NDVI Differencing Tool – User Guide

Purpose

The purpose of this tool is to assist in measuring the change in vegetation in a given area over time.

The tool uses Landsat imagery from two different dates to calculate the difference in the Normalised Difference Vegetation Index (NDVI) within the area designated by a polygon. It then measures the areas of greatest change determined by thresholds which may be either set by the user or determined automatically.

Finally, a map is produced visualising the results. Additionally, at various stages geolocated image files will be generated which can be used for further analysis with other programs.

Installation

1. Prepare for install

The following instructions assume you will be using **git** and **conda**. Directions for installing git can be found here, and for Anaconda here.

2. Clone the repository with git

Open **Git Bash** (**Start** > **Git** > **Git Bash**) and navigate to the directory you want to use for the repository. Execute the following:

```
git clone https://github.com/ejleighton/ejl-pyproj.git
```

This will download and create a copy of the repository in a folder named **ejl-pyproj**.

Alternatively, if you do not wish to use git, you can go to the repository homepage on GitHub, click the green *"Code"* button and select *"Download ZIP"*. Once complete, unzip the folder into your desired directory.

3. Create a conda environment

The folder installed in step 2 contains an environment.yml file which can be used to reproduce the required environment to run the script.

Using Anaconda Navigator the environment can easily be imported by going to the **Environments** panel and clicking the **Import** button at the bottom, then navigating to the environment.yml file from the import wizard.

To create the environment from the command line (Windows users should run **Anaconda Prompt** from the **Start** menu), first navigate to the directory you unpacked/cloned the repository to and run:

```
conda env create -f environment.yml
```

This will install the following packages and their dependencies:

- python 3.8.8
- cartopy 0.18.0
- rasterio 1.2.0
- geopandas 0.9.0
- scikit-image 0.18.1

Configuration

1. Prepare your data

Your data should consist of:

- A shapefile of the study area
- Single band geolocated images for both the visible red and near infra-red bands in a Universal Transverse Mercator (UTM) projection

The tool has been tested with Landsat imagery which is available directly from USGS EarthExplorer. Two datasets have been provided for demonstration, one for the Chernobyl Exclusion Zone in the Ukraine, and one for the Mount St. Helen's National Volcano Monument in Washington State, USA.

2. Set up config file

The script requires a number of variables to be set in a file named **myconfig.py** in the same directory as **ndvidiff.py**. The variables are as follows:

Variable	Type	Description
maptitle	*string*	Text title to appear above the final map
newRed	*string*	Path and filename for the red band of the more recent image
newNIR	*string*	Path and filename for the NIR band of the more recent image
oldRed	*string*	Path and filename for the red band of the older image
oldNIR	*string*	Path and filename for the NIR band of the older image
shapefile	*string*	Path and filename for the shapefile containing the study area boundary
spatialref	*string*	Spatial reference for analysis in format 'epsg:#####' Choose a spatial reference system which uses metres
posthresh, negthresh	*float*	Choose thresholds between **1** and **-1** to apply to the NDVI difference image. Pixels with a value above **posthresh** will be classified as *significant positive change*. Pixels with a value below **negthresh** will be classified as *significant negative change*. If a value of 0 is set for both, thresholds will be determined using Otsu's method.
savepath	*string*	Path to save output images

Two config files are provided for use with the sample data, **myconfig.py**, and **MSHmyconfig.py**. To change which is used simply rename MSHmyconfig.py to myconfig.py. You should rename the original file if you want to keep it.

3. Running the script

From the command line (Windows users should run **Anaconda Prompt**), navigate to the directory you have installed the tool and ensure you are in the correct environment by executing e.g.:

```
conda activate ejlproj
```

Run the script by executing:

```
python ndvidiff.py
```

Analysis Methodology

The Normalised Difference Vegetation Index (NDVI) was developed in 1974 (Rouse et al, 1974) as a method of identifying healthy vegetation using satellite imagery. Photosynthesising vegetation absorbs light in the visible red spectrum for energy, but strongly reflects light in the near infra-red (NIR) (Gates, 1980), and therefore appear bright to instruments measuring NIR reflectance, and dark to instruments measuring visible red reflectance.

The script clips the rasters representing these bands (provided by the user) to the study area bounding box and calculates the NDVI using the following formula:

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

The result is an array of values between 1 and -1. Higher values generally correspond to healthy, dense vegetation, while low or negative values generally indicate barren areas (USGS, nd).

NDVI Differencing is commonly used to detect change in vegetation between images (Chernobyl Exclusion Zone: Davids & Doulgeris, 2007; Mount St. Helens: Marzen et al, 2011). This is achieved in this instance by simply subtracting the NDVI array of the older image from the NDVI array of the newer image. This is output to an image file with pixel values representing the change in NDVI between the two imaging dates.

As we are only interested in values inside the study area, the NDVI difference image is masked by setting pixels that fall outside the study area to an arbitrary value outside the normal range for NDVI difference, in this case -9999. This allows these pixels to be easily identified and flagged as 'no data'.

