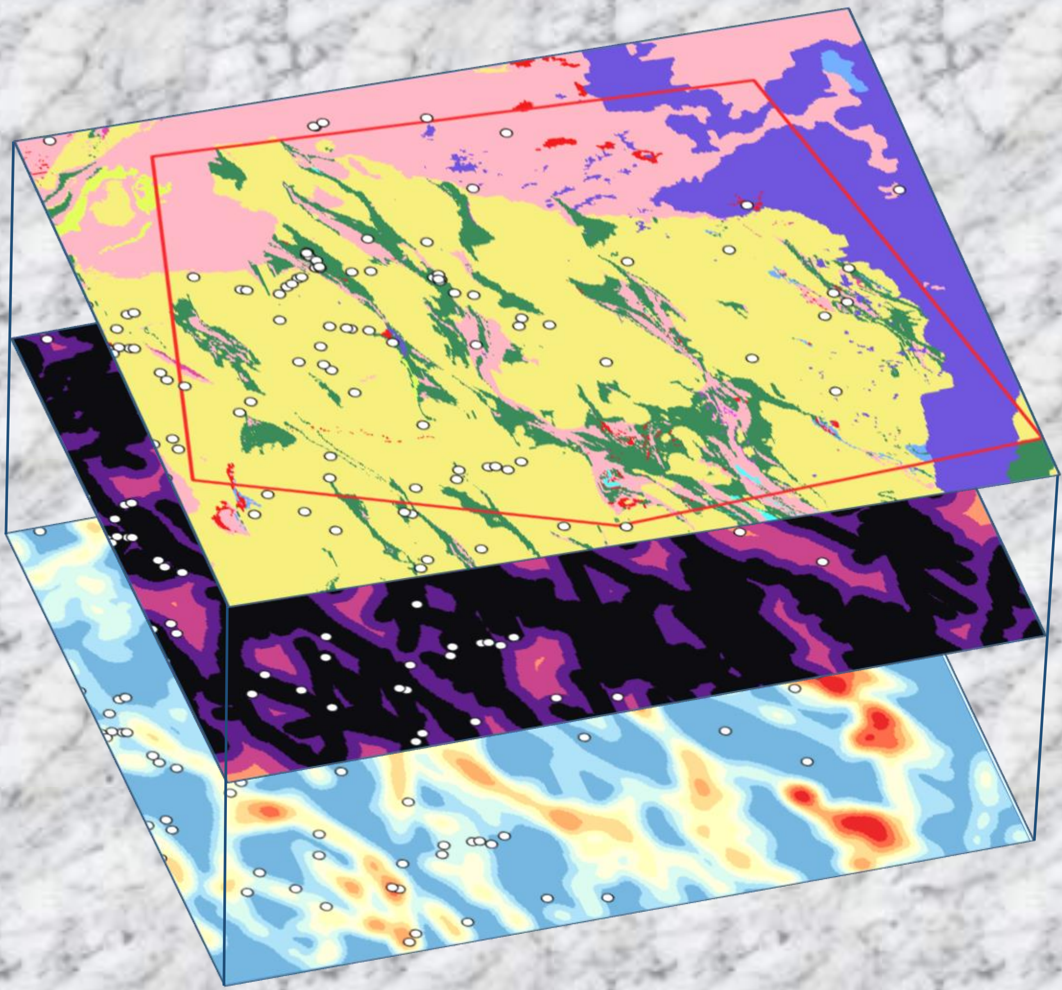
Open Geospatial Information Infrastructure for Research
Academy of Finland Project

QGIS - Plugin for Weights of Evidence Modeling

User Manual and Installation Guide

Bijal Chudasama[a]*, Johanna Torppa[b], Vesa Nykänen[c]

[a] Geological Survey of Finland, Espoo, Finland; [b] Geological Survey of Finland, Kuopio, Finland; [c] Geological Survey of Finland, Rovaniemi, Finland.



* Bijal.chudasama@gtk.fi, www.gtk.fi

May 2020

List of Figures

1. Introduction

The ‘weights-of-evidence’ (WofE) model involves quantification of spatial association (i.e. **the weights**) between the targeted event (the hypothesis) and the evidential event, (**the evidence**). It is a Bayesian method to estimate the probability of a hypothesis (H) based on the knowledge of occurrence of certain evidential events (E). First the prior probability of the hypothesis is calculated and then the same is updated to posterior probabilities of the hypothesis given the occurrence of evidential events.

This model was initially applied to the field of medical diagnosis (Lusted, 1968; Aspinall and Hill, 1983; Spiegelhalter and Knill-Jones, 1984; Reggia and Perricone, 1985; Heckerman et al., 1992). It was then modified for application to mapping the potential of occurrence of mineral deposits by Bonham-Carter et al. (1988) and (1989). Bonham-Carter (1994) provides the details of the WofE approach to exploration targeting of mineral deposits. The WofE model has been widely used for mineral prospectivity modeling (Bonham-Carter et al., 1988; Singh et al.; 1993, Wright and Bonham-Carter, 1996; Raines, 1999; Porwal and Hale, 2000; Boleneus et al., 2001; Carranza and Hale, 2002; Porwal et al., 2010; Chudasama et al., 2018; Jia and Wang, 2019; Tao et al., 2019; Wise, 2019; Williams et al., 2020 - some of the earliest and recent publications). Methods for implementation of the WofE model in GIS environment is described by Agterberg (1989), Agterberg et al. (1990), and Bonham-Carter and Agterberg (1990).

The release of the ArcSDM toolkit (Kemp et al., 2002) allowed implementation of the WofE model in ArcGIS, a commercial GIS-software. The ArcSDM toolkit was developed and released by the U.S. Geological Survey and Geological Survey of Canada (Kemp et al. 2002), and is currently maintained by the Geological Survey of Finland (http://projects.gtk.fi/mpm/ArcSDM5/, https://github.com/gtkfi/ArcSDM). The WofE plugin, discussed in this document, implements the WofE model in QGIS, an open source GIS software. It is developed as a QGIS-equivalent of the WofE modeling tools available in the ArcSDM toolkit. It is created particularly for mineral prospectivity modeling, hence this document will explain plugin implementation, mathematical calculations and the results from the mineral prospectivity modeling perspective. However it can be applied to predictions of other spatial events such as weather forecast, landslide occurrence, forest fires, climate changes, habitat loss, wildlife extinction, infrastructure planning, land-use assessments-planning, oil and natural gas exploration etc. subject to availability of relevant datasets.

2. The WofE Model – Bayes’ Theorem

The Bayes’ Theorem forms the core of the WofE approach. The Bayes’ theorem is about calculations of conditional probabilities. Conditional probability is the probability that an event will occur given the knowledge that another event has already occurred. Conditional probability that can be **calculated** from the data is called the **‘Likelihood’**. And conditional probability that is required to **be estimated** from the ‘Likelihood’ calculations is referred to as ‘**posterior probability’**. Bayes’ theorem states that posterior probability of a hypothesis H, given an evidential event E, can be estimated from the prior belief (the unconditional probability) of the hypothesis combined with the likelihood (conditional probability) and the marginal probability of the evidential event. It can be expressed as below:

Pior Probability *Likelihood*

Posterior Probability

$$P\ (H|E) = P\ (H)\frac{P(E|H)}{P(E)} \qquad (1)$$

Marginal probability

where, *P (H|E) = Posterior probability of hypothesis, H, given the presence of the evidential event E,*
P (H) = Prior belief/ prior (unconditional) probability of the hypothesis H,
P (E|H) = Likelihood (conditional probability) of the evidential event E given the presence of the hypothesis H, and
P (E) = Marginal probability of evidential event E.

In the above equation the terms P (H|E) and P (E|H) both represent conditional probabilities. P (H|E) is unknown and needs to be estimated while P (E|H) can be calculated from known data. Hence P (H|E) is called the posterior probability and P (E|H) is referred to as likelihood.
For multiple evidential events (E_1 - E_n), the Bayes' equation can be derived from the generalized product rule of probability of conditionally dependent events:

$$P(H|E_1, E_2, \cdots\cdots, E_n) = \frac{P(H) \times P(E_1|H) \times \cdots\cdots \times P(E_n|E_{n-1}, \cdots\cdots, E_2, E_1)}{P(E_1) \times P(E_2|E_1) \times \cdots\cdots \times P(E_n|E_{n-1}, \cdots\cdots, E_2, E_1)} \quad (2)$$

However, the influence of each new evidence is conditional on all previously available evidence, which makes the problem non-deterministic in polynomial-time (or np hard). We can overcome this difficulty by assuming that all evidences are conditionally independent, in which case the multiple updating of a-priori probability using Bayes' equation reduces to:

$$P(H|E_1, E_2, \cdots\cdots, E_n) = P(H) \prod_{i=1}^{n} \frac{P(E_i|H)}{P(E_i)} \quad (3)$$

Probabilities can also be expressed in terms of 'Odds'. Odds of the hypothesis is the ratio of the probability of the hypothesis present, P (H), and the probability of the hypothesis absent, *P* $(\bar{H})$. Eq. 3 above expresses the posterior probability of the presence of the hypothesis H, given the presence of evidential events E_i to E_n. Similarly we can derive an equation for the posterior probability of absence of the hypothesis, i.e $\bar{H}$, given the evidential events E_i to E_n:

$$P(\bar{H}|E_1, E_2, \cdots\cdots, E_n) = P(\bar{H}) \prod_{i=1}^{n} \frac{P(E_i|\bar{H})}{P(E_i)} \quad (4)$$

The odds of the hypothesis H is calculated by dividing Eq. 3 and Eq. 4:

$$O(H|E_1, E_2, \cdots\cdots, E_n) = O(H) \prod_{i=1}^{n} \frac{P(E_i|H)}{P(E_i|\bar{H})} \quad (5)$$

Odds take into consideration the presence as well as the absence of the hypothesis given the presence of an evidential event. Therefore odds can be used to assess the spatial association between the hypothesis and the evidential events.

Taking natural logarithms on both sides the odds equation (Eq. 5) becomes:

$$log_e O(H|E_1, E_2, \cdots\cdots, E_n) = log_e O(H) + \sum_{i=1}^{n} log_e \frac{P(E_i|H)}{P(E_i|\bar{H})} \quad (6)$$

In the above equation the term $log_e \frac{P(E_i|H)}{P(E_i|\bar{H})}$ defines the 'weights-of-evidence' of the evidential event E_i Similarly for n evidential events the final odds of the hypothesis given the presence of evidential events E_i to E_n are calculated by addition of the 'weights-of-evidences' of each evidential event E_i and the prior odds of the hypothesis:

$$log_e O(H|E_1, E_2, \cdots\cdots, E_n) = \ log_e O(H) + \sum_{i=1}^{n} W_i \tag{7}$$

The updated posterior probability of the hypothesis given the evidential events E_1 to E_n can finally be calculated as:

$$P(H|E_1, E_2, \cdots\cdots, E_n) = \frac{e^{log_e[O(H|E_1, E_2, \cdots\cdots, E_n)]}}{1 + \ e^{log_e[O(H|E_1, E_2, \cdots\cdots, E_n)]}} \tag{8}$$

3. Mineral Prospectivity Modeling using WofE:

On application of Bayes' Theorem to mineral prospectivity modeling, the hypothesis to be predicted is the occurrence of the targeted mineral deposit (D) and the 'evidential events', (E) are the datasets representing geological features such as lithology, structures, whole rock geochemistry etc. The occurrence of a mineral deposit can be associated with either the presence or the absence of an evidential event. In WofE model the spatial association (i.e. weights) between the deposit and (the presence/absence of) a particular evidential event is assessed by estimation/prediction of the following four posterior (conditional) probabilities using Bayes' Theorem:

i. *P (D|E) = Posterior probability of deposit D given the presence of the evidential event E,*
ii. *P ($\bar{D}$|E) = Posterior probability of absence of the deposit D, given the presence of the evidential event E,*
iii. *P (D|$\bar{E}$) = Posterior probability of deposit D given the absence of the evidential event E, and*
iv. *P ($\bar{D}$|$\bar{E}$) = Posterior probability of absence of the deposit D given the absence of the evidential event E.*

It is mentioned in Section 2, that probabilities expressed in terms of 'Odds', takes into consideration both the presence and the absence of the deposit given the presence of an evidential event. Similarly odds can also be calculated for the deposit given the absence of the evidential event. Using odds calculations of the deposit for the instances of event present and absent, spatial association between the deposit and the evidential event can be quantified. Spatial association can be either positive or negative. Positive spatial association occurs when the deposit occurrence is favored by the presence of the evidential event. In this case the value of odds of deposit given the feature would be higher than the odds of deposit given the absence of the feature. Negative spatial association occurs when the deposit occurrence is favored by the absence of the evidential event. In this case the value of odds of deposit given the absence of feature would be higher than the odds of deposit given the presence of the feature. Hence using the 'Odds' of deposit given the presence of an evidential event and 'Odds' of deposit given the absence of evidential event, we derive at the following equations for quantifying the spatial associations between the deposit and the evidential event:

$P(D\|E) = P(D)\,\frac{P(E\|D)}{P(E)}$ (9)	$P(\bar{D}\|E) = P(\bar{D})\,\frac{P(E\|\bar{D})}{P(E)}$ (10)	$P(D\|\bar{E}) = P(D)\,\frac{P(\bar{E}\|D)}{P(\bar{E})}$ (11)	$P(\bar{D}\|\bar{E}) = P(\bar{D})\,\frac{P(\bar{E}\|\bar{D})}{P(\bar{E})}$ (12)
$P(D\|E)$ = Posterior probability of deposit 'D' given the presence of the evidential event 'E'	*$P(\bar{D}\|E)$ = Posterior probability of absence of the deposit 'D' given the presence of the evidential event 'E'*	*$P(D\|\bar{E})$ = Posterior probability of deposit 'D' given the absence of the evidential event 'E'*	*$P(\bar{D}\|\bar{E})$ = Posterior probability of absence of the deposit 'D' given the absence of the evidential event 'E'*
$P(D)$ = Prior (unconditional) probability of the presence of the deposit 'D'; $P(\bar{D})$ = Prior (unconditional) probability of the absence of the deposit 'D'			
$P(E\|D)$ = Likelihood (conditional probability) of the evidential event 'E' given the presence of the deposit, 'D'	*$P(E\|\bar{D})$ = Likelihood (conditional probability) of the evidential event 'E' given the absence of the deposit, 'D'*	*$P(\bar{E}\|D)$ = Likelihood (conditional probability) of the absence of the evidential event 'E' given the presence of the deposit, 'D'*	*$P(\bar{E}\|\bar{D})$ = Likelihood (conditional probability) of the absence of evidential event 'E' given the absence of the deposit, 'D'*
$P(E)$ = Marginal probability of evidential event 'E'		*$P(\bar{E})$ = Marginal probability of the absence of the evidential event 'E'*	
Equations 9 and 10 are used to calculate spatial association in the instance of the event present		***Equations 11 and 12 are used to calculate spatial association in the instance of the event absent***	

O (D|E) = Odds of deposit given E

Eq. 9 divided by Eq. 10 = O (D|E)

$$O(D|E) = \frac{P(D|E)}{P(\bar{D}|E)} \qquad (13a)$$

Substituting Eq. 9 in the numerator and Eq. 10 in the denominator of Eq. 13a:

$$O(D|E) = \frac{P(D)\frac{P(E|D)}{P(E)}}{P(\bar{D})\frac{P(E|\bar{D})}{P(E)}} \qquad (14a)$$

Rearranging the terms in Eq. 14a

$$O(D|E) = \frac{P(D)}{P(\bar{D})} \times \frac{P(E|D)}{P(E|\bar{D})} \qquad (15a)$$

where, $\frac{P(D)}{P(\bar{D})}$ are prior odds of the deposit = *O (D)*

And $\frac{P(E|D)}{P(E|\bar{D})}$ is the **sufficiency ratio**

$$O(D|E) = O(D) \times \frac{P(E|D)}{P(E|\bar{D})} \qquad (16a)$$

Taking natural logarithms on both sides

$$Log_e O(D|E) = log_e O(D) + log_e \frac{P(E|D)}{P(E|\bar{D})} \qquad (17a)$$

Weights-of-evidences

$log_e \frac{P(E|D)}{P(E|\bar{D})}$ in (Eq. 17a) is the **positive weights of evidence**

$$Log_e O(D|E) = log_e O(D) + W^+ \qquad (18a)$$

O (D|Ē) = Odds of deposit given absence of E

Eq. 11 divided by Eq. 12 = O (D|Ē)

$$O(D|\bar{E}) = \frac{P(D|\bar{E})}{P(\bar{D}|\bar{E})} \qquad (13b)$$

Substituting Eq. 11 in the numerator and Eq. 12 in the denominator of Eq. 13b we get

$$O(D|\bar{E}) = \frac{P(D)\frac{P(\bar{E}|D)}{P(\bar{E})}}{P(\bar{D})\frac{P(\bar{E}|\bar{D})}{P(\bar{E})}} \qquad (14b)$$

Rearranging the terms in Eq. 14b

$$O(D|\bar{E}) = \frac{P(D)}{P(\bar{D})} \times \frac{P(\bar{E}|D)}{P(\bar{E}|\bar{D})} \qquad (15b)$$

where, $\frac{P(D)}{P(\bar{D})}$ are prior odds of the deposit = *O (D)*

And $\frac{P(\bar{E}|D)}{P(\bar{E}|\bar{D})}$ is the **necessity ratio**

$$O(D|\bar{E}) = O(D) \times \frac{P(\bar{E}|D)}{P(\bar{E}|\bar{D})} \qquad (16b)$$

Taking natural logarithms on both sides

$$Log_e O(D|\bar{E}) = log_e O(D) + log_e \frac{P(\bar{E}|D)}{P(\bar{E}|\bar{D})} \qquad (17b)$$

Weights-of-evidences

$log_e \frac{P(\bar{E}|D)}{P(\bar{E}|\bar{D})}$ in (Eq. 17b) is the **negative weights of evidence**

$$Log_e O(D|\bar{E}) = log_e O(D) + W^- \qquad (18b)$$

The net strength of spatial association between the deposit and the evidential event can be measured using 'Contrast', (C). Contrast is calculated from the positive and negative weights as follows:

$$C = W^+ - W^- \qquad (19)$$

Contrast value is positive when the positive weights are higher than the negative weights ($W^+ > W^-$). Similarly negative contrast values occur when the positive weights are lower than the negative weights ($W^+ < W^-$). A high positive contrast implies strong positive spatial association ($W^+ >>> W^-$), while a high negative contrast implies strong negative spatial association ($W^+ <<< W^-$). At contrast value 0, there would exist no spatial association between the deposit and the event.

