Interim Report

Topic: Understanding the sensitivity of parameters while estimating the streamflow using HEC-HMS

Rishitha Bikkumalla, Lyles School of Civil Engineering, Purdue University

Date of Completion: 04/10/2024

Introduction:

Publication synopsis

A comprehensive study on the impact of varying parameters on streamflow estimation accuracy using HEC-HMS.

Publication description

This research delves into a systematic sensitivity analysis of critical parameters in the HEC-HMS for estimating streamflow. By utilizing datasets like precipitation, streamflow, etc, the study seeks to understand how different parameters influence the accuracy of streamflow predictions, aiming to contribute significantly to water resources management and flood forecasting.

Tags

Precipitation, Streamflow, Watershed, Modeling (1952, 4316), prediction, sensitivity analysis

Source data overview:

Provenance of the source data

The research employs data from multiple sources. All the data is for the Upper Wabash Basin. The rainfall data was obtained by an ex-group student 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. Daily streamflow data is sourced from USGS for 1965-2004, covering the Upper Wabash watershed monitored by four-gauge stations. Only the interested timeframe, 1980-1982 was studied.

Additional datasets include DEM, SSURGO soil data, and Land-use Land-cover (LULC). DEM is obtained from the Earth Explorer of USGS. SSURGO, a comprehensive database on soil characteristics, is accessed from the Natural Resources Conservation Service (USDA). LULC data is downloaded from the National Land Cover Database 2021(NLCD).

Furthermore, streamflow data was predicted through HEC-HMS for various combinations of the parameters: Time of Concentration (Tc) and Storage Coefficient(S). These parameters were selected from a uniform distribution. This predicted streamflow was compared with the observed streamflow, and Nash-Sutcliffe efficiency (NSE) was calculated for each simulation.

Format of the source data

Data	Format
Precipitation	.txt
Streamflow	.txt
DEM	.tif
Soil Data	.gdb
LULC data	.tif
Predicted Streamflow	.csv

Methods:

Overview of Processing Completed as Dataset Preparation

The data preparation involved aggregating rainfall and precipitation data and checking their gaps or abnormalities. A specific rainfall event was then selected and the streamflow corresponding to it was extracted for all the nodes. Next, SSURGO and LULC data were used to find the Curve Number of the basin, using which an empirical basin model was prepared from the DEM. This model is then automated through Python by varying Tc and S to predict the streamflow.

Programs/Scripts used.

ArcGIS was employed for spatial data analysis and preparation. It helped understand the behavior of the river in the basin. HEC-HMS was used to perform continuous hydrologic simulations, model streamflow, and conduct sensitivity analysis. Python scripts facilitated data aggregation, preprocessing, quality checking, automating HEC-HMS, and graphically analyzing the data and results.

Graphical data analysis:

Overview of what graphical data analysis was conducted and a summary of what was found

Graphical analysis was conducted to visualize the source data and the impact of different parameters on streamflow estimation. Key findings include understanding the parameter sensitivity range and it’s impact on model accuracy.

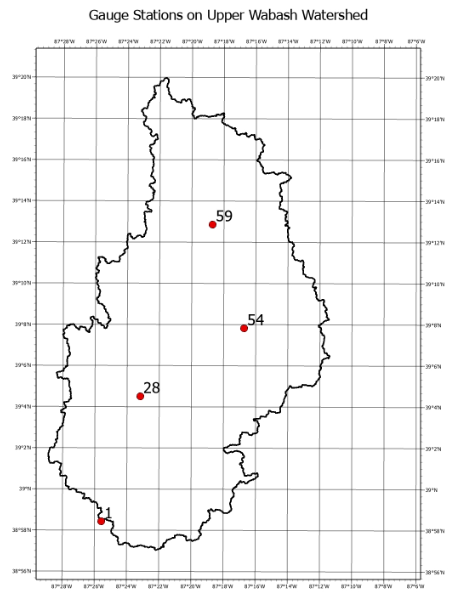

Figure 1: Upper Wabash watershed with 4 different streamflow nodes.

