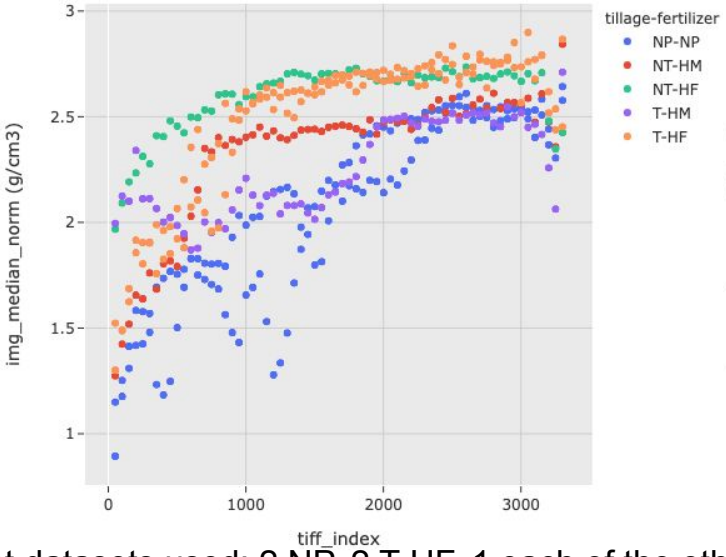
X-Ray Statistical Heterogeneity Metrics

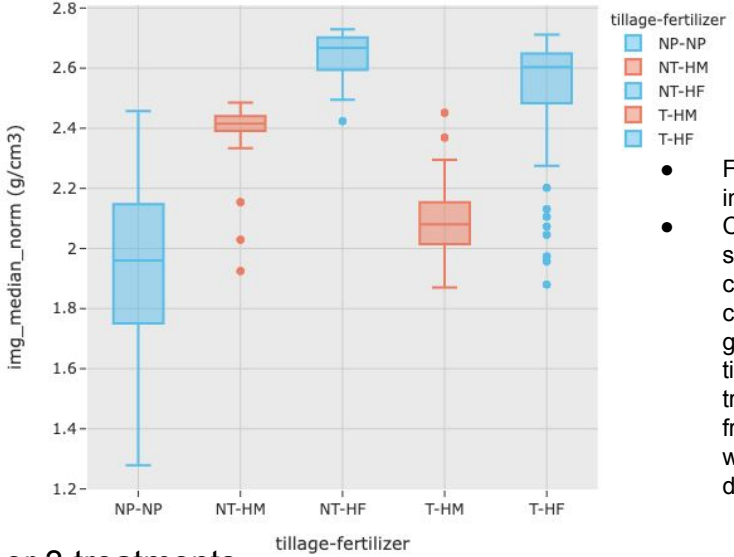
5/6/23

Overview:

- Metrics:
  - Mean, Median, standard dev and outliers (5th and 95th perc) of:
    - **Pixel Intensity (PI)**
    - **Pixel Intensity (PI*n*) normalized to approximate density.**
      - **$PI_n = (PI - PI_{outside\_of\_container})/(PI_{container} - PI_{outside\_of\_container}) * 1.022$, units: g/ml**
    - Sliding Window Skew
    - Sliding Window Kurtosis
    - Sliding Window Variance
    - Sobel Edge Convolution
  - For sliding window metrics, 3 different windows sizes, 10x10, 30x30, 50x50
  - For all metrics, calculated with and without tv denoising
  - Currently sampling every 50th tiff image in a tiff stack for faster processing

Pixel Intensity (PI*n*) normalized to approximate density





- For data between tiff indices 500 and 2000
- Currently not separated by soil core, so all tiff image calculations are grouped within a tillage-fertilizer treatment, even if from a different stack, will calculate differently later on

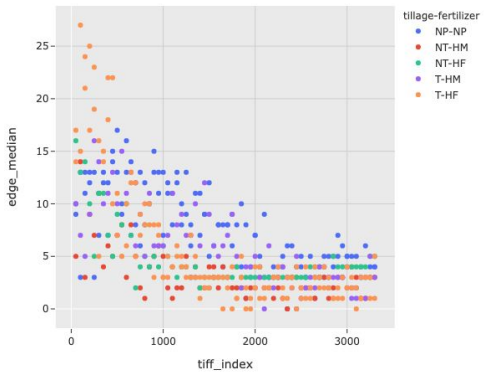
- Current datasets used: 2 NP, 2 T-HF, 1 each of the other 3 treatments
- Interesting differences in density so far, No Till > Till, High Fertilizer > High Manure, Native Prairie with lowest density.
- Density normalization used is 1.27 g/ml - from a few internet sources for PETG, **using Keith's measurement (1.022 g/ml)? See Notebook N1 towards the end regarding literature values: validation needed**

