Instructions for preparing the model

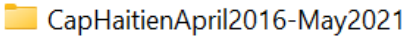
June 2022

Names: Jakko-Jan van Ek, Julia Sipkema, Mark Boeve, Niels Verouden & Raimon Bach Pareja
Commissioner: S.M.J. Arts, Ministerie van Defensie
Coach: Sytze de Bruin

Contents

1. Folder structure

Before running the model, all the folders and files that are needed by the model should be in the correct location. This structure is not mandatory, but recommended as the scripts are based on this structure and will not be able to find the correct files if the structure of folder names are different. When the structure is altered or names are different, the script needs to be altered as well.

The main folder, in which all the folders and files are stored is called "*Study_areaBegin_monthBegin_year-End_yearEnd_year*". For example: "CapHaitienApril2016-May2021" which will look like: 

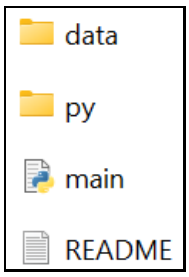


Figure 1: Folder structure main folder

In the main folder there are two folders "data" and "py", one python file "main.py" one .txt file "README.txt". The README contains additional information about the model. The main.py file is the only file that needs to be opened to run the model, all the other files are opened by this file when needed. These other python files are stored in the py folder. All the data is stored in the folder data and will be opened by the main and other python functions automatically, provided that all the data is stored in the correct folder and with the correct names.

In the data folder there should be the input folder named: "Study_areaDownloadsBegin_year-End_year". For example: "CapHaitienDownloadsApril2016-May2021". This folder should contain all the ZIP files with VV and VH Sentinel-1 data. This data can be downloaded from EO Browser (2022), see references. The names of the downloaded files do not need to be changed, because the model is created to be able to read these file names. Multiple years tif files of different years and months that want to be checked need to be stored inside this folder.Next, there should be a DEM folder containing a DEM tiff file of the study area. The folder is called "DEM" and the name of the tiff file does not have to be changed when it is downloaded from EO Browser (2022). Further there should be a folder containing the human settlement data "GlobalHumanSettlement" and a folder containing the water dataset "WaterBodies". These names do not have to be altered as well. At last, there is a folder called "training polygons" with another folder called "xxxx_xx_xx" (where xxxx_xx_xx is the date (yyyy-mm-dd), for example "2020_11_02") in which the training polygons are stored as one .shp file. This .shp file is needed to train the model, what these polygons are and how they need to be created is explained below in the section Creating training polygons.

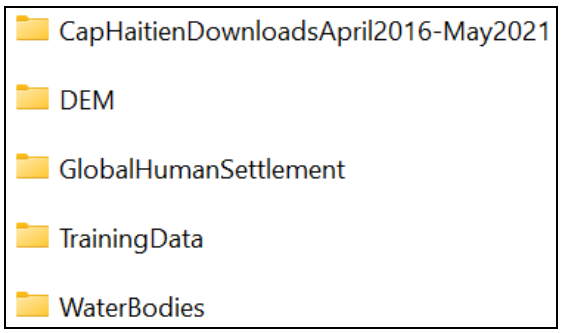


Figure 2: Folder structure of the data folder

2. Get data

2.1 EO Browser

2.1.1 Sentinel-1

From the EO Browser the *Sentinel-1* and *DEM data* are downloaded. This is done with an easy-to-use interface. First the right study area will be searched for, the "Go to Place" search bar can be used for this (See *Figure 3*, 1). Then, a bounding box is created over the complete study are to be able to download only this area to save storage space (*Figure 3*, 2).

