General workflow outline for joining yield values with independent variables, cleaning fertilizer data and, and final yield point selection.

Updated 2018-05-14 by Phil Davis

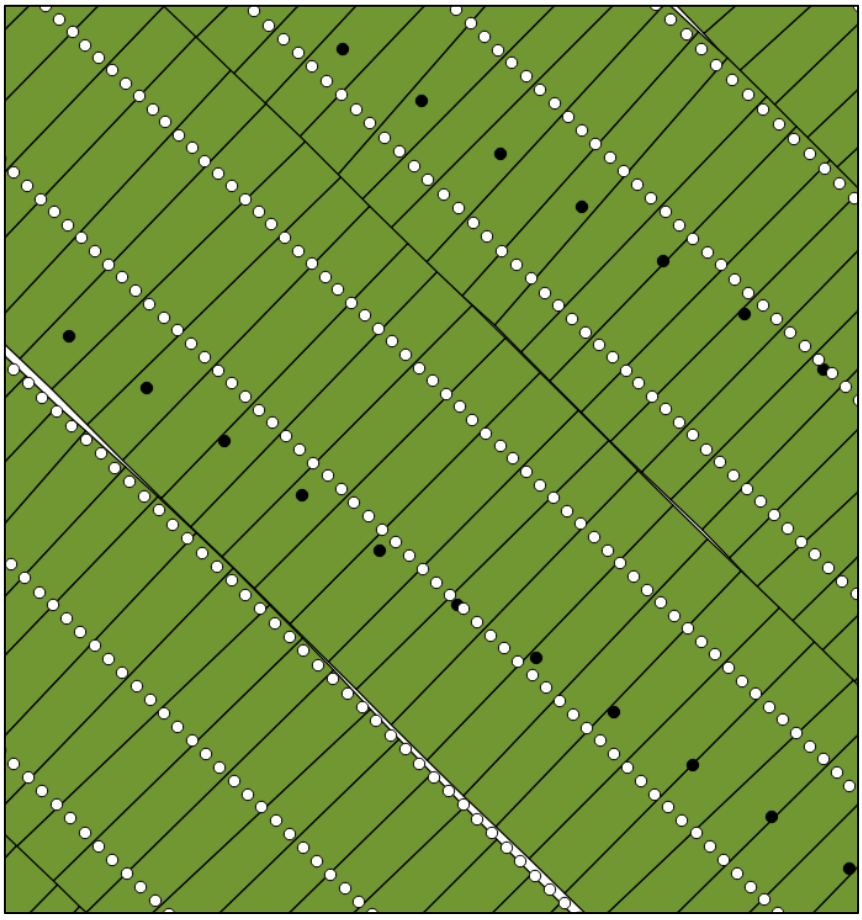
There are three main steps to the process:

1) Relate yield points to all independent variables.
2) Process fertilizer application data to extract areas where fertilizer rate data may be inaccurate.
3) Remove yield points based on fertilizer application file cleaning and proximity to field edges.