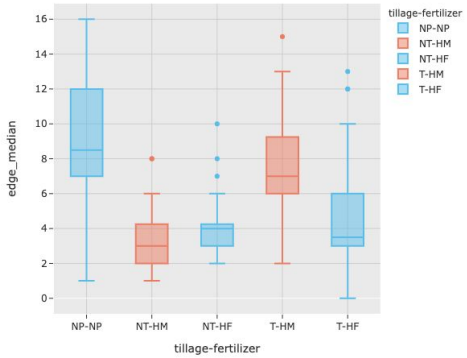
Overview:

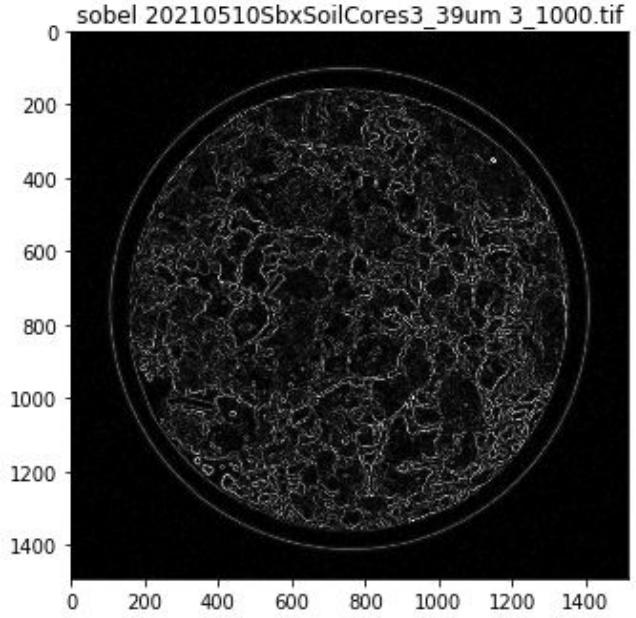
- Metrics:
  - Mean, Median, standard dev and outliers (5th and 95th perc) of:
    - Pixel Intensity (PI)
    - Pixel Intensity (PI*n*) normalized to approximate density.
      - $PI_n = (PI-PI_{outside\_of\_container})/(PI_{container}-PI_{outside\_of\_container})*1.022$, units: g/ml
    - **Sliding Window Skew**
    - **Sliding Window Kurtosis**
    - **Sliding Window Variance**
    - **Sobel Edge Convolution**
  - **For sliding window metrics, 3 different windows sizes, 10x10, 30x30, 50x50**
  - For all metrics, calculated with and without tv denoising
  - Currently sampling every 50th tiff image in a tiff stack for faster processing

Sobel Edge Detection

An image convolution that calculates the vertical and horizontal gradients in an image, good for finding edges. Can describe an aspect of heterogeneity, i.e. the extent of distinct soil regions based on finding their edges.







- For data between tiff indices 500 and 2000
- Currently not separated by soil core, so all tiff image calculations are grouped within a tillage-fertilizer treatment, even if from a different stack, will calculate differently later on