Significance of Contrast

The above equations are applicable to several conditionally independent evidential events E_n, assuming that each such evidential event contains only one evidential feature class. The weights are calculated for evidential feature class present (Eqs. 13a – 18a) and evidential feature class absent (Eqs. 13b – 18b). Real world datasets representative of evidential events, on the contrary, are seldom binary with a single feature class. These are recorded as multiclass categorical or numerical datasets. Therefore, each evidential event may comprise several feature classes. For instance, a GIS layer of lithology in an area is a categorical nominal multiclass dataset (event). Similarly, distance to a geological feature is a continuous numerical dataset (event). Datasets with continuous numerical values can be converted to multiclass by data binning. But for application of the WofE model to such numerical datasets, it is essential to classify the datasets representing the evidential events into binary classes. And in such cases an appropriate threshold value of maximum influence of the event is needed to convert it to a binary dataset. This can be achieved by reclassifying the data based on the contrast calculations for each feature of the evidential event. Further, calculation of the studentized contrast gives the confidence level for the calculated contrast values. Studentized contrast is calculated by dividing the contrast by the standard deviation of contrast (Bonham-Carter, 1994). High values of studentized contrast indicate less deviations in the contrast and the weights values for that class. Hence the class with high contrast and high studentized contrast values can be used as a threshold for binary reclassification of multiclass datasets. The studentized contrast therefore is an approximate Student T test of the hypothesis that the contrast is not zero. When the contrast is not 0, values >0 imply positive spatial association and values <0 imply negative spatial association. Classes with positive spatial association 'favor' the presence of the deposit while classes with negative spatial association do not 'favor' the presence of the deposit. So accordingly the multiclass numerical datasets are reclassified into binary datasets with two feature classes – 'Favorable' and 'Unfavorable'.

From these binary datasets the posterior probabilities of the deposit given the evidential event can be calculated using Eqs. 3 - 8. And these feature classes are mutually exclusive, i.e., where feature class 'Favorable' is present, feature class 'Unfavorable' is absent, and where feature class 'Unfavorable' is present, feature class 'Favorable' is absent (See Fig. 1 below). Therefore, at a given spatial location either of the feature classes can exist. Final calculations of posterior probabilities from reclassified binary evidential datasets can be limited to the calculations of the positive weights of both classes (Eqs. 13a – 18a). Mathematically, the positive weights for feature class 'Favorable' are equal to the negative weights of feature class 'Unfavorable' and vice versa. The positive weights of evidences are calculated for the generalized 'Favorable' and 'Unfavorable' classes of each evidential event. These are then combined using Eq. 7 and posterior probabilities are calculated according to Eq. 8.

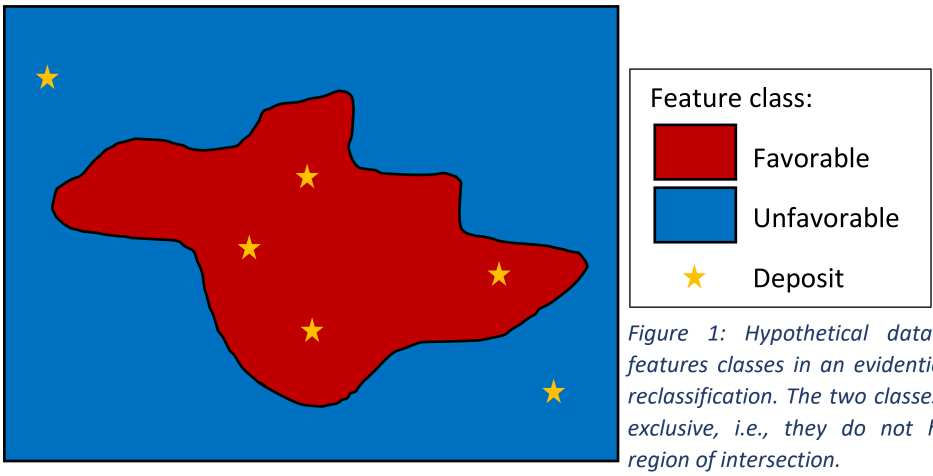


Figure 1: Hypothetical data representing features classes in an evidential event, after reclassification. The two classes are mutually exclusive, i.e., they do not have common region of intersection.

To summarize, mathematically, the WofE model has two parts:

i. Quantifying the spatial association (i.e. weights) between the deposit and the evidential events.
ii. Updating the posterior probabilities of the deposit occurrence by combining the weights-of-evidences of all the events.

And, implementation of the WofE model to mapping the potential of a mineral deposit is accomplished through the following steps in this WofE plugin:

i. Calculation of weights-of-evidences for multiclass (multi-feature) evidential events,
ii. Reclassification of multiclass (multi-feature) evidential events to binary evidential events, based on the weights-of-evidences,
iii. Recalculation of generalized weights-of-evidences after reclassification, and
iv. Calculating posterior probabilities by combining the generalized weights of all the evidential events.

Flow chart in Figure 2 illustrates the generalized approach of the WofE plugin.

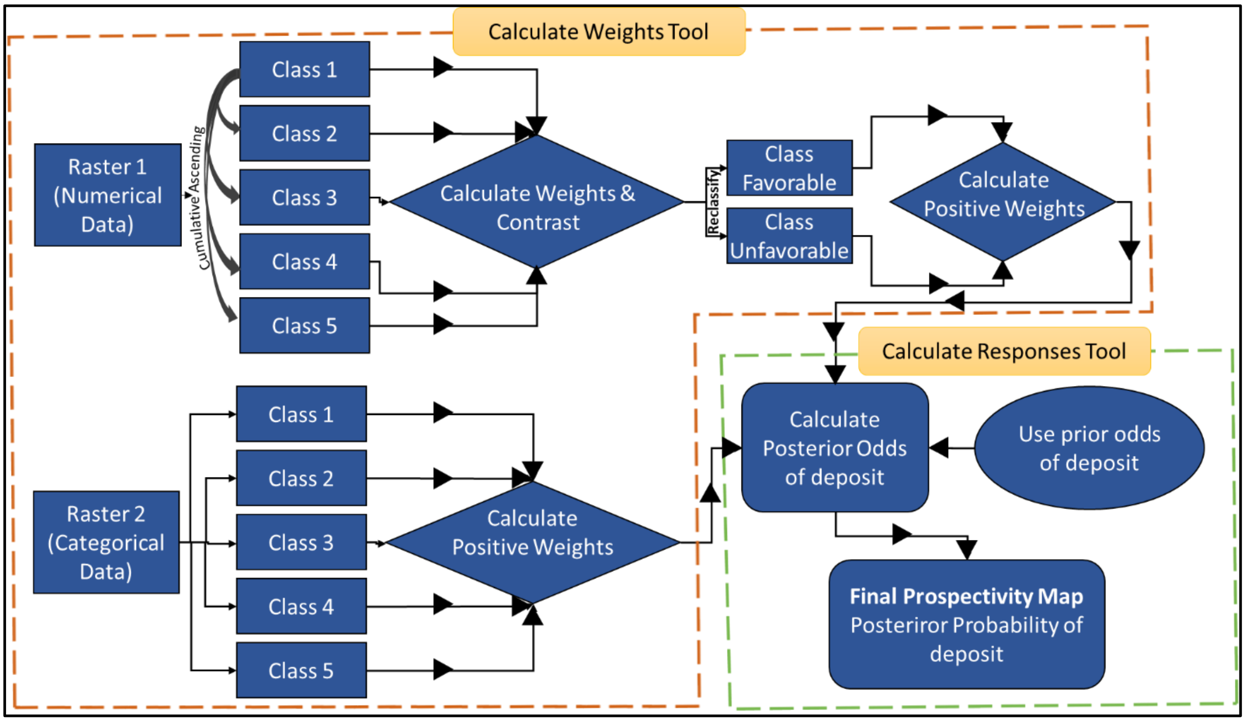


Figure 2: Work flow of the WofE plugin

4. Installation

This pluigin is created and tested for QGIS 3.6 (Python 3.7.0). Python libraries required for calculations in this plugin are:

- GDAL
- Numpy
- Pandas

GDAL and Numpy libraries exist in QGIS. However the ‘pandas’ library needs to be installed separately and imported in to the QGIS platform. If the user does not have ‘pandas’ library installed for QGIS then it is recommended to add to QGIS the one provided in the installation package of this plugin (more instructions as below).

4.1 Plugin installation files

The files included in the installation package are as follows:

i. wofe_module.zip: Plugin installation file.

ii. Pandas.zip: Pandas library, downloaded from: https://pandas.pydata.org/
iii. Sample_data.zip: Test data for plugin implementation.
iv. Results.zip: Results from the plugin (for sample datasets provided for plugin demonstration).

It is recommended to save the wofe_module and Pandas folders as their respective .zip files (i.e. wofe_module.zip and Pandas.zip)

4.2 Installing the WofE plugin

i. In a QGIS project go to 'Manage and Install Plugins' from the 'Plugins' Menu

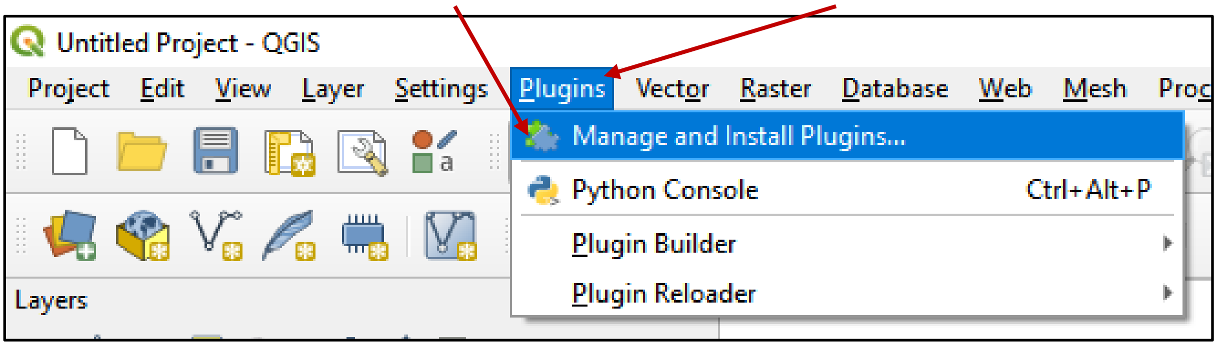


Figure 3: WofE plugin installation – Plugins Menu

ii. In the **'Plugins'** window that opens, select the **'Install from Zip'** option. Navigate to the directory where the installation package is saved and select the **'wofe_module.zip'** file. Click **'Install Plugin'**.

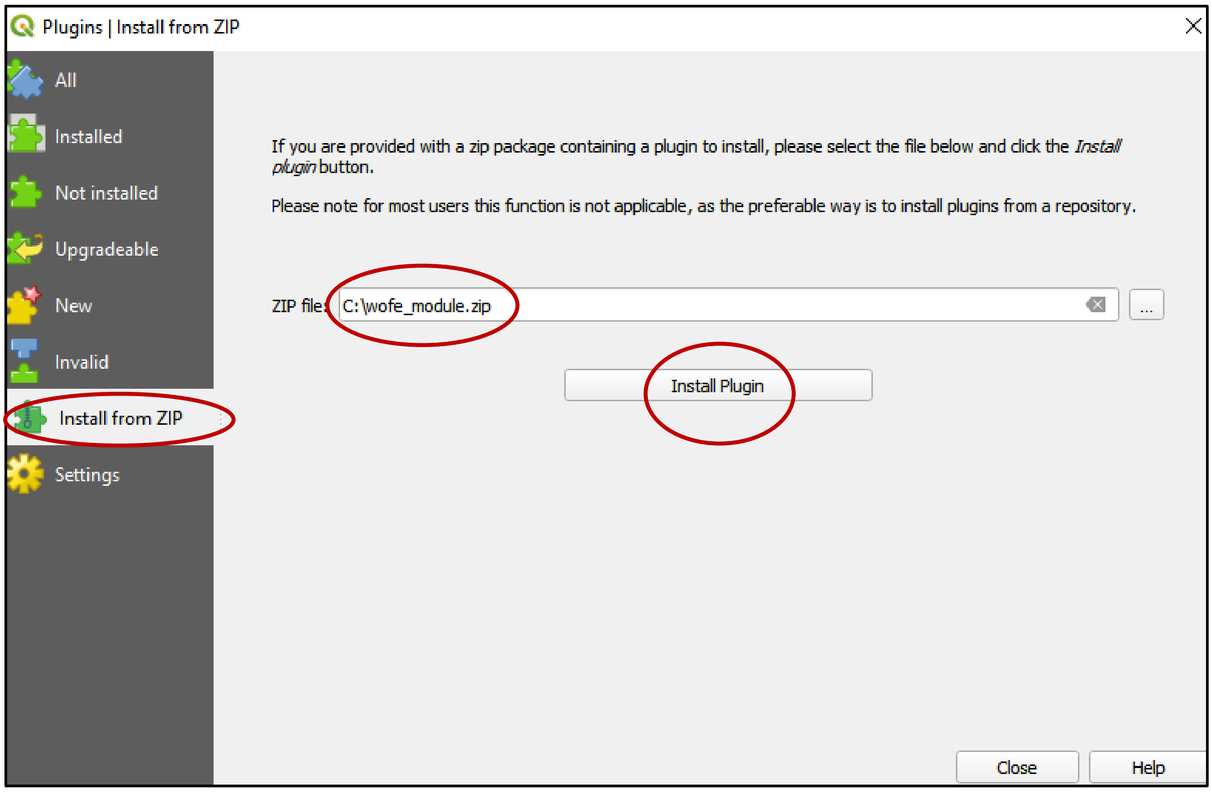


Figure 4: WofE plugin installation – Plugins Window; Installing the plugin using wofe_module.zip

iii. QGIS will issue a **security warning** (Fig. 5). This is because this plugin is not yet uploaded to QGIS plugins repository. This plugin is in experimental stage and will soon be uploaded to the QGIS plugins repository. If the user wishes to use this plugin then they need to continue with plugin installation by selecting **'Yes'**.

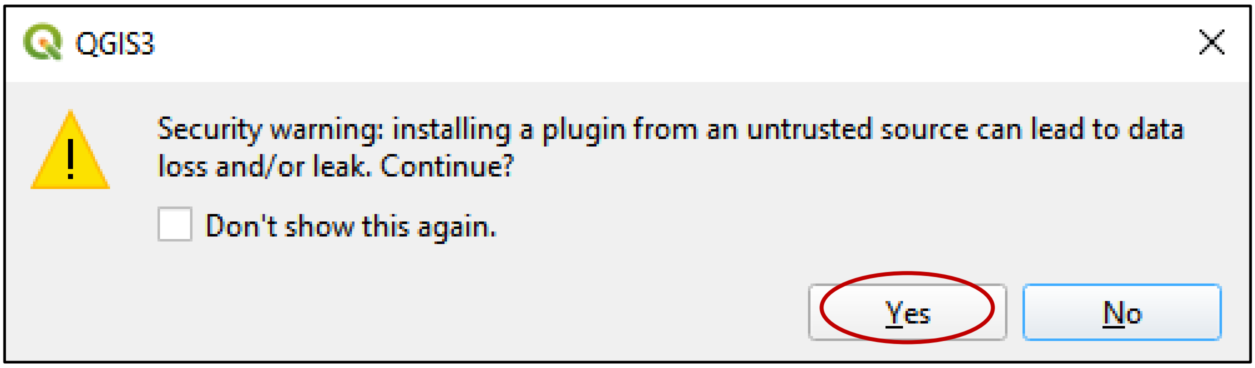


Figure 5: WofE plugin installation – Security warning

iv. On successful installation of the plugin the below message (Fig. 6) is displayed on the window.

Figure 6: WofE plugin installed successfully

v. Close the 'Plugins' window. WofE plugin installation is complete.

vi. In the QGIS main interface, the plugin should now appear in the **'Raster' menu**. It is called the **'Weights of Evidence (WofE) Model'.** (Fig. 7)

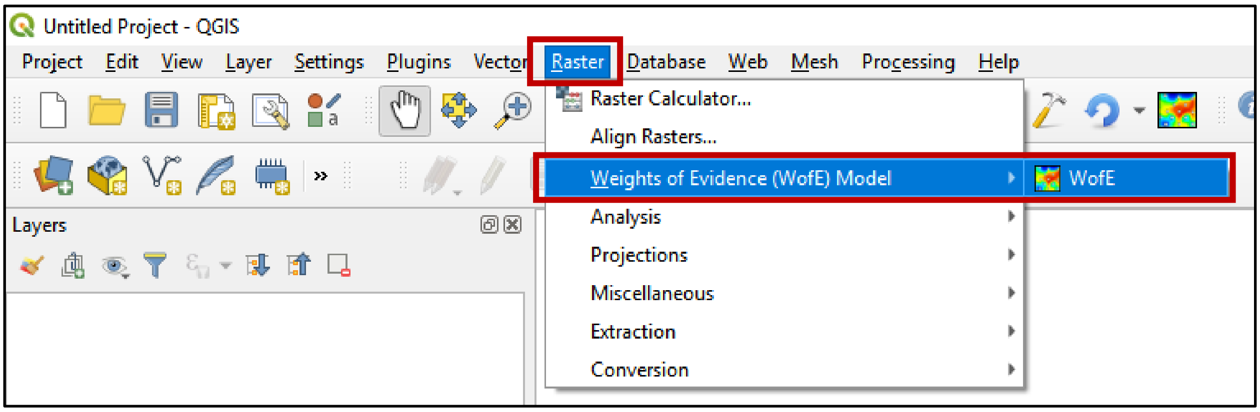


Figure 7: Raster Menu in QGIS contains the WofE plugin.

vii. It can also be accessed directly from the **'Plugins' toolbar**. If the **'Plugins' toolbar** is not loaded to the QGIS interface, it can be added as shown in Figure 8