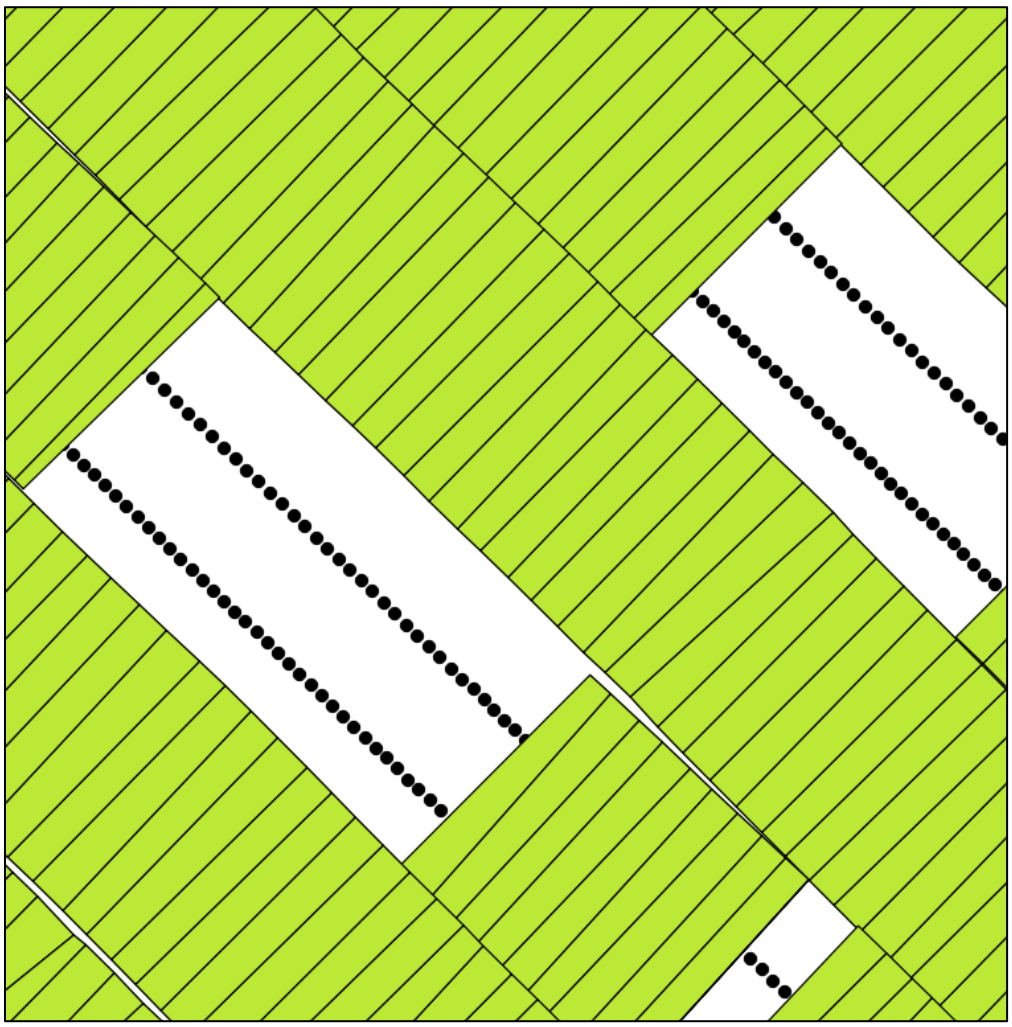
Detailed descriptions of the three steps:

1) **Relate current crop yield values to independent variables. Polygon and raster values are related to yield values by 'point sampling' and independent points values, such as Veris data and previous yield, are spatially joined to their nearest yield neighbor.**

The white points are yield values, black points are Veris points, with multiple attributes, and the polygons are from the fertilizer application data. Using the yield points, the polygons are 'point sampled' and the Veris attributes are joined to the yield values by using a nearest neighbor approach. This same method is used for all other raster files, such as DEM, slope, vegetation indices, etc., and for points such as previous year's yield:



IMPORTANT: Fertilizer application equipment does not create polygons where 0 fertilizer was applied, so point sampling in these areas will result in ‘null’ values. Those then need to be population with ‘0’ for rate:

