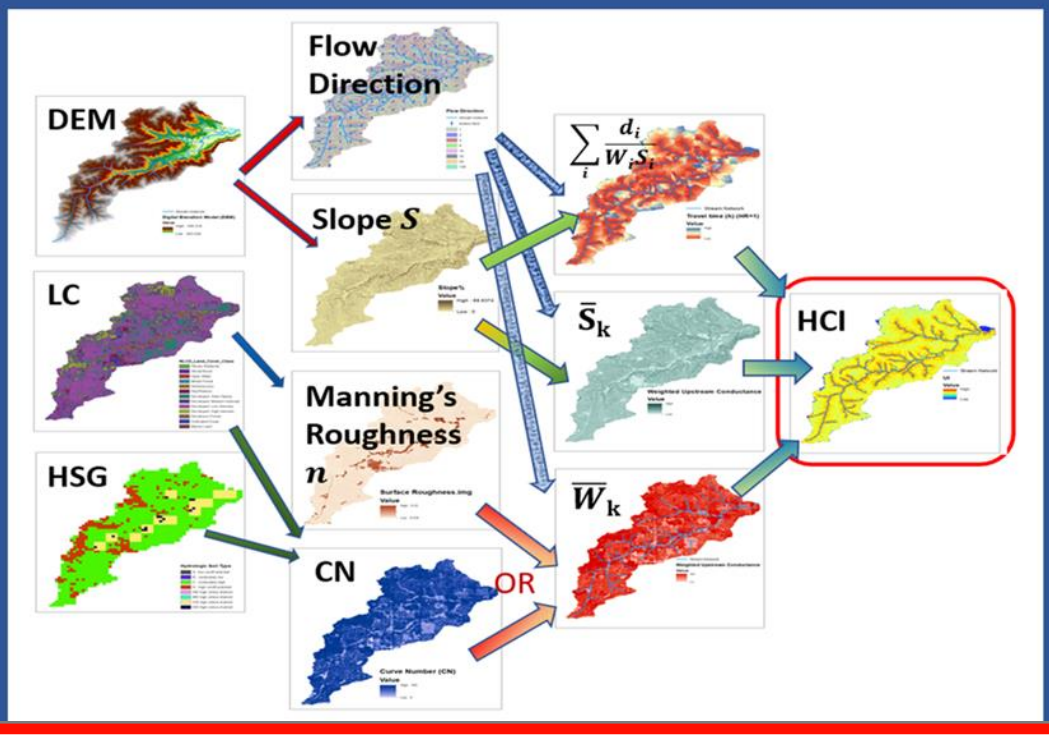
TDOT Department of Transportation

U.S. Department of Transportation
Federal Highway Administration



Updating Equations for Peak Flow Estimation in Urban Creeks and Streams of Tennessee

(Part 1)

Research Final Report from the University of Memphis | Claudio I. Meier, Francesco Dell'Aira, Nischal Kafle, and Dorian J. Burnette | January 26, 2024

Sponsored by Tennessee Department of Transportation Long Range Planning Research Office & Federal Highway Administration

THE UNIVERSITY OF MEMPHIS

DISCLAIMER

This research was funded through the State Planning and Research (SPR) Program by the Tennessee Department of Transportation and the Federal Highway Administration under ***RES # 2020-23: Updating Equations for Peak Flow Estimation in Urban Creeks and Streams of Tennessee.***

Technical Report Documentation Page

1. Report No. RES 2020-23	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle *Updating Equations for Peak Flow Estimation in Urban Creeks and Streams of Tennessee*		5. Report Date January 2024 6. Performing Organization Code
7. Author(s) Claudio I. Meier, Francesco Dell'Aira, Nischal Kafle, and Dorian J. Burnette		8. Performing Organization Report No.
9. Performing Organization Name and Address The University of Memphis 3720 Alumni Avenue Memphis, TN, 38152		10. Work Unit No. (TRAIS) 11. Contract or Grant No. RES 2020-23
12. Sponsoring Agency Name and Address Tennessee Department of Transportation 505 Deaderick Street, Suite 900 Nashville, TN 37243		13. Type of Report and Period Covered 14. Sponsoring Agency Code

15. Supplementary Notes
Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.

16. Abstract

This project addressed the need for updating the existing peak flow equations for urban basins in Tennessee. After reviewing the current state of the art, the work reported herein focuses on unraveling the complex, interacting effects that non-stationary precipitation, evolving urbanization levels, and spatial patterns in land development, rainfall, as well as antecedent conditions, all have on the hydrologic response or urbanizing basins. Potential uncertainties and biases in the estimation of extreme rainfall quantiles (IDF-DDF values), due to the low density of weather stations and the use of totalized rainfall data, and in the frequency analysis of frequent floods, due to using annual maxima instead of partial duration (peaks over threshold) series, are also investigated.

All urban basins in Tennessee have experienced growth in the amounts of developed areas in the past 20 years, and there is a significant increase in the frequency of extreme rainfall events in the region. Using rainfall data with the 15-minute resolution typically available in the U.S. introduces a negative bias that is highly variable across stations, while the low density of rain gauges increases the uncertainty in IDF-DDF values.

We derive a novel urbanization index based on hydrologic connectivity that, in contrast with the traditional approach of using percentage of impervious area (IA), is able to reflect the hydrologic effects of different spatial distributions of urbanized patches within a watershed. This index outperforms IA when used as an explanatory variable in regression equations for predicting urban peak flows.

A methodology to perform continuous hydrologic simulation with artificial neural networks is also proposed to investigate the effects of changing land cover, excluding concurrent effects of trends in precipitation.

17. Key Words	18. Distribution Statement
HYDROLOGIC CONNECTIVITY, NON-STATIONARITY, URBANIZATION INDEX, UNCERTAINTY IN EXTREME RAINFALL, FLOODS	No restriction. This document is available to the public from the sponsoring agency at the website http://www.tn.gov/.

19. Security Classification (of this report)	20. Security Classification. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	xxx	

Acknowledgement

We are grateful to TDOT personnel for their help throughout this project. Specifically, Mr. Wesley Peck, PE, supported this research continuously, helping us with any data needs, while the administrative staff at the Research Office, Mmes. Pamela Boyd-Walker, Sue Laaser, Lia Prince and Melanie Murphy, and Mr. Paweł Polazcyk, were always ready to promptly answer all our questions.

We thank the USGS personnel for always answering any enquiries about data; we are particularly grateful to David Ladd, Daniel Wagner, Peter McCarthy, and Katharine Kolb.

Dr. Antonio Cancelliere, University of Catania, Italy, helped with some of theoretical derivations related to the underestimation of frequent floods.

We are grateful to Eric Spangler and Kristian Skjervold for their help in using the High-Performance Computer (HPC) at the University of Memphis.

Executive Summary

The main objective of this research project centered on updating the equations currently used for estimating peak flows in urban watersheds of Tennessee and incorporating them into the StreamStats web application. The USGS was charged with the operational aspects of this goal, so the tasks assigned to the University of Memphis group mainly related to finding novel ways of accounting for non-stationarity in the hydrologic input (i.e., trends in extreme rainfall) and urbanization levels. These aspects involved broader research in hydrologic engineering, aimed at unraveling the concurrent effects of trends in both land development and increased extreme precipitation on urban peak flows, as well as possible sources of uncertainty or bias, when predicting rainfall Intensity (or Depth)-Duration-Frequency (IDF-DDF) values, as well as urban peak flows.

The University of Memphis group structured their research around the following aspects:

- An in-depth review of existing equations for peak flow estimation in the state of Tennessee, both for rural and urban basins, identifying issues and limitations of the approaches
- A preliminary analysis of trends in the frequency of extreme rainfall events in Tennessee and the Southeastern U.S.
- Formulating, testing, and benchmarking a novel watershed-scale, lumped urbanization index based on hydrologic connectivity, with the aim of increasing the explanatory power of regression equations for predicting peak discharges at ungauged basins
- Studying the interactions between the spatial distribution of land-cover patches and that of rainfall, implicitly accounting for antecedent moisture conditions, via continuous simulation with data-driven, neural-network models
- Using high-resolution precipitation data from Germany, to quantify the underestimation and uncertainty introduced by the use of NWS-COOP rainfall data totalized every 15-minutes, when deriving the currently used IDF-DDF values
- A preliminary analysis of the potential biases and variability in estimations of extreme rainfall due to the low density of the weather station network in Tennessee
- Analyzing how using annual maxima for flood frequency analyses underestimates frequent floods, as compared to choosing partial-duration (peaks-over-threshold) approaches

Key Findings

- Using rainfall data from 473 weather stations in the Southeastern U.S., of which 44 are in Tennessee, we performed a preliminary trend analysis of extreme rainfall frequency. Overall, we find that many more stations report significant increases in extreme storm frequency than decreases, and there is a spatial pattern to these trends: Most gauges with decreased frequency are located in Louisiana or Arkansas, while those reporting the most significant increases tend to be located throughout the other Southeastern states, but with increasing density moving inland. In Tennessee, we detect increasing or significantly

increasing trends in 23 to 45% of the stations, depending on duration, while we find a decreasing trend only for one single combination of station and duration. We conclude that these increasing trends in extreme rainfall frequency should be further studied, also incorporating rainfall magnitude, to improve our estimation of peak flows, but the lack of adequate precipitation data will require innovative approaches.

- Based on a conceptualization of the main factors that influence the hydrologic (flooding) response of urbanizing watersheds, we propose a new urbanization index ***UI*** at the basin scale, based on considerations of hydrologic connectivity. ***UI*** can be used as an explanatory variable when deriving regression equations for urban peak flows. Using USGS study cases, we demonstrate that ***UI*** outperforms the percentage of impervious area ***IA*** as a variable for use in peak flow equations. The index is obtained in a GIS environment, requiring the raster maps for elevation, Curve Number, and land-use cover. We are currently testing this novel concept on a larger number of case-study basins, before publishing it. The USGS StreamStats team have manifested their interest in implementing our index as an explanatory variable into regression equations for estimating urban peak flows and other flood-related hydrologic variables, such as time of concentration. Using ***UI*** would reduce the uncertainty in the estimated floods and would allow for a continuous adjustment of the predictive equations as urbanization levels increase in a basin, directly tackling a main source of non-stationarity. For future research and large-scale implementation, we have identified a range of ideas that could improve the predictive power of the urbanization index and reduce the computational effort necessary to map connectivity indices across vast territories.
- Urbanization rates over the last 20 years were computed for all watersheds in major metropolitan areas in Tennessee; all basins have experienced urban expansion, but there is a large range of variability.
- For four watersheds in Northern Georgia, with flow data from USGS gauging stations and gridded, 1-hour, gauge-corrected radar products from the NWS, we used a data-driven, neural-network model to demonstrate that it is possible to study the effects of changes in land-cover, net of any precipitation trends. This is relevant for TDOT, as it allows dealing with the flow regime attribution issue in future regional models for peak-flow equations.
- Using 13 years of rainfall data with 1-minute resolution at 862 German weather stations, we document the large variability in the negative biases introduced in IDF-DDF estimations, due to the use of totalized data such as those from the 15-minute-resolution NWS-COOP network. These findings raise difficult but serious questions regarding the uncertainties and biases in our current estimates of IDF-DDF values in the U.S. in general, and Tennessee in particular, for the short durations of interest in urban hydrology. Furthermore, emulating as much as possible the methods in Atlas 14, we attempted a preliminary effort to understand potential effects due to the low density of rain gauges, by using the same German data. We find that low weather-station densities, similar to those in Tennessee, result in more cases of underestimation than positive bias, and noticeably increase the variability in IDF-DDF values. These findings are highly relevant for TDOT and other agencies that require accurate estimates of extreme rainfall.
- Based on a rigorous analysis of more than four hundred basins across the U.S., with minimum human impact, we systematically describe the underestimation of frequent floods

when performing flood frequency analyses with annual maxima instead of peaks-over-threshold (partial duration) approaches, which are conceptually more appropriate. We observe that the underestimation for return periods equal to or less than 5 years can be severe (reaching 50% or more) at some locations, depending mostly on climate. We provide a theoretical framework for analyzing what drives this issue, as well as correcting biased estimates obtained with annual maxima This methodology can benefit a variety of scientific and applied disciplines, such as fluvial geomorphology and river restoration, that need accurate predictions of frequent floods. As the use of annual maxima is the most established technique for flood frequency analysis, given its advantages compared to peaks-over-threshold analyses, we suggest that our methodology should be incorporated in the national guidelines for flood frequency analysis.

Key Recommendations

- Further pursue the development of the urbanization index ***UI***, also including the effects of drainage infrastructure on hydrologic connectivity. The current ***UI*** proposed in this report only considers surface hydrologic connectivity, driven by relief and the nature of surface patches (their slope, roughness, infiltrability, etc.), but most urbanized and urbanizing areas contain stormwater sewerage systems, with their own, separate effects on hydrologic connectivity. Including information on both types of connectivity should enhance the explanatory power of ***UI*** in peak flow equations for urbanizing areas, further reducing estimation uncertainty.
- Simultaneously increase the number of gauged basins as well as the density of rain gauges across urban areas of Tennessee. The design for such enhanced hydrometeorological monitoring networks (sometimes known as "mesonets") should consider a more distributed coverage across the different urban areas of the state, as well as a wider range of watershed characteristics. Rain gauging equipment for the monitored watersheds should be state-of-the-art, with 1-minute resolution, and a spatial density that ensures coverage of extreme, small-scale convective events (i.e., thunderstorms). This would benefit TDOT by allowing for future development of locally tailored peak flow equations, better estimation of IDF-DDF values, and improvements in methods for engineering design, that would increase urban resilience to extreme hydrometeorological events.
- Perform further, more detailed studies of: (i) how totalized, 15-min data underestimate actual maxima in the U.S., (ii) what uncertainty or bias is introduced due to the current, low density of weather stations, (iii) trends in both frequency and magnitude of extreme precipitation, and (iv) how the development of novel, event-based rainfall analysis techniques can help extract more information from the sparse, short precipitation records available in Tennessee. Collectively, results from these studies would allow for updating IDF-DDF values in Tennessee, so that the present, higher values of rainfall are reflected in engineering designs. In this respect, it should be noted that there is increasing evidence, including results from this research, that the currently accepted extreme rainfall values for design, from NOAA's Atlas 14, are underestimated.
- Convey to practitioners that commonly performed flood frequency analyses based on annual maxima can severely underestimate frequent floods (say, those with return period ≤

5 years), and thus should not be used in engineering applications requiring such estimates, like the determination of bankfull floods or river restoration projects. This would ensure that designs are not based on biased estimations, which increases their risk of failure.

- Pursue the study of machine learning models for explaining the hydrologic effects of urbanization, disentangling trends in land cover from those in precipitation. We think that this is the next frontier for meaningful improvement in hydrologic estimates needed for design.

Table of Contents

List of Tables

List of Figures

Glossary of Key Terms and Acronyms

AM: Annual Maxima

ANN: Artificial Neural Network

ASM: Antecedent Soil Moisture

ASOS: Automated Surface Observing System

CDC: Climate Data Centre

CDF: Cumulative Distribution Function

COV: Coefficient of Variation

CI: Confidence Interval

CN: Curve Number

COOP: Cooperative Observer Program

CONUS: Conterminous United States

DDF: Depth Duration Frequency

DEM: Digital Elevation Model

DCIA: Directly Connected Impervious Area

DOD: Department of Defense

DWD: German Weather Service

EIA: Effective Impervious Area

FAA: Federal Aviation Administration

FA: Frequency Analysis

FFA: Flood Frequency Analysis

GIS: Geographic Information System

GPA: Generalized Pareto Distribution

HC: Hydrologic Connectivity

HCI: Hydrologic Connectivity Index

HCIA: Hydrologically Connected Impervious Area

HHR: Hydrologically Homogenous Region

HPD: Hourly Precipitation Data

HSG: Hydrologic Soil Group

IA: (proportion or percentage of) Impervious Area

IC: Index of Connectivity

IDF: Intensity Duration Frequency

IEM: Iowa Environmental Mesonet

i.i.d.: identically and independently distributed (random variable)

IPCC: Intergovernmental Panel on Climate Change

LC: Land Cover

LSTM: Long-Short Term Memory (neural network)

M-K: Mann-Kendall (statistical test)

MLP: Multi-Layer Perceptron (neural network)

NCEI: National Centers for Environmental Information

NWS: National Weather Service

OLS: Ordinary Least Squares (regression)

PDS: Partial Duration Series

PILF: Potentially Influential Low Flow

POT: Peaks-Over-Threshold

RFA: Regional Frequency Analysis

RFP: Request for Proposals

ROI: Region of Influence

RNN: Recurrent Neural Network

RUSLE: Revised Universal Soil Loss Equation

SAF: (rainfall) Sampling Adjustment Factor

SN: Stream Network

TDOT: Tennessee Department of Transportation

TN: Tennessee

UI: Urbanization Index

U.S.: The United States

USGS: The United States Geological Survey

UofM: The University of Memphis

WBD: Watershed Boundary Dataset

Chapter 1 Introduction

To design infrastructure such as bridges, culverts, and levees, for management and zoning purposes, for emergency response, to understand channel instability and other environmental impacts, and to design river restoration projects, civil engineers need to predict the magnitude of floods for a range of return periods at ungauged locations along ditches, creeks, and streams. Typically, the methods used for estimating flood frequency are regional, transposing information from gauged to ungauged watersheds under the assumptions of stationarity and hydrologic homogeneity (i.e., the hydrologic response does not change in time, and it is similar across basins, respectively). The issue in cities, where the risk to lives and property is higher, is that the hydrologic response of urban watersheds can change in time due to the simultaneous effects of urbanizing trends, changes in extreme precipitation, or the development of stormwater control measures.

*The issue in cities, where the risk to lives and property is higher, **is that the hydrologic response of urban watersheds can change in time** due to the simultaneous effects of urbanizing trends, changes in extreme precipitation, or the development of stormwater control measures.*

This work focuses on updating the equations used for predicting peak flows in urban basins of Tennessee, accounting for non-stationarity in the hydrologic input (i.e., trends in extreme rainfall) and urbanization levels. It also involves broader investigations in hydrologic engineering, aimed at unraveling the concurrent effects on urban peak flows due to trends in both land development and extreme precipitation, as well as possible sources of bias or uncertainty when predicting Intensity (or Depth)-Duration-Frequency values and urban peak flows.

In the remainder of this introductory chapter, we present the global and local contexts for this research and then introduce its general objectives, before describing the ten different tasks that must be completed to achieve them. The subsequent chapters include the literature review, the methodology, the results, and the conclusions from the work performed.

1.1 General and local context

Traditionally, engineers have used regional equations to estimate peak flows at ungauged locations. These are derived by computing flood quantiles from flood frequency analyses at gauged sites, and then regressing them on a range of lumped basin descriptors, such as drainage area or percentage of impervious area, among others. In theory, the gauged basins used to derive the equations as well as the ungauged basins in which they are then applied should be "hydrologically similar." Because this is not easy to assess, regional equations are typically created for administrative regions, instead of hydrologically homogeneous regions. This is the common case when deriving peak flow equations for urban areas, as it is assumed that the higher levels of urbanization result in watersheds that are more similar hydrologically, independently of the actual physiographic (geology, soils, relief, etc.) and climatic contexts.

In Tennessee, the overall problem of urban flood estimation involves varied aspects:

a. Previous work on the topic is aged (> 40 yrs.), was derived from data for basins that were still undergoing urbanization and did not include watersheds with intense development.
b. Current equations were derived for a small number of gauges with very short flow records (< 8 yrs.), so they are based on synthetic (model-generated) floods, instead of actual data.
c. Existing equations do not consider the possibility of non-stationarity in either land use or rainfall extremes, nor do they include any variables reflecting these urban-flood drivers.
d. As a result, the equations presently in use may not represent the current flood regime adequately; moreover, their predictive uncertainty is probably large.
e. Few cities in or around Tennessee have gauged urban watersheds, and those that do (e.g., Nashville and Huntsville, AL) are located within only three of the five main physiographic zones in the state. Additionally, all gauged urban basins are highly clustered in space, so that their spatial distribution does not fully represent the expected range of variability.
f. The effects of urbanization trends must be disentangled from any potential future changes in extreme precipitation, so that predictive equations can explicitly incorporate both these drivers as separate variables.

In 2017, responding to this conundrum, TDOT issued Research Needs Statement Number 25: "Peak Flow Estimation in Urban Drainage Areas," with the goal of updating equations for peak flow equations in Tennessee. This request for proposals stated that "TDOT uses equations developed by the United States Geological Survey (USGS) to estimate peak flows through all structures with a drainage area greater than approximately 500 acres (0.78 sq.mi.). Equations for urban areas have not been updated since 1984 and do not take advantage of updates in estimation techniques and an additional 30 years of observed flow data to increase accuracy. Tennessee is currently experiencing exponential urban growth and updating the equations will provide more accurate peak flows resulting in more accurate drainage structure design." In turn, the stated goal for the research was to "update the peak flow estimation equations for urban drainage basins in Tennessee, implement them in the StreamStats computer program currently used and enhance StreamStats to provide parameters used in urban equations."

This research offers innovative approaches to better understand the hydrologic response of urbanizing watersheds, supporting improved predictive equations for urban peak flows in a context of limited information, both in time (short records) and space (few, clustered stations).

It should be noted that the proposal that was selected originally considered a single award to the UofM with a sub-award to the USGS. Because of this, tasks, deliverables, and milestones were all written to reflect a joint UofM-USGS project, so that some of them were collaborative, while others needed to be performed by UofM for the USGS. In the end though, the work was clearly split between the USGS and UofM, with no direct feedback or overlap, and two separate grants were awarded. Because of this, the main goals and scope of work for the two research groups are different, while some of the original subtasks and deliverables became moot.

UofM was tasked with more general research in hydrologic engineering, aimed at better predicting rainfall-runoff patterns in urbanizing watersheds, including non-stationary effects and uncertainties or biases in estimation.

Thus, UofM's proposal initially included aspects such as (i) developing new indices of urbanization that consider hydrologic connectivity, so as to improve the explanatory power of regression-based peak flow equations, (ii) reflecting the effects of urbanization trends on hydrologic response, by considering event-based analyses, (iii) analyzing whether there are recent trends in extreme precipitation in Tennessee and surrounding states, and (iv) getting a preliminary notion of possible sources of uncertainty and potential biases in the estimation of both extreme rainfall and urban floods in Tennessee.