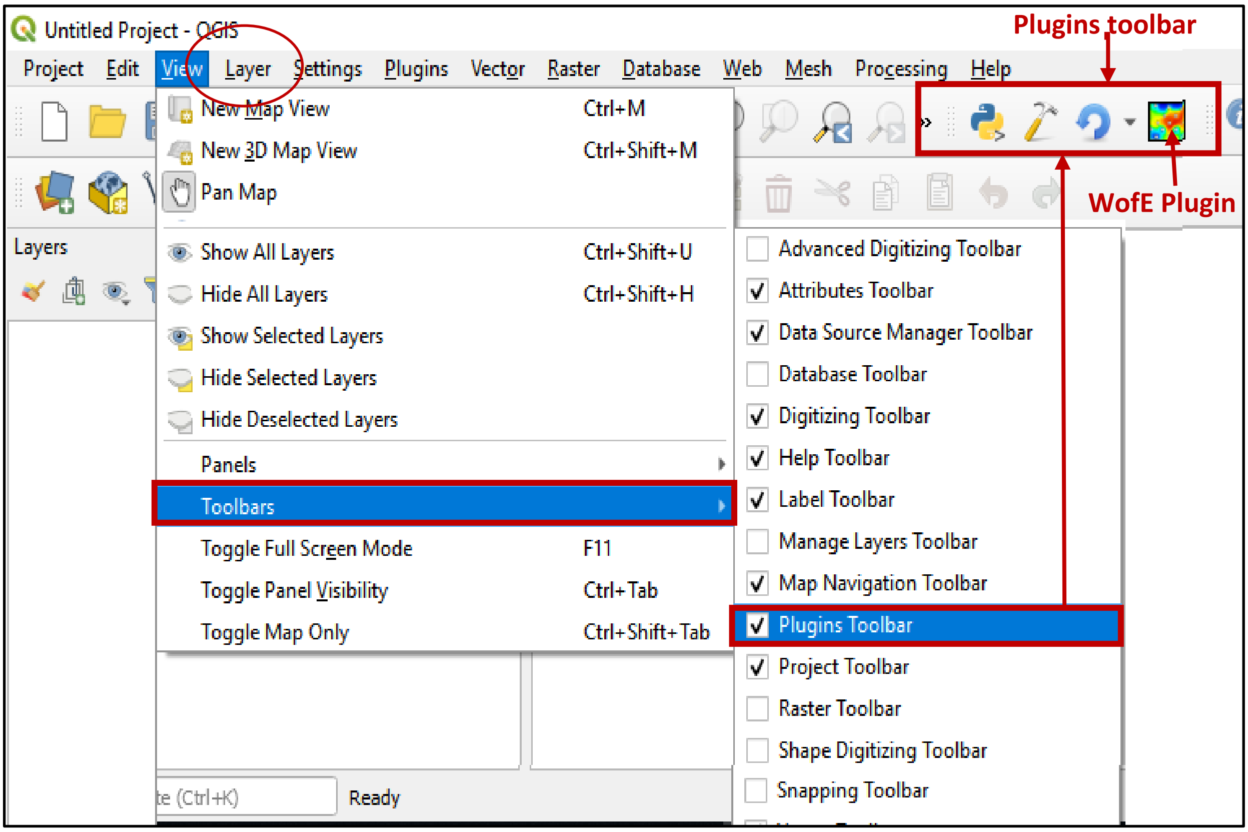


Figure 8: Adding the Plugins Toolbar. WofE plugin appears in the Plugins Toolbar.

viii. However if the plugin doesn't appear in the 'Raster' menu and in the 'Plugins' toolbar, **ensure that the plugin is activated in the 'Manage and Install Plugins' window.** Open the 'Manage and Install Plugins' window (Fig. 3). In the **'Installed'** tab of the **'Plugins'** window look for **'Weights of Evidence (WofE) Model'** plugin. Tick the checkbox next to it. (Fig. 9)

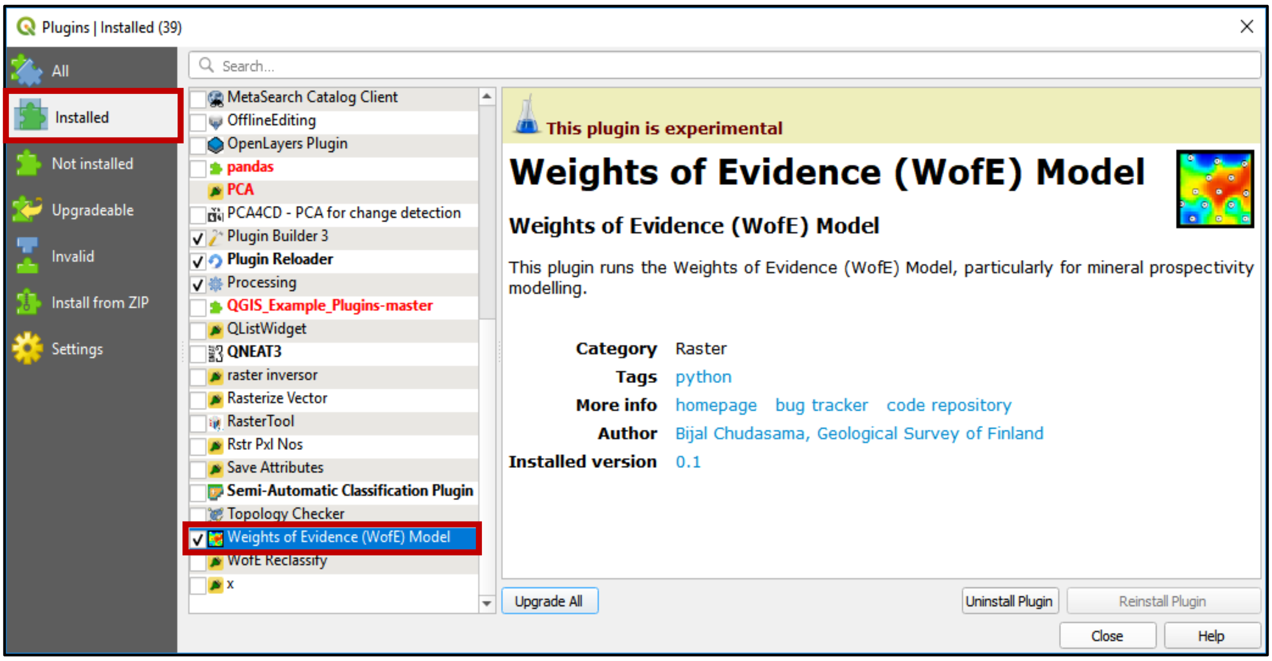


Figure 9: WofE plugin – Activation in Plugins window

ix. If the plugin still does not appear in the Raster Menu then restarting QGIS is recommended to load the newly installed plugin.

4.3 Installing ‘Pandas’ library for QGIS

i. Repeat steps (i) and (ii) of WofE plugin installation (Fig. 3 and 4) for **‘pandas.zip’** file: Go to ‘Manage and Install Plugins’ from the ‘Plugins’ Menu. In the **‘Plugins’** window, select the **‘Install from Zip’** option. Navigate to the directory where the installation package is saved and select the **‘pandas.zip’** file. Click **‘Install Plugin’**.

ii. QGIS will again issue a **security warning** (Fig. 5). This is because ‘pandas’ library is being installed from the .zip file. However it is not available in the QGIS plugin repository for direct installation. If the user wishes to use this plugin then they need to continue to ‘pandas’ library installation by selecting **‘Yes’**. (Alternatively, if users are familiar with python packages installation and QGIS plugins then the users can install pandas library on their computers using ‘pip install’ command and then add the ‘pandas’ folder to the QGIS plugins folder.)

iii. After the pandas files have been extracted, QGIS will issue a **‘Python error’ or ‘python warning’** (Fig. 10)**.** This error/warning can be ignored. Close the error window. The next steps illustrate how to check if the ‘pandas’ library is installed successfully.

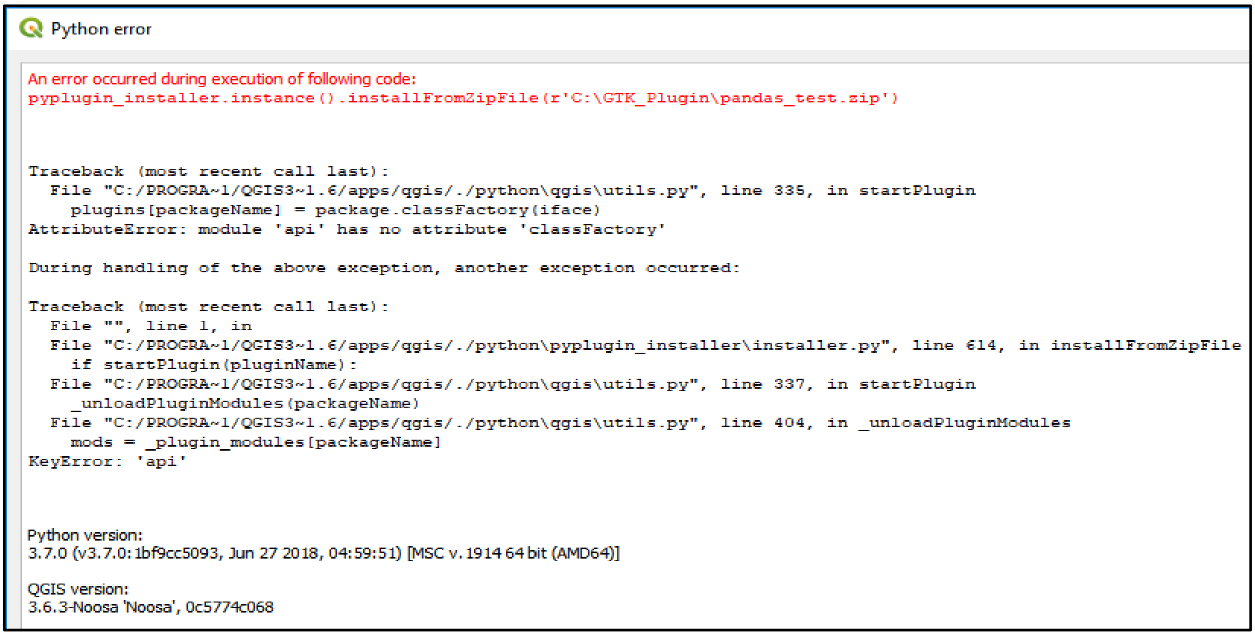


Figure 10: Pandas library installation – Python error/warning.

iv. To check if the library is successfully extracted go to QGIS plugins directory (Fig. 11):

Settings – User Profile – Open Active Profile Folder

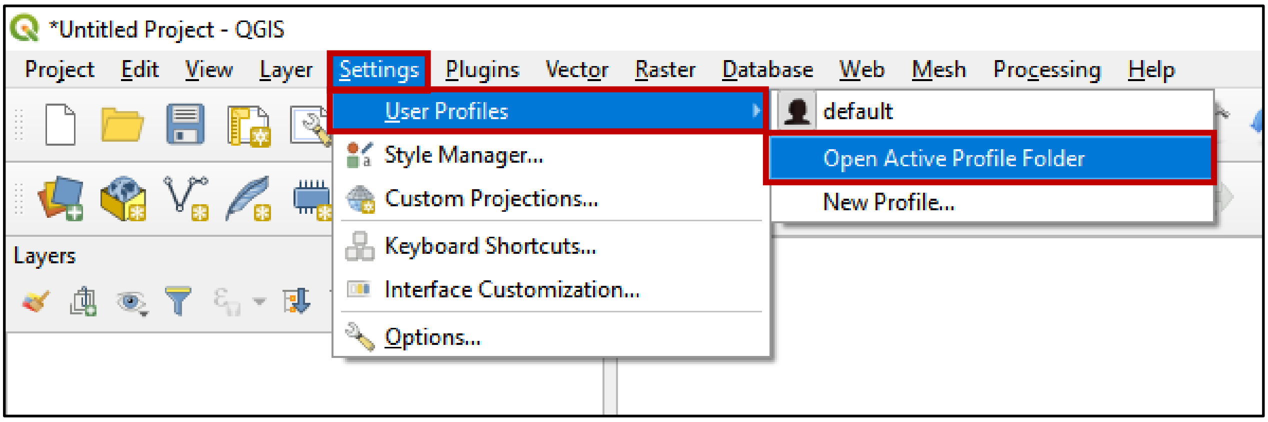


Figure 11: QGIS path

v. The directory where QGIS is installed in the computer is opens up. The path will be like: C:\Users\Admin\AppData\Roaming\QGIS\QGIS3\profiles\default

vi. From here navigate to the plugins directory as follows: C:\Users\Admin\AppData\Roaming\QGIS\QGIS3\profiles\default**python\plugins.**

vii. Check if a **'pandas'** folder is created in the '**plugins'** directory. If such a directory exists then 'pandas' library is successfully added to QGIS.

viii. Next, open **QGIS python console** (Fig. 12)**.**

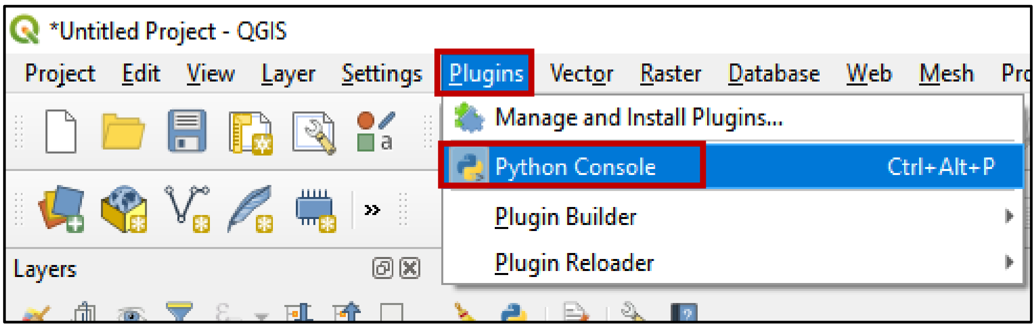


Figure 12: QGIS- Open Python Console

ix. In the python console type the following and click 'enter' on the keyboard (Fig. 13)

import pandas

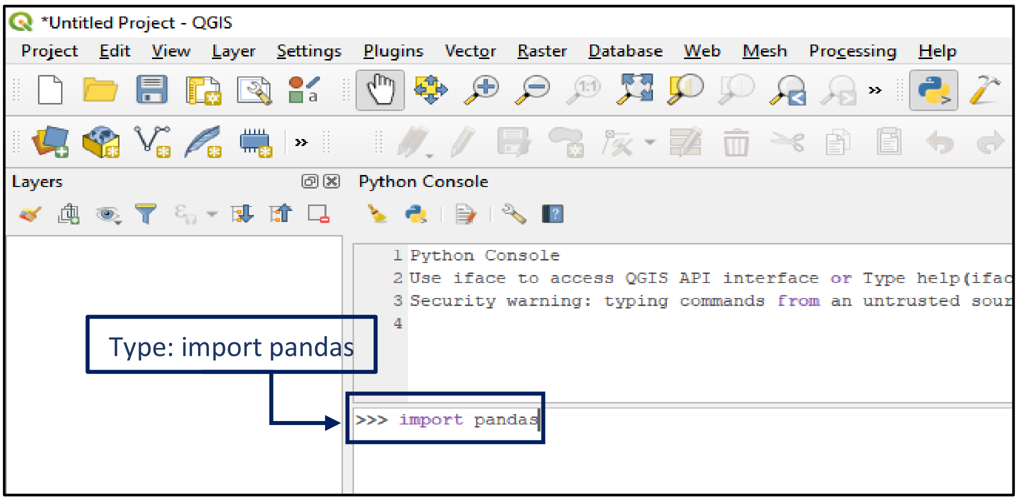


Figure 13: QGIS Python Console –typing the code: import pandas

x. The code does not give any error (Fig. 14). This means that the library is successfully imported to QGIS. Now the plugin is ready for use.

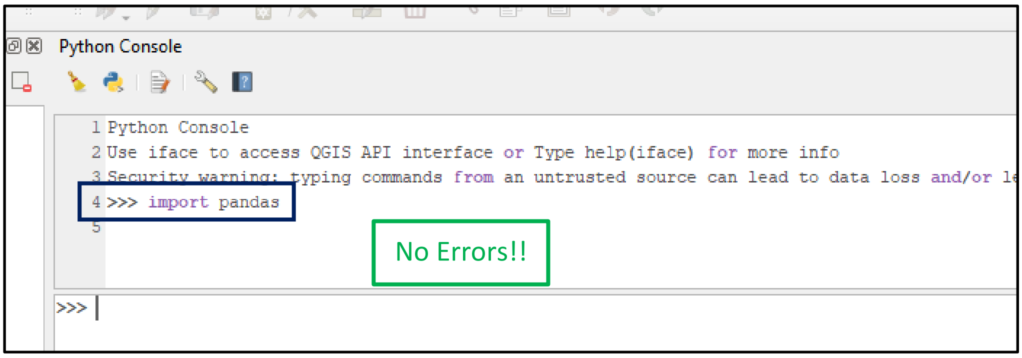


Figure 14: QGIS Python Console – importing pandas library

5 Running the WofE Plugin

5.1 Data requirements

Table 1 below describes the data requirements and pre-requisites:

Input data	Format	Description
Training Sites	Point Shapefile (.shp)	This data is spatial locations of the known deposits/occurrences represented as points.
Extent of analyses	Polygon Shapefile (.shp)	This layer defines the extent of the study area. All calculations are performed within the area of this layer.
Evidential rasters	Multiclasss GeoTIIF (.tif)	Rasters representing the evidence events or patterns with which the spatial association of the training sites is assessed.

- All input data should have the same projected co-ordinate system.
- There should not be any missing or no-data values in the evidence rasters.
- The numerical evidence rasters should be multiclass or binary (not continuous).
- The input **'Training Sites'** layer should have a field called **'Dep_ID'** with value '**1**' for each training point.

Additionally, except the **'Extent Mask Layer'**, the outputs of the 'Pre-process data' tool become the inputs to the 'Calculate Weights' tool. Similarly, the outputs from the 'Calculate Weights' tool are used as the inputs for the 'Calculate Responses' tool (Fig. 15:)

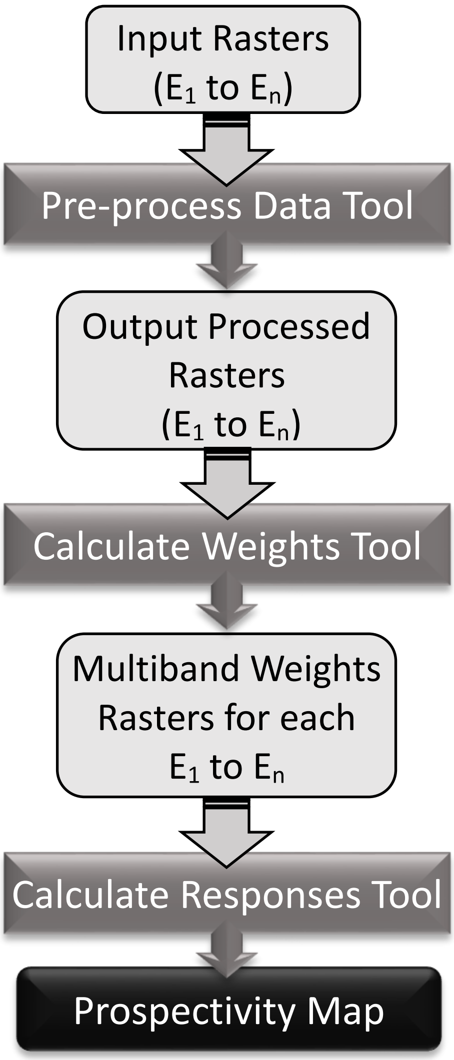


Figure 15: Simplified flowchart of data input-output relation between the tools in the WofE plugin.