The image is then classified to identify areas of significant change, both positive and negative. This is done using thresholds set by the user, however if the user sets thresholds to zero, Otsu's method (Otsu, 1979) will be used to determine the thresholds. As Otsu's method seeks the best threshold to divide grey histogram value peaks it will tend towards including all change, therefore it is recommended to manually set thresholds.

The image of classified values is then polygonised and transformed into the user set spatial reference system, allowing the total area of significant change to be calculated.

Expected Results

The script provides two outputs. The first is in the form of printing to the command line, for example you should see something like the figure 1 when running the Chernobyl dataset (threshold dependent)

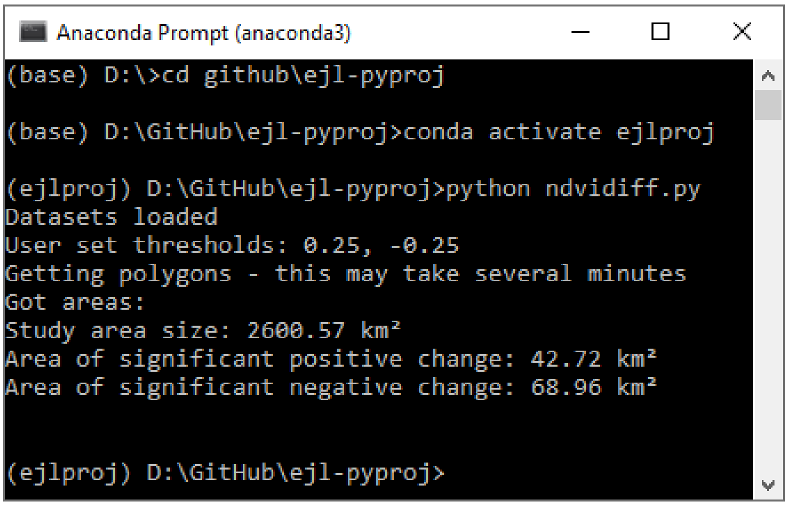

Figure 1 – CEZ Command Line output

It will also output a map of the data as demonstrated in figure 2 below. The map will be in the UTM projection of the original input rasters and shows the NDVI Difference raster colormapped with a gradient, and the polygons indicating the areas of significant change. The areas calculated from the polygons are displayed in an information box below the map.

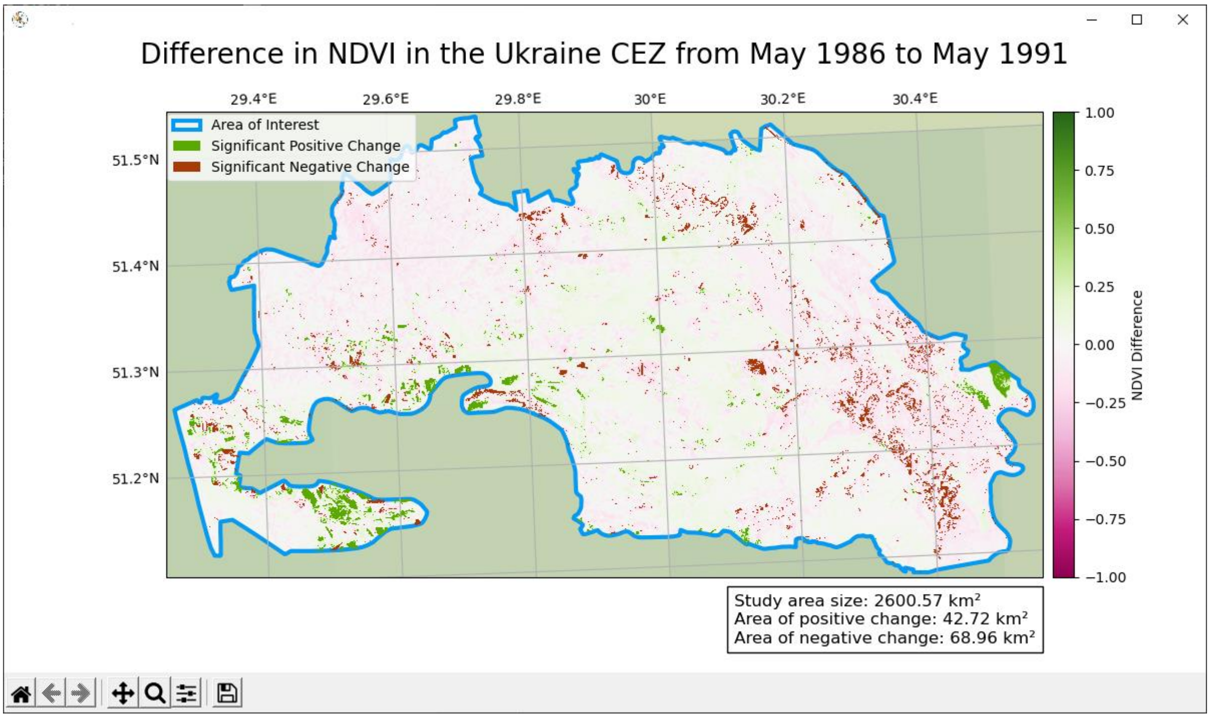

Figure 2 – CEZ Map output

Running the script on the Mount St. Helens sample data returns the output in figure 3. Note that in this case the thresholds were determined automatically. The map output generated (figure 4) is also saved as a PNG to the output folder for publication or sharing. In this folder you will also find some GeoTiff files generated during the analysis, which may be useful for analysis in other GIS or image processing applications.

