**ReadMe file to run python scripts to obtain quality controlled groundwater levels and corresponding Sy**

All essential data and scripts are kept in the “example\_data” folder. Folder contains one seven items which are two ipynb files:

1. 1\_CGWB\_all\_well\_type\_filteration.ipynb
2. 2\_Specific\_yield\_extraction.ipynb

one tiff file:

1. 2\_hydrolgeology\_map\_georeferenced.tif

three csv files:

1. 1\_example\_unfiltered\_CGWB\_data.csv
2. 2\_filtered\_CGWB\_all\_wells.csv
3. Final\_2265\_GWL\_ts\_with\_ref\_sy.csv

and one docx file:

1. ReadMe.docx

Download the folder to run the following tasks.

1. To obtain quality-controlled groundwater level data.

* Open “1\_CGWB\_all\_well\_type\_filteration.ipynb” in Jupyter notebook.
* Give “1\_example\_unfiltered\_CGWB\_data.csv” as input.

It has 2000 rows × 55 columns. Rows denotes well locations. Columns are as follows:

STATE, DISTRICT, LAT, LON, SITE\_TYPE, WLCODE, Jan 2004 ...... Nov 2015, geometry.

Then, run the code.

* Output is quality controlled groundwater levels saved in “filtered\_df” dataframe containing 171 rows x 55 columns.

Note: “1\_example\_unfiltered\_CGWB\_data.csv” is a sample dataset provided to run this code.

1. To obtain specific yield values for quality-controlled groundwater levels.

* Open “2\_Specific\_yield\_extraction.ipynb” in Jupyter notebook.

To run the code two input files are required.

1. “2\_filtered\_CGWB\_all\_wells.csv”

This is the final quality-controlled groundwater levels obtained over entire India. It has 2265 rows × 55 columns. Rows denotes well locations. Columns are as follows:

STATE, DISTRICT, LAT, LON, SITE\_TYPE, WLCODE, Jan 2004 ...... Nov 2015, geometry

2. “2\_hydrolgeology\_map\_georeferenced.tif”

Then, run the code.

Output is specific yield values saved in “reclassify\_kmeans\_label” array. This array is merged as 56th column ("reference\_sy") with input file 1 (2\_filtered\_CGWB\_all\_wells.csv). Final output with quality controlled groundwater level along with reference Sy are saved in “Quality\_controlled\_well” dataframe.

Note: The “Quality\_controlled\_well” dataframe is saved as “Final\_2265\_GWL\_ts\_with\_ref\_sy.csv” and already provided in the “example \_data” folder.