5.2 Plugin Graphical User Interface (GUI)

The Plugin GUI comprises three main 'tab widgets'/ tools:

i. Pre-process Data (Fig. 16 a)
ii. Calculate Weights (Fig. 16 b)
iii. Calculate Responses (Fig. 16 c)

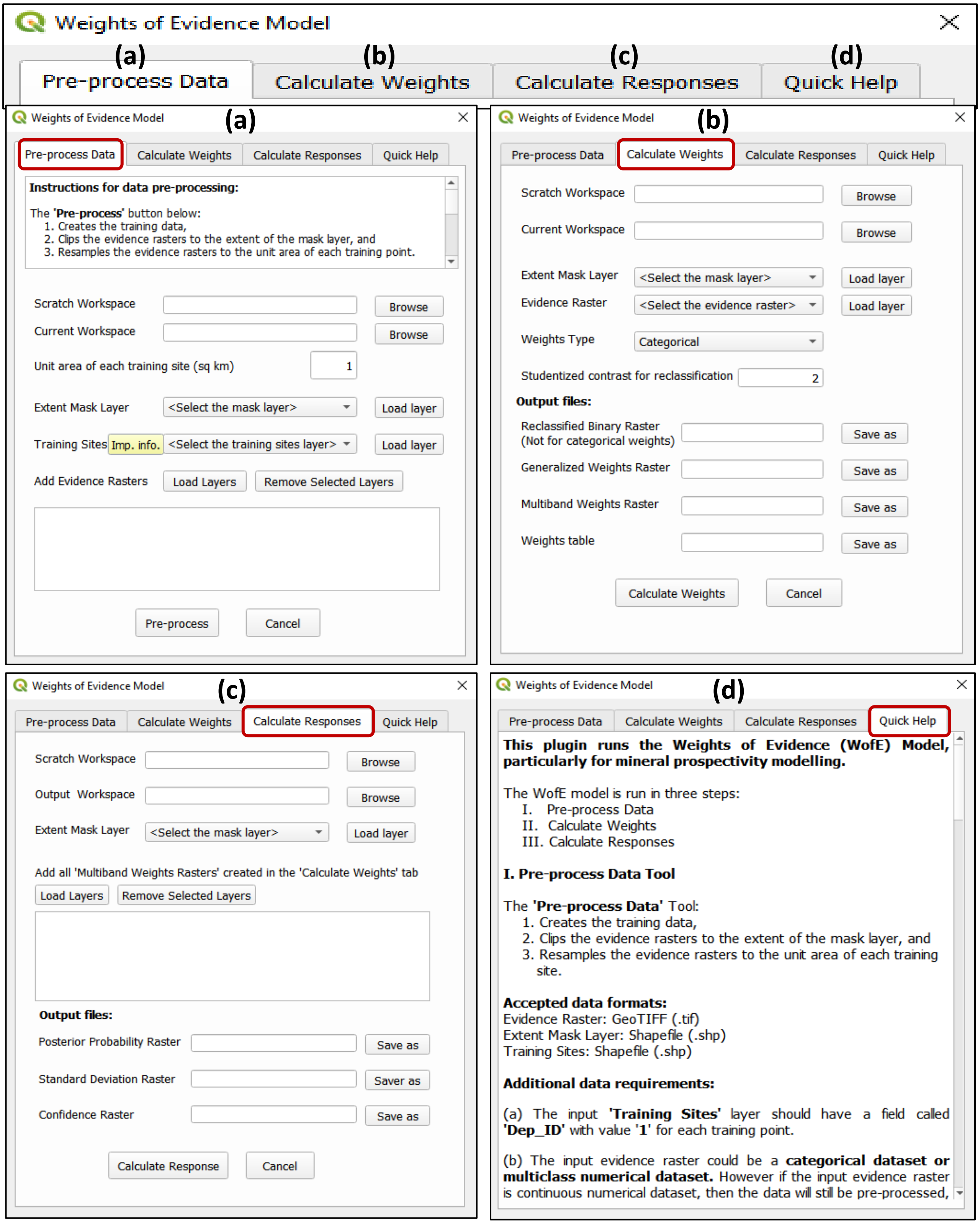


Figure 16: WofE Plugin user interface and tools. (a) Pre-process Data; (b) Calculate Weights; (c) Calculate Responses; (d) Quick Help.

5.3 Test data

The file Sample_data.zip contains the datasets that can be used to implement and test the functioning of the WofE plugin. It contains two shapefiles (.shp) and three GeoTiffs (.tif) as follows (Fig. 17):

- 1_Training_sites.shp: Deposit data (Fig. 17a).
- 2_Extent mask.shp: Polygon defining the boundary of area of analysis (Fig. 17a).
- 3_Lithology_Categorical.tif: Lithological units are represented in this raster. This raster is a categorical dataset, and so will be used to demonstrate categorical weights calculations (Fig. 17a).
- 4_Distance_Ascending.tif: Distance to geological structures such as faults, folds, shear zones, etc. is represented in this raster. It is multiclass numerical dataset. This raster will be used to demonstrate (cumulative) ascending weights calculations (Fig. 17b).
- 5_Density_Descending.tif: Density of geochemically anomalous samples is represented in this raster. It is multiclass numerical dataset. This raster will be used to demonstrate (cumulative) descending weights calculations (Fig. 17c).

The sample datasets are derived from primary geological data available from the Geological Survey of Finland.

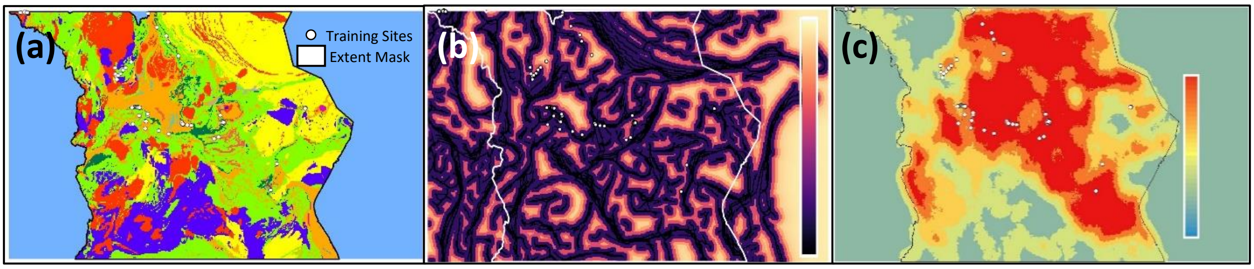


Figure 17: Sample data for plugin testing and implementation. (a) 3_Lithology_Categorical.tif; (b) 4_Distance_Ascending.tif; (c) 5_Density_Descending.tif. Refer to Section 5.3 for description of data.

5.4 Implementation

5.4.1 Pre-process Data Tool

The **'Pre-process Data'** Tool (Fig. 16a):

i. Creates the training data,
ii. Clips the evidence rasters to the extent of the mask layer, and
iii. Resamples the evidence rasters to the unit area of each training site.

Run this tool with the sample data according to the instructions given in Figure 18.

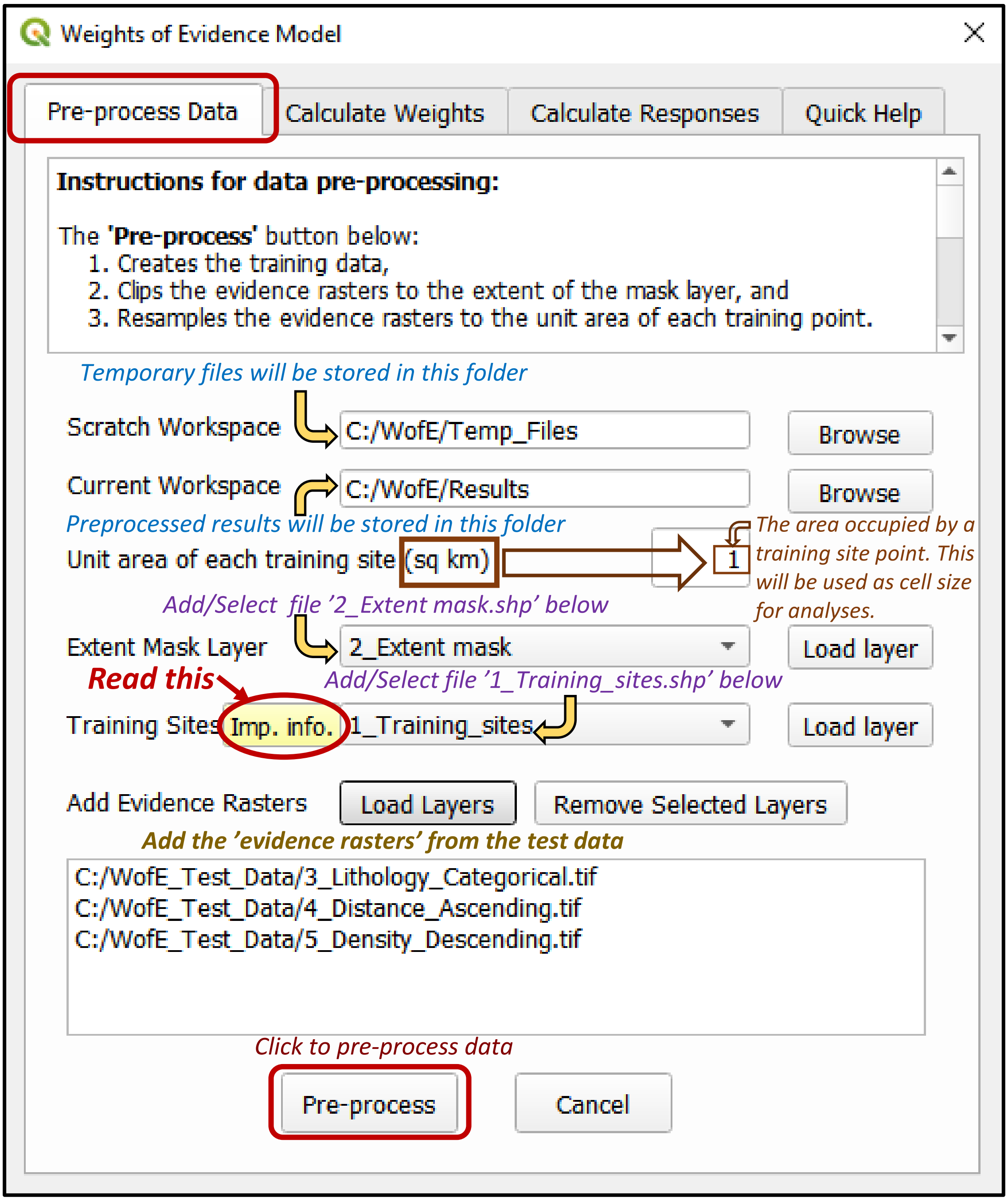


Figure 18: WofE Plugin – Instructions for 'Pre-process Data' tool.

On succesful completion of data pre-processing, a 'message box' will open (Fig. 19). Read and close the 'message box'. Check the 'Layers' list of the open QGIS project. The preprocessed data and the training data files are added to the project (Figs. 20 and 21). Use these files for calculation of weights in the 'Calculate Weights' tool.

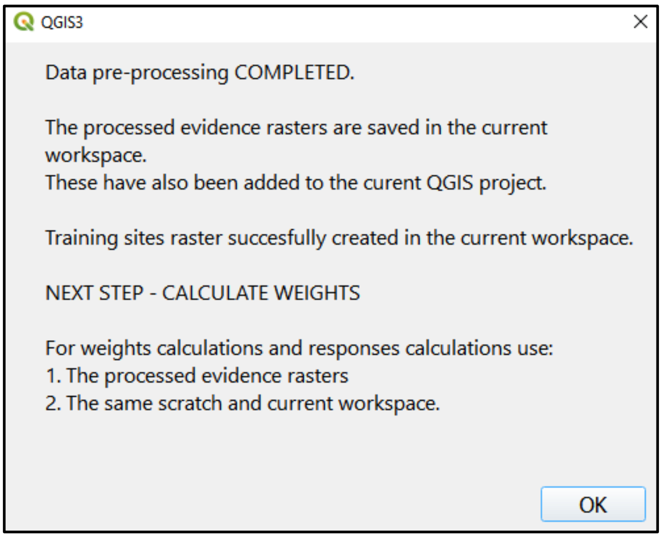


Figure 19: WofE Plugin - Message Box displayed after completion of the Pre-process Data Tool.

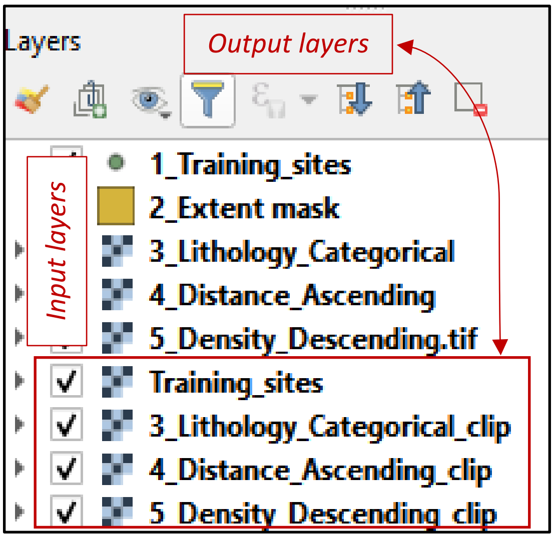


Figure 20: QGIS Layer List - Outputs from 'Pre-process Data' tool are added to QGIS.

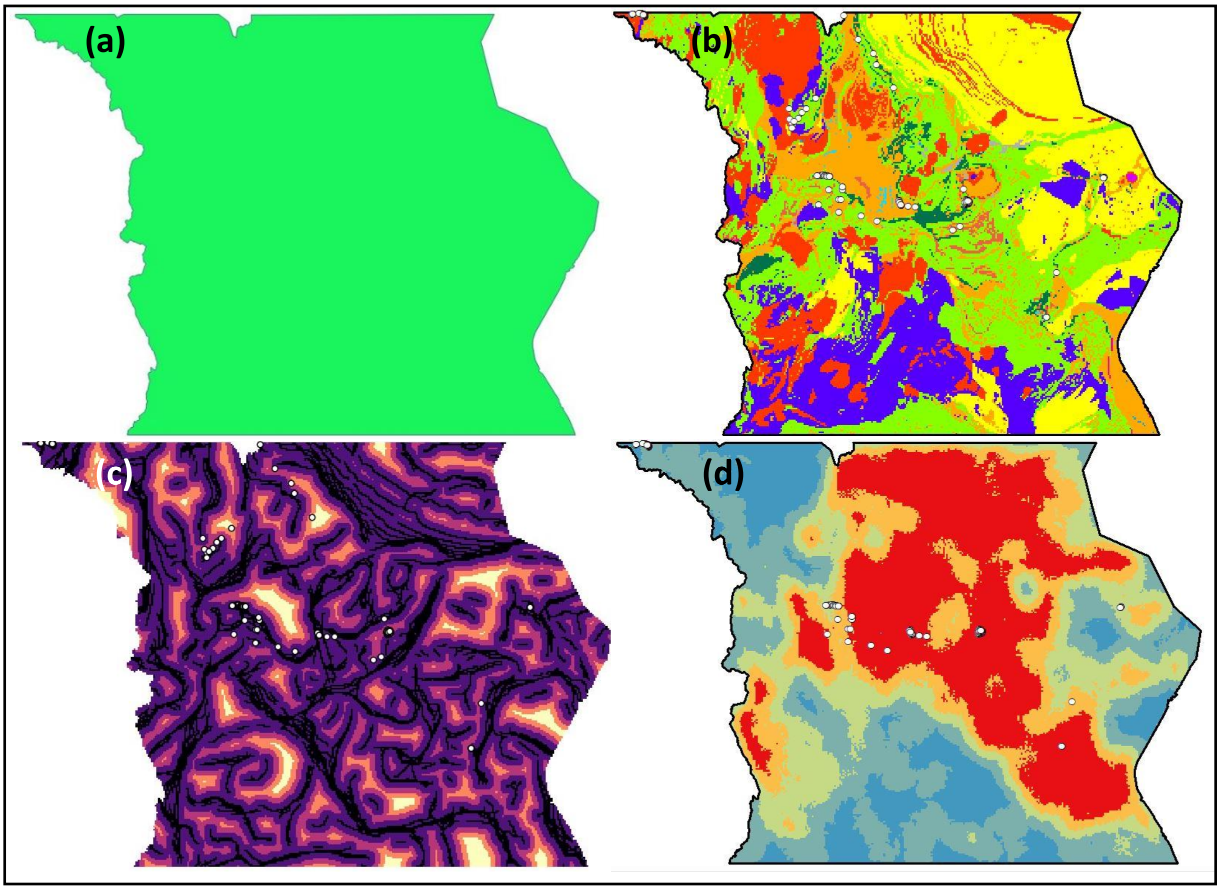


Figure 21: Outputs from 'Pre-process Data' tool. (a) Training_sites.tif; (b) 3_Lithology_Categorical_clip.tif; (c) 4_Distance_Ascending_clip.tif; (d) 5_Density_Descending_clip.tif.

5.4.2 Calculate Weights Tool

The **'Calculate Weights'** tool (Fig. 16b) should be implemented individually for each evidence raster. This tool executes the following calculations:

i. Estimation of prior odds for the mineral deposit per unit cell of the study area,
ii. Reclassification of numerical multi-class evidence raster into the 'Favorable' (Class value: 2) and 'Unfavorable' (Class value: 1) binary evidence raster using each class of the multiclass raster as a threshold in turn, and
iii. Estimation of the 'weights-of-evidence' for each binary evidence raster (from step ii) using the Bayes' Rule.

In this tool the input '**Evidence Raster'** is the raster generated from the **'Pre-process Data'** tool.

This tool will not run for evidence rasters with continuous numerical values, even if such rasters have been successfully pre-processed by the 'Pre-process Data' Tool.

Tool parameters
Workspace: Use the 'Scratch' and 'Current' workspace that was used for data pre-processing.

Extent Mask Layer: The polygon shapefile that marks the area of analysis.

Evidence Raster: Raster generated from the **'Pre-process Data'** tool.

Weights Type: Categorical, Ascending, Descending

- **Categorical:** Use this for weights calculations of categorical data. Example: Different lithological units, different watersheds etc. The data is not reclassified in to binary raster with generalized 'Favorable' and 'Unfavorable' classes.
- **Ascending:** Use this for calculation of cumulative ascending weights for multiclass-numerical data where lower threshold is required for data reclassification. Example: Proximity to a geological feature, where closer the pixel is to the feature, higher is the positive spatial association.
- **Descending:** Use this for calculation of cumulative descending weights for multiclass-numerical data where higher threshold is required for data reclassification. Example: Elemental concentrations in geochemical datasets, where higher the concentration of an element in the pixel, higher is the positive spatial association.

In ascending and descending weights calculations the classes are arranged in ascending or descending order, respectively, and the favorable classes are the ones below the threshold for ascending and above the threshold for descending. For instance in ascending weights, the count of class 1 is added to that of class 2, for class 2 calculations and count of classes 1 and 2 is added to that of class 3, for class 3 calculations and so on. The same applies for descending weights, but the order is reversed.

Studentized contrast: The studentized contrast is calculated for numerical data by dividing the **'Contrast'** by the **'Standard deviation of the contrast'**. The user needs to specify a studentized contrast value that would be suitable for binary reclassification of their numerical datasets. Higher values of studentized contrast indicate less deviations in the weights for that class and lower values indicate more deviations in the weights for the class. Studentized contrast gives the confidence level for the calculated contrast values and the weights. Binary classification with studentized contrast higher than the specified value are classified as '**Favorable'** and the ones with lower values are classified as '**Unfavorable'.** Default value is 2.0, which is approximately 98% confidence. The user can change it to a suitable value that quantifies the spatial association with appropriate confidence.