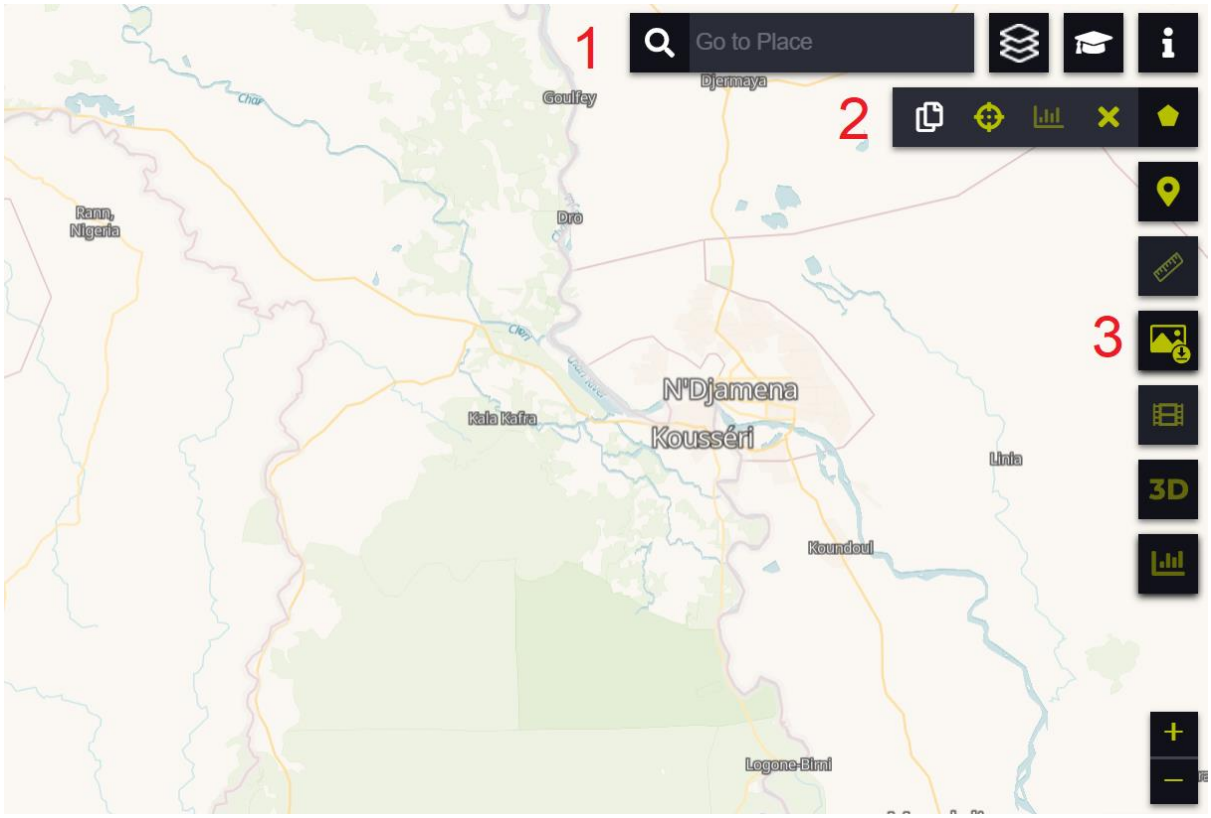


Figure 3: Interface of EO Browser

Next, the right date and datatype is selected in the "Discover" tab. Select 'Sentinel-1" under the "search" tab (Figure 4). Next select the correct begin and end date to search in a specific period

(Figure 5). A lot of data from different months and years is wanted to obtain the best results from the model.

