Diploma in Python for Water Resources and Geoscience - Asynchronous
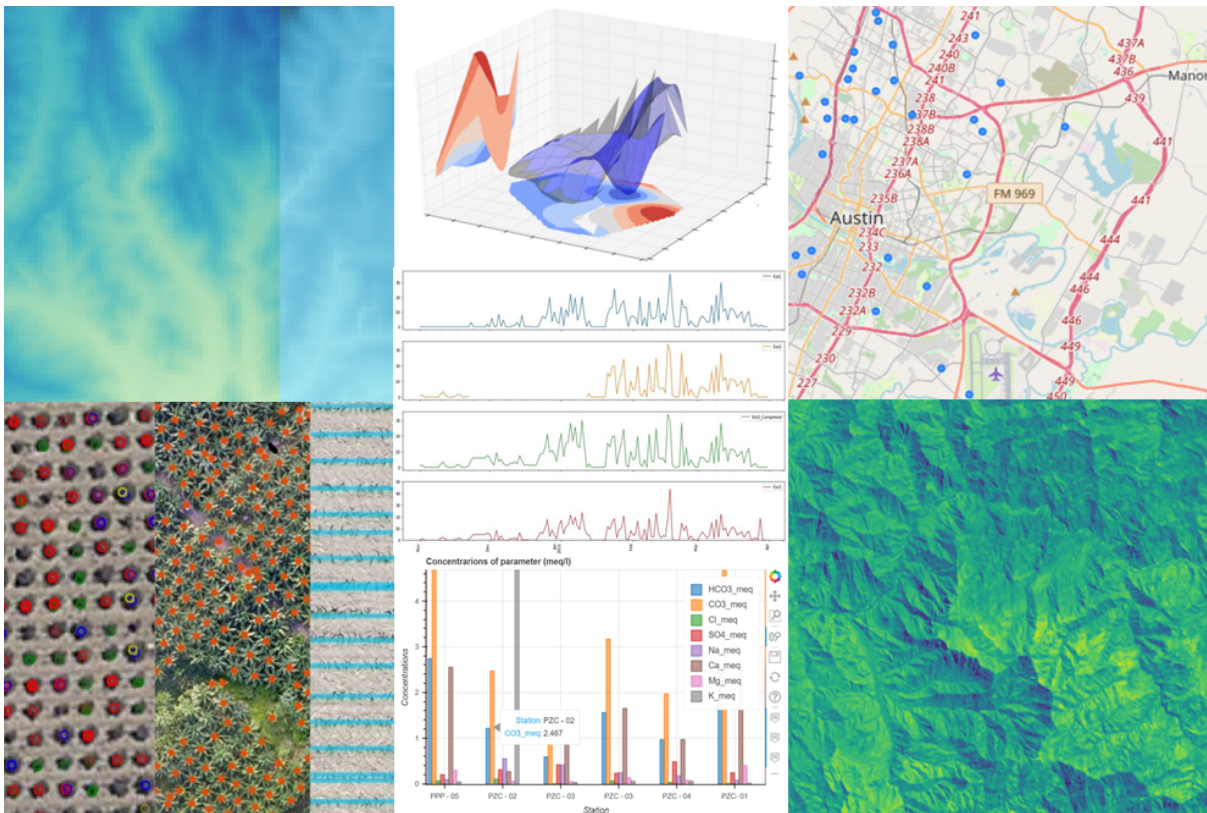
Diploma in Python for Water Resources and Geoscience - Asynchronous

1. Description

2. Objectives

3. Content

    Module 1: Python for Hydrology - Part 1

        Session 1: Anaconda interface

        Session 2: Python data types

        Session 3: Python loops and data structures

        Session 4: Numpy and matplotlib for water resources

        Session 5: Precipitation data analysis with Pandas

        Session 6: Precipitation and streamflow data analysis and visualization

    Module 2: Python for Hydrology - Part 2

        Session 1: Rainfall statistics with Scipy I

        Session 2: Rainfall statistics with Scipy II

        Session 3: Filling missing precipitation data

        Session 4: Long term temperature data analysis

        Session 5: Interpolation of Precipitation Data with Python and Matplotlib

        Session 6: Climate variable exploration from multiple climate stations

    Module 3: Data visualization in Python

        Session 01: Matplotlib

        Session 02: Seaborn

        Session 03: Bokeh

        Session 04: Pyvista

        Session 05: Folium

        Session 06: Altair

    Module 4: Applied geospatial 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: Raster data management with Rasterio and Python

        Session 05: Introduction to Geopandas for flooded areas analysis

        Session 06: Glacier delimitations with Python and Rasterio.