Tool Results

Reclassified Binary Raster: This raster is created only for **‘ascending’** and **‘descending’** weights calculations. The reclassified raster has two classes:

Class	Value
Favorable class:	2
Unfavorable class:	1

Generalized Weights Raster: After reclassification of numerical data, the weights-of-evidence are recalculated. This raster contains the weights-of-evidence for the **'Favorable'** and **'Unfavorable'** classes of the ‘**Reclassified Binary Raster’**. This raster is also generated for **‘Categorical’** weights calculations, the weights-of-evidence being the positive weights of each class.

Multiband Weights Raster: The reclassified binary (or original categorical) raster, the generalized weights raster and the variance of weights raster are combined in to this 3-band composite raster. Use this ‘**Multiband Weights Raster**’ for calculation of responses and generation of final prospectivity map in the **‘Calculate Responses’** tool.

Weights Table: Results from calculated weights are stored in this **.csv file**. The file contains values of the positive and negative weights, contrast and their standard deviations, studentized contrast, generalized class values and generalized weights and its standard deviation. Check these values and reclassification of the data. If reclassification of the data seems unacceptable for the given studentized contrast value either run ‘Calculate Weights’ tool again with suitable studentized contrast or reclassify the data using a QGIS tool and then run the ‘Calculate Weights’ tool as ‘Categorical’ type.

Categorical Weights Calculations

Figure 22 gives instructions for calculation of categorical weights. Use the raster ‘3_Lithology_Categorical_clip.tif’ obtained from the ‘Pre-process Data’ tool. Run the tool as directed in Figure 22. Keep the workspace same as the one used for ‘Pre-process Data’ tool. On successful completion of weights calculations, a ‘message box’ (Fig. 23) will open. Read the message and close the ‘message box’. Check the ‘Layers’ list of the open QGIS project. The results from the tool are added to the current QGIS project. The results obtained for categorical weights are listed in Figure 24a. These are Lith_cat_wgts.csv (Fig. 25), Lith_cat_wgts.tif (Fig. 26a), Lith_cat_wgts_std.tif, Lith_cat_wgts_var.tif (Fig. 26b) and Lith_merged_wgts.tif (Fig. 26c).

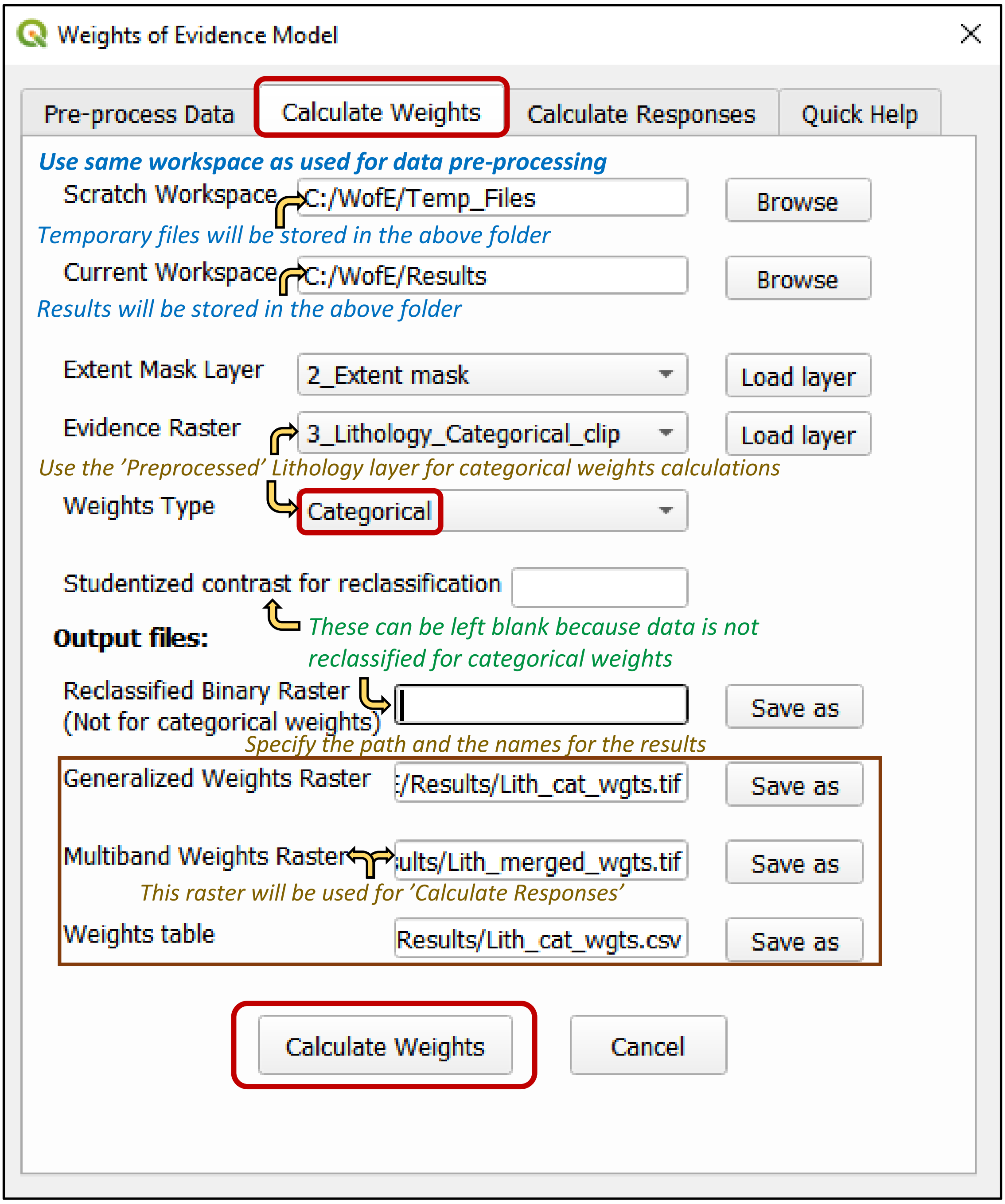


Figure 22: WofE Plugin – Instructions for 'Calculate Weights' tool. This Figure demonstrates calculation of categorical weights.

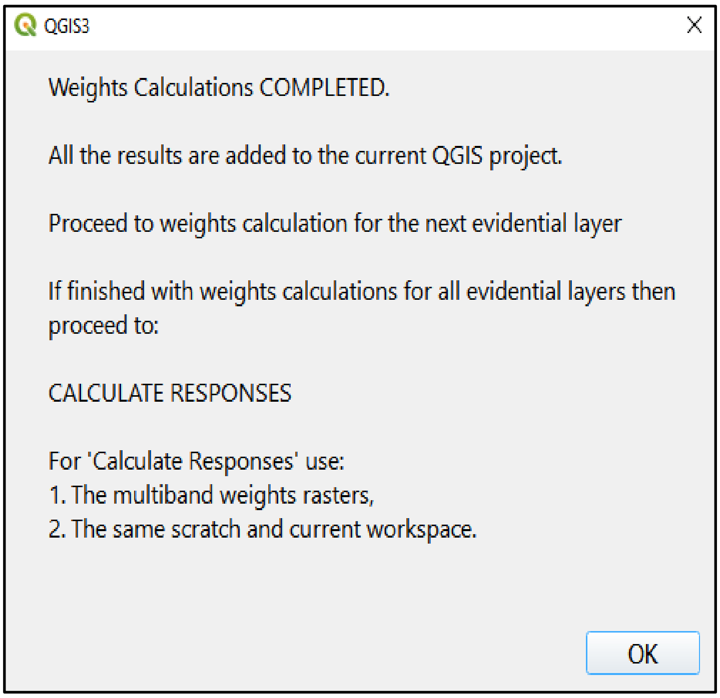


Figure 23: Message box that appears on completion of weights calculation.

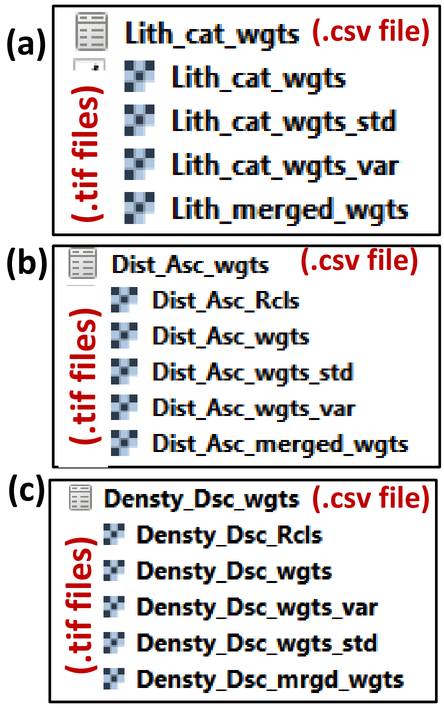
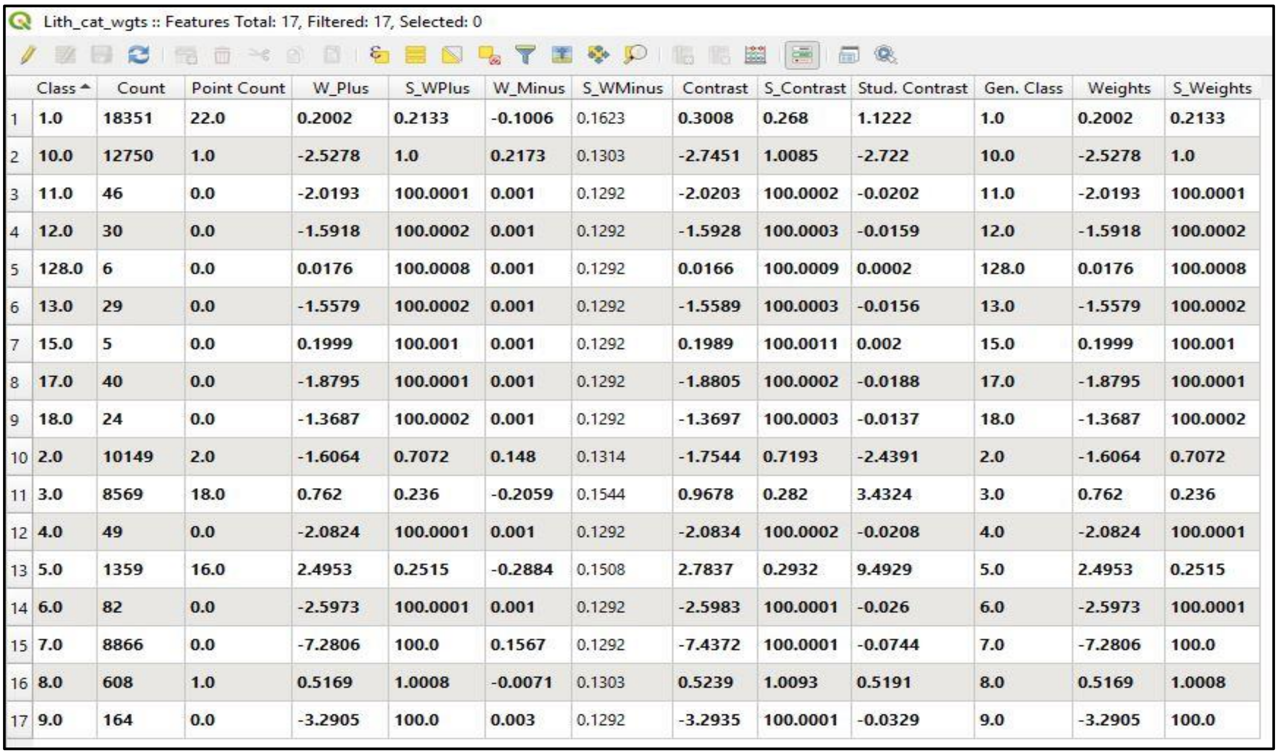


Figure 24: List of files created from Calculate Weights tools for: (a) Categorical, (b) Ascending and (c) Descending weights calculations.


Lith_cat_wgts :: Features Total: 17, Filtered: 17, Selected: 0

	Class	Count	Point Count	W_Plus	S_WPlus	W_Minus	S_WMinus	Contrast	S_Contrast	Stud. Contrast	Gen. Class	Weights	S_Weights
1	1.0	18351	22.0	0.2002	0.2133	-0.1006	0.1623	0.3008	0.268	1.1222	1.0	0.2002	0.2133
2	10.0	12750	1.0	-2.5278	1.0	0.2173	0.1303	-2.7451	1.0085	-2.722	10.0	-2.5278	1.0
3	11.0	46	0.0	-2.0193	100.0001	0.001	0.1292	-2.0203	100.0002	-0.0202	11.0	-2.0193	100.0001
4	12.0	30	0.0	-1.5918	100.0002	0.001	0.1292	-1.5928	100.0003	-0.0159	12.0	-1.5918	100.0002
5	128.0	6	0.0	0.0176	100.0008	0.001	0.1292	0.0166	100.0009	0.0002	128.0	0.0176	100.0008
6	13.0	29	0.0	-1.5579	100.0002	0.001	0.1292	-1.5589	100.0003	-0.0156	13.0	-1.5579	100.0002
7	15.0	5	0.0	0.1999	100.001	0.001	0.1292	0.1989	100.0011	0.002	15.0	0.1999	100.001
8	17.0	40	0.0	-1.8795	100.0001	0.001	0.1292	-1.8805	100.0002	-0.0188	17.0	-1.8795	100.0001
9	18.0	24	0.0	-1.3687	100.0002	0.001	0.1292	-1.3697	100.0003	-0.0137	18.0	-1.3687	100.0002
10	2.0	10149	2.0	-1.6064	0.7072	0.148	0.1314	-1.7544	0.7193	-2.4391	2.0	-1.6064	0.7072
11	3.0	8569	18.0	0.762	0.236	-0.2059	0.1544	0.9678	0.282	3.4324	3.0	0.762	0.236
12	4.0	49	0.0	-2.0824	100.0001	0.001	0.1292	-2.0834	100.0002	-0.0208	4.0	-2.0824	100.0001
13	5.0	1359	16.0	2.4953	0.2515	-0.2884	0.1508	2.7837	0.2932	9.4929	5.0	2.4953	0.2515
14	6.0	82	0.0	-2.5973	100.0001	0.001	0.1292	-2.5983	100.0001	-0.026	6.0	-2.5973	100.0001
15	7.0	8866	0.0	-7.2806	100.0	0.1567	0.1292	-7.4372	100.0001	-0.0744	7.0	-7.2806	100.0
16	8.0	608	1.0	0.5169	1.0008	-0.0071	0.1303	0.5239	1.0093	0.5191	8.0	0.5169	1.0008
17	9.0	164	0.0	-3.2905	100.0	0.003	0.1292	-3.2935	100.0001	-0.0329	9.0	-3.2905	100.0

Figure 25: Weights table, Lith_cat_wgts.csv, created from categorical weights calculation of raster 3_Lithology_Categorical_clip.tif.
Index to table columns: Class = Value of class of input data; Count = Count of unit-area cells occupied by each class. Point count = Count of unit-area cells occupied by training sites in each class; W_Plus = Positive weights-of-evidence; S_Wplus = Standard deviation of W_Plus; W_Minus = Negative weights-of-evidence; S_WMinus = Standard deviation of W_MInus; Contrast = W_Plus - W_Minus; S_Contrast = Standard deviation of Contrast; Stud. Contrast = Studentized contrast, i.e, Contrast ÷ S_Contrast; Gen. Class = Generalized classes. For categorical weights calculations these are the same as original class values; Weights = Weights-of-evidences of the generalized classes; S_Weights = Standard deviation of the weights of generalized classes.

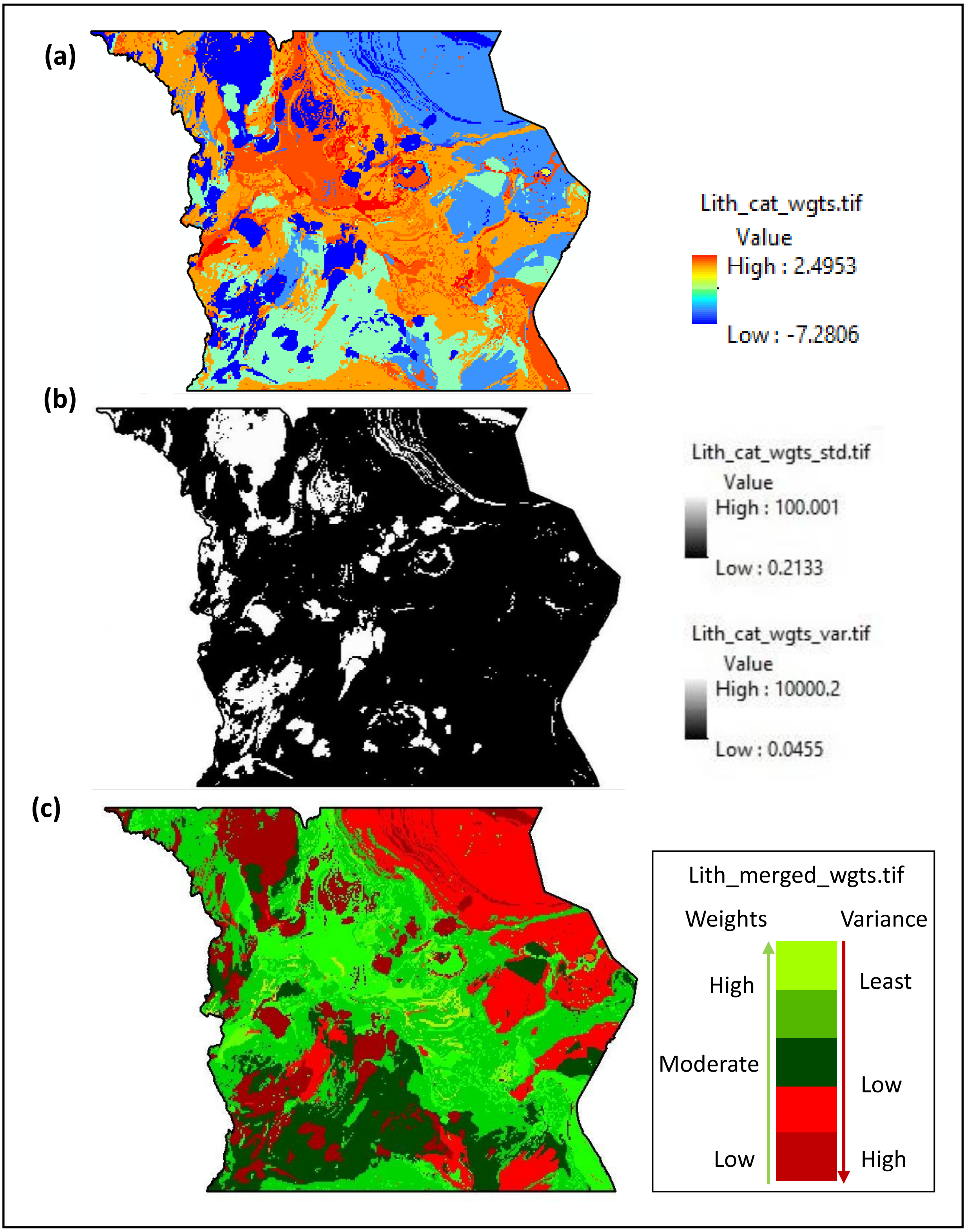


Figure 26: Raster files generated from categorical weights calculation of raster 3_Lithology_Categorical_clip.tif. (a) Weights Raster - Lith_cat_wgts.tif: Represents the positive weights-of-evidence for each class of the evidence raster; (b) Standard deviation and the variance of weights: Variance is square of standard deviation. Hence both attributes are represented by same values on the greyscale color bar; (c) Multiband Weights Raster - Lith_merged_wgts.tif: This raster is comprised of rasters 3_Lithology_Categorical_clip.tif, Lith_cat_wgts.tif and Lith_cat_wgts_var.tif as Bands 1, 2 and 3 respectively. This raster will be used in the 'Calculate Response' tool (Section 5.4.3).

