Chapter 4

THE PLUG-IN STRUCTURE AND GUIDANCE

4.1 General

The QGIS plug-in is developed in python language. It is named as 'RS&GIS'. It is a cross-platform plug-in i.e. it can be used in Windows, Linux and Mac operating system inside a QGIS environment. The plug-in is available in the official QGIS repository which makes it easy to download and use in QGIS environment. The source code of the plug-in can be referred and downloaded using the link (a GitHub repository link): https://github.com/PrathamGitHub/NITK_RS-GIS_17.git. This source is free software, anyone can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation.

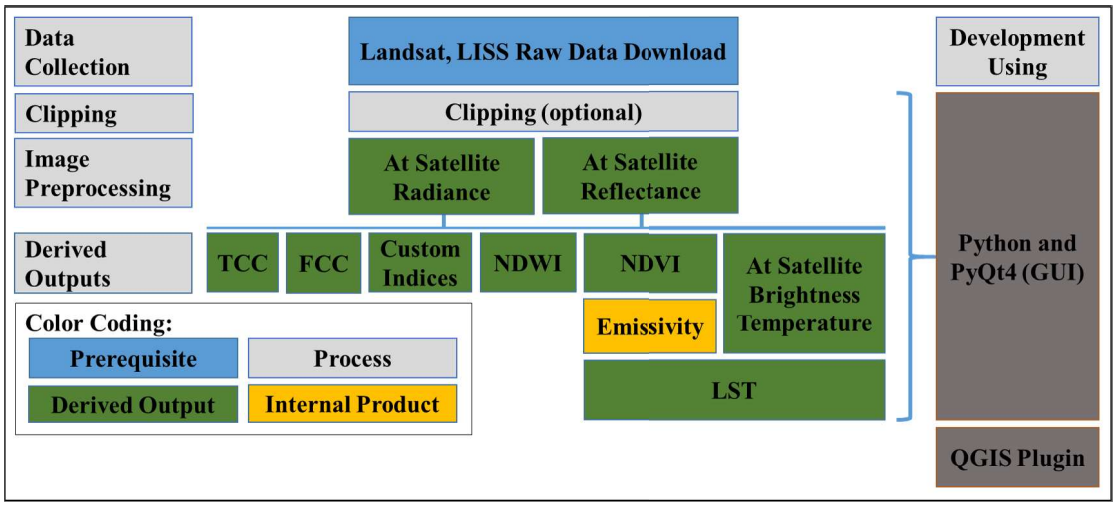

Figure 4.1: The plug-in flow chart

4.2 The methodology adopted

Figure 4.1 shows the flow chart of the plug-in. The only prerequisite for the use of the plug-in is the raw satellite data. The raw Landsat data (Landsat 1 to 8 Level 1 data) can be downloaded using the link (NASA's USGS Earth Explorer website link): http://earthexplorer.usgs.gov/

IRS data (LISS data) can be downloaded using the link (ISRO's Bhuvan website link): http://bhuvan.nrsc.gov.in/data/download/index.php

