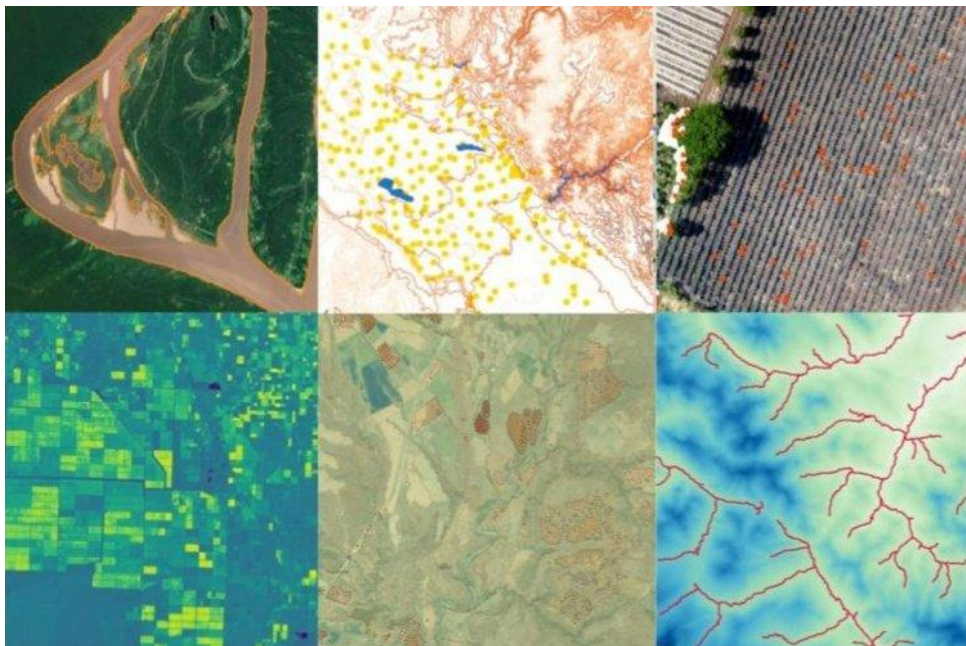




Diploma in Geospatial Analysis, Land Cover Evaluation, Geomachine Learning & Surface Modeling with Python



August 2023 - January 2024



sustainable water management

Diploma in Geospatial Analysis, Land Cover Evaluation, Geomachine Learning & Surface Modeling with Python

1. Description
2. Objectives
3. Content

Module 1: Python Fundamentals

- Session 01: Anaconda distribution and Jupyter interface
- Session 02: Python data types
- Session 03: Python loops and data structures
- Session 04: Numpy and matplotlib for water resources
- Session 05: Precipitation data analysis with Pandas
- Session 06: Precipitation and streamflow data analysis and visualization

Module 2: Vectorial data analysis with Python

- Session 01: Introduction to Fiona
- Session 02: Spatial Analysis of Total Coliforms with Fiona
- Session 03: Introduction to Shapely
- Session 04: Processing spatial vector data
- Session 05: Introduction to geopandas
- Session 06: Analysis of flooded areas with Geopandas

Module 3: Raster data analysis with Python

- Session 01: Introduction to Rasterio
- Session 02: Basic operation with rasters images - Glacier delimitation
- Session 03: Landsat 8 imagery processing and analysis with Python and Rasterio
- Session 04: Geospatial interpolation with Python, Scipy, Geopandas and Rasterio
- Session 05: Creation of an Elevation Raster from Contour Lines
- Session 06: Watershed and Stream Network Delimitation with Python and Pysheds

Module 4: Geomachine Learning for Crop Identification with Python

- Session 01: Tree Counting classification with Scikit Image and Python
- Session 02: Geospatial crop counting of palm trees
- Session 03: Crop line detection for corn crops
- Session 04: Spatial Python class for crop recognition
- Session 05: Interactive crop identification for olive trees
- Session 06: Crop fields delimitation using marching squares method with Python

Module 5: Land cover analysis with Python and geospatial libraries

- Session 01: Land cover change analysis with Python and Rasterio



sustainable water management

Session 02: Land use change analysis from vector data with Python

Session 03: Delimitation of water bodies with Canny filters

Session 04: Land cover classification using a Naives Bayes algorithm with Python

Session 05: Glacier delimitation using image segmentation with Scikit Image

Session 06: River and riparian zone delimitation based on segment analysis with Scikit Learn

Module 6: Surface process modeling with Landlab and Python

Session 01: Introduction to surface process modeling with Landlab

Session 02: River erosion simulation with Landlab

Session 03: Stream network delimitation with Landlab

Session 04: Modeling Land Evolution at Basin Scale with Python and Landlab

Session 05: Flood simulation over a surface with Landlab

Session 06: Simulation of surface and groundwater flow on a conceptual catchment

4. Trainer
5. Methodology / Examination
6. Date and time
7. Cost and payment method



sustainable water management

1. Description

Hatarilabs has developed a complete program that offers a comprehensive exploration of the basic concepts to cutting-edge techniques for geospatial analysis, land cover evaluation, geomachine learning and surface modeling with Python. The program was designed to give the student the fundamental knowledge necessary to excel in advanced environmental modeling, spatial land analytics or applied machine learning of surface features. The program content incorporates the research findings on the latest versions from the most powerful spatial analysis libraries in Python as Shapely, Geopandas, Rasterio, Scikit Image and Landlab.

The program has an approach of step-by-step learning coupled with applied examples that covers from basic Python concepts to most complex use of geospatial and machine learning libraries.

2. Objectives

The objectives of this diploma include the following:

- Understand the principles and concepts of geospatial analysis, including spatial data types, coordinate systems, projections, and spatial analysis techniques.
- Develop proficiency in Python programming language for data analysis and geospatial analysis tasks.
- Develop skills to assess and analyze land cover types using remote sensing data and satellite imagery.
- Explore the application of machine learning algorithms and techniques to geospatial data.
- Acquire knowledge and skills in surface analysis and modeling of surface processes.
- Apply the acquired knowledge to solve real-world problems related to land cover evaluation, surface modeling, and geospatial analysis.



3. Content

The diploma is divided into six modules, described below.

Module 1: Python Fundamentals

This course develops the basic concepts of Python programming under Anaconda and Jupyter. Exercises will cover the basic Python data structures, conditional statements, loops coupled with an introduction to array manipulation in Numpy, tabular data management with Pandas and applied exercises with precipitation data.

Session 01: Anaconda distribution and Jupyter interface

A session dedicated to the Anaconda distribution that runs all the capabilities of Python in Windows under the interactive environment of JupyterLab.

- Basic concepts of the Anaconda distribution.
- Jupyter Lab tools.
- Interaction with the operating system.
- Advanced features.

Session 02: Python data types

Understanding the way Python works and exploring the numerical types together with mathematical operations. Strings and boolean expressions are also reviewed. The topics covered in the session are:

- Lexical and syntax analysis.
- Types and objects (strings, list, tuples and dictionaries).
- Expressions and operators.
- Conditions and iterations.

Session 03: Python loops and data structures

A review of the conditional forms and functions in Python with practical examples of creation, indexing and management of lists, tuples, dictionaries and sets. The topics covered are:

- Lists.
- Functions.
- Loop with while and for.
- Dictionaries.



sustainable water management

Session 04: Numpy and matplotlib for water resources

This session covers the key concepts of multidimensional arrays management with Numpy together with applied examples of data visualization with Matplotlib.

- Numpy array creation.
- Operation with numpy arrays.
- Indexing and reshaping of numpy arrays.
- Introduction to data visualization with Matplotlib.

Session 05: Precipitation data analysis with Pandas

Exploration of the Pandas library for the analysis and management of data in tabular format, review of the available operations among columns, creation of dataframes and export options. The session will cover:

- Read data from text and MS Excel files.
- Filtering temporal series.
- Columns and rows manipulation.
- Export dataframes to other formats.

Session 06: Precipitation and streamflow data analysis and visualization

This session covers example scripts of long term precipitation and streamflow data collection, filtering, visualization and correlation analysis.

- Historical analysis of precipitation.
- Analysis of streamflow and rainfall relationship.
- Precipitation based calculations.
- Double axis precipitation and streamflow visualization.



sustainable water management

Module 2: Vectorial data analysis with Python

Once we have covered the fundamentals topic on Python, we will learn how to use Python for vector data analysis with different libraries such as Shapely, Geopandas and Fiona.

Session 01: Introduction to Fiona

Fiona is a Python package that reads and writes vector spatial data as ESRI shapefiles and other formats. In this session we will explore the fundamentals of Fiona. The session will cover:

- Reading of geospatial data.
- Reading multilayer data.
- Checking format drivers.
- Conversion geospatial data.

Session 02: Spatial Analysis of Total Coliforms with Fiona

Applied example of the spatial distribution of coliforms in the sea with Pandas, results are exported with Fiona as OGC vector format: geopackage. The topics covered are:

- Analysis of the value and ranges of parameters.
- Spatial visualization of concentrations.
- Exporting resulting data as geopackage.
- Calculation of correlation between variables.
- Correlation heatmap visualization using the Seaborn package.

Session 03: Introduction to Shapely

In this session we will explore the manipulation and analysis of geometries from spatial objects on a cartesian plane with Shapely. The session will include:

- Creating geometries
- Distinction between constructive (buffer, convex hull) and set-theoretic operations (intersection, union).
- Spatial analysis over elements: flooded schools, highways.
- Saving the results as an ESRI shapefile with the corresponding metadata.



Session 04: Processing spatial vector data

The fourth session covers how to clip spatial vector data, join lines and densify vertices with Fiona and Shapely libraries. This session includes:

- Opening a polygon and clip layer as Fiona elements.
- Interpretation of the geometries as Shapely polygon data types.
- Clipping the polygons and store results as an output ESRI Shapefile with the corresponding metadata.
- Selective filtering of a determined roads from a road geopackage.
- Distribution of a set of points along the merged line path.

Session 05: Introduction to geopandas

This session covers examples of geospatial data listing, indexing and manipulation with GeoPandas for polygons and lines. This session will cover:

- Reading spatial data as Geopandas dataframes.
- Working with polygons and lines.
- Spatial operation and analysis using city road data.
- Plotting data and export results as ESRI shapefiles.

Session 06: Analysis of flooded areas with Geopandas

This session will cover an applied analysis of flooded areas impact on infrastructure with Python and Geopandas. This session will cover:

- Reading data into Geopandas.
- Working with linestrings and points.
- Spatial operations among geospatial objects.
- Plotting spatial data with Folium.
- Clipping and exporting flooded areas.



Module 3: Raster data analysis with Python

Raster data processing is an important task in geospatial analysis that include cropping and reprojecting raster data, using raster math to derive new rasters, and reclassifying rasters using a set of values.

Session 01: Introduction to Rasterio

This session focuses on work with rasters using Rasterio package which allows reading raster files, plotting and make calculations with raster values. This session covers:

- Opening monoband and multiband Tiff images.
- Exploring raster dataset info and attributes.
- Analyzing spatial information (georeference).
- Reprojection for single and multiple rasters from WGS 84 UTM to WGS 84 Geographic

Session 02: Basic operation with rasters images - Glacier delimitation

Working with Python enhances the exploration of correlations between bands and values. This session develops an example of land classification using high contrast images with Python and Rasterio. This session covers:

- Opening a raster in Tif format.
- Reading and changing raster band data.
- Performing an index for glacier detection.
- Polygonizing to vector formats: lines and polygons.

Session 03: Landsat 8 imagery processing and analysis with Python and Rasterio

The session shows the commands to identify the raster array dimensions and the geospatial referencing parameters, make representations of each visible band, export band composites as geospatial rasters in Tiff format, and perform spatial operations. Here we will cover:

- Read monoband and panchromatic Tiff images.
- Explore raster spatial reference and array information.
- Plot rasters with rasterio and matplotlib tools.
- Clip rasters with shapefile in Python.
- Raster algebra example: NDVI vegetation index calculation.



Session 04: Geospatial interpolation with Python, Scipy, Geopandas and Rasterio

This session is focused on making examples about interpolation using locations with sampled values and producing geospatial rasters with the interpolate data. Here we will cover:

- Import points with elevation as an attribute.
- Creation of a triangular interpolation function.
- Generation interpolated geospatial raster in TIFF format.
- Determination of elevation from raster for another set of points.

Session 05: Creation of an Elevation Raster from Contour Lines

This session is focused on to generate an elevation raster file from a contour line with several steps of data processing. Here we will cover:

- Identification of invalid geometries
- Simplification of the polylines
- Extraction of vertices
- Creation a point geodataframe interpolated
- Geotransformation to geospatial raster in .tiff format.