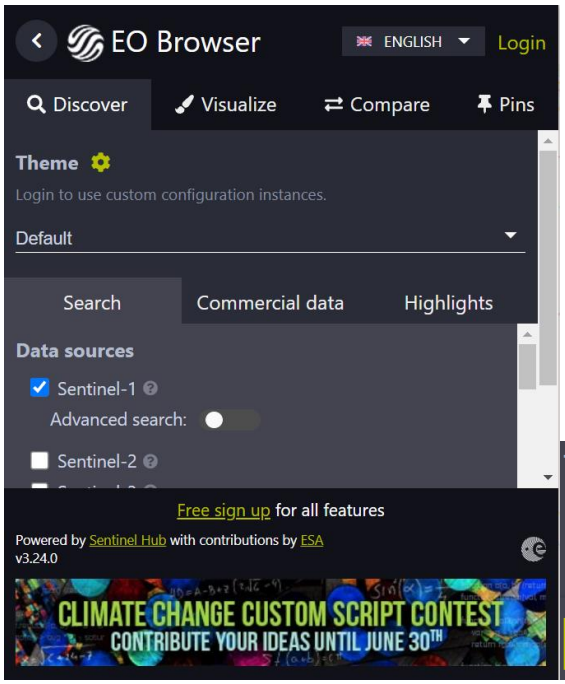


Figure 4: Discover tab of EO Browser

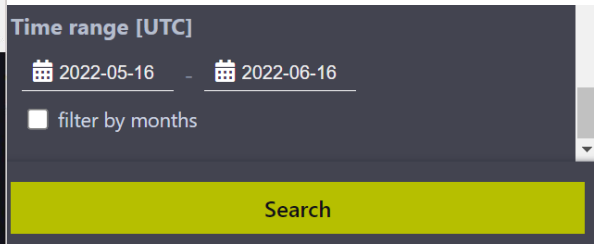


Figure 5: Discover tab of EO Browser (2)

When the data is searched for, transparent polygons appear on the map. When you click on these polygons a list of datasets appear. Check the date and click on "Visualize". Next, click on the download button on the right (*Figure 3*, 3). Go to the "Analytical" tab and set the correct settings to download the data in the wanted form:

- **Image format:** TIFF (16-bit)
- **Coordinate system:** WGS 84 (EPSG:4326)
- **Layers:**
  - VV - decibel gamma0 - radiometric terrain corrected
  - VH - decibel gamma0 - radiometric terrain corrected

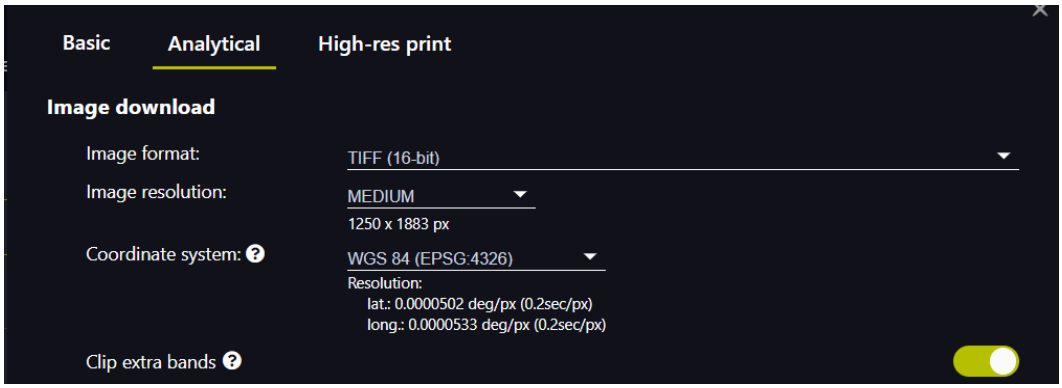


Figure 6: Download tab of EO Browser for downloading Sentinel-1 data

Figure 7: Download tab of EO Browser for downloading Sentinel-1 data (2)

When the files are downloaded, place them in correct folders according to the section *Folder structure*. Be aware that the data should be stored as .zip file, do not unzip them.

2.1.2 DEM

The DEM is downloaded from EO Browser as well, with the same bounding box. Just as with the Sentinel-1, go to the "Analytical" tab and set the correct settings to download the data in the wanted form:

- **Image format**: TIFF (16-bit)
- **Coordinate system**: WGS 84 (EPSG:4326)
- **Layers**: Grayscale

