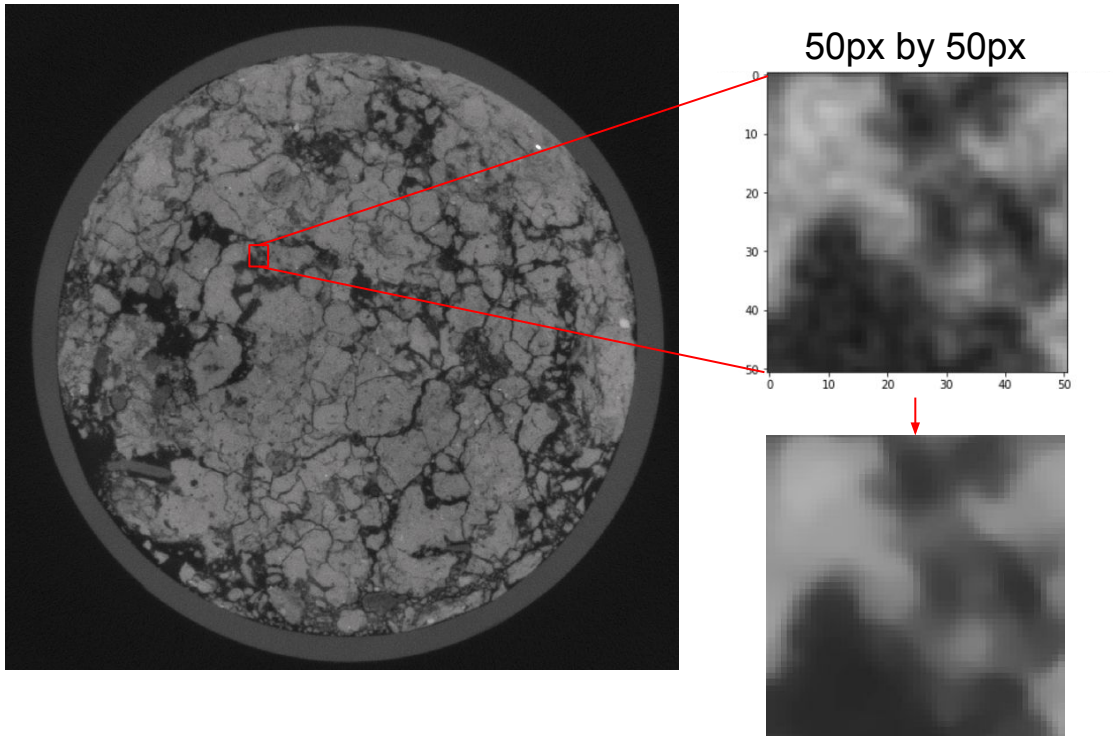
Local Soil Density



Greater difference in density by fertilizer treatment than tillage.

Bulk density measurements for these scanned soils?

Characterizing Heterogeneity



Denoised with total variation technique

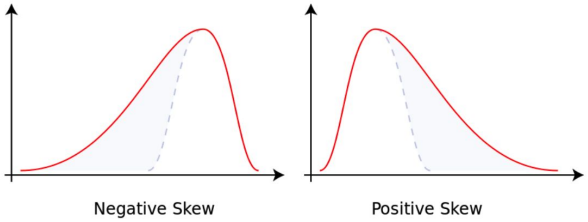
What can **local density statistics** tell us about the soil's heterogeneity and differences between soil treatments/management practices?

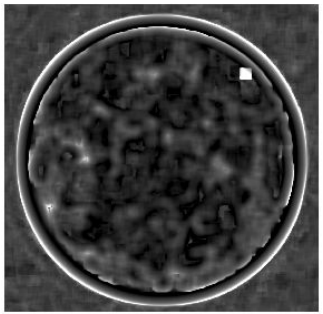
Characterizing Heterogeneity

Focusing on metrics robust against x-ray related noise

Skewness

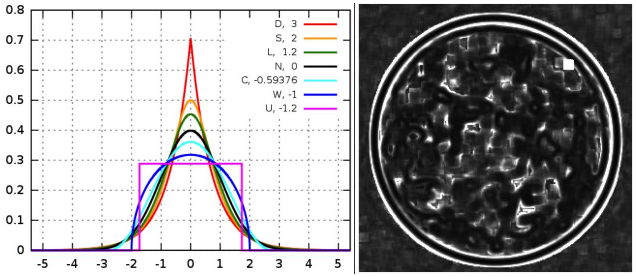
Asymmetry in pixel density distribution.





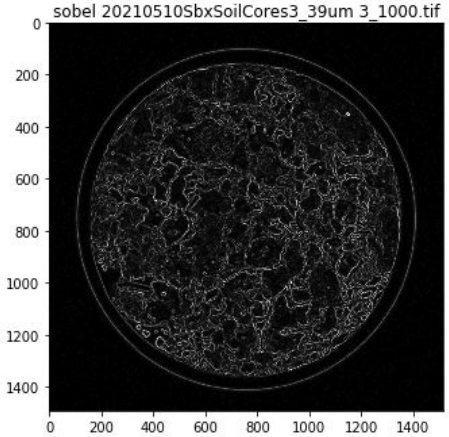
Kurtosis

“peakedness” or “flatness” of distribution compared to the normal distribution, specifically, how much of the distribution is in the tails.



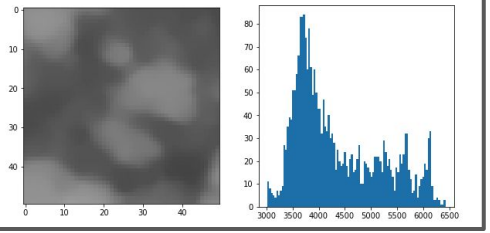
Sobel Edges

Edge finding image convolution that calculates the vertical and horizontal gradients in an image.

