HTP Geoprocessor: An open-source tool for geoprocessing data from vehicle-based sensing platforms

Dr. Kelly R. Thorp
USDA-ARS, Arid Land Agricultural Research Center
kelly.thorp@ars.usda.gov
February 27, 2018

For more information, see the following:

Wang, X., Thorp, K. R., White, J. W., French, A. N., Poland, J. A., 2016. Approaches for geospatial processing of field-based high-throughput plant phenomics data from ground vehicle platforms. Transactions of the ASABE 59 (5), 1053–1067. doi:10.13031/trans.59.11502

Andrade-Sanchez, P., Gore, M. A., Heun, J. T., Thorp, K. R., Carmo-Silva, A. E., French, A. N., Salvucci, M. E., White, J. W., 2014. Development and evaluation of a field-based, high-throughput phenotyping platform. Functional Plant Biology 41 (1), 68-79. doi:10.1071/FP13126

Also, see the webinar here:
www.extension.org/pages/68270

Overview

Sensing systems on ground-based vehicles are now commonly utilized to collect data for applications in field-based high-throughput phenotyping. However, geoprocessing capability is needed to relate the sensor data to other geospatial characteristics, such as treatment plots in the field. HTP Geoprocessor is a flexible geospatial software program created for this purpose. The algorithms and user interfaces were developed as a plug-in for the open-source Quantum GIS environment (www.qgis.org). The plug-in incorporates three main tools: the Map Creator, the Preprocessor, and the Geoprocessor. The software design makes these geospatial tools readily transferable to a variety of vehicle-based sensor platforms.

Map Creator

The main purpose of the Map Creator (Fig. 1) is to create a polygon shapefile to delineate field plots and map the layout of field experiments. The tool requires knowledge of the geographic coordinates outlining each field plot. The tool calculates the convex hull of these coordinates, which is the smallest area that contains all the points. For rectangular field plots, only the four corner coordinates are necessary. However, the tool is also able to create polygons for non-rectangular plots. Required coordinates can be obtained from calculations made during the planning phases of field trails, from field surveys with a GPS unit during field trials, or from post-processing activities on the phenotyping data itself.

The Map Creator expects to process coordinate information from a comma delimited (CSV) file, which must be prepared by the user. An example of the required format is shown in Figure 2. The first line of the file contains header information, and the remaining lines contain the plot coordinate and plot name information. The tool expects the first column to provide a set of unique text strings to identify each plot. This plot name will be used in later geoprocessing steps to tie sensor information to plot information. Additional columns of plot information can be added to the shapefile that results from the map creation step; however, the Map Creator tool expects only a unique plot name identifier. Additional column in the CSV file provide information on the coordinate pairs, grouped in consecutive columns. Horizontal coordinates should appear before vertical coordinates. There should be three or more coordinate pairs. There can be different numbers of coordinate pairs per plot. The user will be asked to specify the coordinate reference system of the data in the CSV file. Thus, the tool can process latitude/longitude coordinates, Universal Transverse Mercator (UTM) coordinates, or other coordinate systems available in the Quantum GIS database. To use the Map Creator (Fig. 1) tool, do the following:

1. Select Plugins->HTP Geoprocessor-> Map Creator
2. Specify the file path of the CSV input coordinate file.
3. Specify the file path of the shapefile to store plot map information
4. The "Run" button will load the CSV input coordinate file and create a polygon shapefile of the plot boundary map. The user will be asked to specify the coordinate reference system of the data.