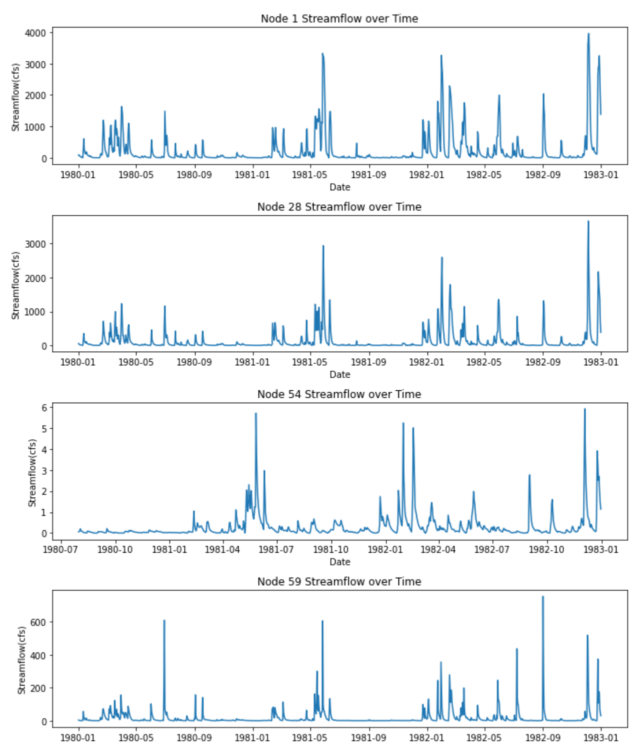

Figure 2: Streamflow of 4 gauge stations of Upper Wabash watershed for 3 years

Further, the data of the selected rainfall event was analyzed.

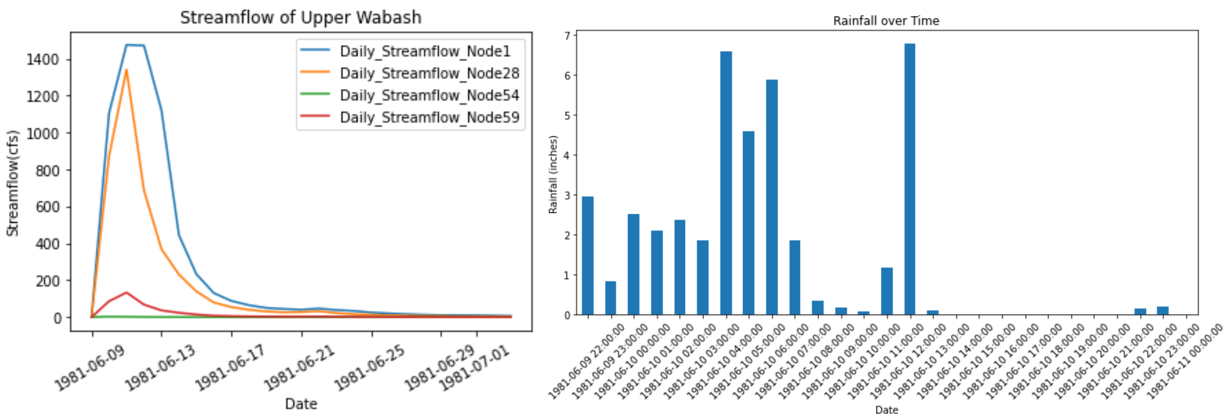

Figure 3(a): Streamflow of The Upper Wabash for all nodes for a specific rainfall event of 1981, Figure 3(b): Precipitation corresponding to the specific rainfall event.

Next, a simulation was made with 1000 samples of Tc and S from a uniform distribution. The simulations predicted streamflow for each pair of samples. The predicted streamflow is compared with the observed streamflow through NSE. The NSE variation is then analyzed.

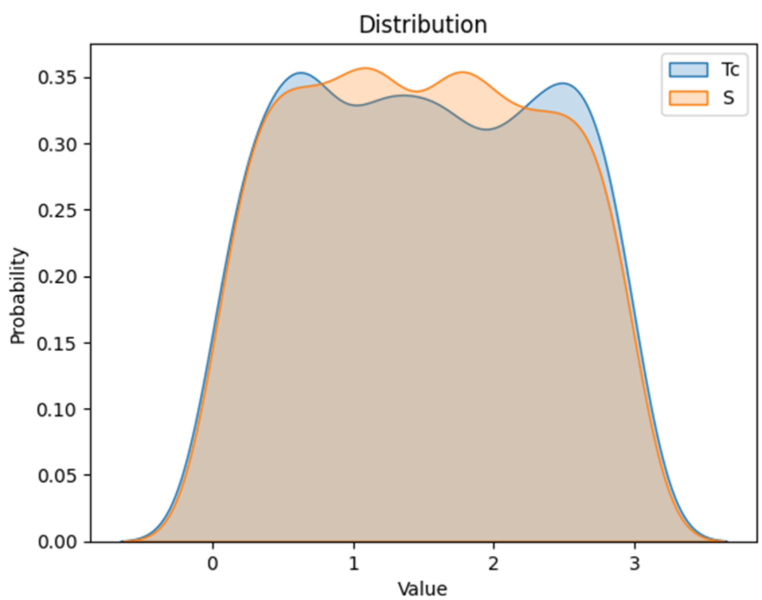

Figure 4: KDE plot of Tc and S from a uniform distribution.