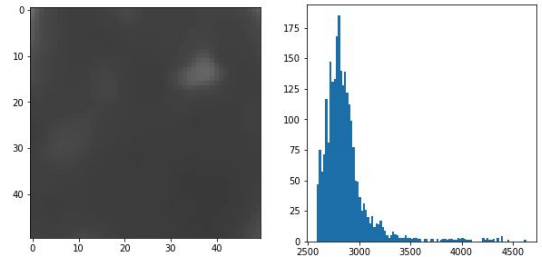


Skewness vs Kurtosis

kurtosis: -0.7035735290584433, skew: 0.7068692420944443

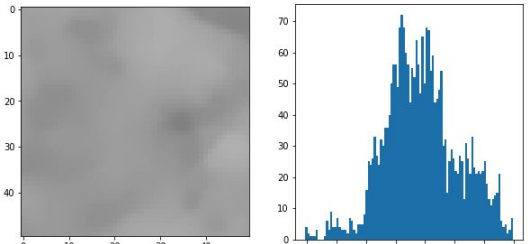


kurtosis: 13.761542936508711, skew: 3.139147400978889

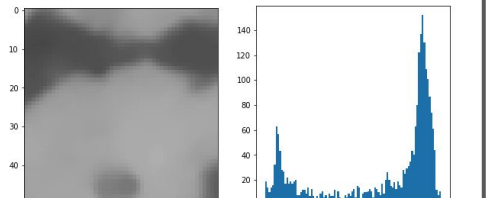


Kurtosis

kurtosis: -0.05701518338358991, skew: 0.07986803020788472

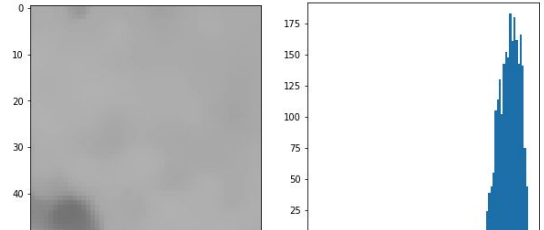


kurtosis: -0.8792197836069437, skew: -0.8676381645358686

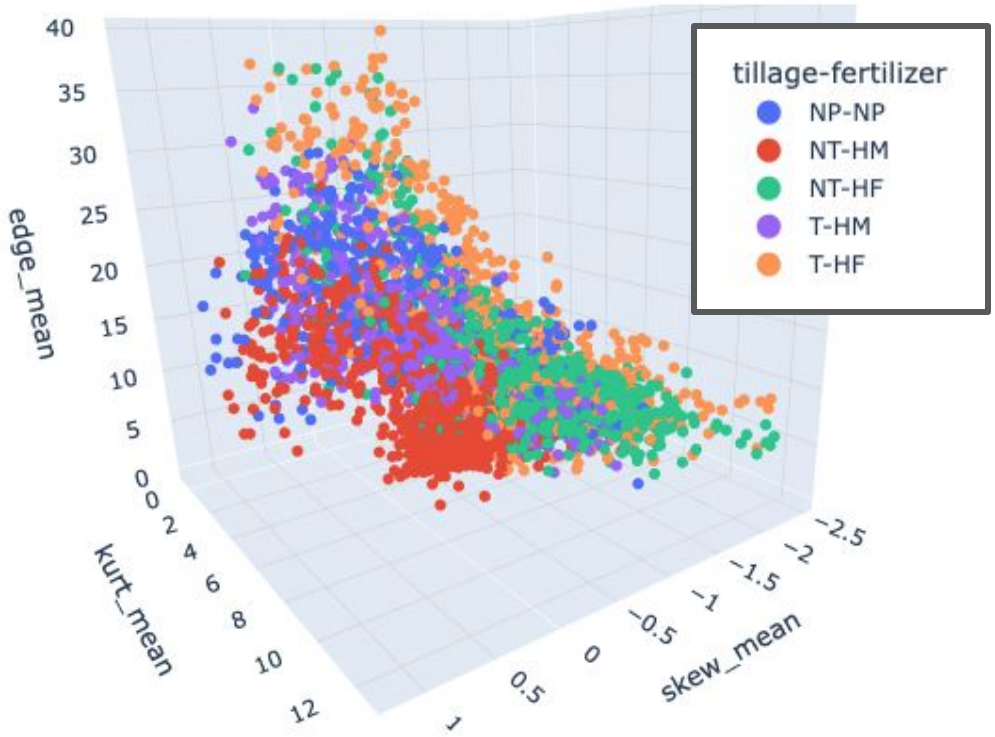


Skewness

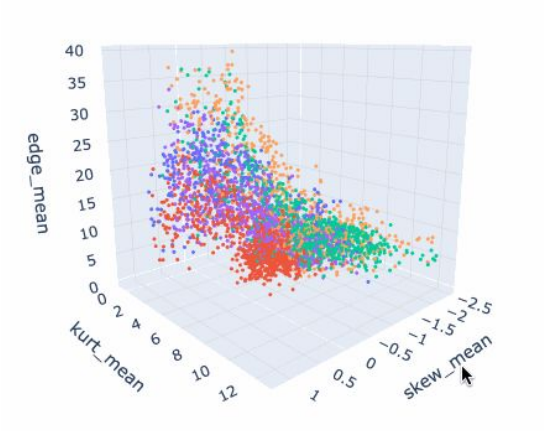
kurtosis: 15.198680629169704, skew: -3.76185317052183



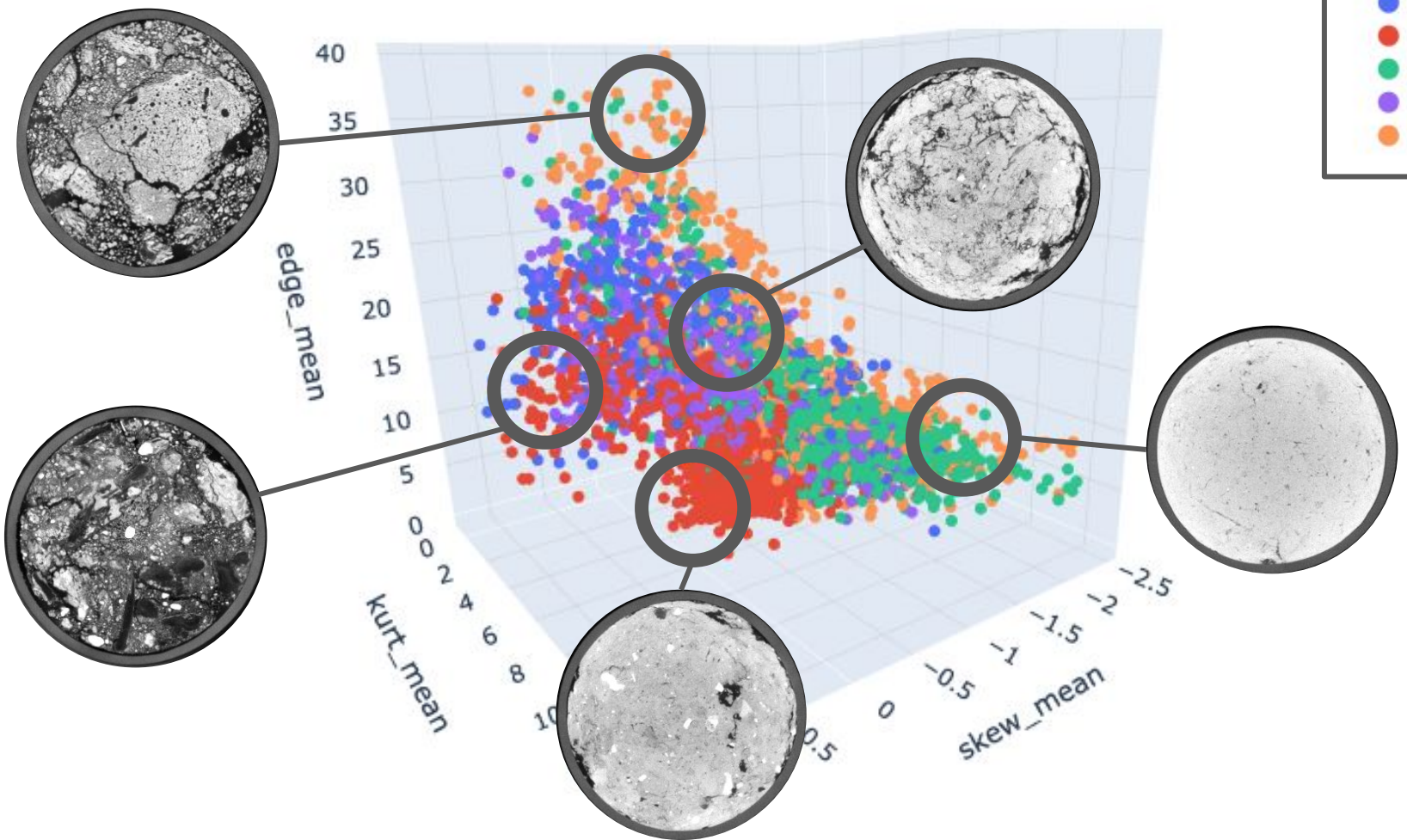
Skewness vs Kurtosis vs Sobel Edge Metric



- Each point is in an average from an **individual horizontal xray slice**
- Sampled **every 50** slices
- Data seems fall roughly on a curved surface

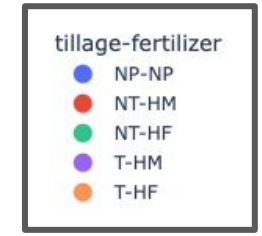


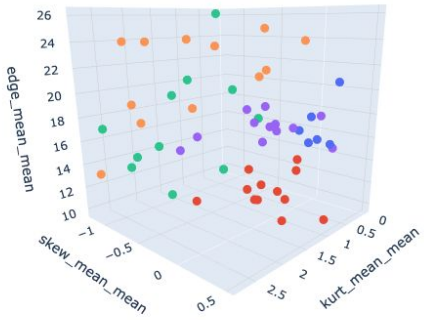
What do different regions look like?