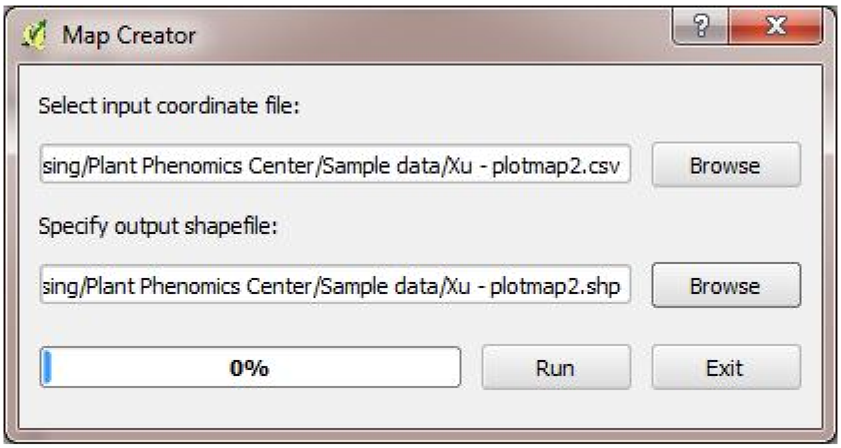
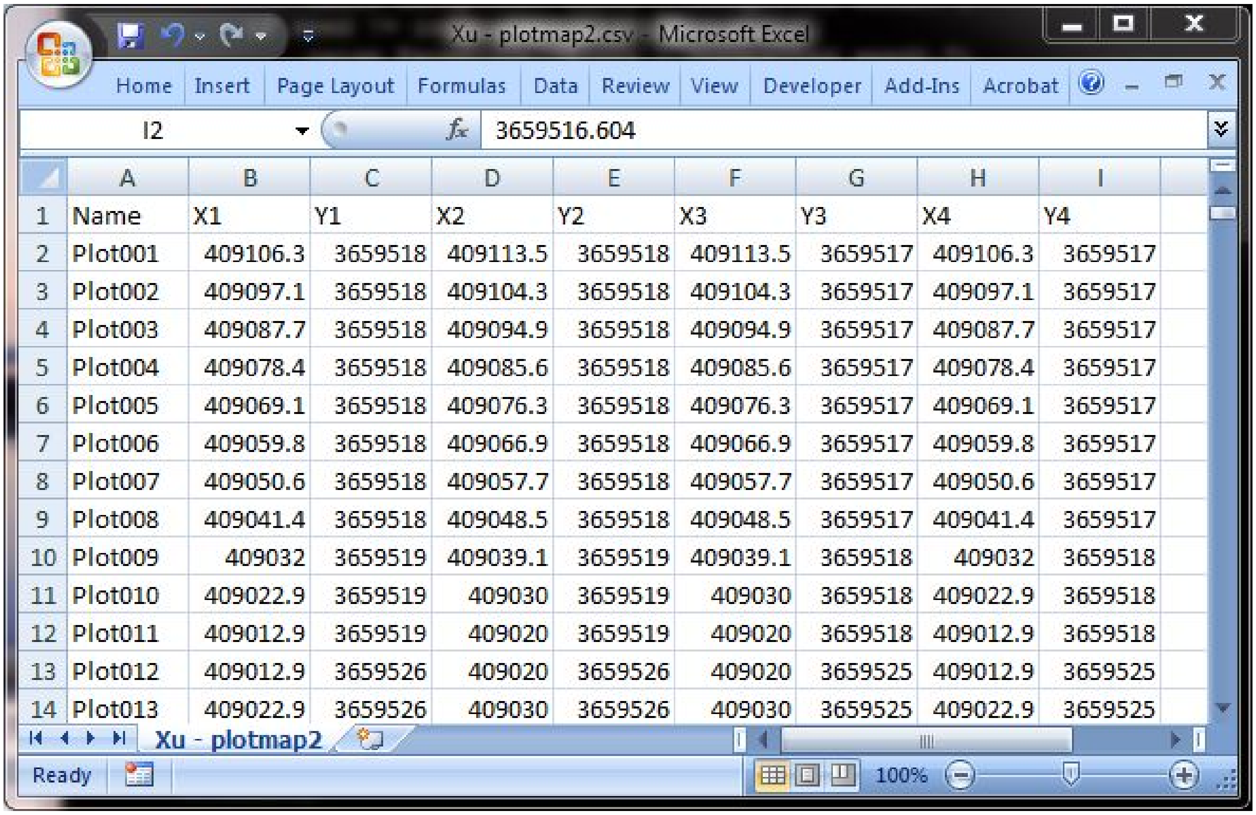