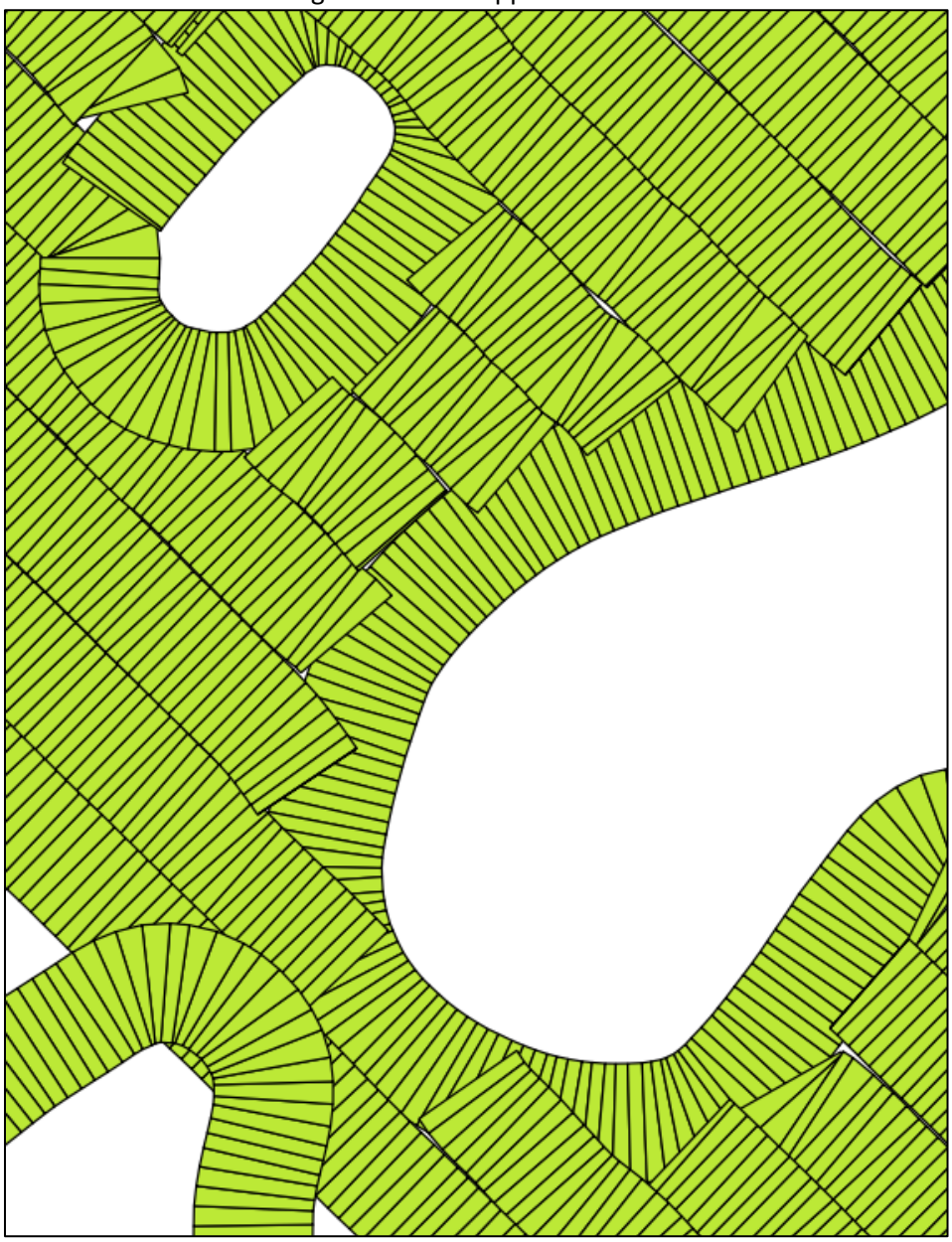


2) Cleaning of the fertilizer application file.