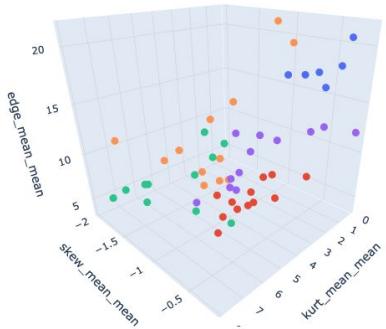
Averaging metrics by x-ray scan

12 total features for each scan: skew, kurt and sobel averages at 4 depth bins.

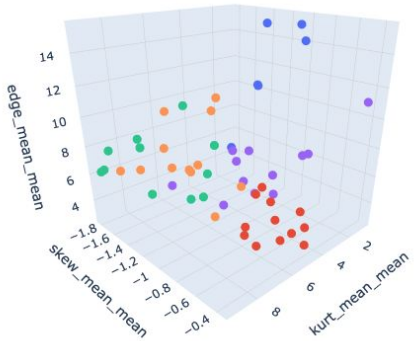




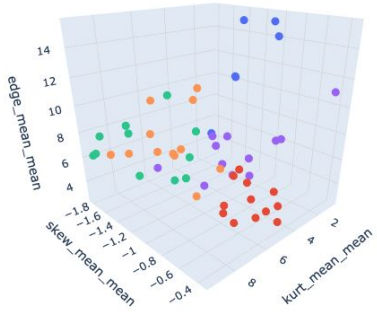
0 - 3.3 cm



3.3 - 6.6 cm

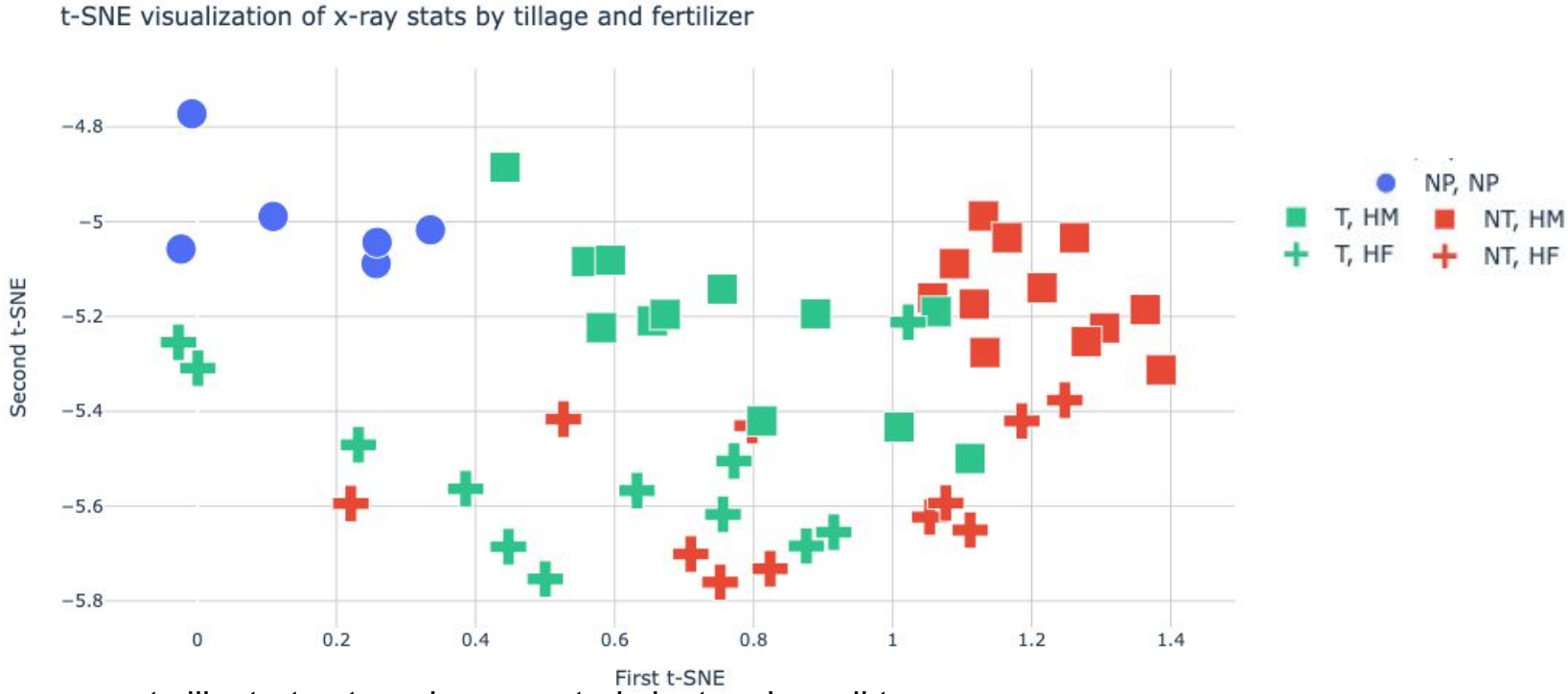


6.6 - 9.9 cm



9.9 - 12.9 cm

Reducing dimensionality to extract qualitative insights



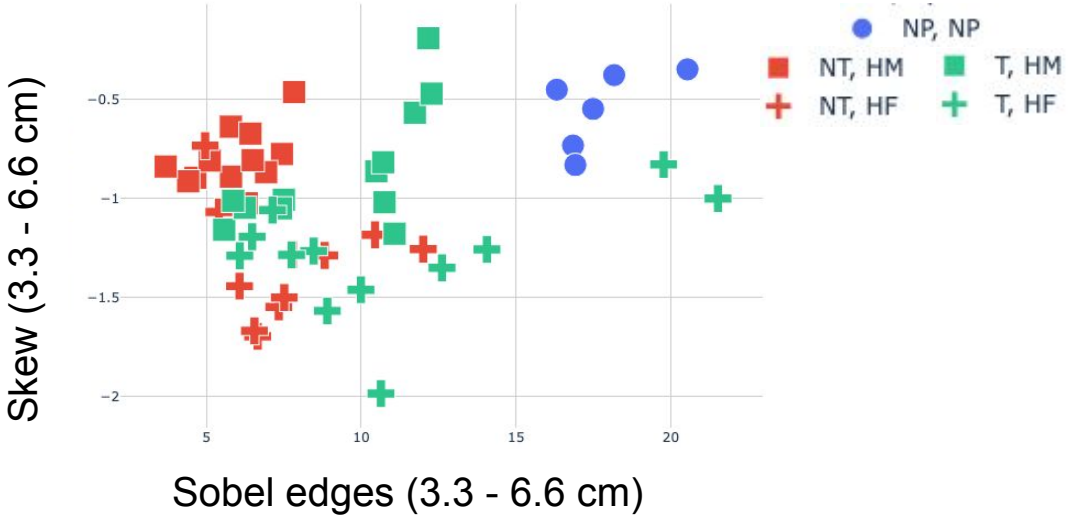
t-SNE components illustrate strongly separated clusters by soil type.