Figure 1. The graphics users interface for the Map Creator tool



Name	X1	Y1	X2	Y2	X3	Y3	X4	Y4
Plot001	409106.3	3659518	409113.5	3659518	409113.5	3659517	409106.3	3659517
Plot002	409097.1	3659518	409104.3	3659518	409104.3	3659517	409097.1	3659517
Plot003	409087.7	3659518	409094.9	3659518	409094.9	3659517	409087.7	3659517
Plot004	409078.4	3659518	409085.6	3659518	409085.6	3659517	409078.4	3659517
Plot005	409069.1	3659518	409076.3	3659518	409076.3	3659517	409069.1	3659517
Plot006	409059.8	3659518	409066.9	3659518	409066.9	3659517	409059.8	3659517
Plot007	409050.6	3659518	409057.7	3659518	409057.7	3659517	409050.6	3659517
Plot008	409041.4	3659518	409048.5	3659518	409048.5	3659517	409041.4	3659517
Plot009	409032	3659519	409039.1	3659519	409039.1	3659518	409032	3659518
Plot010	409022.9	3659519	409030	3659519	409030	3659518	409022.9	3659518
Plot011	409012.9	3659519	409020	3659519	409020	3659518	409012.9	3659518
Plot012	409012.9	3659526	409020	3659526	409020	3659525	409012.9	3659525
Plot013	409022.9	3659526	409030	3659526	409030	3659525	409022.9	3659525

Figure 2. A comma delimited file of plot coordinates in Microsoft Excel.

Preprocessor

The main purposes of the Preprocessor are 1) to convert latitude and longitude coordinates to the Universal Transverse Mercator (UTM) coordinate system and 2) to calculate coordinate transformations from the GPS receiver to the sensor positions depending on the vehicle's direction of travel. Proper coordinate transformations are particularly important for high-throughput phenotyping, because the sensors at each position along the boom may collect data from different treatment plots. The Preprocessor tool was designed for flexibility to read sensor data in a variety of formats from comma delimited sensor data files. The user interface requires entries that provide instruction for properly reading sensor data from a file, such as the column numbers for latitude, longitude, vehicle heading, and sensor information. Offsets (m) from the GPS receiver to each sensor position are also required.

The Preprocessor expects to process sensor data from comma delimited (CSV) files only. Header information should appear in the first line of the file (this is a modification from the previous version of the tool). To use the Preprocessor (Fig. 3) tool, do the following:

1. Select Plugins->HTP Geoprocessor-> Preprocessor
2. Enter the total number of columns in the CSV sensor data file. The tool will not process any lines of data that do not have the specified number of columns. Blank lines will also be eliminated.
3. In the table under "Sensor Name," enter a name for each sensor included in the file. There should be entries on one line per sensor. For example, if there are four sensors, four lines of information should be completed.
4. In the table under "Latitude Column #," enter the column number containing the latitude information for each sensor (counted from left to right starting at 1).

5. In the table under "Longitude Column #," enter the column number containing the longitude information for each sensor (counted from left to right starting at 1).
6. In the table under "Heading Column #," enter the column number containing the heading information for each sensor (counted from left to right starting at 1).
7. In the table under "Sensor ID Column #," do one the following:
    a. If the data from each sensor is written on consecutive lines, enter the column number containing the sensor ID for each sensor (counted from left to right starting at 1).
    b. If the data from each sensor is written on the same line, leave blank.
8. In the table under "Sensor ID Text," do one the following:
    a. If the data from each sensor is written on consecutive lines, enter the text in the sensor ID column that denotes each sensor.
    b. If the data from each sensor is written on the same line, leave blank.
9. In the table under "Data Column #(s),", enter the column numbers containing data that is specific to each sensor. More than one column can be entered. The columns do not have to be consecutive. Each entered column number should be separated by a comma.
10. In the table under "Easting Offset (m)," enter the easting offset from the GPS receiver to each sensor. Assume the vehicle is facing north, such that west is negative and east is positive.
11. In the table under "Northing Offset (m)," enter the northing offset from the GPS receiver to each sensor. Assume the vehicle is facing north, such that north is positive and south is negative.
12. If there is any offset between the GPS system used to collect the sensor data and the GPS system used to make the plot map, enter those offsets in the boxes at the bottom left. Offsets are considered <u>from the vehicle GPS to the field map GPS</u>. East and north are positive. West and south are negative.
13. The "Save File" button will save this information to a file for later use.
14. The "Load File" button will load the information back into the window.
15. The "Run" button will preprocess the sensor data in a data file. When "Run" is pressed, the user is asked to select the file for processing. The Preprocessor will then process the data and save a new file with the same name, but appended with "-preprocess" between the file name and file extension. For example, "temp.csv" would become "temp-preprocess.csv". The processed file is written at the same path as the original file. If it already exists, it will be overwritten. Sensor data files will be reorganized to provided UTM coordinates at sensor locations, sensor data collected and that location, and other information originally recorded in the sensor data file.