In turn, the USGS was tasked with deriving updated equations for peak flows and incorporating them into the StreamStats web application, which involves compiling streamflow data for urban areas, determining flood quantiles at these gauged locations, computing a range of potential basin descriptors to be used as explanatory variables, and finally performing the multiple regressions to obtain the updated equations, before implementing them in StreamStats.

1.2 Objective of the research

This project's specific objective is to develop improved sets of equations for estimating peak flows in urban areas of Tennessee and neighboring states, and then implement them into the StreamStats web application, ensuring that the program can automatically compute the different parameters and variables included in the equations.

As mentioned above, UofM's tasks consider broader objectives related to improving our understanding of the hydrologic response in urbanizing basins, dealing with non-stationary behavior (trends) in the rainfall and the system's urbanization levels, and discerning potential sources of uncertainty or biases in the estimation of extreme rainfall and floods in Tennessee.

1.3 Tasks to achieve the objective

The proposal considered the ten following tasks, split between the USGS and UofM as indicated in each case. Please note that Task 6 was originally a joint USGS-UofM task, but ended up being performed solely by the USGS, using traditional basin descriptors, after the project was split into separate awards. Nonetheless, we are actively conversing with the StreamStats team at the USGS, due to their interest in including the new urbanization index **UI** into StreamStats.

Task 1 - USGS: Compile and analyze available flow and stage data for urban areas in Tennessee and surrounding urban areas of neighboring states in order to understand the spatial coverage and record lengths of streamflow data, assess the ranges of variability for the different watershed characteristics (area, slope, etc.), and determine flood magnitudes for annual exceedance probabilities (AEPs) at stream gauge locations using methods described in the Subcommittee on Hydrology of the Advisory Committee on Water Information's Guidelines for Determining Flood Flow Frequency Bulletin 17C (England et al., 2019).

Task 2 - UofM: Review existing peak flow equations for both rural-unregulated and urban watersheds in Tennessee and immediately surrounding areas of neighboring states. Document the variability in physiographic factors (geology, soils) across cities in the state. This will help in determining candidate watershed (and climate) characteristics to potentially include in the analyses to develop the prediction equations, and in defining homogeneous "regions" (groups of cities, in this case) for equation development and range of applicability.

Task 3 - UofM: Obtain all continuously recorded point rainfall data at and near urban areas in Tennessee and analyze the records to detect jumps or trends in extreme precipitation for the range of durations relevant to the watershed sizes considered in this proposal (1 to 100 mi^2). These precipitation records, in conjunction with rainfall radar data, will also help in determining water inputs for the analysis of specific rainfall-runoff events.

Task 4 - UofM: Utilize existing land cover datasets (e.g., National Land Cover Database; Homer et al., 2015) to assess urban and urbanizing watersheds, in order to reflect the non-stationary nature of urban environments and the impact that this has on total impervious area, hydrologic connectivity, and resulting runoff rates. Determine applicable ways of better incorporating indices of urbanization as variables in the predictive equations (e.g., total vs. effective vs. directly connected impervious areas).

Task 5 - UofM: Perform event-based rainfall-runoff analyses for all flow events above a certain threshold (locations with streamflow data), or all rainfall events above a certain threshold (locations with only stage data), in order to understand the effects of antecedent conditions and time-varying urbanization levels on hydrologic response (i.e., on lag times and/or peak flows).

Task 6 - USGS: Based on results from all previous tasks, determine homogeneous groups of urban areas, and select the basin characteristics and rainfall event variables to be included in the peak flow estimation equations.

Task 7 - USGS: Calculate potential basin characteristics (explanatory variables) using GIS techniques for improved accuracy and consistency over those used in previous urban peak-flow studies in Tennessee and to ensure that the basin characteristics can be used in StreamStats.

Task 8 - USGS: Using statistically significant basin characteristics and flood-frequency estimates at gauged locations, regression equations will be developed using generalized or weighted least squares methods to estimate flood magnitudes for selected annual exceedance probabilities.

Task 9 - USGS and UofM: Deliver final report summarizing the findings of the study.

Task 10 - USGS: Develop an urban basin flood-frequency decision support tool for TDOT staff, comprising a final report and incorporation of the updated regression equations into the StreamStats web application.

Chapter 2 Literature Review

Besides a general outlook to urbanization effects on hydrologic response, this chapter provides background about previous research on the UofM-assigned Tasks (2 to 5). We discuss existing peak flow equations for both rural and urban streams in Tennessee, and their limitations. We then review the state of the knowledge about rainfall trends in TN and the U.S. Southeast in general, for sub-daily durations that are of interest when designing stormwater or hydraulic infrastructure for urban basins under 100 sq.mi. Next, we cover the previous attempts at developing hydrologic connectivity indices, to better describe the rainfall-runoff response in urbanizing watersheds, before discussing potential approaches to represent the highly complex, interacting effects of land-cover changes and rainfall-runoff response. Finally, we briefly discuss a range of possible sources that introduce uncertainty or bias in our estimates of extreme rainfall as well as floods.

2.1 Hydrologic response of urbanizing watersheds

It has been recognized since at least the 1960s that land development has strong effects on the hydrologic response of urban basins (Leopold, 1968; Hollis, 1975; Konrad 2003). Typically, runoff volume and peak flow rates increase, lag times are reduced, discharge decreases during low-flow periods (baseflows), and flood hydrographs become peakier, as compared to pre-development conditions (Konrad 2003; Sillanpää & Koivasulo, 2015). Recent studies on urbanization impacts to flood hydrology include Shuster et al. (2005), Schueler et al. (2009), and Jacobson (2011).

Any Engineering Hydrology textbook (e.g., Viessman & Lewis, 1995) explains that the variation in rainfall floods in rural, unregulated streams within a physiographically homogeneous area is explained primarily by basin area and rainfall, modulated by antecedent conditions. Secondary variables include basin shape (indexed by various shape factors, or else through a combination of basin length with area) and either basin or channel slope. These are, for example, the typical watershed variables included in regression equations for synthetic unit hydrographs, the most common engineering tool for flood prediction in ungauged basins (see, e.g., Chapters 9 & 10 in Gupta, 2017). In contrast, other variables can become more relevant in urban and urbanizing watersheds, such as impervious area (amount and spatial distribution), its location with respect to the basin's outlet, the degree of connectivity between impervious areas and streams and between impervious and pervious patches, and the presence of drainage infrastructure.

Impervious area (IA) is typically included as an explanatory variable in predictive equations, as percentage of the total basin area. But some authors (Roy and Shuster, 2009; Jacobson, 2011) suggest that effective impervious area (EIA), that portion of IA that is hydraulically connected to the stream network by the drainage system, better explains urbanization effects on peak flows. Following Ebrahimian et al. (2016), this occurs when surface runoff generated over impervious areas flows over paved surfaces, through pipes, or in any other type of structure that does not reduce runoff volume. Ebrahimian et al. (2016) emphasizes the difference between EIA and a similar concept: directly connected impervious area (DCIA); while both are attempts at indexing that part of impervious area that contributes runoff to the stream network through stormwater infrastructure, DCIA is obtained from the characteristics of the basin surface and its drainage system (Seo et al. 2013), while EIA is derived from retrospective analyses of rainfall-runoff events.

Assuming that the basin's portions that actively contribute to its response grow in extent when considering increasing rainfall events, DCIA should be the area that contributes first, for small events, followed by the indirectly connected IA and then the pervious patches, which add their contributions for increasingly intense rainfall events, respectively (Boyd et al. 1993). Determining EIA is difficult, because it depends on rainfall intensity, which changes in space and time. On the other hand, DCIA can be defined if high-resolution spatial information is available, or else through field visits. Epps and Hathaway (2018) quantify EIA for three urban watersheds in Knoxville, TN, with methods that are not applicable at the scale of this project.

Sytsma et al. (2020) highlight the importance of the drainage network and the effects of antecedent soil moisture (ASM) conditions of pervious patches in affecting the hydraulic connectivity of impervious areas. This implies that the ways in which impervious areas exchange water with pervious patches and with each other should be regarded as a dynamic property, depending on past rainfall amounts. Sytsma et al. (2020) quantify such effects by defining the so-called hydrologically connected impervious areas (HCIA). However, this methodology is also impracticable in the framework of this project, due to issues in accessing information on the stormwater network for large portions of territory.

Regardless of whether imperviousness is better represented by IA, EIA, DCIA, or HCIA, multiple authors show that its spatial distribution, interacting with the spatio-temporal variability of rainfall, affect the flood response of urban basins (Mejía and Moglen, 2010a and 2010b; Yang et al., 2011; Yao et al., 2016; Miller and Brewer, 2018). It should be noted though that Bell et al. (2016) found IA to be the best predictor of peak flow, stating that even though "there is a strong theoretical basis for EIA as a predictor of urban hydrologic response, and numerous modeling [...] studies have demonstrated the importance of including connectivity [...] no empirical studies known to the authors have actually demonstrated the predictive power of EIA."

The flow regime attribution problem

Urbanizing watersheds are non-stationary by definition; we also know that extreme rainfall events are increasing in the U.S. over a range of durations (e.g., Douglas and Fairbanks, 2011, for the Northeast; Pryor et al., 2009, for the Central Plains and northwestern Midwest), which also introduces non-stationarity. For these reasons, when attempting to predict peak flows for urbanizing basins under changing climatic conditions, it would be helpful to first disentangle the effects of these different factors, what Jovanovic et al. (2018) term the "flow regime attribution problem;" in our case, we are specifically interested in "flood regime attribution."

As done by Miller et al. (2014) and Sillanpää and Koivasulo (2015), this would require documenting the historic changes in land use across a study watershed and attempting to correlate them to hydrologic change. In a similar context to this project – estimating peak flow quantiles in heavily urbanized Northeastern Illinois basins, Over et al. (2017) proposed two different methods to explicitly incorporate changing urbanization rates. Application of these methods to Tennessee may be difficult due to a relative lack of streamflow data, though; Over et al. (2017) had access to much longer streamflow records with a more uniform geographic coverage and a larger number of gauging station (117 stations).

A third potential source of trends in hydrologic response is the introduction of stormwater control measures. Even though Bell et al. (2016) found no influence on peak flows, Smith et al. (2013) concluded that differences in hydrologic response among urban basins are indeed

linked to drainage infrastructure, stating that "detention basins in these watersheds appear to operate as intended by stormwater legislation to lower peak discharges but not runoff volumes." Smith et al. (2013) work was performed in Baltimore though, where there is a much higher penetration of stormwater control measures than in Tennessee. For this reason, we do not further consider this possible source of non-stationarity in what follows.

Dealing with the lack of long-duration records

Predictive equations for peak flows are derived by applying regression techniques to long flow records at multiple gauges. This purely statistical way of relating floods to different explanatory variables works well when there is a large sample of basins and when the different watershed descriptors included in the equations (such as area, slope, ***IA***, etc.) adequately cover the range of variability, with little collinearity. We posit that in the case of Tennessee, there is not a sufficiently large number of urban gauging stations with sufficiently long records to sustain this approach beyond a preliminary level, so that alternatives are needed.

One such alternatives is using partial duration (also known as peaks-over-threshold or POT) methods for performing flood frequency analyses. This has not been commonly used in the context of peak flow estimation (but consider Soong et al., 2004), even though it is a good way of increasing sample size without incorporating potentially influential low floods (PILFs), as frequently happens in annual maxima analyses (Cohn et al. 2013). Keeping PILFs would introduce heterogeneity in the flood population, violating a basic assumption of frequency analysis. On the other hand, partial-duration analyses only consider the larger events from the beginning, so that if the threshold is chosen correctly, the need for low-outlier analyses, such as Grubbs-Beck type tests (Grubbs and Beck, 1972; England et al., 2019) is precluded.

Another way of increasing sample size when dealing with short records is by performing event-based analyses, seeking to relate different characteristics of the response of individual floods (peak flow, lag time, duration above a certain discharge, etc.) to rainfall characteristics as well as antecedent conditions. This approach was used by Boning (1976) in Shelby County, Tennessee, albeit with very limited data (2 years in the best cases), as well as by Kermadi et al. (2012), Braud et al. (2013), and Sillanpää and Koivasulo (2015). Whereas 10 annual peak flows are insufficient to estimate flood quantiles adequately, the 10 yr-long record could contain on the order of fifty large storm-flood events. Analyzed in conjunction with antecedent conditions and storm characteristics, such a larger sample size can yield valuable information about the response of urbanizing watersheds that would not be contained in only 10 annual maxima.

2.1 Existing peak flow equations for Tennessee

As mandated in Task 2, we review the peak flow equations currently used in Tennessee, discussing their limitations, as well as flow-data issues related to the short record durations, the scarcity and clustering of urban gauging stations, and the lack of consistent data. We first review equations for rural streams and rivers, before focusing on the urban case. Because this research focuses on estimating discharges, we do not include past literature on the estimation of flood depths in Tennessee (e.g., Gamble and Lewis 1977; Gamble 1983), beyond noting that they derived their equations exclusively from depth data measured at gauging stations.

This probably gives a biased view of the possible flood depths at cross-sections far from measurement transects, as the locations for gauging are not randomly selected along streams, but typically focus on bridge and narrow crossings, for operational reasons.

Peak flow equations for rural streams

Over the last 65 years, seven studies have been published on the magnitude of floods in rural, unregulated streams of Tennessee and nearby areas, by Jenkins (1960), Patterson (1964; this is a graphical method, so it is not mentioned further), Speer & Gamble (1964), Randolph & Gamble (1976), Weaver & Gamble (1993), Law & Tasker (2003), and Yan (2012). Given that the present work focuses on urban and urbanizing watersheds, it is out of scope to compare these approaches in detail; moreover, all this literature is already summarized meticulously in Yan (2012). Summing up these studies for rural watersheds in Tennessee, we can state that:

a. They are all based on hundreds of gauging stations, from a low of 216 stations in Speer and Gamble (1964), who focused on the Cumberland and Tennessee basins, to a high of 453 gauges in Law & Tasker (2003), who considered all of TN and its adjacent states.
b. Only rural or "lightly developed" basins were used, even though the latter is defined in different ways, all the way from "no more than 10% urban land use" (Yan, 2012) to "up to about 30 percent total impervious area" (Law and Tasker, 2003). It should be clear that a basin with 30% of ***IA*** cannot be classified as rural or "lightly developed."
c. Most consider gauges with 10 years of data or more, a very small sample size for estimating any flood quantile beyond a return period of 10 years, so that uncertainties must be large.
d. They split their coverage area in a number of hydrologically homogeneous regions (HHRs), ranging from a maximum of 27 zones (Patterson, 1964) to a minimum of four, which is the case for all recent (since 1976) reports. These four regions roughly correspond, with some slight differences, to: HHR4 - western Tennessee (Mississippi Alluvial Plain and East Gulf Coastal Plain physiographic provinces), HHR2 - the Highland Rim of central Tennessee, HHR3 - the Nashville Basin physiographic province, and HHR-1 eastern Tennessee (Cumberland Plateau, Cumberland Mountain, Southern, and Tennessee physiographic provinces).
e. The range in basin areas is huge, covering from about 100 acres to many thousands of square miles, clearly crossing the thresholds for the dependence of extreme floods on different precipitation mechanisms, from single-cell convective events to organized convection to synoptic rainfall. Moreover, results for watersheds larger than about 100 sq.mi. are probably irrelevant for the research at hand, which focuses on urban basins.
f. Most of the predictive equations depend exclusively on a single variable, watershed drainage area, except for the following two exceptions:
    i. For one of their HHRs, Law and Tasker (2003) propose equations where peak flows are a function of drainage area, main-channel slope (albeit to a very low exponent ranging from 0.112 to 0.117, so that results are highly insensitive to it), and a climate factor (with an extremely large, highly variable exponent ranging between 2.904 and 4.581, which makes one question how sensitive results are to variability in this factor). In the other three HHRs, their equations are functions of drainage area and main-channel slope, the latter with an exponent which is again very low, ranging from 0.054 to 0.160.
    ii. Yan (2012) derived equations based on a combination of the following variables:

drainage area, mean slope, shape factor, mean basin elevation, and mean annual rain. For some HHRs and return periods, his equations incorporate all five independent variables together. Although he checked for correlations between the explanatory variables, the statistical criterion was probably too liberal; this might explain the incorporation of both mean basin elevation and mean annual precipitation – which should be correlated - in some of his equations. Beyond this, it is not clear why one would use mean annual precipitation to explain extreme floods, given the availability of IDF-DDF values over the area of interest. It should be noted that some of Yan's (2012) equations for HHR 3 have a higher exponent for the shape factor than for the drainage area, which is atypical and clearly unexpected, raising questions about his procedures.

In all these equations, basin drainage area is (almost) always the variable with the highest explanatory power, so it is important to analyze the variability in its scaling exponent, which we will call α. Jenkins (1960) proposed a value of 0.77 for α, in all regions, while Speer and Gamble (1964) consider an exponent of 0.793. Randolph & Gamble (1976) suggested that α ranges from 0.709 to 0.752 in HHRs 1 to 3, while it is lower, between 0.515 and 0.575 for their HHR 4, which corresponds to western Tennessee (roughly, west of the Tennessee River). Rather similarly, Weaver & Gamble (1993) advance values between 0.666 and 0.753 for HHRs 1 to 3, and between 0.523 and 0.612 for HHR 4, while Yan (2012) finds that α ranges from 0.64 to 0.77, and from 0.52 to 0.55, for the same respective HHRs. The only deviation from this rather consistent behavior for α is found in Law and Tasker (2003), who propose higher values, in the range 0.742 to 0.815, for HHRs 1 and 2, as well as for smaller (< 30 sq.mi.) watersheds in HHR 3, while finding that $0.527 \leq \alpha \leq 0.587$ for HHR 4 and larger (30 to 2000 sq.mi.) basins within HHR 3.

All recent studies concur in that small (order of magnitude O.M.~ 1 sq.mi.) rural basins of western Tennessee (HHR 4) produce the largest flood discharges per unit area (specific flow) in the state, most probably reflecting a combination of more impervious soils and higher rainfall intensities, for those shorter durations related to the expected response times of such small basins (say, from 30 minutes to 1 or 2 hours). Indeed, spatial plots of the 10-yr or 25-yr DDF values for such durations display the steepest increasing gradients moving from eastern Tennessee to southwestern Tennessee. For these smaller basins, the largest peak flows in HHR4 are followed closely by HHR3, corresponding to the Nashville Basin. When scaling up to larger basins though, O.M. ~ 10 to 100 sq.mi., floods in western Tennessee increase much less than in other regions, as reflected by the significantly lower values of the scaling exponent α, as discussed before. This is probably due to the much flatter relief in this part of Tennessee, which should result in longer travel times and stronger diffusion effects on the flood hydrographs from larger basins. For larger watersheds, the highest specific flood discharges in Tennessee occur in HHR3, followed by HHR2, and then by HHR4. Peak floods (expressed as specific discharges, i.e., flows per unit area) in HHR1 are smaller, across all basin sizes.

Peak flow equations for urban areas

The only statewide study with equations for peak flows in small urban streams of Tennessee is that of Robbins (1984), who used data from 22 basins with areas ranging from 0.21 to 24.3 sq.mi., located in cities with populations between 5000 and 100,000 inhabitants, with ***IA*** ranging from 4.7 to 74.0 % of the drainage area. These gauges were established specifically for the purposes of the study, and thus had a clear focus towards smaller urban basins.

Because the records used in the analysis were all very short, between 4 and 8 years, frequency analyses could not be performed on gauged annual peak flows. Instead, the observed flow data together with measured precipitation at co-located rain gauges were used to calibrate rainfall-runoff models for each basin, which were in turn used to derive synthetic flood-frequency estimates. The simulation models required gauge-measured rainfall and pan evaporation, but there were only four evaporation stations across the whole state, sometimes located very far from the site of interest. The use of synthetic data evidently increases the estimation uncertainty.

Robbins (1984) included nine basin descriptors in his analysis: drainage area, main-channel slope, main-channel length, the 2-year/24-hr rainfall, a "basin development factor" (accounting for storm sewers, channel "improvements" and revetments, and curbs and gutters on streets), the % of ***IA***, the basin's lag time, the mean annual rainfall, and the peak discharge for rural conditions (from Randolph & Gamble's (1976) equations). His equations for urban peak flow depend only on three of these characteristics: watershed area to an exponent ***a*** of 0.74 or 0.75, ***IA*** to a power of 0.39 to 0.48, and the 24-h rain for T= 2 yrs, with widely varying exponents that decrease from a high of 3.01 (for T = 2 yrs) to a low of 1.10 (for T = 100 yrs). Because of these large exponents, the resulting peak flows are highly sensitive to the 2-yr/24-hr rain, even though this variable does not display much variability at all across Tennessee. This is unfortunate, because in physical terms, one should not expect the response of such small basins, which should be driven by extreme convective events, to be directly related to the rain over 24 hours, which depends on stratiform (i.e., frontal) events. Here, the 2-yr/24-hr rain acts as an index for the actual rainfall over shorter times, which would cause such small basins to flood. Even though the 2-yr/24-hr rainfall also displays a general increase when going from NE to SW Tennessee, its spatial behavior is quite different from that exhibited by rainfall over shorter durations.

For Nashville-Davidson Co., Robbins (1984) refers to Wibben (1976), who analyzed 14 basins with surface areas between 1.58 and 64.0 sq.mi. and ***IA*** between 3 and 37%, concluding that "*in a fully developed residential area, the flood peaks and the basin lag times will not be significantly different from those expected from an undeveloped area*." For predicting urban peak flows in Nashville-Davidson Co., Wibben (1976) thus recommended using Randolph & Gamble's (1976) equations for rural basins. He did warn that "*data were not sufficient to determine if an increase in flood peaks would occur from extremely small basins with extremely intensive development*." Like Robbins (1984), Wibben's (1976) study is based on synthetic flood data.

Crucially, Robbins (1984) analyzed in greater detail five of his basins, using two different methods. In all but one, urban floods derived with both techniques were quite larger than those predicted using rural equations; he explained the single case in which the rural values exceeded the urban ones (Pistol Creek at Alcoa, TN), by stating that it "*is probably due to the location of the urban development within the basin*." Thus, Robbins (1984) results do not support Wibben's (1976) suggestion that there are no significant differences between rural and urban basins.