ANOVA:

- Significant differences in t-SNE_1 by both fertilizer (P~1E-8), and tillage (P~1.8E-5)
- Significant differences in t-SNE_2 by fertilizer (P~1E-12)

Reducing dimensionality to extract qualitative insights

- t-SNE great for visualization, though no direct mapping between 12 original features and t-SNE components.
- But:
  - t-SNE_1 - strongly correlated with sobel edge features
  - t-SNE_2 - strongly correlated with skew and kurtosis



Similar separation found just with just mid-depth skew v edge metrics

Predicting soil type by depth-binned metrics

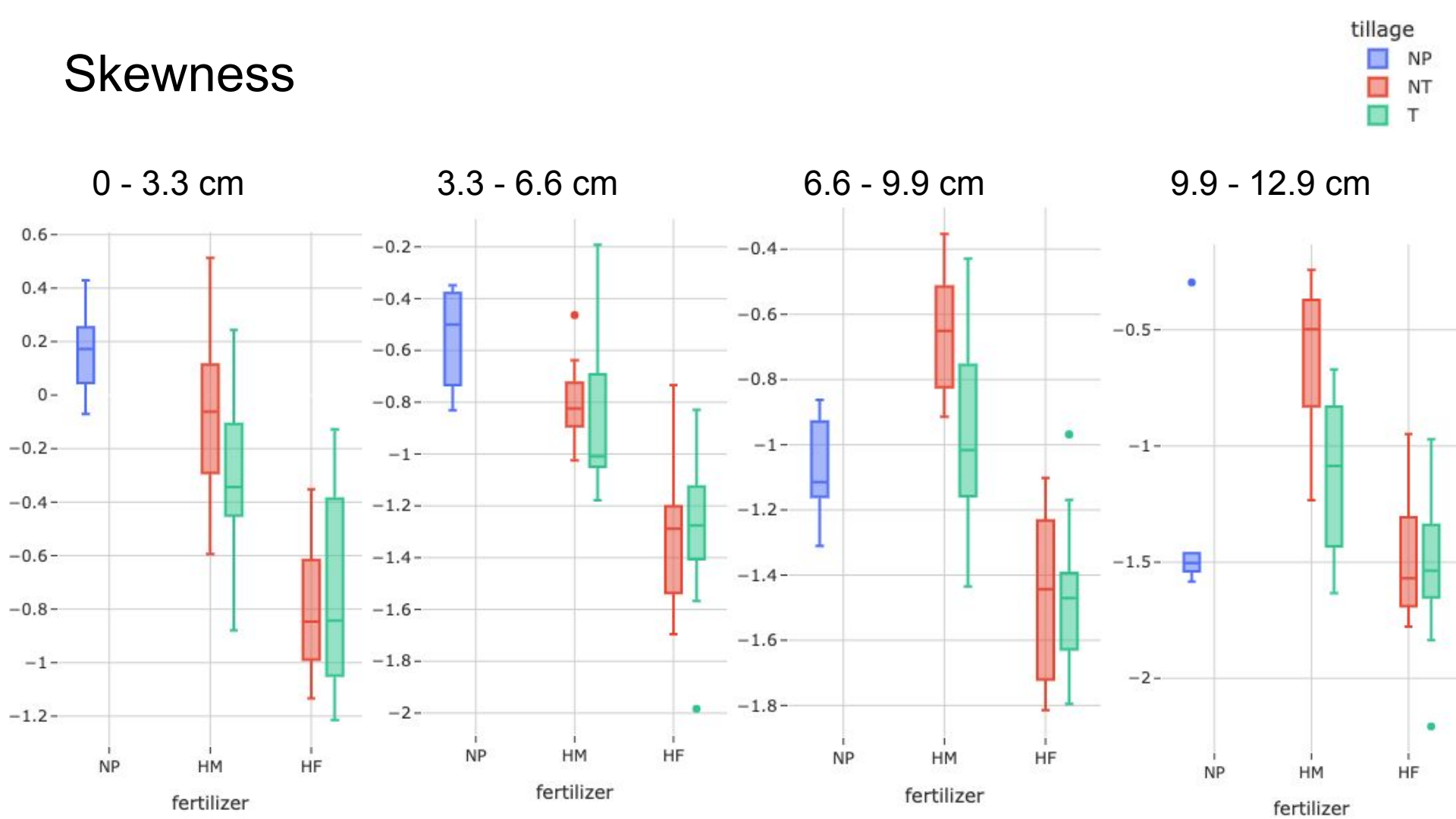
- I trained a support vector machine (SVM, a classification technique) to **predict soil type** from the 12 statistical features calculated for each scan:
  - To determine the accuracy of the SVM model, I used **cross-validation with 5 folds** (for each fold, the model was trained on 80% of the scans to predict the soil type of other 20% of the scans), and compared the accuracy of the predictions.
    - I used stratified group k-folds to preserve the distribution of each soil type in the folds.
  - Accuracy = total correct predictions / total predictions
  - Results:

Classification Type:	Tillage-Fertilizer	Tillage	Fertilizer
Fold 1 accuracy	0.73	0.91	1.00
Fold 2 accuracy	0.82	0.91	0.91
Fold 3 accuracy	0.73	0.91	0.73
Fold 4 accuracy	0.90	0.70	0.90
Fold 5 accuracy	0.90	0.90	0.90
Mean accuracy	**0.81**	**0.87**	**0.89**

