

## Data Management Plan

### 1) Basic information & context:

The focus of my current research is to conduct a comprehensive sensitivity analysis of key parameters involved in estimating streamflow using the Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS). This study aims to evaluate various datasets to discern the impact of different parameters on the accuracy of streamflow estimation. The project is my current Ph.D. project but it is not funded externally.

### 2) Data types, formats, standards & capture methods:

The watershed under investigation is monitored by four gauge stations. Hourly streamflow data were sourced from the U.S. Geological Survey (USGS) in 2016. These gauge stations provide crucial information on both rainfall and discharge, which can be easily downloaded in text format.

Two forms of rainfall data are considered. Firstly, previously available data was obtained from the NASA Land Data Assimilation System, where hourly gridded rainfall data were aggregated over the entire watershed to derive a single rainfall time series. Secondly, 15-minute precipitation data is extracted from the USGS website for various gauges, resulting in datasets available in text format.

In addition to precipitation and streamflow data, DEM, SSURGO soil data, and Land-use Land-cover (LULC) data are incorporated. DEM is obtained as a GeoTIFF file from the Earth Explorer of USGS. SSURGO, a comprehensive database on soil characteristics, is accessed as a geodatabase (gdb). LULC data, in raster format(.tif), is downloaded from the National Land Cover Database 2021.

### 3) Ethics & intellectual property

Given that the datasets employed consist of publicly available environmental measurements provided by government agencies, specific ethical concerns regarding privacy or confidentiality are minimal. However, adherence to the terms and conditions of data use set forth by the data providers is mandatory. Intellectual property rights are to be respected, with proper attribution and citation provided for each dataset according to the guidelines prescribed by the USGS, NASA, and other data sources.

### 4) Access, data-sharing & re-use

The data will be made available for access and sharing upon request. It will be accessible through GitHub and interested public can request us through GitHub or email.

The re-use of the data will be allowed, and appropriate licenses will be applied to ensure that the data can be used by others under clearly defined conditions.

**5) Short-term storage & data management**

Data will be stored on GitHub and OneDrive cloud throughout the research project, with regular backups. A clear naming convention and version control will be implemented for efficient organization and management.

**6) Deposit & long-term preservation**

After the project's conclusion, the datasets will be deposited in GitHub for long-term preservation. Metadata will be provided in accordance with the repository's standards to ensure discoverability and proper context.

**7) Resourcing**

The management of the data will be my responsibility, with oversight from the dissertation committee. The university's IT services will provide the necessary infrastructure for short-term storage, and the selected data repository will require no additional funding for deposit and preservation.

**8) Adherence & review**

Compliance with this Data Management Plan (DMP) will be regularly reviewed to ensure continued relevance and effectiveness. The DMP will also be reviewed in the context of any changes in data management best practices, institutional policies, or relevant legislation.