Ascending Weights Calculations

Figure 27 gives instructions for calculation of cumulative ascending weights. Use the raster '4_Distance_Ascending_clip.tif' obtained from the 'Pre-process Data' tool. Run the tool as directed in Figure 27. Keep the workspace same as that used for 'Pre-process Data' tool.

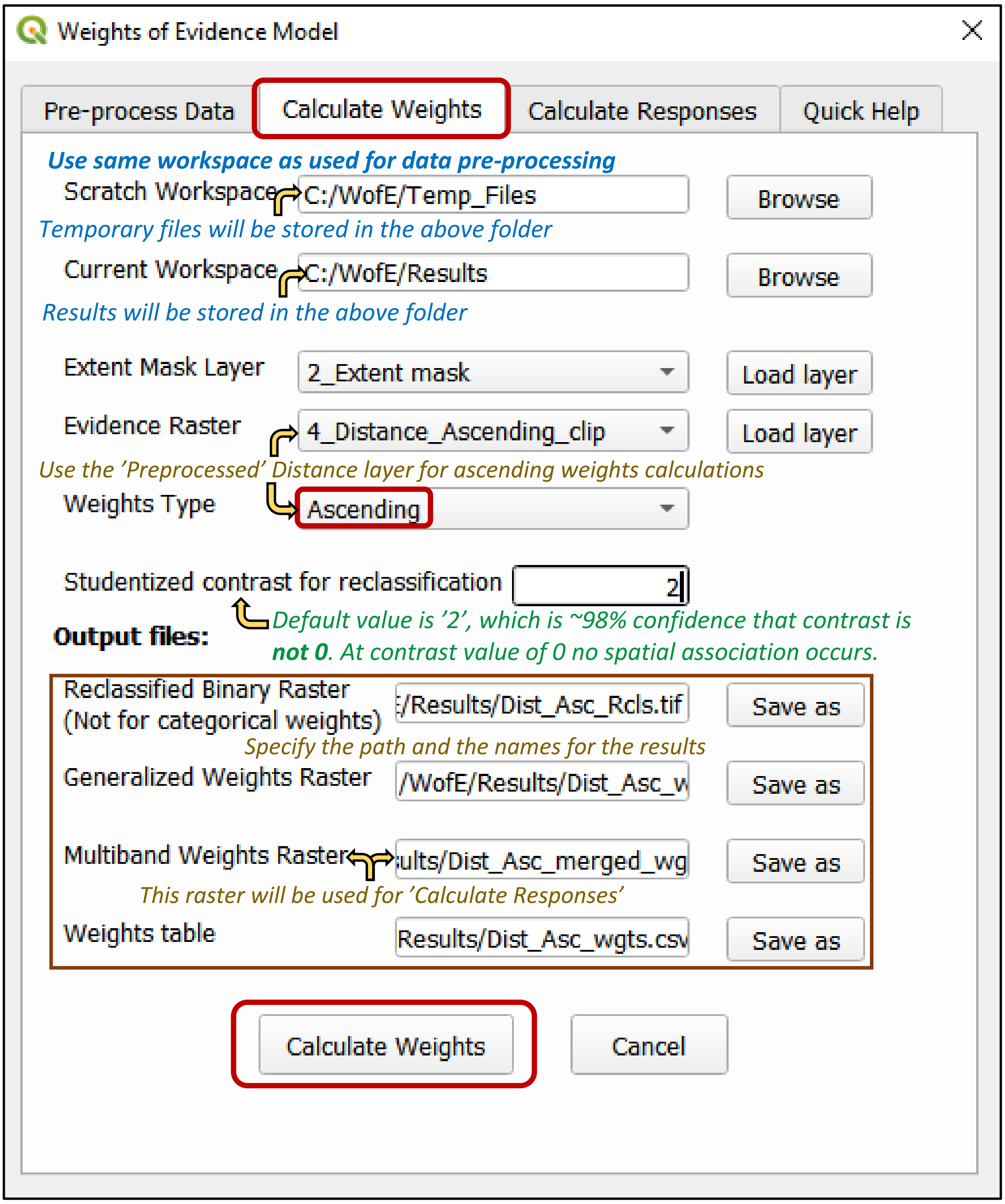


Figure 27: WofE Plugin – Instructions for 'Calculate Weights' tool. This Figure demonstrates calculation of ascending weights.

However now instead of a message box displaying 'weights calculation completed' (Fig. 23), a different message box appears (Fig. 28a). It indicates an error in calculation of the weights. When this error occurs, a temporary file – 'Calc_Stud_Cont.csv' is created and added to the QGIS project (Fig. 29). This error appears when the studentized contrast value given by the user is too high/low for reclassification. For instance, Calc_Stud_Cont.csv shows that none of the classes had studentized contrast value of ≥2. Therefore all the classes were classified 'Unfavorable', i.e. Class 1. This is mentioned in the message box (Fig. 28a). After closing the message box, QGIS python error (Fig. 28b) is also displayed and the plugin also closes. Open Calc_Stud_Cont.csv (Fig. 29). It shows that the maximum studentized contrast value is 1.686 and occurs for Class 3. But the maximum contrast of >3 occurs for Class 4. However with very low studentized contrast, (which implies low confidence). Therefore, it is at the discretion of the user to decide the studentized contrast values based on the calculated values displayed in the table, and run the tool again. For this dataset, calculate weights again using the studentized contrast value 1.5 (Fig. 30).

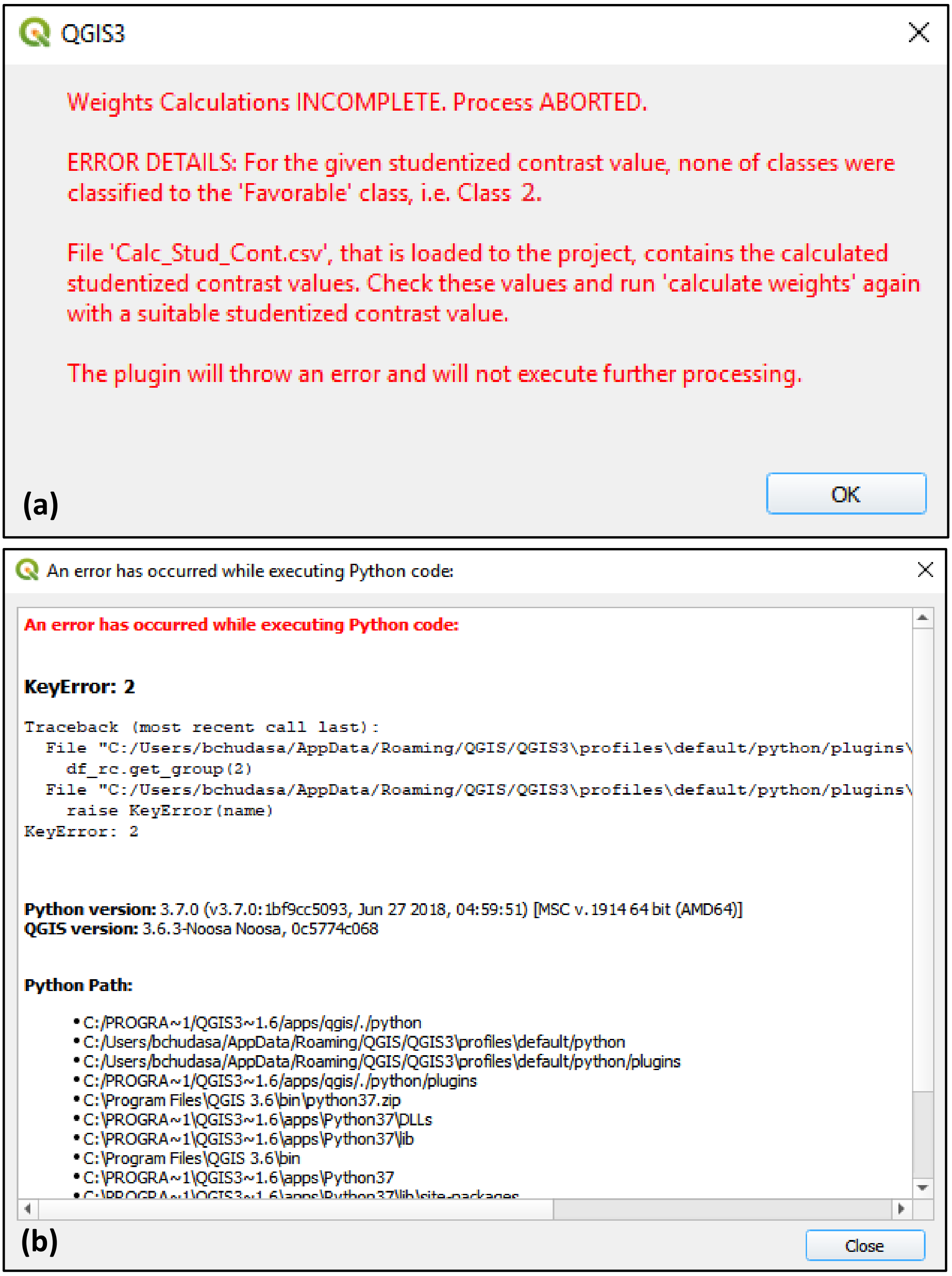
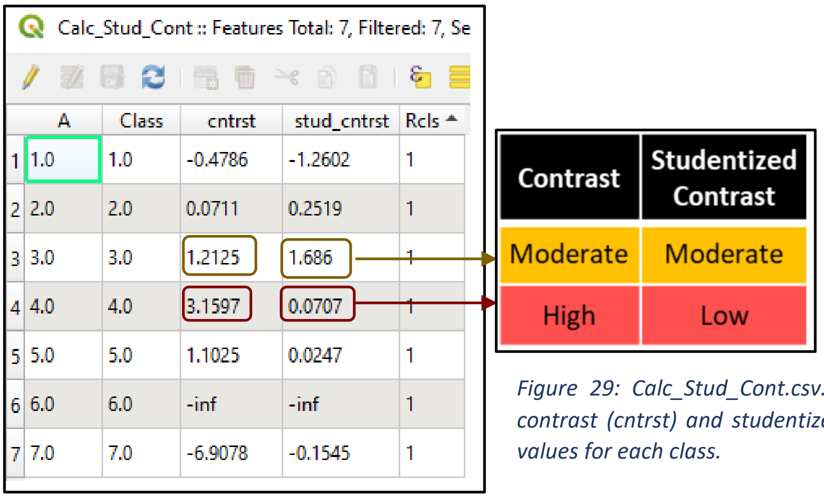


Figure 28: Errors in calculation of weights. (a) WOfE plugin message box giving error details; (b) QGIS error window.



Calc_Stud_Cont :: Features Total: 7, Filtered: 7, Se

	A	Class	cntrst	stud_cntrst	Rcls
1	1.0	1.0	-0.4786	-1.2602	1
2	2.0	2.0	0.0711	0.2519	1
3	3.0	3.0	1.2125	1.686	1
4	4.0	4.0	3.1597	0.0707	1
5	5.0	5.0	1.1025	0.0247	1
6	6.0	6.0	-inf	-inf	1
7	7.0	7.0	-6.9078	-0.1545	1

Contrast	Studentized Contrast
Moderate	Moderate
High	Low

Figure 29: Calc_Stud_Cont.csv. The table displays the contrast (cntrst) and studentized contrast (stud_cntrst) values for each class.

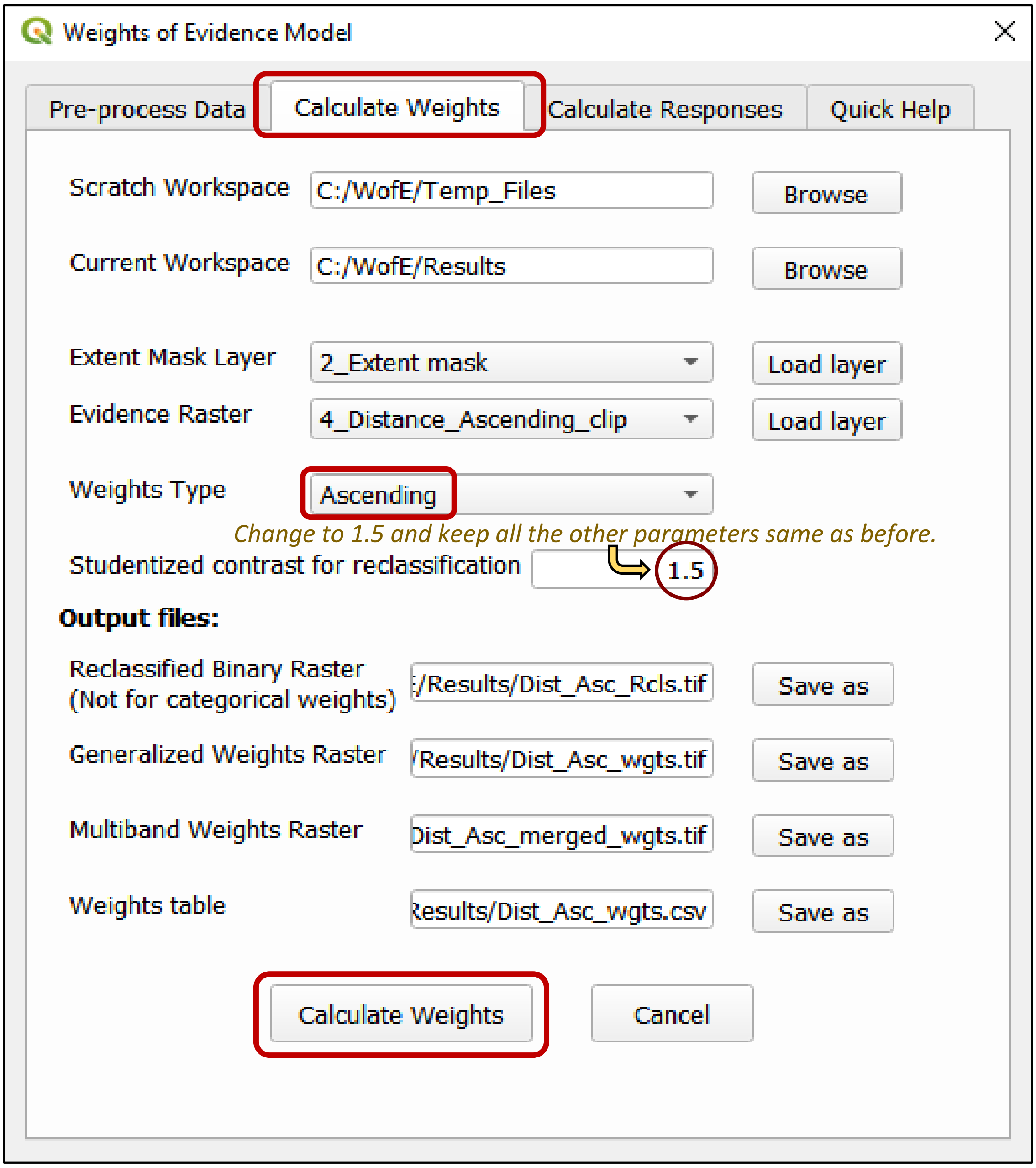


Figure 30: Ascending weights calculations using a lower studentized contrast value.

The weights calculation runs to completion. The message box (Fig. 23) is displayed to convey the same. The results added to the current QGIS project are listed in Figure 24b. These are Dist_Asc_wgts.csv (Fig. 31a), Dist_Asc_Rcls.tif (Fig. 32a), Dist_Asc_wgts.tif (Fig. 32b), Dist_Asc_wgts_std.tif, Dist_Asc_wgts_var.tif (Fig. 32c) and Dist_Asc_merged_wgts.tif. The 3-band composite of the reclassified-, the weights- and the variance-rasters (Dist_Asc_merged_wgts.tif) will be used in the 'Calculate Response' tool (Section 5.4.3).