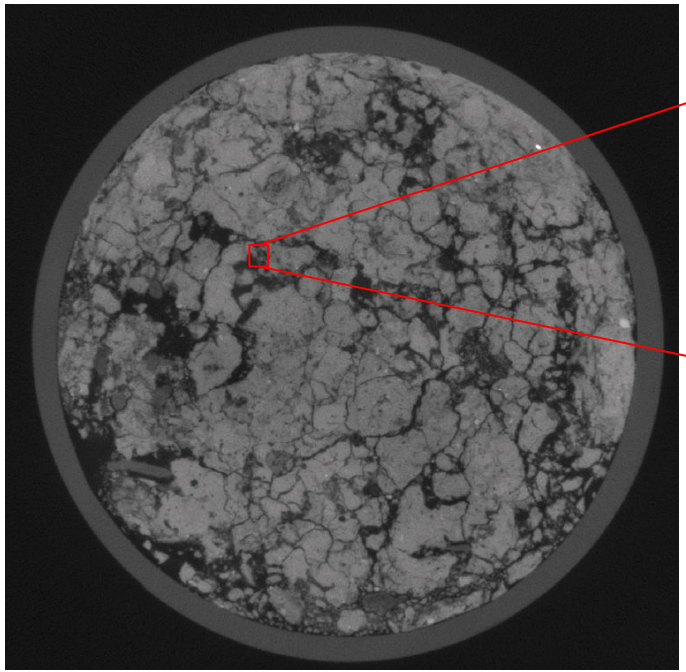
Somewhat inverse relationship to density (image intensity) measurements on average, but not perfectly equivalent: higher density doesn't always mean fewer edges or heterogeneity

Sliding Window Techniques

Calculating **skew, kurtosis and variance** across all pixels of the soil core in an image can hide interesting features at the local level that may be relevant to characterizing heterogeneity.

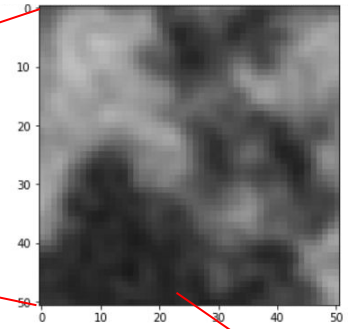
Calculating on **smaller window subsets** and moving across the image will produce metrics for each pixel (using its corresponding window), after which we can capture statistics on each metric for that image.

Sliding Window Sizes

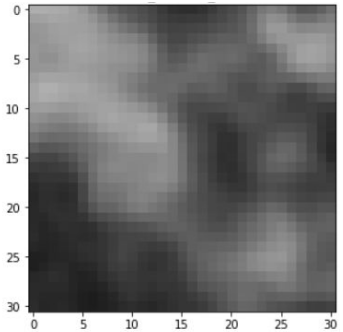


Window Size:

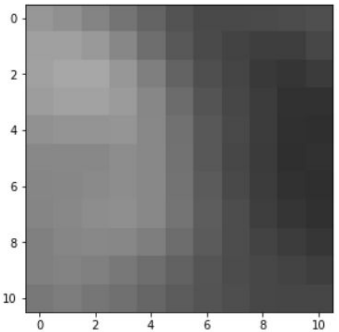
50

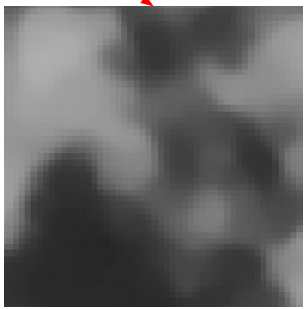


30



10



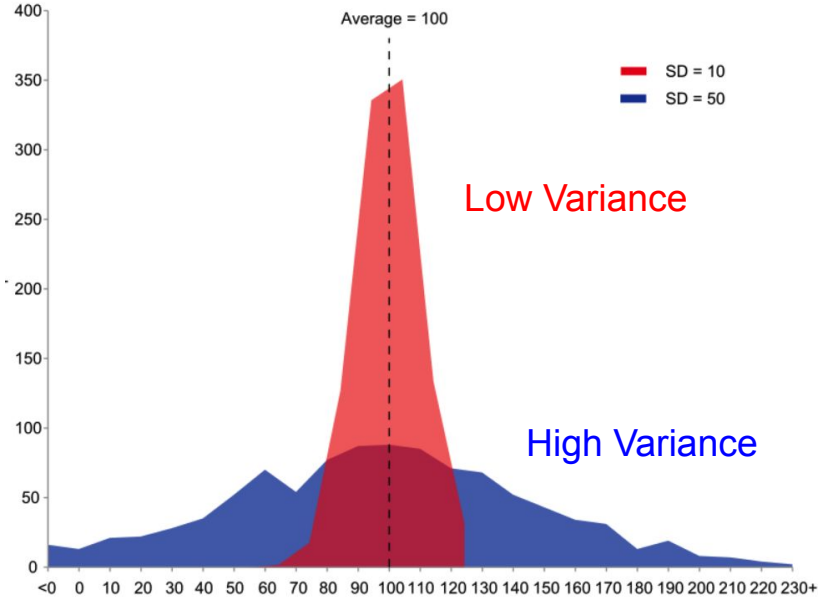


Using 50, 100, 150 pixel windows

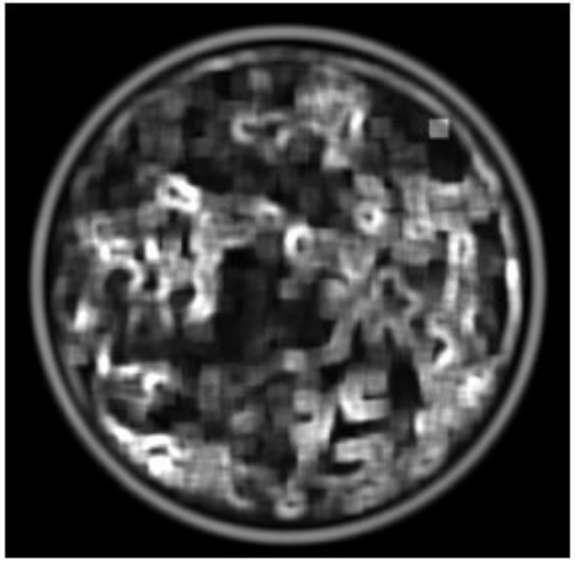
Denoised with total variation technique

Variance

How spread out is the distribution of intensities? Somewhat of an edge detector, like sobel.

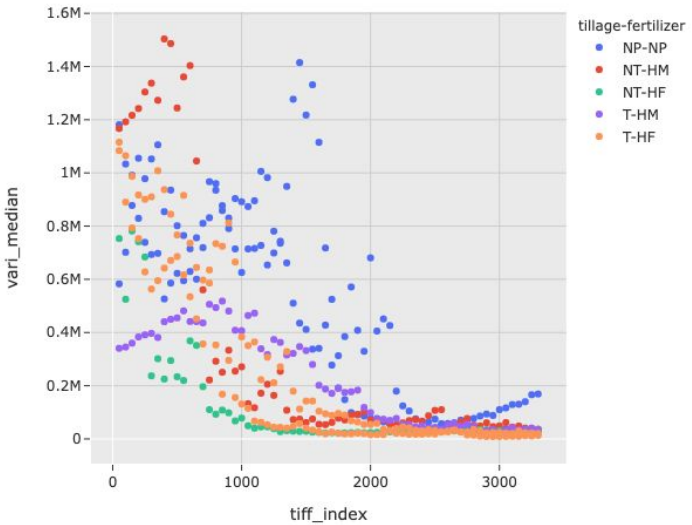


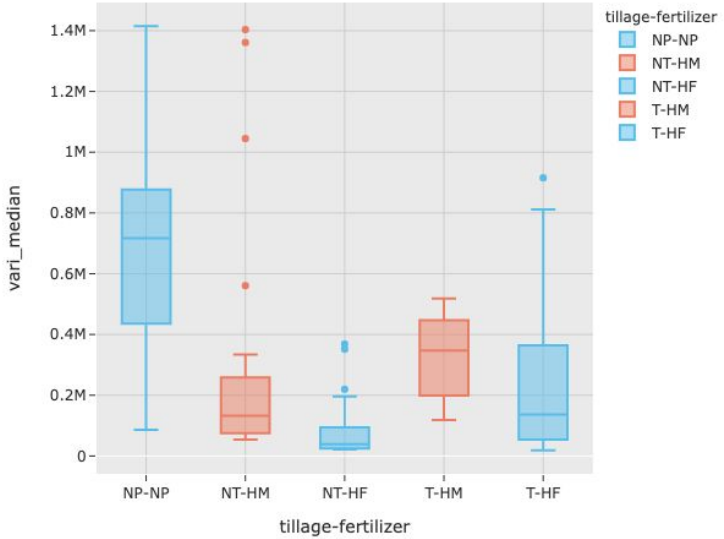
Sliding Window Size: 50



Variance

How spread out is the distribution of intensities? Somewhat of an edge detector, like Sobel operator.