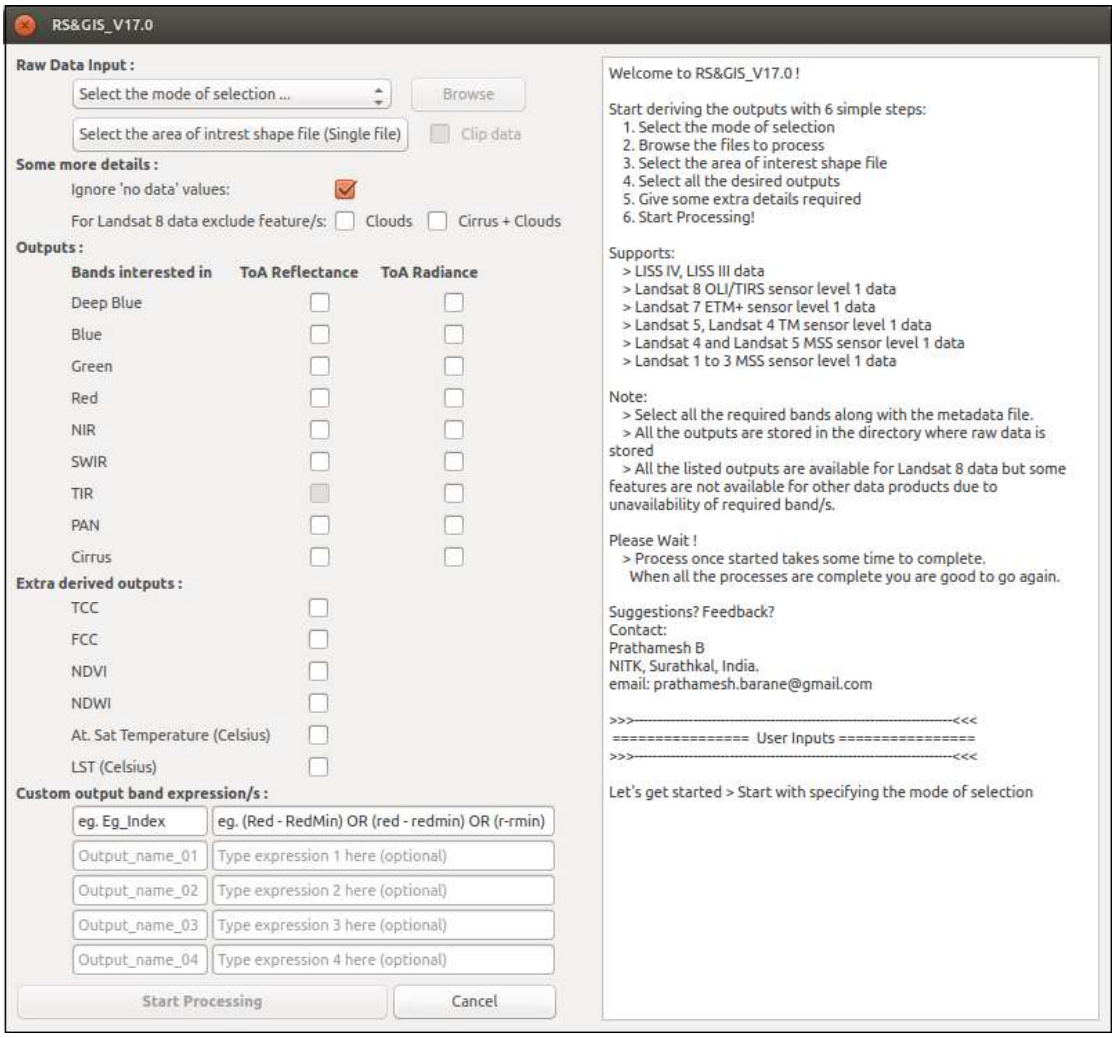

Figure 4.2: The plug-in main window

Fig. 4.2 shows the main window of the plug-in. The main window is divided into six sections: (1) Raw data input, (2) Some more details, (3) Outputs, (4) Extra derived outputs, (5) Custom output band expression/s and (6) Event log window. Event log window is an interactive window that gives the information about the plug-in at the start. After that, it guides the user at every step of giving inputs to the plug-in and when 'Start Processing' button is clicked it gives an interactive summary of all the processes performed in the background in the form of dynamic console outputs. Thus the user gets an idea about the processes performed and time required for the process. In the end of the processes, it shows

total time taken for that particular batch operation. The 'Raw Data Input' section is to input the geospatial data which includes the raw satellite data and the Area of Interest (AoI) shape file. Giving AoI shape file is optional. If the user wants to subset the satellite data to the AoI shape file, the user has to select a single shape file irrespective of its coordinate reference system. The raw satellite data can be input in 3 selection modes: (1) Compressed file/s, (2) Extracted files and (3) Folder containing the extracted data folder/s. To process on single satellite tile selection mode '(2) Extracted files' is preferred where supported file formats are 'tif' and 'txt'. The user has to browse and select all the required band ('tif') files along with its metadata ('txt') file. To perform batch process i.e. to perform operations on multiple satellite tiles, selection mode '(3) Folder containing the extracted data folder/s' is preferred where the user has to select a folder. The selected folder contains one or more than one satellite data sub-folders. Every sub-folder contains extracted files of a downloaded satellite data. But if the user has downloaded satellite data in compressed file format i.e. in 'tar' or 'zip' file format, selection mode '(1) Compressed file/s' is preferred where the user can browse and select single or multiple satellite tiles in compressed file format.

The sections (2) Some more details, (3) Outputs and (4) Extra derived outputs has check boxes in front of every option or functionality available. All the listed functionalities are available for Landsat 8 data but some are not available for other data products due to unavailability of required band/s. The available functionalities include: Ignore 'no data' values, Cloud and Cirrus removal, Top of Atmospheric (ToA) Reflectance, and ToA Radiance for the available bands, True Color Composite (TCC) image, False Color Composite (FCC) image, Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), At-Satellite Brightness Temperature in degree Celsius and Land Surface Temperature (LST) in degree Celsius. The section '(5) Custom output band expression/s' covers almost all the user defined band combinations where the user is given the functionality to input the desired band combination expression/s.