    Module 5: Machine learning in Python for water resources and geosciences

        Session 01: Water chemistry cluster analysis

        Session 02: Crop identification

        Session 03: Fill missing precipitation from multiple stations and climate variables

        Session 04: Geological modeling

        Session 05: Delimitation of water bodies with Canny filters

sustainable water management

1. Description

Hatarilabs presents its educational program designed for mastering Python in real professional and academic environments. The program has an extensive practical work that goes from the basic concepts of Python, Numpy and Pandas to specific applications in water resources and geosciences coupled with geospatial analysis and machine learning.

We are sure that Python is a tool / asset for hydrologists, hydrogeologist, geoscientist or related professionals and we know that only practice can give you the Python level you need to apply in everyday data analysis or numerical modeling. The amount of hours, the topics covered together with the examination and certification processes give you a strong reference of Python on your professional capabilities.

This diploma is developed in asynchronous mode that allows students to be more flexible in their schedule and enjoy of instructional materials during 06 months.

Note: This diploma is the recorded version of our last synchronous diploma.

2. Objectives

This diploma is designed to provide you with the following capabilities:

- Master the basic concepts of Python and the Jupyter environment
- Become proficient in the common tools of the scientific oriented Python packages as Numpy, Pandas and Scipy
- Create full feature data visualizations for tabular, geospatial and 3D data.
- Learn and apply the most common geospatial tools for vector and raster data analysis in Python
- Have a perspective on the application of machine learning tools in Python for water resources and related fields.
- Get experience on the application of Python in numerical modeling.
- Understand the Python packages development and have the capabilities to apply new packages.

3. Content

The diploma is divided into 6 modules, each module is divided into 6 sessions. The content of every module and description of every session is described below:

Module 1: Python for Hydrology - Part 1

This course develops the basic concepts of Python programming under 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 1: Anaconda 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 2: Python data types

Understanding the way Python works and exploring the numerical types together with mathematical operations. Strings and boolean expressions are also reviewed.

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

Session 3: 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.

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

Session 4: 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 redimension of numpy arrays
- Introduction to data visualization with Matplotlib

Session 5: 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.

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

Session 6: 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

Module 2: Python for Hydrology - Part 2

Once we have covered the basic concepts of Python programming and introductory examples with water resources data we will move to more specific precipitation statistics with Scipy, analysis of long term climate data with temporal queries, spatial interpolations, multiple station data exploration and a of machine learning example for filling missing precipitation data.

Session 1: Rainfall statistics with Scipy I

This session will develop regression analysis over rainfall data, develop statistical distributions for different climate stations, and calculate rainfall values for different return periods.

- Regression analysis for rainfall
- Statistical distributions for precipitation
- Determining returning periods for rainfall

Session 2: Rainfall statistics with Scipy II

Examples of linear interpolations, evaluation of correlation among variables, determination of correlation factors and analysis of confidence intervals.

- Data linear interpolation
- Correlation analysis and coefficients
- Confidence intervals
- Multivariate frequency distributions

Session 3: Filling missing precipitation data

Implementation of different methods in Python to fill gaps in precipitation or other water related variables such as simple AA, normal ratio, multi-linear regression, and neural networks.

- Simple Arithmetic Average (AA)
- Multiple linear regressions
- Neural networks

Session 4: Long term temperature data analysis

Analysis of meteorological station data as precipitation, snow and temperature for a period of 150 years. Data exploration, metric generation and diagnose the development of the weather parameters.

- Exploring climate data with Python and Pandas.
- Temporal distribution of climate data
- Analyzing minimum and maximum temperature data.
- Implementation linear regression model to diagnose the development of the weather parameters regarding climate change.
- Checking the variability for decades.

Session 5: Interpolation of Precipitation Data with Python and Matplotlib

Example of measured precipitation analysis registered over the Tropical Storm Ida in the state of Louisiana, US. Interactive plotting of interpolated precipitation over 9 days.

- Exploring precipitation data with Python.
- Define stations location with Folium.
- Plot precipitation values with following methods: Linear, Cubic and Nearest
- Creation of a plot function for interpolation precipitation data.

Session 6: Climate variable exploration from multiple climate stations

Advanced example of exploratory data analysis of 5 climate variables in 31 stations with Python. Stations location and data was analyzed using scatter plots, radial plots and histograms.

- Read climate data as Pandas dataframe.
- Filter locations from tabular data.
- Interactive radial plots for climate variable visualization.
- Plots de la red neural

Module 3: Data visualization in Python

Following the learning process of Python programming for water resources we will develop a course focused on data visualization using different graphical libraries like Matplotlib, Seaborn, Bokeh, Pyvista, Folium and Altair. This module is aimed to learn the creation and control process of plots for an efficient and interactive data analysis.

Session 01: Matplotlib

This session is focused on learning how to use the Matplotlib library that provides blocks to create visualizations for different kinds of data and how to customize them.

- Introduction to Matplotlib.
- Creation of a simple plot using Matplotlib.
- Customizing plots.
- Bar and histogram plots.
- Scatter and 3D scatter plots.

Session 02: Seaborn

Introduction to the Seaborn library and how to visualize chemistry data using Seaborn plots as bars, histograms, scatters and 3D scatters.

- Introduction to Seaborn
- Reading water chemistry data on csv format
- Plotting graphs with Seaborn: Histograms, bar plots, box plots, pair plots and subplots.

Session 03: Bokeh

Another powerful library is Bokeh that allows creating interactive visualizations. In this session, we will learn the customization of a bokeh plot as well as focusing on how to introduce interactivity into an object.

- Introduction to Bokeh.
- Creating a line chart and combining multiple graphs.
- Customizing axes, creating legends and applying Boke's style in: Line plot, bar plot, pie plot, scatter plot, histogram.
- Creation Bokeh DataTables with scatter plot.

Session 04: Pyvista

Learn how to visualize data in Pyvista. We will develop examples to make a 3D surface and convert geospatial data ESRI shapefiles to unstructured grid Vtk format type (Vtu).

- Introduction to Pyvista.
- Making a 3D surface using a DEM file.
- Conversion geospatial data(shapefile) to 3D data using geopandas and Pyvista: definition of the elevation to 3D geometries, definition line shapefile as individual parts and creation of geospatial polygons.

Session 05: Folium

We will explore the different capabilities of the Folium library that allows creating maps quickly using a tap water and groundwater dataset.

- Introduction to Folium.
- Markers and creation point maps.
- Processing tap water data and groundwater wells.
- Visualization of geospatial data with multiple popup lines.

Session 06: Altair

In this session, we will learn to create visualizations with the Altair Python library and know how to manipulate and interact with those charts using weather data.

- Introduction to Altair
- Exploring weather data:
  - Calculation the total precipitation of each month.
  - Visualization of mean and range temperature using bar plot.
- Relationship between weather, precipitation, maximum temperature:
- Configuration to use a larger canvas and to allow interactive panning and zooming with the mouse.

Module 4: Applied geospatial data analysis with Python

Modeling surface flow, groundwater flow or any physical process on the environment is by itself a distributed process where analytical tools need to be combined with geospatial tools on a programming level. We have compiled the basic information and applied examples of the most common geospatial tools available in Python while assuring functionality on any operating system.