Session 06: Watershed and Stream Network Delimitation with Python and Pysheds

In this session, we will learn how to delineate watersheds and stream networks with Python and Pysheds. This library requires a set of advanced data processing and spatial analysis libraries as Numpy, Pandas, Scipy, Scikit-Image, Rasterio and others. This session covers:

- Import a digital elevation model.
- Determination of flow direction raster.
- Watershed delimitation.
- Analysis of flow accumulation raster.
- Extraction of stream network.
- Basin vector/raster generation.



Module 4: Geomachine Learning for Crop Identification with Python

This module develops topics focused on crop identification / delineation with geospatial and machine learning tools of Python. The course covers the introductory concepts of the geospatial libraries and a series of machine learning applications for crop identification in olive trees, palms, agave and corn fields.

Session 01: Tree Counting classification with Scikit Image and Python

The first session covers a tree counting example based on the classification of crops in between small, medium, large and extra. This session includes:

- Explore single crop and multiple crop templates.
- Perform match template for a single and multiple crop.
- Analyze the match template array.
- Define filter criteria and count crops.
- Representation of identified crops over raster image.

Session 02: Geospatial crop counting of palm trees

The second session is focused on plant recognition from an orthophoto with Python and the machine learning libraries Scikit Learn and Scikit Image. Here we will cover:

- Coupled representation of spatial data with Matplotlib and Rasterio.
- Template extraction for spatial point data.
- Match template for a group of templates.
- Representation of the interpreted points.
- Cluster analysis with Birch algorithm

Session 03: Crop line detection for corn crops

In this session we will learn how to get information from drone imagery with Python and Scikit Learn. The exercise shows how to delineate crop rows on a corn field and provides results as a vector spatial file. This session will include:

- Perform Canny filter for edge detection.
- Identification of lines with Hough line transform.
- Convert results to geospatial data.
- Save resulting line to shapefile.
- Representation of interpreted crop rows.



Session 04: Spatial Python class for crop recognition

This session covers the complete procedure to use a Python class that recognizes plants from an orthophoto based on sample points. The Python class also creates intermediate plots and identifies plants as point shapefiles. This session will include:

- Overview of the Python class coding
- Define crop orthophoto and point plant location
- Define parameter for the match template and raster band
- Perform single match templates
- Match template for all points
- Cluster analysis and save results as shapefile

Session 05: Interactive crop identification for olive trees

Based on a modified Python class together with the interactive features of Jupyter Notebooks, a straightforward procedure was established to identify olive trees from a drone orthophoto. This session will include:

- Set up an interactive raster representation for sample olive identification
- Capture olive coordinates and identify corresponding row and cols
- Fix match template on edges
- Export crop points as geospatial files

Session 06: Crop fields delimitation using marching squares method with Python

This session deals with an applied case of crop field delimitation using Scikit Learn, Rasterio and Geopandas. Contours have been identified from an enhanced combination of raster bands and exported as polygons and linestrings. This session covers:

- Open the raster file band and find contours
- Approximate polygons with a given tolerance
- Export contours as lines and polygons
- Final representation of the resulting spatial objects.



Module 5: Land cover analysis with Python and geospatial libraries

This course is oriented to the application of Python and its spatial and machine learning libraries for applied cases of land cover analysis and land cover dynamics. The module is specifically oriented on land cover analysis techniques on applied cases.

Session 01: Land cover change analysis with Python and Rasterio

This session covers the complete procedure to create a land cover change raster from a comparison of generated vegetation index (NDVI) rasters by the use of Python, Numpy and Rasterio. This session covers:

- Opening Landsat image bands.
- Comparison of matrix sizes and geotransformation parameters.
- Compute and plot NDVI image.
- Generation of contours of land cover change.
- Analysis of deforestation.

Session 02: Land use change analysis from vector data with Python

This session covers the trend of increasing or decreasing change of certain land covers or land uses using vector data on different periods over a determined zone. Here we will:

- Spatial land use information processing at regional level.
- Quantification of land use type area.
- Analysis of increasing or decreasing land use types.