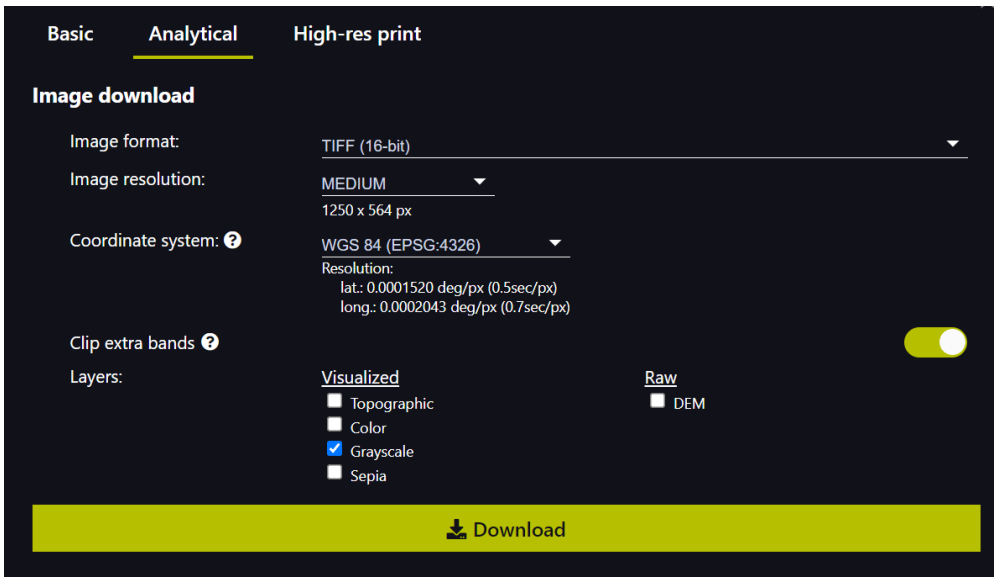


Figure 8: Download tab of EO Browser for downloading DEM data

After the downloading is finished, place the DEM in the correct folder.

2.2 Human Settlement

The human settlement data is obtained from “GHSL - Global Human Settlement Layer”, provided by the European Commission (2019). This data visualizes on a map how much people live in a specific area, which is useful for evaluating danger and risks of floods on a global scale. The following settings need to set to receive the right data:

- **Product**: GHS-POP
- **Epoch**: 2015
- **Resolution**: 9 arcsec
- **Coord. system**: WGS84

You can set these settings on the left side of the screen. After setting the right settings, click on the tile on the map you want to download and the download starts immediately.

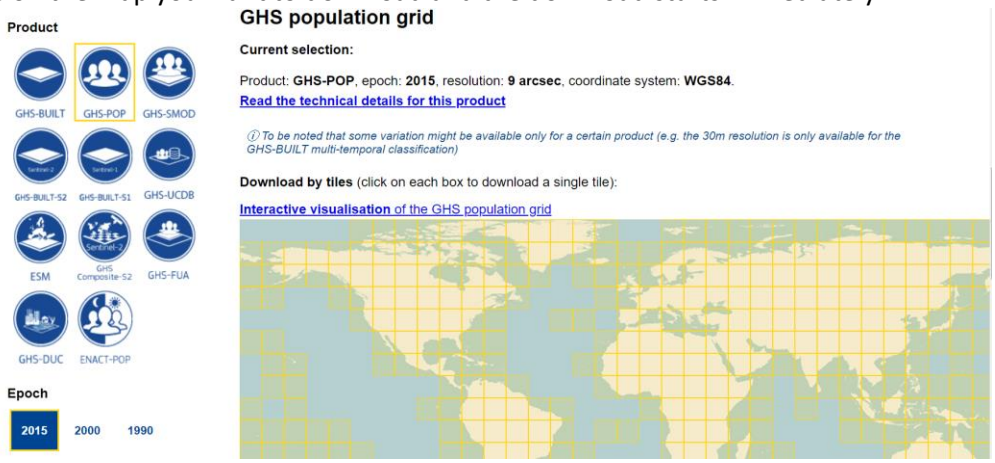
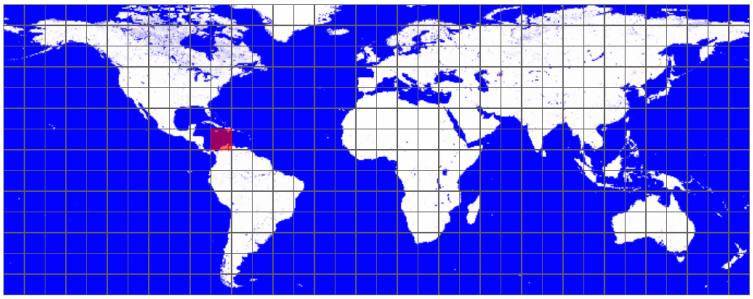


Figure 9: Interface of the Global Human Settlement Layer website

When the download is finished, place the .zip file in the correct folder: see section *Folder structure*.