Session 01: Introduction to Fiona

Fiona is a Python package for the read and write of ESRI shapefiles and other formars. In this session we will explore the fundamentals of Fiona

- Reading of geospatial data:
    - Exploring metadata and geometry
    - Working with point and polygon data.
    - Creation composite plot.
- Reading multilayer data.
- Checking format drivers.
- Conversion geospatial data:
    - Shapefile to geojson.
    - Geopackage to geojson

Session 02: Spatial Analysis of Total Coliforms with Fiona

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

- Analyzing the value and ranges of parameters
- Visualization of concentrations in a map.
- Exporting resulting data as geopackage.
- Calculation of correlation between variables.
- Visualization of correlation 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.

- 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 a shapefile with the corresponding metadata.

Session 04: Raster data management with Rasterio and Python

We will explore Rasterio, a Python package for raster management and analysis, exploring its capabilities for raster read / write, extract metadata, plot and band algebra.

- Reading monoband and panchromatic Tiff images.
- Exploring raster dataset info and attributes.
- Analyzing spatial information.
- Reading raster band data.
- Raster plotting options with Rasterio and Matplotlib.
- Raster algebra example: NDVI vegetation index calculation.

Session 05: Introduction to Geopandas for flooded areas analysis

This session will cover an analysis of flooded areas impact on infrastructure with Python and Geopandas. The exercise will show how Geopandas can manage both the spatial analysis and operation among columns.

- Reading data into Geopandas.
- Working with linestrings and points.
- Spatial operation.
- Plotting data with Folium.
- Clipping and exporting flooded areas.

Session 06: Glacier delimitations with Python and Rasterio.

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.

- Importing libraries required.
- Opening a raster in format ".tif".
- Reading and changing raster bands data.
- Performing an Index for glacier detection.
- Poligonizing to geospatial data as line and polygon.

sustainable water management

Module 5: Machine learning in Python for water resources and geosciences

Algorithms of machine learning in Python are simple and efficient tools for predictive data analysis and can be applied to any field of water resources related analysis. We have developed some applied cases of machine learning prediction with the Scikit Learn and Scikit Image of a variety of topics that range from water chemistry, fill missing precipitation, crop identification, geological modeling, and land cover classification.

Session 01: Water chemistry cluster analysis

This session will apply the Principal Component Analysis algorithm to a set of water chemistry data extracted from a tabular file and will perform dendrograms.

- Reading data into a pandas dataframe.
- Describing component samples.
- Principal Component Analysis (PCA) using Scikit-Learn
- Performing hierarchical clustering:
  - Treating each data as one cluster.
  - Forming a cluster by joining the two closest data.
  - Plotting a dendrogram to divide into multiple clusters.

Session 02: Crop identification

An applied example of plant recognition and counting from a drone orthophoto with Python and the machine learning libraries Scikit Learn and Scikit Image.

- Open raster and vector files with Python.
- 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.
- Representation of clustered points.

Session 03: Fill missing precipitation from multiple stations and climate variables

This session shows an applied procedure to run a complete script for the filling of missing precipitation in one station by the use of data from multiple stations and climate variables.

- Multiple clips to stations data.
- Exploring station and climate variables.

www.hatarilabs.com   13

- Creation of a plot function for visualizing the temporal distribution of climate variables.
- Generation of a neural network.
- Dividing in train and test parameters.
- Implementing metrics to evaluate the predictions.

Session 04: Geological modeling

A geological model will be done in Python with the machine learning library Scikit Learn based on lithology from a dataset of 266 drillings.

- Importing well location and lithology.
- Definition of point cloud lithologies.
- Implementation of coordinate transformation and neural network classifier.
- Generation output grid refinement.
- Generation of regular grid VTK.

Session 05: 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.

- Importing required packages.
- Reading bands.
- Canny filter.
- Exporting edges as geospatial shapefile.

Session 06: Soil classifications machine learning

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.

- Reading a panchromatic image in Python using the Naives Bayes algorithm implemented on the Scikit Learn package.
- Analysis of statistics about the classification precision.
- Generation of classification of categories as: rivers, river , woods and pastures.
- Exporting the predicted raster.

Module 6: Applied numerical modeling with Python

Python is a general purpose language for data analysis and it has extended and specific tools to work and interact with other models, algorithms and softwares. The interaction among Python and other tools can be on the data preprocessing, model simulation, output visualization. We have researched practical examples in Python to simulate groundwater flow, land evolutions, geochemical speciation, hydraulic and hydrogeological modeling.