For Memphis and Shelby County, Neely (1984) used data from 27 basins, ranging from 0.043 to 19.4 sq.mi. He considered nine potential basin descriptors for his regressions: drainage area, ***IA***, mean basin slope, mean basin length, mean channel slope, main channel length, channel condition (a measure of what proportion of the main ditch is lined), channel width, and basin shape. Some of these variables are evidently correlated (e.g., basin length and channel length);

moreover, shape can be inferred by a regression model given basin area and some measure of its length. Also, channel width is a hydraulic variable of little if any interest in explaining peak discharges. Neely's (1984) equations, which are also derived using synthetic (i.e., modeled) flood data, are function of basin area to a power ***a*** between 0.76 and 0.81, and an "average channel condition" to an exponent 1.04 to 1.11. This latter characteristic is supposed to reflect the degree of urbanization of the watershed, but it clearly is only a very rough descriptor or index. Neely (1984) did not use any precipitation variable in his analysis, correctly reasoning that the same rainfall climate applied to all his locations within the greater Memphis area.

Previously, based on the same gauging stations used by Neely (1984), but with much shorter records (2 years at most), Boning (1977) had proposed urban equations for Memphis/Shelby County as a function of drainage area, ***IA***, and maximum rainfall over 120 min. Even though his study was preliminary, as it was based on scant data, it used an event-based approach instead of synthetic data, extracting as much information as possible from all sizeable flood events observed over the two-year period. It also focused on those basin descriptors that one would a priori expect to influence urban floods the most, including the role of the rainfall over a duration that makes conceptual sense (2 hours), considering the range of basin sizes.

There are several consistent issues with the few existing urban equations for Tennessee:

a. Except for Robbins (1984), who focused on cities with 5,000 – 100,000 people, across the state, current estimation methods for urban peak flows are limited to Memphis-Shelby County or Nashville-Davidson County. Sadly, only one of the 22 stations that he used is still under operation; all others were discontinued at the end of the study, or shortly thereafter.
b. Existing methods for urban peak flows are based on short records so all authors used rainfall-runoff models to extend them, generating synthetic annual maxima based on rainfall records. This introduces a large uncertainty in the estimates. Boning (1977) is the exception, as he uses an event-based approach, although for a very small sample size of events.
c. Neely (1984) did not include precipitation as an input variable, so his equations cannot be used for analyzing the effects of rainfall trends. This is a problem because of the attribution issue: if peak flows are changing in time, it will not be clear whether this is due to changes in urbanization, or in extreme rainfall, or both. Also, Neely's (1984) equations do not explicitly consider any variable closely correlated with urbanization rates; "average channel condition" is a very general descriptor that only describes the condition of the mainstream channel at a few cross-sections. It can be assumed that it will tend to be higher in highly urbanized areas, but it is difficult to visualize how it actually relates to typical measures of urbanization, and how will it change in the future, for any specific watershed that is being urbanized.
d. Robbins (1984) used the maximum 24-hr depth as an index for rainfall at each location. This duration is too large for his range of urban basin sizes, as most events causing peak flows in such small watersheds are convective (i.e., thunderstorms), and thus better represented by durations in the range of 6 hours or less (e.g, Smith et al., 2013). This probably explains for the huge range of fluctuation for the rainfall exponent in Robbins' equations as return period is increased. It is important to note that convective events do not scale similarly to frontal (also known as cyclonic or stratiform) storms, neither in time nor space.
e. All of the existing equations are based on data and procedures that are 40 years or older,

and thus are outdated regarding (a) the last few decades of streamflow records (even though there are not many urban stations in Tennessee), (b) the magnitude of recent precipitation extremes, (c) the type of urbanization patterns, (d) the introduction of requirements for flow detention and other storm water control measures, and (e) newer knowledge regarding both the hydrologic response of urban watersheds and data analysis.

2.3 Precipitation trends in Tennessee

Is Tennessee seeing trends in extreme rainfall for those durations relevant to designing urban infrastructure? The frequency and magnitude of extreme events appear to be increasing in many countries, including the U.S. (Wuebbles et al., 2017). The Intergovernmental Panel on Climate Change (IPCC, 2023) has concluded that increases in frequency and intensity of heavy rainfall events are likely in the future, over North America, Europe, and Asia. Regionally, increasing trends in heavy precipitation have already been observed with at least medium confidence for half of the IPCC's (2023) climatic regions.

For the Southeast U.S., Powell & Keim (2015) documented a general increase in the intensity of extreme rainfall events between 1948 and 2012, except for more easterly locations. Skeeter et al. (2019) found that intense precipitation events, defined as the top 1% (or above the 99th percentile), have significantly increased in the region, both in numbers (but only in the fall) and magnitude. Specifically, the Ohio-Mississippi Valley physiographic region, which covers most of Tennessee, has seen one of the highest and most significant increases in event frequency. These results concur with those of Mallakpour & Villarini (2017), who also note that the increasing trend in frequency is stronger than that in magnitude, even though they use gridded instead of weather-station daily data. Brown et al. (2020) analyzed the trends in hourly precipitation at 50 weather stations across the Southeast U.S., finding significant positive increases at 44% of them, but this refers to the mean and the 90th percentile hourly precipitation, not the actual maxima. Rahman et al. (2023) characterized the spatial and temporal changes in all precipitation events with an intensity above a threshold of 1 in/h, finding that these have increased at 53% of the 61 analyzed stations, with the change being significant at 34% of the stations.

The issue with all these studies is that they are either at the event or the daily scale, which is not useful for the purposes of the present work, or else, if hourly, they do not consider the maxima but only changes in the distribution of hourly rainfall values. The study of precipitation extremes for sub-daily durations to has been more limited. Barbero et al. (2017) used hourly precipitation data from the National Centers for Environmental Information (NCEI) to compute trends in hourly and daily precipitation extremes across the contiguous U.S., noting that changes in sub-daily precipitation extremes may emerge more slowly than for daily extremes.

Usually, rainfall extremes are studied with the annual maxima approach, extracting yearly maxima for the different durations of interest (e.g., one hour) before analyzing them. But given that many rainfall data are now available at finer resolution (e.g., the 15-min data available for most weather stations in the contiguous U.S., as well as the 1-min data at ASOS - Automated Surface Observing Systems - rain gauges), it is possible to individualize independent events (or storms), and then investigate whether trends in their characteristics (internal and external) are occurring. In his seminal work on a statistical formulation of the water budget, Eagleson (1978) used derived distributions to represent rainfall variability based on the attributes of individual

events. Meier et al. (2016) showed that this method only requires a few years of rainfall data in order to obtain an adequate climatological description of rainfall totals. Thus, we propose that event-based approaches hold promise for detecting non-stationarity over shorter periods, as well as for shorter records. At the same time, using an event-based framework allows for partial duration (POT) approaches to be used for estimating extremes. As high-quality records of point precipitation with a fine temporal resolution are typically short, it seems important to extract as much information as possible from them, beyond the annual maxima.

We conclude that there is a research gap about possible recent trends in hourly and sub-hourly extreme precipitation in the Southeastern U.S.A., and specifically in Tennessee.

2.4 Descriptors of urbanization at the basin scale

This part of the research work focuses mainly on studying the effects of urbanization on the generation of peak flows. Specifically, in Task 4, a large effort was devoted to the derivation of improved indices of urbanization to be used in regional models for estimating peak flows.

General context for urbanization indices

Most regional peak-flow equations for urban basins (e.g., Southard, 2010; Austin, 2014; Feaster et al., 2014; Kennedy & Paretti, 2014) include ***IA*** (% of impervious area) as a lumped descriptor of the degree of urbanization. This is sound, as developing the land should affect a basin's response, decreasing infiltrability while raising runoff volumes and speeds, which ultimately yields increased peak flows (Konrad, 2003). However, ***IA*** alone cannot fully characterize the effects of urbanization on peak flows, as it does not consider the spatial distribution of the impervious surfaces. E.g., Fig. 2.1 shows two contrasting spatial configurations of pervious and impervious areas, for the same, simple "open-book" basin. Both have the same area, slope, shape, and ***IA*** (50%), but they are expected to produce quite different flood hydrographs for the same rainfall event: Case (a) will result in higher peak flows and shorter lag-times than (b).

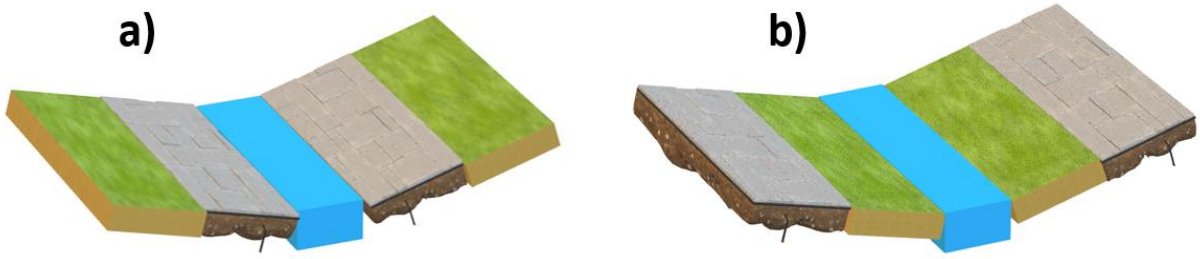


Figure 2.1. Comparison of two "open textbook" basins with the same area, shape, slope, and percentage of impervious areas, but with different spatial distribution of the impervious areas with respect to the main channel and the outlet.

In Case (a), all rain falling on the paved strips will be fully converted to runoff, quickly flowing to the channel and then towards the outlet, without suffering infiltration losses; instead, in Case (b), as runoff produced over the paved areas flows over pervious land, it would be decreased in volume as well as delayed, due to infiltration losses and slower speeds caused by higher surface roughness (Liu & De Smedt 2004), respectively.

This simplistic example clearly shows how ***IA*** ignores important information about how different impervious patches are located with respect to each other, the stream network, and the basin's outlet, which is relevant in predicting peak flows. Any model that only considers ***IA*** as an index for the degree of urbanization is unnecessarily discarding a potentially large amount of explanatory power, by omitting the spatial arrangement of basin surfaces.

A more informative descriptor of urbanization should consider how impervious areas are hydrologically connected to each other, to the pervious patches located downslope, and to the stream network and its outlet. Hydrologic Connectivity (HC), the spatially distributed, directional property induced by topographic gradients and affected by the spatial arrangement of land-cover patches with different infiltrability and distance to the stream network, may serve this purpose, providing novel ways of deriving lumped connectivity-based urbanization indices.

State-of-the-art for connectivity indices

A well-known, GIS-based formulation for basin connectivity was proposed by Borselli et al. (2008) for characterizing the potential for sediment erosion and transport at the basin scale, and was then modified by others (Cavalli et al., 2013; Persichillo et al., 2018; Zanandrea et al., 2019) to focus on other watershed dynamics such as runoff (Hooke et al., 2021) or landslide occurrence (Husic & Michalek, 2022). Its general formulation defines, for each raster cell *k* of the basin hillslope component (i.e., excluding cells from the stream network), an upslope and a downslope component ($\boldsymbol{D_{up,k}}$ and $\boldsymbol{D_{dn,k}}$, respectively; see Fig. 2.2), from which the index of connectivity $\boldsymbol{IC_k}$ is then computed as the logarithm of their ratio (Eq. 2.1; Borselli et al. 2008).

$$IC_k = log_{10}\left(\frac{D_{up,k}}{D_{dn,k}}\right) = \log_{10}\left(\frac{\overline{W}_k\,\bar{S}_k\,\sqrt{A_k}}{\sum_{i=k}^{n_k}\frac{d_i}{W_i\,S_i}}\right) \tag{2.1}$$

Eq. (2.1) provides a connectivity index $\boldsymbol{IC_k}$ for a generic hillslope cell $\boldsymbol{k}$; applying it to all hillslope cells, a connectivity raster is obtained for the basin. Conceptually, the upslope component $\boldsymbol{D_{up,k}}$ denotes the potential for the drainage area upstream of cell $\boldsymbol{k}$ to generate runoff, which is assumed to be proportional to its length scale (square root of its area) and mean slope; the downslope component $\boldsymbol{D_{dn,k}}$ accounts instead for the ease with which runoff can travel from the reference cell $\boldsymbol{k}$ to the stream network, along a travel path imposed by topography.

The expressions for $\boldsymbol{D_{up,k}}$ and $\boldsymbol{D_{dn,k}}$ (Fig. 2.2) contain weighting factors $\overline{\boldsymbol{W}}$ and $\boldsymbol{W}$ for the upslope and downslope components, respectively. A larger or steeper drainage area upstream of reference cell $\boldsymbol{k}$ would result in higher connectivity values for that cell, reflecting conditions associated with potentially larger runoff (or sediment, in the original formulation) volumes or faster runoff travel times. In turn, longer travel distances or flatter paths in the downslope component decrease the connectivity, reflecting a lower potential of cell $\boldsymbol{k}$ to readily contribute runoff to the stream network. $\boldsymbol{S_i}$, $\boldsymbol{d_i}$, and $\boldsymbol{W_i}$ in the downslope component represent the slope, travel distance, and a weighting factor, respectively, for each cell $\boldsymbol{i}$ located on the travel path between reference cell $\boldsymbol{k}$ and the location where that path pours to the stream (or more generally, drainage) network; the summation that defines the downslope component covers all cells 1, ..., $\boldsymbol{i}$, ..., $\boldsymbol{n_k}$ located downstream of cell $\boldsymbol{k}$, along the travel path to the stream network, which is generally obtained by the D8 or D-infinity algorithm (Tarboton 1997; Hooke et al. 2021).

In the upslope component, $\boldsymbol{A_k}$, $\overline{\boldsymbol{S}}_{\boldsymbol{k}}$, and $\overline{\boldsymbol{W}}_{\boldsymbol{k}}$ represent the area, average slope, and average weighting factor of the drainage area located upstream of reference cell $\boldsymbol{k}$. $\overline{\boldsymbol{S}}_{\boldsymbol{k}}$ and $\overline{\boldsymbol{W}}_{\boldsymbol{k}}$ are the arithmetic means of the slopes and weighting factors, respectively, over the upstream area.

The weighting factor may change depending on the process for which the basin connectivity is being established. For instance, Borselli et al. (2008) considered a measure of the potential for erosion, namely the RUSLE C-factor (Renard et al., 1997), as the weighting coefficient for their sediment transport model. On the other hand, morphologic characterizations (Cavalli & Marchi, 2008) or landslide risk assessments (Husic & Michalek, 2022) typically use some measure of topographic roughness. A possible weighting factor of hydrologic interest can be a function of the Manning's surface roughness coefficient $\boldsymbol{n}$, which depends primarily on the land-cover (LC) type of each basin cell (Hooke et al., 2021). Manning's $\boldsymbol{n}$ has been used in different types of applications, such as the study of anthropogenic effects on landscape and sediment transport changes (Persichillo et al., 2018), landslide occurrence (Zanandrea et al., 2019), and runoff paths (Hooke et al., 2021); a typical weighting factor is $\boldsymbol{W} = 1 - \boldsymbol{n}$ (Persichillo et al., 2018; Zanandrea et al., 2019; Hooke et al., 2021), so that rougher cells (with higher $\boldsymbol{n}$) get a lower weight.

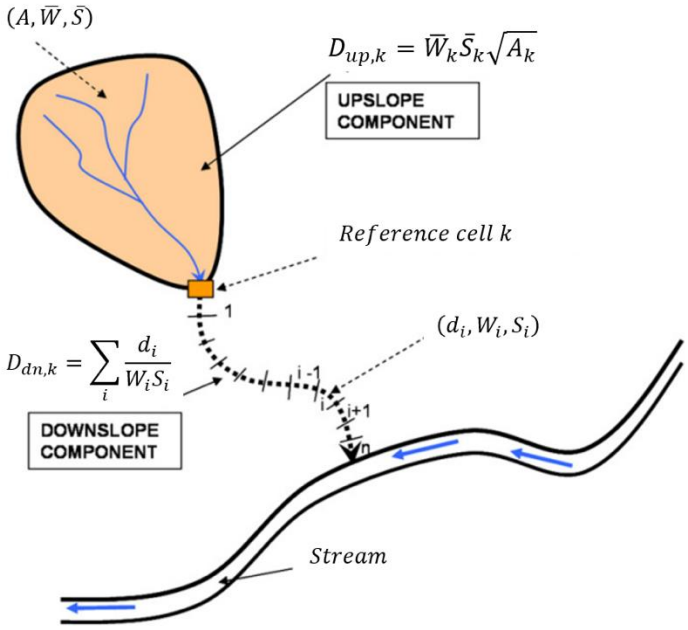


Figure 2.2. Scheme for determining the connectivity of generic cell $\boldsymbol{k}$ (from Borselli et al., 2008).

2.5 Overlapping effects of the variability in rainfall, antecedent conditions, and land cover

As stated in Task 5, event-based analyses must account for the effects of antecedent conditions on urban floods, and how these interact with the spatio-temporal variability in the rainfall input and changing urbanization. Next, we introduce the general problem, sketch a possible methodological approach, and review recent data-driven rainfall-runoff modeling techniques.

The general problem

Antecedent soil moisture (ASM), changes in urbanization, as well as rainfall variability in space and time affect hydrologic response in highly nonlinear ways, varying across basins of different sizes and geomorphic conditions. E.g., for a given drainage, more intense storms should cause floods with shorter lag-times and larger peaks, for the same land-cover (LC) and ASM conditions, particularly in small urban basins affected by intense thunderstorms (common in Tennessee). However, urbanization may also lead to larger floods as compared to pre-development conditions, especially in basins with high-infiltrability soils. This also affects the sensitivity to ASM conditions, so that, e.g., clay-dominated basins should display less variability in their hydrologic response as compared to high-infiltrability watersheds, with respect to ASM.

The complex, highly site-specific nature of how all these aspects interact to affect a basin's hydrologic response made Task 5 one of the most challenging of the project. This is especially true when considering the observation-driven nature of the "experiments" (those basins that happen to have flow data), which precludes considering single scenarios of ASM, LC changes, and rainfall intensity, while controlling for the other factors. Not only is there a range of different combinations of ASM, LC changes, and precipitation conditions in the historic records available for each candidate case-study watershed, but all these drainage basins differ from each other in terms of the sensitivity of their hydrologic response to these interacting aspects.

A possible methodological approach

Initially, these aspects led to major difficulties in determining an experimental methodology for studying the effects of ASM, LC variations, and precipitation intensity, considered individually, and then generalizing the findings across a range of watersheds with different characteristics. However, after an in-depth review of recent rainfall-runoff modeling techniques (Abrahart et al., 1999; Campolo et al., 1999; Dawson and Wilby, 2001; Kumar et al., 2004; Aqil et al., 2007; Shamseldin, 2010; Kratzert et al., 2018; Couta et al., 2019; Hu et al., 2020; Xu et al., 2020), discussed below, we realized that a shift from the originally proposed "event-based" framework to a methodology based on performing continuous flow predictions based on past flow and rainfall records could circumvent some of the above-mentioned difficulties. Such approach focuses our analysis on the effects of LC and LC changes over time, even though at the price of neglecting the influence of ASM conditions, which are *implicitly* accounted for by the model, when considering continuous precipitation time series as the model input. This became our preferred line of action, as this project mainly focuses on the study of the effects of land-cover (degree of urbanization, in particular), while ASM conditions may be regarded as a secondary target; furthermore, this methodology intrinsically presents continuity with, and expands the range of applications of the connectivity-based index of urbanization presented in Section 3.3.

Review of data-driven rainfall-runoff modeling techniques

Recent data-driven applications in hydrology have framed rainfall-runoff modeling as a continuous time-series prediction problem, where flow at future time steps is inferred from the past values of two variables, the discharge itself and the rainfall over the basin (Dawson and Wilby, 2001). In general, the rainfall-runoff process at the basin scale is highly non-linear (Xu et al., 2020), determined by dynamic aspects such as point infiltration rates, the spatio-temporal variability of the rainfall intensity, and the interaction of runoff generated from different

landscape patches. In turn, infiltration is affected by the antecedent soil moisture (ASM) conditions, dependent on past precipitation and evapotranspiration rates (Ponce, 2014).

Because of their versatility in approximating non-linear functions (Hornik et al., 1989; Schäfer & Zimmermann, 2006), artificial neural networks (ANNs) have been used recently to obtain basin rainfall-runoff models calibrated on data (Dawson & Wilby, 2001). Xu et al. (2020) showed that ANN models achieve forecast performances comparable to distributed physical models calibrated and tested on the same basin. However, while the latter require characterizing the basin (Beven, 2012), ANNs allow for quickly developing powerful river-flow forecasting models, only from concurrent flow and rainfall data (Abrahart et al., 1999; Campolo et al., 1999; Aqil et al., 2007; Hu et al., 2020; Xu et al., 2020). Their aptitude for flow forecasting spans a wide range in terms of basin size (from a few to hundreds of thousands of square kilometers; Abrahart et al., 1999; Campolo et al., 1999; Dawson and Wilby, 1999; Kumar et al., 2004; Aqil et al., 2007; Shamseldin, 2010; Couta et al., 2019; Hu et al., 2020; Xu et al., 2020), structure of the rainfall input (i.e., uniform vs. spatially variable; Abrahart et al., 1999; Campolo et al., 1999; Dawson and Wilby, 1999; Aqil et al., 2007, Hu et al., 2020; Xu et al., 2020), as well as the temporal resolution of the data, from a few minutes to hours, days, or even months (Abrahart et al., 1999; Dawson and Wilby, 1999; Kumar et al., 2004; Aqil et al., 2017; Kratzert et al., 2018). Typically, mean flows are used over each time block for larger time-steps (e.g., one day or one month), while instantaneous flows are considered for hourly or smaller time steps (Dell'Aira et al., 2022). In turn, the heterogeneity in the time-series resolution affects the lead times for the forecasts.