After input of all the mandatory information, the button 'Start Processing'

become operational (active). When the 'Start Processing' button is clicked the plug-in performs the operations on raw data and stores the desired outputs in 'geotiff' file format which can be opened by all the GIS Softwares like ENVI, Erdas Imagine, QGIS, ArcGIS, etc. All these files are stored in the folder where the raw satellite data is stored. The user has to make sure that there is enough space in the directory to store the outputs. The names to the outputs produced are given in such a way that it consists of the type of sensor, path and row information, date of capture of the data, band name and band number information. There is a possibility that the user may commit mistakes in selecting the file or selecting the output/s required but these mistakes can be easily identified by the name of the output file and the user always has the chance to rerun the process with correct input. Therefore, by use of this plug-in possibilities of human error are almost negligible.

4.3 The Plug-in structure

This section describes in brief the internal structure of the Plug-in developed and the order of operations.

4.3.1 Mode of selection

The first step involves giving the location of the files to process. User has to select the mode of selection. It is a drop down menu which has three modes of selection. [Figure 4.2]

- Compressed file/s
- Extracted files
- Folder containing the extracted data folder/s

4.3.1.1 Compressed file/s

When user downloads the raw freely available satellite data e.g. Landsat 8 data from "earthexplorer.usgs.gov" website, it is in ".tar.gz" file format. If user

wants to process the data directly without manually decompressing it, the user can select this option.

- Some of the satellite sensor data are available in ".zip" or other compressed file formats. So all the possible compressed file extensions are considered to facilitate browsing and selection.
- If user selects this option, user can browse and select a single compressed file for single satellite tile or multiple compressed files for multiple satellite image tiles (provided that all the files are stored in same directory).
  - Satellite tile is a geographical area extent which a satellite sensor can capture in a single image.
  - Satellite data is available in the form of tiles, designated by path and row of satellite movement.

4.3.1.2 Extracted files

If user has extracted the raw data from the compressed file and he wants to process only a single tile, he can select this option.

- If user selects this option, he can select the specific band images (in ".tiff" file format) which he wants to process along with the metadata file (in ".txt" file format)

4.3.1.3 Folder containing the extracted data folder/s

User can select this option, if user has extracted the raw data from the compressed file/s and he wants to work on multiple satellite tiles at a time.

- If user selects this option, he has to browse and select a single folder which consists single or multiple folders. Each folder contains extracted data of a single tile.

4.3.2 Browse button

User has to click "Browse" button which opens a browser window using which user can select the files to be processed using the tool. [Figure 4.2]