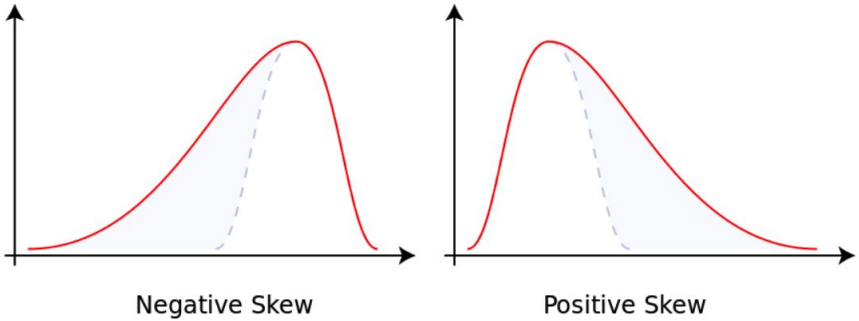




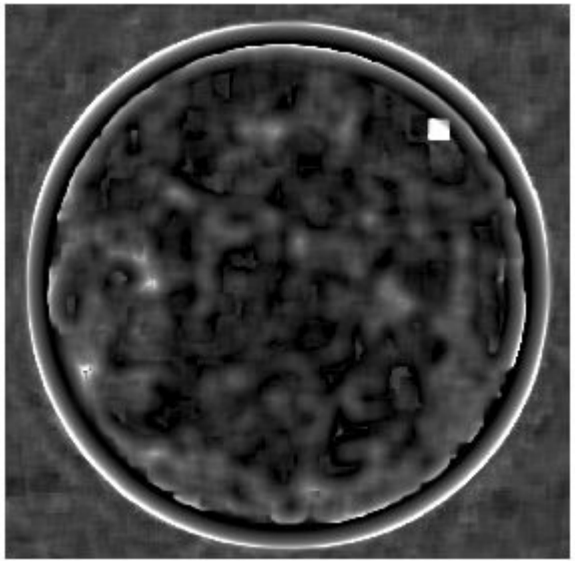
Similar differences between the different tillage-fertilizer treatments to that found with Sobel transform, although NT-HM vs NT-HF looks different between the two techniques.

Skewness

Measure of asymmetry in distribution of pixel intensities.