(a)

dist_Asc_wgts :: Features Total: 7, Filtered: 7, Selected: 0

	Class ▲	Count	Cmltv. Count	Point Count	Cmltv. Point Count	No Dep Cnt	WPlus	S_WPlus	WMinus	S_WMinus	Contrast	S_Contrast	Stud Contrast	Gen. Class	Weights	S_Weights
1	1.0	12148	12148.0	8.0	8.0	12140.0	-0.3994	0.354	0.0792	0.1387	-0.4786	0.3798	-1.2602	2.0	0.0743	0.1314
2	2.0	29699	41847.0	34.0	42.0	29665.0	0.0223	0.154	-0.0488	0.2358	0.0711	0.2821	0.2519	2.0	0.0743	0.1314
3	3.0	13017	54864.0	16.0	58.0	13001.0	0.0743	0.131	-1.1382	0.7072	1.2125	0.7191	1.686	2.0	0.0743	0.1314
4	4.0	4861	59725.0	2.0	60.0	4859.0	0.0242	0.129	-3.1355	44.7214	3.1597	44.7215	0.0707	1.0	-1.1236	0.7072
5	5.0	1223	60948.0	0.0	60.0	1223.0	0.0039	0.129	-1.0986	44.7214	1.1025	44.7215	0.0247	1.0	-1.1236	0.7072
6	6.0	176	61124.0	0.0	60.0	176.0	0.001	0.129	inf	44.7214	-inf	44.7215	-inf	1.0	-1.1236	0.7072
7	7.0	3	61127.0	0.0	60.0	3.0	0.001	0.129	6.9088	44.7214	-6.9078	44.7215	-0.1545	1.0	-1.1236	0.7072

(b)

Density_Dsc_wgts :: Features Total: 6, Filtered: 6, Selected: 0

	Class ▼	Count	Cmltv. Count	Point Count	Cmltv. Point Count	No Dep Cnt	WPlus	S_WPlus	WMinus	S_WMinus	Contrast	S_Contrast	Stud Contrast	Gen. Class	Weights	S_Weights
1	5.0	17530	17530.0	27.0	27.0	17503.0	0.4511	0.193	-0.2596	0.1741	0.7107	0.2595	2.7388	2.0	0.116	0.1303
2	4.0	10423	27953.0	11.0	38.0	10412.0	0.326	0.162	-0.3917	0.2133	0.7178	0.2674	2.6847	2.0	0.116	0.1303
3	3.0	11097	39050.0	15.0	53.0	11082.0	0.3244	0.137	-1.1267	0.378	1.4511	0.4024	3.6065	2.0	0.116	0.1303
4	2.0	14479	53529.0	6.0	59.0	14473.0	0.116	0.13	-1.9871	1.0001	2.1031	1.0085	2.0853	2.0	0.116	0.1303
5	1.0	7597	61126.0	1.0	60.0	7596.0	0.001	0.129	inf	44.7214	-inf	44.7215	-inf	1.0	-2.0101	1.0001
6	0.0	1	61127.0	0.0	60.0	1.0	0.001	0.129	6.9088	44.7214	-6.9078	44.7215	-0.1545	1.0	-2.0101	1.0001

Figure 31: Weights tables (a) Dist_Asc_wgts.csv, created from ascending weights calculation of raster 4_Distance_Ascending_clip.tif, and (b) Densty_Dsc_wgts.csv created from descending weights calculation of raster 5_Density_Descending_clip.tif. In ascending and descending weights calculations the classes are arranged in ascending or descending order, respectively. Calculations are then performed cumulatively. For instance in ascending weights, the count of class 1 is added to that of class 2, for class 2 calculations and count of classes 1 and 2 is added to that of class 3, for class 3 calculation and so on. The same applies for descending weights, but the order is reversed. Index to table columns: Class = Value of class of input data; Count = Count of unit-area cells occupied by each class. Cmltv. Count = Cumulative count of unit-area cells occupied by the class. Point count = Count of unit-area cells occupied by training sites in each class; Cmltv. Point Count = Cumulative count of unit-area cells occupied by training sites in the classes; No Dep Cnt = Cumulative count of unit-area cells without training sites in the classes; W_Plus = Positive weights-of-evidence; S_Wplus = Standard deviation of W_Plus; W_Minus = Negative weights-of-evidence; S_WMinus = Standard deviation of W_MInus; Contrast = W_Plus - W_Minus; S_Contrast = Standard deviation of Contrast; Stud. Contrast = Studentized contrast, i.e, Contrast ÷ S_Contrast; Gen. Class = Generalized classes after reclassification, 2 – Favorable, 1 - Unfavorable; Weights = Weights-of-evidences of the generalized classes; S_Weights = Standard deviation of the weights of generalized classes.

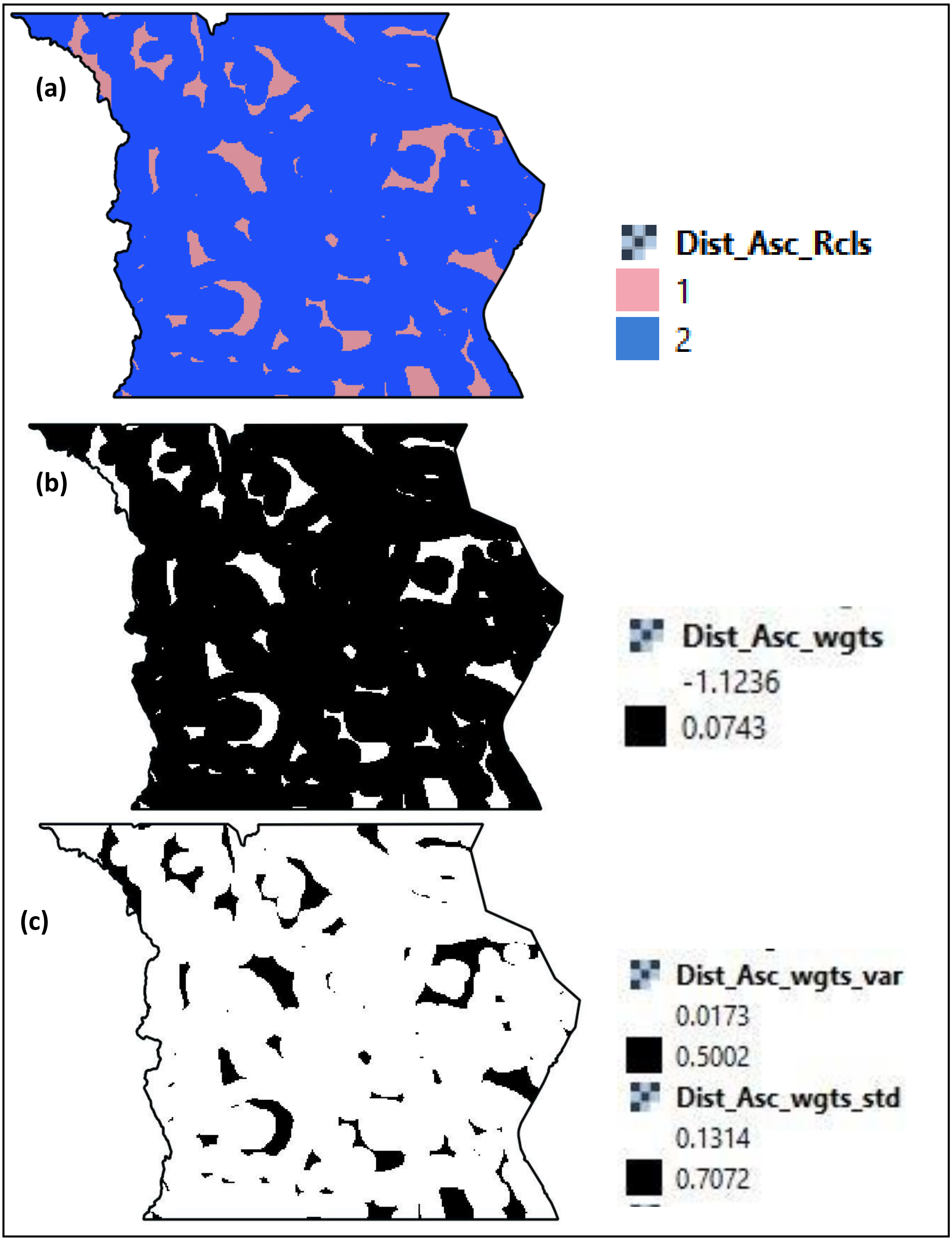


Figure 32: Raster files generated from ascending weights calculation of raster 4_Distance_Ascending_clip.tif. (a) Reclassified binary raster – Dist_Asc_Rcls.tif: Represents the generalized favorable and unfavorable classes (b) Weights Raster – Dist_Asc_wgts.tif: Represents the positive weights-of-evidence for the generalized classes; (c) Standard deviation and the variance of weights: Variance is square of standard deviation. Hence both attributes are represented by same values on the greyscale color bar.

N.B.: Dist_Asc_merged_wgts.tif is also created. It is a 3-band composite of the reclassified (32a), the weights (32b) and the variance (32c) rasters. This raster will be used in the 'Calculate Responses' tool (Section 5.4.3).

Descending Weights Calculations

Figure 33 gives instructions for calculation of cumulative descending weights. Use the raster '5_Density_Descending_clip.tif' obtained from the 'Pre-process Data' tool. Run the tool as directed in Figure 33. Keep the workspace same as that used for 'Pre-process Data' tool.

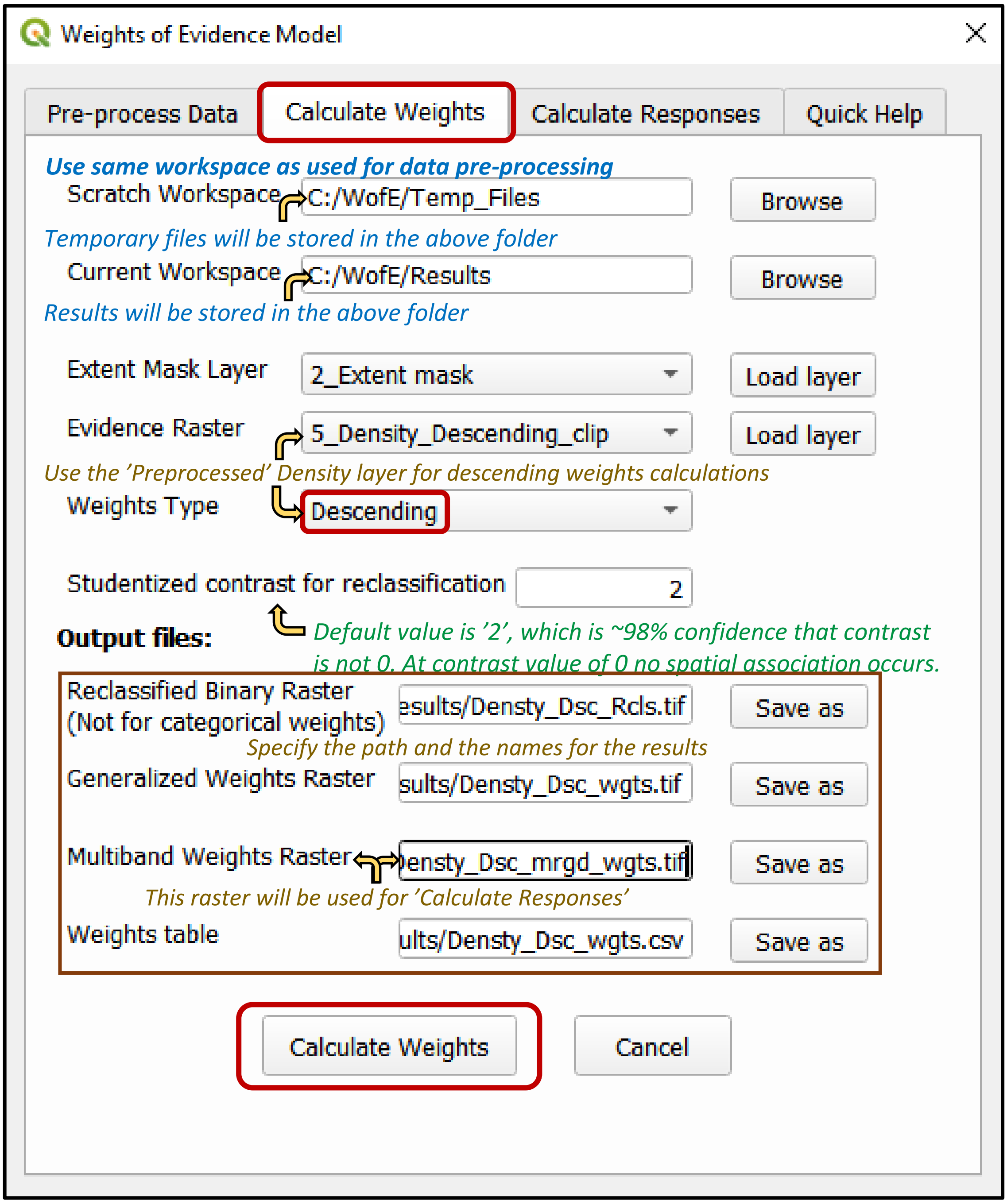


Figure 33: WofE Plugin – Instructions for 'Calculate Weights' tool. This Figure demonstrates calculation of descending weights.

Weights calculation runs to completion. The message box (Fig. 23) is displayed to convey the same. The results added to the current QGIS project are listed in Figure 24c. These are Densty_Asc_wgts.csv (Fig. 31b), Densty_Dsc_Rcls.tif (Fig. 34a), Densty_Dsc_wgts.tif (Fig. 34b), Densty_Dsc_wgts_std.tif, Densty_Dsc_wgts_var.tif (Fig. 34c) and Dist_Asc_merged_wgts.tif. The 3-band composite of the reclassified-, the weights- and the variance-rasters (Densty_Dsc_mrgd_wgts.tif) will be used in the 'Calculate Responses' tool (Section 5.4.3).

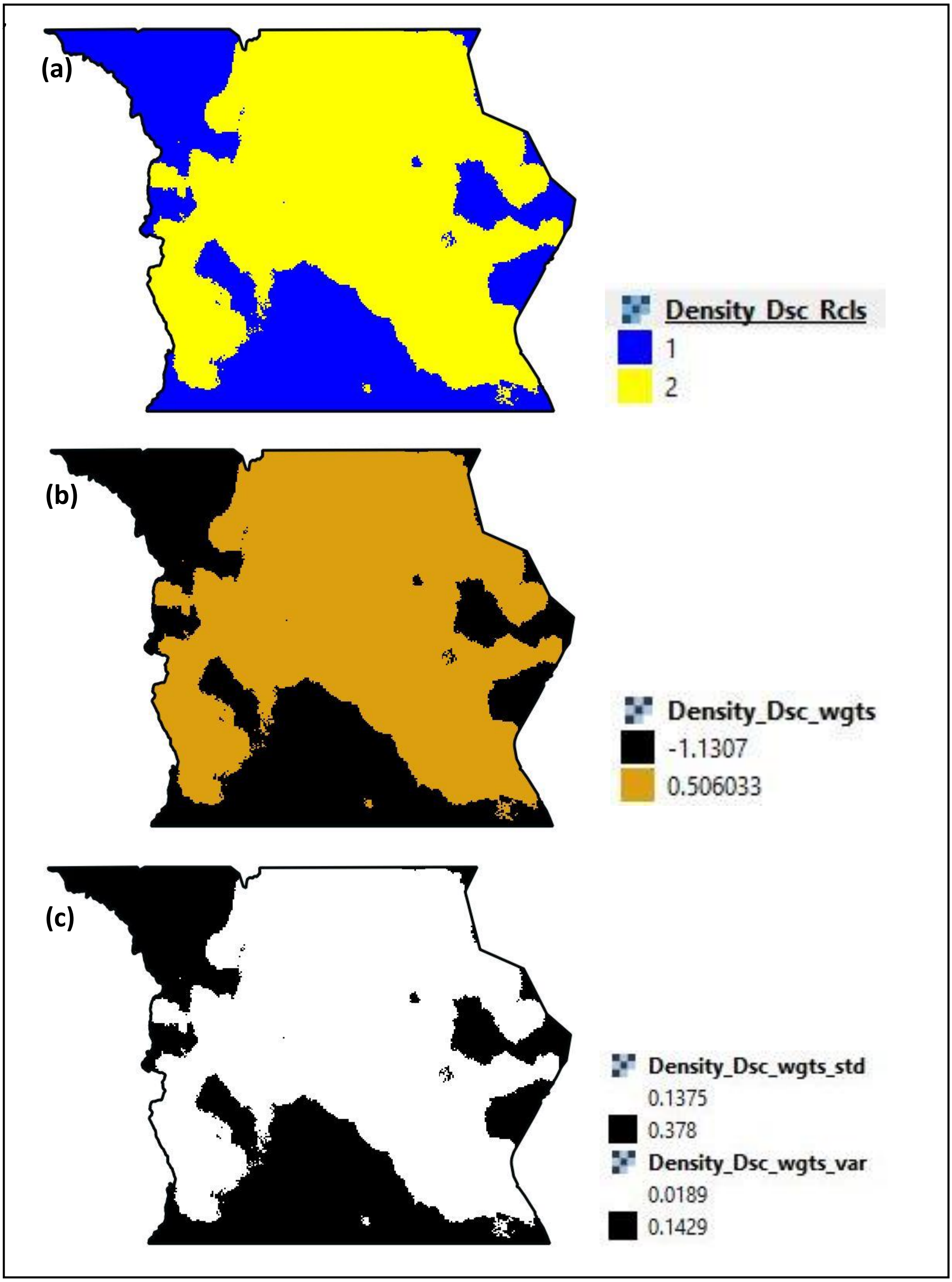


Figure 34: Raster files generated from descending weights calculation of raster 5_Density_Descending_clip.tif. (a) Reclassified binary raster – Densty_Dsc_Rcls.tif: Represents the generalized favorable and unfavorable classes (b) Weights Raster - Densty_Dsc_wgts.tif: Represents the positive weights-of-evidence for the generalized classes; (c) Standard deviation and the variance of weights: Variance is square of standard deviation. Hence both attributes are represented by same values on the greyscale color bar.

N.B.: Densty_Dsc_mrgd_wgts.tif is also created. It is a 3-band composite of the reclassified, weights and the variance rasters. This raster will be used in the 'Calculate Response' tool (Section 5.4.3).

5.4.3 Calculate Responses Tool

The 'Calculate Responses' tool (Fig. 16 c) is the last step of WofE model. This tool calculates the posterior probability for occurrence of a deposit by combining the prior odds and the weights of the evidence rasters calculated in the 'Calculate Weights' tool.

This tool works under the assumption of conditional independence of the input evidence rasters. Conditional independence implies that the occurrence of features in a given evidence raster does not have influence on and are independent of the occurrence of features in any other evidence rasters.

Input to this tool is the **'Multiband Weights Raster'** of each evidence raster created from the **'Calculate Weights'** tool

Results:

Posterior Probability Raster: The values in this raster are the prospectivity values for occurrence of a deposit. It is generated by combining the prior odds and the weights of individual evidence rasters generated in the 'Calculate Weights' tool.

Standard Deviation Raster: This raster contains the standard deviations in the posterior probability calculations because of the deviations in weights of the evidence rasters. The plugin cannot handle evidence rasters with missing/no data values. Hence this raster is also considered as the total standard deviation raster.

Confidence Raster: This raster represents the confidence of the prospectivity values obtained in the **Posterior Probability Raster.** It is ratio of the Posterior Probability Raster and the Standard Deviation Raster. High posterior probability values with low standard deviation values will have high confidence values. Similarly low posterior probability values with high standard deviation values will have low confidence values.

Figure 35 gives instructions for calculation of responses. Keep the workspace same as that used for 'Pre-process Data' tool. Use the **multiband weights rasters created from the Calculate Weights tool (Lith_merged_wgts.tif, Dist_Asc_merged_wgts.tif and Densty_Dsc_mrgd_wgts.tif)** as input files. Run the tool. On successful completion of the calculations, a 'message box' (Fig. 36) will show up. All modeling results will be added to the open QGIS project. Figure 37 summarizes the entire framework of the WofE plugin from the input data pre-processing to the final calculate responses. The results are shown in Figure 38.