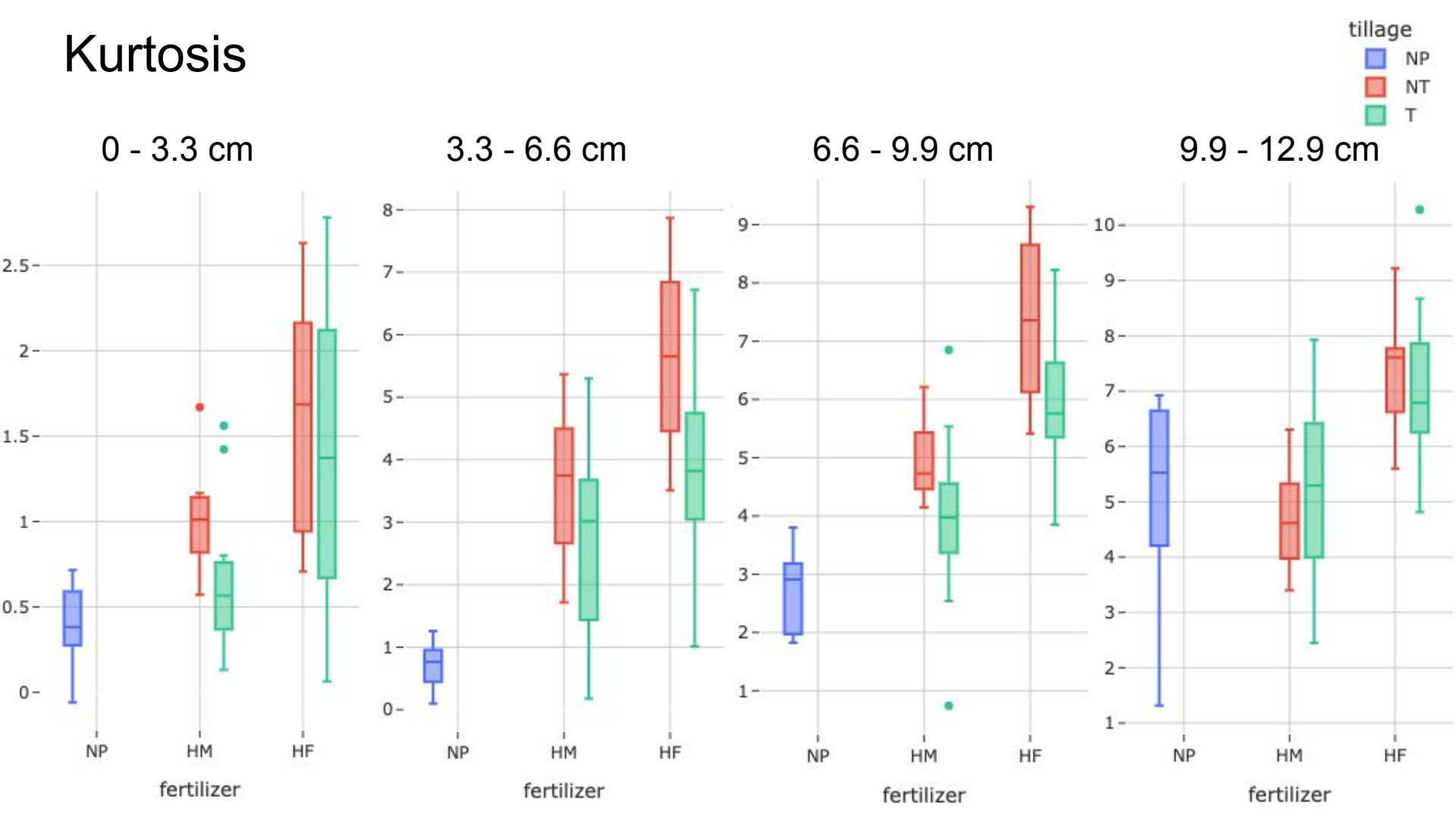
Conclusion

- Skew, Kurtosis, and Sobel edge seem to be good metrics to differentiate soils
- Changes in skew and kurtosis seem indicative of fertilizer vs manure.
- Changes in edges seems indicative of tillage differences
- Depth seems important, especially to differentiate tillage of high fertilizer soils
- Future directions:
    - How do these metrics connect with physical metrics like drainage, nutrient transport, etc? More experiments required.
    - How do these metrics connect to pore structure metrics?
    - How do bulk density measurements compare against x-ray calculated values?

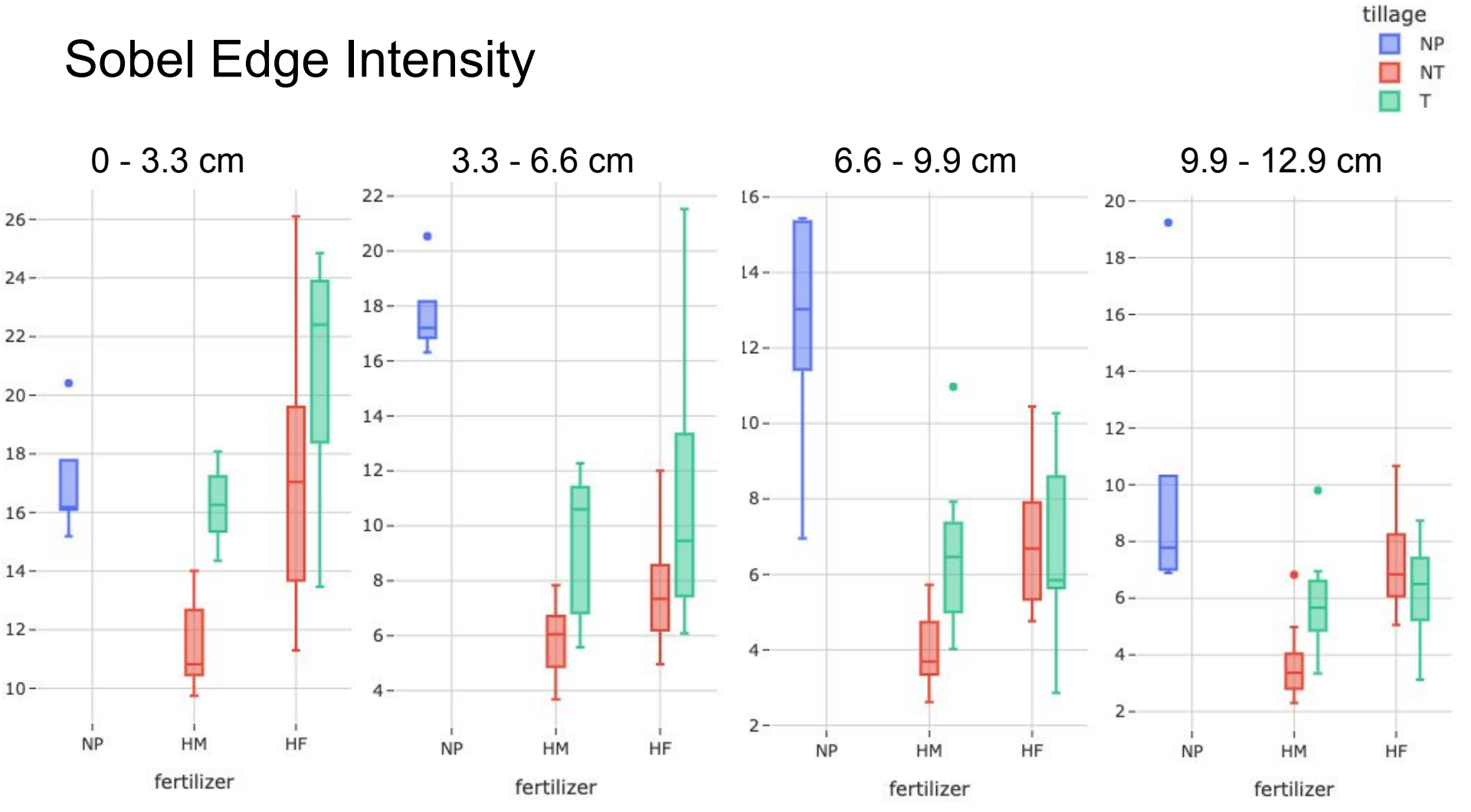
Skewness


tillage
NP
NT
T
0 - 3.3 cm
3.3 - 6.6 cm
6.6 - 9.9 cm
9.9 - 12.9 cm
NP
HM
HF
fertilizer

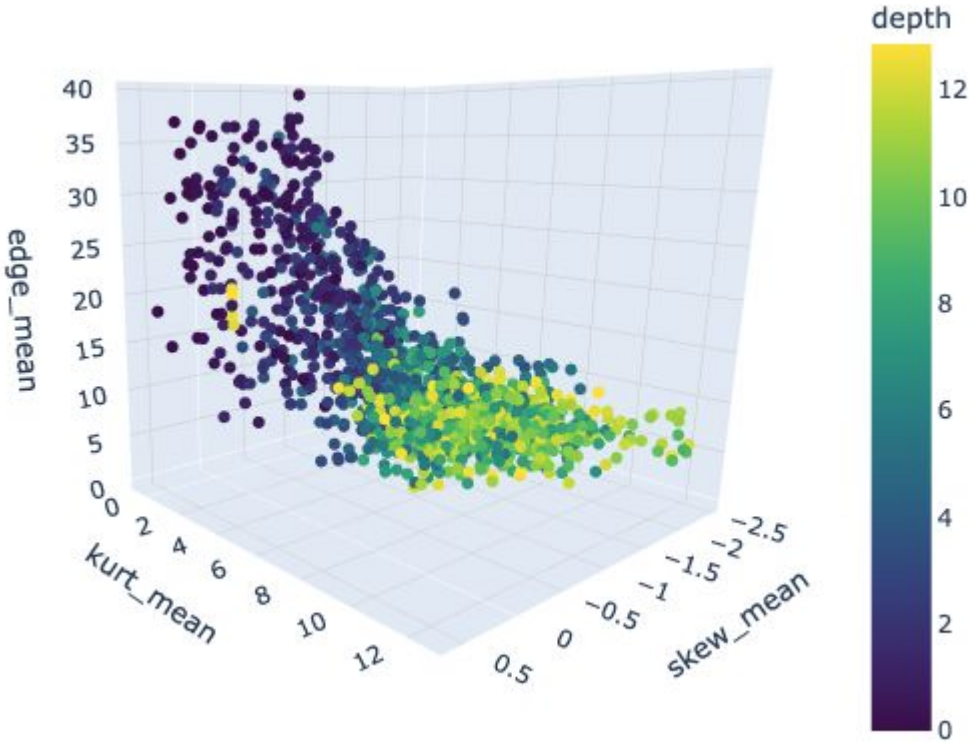
Kurtosis


tillage
NP
NT
T
0 - 3.3 cm
3.3 - 6.6 cm
6.6 - 9.9 cm
9.9 - 12.9 cm
NP
HM
HF
fertilizer