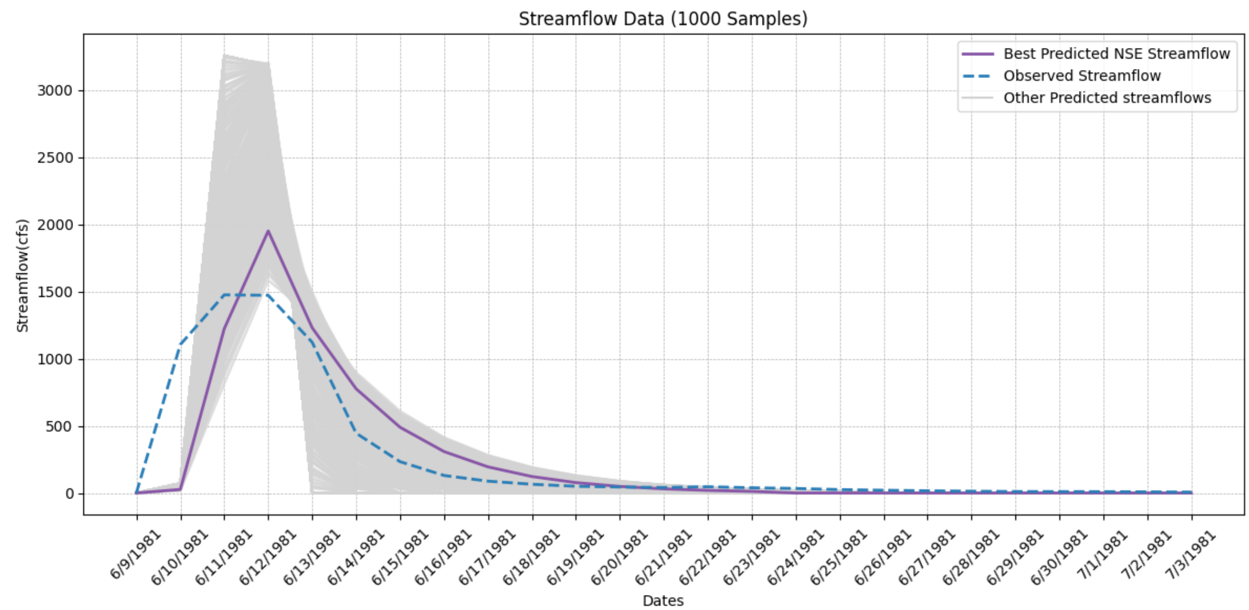

Figure 5: The predicted streamflow from the simulations. The grey region contains multiple streamflow plots. The Best predicted streamflow has the highest NSE of 0.687489024. The dashed line shows the observed streamflow values.