2.3 Global Surface water

In the model, water bodies are masked out to prevent that water bodies are classified as frequently flooded areas. The global water data is provided by the European Commission’s Joint Research Centre (2022). For downloading the data, scroll to the world map consisting of tiles. Select the tile you want to download and select the link after *Occurrence*. After the download is finished, place this the data will be downloaded. Place this file in the correct folder, see section *Folder structure*.



Granule with top-left corner at 80W, 20N:

Occurrence: https://storage.googleapis.com/global-surface-water/downloads2020/occurrence/occurrence_80W_20Nv1_3_2020.tif
Change: https://storage.googleapis.com/global-surface-water/downloads2020/change/change_80W_20Nv1_3_2020.tif
Seasonality 2020: https://storage.googleapis.com/global-surface-water/downloads2020/seasonality/seasonality_80W_20Nv1_3_2020.tif
Recurrence: https://storage.googleapis.com/global-surface-water/downloads2020/recurrence/recurrence_80W_20Nv1_3_2020.tif
Transitions: https://storage.googleapis.com/global-surface-water/downloads2020/transitions/transitions_80W_20Nv1_3_2020.tif
Maximum extent: https://storage.googleapis.com/global-surface-water/downloads2020/extent/extent_80W_20Nv1_3_2020.tif

Figure 10: Interface of the Global Surface Water Date website

3. Creating training polygons

The model needs to be trained to be able to know which part of the area are flooded and which part is not. For training the data, there are polygons needed which are classified as “Flooded”, “UrbanFlooded”, and “Dry”. The creation of these polygons is done in ArcGIS Pro, because the selection and creation of polygons need visual assessment.

The first step is to find a date in which the study area is flooded. Next, the Sentinel-1 data is downloaded from that data, or just after, and just before that data. First, it is important to set the environment to the same CRS: WGS 84. In the Geoprocessing tab, search for “Environment” and set the output Coordinate System to GCS_WGS_1984 (Analysis → Geoprocessing → Environment (*Figure 11*).