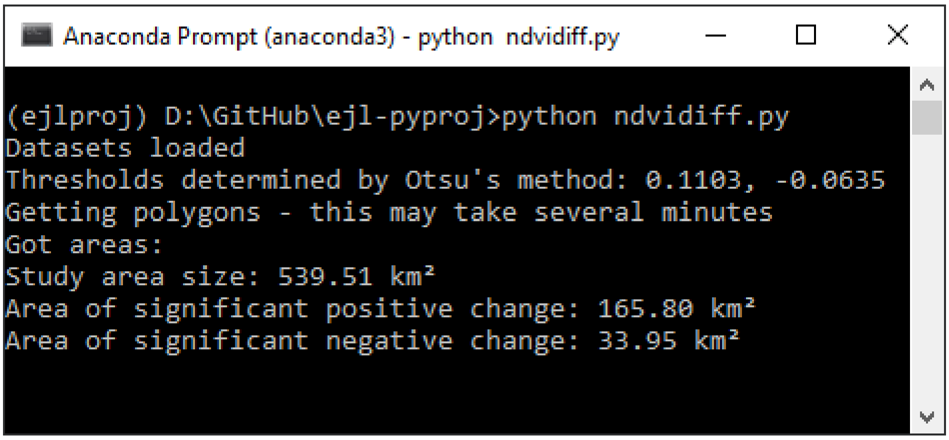

Figure 3 - Mount St. Helens CLI output

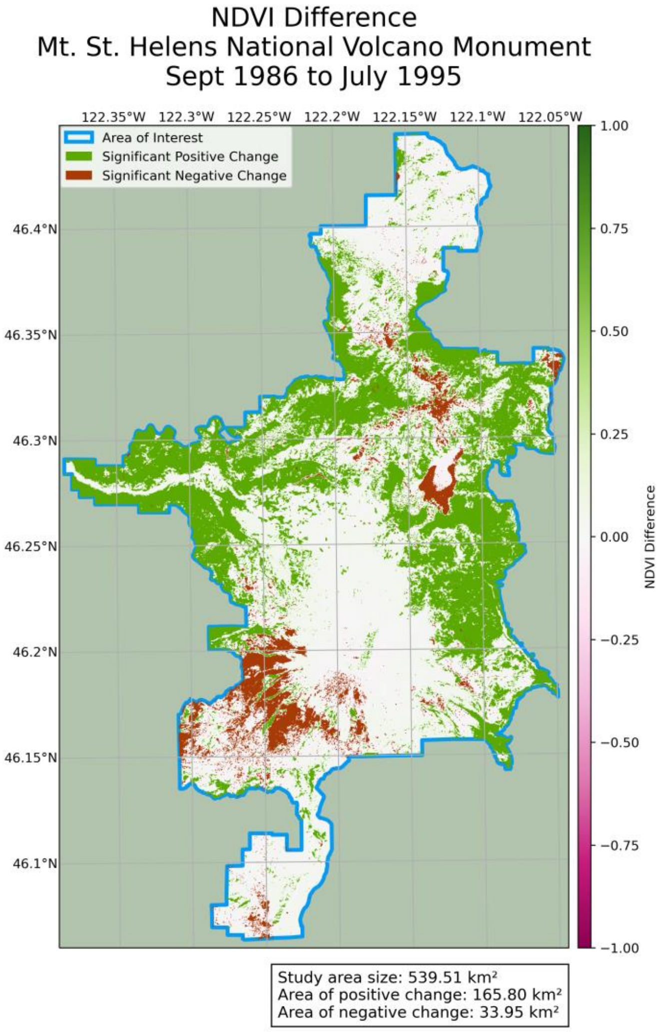

Figure 4 - Mount St. Helens map output

Troubleshooting

There are not currently any known issues. If you do have problems running the script, please ensure that you have provided the expected input datasets and configured **myconfig.py** as described in the configuration section above. If the issue persists, or you believe you have identified a bug, you can raise it on the repository issues page.

References

Davids, C. and Doulgeris, A., 2007. Unsupervised change detection of multitemporal Landsat imagery to identify changes in land cover following the Chernobyl accident. In *2007 IEEE International Geoscience and Remote Sensing Symposium* (pp. 3486-3489). IEEE.

Gates, D.M., 1980. Energy budgets of plants. In *Biophysical Ecology* (pp. 345-381). Springer, New York, NY.

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W. and Harlan, J.C., 1974. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. *NASA/GSFC Type III Final Report, Greenbelt, Md*, *371*.

Marzen, L.J., Szantoi, Z., Harrington, L.M. and Harrington Jr, J.A., 2011. Implications of management strategies and vegetation change in the Mount St. Helens blast zone. *Geocarto International*, *26*(5), pp.359-376.

Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, *9*(1), pp.62-66.

USGS, no date. NDVI, the Foundation for Remote Sensing Phenology [USGS] last accessed 01/05/2021