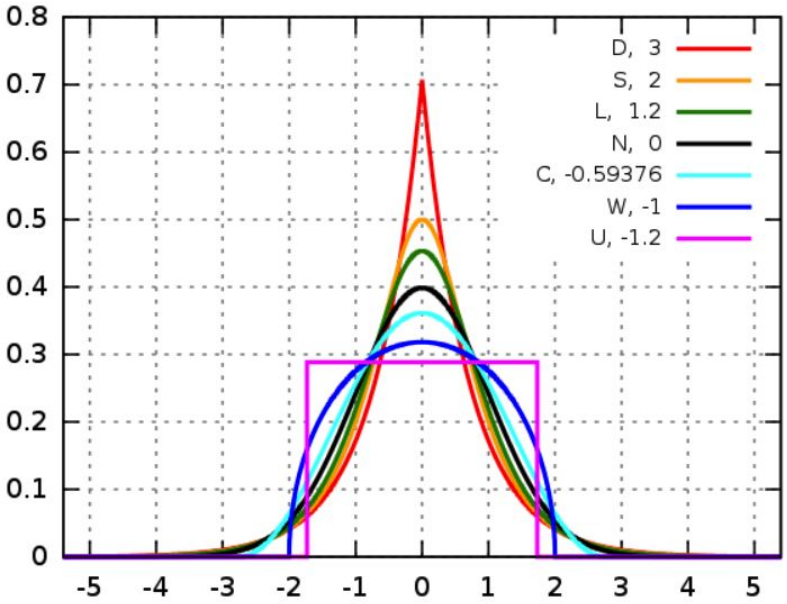


Sliding Window Size: 50

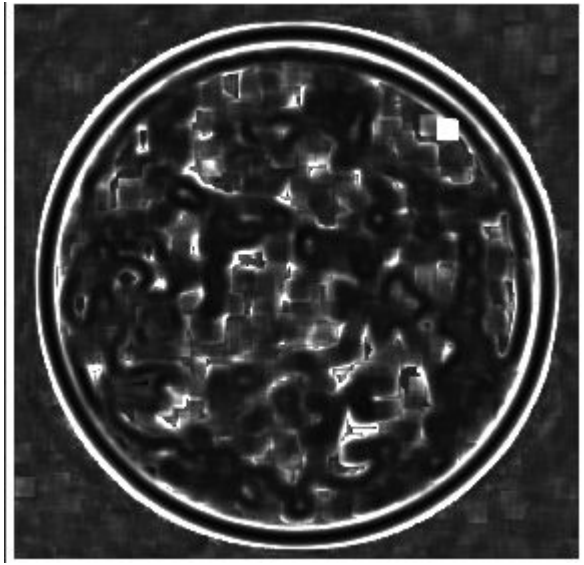


Kurtosis

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

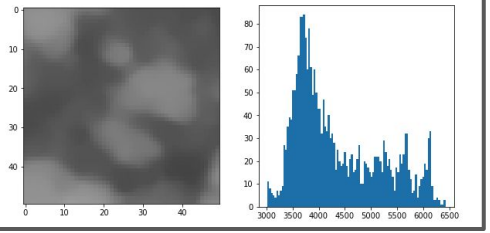


Sliding Window Size: 50

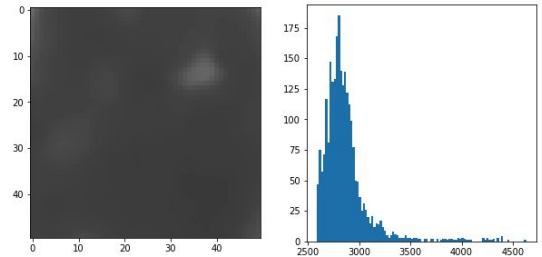


Skewness vs Kurtosis

kurtosis: -0.7035735290584433, skew: 0.7068692420944443

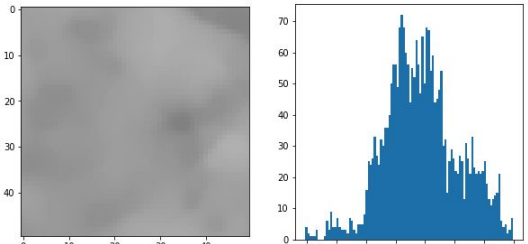


kurtosis: 13.761542936508711, skew: 3.139147400978889

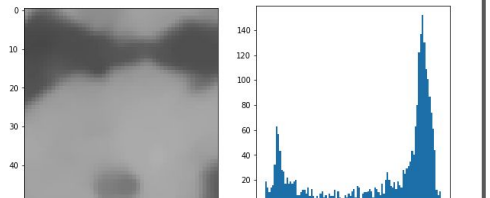


Kurtosis

kurtosis: -0.05701518338358991, skew: 0.07986803020788472

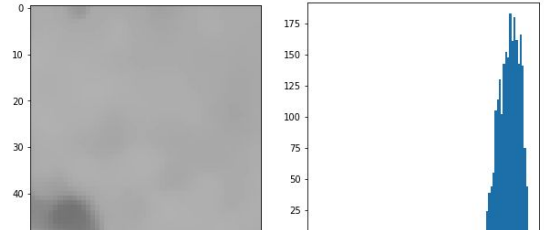


kurtosis: -0.8792197836069437, skew: -0.8676381645358686