The first ANN architectures used in river flow forecasting were multi-layer perceptrons (MLPs) (Abrahart et al., 1999; Campolo et al., 1999; Dawson and Wilby, 1999; Dawson and Wilby, 2001), but after a few years, recurrent neural networks (RNNs) started replacing them (Kumar et al., 2004; Aqil et al., 2007; Kratzert et al., 2018; Couta et al., 2019; Hu et al., 2020; Xu et al., 2020). Among the latter, Kumar et al. (2004) and Aqil et al. (2007) used a simple RNN architecture that recirculates the output and the hidden-layer activations at each prediction to the input layer in the next one, giving the network a short-term memory to learn short temporal patterns in sequential data (e.g., ASM conditions from rainfall time series, in our case). In the most recent applications, however, (e.g., Kratzert et al., 2018; Couta et al., 2019; Hu et al., 2020; Xu et al., 2020), a more advanced type of RNN, namely the long-short term memory (LSTM) network has become widespread. LSTM networks differ from the traditional ANNs and simple RNNs in the type of "neurons" (or units) that are deployed to build their architecture. Traditional ANNs have units that work as a simple input-output function, where the input signal from other units (through "dendrites", using the terminology from biology) is algebraically combined and then transformed by an activation function (popular examples are the sigmoid and the hyperbolic tangent functions), before being sent out as the neuron's output signal (through the "axon") to other units, that in turn receive it as part of their inputs. In contrast, LSTM networks are made of more sophisticated units that, besides exchanging input and output signals with other units, also have an internal state that can store information from past inputs. This gives each LSTM unit (and the network made from many of these units) the ability to produce output signals not only based on the current input, but also incorporating information from past inputs. The rules for the storing/update of the internal state, as well as the way current input and internal state information are combined to produce the output signal at each step are both calibrated during training; hence they depend on the specific process approximated by the model.

In this way, LSTM networks can learn from patterns in sequential data. In the field of hydrology, Kratzert et al. (2019) have shown that the internal states in properly trained LSTM rainfall-runoff models do correspond to hydrologic states from our qualitative understanding of the physical processes related to runoff generation, such as water storage in the groundwater (which controls baseflow), in the upper-layer soil moisture (i.e., ASM conditions, affecting the direct surface runoff), and even water volumes stored as snowpack, when the temperature is below the freezing point.

Based on this, we propose that using an LSTM model trained with spatially distributed rainfall inputs could be used as an investigation tool to analyze the effects of changes in land-cover and how these interact with rainfall-runoff dynamics.

2.6 Uncertainty and bias in extreme rainfall and flood estimation

Here, we discuss possible sources of uncertainty in our traditional estimates of extreme rainfall and floods, and briefly review recent evidence indicating that these could be negatively biased, due to a range of methodological reasons. We first discuss the literature, including our own work on rainfall extremes at the sub-hourly and hourly scales, before focusing on floods.

Errors and biases in the estimation of extreme rainfall

Many authors have recently suggested that standard estimates for IDF-DDF values are outdated in the U.S., due to the non-stationarity of extreme rainfall (e.g., Wright et al., 2021; Kim et al., 2022; First Street Foundation, 2023). Basically, this relates to the fact that IDF-DDF values estimated from long rainfall records, which are necessarily weighted by the older data, do not reflect the current climate. This negative bias evidently affects infrastructure design.

We mostly agree with the issues raised in these studies but opine that there could be further reasons why published IDF-DDF values should be expected to either underestimate actual rainfall intensities in the U.S., and specifically in Tennessee, or at least be more uncertain than typically reported. These issues, which to our knowledge have not been pursued in the literature but for which we have preliminary results from our own research, are related to the fact that: (i) weather stations are sparse in the U.S. (e.g., density is about 17 times smaller in the Southeastern U.S. than it is in Germany), (ii) this low density of rain gauges should affect the regionalization procedures that are used to perform frequency analysis of extreme rainfall, and (iii) most rain gauges in the U.S. have a resolution of 15 min, which introduces a negative bias, as rainfall amounts are totalized over clock or fixed periods, lowering the true, continuous-time maxima, at least for those shorter rainfall durations close to the gauge resolution.

Effects of rain gauge density

Rain-gauge density effects on extreme rainfall estimation have not been systematically studied before, to our knowledge; we propose that they should at least affect the uncertainty of IDF-DDF estimates. Traditionally, confidence intervals for extreme rainfall estimates are obtained by resampling the point-scale data; as frequency analysis procedures are regional though, it should be the case that a higher density of stations should decrease uncertainty in local, point estimates, and vice-versa. We raise this issue because of the sparsity of weather stations in the U.S. - with large surface areas per rain gauge - as compared to the typical spatial scale of the rainfall mechanisms that are expected to result in urban floods in this part of the world.

Herrera et al. (2019) showed that station density has a greater influence on explaining the variance in gridded precipitation products, as compared to how interpolation is performed. This study was done for the mean annual and 50-year return period values for daily rainfall; effects on sub-daily precipitation and event maxima are still unexplored. Berndt et al. (2014) revealed that a higher density of rain gauges decreases the error in various interpolation techniques used to merge radar and gauge data. Lengfeld et al. (2020) showed that in Germany, only 17.3% of the heavy, hourly precipitation events observed with weather radar were detected by the rain gauge network, while weather stations only perceived 52% of the daily, heavy (above a threshold) events in the U.K. Given that Germany and the U.K. have two of the densest gauge networks of any country, we should possibly expect different IDF-DDF values if the station density were higher than what it currently is in the U.S. Pöschmann et al. (2021) suggest that the spatial distribution of 5 min to 1 hour extreme precipitation is not necessarily dependent on relief, as tends to be the case for longer durations. This highlights the importance of a dense gauge network when attempting to estimate DDF values for extreme rainfall.

The most widely adopted approach for developing spatial maps of DDF values is the Regional Frequency Analysis (RFA) introduced by Hosking and Wallis (1997). This method requires either clustering gauges into "homogeneous regions" or else defining a Region of Influence (ROI) around each gauge, as explained by Burn (1990), in order to estimate regional parameters for the selected distribution. But if the weather stations are too far apart, there is a higher probability that the parent distributions will be different, which will affect estimation.

Summarizing, previous studies have underscored the impact of rain-gauge network density on precipitation estimation. However, they have focused on time scales longer than sub-hourly and have not studied precipitation maxima over a range of durations. Thus, current knowledge fails to explain the effects that the spatial density of the gauge network can have on the estimation of the distribution parameters and the resulting DDF values. Additionally, prior studies have often relied on rainfall data with low temporal resolution, which introduces uncertainty linked to sampling maxima from discretized or totalized series, as discussed below.

Effects of rain gauge temporal resolution

These arise because most long-term "continuous" rain gauges in the U.S. actually aggregate rainfall over rather long 15-minutes clock intervals. Thus, during an actual event, we only know the depths collected between 8:00 and 8:15, and then between 8:15 and 8:30, and so on. This creates a problem in engineering practice, when attempting to estimate extreme rainfall for shorter durations, say 2 hours or less: we always underestimate, because the fixed maxima that we extract from the record must be smaller than the true ones, that would have occurred in continuous time. To compensate for this, we use rainfall Sampling Adjustment Factors (SAFs). These corrective factors are derived at those few rain gauges with finer temporal resolutions, say 1-min, where we can compare the true maxima, in (almost) continuous time, with the fixed maxima that would be obtained from a gauge with lower (i.e., 15-min) resolution. Typically, when the duration is the same as the instrument resolution (e.g., obtaining the IDF-DDF values for a duration of 15 min at gauges that totalize every 15 min– we refer to this case as having a sampling ratio of 1), we term the SAFs as Hershfield factors, as he first proposed this method, in Technical Paper (TP) N. 24 (U.S. Weather Bureau, 1953).

When computing the ratio of the DDF value obtained for continuous, 24-hour rainfall data to that for fixed, daily precipitation, as well as the ratio of DDFs obtained from continuous, 60-minute rainfall to that for fixed, 1-hour, Hershfield found an average value of 1.13 in both cases, for different locations.

Papalexiou et al. (2016), Morbidelli et al. (2018), and Llabrés-Brustenga et al. (2020) review previous literature on SAFs. The main issues in relation to this project, is that up to now, no study has looked at SAFs for sub-hourly durations, nor considered their variability. Averaging across locations, for a sampling ratio of 1, most studies have found values that are very close to the 1.13 originally proposed by Hershfield in TP N. 24 (U.S. Weather Bureau, 1953). This means that, on average, DDF values obtained from totalized data should be multiplied by 1.13 in order to obtain the "true" precipitation maxima. Current IDF-DDF values derived for the US (Atlas 14), mostly from gauges that totalize every 15 minutes, use such fixed, mean SAFs uniformly, but this raises questions about their variability. The only research about SAF variability that we are aware of, that also happens to focus on short rainfall durations of interest in urban floods, comes from our group and is discussed next.

Muñoz et al. (2018) and Meier et al. (2019), using 34 concurrent years with 10-min rainfall data from a dense network in Switzerland (52 gauges over 15,940 sq.mi, so 307 sq.mi per station, nine times higher than in the Southeastern U.S.), found a large variability, both at the point (at-a-gauge) scale as well as spatially, in the factors used to "correct" extreme rainfall estimates obtained from totalized precipitation data. For a sampling ratio of 1 (e.g., when estimating the 30-min rain maxima with data totalized every 30 minutes), we found mean SAF values between 1.11 and 1.12, very similar to Hershfield's; but at some sites, for some durations, SAFs could reach values as high as 1.22. In such cases, correcting by the mean SAF, as usually done, would result in a design rainfall underestimated by about 10%.

Moreover, using concurrent lightning strike data, Muñoz et al. (2018) and Meier et al. (2019) demonstrated for the same 52 Swiss stations that these correction factors are significantly larger for the case of convective (thunderstorm) rainfall, than for synoptic or frontal events. It is important to note that most research on this topic has been performed for daily, or at the most hourly data; our group's research has focused on assessing these effects at the sub-hourly durations that are critical to many urban infrastructure designs. As part of the present project, we have retaken this research with a German dataset that includes 862 stations (1 gauge every 160 sq.mi., 17 times higher than in the SE U.S.), with rainfall data at 1-minute time steps. This allows for a much better understanding of the biases introduced by using fixed, clock rainfall data when determining DDF-IDF values for the short durations that are relevant to urban hydrology. Specifically, this information should help us estimate by how much we could err when using 15-min rainfall data to compute DDF-IDF values for durations between 15 minutes and two hours (i.e., for small sampling ratios).

Potential underestimation of frequent floods

One goal of this project was to perform event-based approaches for analyzing rainfall-runoff data, including the use of partial-duration ("peaks-over-threshold", POT) methods for frequency analysis.

As part of this, we reviewed the methodological approaches applied in such type of analyses, ranging from the most popular criteria for event separation to the advantages (and disadvantages) of performing POT frequency analysis (FA), as compared to annual maxima (AM) frequency analysis, including also state-of-the-art techniques for the choice of the threshold.

Pan et al. (2022) provides a review of the technical aspects, including the theoretical framework and established independence criteria, while a list of threshold selection algorithms can be found in Kiran & Srinivas (2021). For the purposes of this project, it is worth mentioning that POT-FA presents multiple advantages as compared to AM-FA, at the price of a more cumbersome procedure that requires the preliminary identification of each independent event in the discharge record, as well as the choice of a threshold for extreme events. The advantages of POT-FA over AM-FA involve: (i) a larger number of peak flow values for the same record length (Tavares & Da Silva, 1983; Robson & Reed, 1999; Bezak et al., 2014; Pan et al., 2022), (ii) the fact that potentially influential low flows (PILFs) are automatically excluded from the peak series (Cohn et al., 2013; Plavšić et al., 2016; England et al., 2019), and (iii) the possibility of predicting unbiased quantiles for small-return-period (i.e. frequent) events (Langbein, 1949; Wyżga, 1995; Keast & Ellison, 2013; Karim et al., 2017). While (i) and (ii) only represent practical advantages, the negative biases introduced when estimating frequent floods with annual maxima (Langbein, 1949) deserve more attention.

This underestimation comes from the conceptual difference in the average interarrival time (AIT) that each method predicts for a given flood value (Wang & Holmes, 2020; Dell'Aira et al., 2023). Specifically, AM-FA estimates a return period (R), i.e., the average number of years with no events between years in which the event did occur. This means that the minimum AIT value that AM-FA can predict is one year, in theory at least. In contrast, POT-FA predicts an average recurrence interval (ARI) independent of any temporal sampling into distinct time blocks (years, in the case of AM), which means that AIT values can also be obtained at the sub-annual scale. This makes POT-FA conceptually more appropriate than AM-FA for predicting frequent floods (i.e., with return periods less than 5 years; Karim et al., 2017; Ball et al., 2019; Dell'Aira et al., 2023), explaining why national guidelines from other countries (e.g., Australia; Ball et al., 2019) recommend the use of POT series for the estimation of this kind of events. In the U.S., however, it seems that the issue of frequent flood (FF) underestimation has not received the same attention, since AM-FA is systematically performed even when predicting events with return periods as low as 2 years, or even less (e.g., Southard, 2010; Austin, 2014; Feaster et al., 2014; Kennedy & Paretti, 2014).

The reasons why AM-FA enjoys greater popularity than POT-FA are clear. The main one is probably the wider availability (both in space and time) of AM series, as compared to POT series. Another aspect that likely plays a role is the lack of systematic research efforts on FF underestimation, with only a few studies involving a limited number of regions in the world, such as Poland (Wyżga, 1995) and Australia (Keast & Ellison, 2013; Karim et al., 2017). Findings from these authors are consistent, and mainly involve the observation that the amount of underestimation is site-dependent and affected by the type of climate and resulting flow regime, as rivers with flashy behavior, typical of dry regions, display greater FF underestimation than rivers in humid regions, with more stable flows (Wyżga ,1995; Karim et al., 2017).

As the search for better descriptors of the degree of urbanization involves testing the performance of regional models in predicting flood quantiles that are generally derived from AM-FA, the issue of frequent flood underestimation has potential implications for this project; it may well be the case that estimates for frequent floods may require some type of bias-adjustment, in order to prevent that the testing is performed on underestimated quantiles.

Chapter 3 Methodology

We describe the methods used for obtaining results related to UofM Task 3 - precipitation trends, Task 4 – urbanization indices, and Task 5 – interacting effects of the variability in rainfall, antecedent conditions, and land-cover, presented in the following order: (i) trends in extreme rainfall, (ii) formulation of a connectivity-based urbanization index, (iii) land-cover effects on rainfall-runoff conversion through continuous simulation with data-driven models, (iv) potential biases and uncertainties in extreme rainfall, and (v) frequent flood underestimation.

3.1 Trends in extreme precipitation frequency

The objectives include analyzing temporal trends in the frequency of hourly and sub-hourly extreme precipitation events for Tennessee and surrounding states, as well as investigating their spatial structure. In this preliminary study, we did not look at trends in the magnitude of extremes, as this requires better assurances with respect to the quality of the precipitation data.

Study region and available data

The analysis covers 11 southeastern U.S. states, depicted in Fig. 3.1. Rainfall data were gathered from two distinct gauge networks. The hourly precipitation data (HPD) network with data every 15 min is maintained by the National Weather Service's (NWS) Cooperative Observer Program (COOP), while the Automated Surface Observing Systems (ASOS) network is a joint program of the NWS), the Federal Aviation Administration (FAA), and the Department of Defense (DoD), with gauges at 1-min resolution. For this project, we downloaded the COOP data from the NWS site, and the processed ASOS data from Iowa State University, Iowa Environmental Mesonet.

Data processing

Initially, we downloaded data for 348 COOP and 149 ASOS stations, with respective temporal resolutions of 15 min and 1 min. Dropping rain gauges with < 10 yrs of data, 147 ASOS and 326 COOP stations were analyzed for the study. The study area and 473 stations are depicted in Fig. 3.1, while Fig. 3.2 shows the distribution of record lengths. These rain gauges cover a large area of 501,000 sq.mi., so that the actual average density of stations is rather low, with only one station every 2744 sq.mi. The mean record length is 25.4 years, while the median is 26.3 years.

We caution here that in an earlier analysis of this COOP data we detected many anomalies; for example, data for the same gauges and periods, but downloaded from different official websites displayed clear differences in the maxima. Given the preliminary nature of this part of our work, it would be out of scope to explore this issue in further detail, but this should be investigated in any future research. It is important to note that 74% of the stations use Fisher and Porter rain gauges, with a totalization period of 15 min, which results in potential under-estimation of the reported DDF values, due to the uniform use of mean SAF values (Section 2.6).

The precipitation data files for the COOP stations contain many missing intervals and data flags. We considered missing values to be null if the interval without data was longer than 6 hours. In the case of intervals with missing data shorter than 6 h, when data flags indicating total depth were available, we decided to equally distribute the cumulative accumulation among each missing value.

This preliminary procedure ensures that we do not overestimate our result by replacing missing data with unrealistically large intensities, but it could also result in underestimation if the missing periods did indeed contain high rainfall intensities.

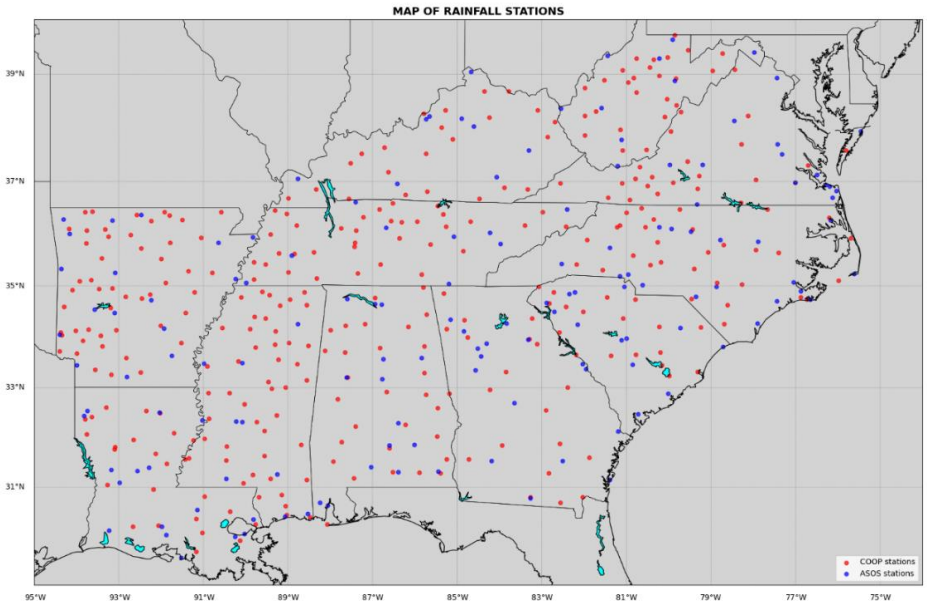


Figure 3.1. Map of the study area with rainfall stations. Blue dots depict weather stations from the 1-min ASOS network while red dots show stations in the COOP network, with 15-min resolution.

Procedures

Because we are interested in event-based analyses, we first identified all independent storms at each gauge by adopting a minimum interevent time ("dry time") of 3 h to distinguish between events. For each individual storm, we first computed its external characteristics (storm duration, total depth) and then proceeded to extract the storm maximum depths for the following 13 different durations: 15 min, 30 min, 45 min, 1 h, 1.5 h, 2 h, 3 h, 4 h, 6 h, 9 h, 12 h, 18 h, and 24 h. Note that maxima were not extracted when the duration of interest was longer than the storm duration, so as to not "dilute" the actual storm depths over longer durations than those over which it actually rained. In this way, we obtained 13 series of rainfall depth maxima at each station, generally of different lengths, one for each duration of interest. For each one of these time series, we then computed the 90^{th} percentile, which was used as a threshold.

Subsequently, for each station and duration of interest, we created a series of exceedances that tracks the number of event maxima per year that exceed the corresponding 90^{th} percentile. Each data point in this time series represents the number of storms exceeding the 90^{th} percentile during that year. Finally, we checked for the presence of a trend in the frequency of extreme events, by applying two test methods to the time series: the parametric Ordinary Least Square (OLS) regression test and the non-parametric Mann-Kendall (M-K) test.

The M-K test analyzes the sign of the differences between combinations of earlier and later data points, and then accumulates these signs to detect a trend. The null hypothesis states that there is no monotonic trend in the series while the alternate hypothesis is that a trend exists;

in such case, the trend can be positive, negative, or non-null. We used the function kendalltau() from the Python library to compute the Kendall tau and ***p***-value for each time series. The value of this statistic lies between -1 to 1 where 1 represents a strong increasing trend while -1 represents a strong decreasing trend. We decided to classify the significance of our trends in two categories: significant trend if $p \leq 0.05$ and less significant trend if $0.1 \leq p < 0.05$. In this way, for each duration of interest, we were able to associate to each station one of the five labels: "significantly increasing," "increasing," "significantly decreasing," "decreasing," and "no trend."

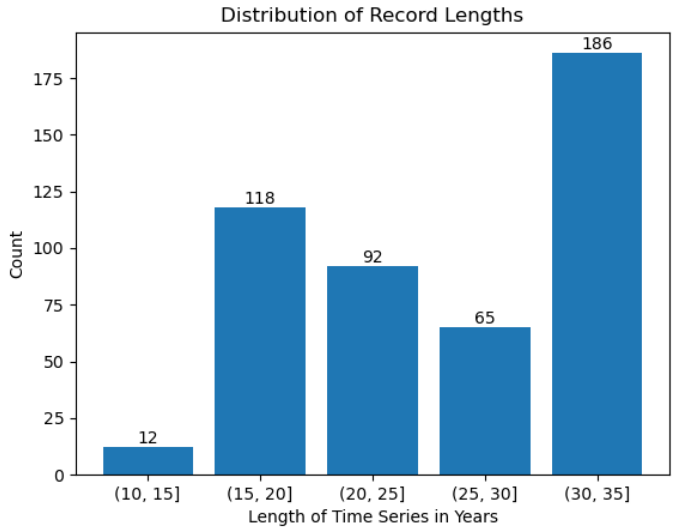


Figure 3.2. Histogram showing the number of rain gauges with different record lengths in years.

The OLS trend test assesses the presence and significance of the trend by fitting a linear regression line on the data. The null hypothesis in the OLS method states that there is no trend, i.e., the slope is not significantly different from 0, while the alternate hypothesis states that the slope significantly differs from 0, i.e., there is a trend. The ***p***-value and confidence intervals of the slope give the significance of the trend. We used the OLS from the "statsmodel" Python library to perform linear regression and compute the required confidence intervals. In this way, for each rainfall duration, each one of the weather stations was classified as having a trend from one of the following four types: "Significant decreasing," "insignificant decreasing," "insignificant increasing," and "significant increasing."