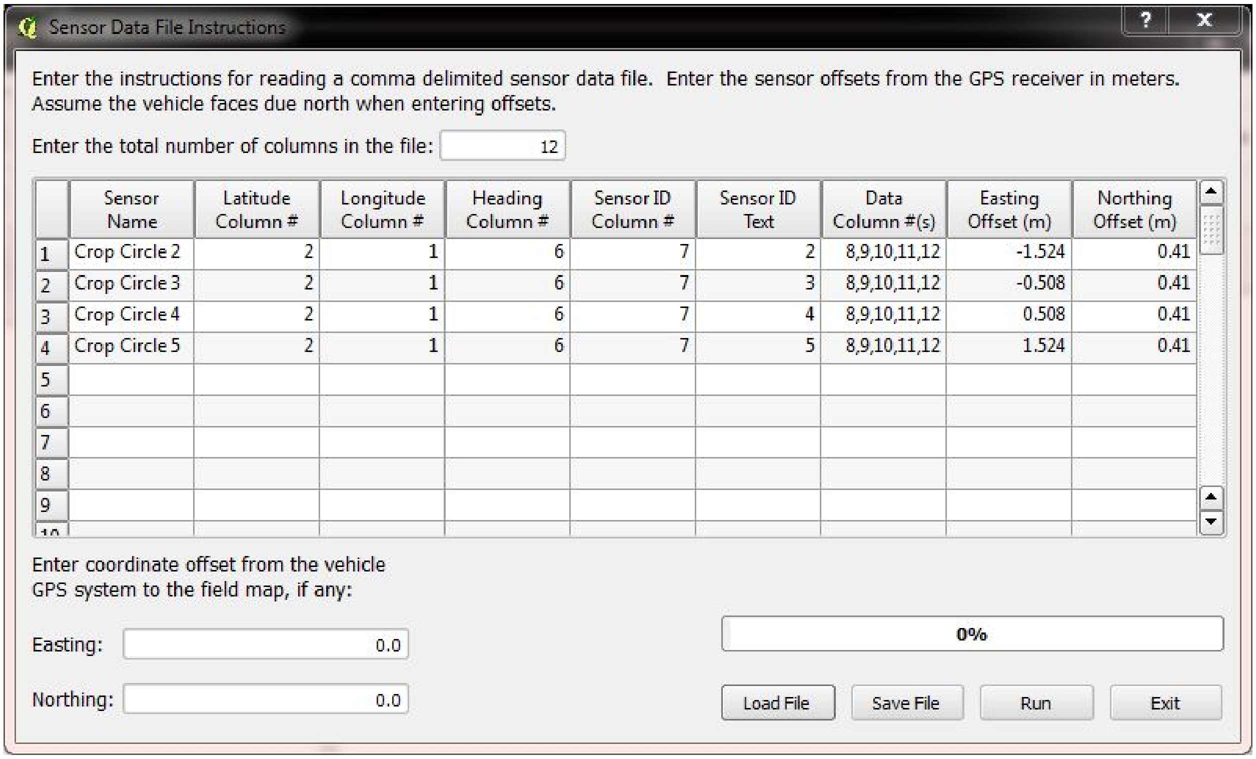


Figure 3. The graphics user interface for the Preprocessor tool.

Geoprocessor

The main purpose of the Geoprocessor is to analyze sensor data within plot boundaries, delimited by rectangular polygons. Thus, the primary requirement for the Geoprocessor is the availability of a plot boundary map, loaded as a polygon shapefile within the GIS. For example, a shapefile created with the Map Creator tool could be used. The Geoprocessor permits the calculation of summary statistics for sensor data collected within each plot boundary, and the statistics are appended to the plot boundary layer as attributes of individual polygons. Contrasted to the default geospatial tools commonly provided with GIS software, the Geoprocessor iteratively summarizes the sensor data independently for each feature in the plot boundary layer, appending the results to the layer. The Geoprocessor also permits the assignment of a plot information to each sensor data point, and the plot names are appended to the sensor data layer. This functionality is particularly important to prepare the data for subsequent genetic analysis. If the only desired geoprocessing task is to assign plot information to sensor information, the default Vector Geoprocessing Tools in Quantum GIS can be used for the same purpose, likely with some advantages in computational efficiency. The Geoprocessor was originally designed for a different application, but it is included here in case it is useful. Other statistical calculations can be added on request.

To use the Geoprocessor tool (Fig. 4), complete the following steps in the Quantum GIS environment.

1. Select Plugins->HTP Geoprocessor-> Geoprocessor
2. Select the processing objective from the list of available options, as follows:
    a. Find mean value of points (process layer) within polygons (base layer).
    b. Find median value of points (process layer) within polygons (base layer).
    c. Find maximum value of points (process layer) within polygons (base layer).
    d. Find minimum value of points (process layer) within polygons (base layer).

   e. Find (area-weighted) mean value of polygons (process layer) within polygons (base layer).
   f. Find maximum area polygon (process layer) within polygons (base layer). This option returns the attribute from polygon having the maximum area within the base layer polygons. If two or more polygons have the same attribute value, their areas are summed.
   g. Add attributes of polygons (process layer) to the points (base layer) falling within the polygon. Likely, this is the option needed for processing the HTP data.
3. Select the base layer polygon shapefile. For processing options 'a' through 'f' above, the combo box will be populated with all polygon shapefiles in the current workspace. For processing option 'g' above, the combo box will be populated with all point shapefiles in the current workspace.
4. Select the layer to be processed. For processing option 'a' through 'd' above, the combo box will be populated with all point shapefiles in the current workspace. For processing options 'e' through 'g' above, the combo box will be populated with all polygon shapefiles in the current workspace.
5. Select the attributes (data fields) from the process layer. For processing options 'a' through 'e' above, the list box will be populated with all numeric data fields in the process layer. For processing options 'f' and 'g' above, the list box will be populated with all text data fields in the process layer.
6. Click Run.

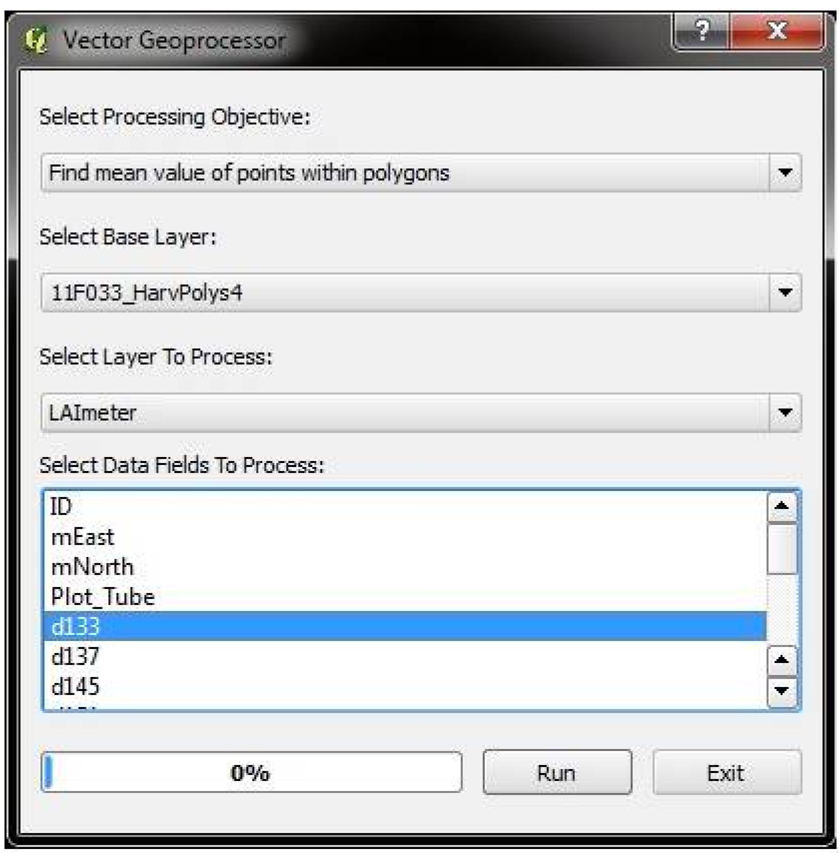


Figure 4. The graphic user interface for the Vector Geoprocessor tool.