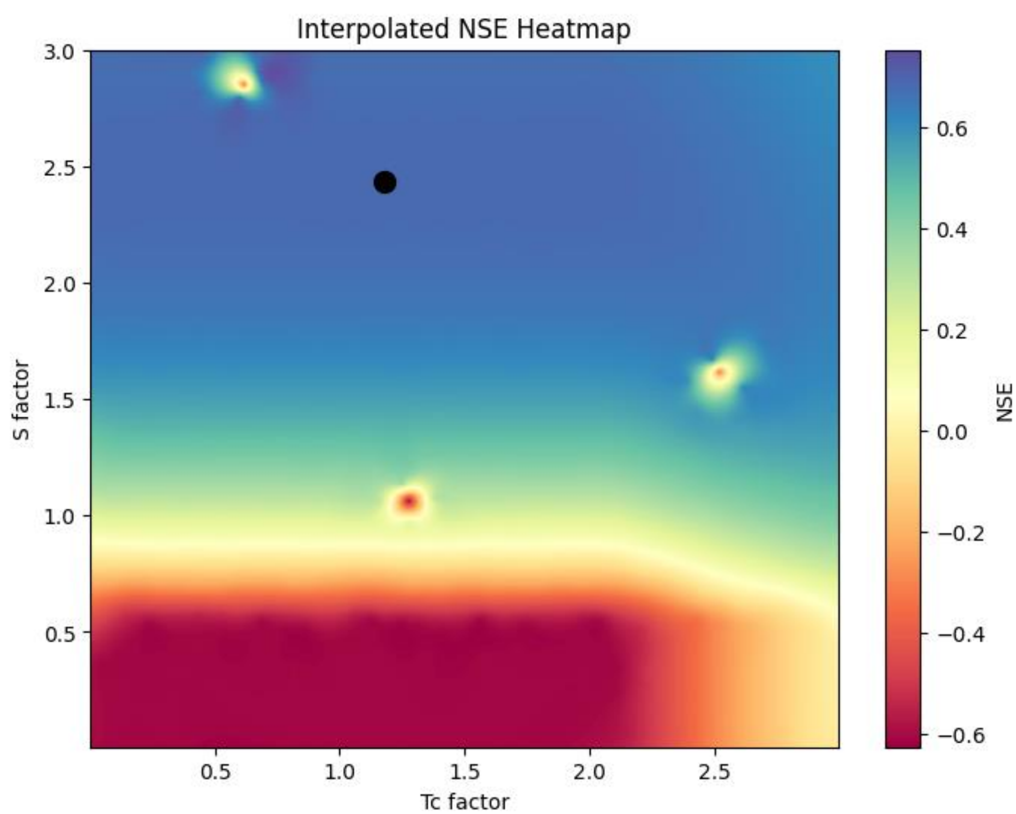

Figure 6: The kriging interpolated heatmap of S and Tc, showing the corresponding NSE. The black point represents the point of Highest NSE at Tc factor =0.57139, S factor =2.43954. At the red gradient regions (Regions where S<0.75), the NSE is negative. As the gradient slowly changes towards blue, the NSE values increase. It can be observed that there are blobs of color gradient present at 3 parts of the map. This might be due to uncertainties in the data. The top left blob has a darker blue as compared to our max NSE point. This might be because of an error in interpolation.

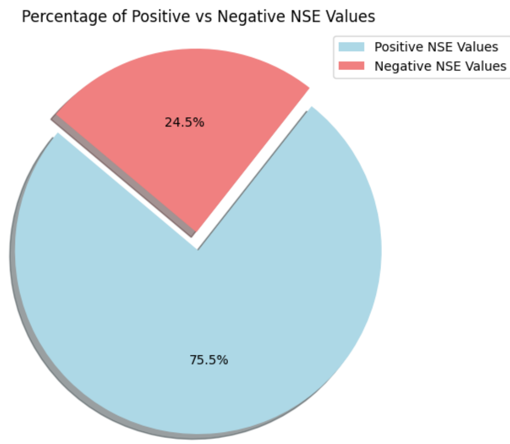

Figure 7: Among the 1000 simulations, 25% of the corresponded to producing a negative NSE whereas, 75.5% of them produced a positive NSE.

Data Quality Checking:

Overview

Data quality checking involved validating the accuracy of input datasets (rainfall, streamflow) through comparison with known benchmarks and identifying any gaps or inconsistencies. Both streamflow and rainfall datasets had zero errors in data. There were no abnormalities or gaps., hence, no corrections were made.

	Precipitation	Streamflow
No Data or Gaps	0	0
Gross Error	0	0

Table 1: Number of Corrections made in data quality checking

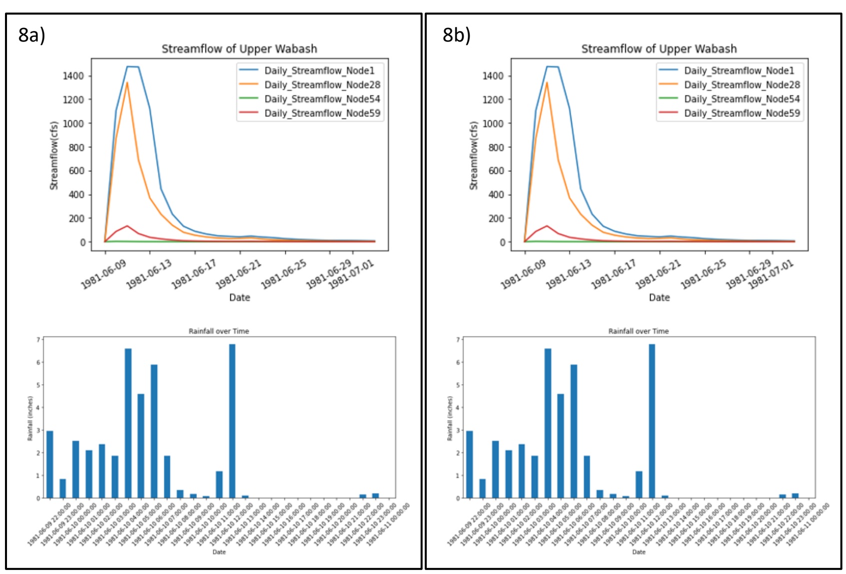

Figure 8: Streamflow and Rainfall of the event before(8a) and after correction(8b)