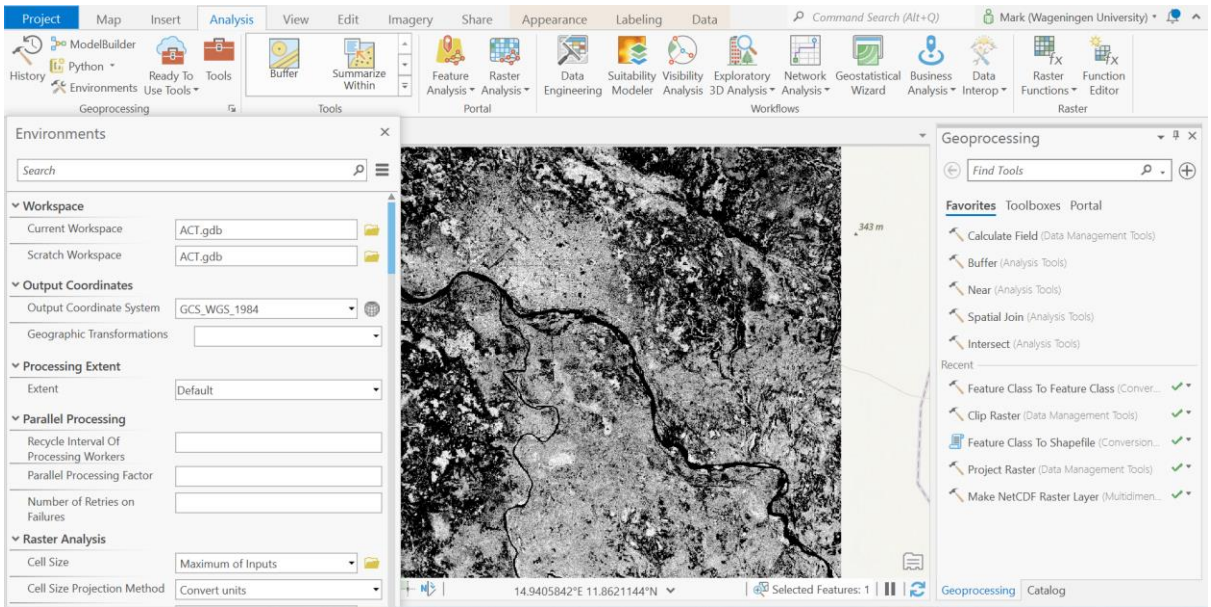


Figure 11: Setting environment CRS to GCS_WGS_1984

Next the .TIFF files are loaded in ArcGIS: one before and one after the flood. To switch between the two files floods can be recognized in relation to the knowledge about the environment, for example where the cities are. When a flood is recognized, a polygon need to be laid upon the specific area.

The first step for this is to add a feature class: Catalog → Databases → *Select the project with right mouse click* → New → Feature class. Give the feature class a name: *Train_Polys*. When this class is created, the name needs to be changed to *Train_Polys_xxxx_xx*_xx, with the date on the x's. this cannot be done while creating the feature class, because ArcGIS Pro does not accept numbers in this field. The polygon are then created by going to the edit tab and draw polygons and add them to the class: Edit → Features → Create (*Figure 12*, 1)→ Select the Train_Polys_xxxx_xx_xx feature class → Draw polygons (*Figure 12*, 2). Select multiple areas of all three classes: "Flooded", "FloodedUrban", and "Dry".

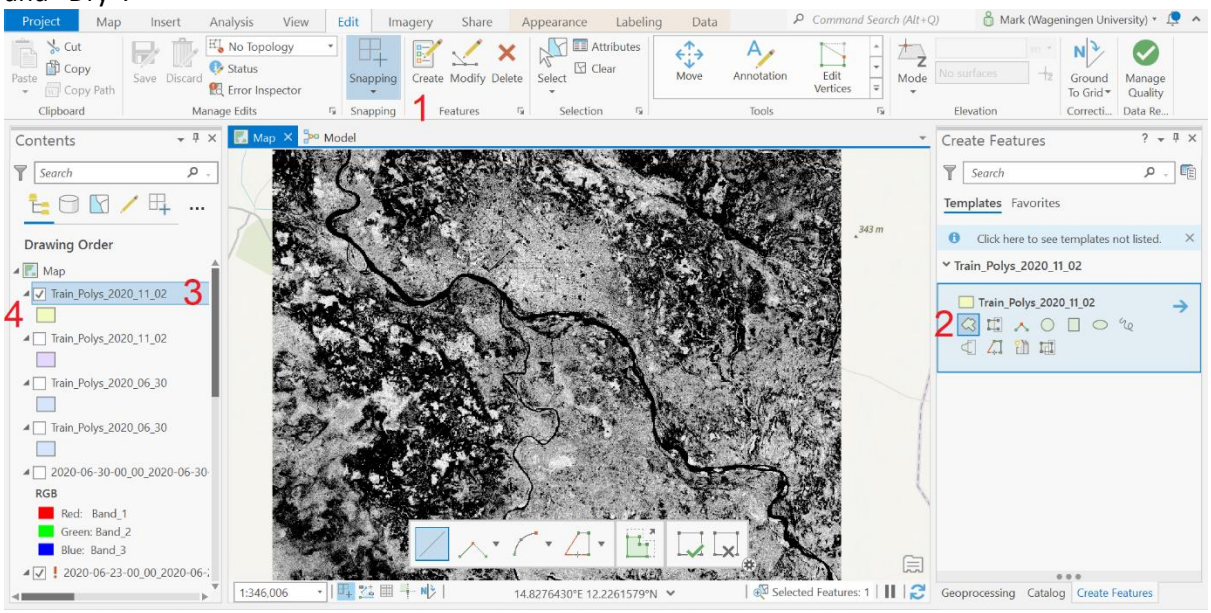


Figure 12: Creating polygons

After the polygons are created, search "Add field" in the Geoprocessing tab. Use the created feature class, Train_Polys_xxxx_xx_xx, as input. For the Field Name enter "Label", set the Field Type to "Text" and press Run (*Figure 13*).