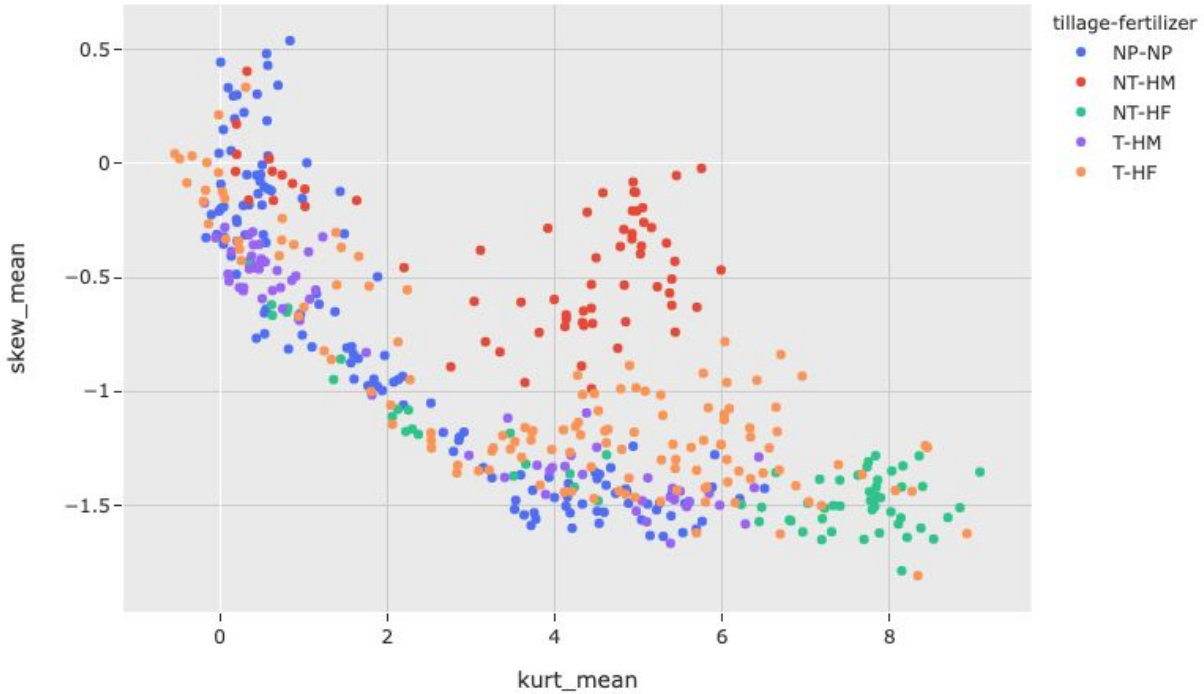


Skewness

kurtosis: 15.198680629169704, skew: -3.76185317052183



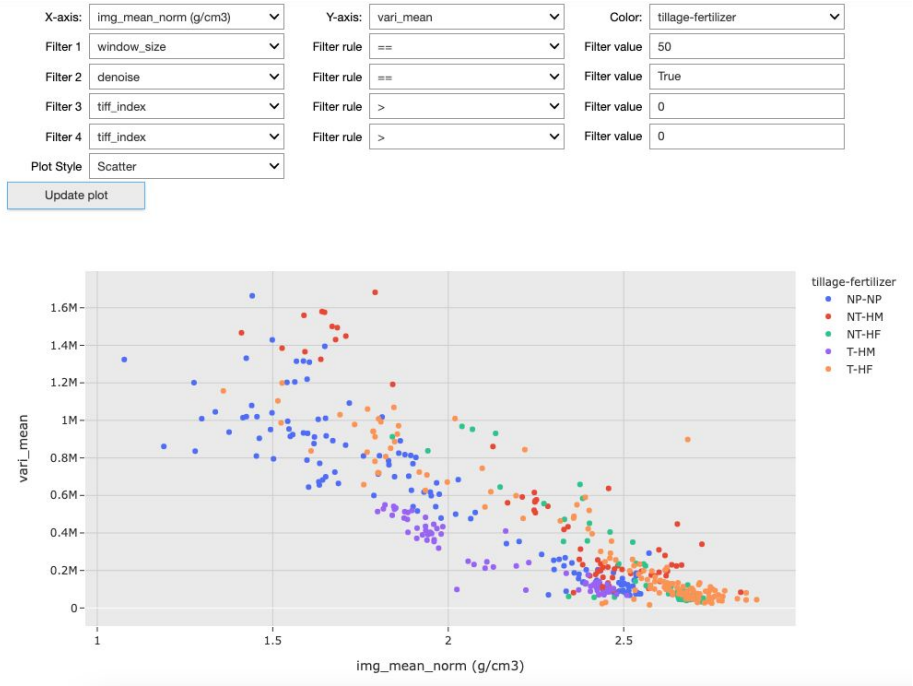
Skewness vs Kurtosis



- Potentially interesting clusters distinguishing the different soils types.
- Will be useful to think about qualitative physical meaning behind these metrics..

Many ways to parse these metrics

Each metric has median, mean, std, 5th and 95th percentiles. 3 different sliding window sizes for skew, kurt and variance



Streamlit app of metrics available

Other approaches to explore:

- Identifying pore space
- Additional quantitative techniques
  - spatial autocorrelation, entropy, fractal analysis