3.2 A hydrologic connectivity index for studying urbanization effects

In this research, we propose new formulations to create a connectivity index that better suits the purpose of deriving a lumped, watershed-scale urbanization index based on hydrologic connectivity. Specifically, we alter existing conceptualizations by proposing (i) new expressions for the weighting coefficient ***W*** and (ii) a generalized formulation for the upslope component.

The new weighting coefficient ***W*** considers the Curve Number (***CN***) of each basin cell, a metric proposed by the USDA (1986) to estimate the potential for runoff generation, based on the Hydrologic Soil Group (HSG) and Land Cover (LC) type of a land patch. ***CN*** values range between

0 and 100, where 0 indicates fully pervious, while 100 refers to no infiltration capacity, so that all precipitation becomes runoff, as for impervious patches with no retention. All possible combinations of LC-types and HSGs receive a ***CN*** value within this range, based on conversion tables proposed by the USDA (1986). Since infiltration during a storm event is affected by soil moisture conditions, three different ***CN*** values are given for each LC-HSG pair, related to dry, average, and wet antecedent conditions, based on rainfall depths in the previous days; the ***CN*** value that is typically tabulated and that will be used in this work corresponds to average antecedent moisture conditions (termed AMC II in USDA, 1986). To work with values in the range between 0 and 1, we define a new weighting coefficient for each basin cell as follows:

$$\boldsymbol{W = CN/100} \tag{3.1}$$

Considering the ***CN*** as an alternative to Manning's roughness coefficient $\boldsymbol{n}$ (Hooke et al., 2021) is a natural choice for the definition of a runoff-connectivity index, similar to the use of the RUSLE C-factor (Renard et al., 1997) by Borselli et al. (2008) in defining their sediment connectivity index. The latter represents the potential for erosion, while the ***CN*** indexes the potential for runoff generation, considering both the infiltrability corresponding to each specific HSG as well as the ease with which runoff can move over the landscape, depending on LC type.

Besides proposing a new $\boldsymbol{W}$, we also introduce a more general formulation for the upslope component, where the characteristics of the drainage area upstream of the reference cell $\boldsymbol{k}$, namely $\overline{\boldsymbol{W}}_{\boldsymbol{k}}$ and $\overline{\boldsymbol{S}}_{\boldsymbol{k}}$, are not necessarily obtained as arithmetic means, but are weighted by considering a custom weighting function $\boldsymbol{w_j = f_w(cell_j)}$ that assigns a weight $\boldsymbol{w_j}$ to each cell $\boldsymbol{j}$ upstream of the reference cell $\boldsymbol{k}$, depending on some of its properties. The traditional formulation with arithmetic means $\overline{\boldsymbol{W}}_{\boldsymbol{k}}$ and $\overline{\boldsymbol{S}}_{\boldsymbol{k}}$ can be regarded as a special case of our approach, with $\boldsymbol{w_j = 1}$ for all cells. By changing the weighting function $\boldsymbol{f_w}$, cells close to the reference cell $\boldsymbol{k}$ can have a stronger influence on its overall connectivity, while cells far from it have a lower effect. This can be modeled as an inverse distance weighting function (Eq. 3.2):

$$\boldsymbol{f_w(d_{j,k}) = \alpha + \beta \frac{1}{d_{j,k}}} \tag{3.2}$$

where $\boldsymbol{d_{j,k}}$ indicates the distance from the upstream cell $\boldsymbol{j}$ to the reference cell $\boldsymbol{k}$, while $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are parameters. The tuning of the weighting function can depend on hydrologic considerations about the correlation scale for runoff dynamics, or else on empirically finding the optimal formulation among alternatives for the connectivity-based urbanization index (for different $\boldsymbol{f_w}$ functions) that, for example, maximizes its predictive power in peak-flow regression models.

Translating all of this into equations, we obtain new formulations for the upslope variables:

$$\overline{\boldsymbol{S}}_{\boldsymbol{k}}^{*} = \frac{\sum_j \boldsymbol{f_w(d_{j,k})S_j}}{\sum_j \boldsymbol{f_w(d_{j,k})}} \tag{3.3}$$

$$\overline{\boldsymbol{W}}_{\boldsymbol{k}}^{*} = \frac{\sum_j \boldsymbol{f_w(d_{j,k})W_j}}{\sum_j \boldsymbol{f_w(d_{j,k})}} \tag{3.4}$$

where $\overline{S}_k^*$ and $\overline{W}_k^*$ are starred to indicate that they are now weighted averages, instead of arithmetic averages as in the traditional formulation given by Eq. (2.1). Clearly, the expressions for $\overline{S}_k^*$ and W_k^* depend on the formulation f_w chosen for performing the weighted averaging.

The resulting, new hydrologic connectivity index, HCI, takes the expression given by Eq. (3.5):

$$HCI_k = log_{10}\left(\frac{D_{up,k}}{D_{dn,k}}\right) = \log_{10}\left(\frac{\overline{W}_k^* \overline{S}_k^* \sqrt{A_k}}{\sum_{i=k}^{n_k} \frac{d_i}{W_i S_i}}\right) = \log_{10}\left(\frac{\overline{CN}_k^* \overline{S}_k^* \sqrt{A_k}}{\sum_{i=k}^{n_k} \frac{d_i}{CN_i S_i}}\right) \quad (3.5)$$

In practical terms, a georeferenced raster of HCI values can be obtained from a basin's digital elevation model (DEM), together with LC and HSG raster maps (Fig. 3.3), once a choice has been selected for the weighting function f_w. All computations are performed in a GIS environment.

Computational aspects regarding the connectivity index

Generalizing the new formulation of the connectivity index (Eq. 3.5), where the characteristics of the upslope component are obtained as weighted averages, necessitates, at some steps, executing operations that cannot be performed using conventional GIS software, because of the lack of a suitable set of GIS functions for computing $\overline{S}_k^*$ and $\overline{W}_k^*$. This is because the weights assigned to the basin cells when aggregating information for the upslope component $D_{up,k}$ at each point k depend on the location, which means that there are virtually as many sets of weights as the number of basin cells, and each weight, given a generic reference cell k (refer to the scheme in Fig. 3.4), must be determined from the distance from the upstream cell j to the reference cell k for which $D_{up,k}$ is being computed at the given iteration.

For operations that could not be run with traditional GIS, we generated our own algorithm (see Algorithm 3.1 and Figs. 3.5 and 3.6), implementing it functionally into a Python program. Algorithm 3.1 summarizes the operations performed in the first step, to derive the set of raster maps of upslope contributions from different distances. The algorithm function takes as input the number of rows ($ySize$) and columns ($xSize$) of the DEM model, the raster rW of weighting coefficients (obtained either as a function of n or CN), the raster of slopes, rS, and the flow direction raster $rFLowDir$ (obtained using the D8 algorithm). It should be noted that all the input raster maps can be derived from the basin DEM using any traditional GIS software. We recommend preliminary preprocessing of the DEM to remove any local depressions.

Note that a dynamic indexing is used in Algorithm 3.1 to keep track of the contributions of upstream cells across the basin, one step at a time. In the first iteration, each basin cell is assigned a unique id (see Fig. 3.5), and all these ids are stored in the $raster\_id$ raster variable. A second raster, $dest\_id$, is produced from the $rFlowDir$ raster, which stores, for each cell, the id of the next adjacent cell (i.e., the id of the destination cell), based on the D8 algorithm. The contributions of upstream cells distant only one cell from the reference point are identified by comparing the ids stored in $raster\_id$ and $dest\_id$. Specifically, a given cell at location $[i][j]$ of the basin DEM, with a certain id stored in $raster\_id[i][j]$, takes its 1-cell-distant contributions from those cells of the basin $[k][l]$ for which the id in the $dest\_id[k][l]$ is the same as in $raster\_id[i][j]$.

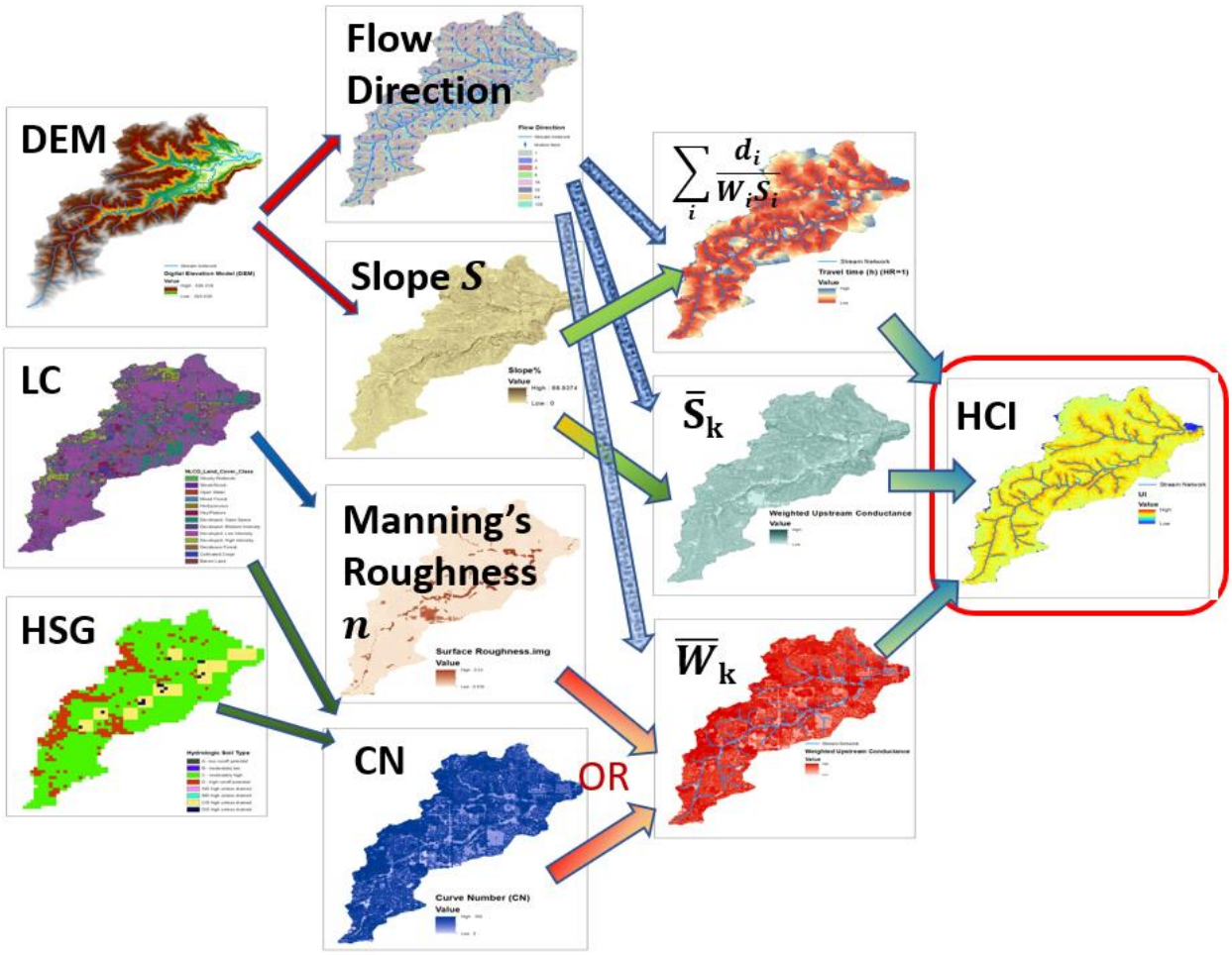


Figure 3.3. Break-down of the sequential GIS operations needed to derive a raster of the connectivity index from DEM, LC, and HSG maps.

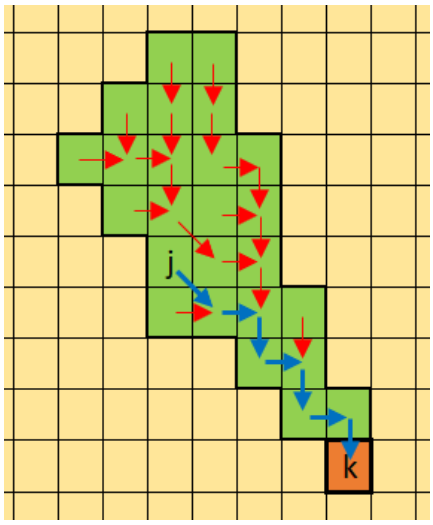


Figure 3.4. Generic reference cell ***k*** and its upstream drainage area, delimited by thick edges. Arrows show flow directions, while the flow path for a generic upstream cell ***j*** is noted with bold arrows.

Algorithm 3.1. Pseudocode of the algorithm to obtain raster files "W_distance*" and "S_distance*" for different distances (in number of cells). Each raster provides the cumulative W and S at each cell, taken from all the cells located at a fixed distance upstream of it. E.g., "W_distance3" provides the cumulative weighting coefficients for all upstream cells that are 3 steps away from reference cell ***k***.

```
1:  void upslope_component_S_W (xSize, ySize, rS, rW, rFlowDir)
2:      define raster_id[1:xSize][1:ySize] = 0 // initialize raster for storing cell id's
3:      id = 0 // initialize id
4:      for each basin cell [i][j] do
5:          raster_id[i][j] = id
6:          id = id+1  // increment the current id so that each basin cell will receive a unique id
7:      cell_dist = 1
8:      repeat
9:          stop = true
10:         define rWstar[1:xSize][1:ySize] = 0// initialize raster for storing the weights
11:         define rSstar[1:xSize][1:ySize] = 0// initialize raster for storing the slopes
12:         define dest_id[1:xSize][1:ySize] = -1 // initialize raster for storing the cell id of the cell downstream
13:         for each cell rFlowDir[i][j] do
14:             determine cell rFlowDir[k][l] downstream of rFlowDir[i][j]
15:             dest_id[i][j] = raster_id[k][l]
16:         end for
17:         for each cell raster_id[i][j] do
18:             for each cell dest_id[k][l]==raster_id[i][j] do
19:                 rWstar[i][j] = rWstar[i][j] + rW[k][l]
20:                 rSstar[i][j] = rSstar[i][j] + rS[k][l]
21:             end for
22:         end for
23:         save rWstar as "rasterW_distance"+str(cell_dist)
24:         save rSstar as "rasterS_distance"+str(cell_dist)
25:         cell_dist = cell_dist+1
26:         raster_id = dest_id // moving one cell upstream along every runoff path
27:         if raster_id[i][j] == raster_id[k][l] for any [i][j]≠[k][l] then // if all flow paths have been tracked up to the outlet cell
28:             stop = false
29:         end if
30:     until stop == true
```

Each subsequent iteration works similarly, with the difference that the $\boldsymbol{dest\_id}$ raster of the previous iteration becomes the $\boldsymbol{raster\_id}$ of the current iteration, in order to consider all the 2-cell-distant contributions in the 2nd iteration, all the 3-cell-distant contributions in the 3rd, and so on. When the flow paths on the hillslope component have all been completed, all cells of $\boldsymbol{raster\_id}$ at the current iteration will have the initialization value of -1, and the algorithm stops.

To avoid repeated computations of the connectivity raster for varying weighting functions while fine-tuning its parameters by trial-and-error, we break down the procedure in two steps, as outlined in Fig. 3.6. We first derive a series of rasters of the basin, each providing the cumulative $\boldsymbol{W}$ and $\boldsymbol{S}$ of upstream cells from a distance of 1, 2, etc., cells from the reference. Second, we compute the weighted averages to obtain $\overline{\boldsymbol{S}}_{\boldsymbol{k}}^{*}$ and $\overline{\boldsymbol{W}}_{\boldsymbol{k}}^{*}$ for different weighting functions $\boldsymbol{f_w(d_{j,k})}$. Note that the contributions (from different distances) stored in each raster file from the first step (Fig. 3.6) are not weighted yet; hence, these computations are performed only once. This is helpful because these are so computationally intense, that they must run in the High-Performance Computing (HPC) infrastructure at UofM. We parallelized these operations by running the computation of each individual raster of contributions on a different CPU (i.e., the

repeat loop in the Algorithm 3.1 pseudocode is parallelized). On the contrary, the weighted summation of raster maps in step 2, to obtain $\overline{W}_k^*$, can be easily performed on a single CPU.

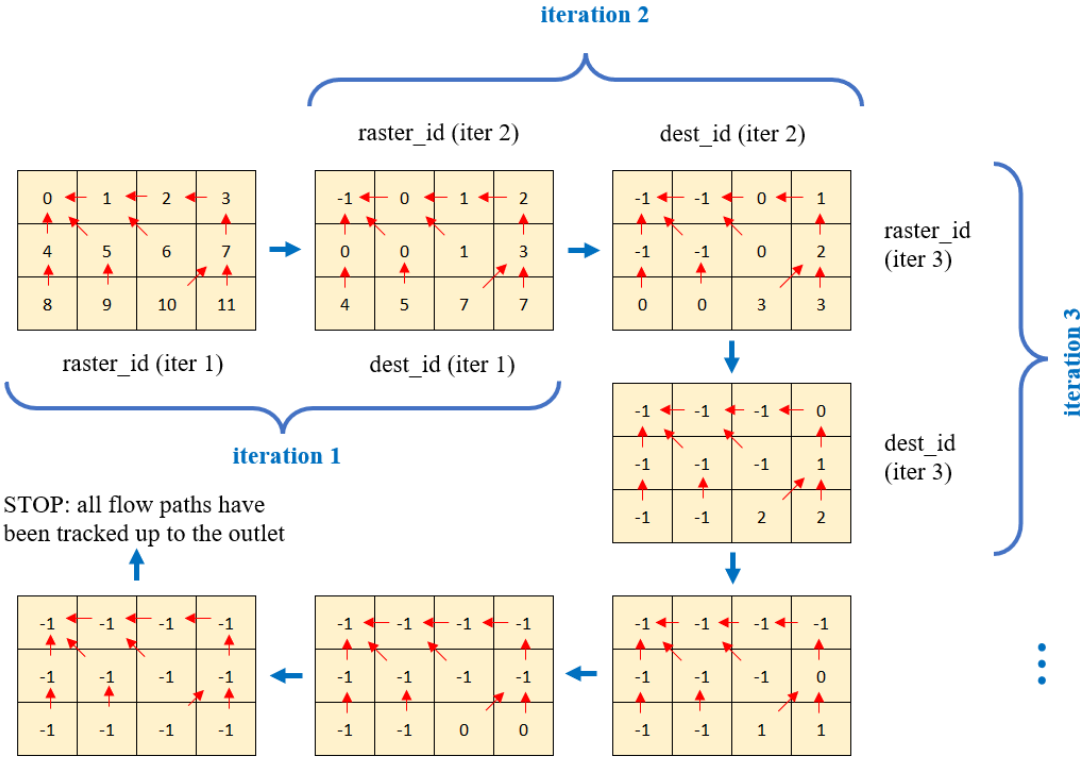


Figure 3.5. Explanation of the dynamic indexing adopted in Algorithm 1 to keep track of the contributions of upstream cells across the basin, one step at a time. An arbitrary flow direction raster (indicated by the arrows) is assumed.

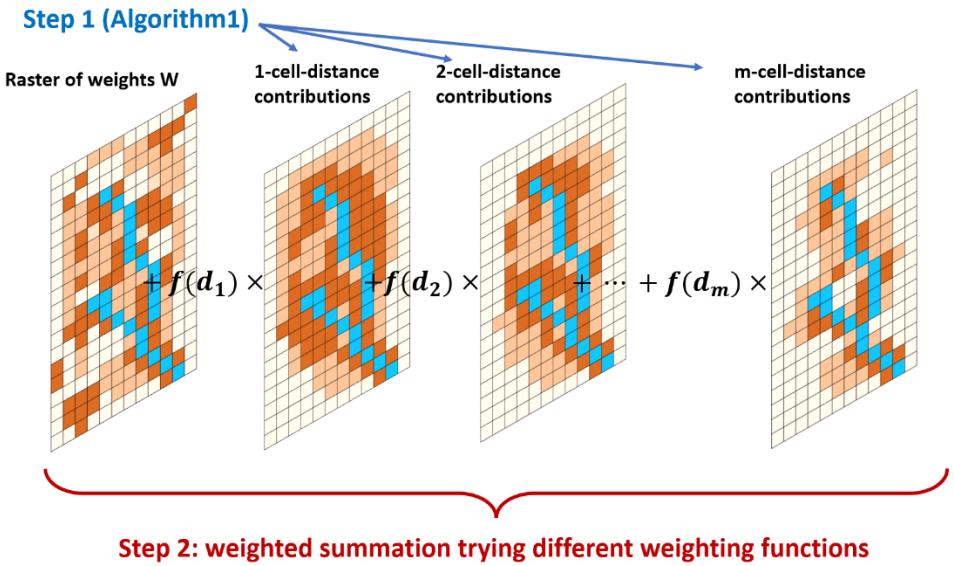


Figure 3.6. Scheme of the procedure to optimize the use of computational resources while calculating the upslope component for a range of weighting functions $f_w(d_{j,k})$. Operations to obtain the raster of contributions from different distances (expressed in number of cells) are performed on a HPC, while the weighted summation considering different weighting functions can be done on a traditional computer.

3.3 Formulating a new connectivity-based urbanization index

From the methods proposed in Section 3.2, an Urbanization Index for the generic reference cell k can be defined as follows:

$$UI_k = \frac{HCI_k}{HCI_{imp,k}} \tag{3.6}$$

The right-hand side of Eq. (3.6) is the ratio of the connectivity HCI_k of the actual basin at cell k to that of a virtual basin with the same relief and stream network but totally impervious surfaces, which would have the highest potential connectivity $HCI_{imp,k}$. This approach recalls the runoff coefficient concept, the proportion of total precipitation that becomes effective rainfall and thus contributes to runoff, net of infiltration or other losses. Thus, the runoff coefficient denotes that rainwater that actually becomes runoff, relative to the maximum runoff that could be potentially achieved if infiltration and other losses were negligible. Similarly, the ratio of the connectivity of the actual basin cell to that for the totally impervious, virtual copy, provides a relative measure of the actual basin connectivity, as compared to the ideal connectivity that would maximize runoff and thus peak flow, of a basin with impervious areas and the smoothest possible travel paths towards the stream network. However, while the runoff coefficient is a lumped quantity defined for an entire watershed, the urbanization index of Eq. (3.6) is specified for each hillslope cell, so that it can be regarded as a distributed watershed property.