Session 03: Delimitation of water bodies with Canny filters

This session will use Python together with Scikit Learn and geospatial libraries to delineate water bodies and provide results as a vector spatial file. This session covers:

- Reading raster bands.
- Perform Canny filter.
- Exporting edges as geospatial shapefile.



sustainable water management

Session 04: Land cover classification using a Naives Bayes algorithm with Python

This session will perform an applied case of land cover classification from a multispectral image in Python using the Naives Bayes algorithm implemented on the Scikit Learn package. This session covers:

- Reading a panchromatic image in Python.
- Raster classify with Naives Bayes algorithm implemented on the Scikit Learn package.
- Analysis of classification accuracy.
- Exporting the predicted raster.

Session 05: Glacier delimitation using image segmentation with Scikit Image

This session covers an example of glacier delimitation from a geospatial raster using the SLIC algorithm (k-means clustering). Segmentation results are exported as ESRI Shapefiles and total glacier area is calculated on hectares.

- Read the geospatial and swap axes to a image format
- Create a mask to extract only the glacier area.
- Perform segmentation over the masked area.
- Create a geopandas dataframe from the glacier area.

Session 06: River and riparian zone delimitation based on segment analysis with Scikit Learn

This is an applied example of river and riparian zone delimitation from a river area on the Amazon rainforest. The unsupervised segmentation works on a geospatial raster where different zones corresponding to the rainforest, river and riparian zone were defined. Here we will cover:

- Open a geospatial raster from the Amazon rainforest.
- Perform a segment analysis with the SLIC algorithm.
- Classification plot with raster average color.
- Export polygon as geospatial ESRI shapefiles.



Module 6: Surface process modeling with Landlab and Python

Landlab is a Python library for the numerical simulation of surface processes. The library is designed for scientific fields and calculates the dynamics of earth surface such as geomorphology, hydrology, glaciology, stratigraphy and others related. This course covers applied examples using Landlab and other libraries for different cases of surface analysis.

Session 01: Introduction to surface process modeling with Landlab

This session covers the basic parts of surface simulation with Landlab. Concepts as raster and hexagonal model grids are discussed and applied to erosion modeling over a faulted surface. This session covers:

- 1D erosion modeling with Python and Numpy.
- Definition of model grids and 2d erosion modeling with Landlab.
- Implementation of the LinearDiffuser class and surface evolution analysis.

Session 02: River erosion simulation with Landlab

This session deals with modeling of the landscape evolution with an active uplifting process. The example helps us understand the setup of an uplifting raster, the patterns of erosion and the formation of flow channels.

- Definition of model grids and setup of boundaries at grid edges.
- Calculate flow routing and simulate stream erosion.
- Plot of final surface and drainage area.
- Create a geopandas dataframe and export shapefile.

Session 03: Stream network delimitation with Landlab

Based on the several components of Landlab together with other Python packages some procedures can be established to extract stream networks from a digital elevation model and export as vector spatial data formats. The session includes:

- Open a raster file and create a Landlab model grid.
- Compute flow accumulation and fill depressions.
- Profile channels with defined accumulation area threshold.



Session 04: Modeling Land Evolution at Basin Scale with Python and Landlab

The session deals with the simulation of land evolution from erosion at basin scale for 100 thousand years; inputs come from geospatial rasters and output data is exported as Ascii raster files.

- Import a geospatial raster and extract array.
- Setup a model grid from raster data and parameters.
- Simulate river erosion over 100 000 years every 1000 years.

Session 05: Flood simulation over a surface with Landlab

This is an applied case of an overland flow simulation from a flood period and post-event period with Landlab. The case covers the spatial analysis of surface water and the flow rates at discharge points.

- Reading a raster dem file with Landlab.
- Define storm rate and duration.
- Simulate overland flow for flood and after.
- Plot final surface water depth and discharge hydrographs.

Session 06: Simulation of surface and groundwater flow on a conceptual catchment

This session covers a simulation example of groundwater flow and groundwater discharge with the GroundwaterDupuitPercolator component of Landlab. Simulation runs for many time steps and results are plotted on charts and grids.