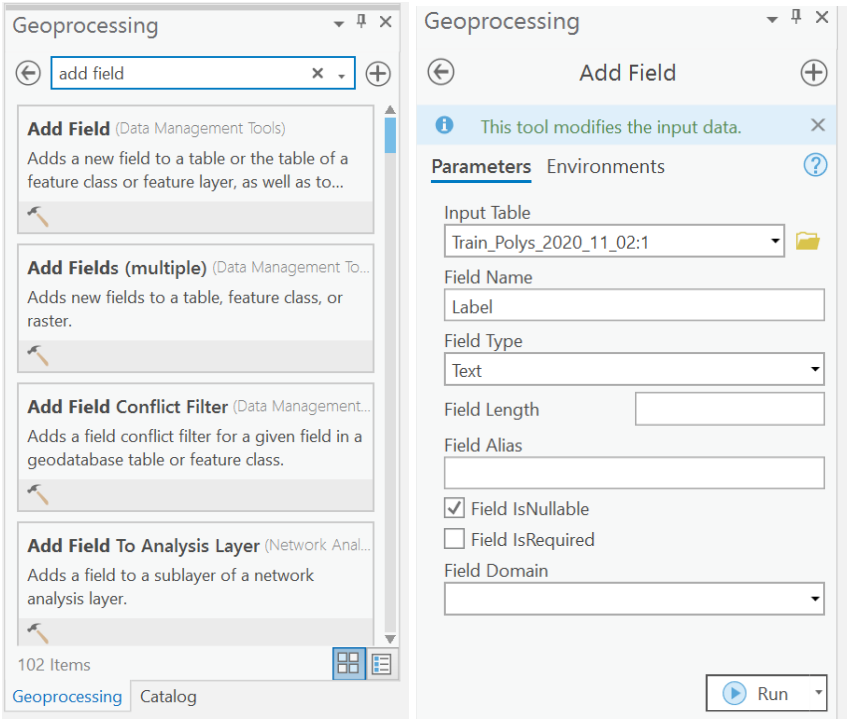


Figure 13: Add field tool in ArcGIS Pro

After adding the field "Label" to the training polyongs, go to the attribute table (Right-click on the training polygons name in the Contents pane and select *Attribute table* (*Figure 12*, 3). In the Attribute table, select one of the polygons (*click on a number on the left side of the attribute table, Figure 14*, 1). As you can see, a polygon now has a light-blue edge, because it is selected (*Figure 14*, 2). Now by turning the training polygons off and on, you can see what the area underneath the polygon is "Flooded", "FloodedUrban", or "Dry" (*Uncheck and check the boxes, Figure 12*, 4). Classify the polygons in the Label field by just typen "Flooded", "FloodedUrban", or "Dry" (Double click on a value field, *Figure 14, 3)*. When all the polygons are classified click on "clear" to clear the selection, otherwise only the selected polygon will be saved (*Figure 14*, 4).