A lumped urbanization index can be obtained by averaging all the connectivity ratios over the basin (Eq. 3.7). It should be noted that the connectivity index considers how far each hillslope cell is from its nearest pour point on the stream network, but does not account for the cell's relative location with respect to the basin outlet. To address this limitation, a weighting function $f_{w,UI}(d_{SNk})$ can be introduced in the definition of the lumped urbanization index. If we indicate as SN_k that stream network cell where reference cell k finally pours to, and as d_{SNk} the travel distance from cell SN_k to the basin outlet, following the stream network (see Fig. 3.7), the weight $f_{w,UI}(d_{SNk})$ given to cell k as a function of the distance of SN_k to the basin outlet would allow to consider greater contributions from those cells that drain to the network closer to the basin outlet, and smaller contributions from cells that pour to stream segments further upstream. This allows to differentiate cells depending on their location relative to the basin outlet.

$$UI = \sum_k f_{w,UI}(d_{SNk}) \frac{HCI_k}{HCI_{imp,k}} \tag{3.7}$$

The tuning of the weighting function $f_{w,UI}(d_{SNk})$ could be based on hydraulic considerations about the travel speed of flood waves along the stream network, or else could be empirically derived by optimizing the predictive power of the urbanization index UI when including it as an explanatory variable while obtaining regression-based urban peak-flow equations.

Testing the predictive power of the new urbanization index

We test the predictive power of the new index ***UI*** against the percentage of impervious area (***IA***), when both are included as alternative proxies for the degree of urbanization in regional regression equations for peak flow estimation in urban basins. As first case study, we consider a regional model for Missouri (Southard, 2010), where 35 urban basins were selected, based on criteria related to (i) minimum length of peak flow records of at least 9 yrs, (ii) a relatively stable level of urbanization within the timeframe for which there were flow data, (iii) no major river diversions, and (iv) minimal presence of retention ponds, to ensure a hydrologic response that mainly reflects surface runoff dynamics. The basins cover a range of sizes from a few to some hundred square kilometers, with ***IA*** ranging from a few percentage points to almost 50%. Their regression equation for $\boldsymbol{Q_T}$, the peak flow with return period ***T***, takes the expression:

$$\log Q_T = \log \beta_0 + \beta_1 \log A + \beta_2 IA \tag{3.8}$$

where ***A*** is the basin area and ***IA*** the percentage of impervious areas, while $\boldsymbol{\beta_0}$, $\boldsymbol{\beta_1}$, and $\boldsymbol{\beta_2}$ represent model parameters. Since Eq. (3.8) is linear in logarithm form, the model coefficients can be determined by the ordinary least squares (OLS) regression method.

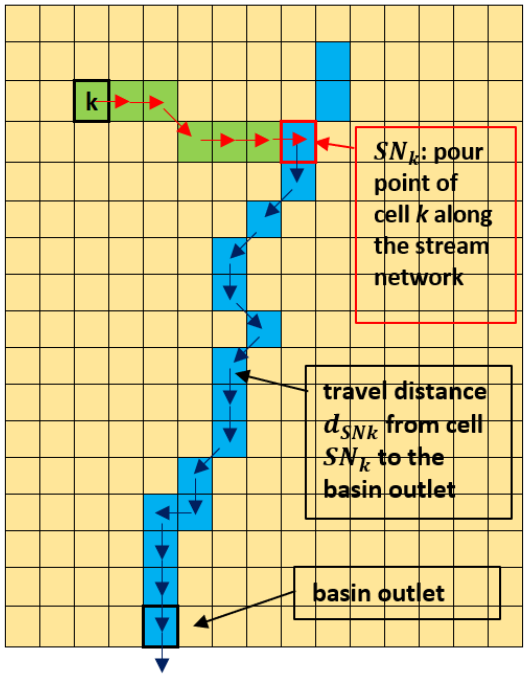


Figure 3.7. Scheme depicting the stream network component $\boldsymbol{d_{SNk}}$ of the total distance from generic cell $\boldsymbol{k}$ to the basin outlet, which is used for weighting that cell when aggregating all information into the lumped, connectivity-based urbanization index $\boldsymbol{UI}$ (Eq. 3.7). Water from cell $\boldsymbol{k}$ first reaches the stream network at cell $\boldsymbol{SN_k}$ through a path along the hillslope component, and then continues its travel to the basin outlet along the channel. The distance $\boldsymbol{d_{SNk}}$ is measured from cell $\boldsymbol{SN_k}$ to the outlet.

3.4 Studying land-cover effects using data-driven, continuous rainfall-runoff models

It is sensible to posit that different portions of a basin may have hierarchically different effects over its response under a given rainfall input, depending on aspects such as their proximity to the stream network and the basin's outlet (Dell'Aira et al., 2022), as well as the land-cover (LC) type and arrangement of upstream drainage areas, as determined by the surface topography; the LC type affects the infiltrability (or imperviousness) levels as well as the runoff travel times, while the size of the upstream draining areas determine potential local runoff volumes.

Identifying those portions of a basin that have a stronger impact in determining variations in the peak flow at the outlet, under different localized rainfall inputs, may help explain the effects that different LC types, as well as their changes, have on hydrologic response. By training a universal, data-driven function approximator (e.g., an artificial neural network model; Hornik et al., 1989; Schäfer & Zimmermann, 2006) to continuously simulate how spatially distributed rainfall inputs are converted into runoff at the outlet, we can investigate the changes in model performance under different combinations of rainfall input. We can also test whether a partial, incomplete depiction of the overall rainfall, considering only some portions of the basin, would provide enough information to match the performance of a benchmark model trained with the full precipitation information. If that is the case, then the watershed portions with the highest predictive power in the model will be those with the strongest influence on the basin's hydrologic response. In other words, if a universal function approximator is able to map rainfall inputs at those basin locations into outlet discharge with satisfactorily high accuracy, then it must mean that the watershed's response is mostly affected by such portions of the watershed.

We believe that an LSTM model is most suited for this purpose, given its ability to learn time patterns in the input time series (see Chapter 2). Specifically, it can adapt the rainfall-runoff simulation to the current antecedent soil moisture (ASM) conditions, as these are determined by prior precipitation patterns. In this way, analyzing those portions of the basin where rainfall inputs yield the highest predictive power will provide insights into the effects of land cover and spatial proximity to the stream network and outlet, net of ASM influences.

3.5 Uncertainty and biases in hourly and sub-hourly extreme rainfall

It is of interest to derive urban peak flow equations that have rainfall as a variable, so they can reflect future changes in design precipitation values. Because of this, we were interested in better assessing as-of-yet little explored potential sources of uncertainty, as well as potential biases, when estimating extreme rainfall quantiles, i.e., DDF-IDF values. We briefly discuss the methods used for preliminary assessment of the effects of rain gauge density, and then explain how gauge resolution introduces underestimation. It is important to note that not only does the Southeastern U.S. have a low density of rain gauges, but most of these measure rainfall at 15-min clock intervals. Thus, we can only explore these effects by using data from locations with much higher station densities and temporal resolutions. The German data we use come from a network with a density that is 17 times larger than that in the Southeastern U.S., allowing us to simulate the effects of estimating DDF-IDF values with sparse rain gauge networks. Of course, there are no tropical cyclones or storms in Germany, but they do get a mixture of convective and frontal precipitation events, as is the case in large parts of the U.S.

Effects of station density

To address the research gaps identified in Section 2.6 (sub-section on "Effects of rain gauge density"), we take advantage of high temporal-resolution data (1-min) from Germany, from a network with a noticeably higher spatial resolution than in the U.S. (about 17 times higher). We aim to fill this research void by simulating numerous scenarios for networks with low rain-gauge density, randomly drawn from this high-resolution data. Through this approach, we intend to quantify the uncertainties and possible biases that arise in DDF values due to lower rain gauge densities.

Precipitation Data

We obtained rainfall data at 1-min resolution for 1093 German stations operated by the German Weather Service (DWD), from their website Wetter und Klima - Deutscher Wetterdienst - CDC (Climate Data Center) (dwd.de). Most stations start after 2000, and the gauge with the longest record covers from 2001 to 2021. Since our main goal relates to the effects of station density, we do not want to discard too many stations from the analysis by imposing too strict conditions on record length. Therefore, we optimized the available number of stations given the lengths of their record, considering the same study period for all stations. Eventually, 862 stations with a common record period from the start of 2009 to the end of 2021 (13 years) were selected for the analysis. The map with the locations of the 862 rain gauges is shown in Fig. 3.8.

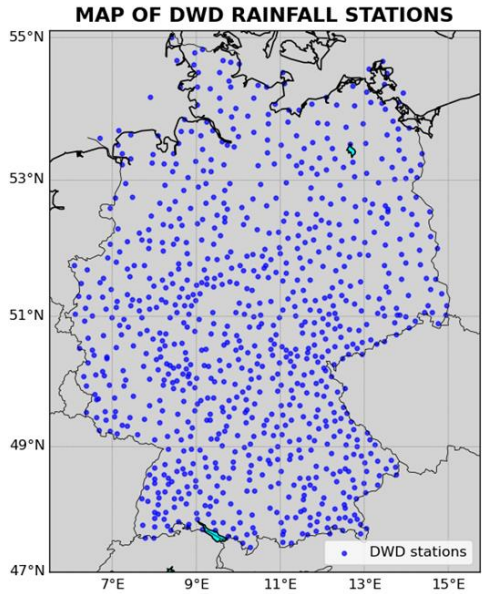


Figure 3.8. Map of the study area (Germany), depicting the 862 weather stations with concurrent rainfall data at 1-minute temporal resolution for the 13-yr long period from 2009 to 2021.

Regional frequency analyses

We first extracted independent events from the 1-min time series, using a minimum inter-event time of 3 hours as criterion. The maxima for durations of 10, 20, 30, 45 min, 1,1.5 and 2 h were then extracted for each storm. We attempt as closely as possible to mimic regional frequency analyses used in Atlas 14, to better understand how the low density of rain gauges in Tennessee (and the U.S.) could be affecting DDF estimates; hence, we used the region of influence (ROI)

approach (Burn, 1990) to determine homogenous areas. Each gauge is surrounded by a unique homogeneous region (HR); we arbitrarily adopted a 60 km (37.3 mi) radius for the ROI. Thus, for each of the study gauges, we assumed that any weather station within 60 km belongs to the HR for that particular gauge, so we can pool their data to estimate the regional probability distribution parameters at that site. Given the number of stations and the size of Germany, there are on average 27.3 gauges per ROI, when using the full density of weather stations. Then, for each duration of interest, we combined the time series of event maxima of all stations in the ROI, obtaining a long series of maxima (on average, 354 years long), and then we extracted the partial duration series (PDS) for the region from this series. On average, we took four values from each year to create the PDS for each ROI. The lengths of the PDSs for the different ROIs are variable, as they depend on the actual number of stations in each one of them.

We adopted the Generalized Pareto (GPA) distribution to fit the PDSs obtained for each gauge's ROI, as well as to perform at-a-site frequency analyses for each gauge. We used a POT approach, as the short records precluded AM. There are two sets of PDSs: those for each of the 862 gauges, with 52 values (13 yrs. x 4 events per yr on average), as well as that for each ROI, with a variable size depending on gauge numbers. As we are interested in the effects of station density, we generated results for the following percentages of the original gauge density: 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, and 100%. For each density, we sampled 100 possible realizations of the network, with replacement. Altogether, we thus computed 10 (densities strictly smaller than 100%) x 7 (durations) x 100 + 7 = 7007 realizations, across all different station densities.

Now, given a density, for each duration of interest and each realization, we extracted the PDS for each ROI, computing the L-moments and the GPA shape parameter for that ROI. Then, a regional flood frequency analysis was performed, using the regional shape parameter along with the normalized location and scale parameters (from the corresponding at-a-station PDS), obtaining normalized DDF values for different return periods. Finally, the index flood procedure (Dalrymple, 1960) was implemented, multiplying the normalized DDF values by the mean of the at-a-station PDS, to get the estimate of the regional DDF value for each gauge. Similarly, for each duration, we obtained the DDF values for each station for all other network realizations.

In this study, we consider that our DDF values calculated at each site using a regionalized shape parameter derived from the ROI at 100% station density are "correct." To assess the potential bias caused by variations in station density, we then computed the percentual difference between each DDF value obtained for lower densities and the corresponding "correct" DDF value. As a result, we generated 100 DDF-difference values (in %) for each station, specific to each combination of return period, duration of interest, and station density.

Bias due to the temporal resolution of the rain gauges

For this work (Meier et al., 2022; Marasini et al., 2023; Meier et al., 2023), we basically reproduced the approach followed by Muñoz et al. (2018) and Meier et al. (2019) for the 10-min Swiss rainfall data, but now using 1-min resolution rainfall data for 862 weather stations across Germany. The raw 1-min data were checked for consistency and corrected, when needed, based on the quality controlled 10-min data. To perform frequency analyses based on partial duration (POT) series, independent events needed to be determined at each gauge. Based on the corrected 1-min series and using a minimum inter-event time of 3 hours, we separated all storms at each location, before extracting the "true" rainfall minima for each storm, for

durations of 5, 10, 15, 20, 30, 45, 60, 75, 90, and 120 minutes. This was done by using a sliding time window, moved "continuously," 1-min time step at a time. Then, the corrected 1-min time series were aggregated over fixed durations for totalization periods of 5, 10, 15, 20, 30, and 60 minutes, to mimic the data that would be generated at a typical gauge that totalizes over fixed, clock-time windows. To estimate the at-a-station variability of SAFs, we introduced a novel resampling methodology, whereby the start time of the series was successively moved by 1-min steps before totalizing. For example, when aggregating the 1-min time series into fixed 15-min series, we totalized the original, "continuous" series in 15 different ways. The resulting maxima are different according to how the 1-min data totalized, adding up to different series for frequency analysis, and thus in different DDF values for the given duration and return period. In this way, dividing the single "true" DDF value by each one of the fixed DDF values, we were able to generate a sample of SAFs, for every combination of rainfall duration and totalization time, which allowed us to understand the at-a-station variability.

After computing SAFs at each site, for all possible combinations of duration, totalization time, and way of totalizing, we analyzed their at-a-station variability (mean, median, range, skewness, distribution), pooled statistics, as well as their spatial structure, using simple geostatistical tests.

3.6 Methodology for correcting AM-based frequent flood estimates

Let us consider the average interarrival time (AIT) $\boldsymbol{T}$ and the corresponding quantiles $\boldsymbol{Q_{AM}(T)}$ and $\boldsymbol{Q_{POT}(T)}$, obtained from the inverse cumulative distribution functions (CDFs) for the annual maxima (AM) and peaks-over-thresholds (POT) series, respectively. Their ratio (Eq. 3.9), which from now on we term the "quantile ratio" (Dell'Aira et al., 2023), is a measure of the under-estimation of the $\boldsymbol{T}$-year flood from AM frequency analysis (AM-FA), given that the actual frequency of frequent floods is better reflected by the POT-based estimates (Karim et al. 2017).

$$\boldsymbol{r(T)} = \frac{\boldsymbol{Q_{AM}(T)}}{\boldsymbol{Q_{POT}(T)}} \tag{3.9}$$

$\boldsymbol{Q_{AM}(T)}$ is smaller than $\boldsymbol{Q_{POT}(T)}$ for small $\boldsymbol{T}$s and becomes closer to $\boldsymbol{Q_{POT}(T)}$ for increasing $\boldsymbol{T}$s, as a direct consequence of the principles behind Langbein's (1949) equation (Dell'Aira et al., 2023). Hence, $\boldsymbol{r(T)}$ must be smaller than 1, theoretically (i.e., ignoring any sampling variability effects). From $\boldsymbol{r(T)}$, the percentage of underestimation introduced by using AM instead of POT to derive frequent flood quantiles can be derived as shown in Eq. (3.10) (Dell'Aira et al., 2023):

$$\boldsymbol{u(T)} = \boldsymbol{[1 - r(T)] \times 100\%} \tag{3.10}$$

Adopting the Generalized Pareto (GP) distribution (with shape, scale and location parameters $\boldsymbol{\xi_p}$, $\boldsymbol{\sigma_p}$, and $\boldsymbol{\mu_p}$,respectively) for modeling the POTs and the Generalized Extreme Value (GEV) distribution (with shape, scale and location parameters $\boldsymbol{\xi_g}$, $\boldsymbol{\sigma_g}$, and $\boldsymbol{\mu_g}$, respectively) for modeling the AM (see Dell'Aira et al., 2023, for the theory and reasons behind such choices), the inverse CDFs are given by the following Eqs. (3.11) and (3.12) (Dell'Aira et al., 2023):

$$Q_{AM}(\xi_g, \sigma_g, \mu_g, T) = \begin{cases} \mu_g + \frac{\sigma_g}{\xi_g}\left\{\left[\ln\left(\frac{T}{T-1}\right)\right]^{-\xi_g} - 1\right\}, & for\ \xi_g \neq 0 \\ \mu_g - \sigma_g \ln\left[\ln\left(\frac{T}{T-1}\right)\right], & for\ \xi_g = 0 \end{cases} \quad (3.11)$$

$$Q_{POT}(\xi_p, \sigma_p, \mu_p, \lambda, T) = \begin{cases} \mu_p + \frac{\sigma_p}{\xi_p}\left[(\lambda T)^{\xi_p} - 1\right], & for\ \xi_p \neq 0 \\ \mu_p + \sigma_p \ln(\lambda T), & for\ \xi_p = 0 \end{cases} \quad (3.12)$$

If we also assume that the number of exceedances follows a Poisson process (Wang & Holmes, 2020; Pan et al., 2022) then, applying the Total Probability Theorem (Dell'Aira et al. 2023) and the consequent duality property between the GP and GEV distributions (Wang and Holmes 2020; Prosdocimi & Kjeldsen, 2022), we can reparametrize the quantile $Q_{POT}(\xi_p, \sigma_p, \mu_p, \lambda, T)$ in Eq. (3.12) using the GEV-parameters, obtaining an expression for the POT-based quantile $Q_{POT}(\xi_g, \sigma_g, \mu_g, T)$ function of the parameters of the AM distribution (Dell'Aira et al., 2023). This allows to rewrite the quantile ratio as a function of the GEV parameters, as in Eq. (3.13):

$$\hat{r}(\xi_g, \sigma_g, \mu_g, T) = \frac{1 + \frac{1}{\xi_g}\frac{\sigma_g}{\mu_g}\left[\left(\ln\frac{T}{T-1}\right)^{-\xi_g} - 1\right]}{1 + \frac{1}{\xi_g}\frac{\sigma_g}{\mu_g}\left[T^{\xi_g} - 1\right]} \quad (3.13)$$

The underestimation $\hat{u}(\xi_g, \sigma_g, \mu_g, T)$ of the T-year event using the duality-derived quantile ratio in Eq. (3.13) can be obtained replacing $\hat{r}(\xi_g, \sigma_g, \mu_g, T)$ into Eq. (3.10) (Dell'Aira et al., 2023).

As the GEV parameters can be easily obtained by performing an AM-FA, the duality-derived quantile ratio $\hat{r}(\xi_g, \sigma_g, \mu_g, T)$ can be used to correct the biased estimates of the T-year quantile from AM-FA, without the need to actually perform POT frequency analyses. Specifically, the unbiased T-year quantile $Q^*_{POT}(T)$ is derived from the following Eq. (3.14).

$$Q^*_{POT}(\xi_g, \sigma_g, \mu_g, T) = \frac{1}{\hat{r}(\xi_g, \sigma_g, \mu_g, T)} Q_{AM}(\xi_g, \sigma_g, \mu_g, T) \quad (3.14)$$

Eq. (3.13) can be used to investigate how the distribution parameters affect the under-estimation. In turn, as climate and basin descriptors determine the distributions of floods, the physical drivers of frequent flood underestimation can be elucidated (Dell'Aira et al., 2023).

3.7 Urbanization trends across Tennessee

Urbanization trends are studied using the National Land-Cover Database (NLCD; Homer et al., 2020), available for the entire CONUS at a resolution of 30 by 30 m, for the years 2001, 2004, 2006, 2008, 2011, 2013, 2016, 2019, and 2021. These land-cover maps provide information on a variety of LC types, including developed land, expressed at four different levels of urbanization: "developed open space," "developed low intensity," "developed medium intensity," and "developed high intensity," corresponding to increasing ranges of IA. E.g., "developed open

space" patches display a percentage of IA ≤ 20%, while "developed high intensity" indicates any basin cell with a percentage of IA > 80%, and potentially up to 100% (i.e., totally impervious).

To decrease the computational burden, national raster maps for each year were preliminarily clipped on the territory of Tennessee, in a GIS environment. Because our project was conducted separately from the USGS (in contrast with the original intentions) and they have not yet defined a case-study subset of Tennessee urban watersheds, we considered all basins at the HUC-12 scale (see Fig. 4.5) from the Watershed Boundary Dataset (WBD) defined by the USGS (Jones et al., 2022). The state of Tennessee contains 1147 HUC-12 watersheds overall. From these, we identified all the urbanized basins, considering a threshold of IA of 10% by year 2021 (i.e., the last year with LC data available at the time this report was written), and studied their urbanization trends in the previous 20-year period, from 2001 to 2021.

Chapter 4 Results and Discussion

This chapter presents the deliverables that were agreed upon for this project in the proposal, lays out the main results and findings of the research as related to the different tasks, sub-tasks, and corresponding deliverables, and lists which sub-tasks were either not pursued or were pursued but only in a preliminary fashion, explaining the reasons. Finally, we briefly propose additional research ideas directly ensuing from the present findings that could be conducted in the future, considering their high potential impact on current engineering practice.

Again, as explained in Chapter 1, UofM's broader mandate included researching novel ideas with the potential for transformative or paradigmatic change in our understanding of urban peak flows and their prediction for engineering purposes, with a secondary focus on Tennessee and the Southeastern U.S. Later, the USGS ended up being fully tasked with generating the updated set of peak flow equations, so that some of the UofM sub-tasks and deliverables that were initially included in the proposal became moot, as they would not be needed by the USGS for their work.