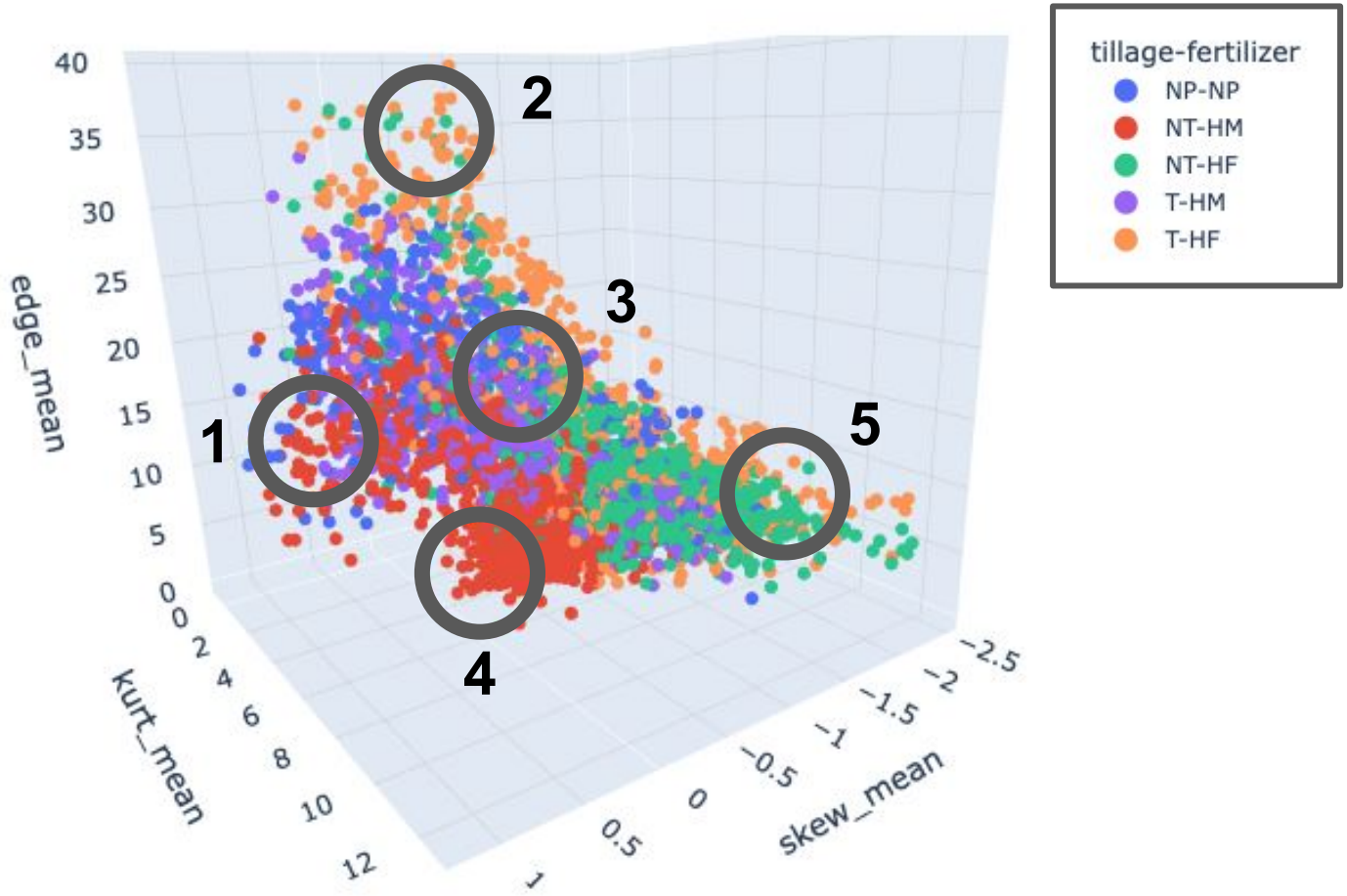
Sobel Edge Intensity



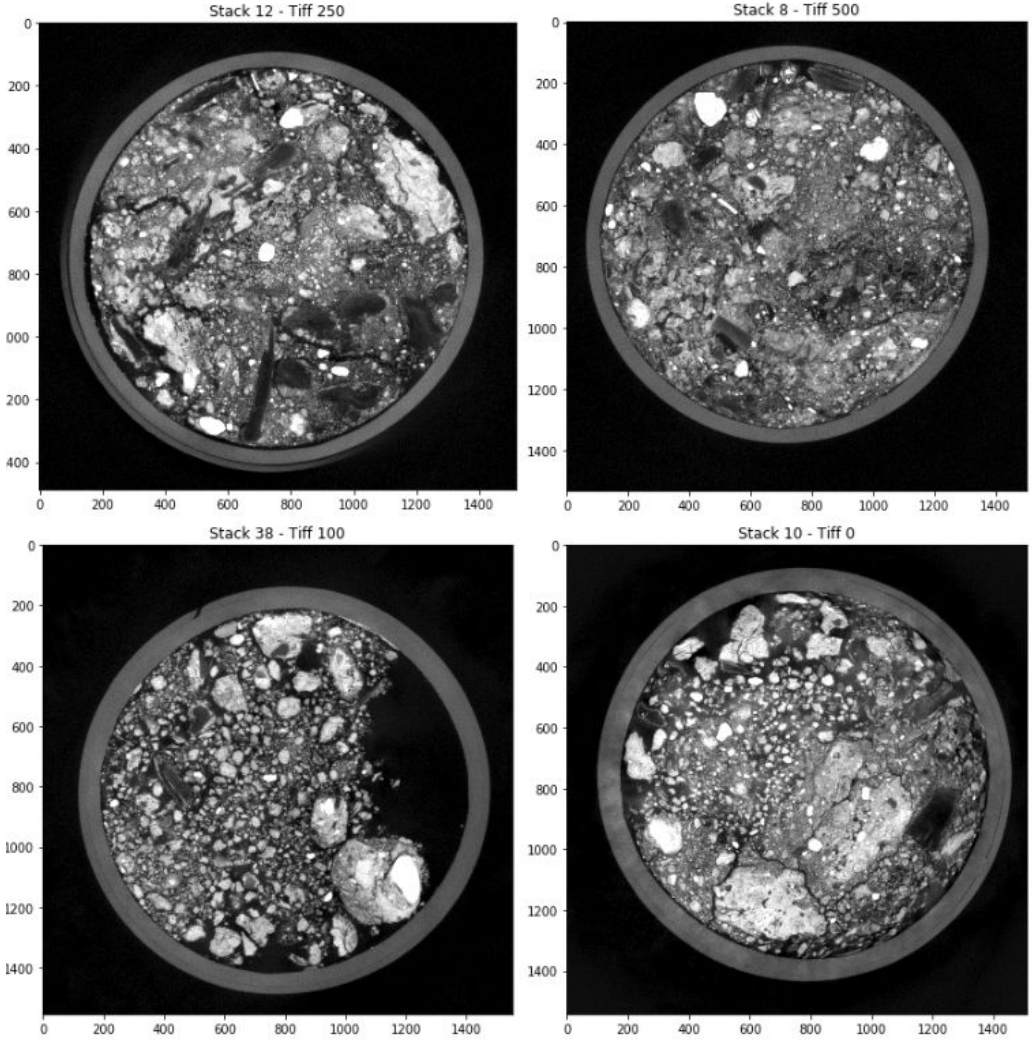
Tilled - High Fertilizer - metrics by depth