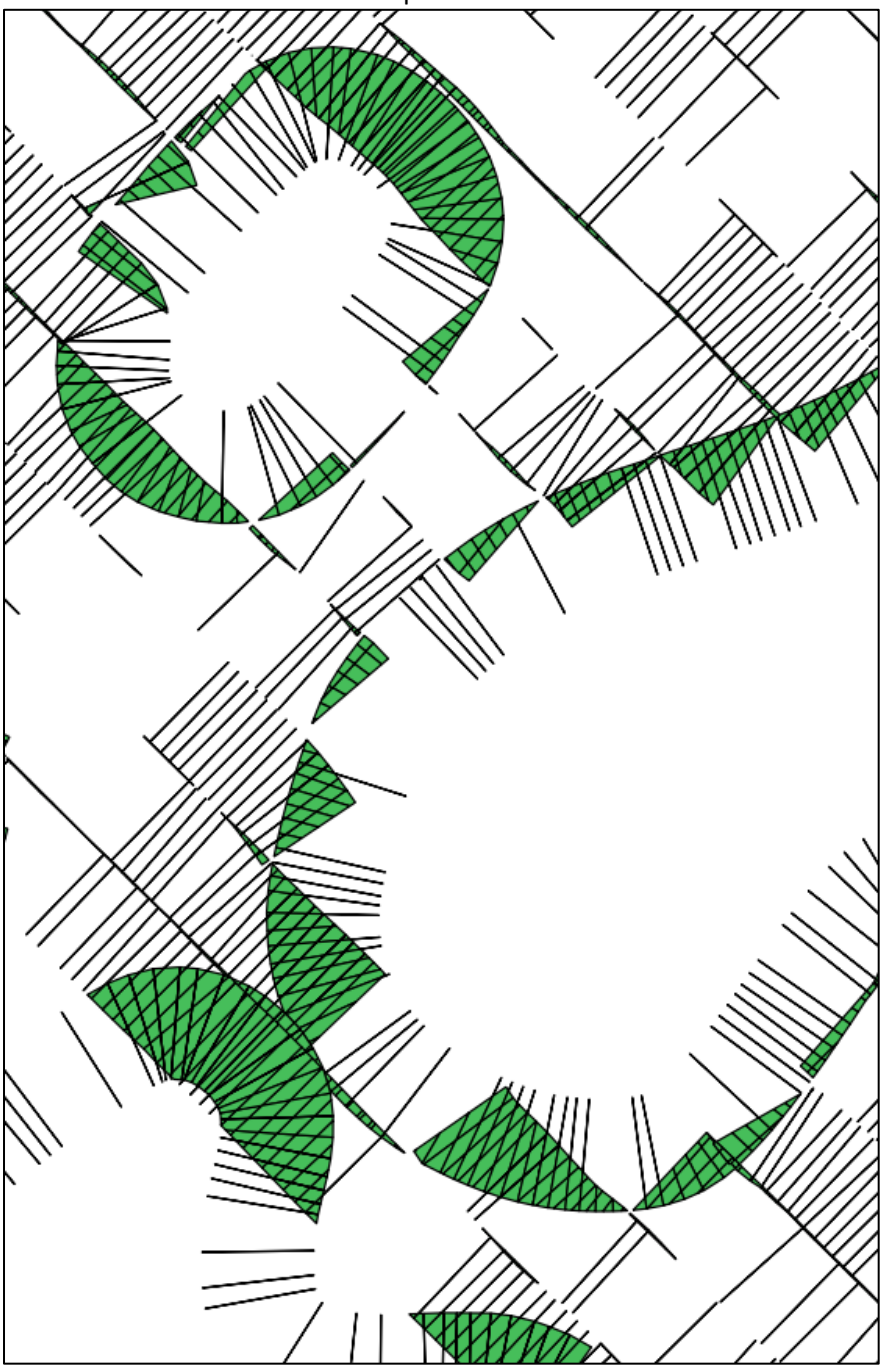
Fertilizer files are cleaned to identify areas where polygons overlapped during application and to extract the lines between adjacent passes of fertilizer equipment. By isolating these two aspects of the fertilizer files we can remove yield points that fall close to the polygons/lines to enhance the accuracy of the aggregated dataset by removing potentially unreliable yield values.

a) Identify overlapped polygons and save as a new file.

Original fertilizer application data:

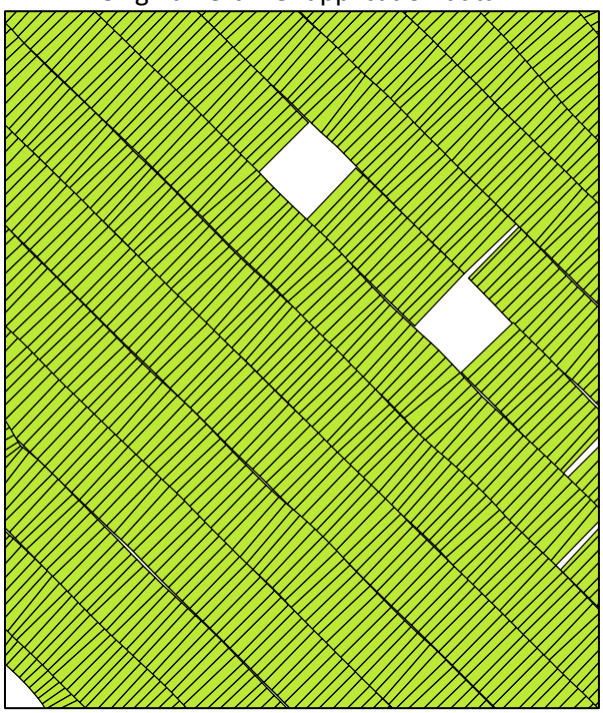


Overlaps extracted:

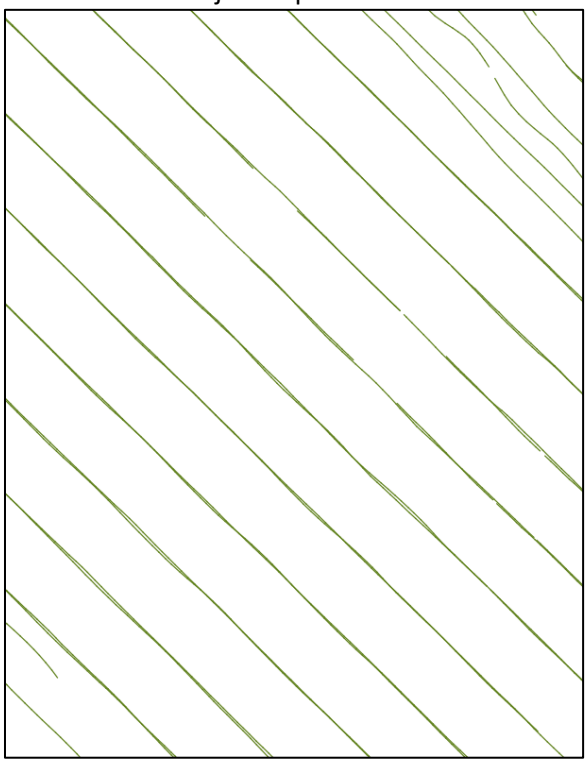


b) Extract lines of adjacent passes and save file. By identifying these, we can eliminate yield points that were collected while the combine was straddling two different fertilizer rates.

Original fertilizer application data:



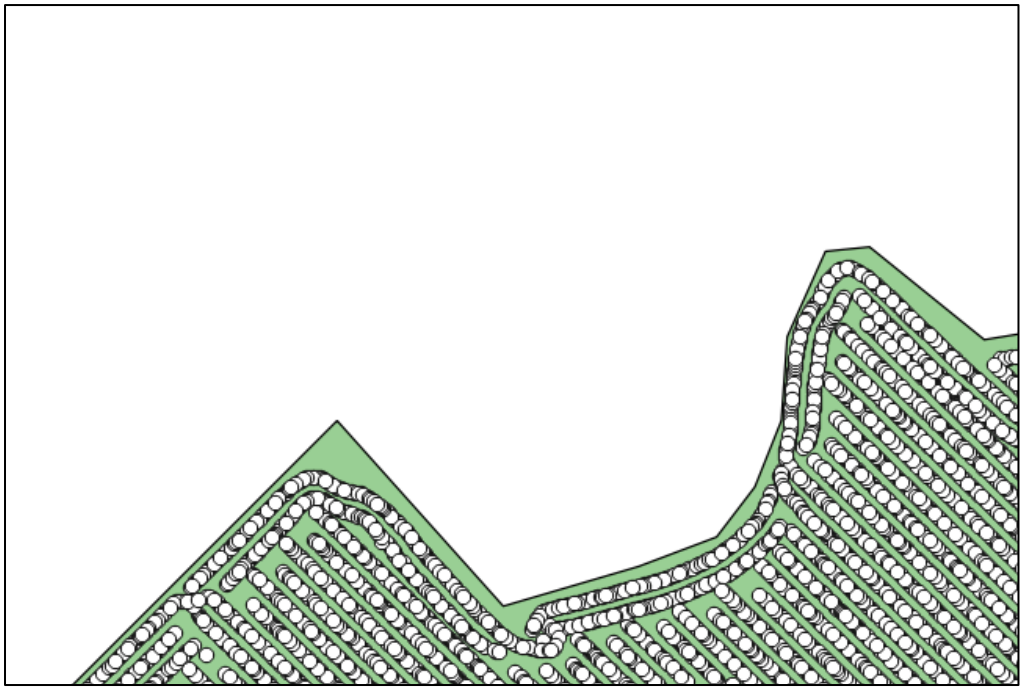
Lines of adjacent passes extracted:



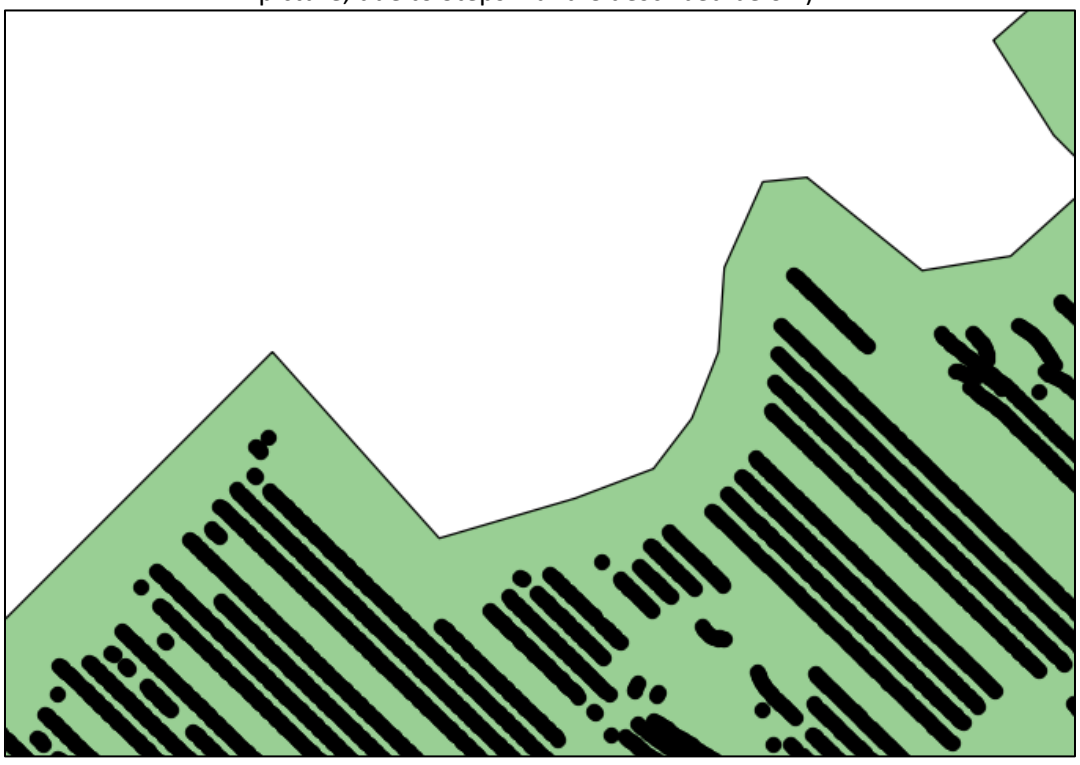
3) Reduce the yield point dataset by eliminating points that were collected in areas that may provide unreliable data. The three main areas where data is often suspect include:

a. Field edges – OFPE's standard is to remove points within 30 m of field edges:

Screenshot of points at field edges:

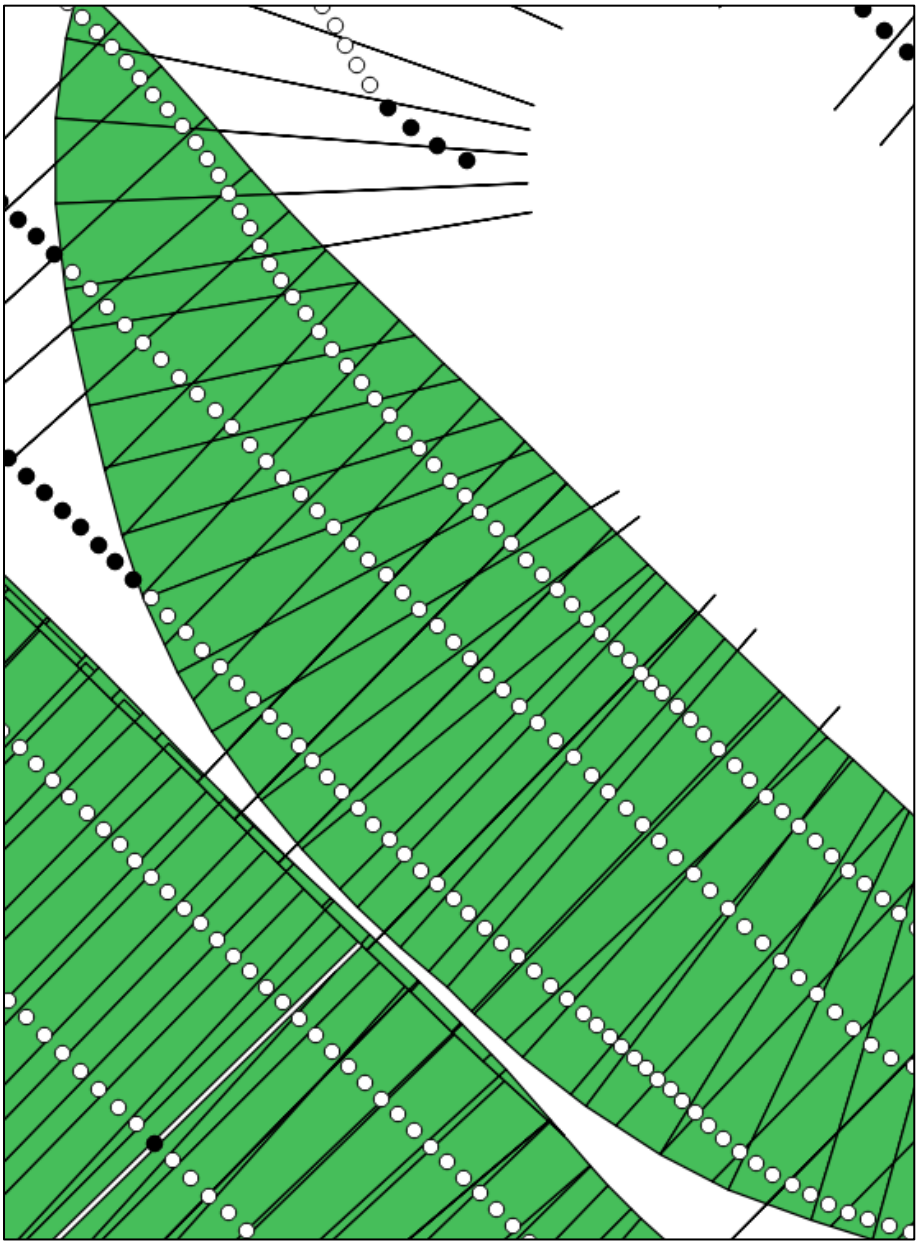


After removal of points within 30m of field edges (some other points are also removed in this picture, due to Steps B and C described below):



b. Overlapping fertilizer polygons – many of the fertilizer values that were 'point sampled' described above in Step 2 are inaccurate because the fertilizer value assigned to each point is only from the 'top' polygon in the dataset and does not include whatever rate was applied first (the 'bottom' polygon). Therefore, those points should be removed from the dataset.

Original points (white) fell in areas of fertilizer overlap and are deleted from the final dataset (black):



c. Edges of adjacent fertilizer passes. Often, the combine harvested while straddling two vastly different fertilizer rates. Those points need to be eliminated. I chose a distance of 3m, which ensures that at least 75% of the width of the combine header collected data in the polygon whose value was picked up by the yield point.

Lines of adjacent fertilizer passes (extracted as described in Step 2) are in green, original points are white, and the reduced dataset includes only black points. White points within 3m of fertilizer lines were removed: