QGIS for Precision Agriculture applications

Corentin LEROUX – Aspexit – www.aspexit.com
Lisa ROLLIER – en stage juillet/août 2020

The functionalities presented in this document allow to **manipulate and process Precision Agriculture data with a QGIS plugin**. This plugin "Precision Agriculture" is directly downloadable from the Github account of Aspexit. Once the folder is dezipped, it can be copied and pasted in the right directory so that it can be read by QGIS :

C:\Users\ASUS\AppData\Roaming\QGIS\QGIS3\profiles\default\python\plugins

This folder can be accessed directly within QGIS by clicking on « Settings », then « User profile » and « Open the folder of active user ».

Once the plugin is in the folder, the functions appear in the processing toolbox in the "Precision Agriculture" folder. These functions are coded in Python, can use existing functions in the Processing toolbox, and do not use any Python package that has not been installed beforehand when installing QGIS (to limit the problems of installations and package versions).

This plugin is complemented by more than 40 R functions that can also be downloaded and used within QGIS (see : R codes to be used within QGIS for Precision Agriculture applications)

Some of the few problems that may arise :

- If possible, try to save the outputs of the functions. Temporary layers can sometimes cause problems when use directly in another algorithms. Some of the outputs do not display well in QGIS while they have been really generated.

TABLE DES MATIERES

1. Classification

1.1. R – Classification

Allows to reclassify a raster into a user-defined number of classes using several classification methods

- **Inputs**
  - Raster layer
  - Classification method
    - Equal : Equal-intervals
    - Quantiles
    - K-means
  - Number of classes

- **Outputs**
  - Classified raster

2. Correlation

2.1. V – Correlation indexes

Allows to calculate a correlation index across numerical columns of the input vector layer. A correlation plot is produced.

- **Inputs :**
  - Vector layer
  - Correlation index : Pearson, Kendall, Spearman

- **Outputs :**
  - Correlation plot

3. Data Conversion

3.1. O – Convert Telepac XML files to Shapefile

Convert TelePAC files from XML format to SHP format. Before running the function, there is just one manual manipulation to make on the XML file. For now, Telepac files comes with a tag that is doubled (which should not be possible with XML files). The following HTML tag should be removed : *xmlns="urn:x-telepac:fr.gouv.agriculture.telepac:echange-producteur"*

- **Inputs :**
  - Telepac XML file
- **Outputs :**
  - Folder to which shapefile will be created

4. Data Manipulation

4.1. V – Convex hull creation

Allows to build a convex envelope (contour) around a point vector layer

- **Inputs :**
  - Point layer
- **Outputs :**
  - Contour of the points

4.2. V- Grid oriented in the direction of rows

For a given point layer oriented in a particular direction (following rows or a machine passage for example), the function builds a grid oriented in the majority direction of the points. The size of the grid is defined by the user. Prerequisite: The order of the points in the layer must follow the data acquisition.

- **Inputs :**
  - Point layer
  - Rotation angle
  - Grid resolution
- **Output**
  - Oriented Grid

4.3. V – Tesselations

Allows to build tesselations from point vector data.

- **Inputs :**
  - Point layer
  - Tesselation type : Delaunay, Thiessen, Voronoi
- **Output**
  - Tesselation

5. Filtering

5.1. V- Local univariate filtering

Detects local/spatial outliers for a given field (column) of a vector layer using several filtering methods. Outliers can either be removed or identified in a new column in the vector layer.

- **Inputs :**
  - Vector layer
  - Filtering method
    - CV : Coefficient of variation (inspired from Spekken et al. (2013). A simple method for filtering spatial data
    - IDW : Each point value (Zi) is compared to interpolated value from neighbours (Zj). If Zi is outside the range Zj+/- n standard deviations of neighbourood values, the point is considered an outlier (inspired from Simbahan et al., 2004 : Screening yield monitor data improves grain yield maps)
    - Normal : outliers removal assuming normal distribution in the data (mean +/- n standard deviations)
  - Confidence intervals (for the filtering method « Normal »)
  - Number of standard deviations
  - Neighbourhood distance : Points separated by a distance lower than the neighbourhood distance are considered neighbours in space
  - Outliers handling : whether outliers are marked or removed.

- **Output**
  - Filtered layer

5.2. V- Global univariate filtering

Detects global outliers for a given field (column) of a vector layer using several filtering methods. Outliers can either be removed or identified in a new column in the vector layer.

- **Inputs :**
  - Vector layer
  - Filtering method :
    - Normal : outliers removal assuming normal distribution in the data. Outliers are removed if outstide the range mean+/- n standard deviations
    - Tukey's rule : Tukey interquartile method

        - Fix threshold : Threshold can be set by user
    - Threshold range (for the method « fix threshold »)
    - Confidence intervals (for the filtering method « Normal »)
    - Outliers handling : whether outliers are marked or removed.