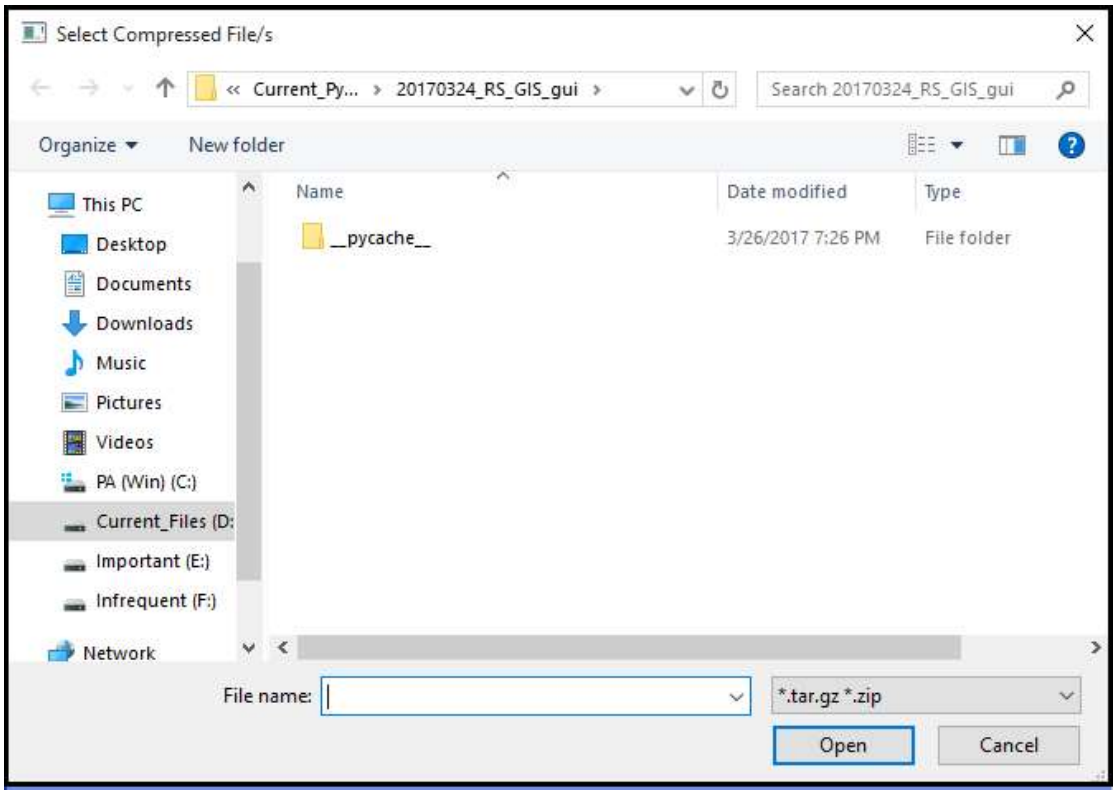

Figure 4.3: Browser window

4.3.3 Select required files (a browser window)

Based on the choice made by user [Figure 4.3]

- "Mode of selection" case [1]: Compressed file/s:

  User has to select single or multiple compressed files and press "Open" button in the browse window.

  - Variable say "browse" stores the location of the folder containing compressed files.
  - Variable say "browseSelected" stores name/s of the selected file/s with their extension as a list.
  - Variable say "browseSelectedExt" stores the extension of one of the compressed file selected.

- "Mode of selection" case [2]: Extracted files:

User has to select ".tiff" file/s of all the required band/s which he wants to process along with the metadata file which is mostly available in ".txt" file format. After selecting the files, he has to click "Open" button on the browse window.

  - Variable say "browse" stores the location of the folder containing the files.
  - Variable say "browseSelected" stores names of all the selected files with their extension as a list.

- "Mode of selection" case [3]: Folder containing the extracted data folder/s:

  User has to select a single folder and press "Open" button on the browse window.

  - Variable say "browse" stores the location of the folder containing extracted data folder/s
  - Variable say "browseSelected" stores name/s of the folder/s contained in the user selected folder as a list.

4.3.4 Gathering the information about the outputs to be processed form the user (multiple check boxes)

Under the heading of "Bands interested in:", user will be given following options: [Figure 4.2]

1. Deep blue
2. Blue
3. Green
4. Red
5. NIR
6. SWIR

7. TIR
8. PAN
9. Cirrus

With every option mentioned above, two check boxes are given:

1. ToA Reflectance
2. ToA Radiance

where, In case of ToA Reflectance, user clicks this if user wants "Top of Atmospheric (ToA) Reflectance raster" of the particular band. Same with the ToA Radiance, user clicks this if user wants "Top of Atmospheric (ToA) Radiance raster" of the particular band.

Under the heading of "Extra derived outputs:", user will be given following options:

1. TCC
2. FCC
3. NDVI
4. NDWI
5. At.Sat.Temperature (Celsius)
6. LST (Celsius)

With every option mentioned above, one check box would be given, to know whether user wants this output or not.

Under the heading of "Some more details:", user will be given following options [Figure 4.2]:

1. Ignore "no data" values: This option is by default checked. If user wants, it can be unchecked.

2. For Landsat 8 data exclude feature/s: This option allows user to remove 'clouds' or 'clouds and cirrus' pixels from the data to avoid their interference while performing the band operations.

4.3.5 A console window

It is interactive window that gives the information about the tool at the start. After that it gives information about the user inputs and when 'Start Processing' button is clicked it gives an interactive summary of all the processes performed in background. Thus user gets an idea about the processes performed and time required for the process. In the end of the processes it shows total time taken for that particular batch operation. [Figure 4.2]

4.3.6 Start Processing (a push button)

Based on all the inputs taken from the user, tool starts to process on the raw data. This button becomes active (ready to be clicked) after tool receives minimum required user inputs. [Figure 4.2]

4.3.7 Process when "Start processing" button is clicked

Tool creates a new folder called "NITK_RSGIS_*data*_*time*" in the folder whose location is the string in variable "browse", i.e. in the folder where all the raw input data is stored.

- If the 'browse' directory already consists a folder named above, tool creates a new name every second in order to avoid overwriting of data.
- Now the variable "browse" replaces the string by the location of the newly created folder.

Extracting the compressed files in case of "Mode of selection" [1] Compressed file/s

- Tool then extracts the selected compressed file/s and stores the extracted data in separate folder/s in that newly created folder ("NITK_RSGIS").

- Name of the folder/s will be same as the name of the compressed file/s without extension.

- Now the variable "browseSelected" stores name/s of the folder/s extracted and the "Mode of selection" case [1] is now same as the case [3].