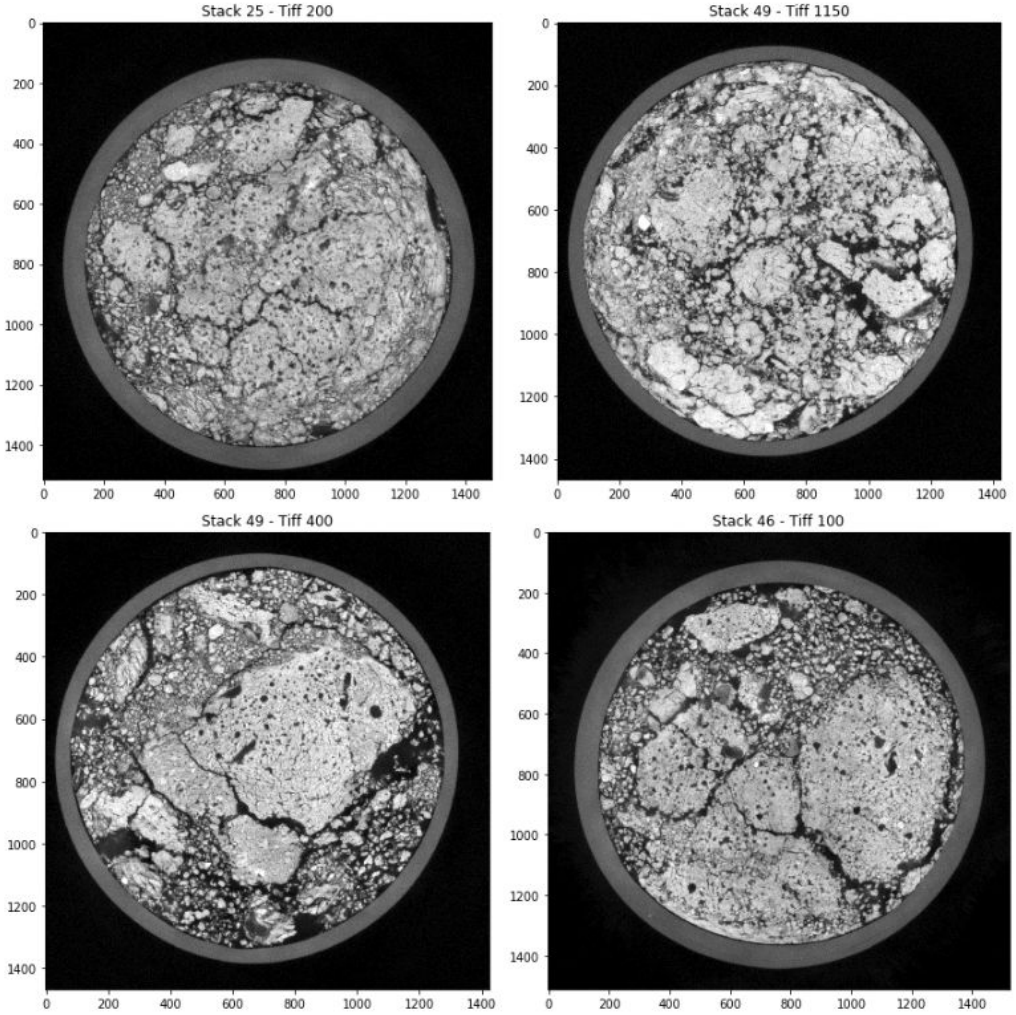
What do different regions look like?