- Definition of the aquifer and initial water table elevation.
- Setup of the hydraulic parameters and groundwater flow component.
- Groundwater simulation for 800 time steps and store partial flows.
- Representation of the final water table elevation and partial and total fluxes.



sustainable water management

4. Trainer

Saul Montoya M.Sc. - Hydrogeologist - Numerical Modeler

Mr Montoya is a Civil Engineer who graduated from the Catholic University in Lima with postgraduate studies in Management and Engineering of Water Resources (WAREM Program) from Stuttgart University – Germany with a mention in Groundwater Engineering and Hydroinformatics. Mr Montoya has a solid analytical capacity for interpreting, conceptualizing and modeling the surface and underground water cycle and their interaction. He is in charge of the numerical modeling for contaminant transport and remediation systems of contaminated sites. Inside his hydrological and hydrogeological investigations, Mr Montoya has developed a holistic comprehension of the water cycle, understanding and quantifying the primary hydrological dynamic process of precipitation, runoff, evaporation and recharge to the groundwater system.

Over the last nine years, Saul has developed two websites for knowledge sharing in water resources: www.gidahatari.com (Spanish) and www.hatarilabs.com (English), that have become relevant due to their applied tutorials on groundwater modeling, spatial analysis and computational fluid mechanics.



sustainable water management

5. Methodology / Examination

Mode: Online with streaming - Synchronous

Some details about the diploma methodology:

- The manuals and files for the exercises will be delivered through our online platform.
- The course will be developed by video streaming with live support and interaction.
- Recorded **videos** will be available on our eLearning platform.
- There is online support for questions regarding the exercises developed through email and meetings.
- Video of the classes will be available for six months.

The exams and certification are organized as follows:

- The program has three exams that comprise the content of 2 courses.
- A digital certificate is available at the end of the program upon exam approval.
- To receive the digital certificate, you must submit the exams on the following date:
 - First exam before 30st September 2023,
 - Second exam before 30st November 2023.
 - Third exam before 30st January 2024.



6. Date and time

The course is offered in sessions of approximately 1.5 to 2 hours. All sessions start at 6 pm Central European Time (CET) - Amsterdam Time.

Module 1 - August 2023

- Week 1: Tuesday 15 and Thursday 17, 6 pm to 8 pm (Amsterdam Time).
- Week 2: Tuesday 22 and Thursday 24, 6 pm to 8 pm (Amsterdam Time).
- Week 3: Tuesday 29 and Thursday 31, 6 pm to 8 pm (Amsterdam Time).

Module 2 - September 2023

- Week 1: Tuesday 12 and Thursday 14, 6 pm to 8 pm (Amsterdam Time).
- Week 2: Tuesday 19 and Thursday 21, 6 pm to 8 pm (Amsterdam Time).
- Week 3: Tuesday 26 and Thursday 28, 6 pm to 8 pm (Amsterdam Time).

Module 3 - October 2023

- Week 1: Tuesday 10 and Thursday 12, 6 pm to 8 pm (Amsterdam Time).
- Week 2: Tuesday 17 and Thursday 19, 6 pm to 8 pm (Amsterdam Time).
- Week 3: Tuesday 24 and Thursday 26, 6 pm to 8 pm (Amsterdam Time).

Module 4 - November 2023

- Week 1: Thursday 02, 6 pm to 8 pm (Amsterdam Time).
- Week 2: Tuesday 07 and Thursday 09, 6 pm to 8 pm (Amsterdam Time).
- Week 3: Tuesday 14 and Thursday 16, 6 pm to 8 pm (Amsterdam Time).
- Week 4: Tuesday 21, 6 pm to 8 pm (Amsterdam Time).

Module 5 - December 2023

- Week 1: Tuesday 05 and Thursday 07, 6 pm to 8 pm (Amsterdam Time).
- Week 2: Tuesday 12 and Thursday 14, 6 pm to 8 pm (Amsterdam Time).
- Week 3: Tuesday, 19 and Thursday, 21, 6 pm to 8 pm (Amsterdam Time).

Module 6 - January 2024

- Week 1: Tuesday 09 and Thursday 11, 6 pm to 8 pm (Amsterdam Time).
- Week 2: Tuesday 16 and Thursday 18, 6 pm to 8 pm (Amsterdam Time).
- Week 3: Tuesday, 23 and Thursday, 25, 6 pm to 8 pm (Amsterdam Time).



sustainable water management

7. Cost and payment method

The normal cost of the program is \$ 1250 dollars.

- Payment with credit card or debit card

Please send an email to saulmontoya@hatarilbas.com to receive a payment form.

- Transfer by Western Union.

Please send an email to gabrielacarita@gidahatari.com for payment details.