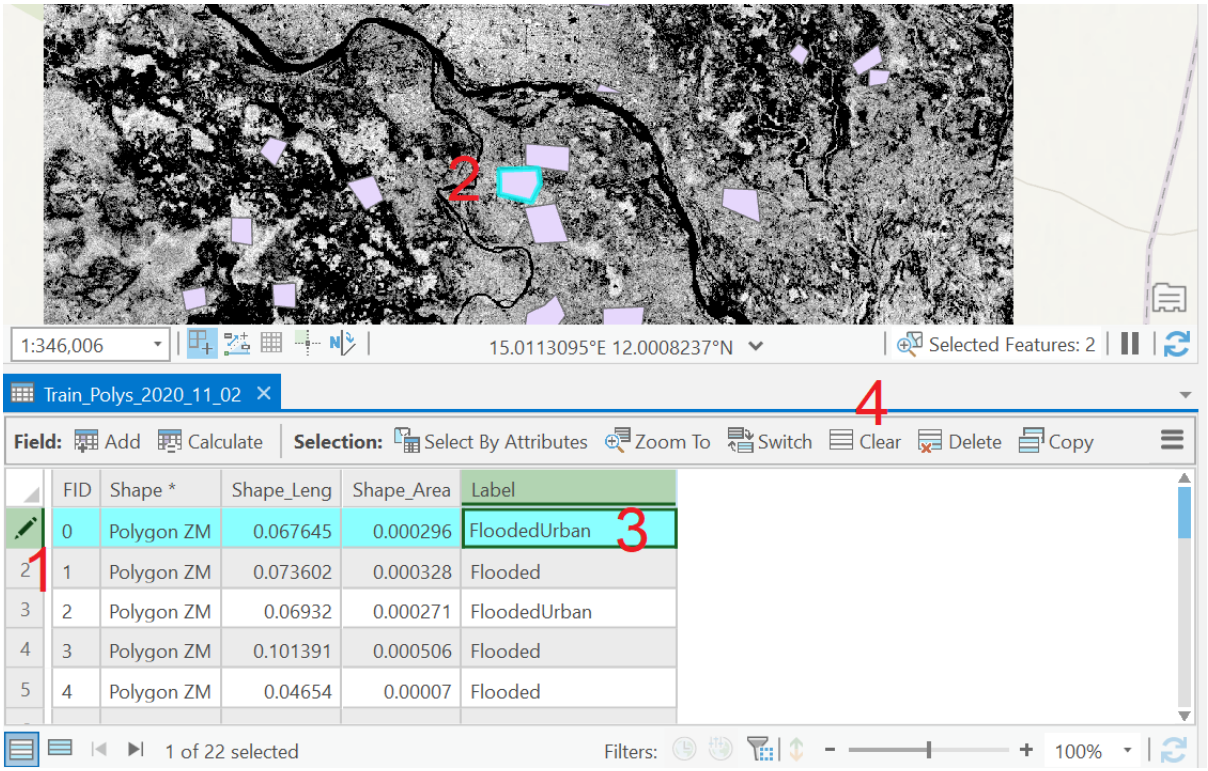


Figure 14: Classifying polygons

When the polygons are classified, they need to be exported to be able to open them in python. This can be done by right-mouse clicking on the feature class in the Contents pane and select data → export features (*Figure 12*, 3). Fill in the fields as follows:

- **Input Features**: select the training polygon class from the drop down menu
- **Output Location**: select the folder wherein the polygon data needs to be stored according to the section *Folder structure*, for example:
    - CapHaitienDownloadsApril2016-May2021\2020_11_02
- **Output Name**: TrainPolys_xxxx_xx_xx, for example TrainPolys_2020_11_02

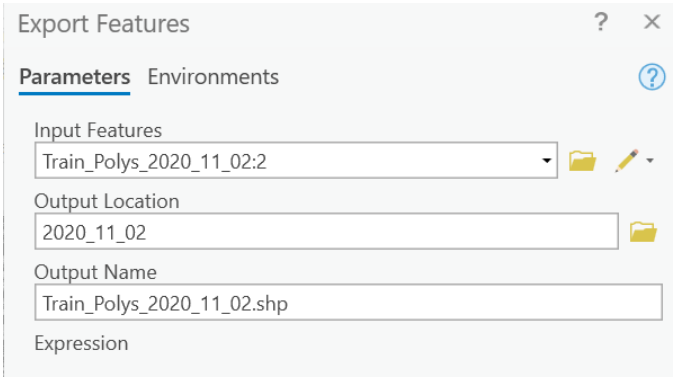


Figure 15: Export features function in ArcGIS Pro

After this is filled in, press “Ok”.

References

European Commission (2019). GHSL - Global Human Settlement Layer, Download the data produced by the GHSL, https://ghsl.jrc.ec.europa.eu/download.php?ds=pop

EO Browser (2022). EO Browser, Home, Explore, derived from, https://www.sentinel-hub.com/explore/eobrowser/

European Commission's Joint Research Centre (2022). Global Surface Water - Data Access, Individual 10°x10° files, derived at 29-06-2022, from https://global-surface-water.appspot.com/download