In case of selection mode [1] and [3], now we have list of the name/s of the folder/s having extracted data. Whereas in case of selection mode [2] the list consists names of all the files (.tiff , .txt) selected by user along with the extension.

A new variable, say "folderFiles" is created which is list of lists. Each list consists of names of all the files present in the folder. First list consists names of files in the folder which is first in the list of "browseSelected" and so on. In case of selection mode [2] "folderFiles" is the list which is same as "browseSelected" list.

Reading the metadata file/s in ".txt" file format and storing all the information available in a variable say "metadata" in the form of list of lists. If any folder does not contain any metadata file, that particular folder will not be considered for the further processing.

- A string in the console: "Particular_folder_name folder does not contain metadata file, so it is not considered for further processing".

- The folder name will be removed from the variable "browseSelected" and list of files in that folder will be removed from the variable "folderFiles".

Creating a new variable called "sensorType". The information is extracted from the metadata file provided with raw satellite data and is then stored in the form of list according to the order of the corresponding folders. For current version of the tool type of sensors considered are:

1. Landsat 8 OLI/TIRS

2. Landsat 7 ETM+

3. Landsat 4-5 TM

4. Landsat 4-5 MSS
5. Landsat 1-3 MSS
6. LISS III
7. LISS IV

If the folder contains the data captured by sensor other than the sensors listed above, that particular folder will not be considered for further processing.

- A string in the console: "Particular_folder_name: sensor type is not listed, so it is not considered for further processing".
- The folder name will be removed from the variable "browseSelected" and list of files in that folder will be removed from the variable "folderFiles" and metadata is removed from the variable "metadata".

New variable called "metadataRequired" is created. According to type of sensor, extracting all the required information from the metadata file and storing it in order in variable "metadataRequired". Different type of sensor have different format of representing the metadata information. Depending on the particular sensor type, a particular method is used to extract the useful information.

Based on the user requirement, determining which band raster image/s from each folder is/are required to be imported. Creating the empty list of all the required bands and then appending each list by the bands corresponding to the folder.

- Eg. If "red" band is required, then a variable called "redBand" is created which is a list of 2D arrays values in float16 format. First array in the list corresponds to first folder in the "browseSelected" list.
- If a particular folder does not contain red band in it, then in pace of 2D array the particular place is kept empty.

Now a list called "spatialRef" is created, each list consists spatial reference information of any band except PAN band in the corresponding folder. Satellite

data being geospatial data is geocoded. All this spatial reference information is stored, in order to use it at the time of writing the processed raster images, so that the output raster image/s is/are geocoded.

In case if user is interested in "PAN" band, a separate variable called "spatialRefPAN" is created and appended for every folder. Spatial reference of PAN band is different from that of all the other bands due to the fact that it has more pixels for same geographical area extent i.e. it has high spatial resolution.

Now we have

- "browseSelected" containing list of names of folder/s
- "folderFiles" containing list of lists, where each list consists names of the files present in corresponding folder.
- "metadataRequired" containing list of lists, where each list consists the useful metadata information corresponding to the folder.
- "sensorType" containing list which shows the type of sensor used to collect the data in the particular folder.
- "spatialRef" containing spatial reference information corresponding to the folder.
- If required "spatialRefPAN" containing spatial reference information of PAN band corresponding to the folder.
- All the required bands stored as separate variables, where each variable contains list of 2D array.

Now separate class for separate sensor type is created. Various methods are created for various operations for getting :

- Top of atmospheric reflectance
- Top of atmospheric radiance
- True color composite

- False color composite
- Normalized difference vegetation index
- Normalized difference water index
- At satellite brightness temperature
- Land surface temperature

  All the above processes involve applying predefined well established algorithms to every element value of the 2D array (pixel value) to get the desired output 2D array which then can be written as a raster image. If bands required for particular operation is not present in the folder and therefore not stored in the particular band variable, console shows the line "Insufficient data available for performing name_of_operation operation for name_of_folder"

- Write the processed 2D array to a geocoded raster image.
  - In "NITK_RSGIS" folder, a new folder called "Outputs" is created.
  - Inside the "Outputs" folder, separate folder for separate input folder is created by the name of the input folder which is taken from "browseSelected" variable.
  - Every specific output is stored in the particular folder to which it belongs.
  - When output image is written, it is geocoded by assigning the spatial reference information to the raster image along with the pixel values. Spatial reference information for the particular image corresponding to a particular folder is retrieved from the variable "spetailRef" and or "spatialRefPAN".

In the end objects of the classes are created based on the type of sensor, such that it directly takes all the required input arguments from the variables available.