Moreover, our preliminary research results strongly suggested that some of the initial ideas we proposed (mainly the new urbanization index and some of the aspects related to extreme rainfall analysis) showed more prospects for successful application to engineering practice. Accordingly, as was periodically informed through the Quarterly Progress Reports for the project, more effort was placed on these promising ideas than originally accounted for in the proposal, which resulted in decreasing the time allocated to some of the other sub-tasks that were initially considered.

Table 4.1 provides a list of the research deliverables of this project, according to the original, approved proposal, with comments and a summary of important findings.

Table 4.1. List of deliverables of the project according to the original proposal.

Task #	*Deliverable*	*Key Findings and Notes*
1	Database compiling all publicly available streamflow and stage data for urban streams in TN and surrounding areas of neighboring states	• USGS task • The notion of using stage data together with rainfall records for documenting basin lag was not pursued further
2	Trend analysis of extreme point rainfall, for different durations, at stations in (and near) TN	• After noting quality issues with the 15-min COOP rainfall data, we only performed a preliminary analysis of trends in extreme rain frequency, for the Southeastern U.S., presented in Section 4.1 • For all durations (except 24 h), many stations show an increased frequency of extreme rainfall events; for durations ≤ 6 h, 18 to 27% of the gauges display a significant increase in numbers of extreme storms • Regarding the two other issues that we studied, related to extreme rainfall, we find that: (i) Low rain gauge densities seem to cause a slight negative bias when estimating DDF values; more importantly, they do result in much increased uncertainty; (ii) Totalizing rain every 15 min (as with COOP gauges) causes a mean SAF slightly above 1.15 when estimating the 15-min DDF values, but there is

		high variability: the 75th percentile of SAFs is almost 1.18 (7.2% more than the 1.10 fixed value used in Atlas 14), the 95th percentile is 1.22, and many stations display SAFs above 1.20, up to 1.30. There is less bias when estimating the 30-min DDF values with 15-min data (sampling ratio = 2): the mean SAF is 1.04, the 75th percentile 1.05, and the 95th percentile 1.06, which compare well with the 1.05 value used in Atlas 14, but again this varies across gauges: many stations have SAFs above 1.07 or 1.08, with a maximum of 1.10. Results are presented in Section 4.1.
3	Analysis of temporal progression in urbanization trends across TN	• For most urban basins in Tennessee, we observed clear increases in the percentage of IA, with rates ranging from less than +0.1% to up to +0.7% per year, in the period from 2001 to 2021. More details in Section 4.2.
4	Detailed analysis of urbanization effects on flood regimes, at the scale of the rainfall-flood event, across TN and neighboring urban areas	• We propose a new, connectivity-based descriptor of urbanization, UI, which outperforms the traditional % impervious area, IA. This lumped watershed-scale descriptor of the hydrologic effects of urbanization shows promise for inclusion as an explanatory variable in regression equations for urban peak flows. See Section 4.3. • We derive a new methodological approach for studying the effects of changes in land-cover net of precipitation trends using LSTM neural network models. Details in Section 4.4.
5	A final report documenting all findings	• Present document
6	Implementation of regression equations and geospatial data required to calculate basin characteristics into the StreamStats web application.	• USGS task. • Our new connectivity-based urbanization index UI (Section 4.3) can support the development of new regression equations. • We are in continuous contact with the USGS StreamStats team, discussing the best ways to implement the urbanization index.

4.1 Trends in the frequency of extreme precipitation

Fig. 4.1 depicts the numbers of rain gauges in the Southeastern U.S. associated with different kinds of trends in extreme rainfall frequency, for each duration of interest, as obtained from applying the Mann-Kendall (M-K) test; Fig. 4.2 shows the same results focusing on Tennessee gages. We only display M-K test results, as they are very similar to those derived with OLS.

It can be seen that the number of gauges with increasing trends (both increasing or significantly increasing) is about 10 times larger than that displaying decreasing trends. Out of 473 stations, 36 % show increasing trends in frequency for the 15-min event, a value that gradually decreases to 14% for the 24-h duration. This suggests that the frequency of short-duration maxima is increasing at more sites than that of longer-duration events. The number of stations with a decreasing or significant decreasing trend is small, never exceeding 4.0% of the gauges, and shows no effect of duration. Again, given the data quality issues, we did not analyze trends in the magnitude of extreme rainfall, even though this is highly relevant for practice.

A careful look into the data for durations larger than 9 hours reveals that most of the values in the series are 0, especially for 18 and 24 h. This is because long storms are much rarer than short-duration events in this part of the U.S.

Because of this, our series for durations > 9 h were not long enough to be able to capture any trend in the frequency of events exceeding the 90th percentile maxima. This implies that results from Figures 4.1 and 4.2 for durations larger than 9 h are less credible.

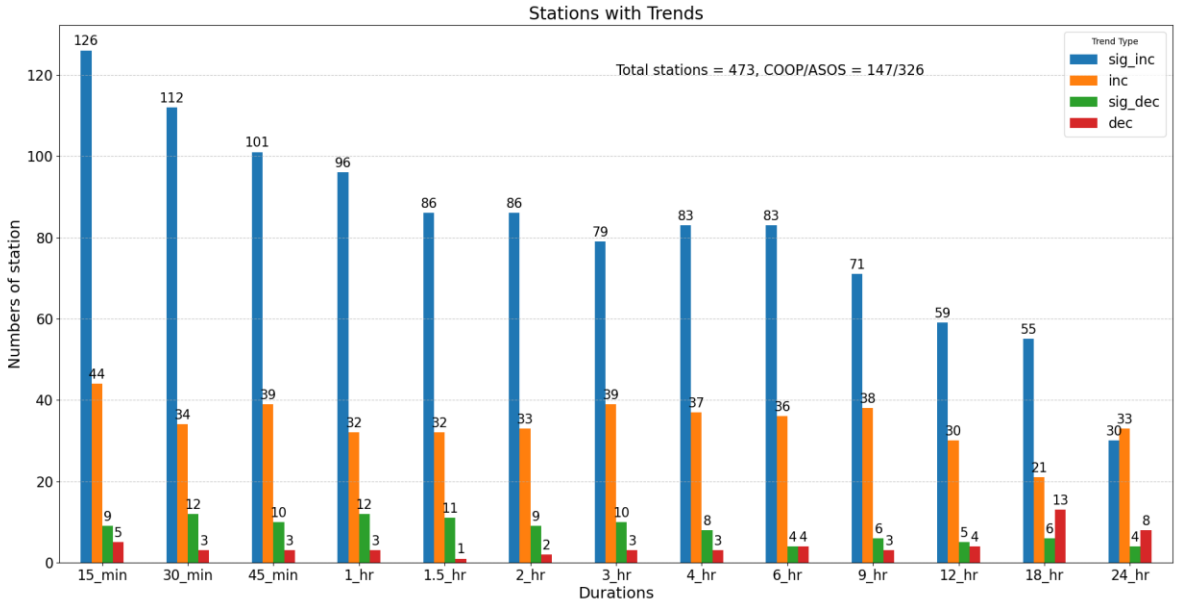


Figure 4.1. Trends in extreme event frequency from M-K test at 473 Southeast U.S. rain gauges for both ASOS and COOP networks. For each duration, the blue bar shows the number of stations with significant increasing trend, the orange bar depicts stations with increasing trend, the green bar shows decreasing trends, and the red bar shows those stations with decreasing trend.

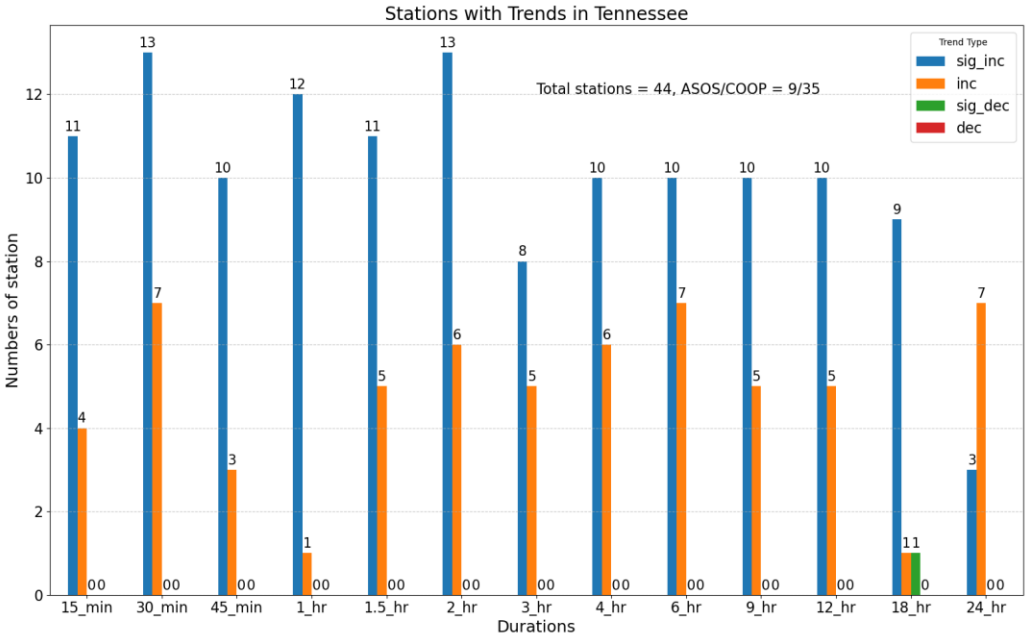


Figure 4.2. Trends in extreme rainfall frequency at both ASOS and COOP networks for Tennessee, from M-K tests.

For each duration, the blue bar shows the number of stations with significant increasing trends, the orange bar depicts gauges with increasing trend, and the green bar shows those stations with significant decreasing trend.

In Tennessee, Fig. 4.2 highlights that among the 44 gauges in the state (9 ASOS/35 COOP), only a single one exhibits a significant negative trend in extreme rainfall frequency, for a duration of 18 h. However, this observation is disregarded due to earlier considerations about the lack of data for durations > 9 h. Overall, 35-45% of the stations exhibit a consistent increasing trend across various durations. Consequently, we can assert with high confidence that the frequency of extreme precipitation is increasing in Tennessee, at least for durations ≤ 9 h.

In turn, Fig. 4.3 depicts the spatial distribution of these trends across the Southeastern U.S., for the shorter durations relevant to urban hydrology (and for which trends were assessed with more power, as explained above). Most gauges with decreasing trends occur in Louisiana and Southern Arkansas. The stations with either no trend or with increasing trend follow a similar spatial pattern across durations. They are more numerous and more randomly distributed than those with decreasing trends, but there is still a noticeable spatial trend towards more significant increases in frequency as one moves inland, which should be a cause of concern.

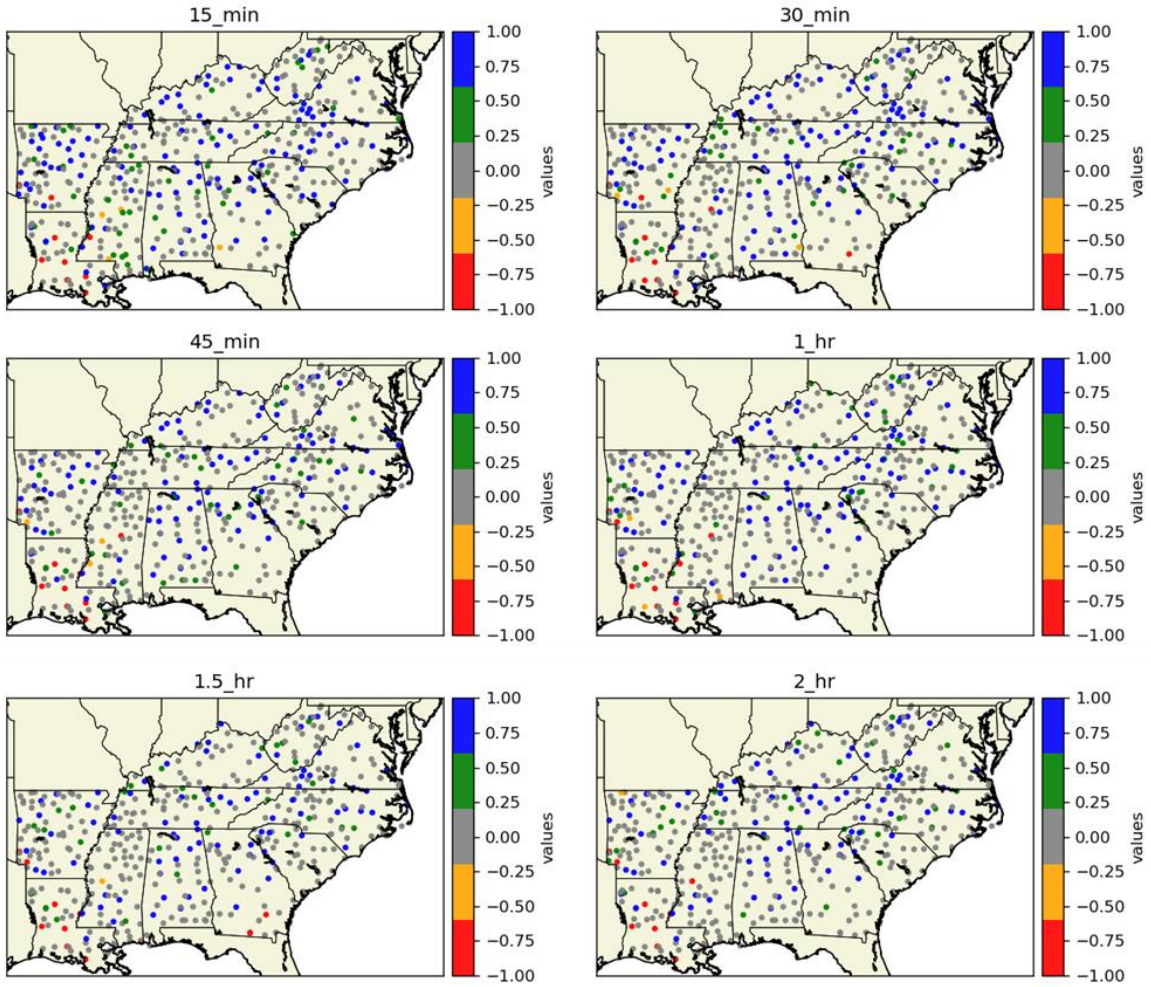


Figure 4.3. Maps depicting spatial trends in extreme precipitation frequency at 473 stations in the Southeastern U.S., for different durations. The color scale is as follows: significantly increasing = blue, increasing = green, no trend = grey, decreasing = orange, significantly decreasing = red.

The spatial distribution of trends in the frequency of extreme storms in Tennessee is depicted in Fig. 4.4. Again, across the 572 combinations of durations (13) and stations (44), the vast majority either reflects no trend (65%) or else increasing trends (34%), with no clear spatial pattern. Only one gauge shows significant decrease, for a single long duration of 18 h.

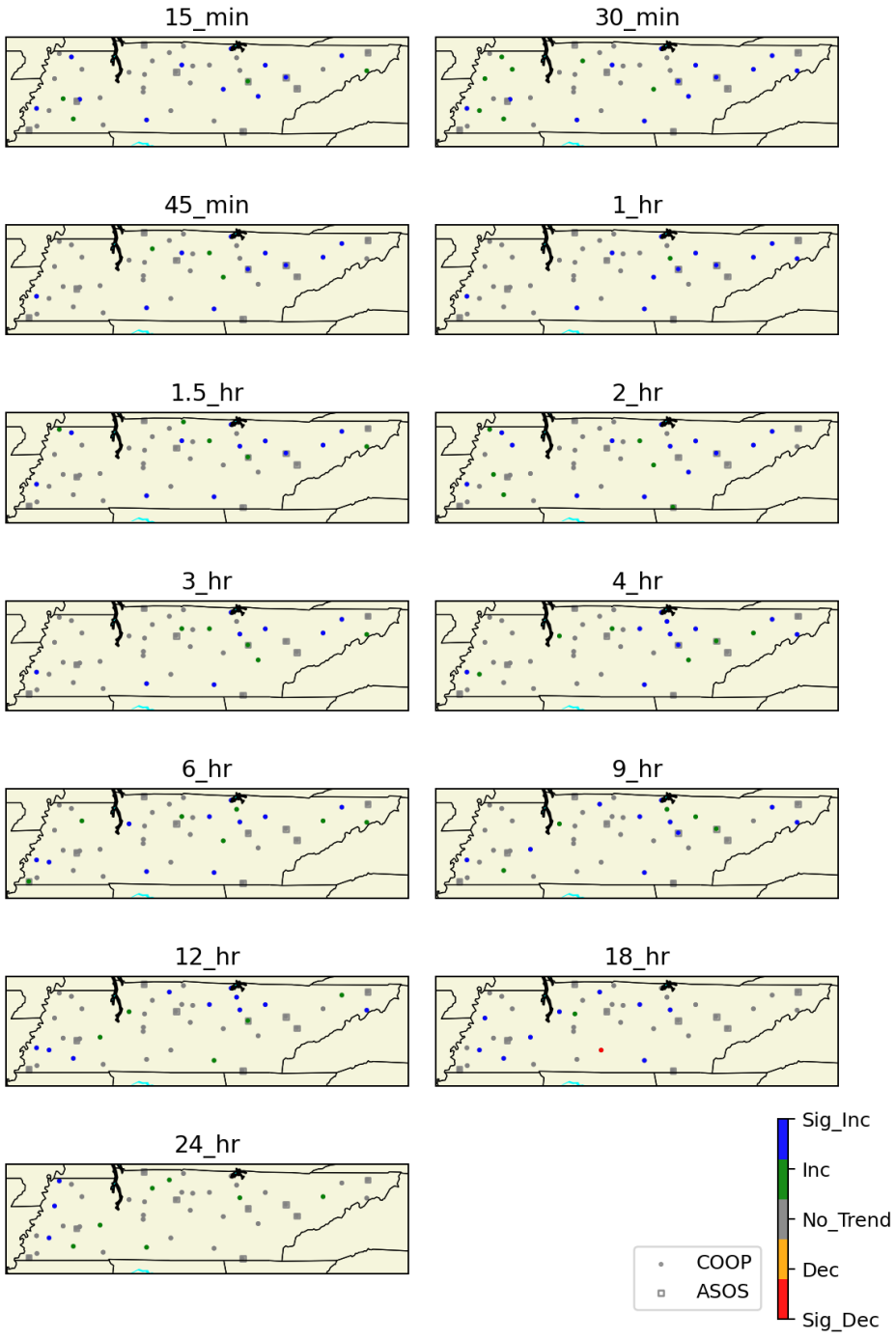


Figure 4.4. Map depicting spatial trends in extreme rainfall frequency at 44 gauges in Tennessee, for a range of durations from 15 min to 24 h. The color scale is as follows: significantly increasing = blue, increasing = green, no trend = grey, decreasing = orange, significantly decreasing = red.

4.2 Urbanization trends across Tennessee

Fig. 4.5 depicts Tennessee with its 1147 HUC-12 basins (Jones et al., 2022). We highlighted all the urbanized watersheds (using the criterion outlined in Section 3.7) using a red boundary. These are mostly clustered at a limited number of locations across the state. The largest groups correspond to the major metropolitan areas of Memphis, Nashville, and Knoxville, moving from West to East. Besides these major clusters, there are a few smaller ones mostly along the state boundaries, associated with the cities of Clarksville, Chattanooga, Kingsport, and Johnson City. There are also a few isolated urban watersheds, e.g., the city of Jackson, as well as the cities of Cookeville, Crossville, and McMinnville, between Nashville and Knoxville.

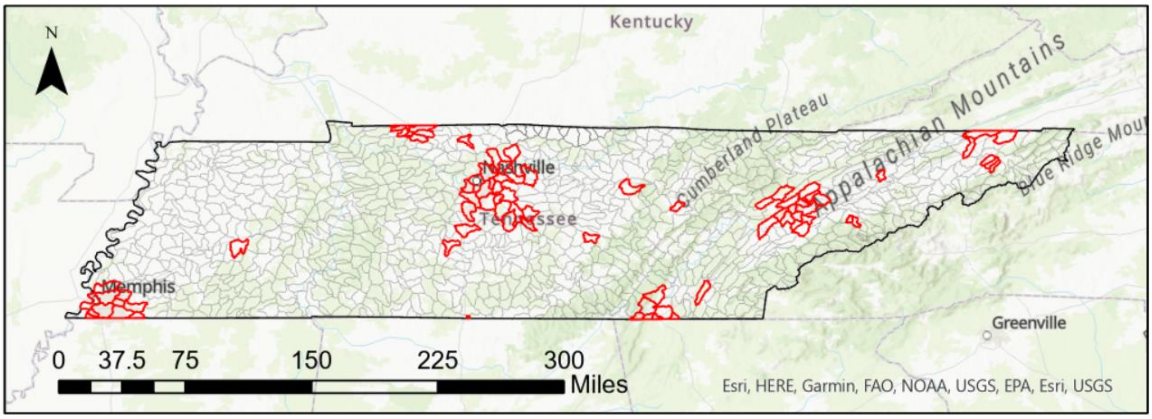


Figure 4.5. HUC-12 basins in TN, from the Watershed Boundary Dataset (Jones et al., 2022). Urbanized watersheds (i.e., those with ***IA*** above 10% in 2021), are marked with red, thicker boundaries.

Fig. 4.6 and Table A1 in the Appendix show the urbanization trends for all developed Tennessee watersheds. Most basins experienced a significant increase in their percentages of impervious surfaces (***IA***), in the 20-year period from 2001 to 2021, with an average overall increase of +6% (which corresponds to an average growth rate of 0.3%/year). However, it is important to note that there is a large heterogeneity in the initial (2001) conditions; most basins started with ***IA*** below 20%, only a few between 30% and 40%, and just two were already above 40% in 2001. By 2021, the number of watersheds with ***IA*** > 40% had doubled. It is interesting to note that no urban watershed in Tennessee shows a decrease or stagnation in ***IA***.

Most developed basins in Tennessee experienced a significant increase in their percentages of impervious surfaces, in the 20-year period from 2001 to 2021.