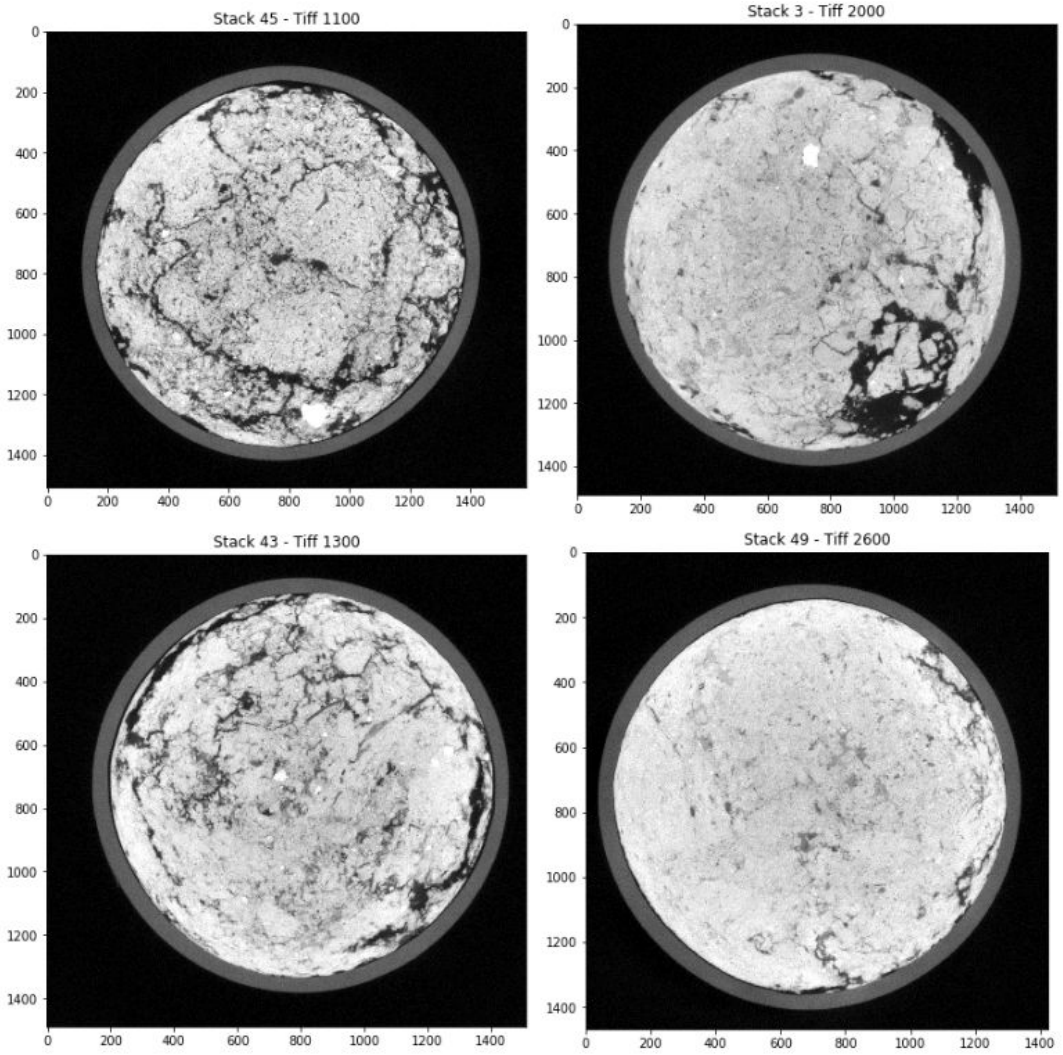
Region 1


Stack 12 - Tiff 250
Stack 8 - Tiff 500
Stack 38 - Tiff 100
Stack 10 - Tiff 0

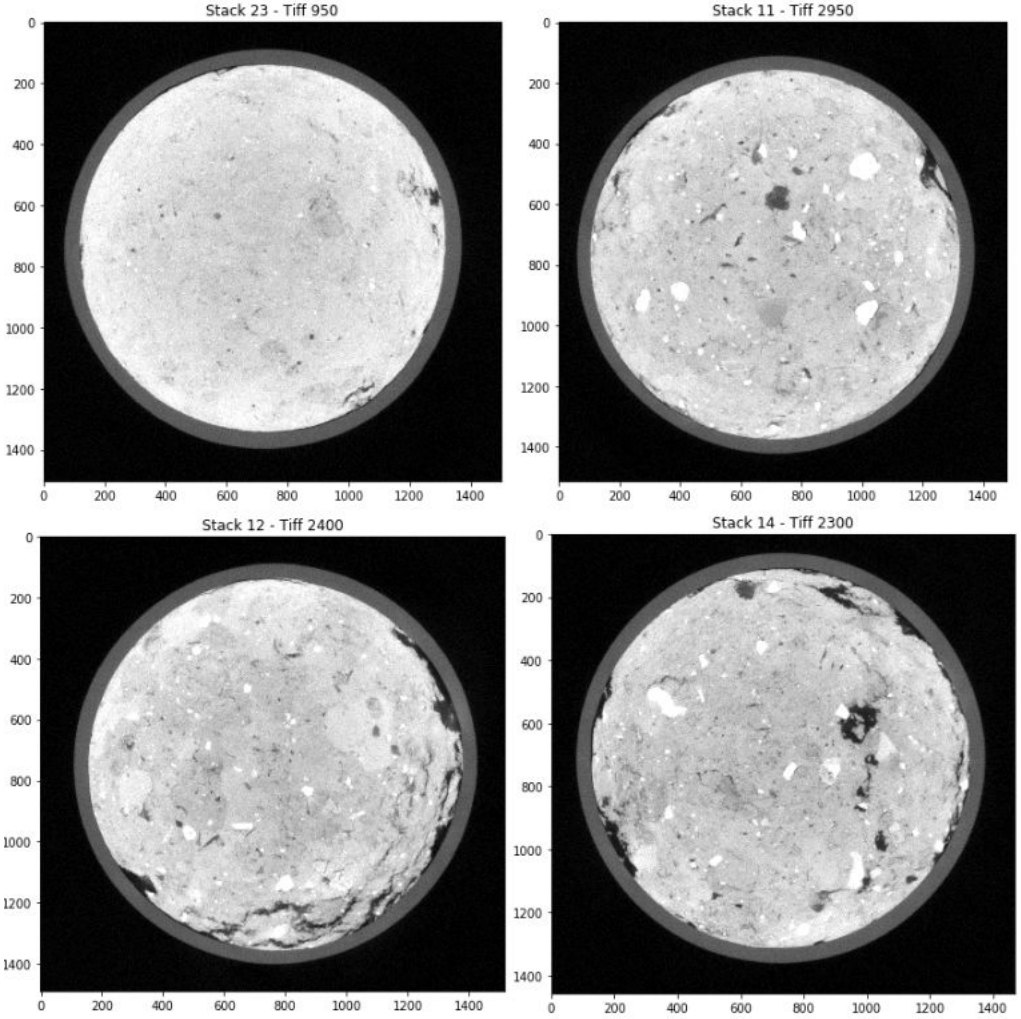
Region 2


Stack 25 - Tiff 200
Stack 49 - Tiff 1150
Stack 49 - Tiff 400
Stack 46 - Tiff 100
0
200
400
600
800
1000
1200
1400

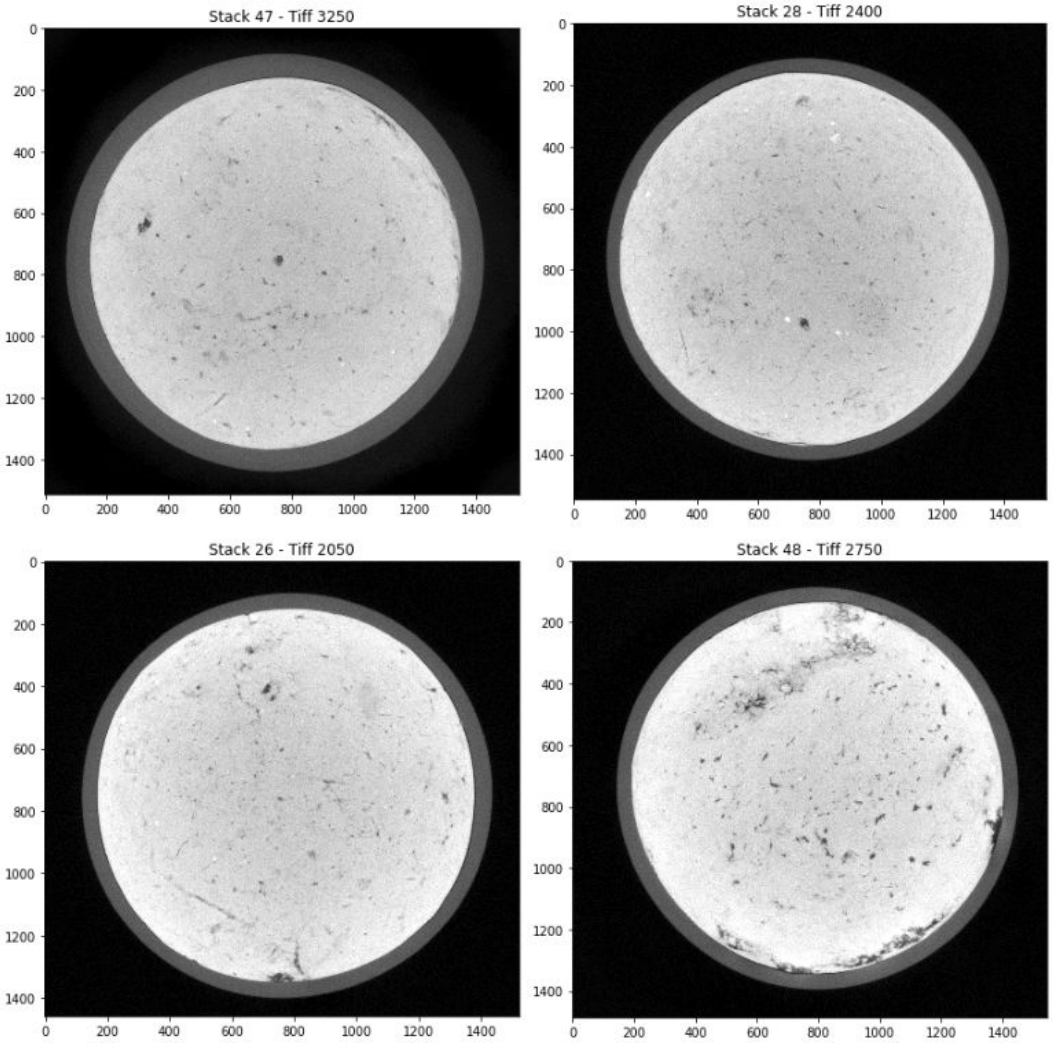
Region 3


Stack 45 - Tiff 1100
Stack 3 - Tiff 2000
Stack 43 - Tiff 1300
Stack 49 - Tiff 2600

Region 4


Stack 23 - Tiff 950
Stack 11 - Tiff 2950
Stack 12 - Tiff 2400
Stack 14 - Tiff 2300
0
200
400
600
800
1000
1200
1400

Region 5


Stack 47 - Tiff 3250
Stack 28 - Tiff 2400
Stack 26 - Tiff 2050
Stack 48 - Tiff 2750
0
200
400
600
800
1000
1200
1400