Software Installation for EBD TRAINING COURSES

Author: Josef Kellndorfer

email:

GITHUB: Software and Training Documents

Training materials and documents are maintained and distributed on the code sharing platform github.com. To obtain training materials and documents, workshop participants should obtain a free github account if they don't already have one and email collaboration request to

Subject: EBD TRAINING, githubusername: YOUR-GITHUB-USERNAME

We will add you as a collaborator to the code distrubution site, which will allow you to access and download all course material ,as well as make your own contributions to the code.

Python

Many aspects of geospatial data analysis can now be performed within the python programming language and its vast suite of open source contributed modules. Many python bindings exist today in major open source and commercial image processing and geographic information systems like QGIS and ArcGIS. Learning and using python for all spatial data analysis tasks prepares a trainee well for the future.

Python is as a scripting/programming language very similar to *Matlab* or *IDL*. As such, users familiar with these languages should readlily be able to read and write python code. Python also has a tight integration with the statistical programming language R via a RPy interface, such that many statistical routines available in R can be called from within a python program. R DataFrames are mimicked in python with the powerful **pandas** package. Numerical computations and matrix operations for image analysis is tightly integrated with the **numpy** (*num*erical *py*thon) package. Raster data stacks are typically loaded with the powerful python implementation of the *Geospatial Data Abstraction Library* **gdal**. Data visualiation in python has advanced quite fast and parallels the capabilities of plotting of R via the main python plotting package **matplotlib** . The python **bokeh** package provides powerful interactive data visualization tools ready for web integration. During this course we will explore both matplotlib and bokeh tools for data visualization.

Anaconda Python Installation with Miniconda

Obtain and install **miniconda** from the Anaconda python distribution. We prefer to work with the latest python 3 version (3.6), 64-bit distribution.

The downlaod and installation instructions are available at:

https://conda.io/miniconda.html

Creating a new Environment

To create a new environment use the conda_ebd.yml file available at our github repository at https://github.com/EarthBigData/openSAR/tree/master/training. If you need to obtain access to the share please email .

After *miniconda* is installed, ensure that the environment variables are set correctly to execute "conda" and start a new terminal. In the new terminal type:

```
> conda env create -f conda_ebd.yml
```

This installs a new python environment "ebd". We will use this environment for execution of all geoprocessing tasks in the workshop. *ebd* is also the environment kernel which we will choose in the jupyter notebooks as described below.

Jupyter Notebook

To start the Jupyter notebook server witin the default webbrowser we change to the directory where the notebooks are.

```
> cd <PATH-TO-NOTEBOOK-DIRECTORY>
```

On a commandline prompt enter:

```
> source activate root   # Linux
> activate root          # Windows
```

Then start the notebook server with:

```
> jupyter notebook
```

This last command starts a local jupyter server (localhost:8888). The default webbrowser will be openeded and the jupyter notebook browser will be active. From the file menu a notebook can be selected and opened via double-click.

To stop the notebook server, use the CRTL+C keystrokes and answer "y". With this keystroke you will also find at any time the notebook server http address with it's token code which you can use to paste into any webbrowser to get access to the server in case the browswer has been closed. E.g.:

```
> [I 22:45:11.589 NotebookApp] Saving file at /github/private/servir_training/notebooks/Kel
> ^C[I 23:07:29.927 NotebookApp] interrupted
> Serving notebooks from local directory: /Users/josefk
> 7 active kernels
> The Jupyter Notebook is running at:
> http://localhost:8888/?token=684a07ecd6e6118075463018a2ea2cec918c124c90c71e4f
> Shutdown this notebook server (y/[n])? No answer for 5s: resuming operation...
```

As you see, if you don't answer "y" the server keeps running . . .

QGIS

During the course we will also make etensive use of the open source Quantum
GIS software **QGIS Version 2.18**, preferably the **64 bit** version.

To install QGIS please see download and installation instructions at qgis.org

Plugins

We will also make use of plugins for QGIS. Install plugins with the menu in Qgis

```
> Plugins > Manage and Install Plugins
```

- MultiQml

Timeseries_vrt Plugin

A plugin for visualization of SAR Time series data is provided for the traning
course. To install the plugin, copy the Windows or Linux version to the QGIS
python plugin directory

- Linux:

    cd ~/.qgis2/python/plugins unzip Timeseries_vrt_linux.zip

- Windows:

    Locate the .qgis2 path and unzip the Plugin into the python/plugins
    subdirectory

DATA SETS FOR THE TRAINING

See the DATA_HOWTO.md for download instructions to obtain the training
data sets.