Session 01: Groundwater Modeling with Modflow 06 and Flopy

An overview of the main features of MODFLOW and Flopy with a practical part of groundwater modeling on steady state with MODFLOW / Flopy on a Jupyter notebook.

- Definition of temporal and spatial discretization
- Specifying solver options
- Insertion hydrogeologic parameters.
- Insertion boundary conditions with Flopy (CHD, WEL)
- Analyze model results.
- Plot hydraulic heads

Session 02: Land evolution modeling with Landlab

An applied case of Python and the Landlab library to simulate the land evolution at basin scale for 100 thousand years; inputs come from geospatial rasters and output data is exported as Ascii raster files.

- Introduction to Landlab.
- Reading raster image.
- Defining parameter fluvial, hillslope and uplift components.
- Creating a grid from raster attributes.
- Initializing the process components of the Landlab equation.
- Simulation at Land evolution at basin scale.

Session 03: Water speciation calculation with Phreeqc

This session will cover a example that calculates the distribution of aqueous species in seawater and the saturation state of seawater relative to a set of minerals

- Introduction to PHREEQC in Python.
- Creation of a PHREEQC object and defining executables and databases.
- Analyzing the solution composition.
- Working with the distribution species and saturation indices
- Running model and show simulation result.

Session 04: Hydraulic Modeling with HEC RAS and Python

We will explore the different capabilities of HEC RAS Controller library that allows opening, running, and extracting model results from HEC-RAS models.

- Installation HEC RAS Controller in Python.
- Opening hydraulic model from HEC-RAS
- Running a steady flow model
- Extracting the model results from HEC-RAS.
- Visualization and analysis of water depth profile.

Session 05: Hydrological Modeling with SWAT and Python

In this session we will learn how to use SQLite to query databases in Python to build /modify a simulation of hydrological modeling with SWAT.

- Installation SQLite library in Python
- Creation a database connection using pandas library.
- Checking meteorological variables.
- Reading specific columns like flow out.
- Carrying out statistics for a river.
- Visualization of discharge flow in time series.

Session 06: Hydrological Modeling with HEC HMS and Jython in HEC DSS Vuee

This session is focused on developing a hydrological model with hec hms and executing simulation using scripts.

- Creation an hydrological model in HEC-HMS
- Introduction to Jython in HEC DSS Vue
- Development of script with Jython in HEC DSS Vue
- Visualization of results with Jython in HEC DSS Vue

4. Trainer

Saul Montoya M.Sc. - Hydrogeologist - Numerical Modeler

Mr. Montoya is a Civil Engineer graduated from the Catholic University in Lima with postgraduate studies in Management and Engineering of Water Resources (WAREM Program) from Stuttgart University – Germany with mention in Groundwater Engineering and Hydroinformatics. Mr Montoya has a strong analytical capacity for the interpretation, conceptualization and modeling of the surface and underground water cycle and their interaction.

He is in charge of 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 main hydrological dynamic process of precipitation, runoff, evaporation and recharge to the groundwater system.

Over the last 9 years Saul has developed 2 websites for knowledge sharing in water resources: www.gidahatari.com (Spanish) and www.hatarilabs.com (English) that have become relevant due to its applied tutorials on groundwater modeling, spatial analysis and computational fluid mechanics.

5. Methodology / Examination

Mode: Offline - Asynchronous

Some details about the diploma methodology:

- Manuals and files for the exercises will be delivered on our elearning platform.
- The course will be developed by **recorded videos** and will be available on our elearning platform.
- There is support for questions regarding the exercises developed through the forum/email.
- Video of the classes will be available for 6 months.

The exams are certification is organized as follows:

- The program has 3 exams that comprise the content of 2 courses.
- Digital certificate available at the end of the program upon the exam approval.
- To receive the digital certificate you must submit the exams.

6. Cost and payment method

The promotional cost of the program is $ 1000 dollars.

Make your payment and register on this link:

https://hatarilabs.com/ht-en/diploma-in-python-for-water-resources-and-geoscience-asynchronous