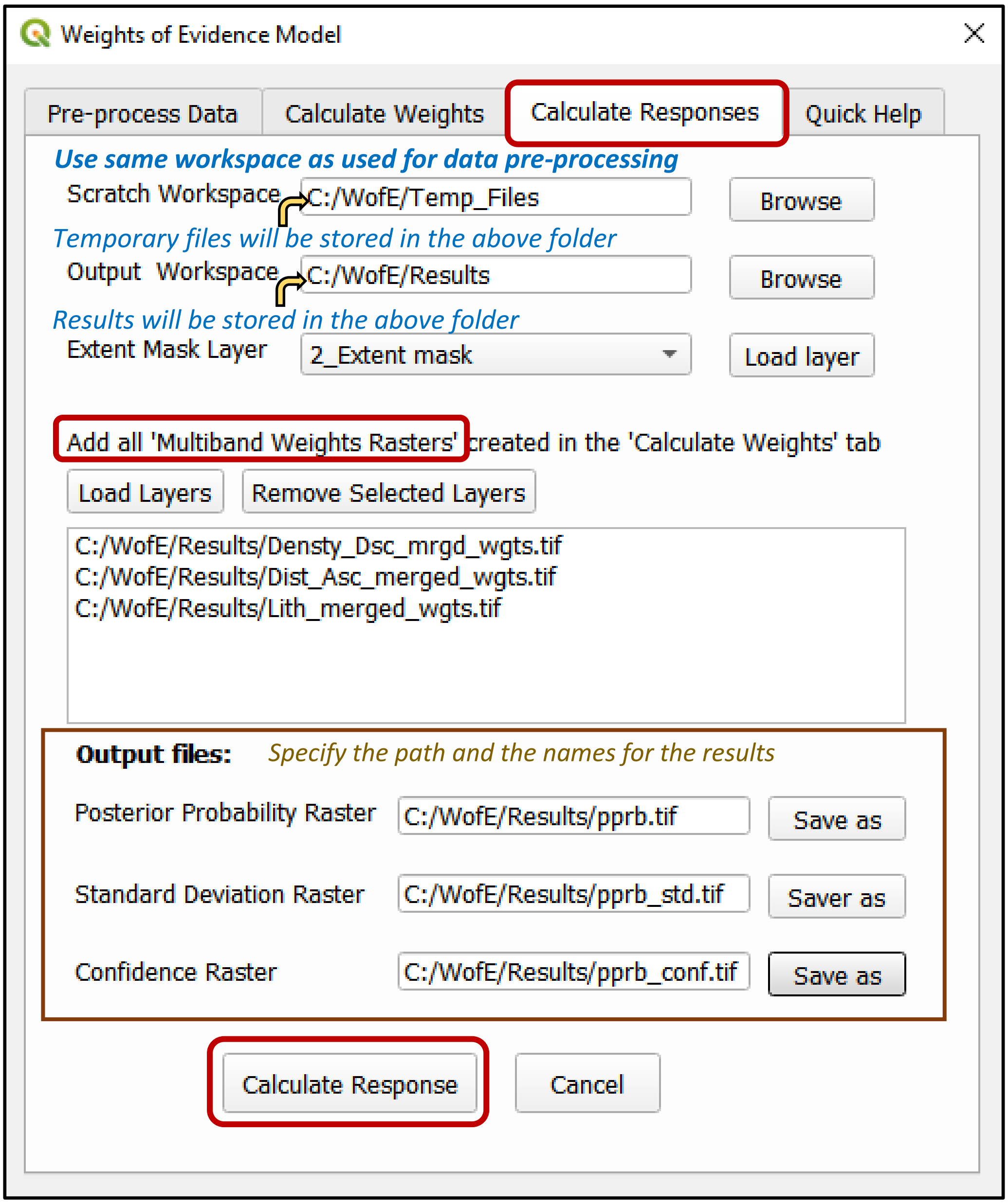


Figure 35: WofE Plugin – Instructions for 'Calculate Responses' tool.

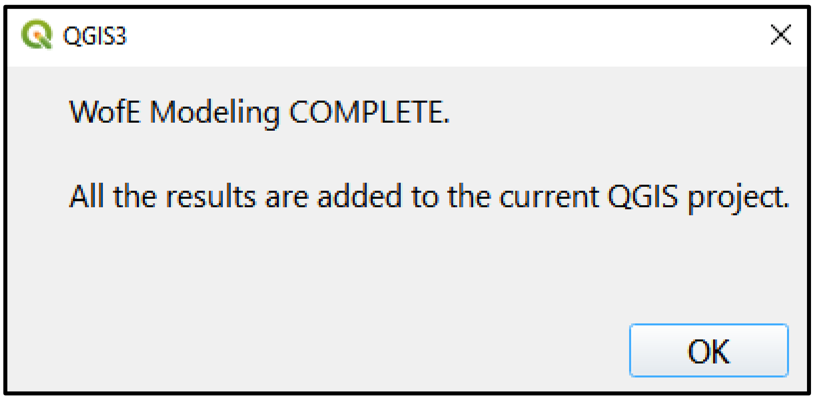


Figure 36: Message box that appears after the completion of responses calculations.

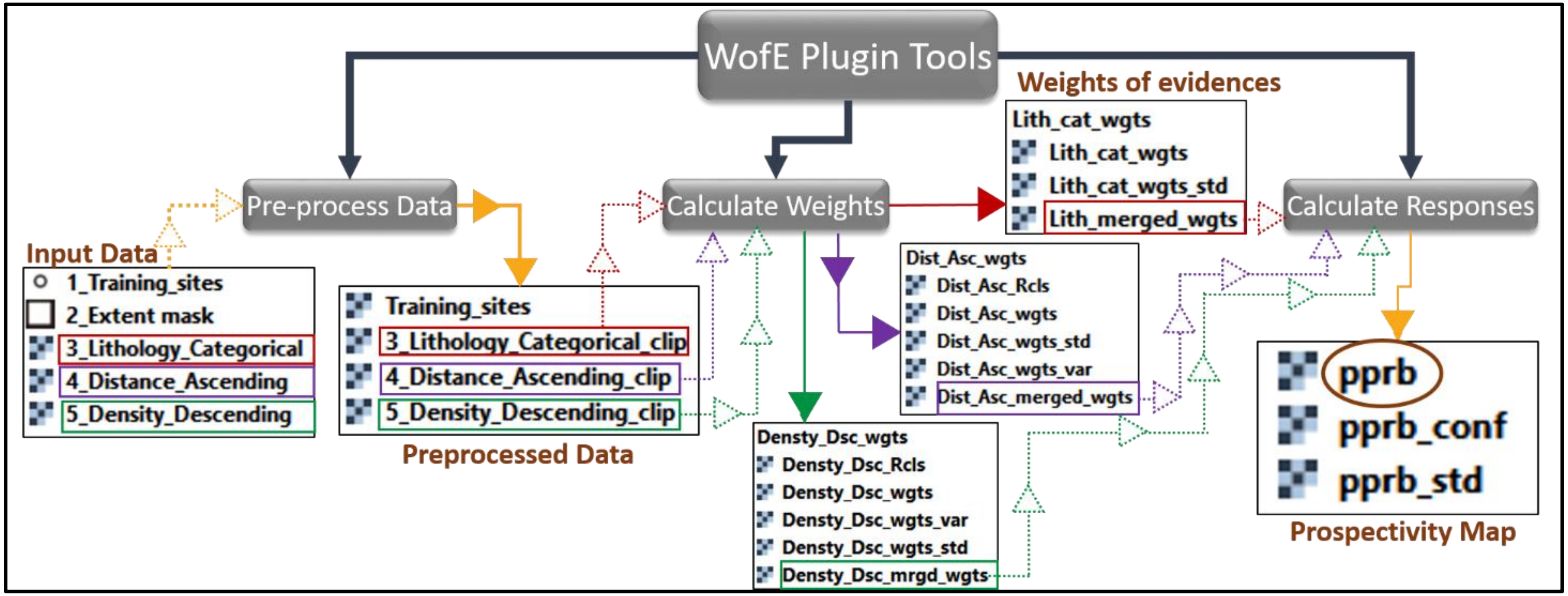


Figure 37: Framework of the WofE Plugin tools. The processing path of the input test data through the tool up to the final results is mapped. Input files to each tool are shown by dotted lines. The straight lines map the outputs from the tool. The path of each input evidence raster is color coded as red, violet and green for 3_Lithology_Categorical, 4_Distance_Ascending and 5_Density_Descending, respectively (see the 'Input Data' block). The result 'pprb' from Calculate Responses tool is the final prospectivity map. It comprises the posterior probabilities values updated from the weights-of-evidences of all the input evidence rasters. The raster 'pprb_conf' gives the confidence in the posterior probabilities values estimated in 'pprb'; and 'pprb_std' represents the standard deviations in the posterior probability calculations because of the deviations in the calculations of the weights of evidences.

5.5 Results – WofE Modeling Prospectivity Map

The prospectivity map (Fig. 38a) is obtained from the 'Calculate Responses' tool of the WofE plugin. This raster contains the posterior probabilities values updated from the weights-of-evidences of all the input evidence rasters. Figure 38b shows the confidence in the estimated posterior probabilities values and Figure 38c represents the standard deviations in the posterior probability calculations because of the deviations in the weights-of-evidences.

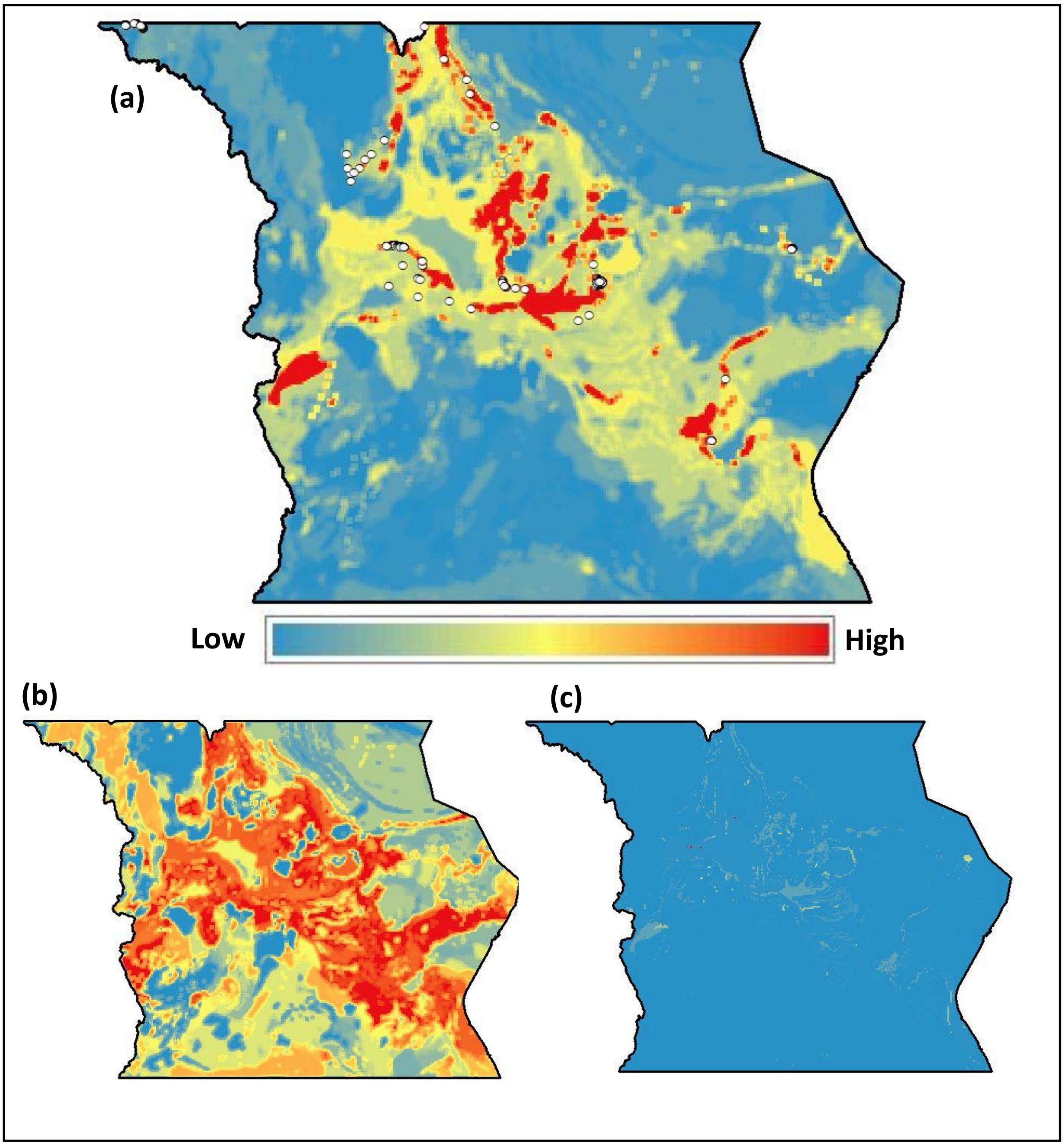


Figure 38: Results of the WofE plugin's 'Calculate Responses' tool. (a) The prospectivity map containing the posterior probabilities values updated from the weights-of-evidences of all the input evidence rasters. (b) Raster showing confidence in the estimated posterior probabilities values. (c) Raster representing the standard deviations in the posterior probability calculations because of the deviations in the weights-of-evidences calculations.

6. Summary and Conclusion

- The spatial association (i.e. **the weights**) between the targeted event, (the hypothesis) and the evidential event, (**the evidence**) can be quantified using the Bayesian 'weights-of-evidence' (WofE) model.
- The WofE plugin, discussed in this document, can implement the WofE model in QGIS, an open source GIS software.
- The plugin functionalities are demonstrated for geological case study of exploration targeting of mineral deposits using geological data.
- The plugin contains three tools – (i) Pre-process Data, (ii) Calculate Weights, and (iii) Calculate Responses, all to be implemented in the order of appearance.
- The input to each tool is the output from the previous tool.
- The final results from the plugin are: (i) the prospectivity map containing the posterior probabilities values updated from the weights-of-evidences of all the input evidence rasters. (ii) raster showing confidence in the estimated posterior probabilities values, and (ii) raster representing the standard deviations in the posterior probability calculations because of the deviations in the weights-of-evidences calculations.
- This plugin has some pre-requisites: (i) the input data should not have any missing/no data values, (ii) accepted raster format is GeoTiff files, (iii) all the input data should have the same projected coordinate system, (iv) The numerical evidence rasters should be multiclass or binary (not continuous).

References

Agterberg, F.P., 1989. Systematic approach to dealing with uncertainty of geoscience information in mineral exploration. Proceedings of the 21st APCOM Symposium, Las Vegas, USA, Chapter 18, p. 165-178.

Agterberg, F.P., Bonham-Carter, G.F. and Wright, D.F., 1990. Statistical pattern integration for mineral exploration. In: Gaál, G. and Merriam, D.F.(Eds.), Computer Applications in Resource Estimation Prediction and Assessment for Metals and Petroleum. Pergamon Press, Oxford-New York, p. 1-21.

Aspinall, P.J., and Hill, A.R., 1983. Clinical inferences and decisions-I: Diagnosis and Bayes' theorem. Ophthalmic and Physiologic Optics, v. 3, p.295-304.

Boleneus, D.E., Raines, G.L., Causey, J.D., Bookstrom, A.A., Frost, T.P., and Hyndman, P.C., 2001. Assessment method for epithermal gold deposits in northeast Washington State using weights-of-evidence GIS modeling. USGS Open-File Report 01- 501, 52 pp.

Bonham-Carter, G.F., 1994. Geographic Information Systems for Geoscientists: Modeling with GIS. Pergamon Press, Ontario, Canada, 398 pp.

Bonham-Carter, G.F., and Agterberg, F.P., 1990. Application of a microcomputer based geographic information system to mineral-potential mapping. In: Hanley, J.T., and Merriam, D.F., (Eds.), Microcomputer-based Applications in Geology, II, Petroleum, Pergamon Press, New York, p. 49-74.

Bonham-Carter, G. F., Agterberg, F. P., & Wright, D. F., 1989. Weights of Evidence modelling: a new approach to map mineral potential. *Geological Survey of Canada Paper*, 8-9.

Bonham-Carter, G.F., Agterberg, F.P. and Wright, D.F., 1988. Integration of geological datasets for gold exploration in Nova Scotia. Photogrammetry and Remote Sensing, v. 54(11), p. 1585-1592.

Carranza, E.J.M., and Hale, M., 2002. Where are porphyry copper deposits spatially localized? a case study in Benguet province, Philippines. Natural Resources Research, v. 11(1), p. 45-59.

Chudasama, B., Kreuzer, O., Thakur, S., Porwal, A., Buckingham, A. 2018a. Surficial uranium mineral systems in Western Australia: Permissive tracts and undiscovered endowment, IAEA Tecdoc – 1861, Vienna, pp. 446–614.

Heckerman, D.E., Horvitz, E.J., and Nathwani, B.N., 1992. Towards normative expert systems: Part I the Pathfinder project. Methods of Information in Medicine, v. 31, p. 90-105.

Jia, W., & Wang, G. (2019). Multiple level prospectivity mapping based on 3D GIS and multiple geoscience dataset analysis: a case study in Luanchuan Pb-Zn district, China. Arabian Journal of Geosciences, 12(11), 332.

Kemp, L.D., Bonham-Carter, G.F., Raines, G.L. and Looney, C.G., 2001, Arc-SDM: Arcview extension for spatial data modeling using weights of evidence, logistic regression, fuzzy logic and neural network analysis.

Lusted, L.B., 1968. Introduction to Medical Decision Making. Charles Thomas, Springfield, 271 pp.

Porwal, A., and Hale, M., 2000. GIS-based weights-of-evidence analysis of multi-class spatial data for predictive mineral mapping: a case study from Aravalli province, western India. Proceedings of XIV International Conference on Applied Geologic Remote Sensing, Las Vegas, Nevada, p. 377-384.

Porwal, A., Gonzalez-Alvarez, I., Markwitz, V., McCuaig, T. C., & Mamuse, A. (2010). Weights-of-evidence and logistic regression modeling of magmatic nickel sulfide prospectivity in the Yilgarn Craton, Western Australia. Ore Geology Reviews, 38(3), 184-196.

Raines, G.L., 1999. Evaluation of weights of evidence to predict epithermal gold deposits in the Great Basin of the western United States. Natural Resources Research, v. 8(4), p. 257-276.

Reggia, J.A., and Perricone, B.T., 1985. Answer justification in medical decision support systems based on Bayesian classification. Computers in Biology and Medicine, v. 15(4), p. 161-167.

Singh, P.K., Grunsky, E.C., Keller, E.V., and Keller, C.P., 1993. Porphyry copper potential mapping using probabilistic methods and geographic information systems in British Columbia. In Eyes on the future: Proceedings of the 1993 International Symposium on GIS, p. 381-394.

Spiegelhalter, D.J., and Knill-Jones, R.P., 1984. Statistical and knowledge-based approaches clinical decision support systems, with an application in gastroenterology. Journal of the Royal Statistical Society, A, Part 1, p. 35-77.

Tao, J., Yuan, F., Zhang, N., & Chang, J. (2019). Three-Dimensional Prospectivity Modeling of Honghai Volcanogenic Massive Sulfide Cu–Zn Deposit, Eastern Tianshan, Northwestern China Using Weights of Evidence and Fuzzy Logic. *Mathematical Geosciences*, 1-32.

Williams, N. D., Elliott, B. A., & Kyle, J. R. (2020). A Predictive Geospatial Exploration Model for Mississippi Valley Type Pb–Zn Mineralization in the Southeast Missouri Lead District. *Natural Resources Research*, 1-26.

Wise, T. (2019). Prospectivity modeling of the Olympic Cu–Au Province. *MESA Journal 90* (2), 36 – 41.

Wright, D.F., and Bonham-Carter, G.F., 1996. VHMS favorability mapping with GIS-based integration models, Chisel-Andersen Lake area. Geological Survey of Canada, Bulletin, v. 426, p. 339-376.