- **Output**
    - Filtered layer

6. Sampling

6.1. R – Univariate sampling

Allows to sample a user-defined number of points in a raster layer (points are the centers of the selected pixels). Several methods are available:

- **Inputs :**
    - Vector layer
    - Classfication method :
        - Random : objects are chosen randomly in the layer
        - Quantiles : The raster is reclassified into a user-defined number of quantiles, and the points are randomly selected in each of the defined quantiles
        - Equal intervals : The raster is reclassified into a user-defined number of equal intervals, and the points are randomly selected from each of the defined intervals
    - Number of groups/classes
    - Number of samples in each class
- **Outputs :**
    - Sampling points

6.2. V – Sampling within polygon

Allows to carry out a sampling in a polygon. The number of points is defined by the user and several sampling schemes are available

- **Inputs :**
    - Vector layer
    - Sampling scheme : Random, Regular
    - Number of points (approximative for Regular sampling)
    - Minimum distance to edges (to avoid border effects)
    - Minimum distance between points (applies only if clicked)

- **Outputs :**
    - Sampling points

6.3. V – Univariate sampling

Allows to sample a user-defined number of points in a vector layer. Several methods are available:

- **Inputs :**
  - Vector layer
  - Classfication method :
    - Random : objects are chosen randomly in the layer
    - Quantiles : The raster is reclassified into a user-defined number of quantiles, and the points are randomly selected in each of the defined quantiles
    - Equal intervals : The raster is reclassified into a user-defined number of equal intervals, and the points are randomly selected from each of the defined intervals
  - Number of groups/classes
  - Number of samples in each class
- **Outputs :**
  - Sampling points

7. Spatial interpolation

7.1. V – Optimal grid size for interpolation

Defines the optimal size of an interpolation grid for a given field (column) of a vector layer. The user must fill in the components of the theoretical variogram related to the chosen field [from Tisseyre et al. (2018) How to define the optimal grid size to map high resolution spatial data? Precision Agriculture]

- **Inputs :**
  - Vector layer
  - Field contour
  - Variogram parameters (model, range, partial sill, nugget effect)

- **Outputs :**
  - Grid size

7.2. V – Spatial interpolation

Generates the interpolation of a user-defined numeric field (column) of a vector layer. A raster is generated. The user can choose the size of the interpolation grid and the interpolation method. For tesselation methods, a tesselation is first built (each polygon takes the value of the point it contains) and the grid of interpolation is then superimposed on the tesselation.

- **Inputs :**
  - Vector layer
  - Interpolation method : IDW, Thiessen, Voronoi, Delaunay
  - Grid resolution
  - IDW power : power of interpolation (applies solely for IDW method)

- **Outputs :**
  - Interpolated map

8. Zoning

8.1. R – Classification-based zoning

Allows to delimit within-field zones on a raster. The raster is first reclassified using a user-defined number of classes and classification method. The classified raster is then post-processed to build zones with the following steps:

- Modal smoothing of the classes to limit the initial data noise [call GRASS r.neigbours function].
- Vectorization of the raster to construct the zones [call QGIS polygonize function].
- Deletion of surface areas below a given threshold [call GRASS v.clean function].
- Calculation of statistics in each of the delimited zones (mean and standard deviation of the raster in each zone) [call QGIS Zone Statistics function].
- Smoothing of the contour of the zones to improve the rendering [calling GRASS v.generalize function]

Each of these steps can be set by the user. This method is considered more like a spatialized classification method than a zoning method because it is based primarily on a classification method whose outputs are post-processed by spatial filters.

- **Inputs :**
  - Raster layer and contour of the field
  - Classification method : Quantiles, Equal-Intervals, Jenks
  - Number of classes
  - Size of modal filter (must be an odd number)
  - Minimum area of the final zones
  - Smoothing method : Douglas, Douglas reduction, Snakes
  - Smoothing parameters : Maximum tolerance value, look ahead parameter, alpha and beta (for Snakes method), reduction (for Douglas reduction method)

- **Outputs :**
  - Within-field zones

8.2. V – Merging two zonings

Allows to merge two zonings of the same field into one single zoning. The final zoning can be considered as a microzoning that combines the zones of both initial zonings

- **Inputs :**
  - First zoning

   - Second zoning

- **Outputs :**
   - Merged zoning

8.3. V – Zoning summary

Allow to compute several zoning descriptors (mean perimeter of the zones, total perimeters of the zones, density of zones..). The variable selected in the function must be the class of each zone (the data originating the zones must have been classified before).

- **Inputs :**
   - Zoning
   - The variable that contains the class of each zone

- **Outputs :**
   - Descriptors of the zones

8.4. V – Zoning quality index

Allows to calculate a zoning quality indicator in relation to available auxiliary information. The indicator should fill in the zoning vector layer, and the vector layer to which the zoning is to be compared.

- **Inputs :**
- **Inputs :**
   - Zoning
   - Point layer for which the quality of zoning need to be adressed
   - The variable that identifies the zones
   - Index : Variance reduction (from Bobryk et al., 2016 - Validating a Digital Soil Map with Corn Yield Data for Precision Agriculture Decision Support)

- **Outputs :**
   - Zoning quality index