README file for the "Natural Flow Regimes in Poland (NFRPL)" data set

Content: Classification of natural flow regimes in Poland and accompanying data on: (1) hydrological metrics (HMs), (2) environmental variables (EVs) (i.e. catchment physiographic descriptors) and (3) principal components (PCs) calculated based on HMs and EVs. All data concerns the set of 147 flow gauges with unregulated flow regime in Poland.

Publisher: Warsaw University of Life Sciences (WULS-SGGW), Poland.

Data citation: Piniewski, M. (2017) Natural Flow Regimes in Poland (NFRPL). Warsaw University of Life Sciences WULS-SGGW. Dataset. DOI: 10.4121/uuid:3cd28c08-8c30-4686-8ff9-d1ecb0a5a29c

Contacts: Mikołaj Piniewski **(**m.piniewski@levis.sggw.pl)

Data description:

The full description of the origin of this data set can be found in Piniewski (2017). A short summary of its contents is presented below.

The data set consists of an Excel spreadsheet (*NFRPL.xlsx*) storing data for 147 flow gauges located in Poland. The spreadsheet is divided into five categories: general characteristics, hydrological metrics, environmental variables, principal components and flow regime classification. General characteristics describe several main features of each gauge, such as the unique seven letter code composed of the river name code and the gauge name code, the code used by the data provider (Institute of Meteorology and Water Management – National Research Institute), mean river flow ($m^3\ s^{-1}$) and a few other features. The HMs fields include the values of all 63 hydrological metrics calculated using the River Analysis Package (Marsh *et al*., 2004). The EVs include the values of 27 environmental variables (catchment descriptors) calculated mostly in the Geographic Information Systems. These variables were used in the Piniewski (2017) paper for predicting class membership in ungauged streams using the random forest model. The PCs category includes the values of the first four PCs for the principal component analysis performed on HMs and on EVs, respectively. Finally, the flow regime classification category contains the results of the cluster analysis using the k-means and k-medoids techniques. Both methods gave almost the same classifications, with only one gauge classified differently.

All descriptions of the particular fields of each category are provided in the second sheet of the *NFRPL.xlsx* file.

There are two accompanying GIS layers in the shapefile format, *NFRPL_point.shp* and *NFRPL_polygon.shp*. One stores the shapes and attributes of the 147 flow gauges (points), whereas the second one concerns their catchments (polygons). Note that the gauges are nested so the catchments have non-zero spatial overlap. The data stored in the spreadsheet can be joined to the data stored in the shapefiles (in order to create a map of classified gauges or catchments) using the seven-character unique code field.

References

Marsh, N. 2004. *Time Series Analysis Module: River Analysis Package*. Tech. rep. Cooperative Research Centre for Catchment Hydrology, Monash University, Melbourne Australia.

Piniewski, M. 2017. Classification of natural flow regimes in Poland. *River Research and Applications*. DOI: 10.1002/rra.3153.