Fig. 4.7 shows the time-averaged changes in urbanization (as a rate in ***IA*** per year) for each developed basin, over the period 2001-2021. The observed average rates of increase in the level of urbanization in Tennessee span one order of magnitude, from minima below +0.1%/year (i.e., about +2% ***IA*** over the complete period of observation) up to about +0.7%/year (i.e.,

about +13% in the 20-year observation period). Many of the basins with the largest changes are located away from the center of the metropolitan areas, i.e., they correspond to suburban areas that have been experiencing significant urban expansion lately.

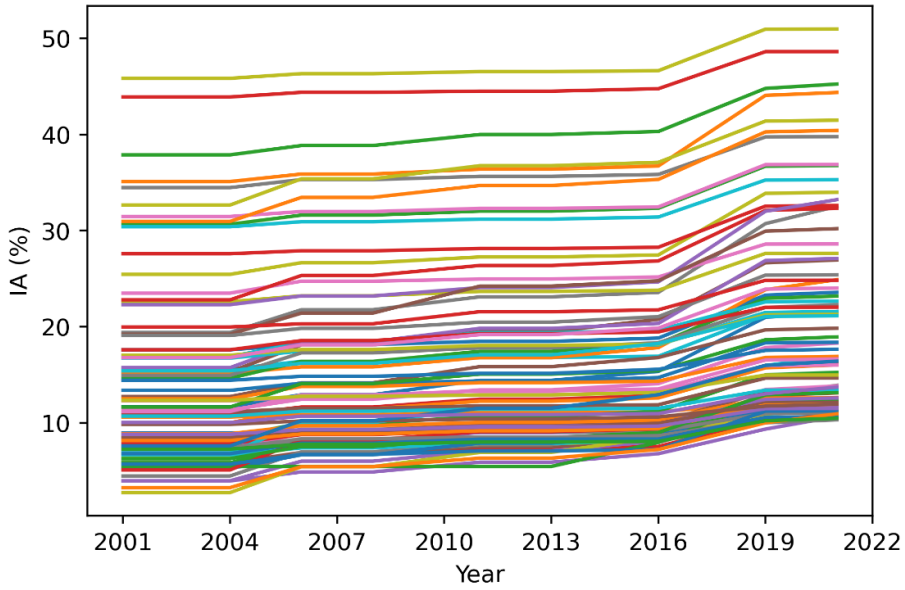


Figure 4.6. Urbanization trends in HUC-12 basins in Tennessee with ***IA*** above 10% by 2021.

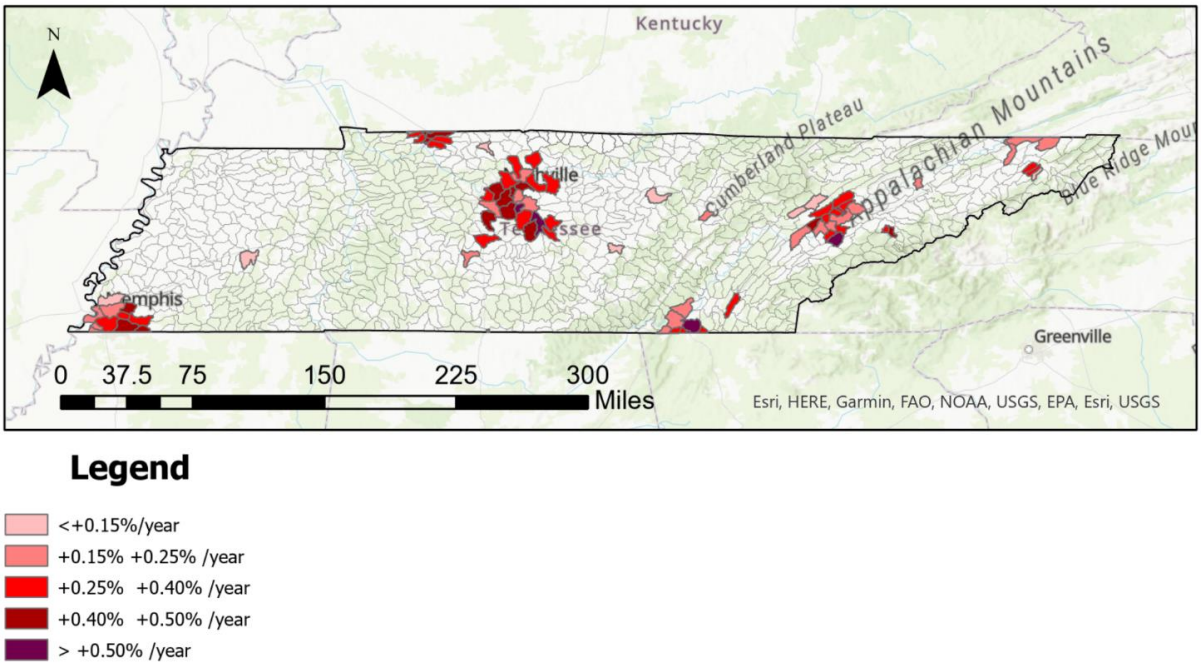


Figure 4.7. Annual rate of increase in the level of urbanization in developed Tennessee watersheds from Fig. 4.5, expressed in increase of the percentage of ***IA*** per year.

4.3 Benchmarking the new connectivity-based urbanization index

We test the new connectivity-based index of urbanization $\boldsymbol{UI}$ (Eq. 3.6) for two different expressions of the weighting coefficient $\boldsymbol{W}$ in the $\boldsymbol{HCI}$ formulation (Eq. 3.5), i.e., $\boldsymbol{W = 1 - n}$ (from existing literature; see Chapter 2) and $\boldsymbol{W = CN/100}$ (proposed in this research work; Eq. 3.1). We also consider a hybrid formulation where the weighting coefficient of the upslope component $\boldsymbol{D_{up}}$ is obtained as $\boldsymbol{W = CN/100}$, while the Manning-based coefficient $\boldsymbol{W = 1 - n}$ is adopted for the downslope component $\boldsymbol{D_{dn}}$.

We consider two alternatives for the weighting function $\boldsymbol{f_w(d_{j,k})}$ needed for deriving the weighted-average terms $\overline{\boldsymbol{W}}_{\boldsymbol{k}}^{*}$ and $\overline{\boldsymbol{S}}_{\boldsymbol{k}}^{*}$ in the upslope component $\boldsymbol{D_{up}}$ (Eqs. 3.3 and 3.4), outlined in Fig. 4.8. Table 4.2 shows the labels used for each combination of W and weighting $\boldsymbol{f_w(d_{j,k})}$.

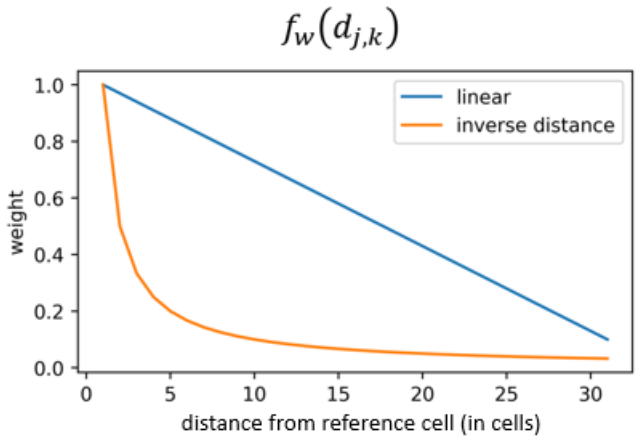


Figure 4.8. Two different weighting functions for the upstream cells when computing the upslope component D_{up}. Cells further upstream receive smaller weights.

Table 4.2. Labels for experiment identification, for different combinations of W and $f_w(d_{j,k})$.

		Weighting coefficient W		
		$W = 1 - n$	$W = CN/100$	$W = CN/100$ for D_{up} $W = 1 - n$ for D_{dn}
Weighting function $f_w(d_{j,k})$	Linear	n_Lin	CN_Lin	$CNhybr\_Lin$
	Inv. Dist.	n_Inv	CN_Inv	$CNhybr\_Inv$

Fig. 4.9 compares the predictive power of bivariate models regressed on $\boldsymbol{A}$ (basin area) and each one of the connectivity-based indices of urbanization $\boldsymbol{UI}$ (six cases shown in Table 4.2), obtained by applying Eq. (3.6) from the different expressions of $\boldsymbol{HCI}$ listed in Table 4.2, against (i) a simple univariate model regressed on basin area $\boldsymbol{A}$ only, and (ii) the state-of-the-art bivariate model in Missouri (Southard, 2010), fitted on $\boldsymbol{A}$ and $\boldsymbol{IA}$ (i.e., the % of impervious areas), considered as the benchmark. All models predict the peak flows $\boldsymbol{Q_T}$ with return period $\boldsymbol{T}$ of 2, 5, 10, 25, 50, 100, or 500 years, for the Missouri case study (Southard, 2010).

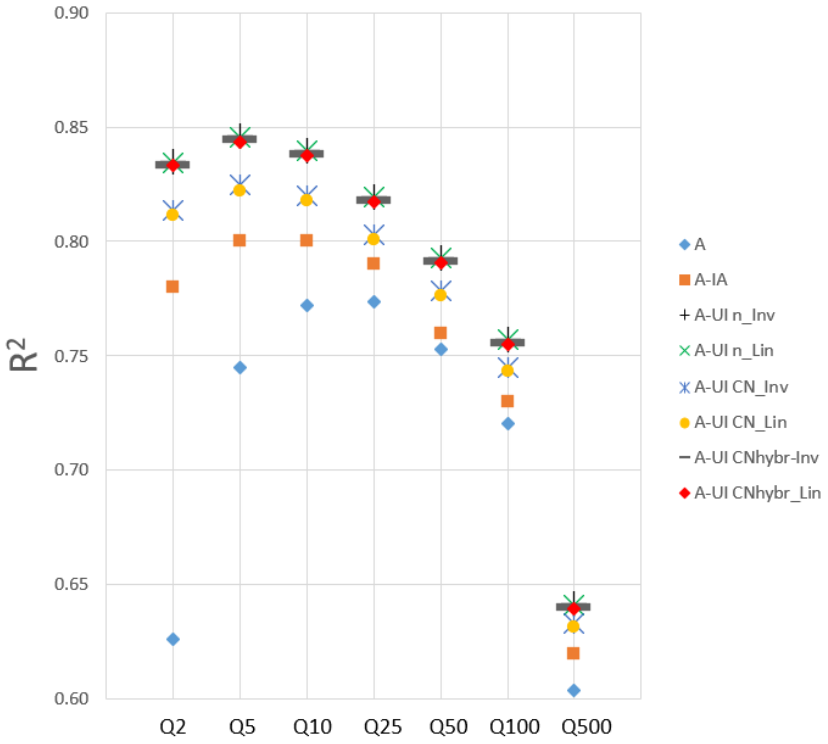


Figure 4.9. R^2 obtained when fitting regression Eq. (3.8) to A and IA, and to A and UI, using the full dataset, for different combinations of the weighting coefficient W and weighting function $f_w(d_{j,k})$.

We notice that all models that consider $\boldsymbol{UI}$ instead of $\boldsymbol{IA}$ perform better than the state-of-the-art model by Southard (2010), which uses $\boldsymbol{A}$ and $\boldsymbol{IA}$ as explanatory variables, highlighting the superiority of connectivity-based descriptors to fit peak-flow regression equations over traditional, lumped ones. This is also the case when predicting the most extreme quantiles (i.e., $\boldsymbol{Q_{50}}$, $\boldsymbol{Q_{100}}$, and $\boldsymbol{Q_{500}}$), for which the dependence on the level of urbanization seems to gradually vanish with increasing return period $\boldsymbol{T}$, as indicated by the convergence of the performances of the benchmark models with (i) only $\boldsymbol{A}$ or (ii) both $\boldsymbol{A}$ and $\boldsymbol{IA}$ as explanatory variables.

We observe a stronger sensitivity to the choice of weighting coefficient (i.e., $\boldsymbol{n}$- or $\boldsymbol{CN}$-based), as compared to the weighting function $\boldsymbol{f_w(d_{j,k})}$. E.g., CN_Inv and CN_Lin have similar performances, as well as n_Inv and n_Lin, or $CNhybr\_Inv$ and $CNhybr\_Lin$. However, the two models that use the n-based UI (i.e., n_Inv and n_Lin) perform better than their CN-based counterparts CN_Inv

and CN_Lin. The hybrid models $CNhybr\_Inv$ and $CNhybr\_Lin$ show similar performances to those of n_Inv and n_Lin.

We also tested the predictive power of $\boldsymbol{UI}$ against $\boldsymbol{IA}$ in a k-fold validation framework. As the purpose of peak-flow prediction equations is the application to ungauged basins, that are not considered when fitting the model to the available data, the average performance on this test set should be a better indicator of the usefulness of our approach to regression equations. A high performance on the test set also ensures that the model could generalize a relationship between input and output variables, without overfitting the original dataset.

… all the models that consider [the new urbanization index] $\boldsymbol{UI}$ instead of [the traditional % of impervious areas] $\boldsymbol{IA}$ perform better than the state-of-the-art model … highlighting the superiority of connectivity-based descriptors to fit peak-flow regression equations, over traditional, lumped ones.

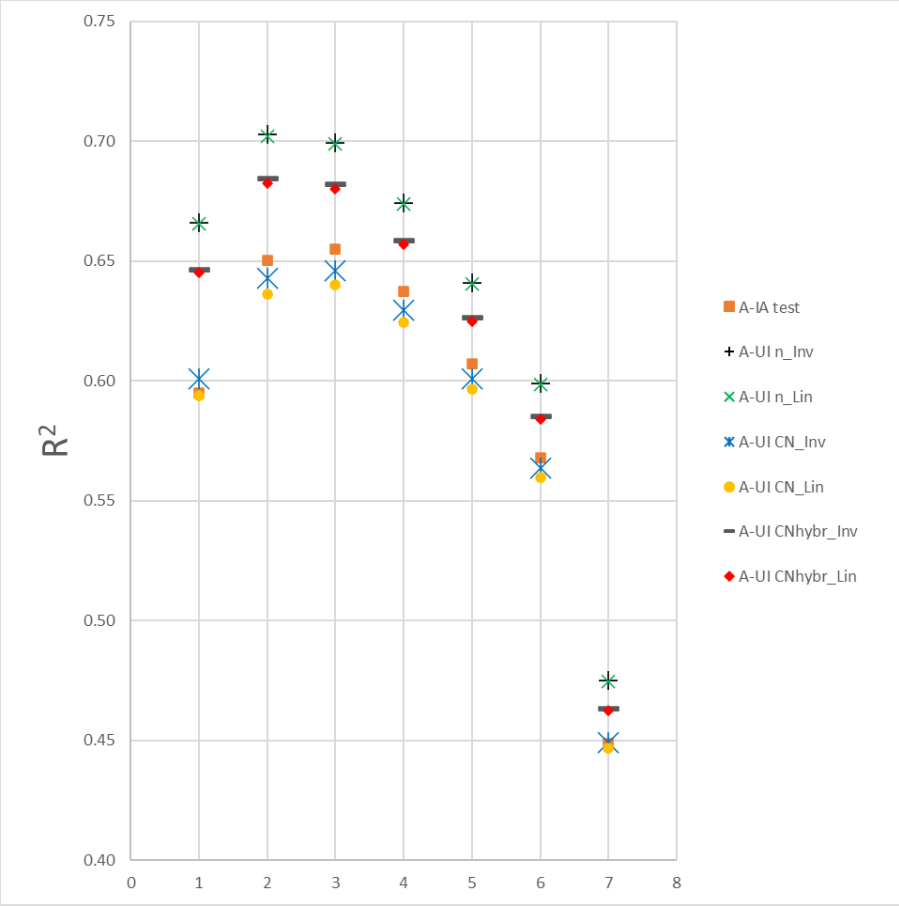


Figure 4.10. R^2 values obtained fitting Eq. (3.8) to $\boldsymbol{A}$ and $\boldsymbol{IA}$, and to $\boldsymbol{A}$ and $\boldsymbol{UI}$ on the test set in a 5-fold validation framework, for different configurations of the weighting coefficient $\boldsymbol{W}$ and the weighting function $\boldsymbol{f_w(d_{j,k})}$.

While Fig. 4.9 shows the R^2 obtained by the models when fitted on the full dataset, Fig. 4.10 reports the average performance on the test sets, highlighting a different picture as compared to the conclusions drawn from looking at Fig. 4.9 only. In Fig. 4.10, the benchmark model with $\boldsymbol{A}$ and $\boldsymbol{IA}$ shows a qualitatively similar performance as in Fig. 4.9, with a peak for intermediate quantiles. However, lower R^2 values are achieved, as expected on a test set.

Some of the competing models using $\boldsymbol{UI}$ as the proxy for urbanization outperform the benchmark, while others slightly underperform. Specifically, both models with a $\boldsymbol{CN}$-based weighting factor are on average less skilled than the benchmark. On the other hand, the n_Inv and n_Lin models both beat the benchmark in predicting all $\boldsymbol{Q_T}$ values. The $CNhybr\_Inv$ and $CNhybr\_Lin$ formulations also systematically outperform the benchmark (on average) on the test sets, with a score in between the models with $\boldsymbol{CN}$- and $\boldsymbol{n}$-based weighting coefficients. In general, model performance shows more sensitivity to the formulation of the weighting coefficient $\boldsymbol{W}$ rather than to the weighting function $\boldsymbol{f_w(d_{j,k})}$.

For the two most extreme quantiles, i.e., $\boldsymbol{Q_{100}}$ and $\boldsymbol{Q_{500}}$, the performance of all models decreases, even though the two models with $\boldsymbol{n}$-based weighting factor and the hybrid ones still outperform the benchmark. For those extreme floods, basin size becomes the preponderant variable, while the effects of urbanization tend to become negligible, as indicated by the convergence of performances of the univariate ($\boldsymbol{A}$) and bivariate ($\boldsymbol{A\text{-}IA}$) models (Fig. 4.9). Besides this, it should be considered that there are inevitably huge uncertainties in estimating extreme quantiles from short flow records, which means that the models are fitted on highly uncertain data points. In conclusion, improvements in prediction accuracy should be pursued in the range of more frequent floods (below 100-year return period).

4.4 Hierarchical contributions of different basin portions to discharge

We explored the potential of LSTM networks for modeling the rainfall-runoff conversion process across a range of basin sizes and ways of structuring the precipitation input (Dell'Aira et al., 2022). For this research's purposes, our main finding is that LSTM models do benefit from a spatially distributed representation of the rainfall input. Hence, this kind of model allows us to investigate what portions of the watershed are most relevant in determining its response to precipitation inputs: They will be those that lead to the best model performance when predicting outlet discharge based on their localized precipitation inputs by, e.g., minimizing for the number of inputs while maximizing for model performance. See, e.g., Dell'Aira et al. (2023) for implementation details of optimization as an investigation tool.

Dell'Aira et al. (2022) show a preliminary application of LSTM rainfall-runoff model training using an incomplete basin representation. Even though only the precipitation at the lowermost quarter of the basin (closest to its outlet) was considered for calibration, the model could match the performance of other LSTM networks trained with the complete rainfall information. It is worth noting that the size of the basin is more than 5000 km^2; thus, there is only low correlation among rainfall inputs from distant portions of the watershed. This indicates that the high performance of the model trained with incomplete localized rainfall series is not imputable to the fact that, e.g., the missing information from the other, neglected portions is highly correlated (and thus redundant) with the precipitation signals that were considered as model input;

it rather confirms that different parts of the watershed have unequal importance in affecting its hydrologic response. For the case study basin analyzed by Dell'Aira et al. (2022), the lowermost quarter has the strongest influence, given its spatial proximity to the outlet.

To extend and generalize this application to distinguish between basin portions that are important, due to their location, from those that owe their strong influence mainly to their LC types, it may be useful to also analyze the average connectivity of those locations, using HCI raster maps derived for the studied basin (see Section 3.3 about the concept of HCI). However, this application has not been completed yet. Future steps should involve the development of a systematic procedure to study the influence of time-varying urbanization levels on the hydrologic response, from cross-analyses of basin connectivity and the portions of the basins whose precipitation signals have the strongest effects on outlet discharge.

4.5 Uncertainties and biases in extreme precipitation

We present the results from our analyses on high-resolution (1-min), high-density rainfall data from Germany, in an attempt at understanding what types and levels of uncertainty and biases we could be having in the Southeastern part of the U.S., and Tennessee specifically, with our COOP 15-min gauges, and with a weather station density that is about 17 times sparser. Again, the distribution of types of storms making up the list of extreme events will be different, as there are no tropical storms nor cyclones in Germany. Moreover, Central Europe does not see as many of the typical isolated, slow-moving, summer-season thunderstorms. Extreme thunderstorms, whether associated to frontal systems or not, are also more frequent in this part of the world. Future research on the effects of totalization could focus on the ASOS 1-min stations with longer records, but their density will always be too low to allow any meaningful study of the gauge density effects on extreme rainfall, justifying the use of the German data.

We first discuss our findings about the potential effects of using totalized (15-min) data, and then present our results related to the potential uncertainties or biases due to low gauge density.

Biases due to totalization

For the German data, when estimating the 15-min extreme rainfall for a sampling ratio (SR) of 1 (i.e., when using original gauge data totalized every 15 min, such as the COOP data), for an ARI of 2 years, we find a median SAF value slightly above 1.15. This is 4.5% more than the 1.10 fixed value used in Atlas 14, so our results suggest that we are consistently underestimating the 15-min rain. But this relatively small bias is not the main issue at hand; it is the large individual, and thus unpredictable, at-a-gauge variability about the mean. For example, the 75th percentile for the 15-min SAF with 15-min data, across all possible combinations of stations (862) and ways of totalizing (15, so a total of 12,930 different SAFs for this duration) has a value of almost 1.18 (7.3% higher than the fixed correction applied in Atlas 14), while the 90th and 95th percentiles amount to 1.20 and 1.22, respectively. Moreover, as can be seen in Fig. 4.11, there are many cases in which the SAFs can be larger than 1.20, up to an absolute maximum of 1.30.

This finding is sobering, because it indicates that at our existing 15-min gauges in the U.S., our current DDF estimates (for at-a-station frequency analyses) for 15-min rain, are already underestimated by 7, 10, or even 15% or more at a sizeable number of these gauges, in comparison to the DDF values we would have computed had we had rain gauges with 1-min resolution,

instead. It is also most probable that if we had such data, the spatial variability in rainfall maxima would be higher locally, which would in turn affect the regional frequency analyses.

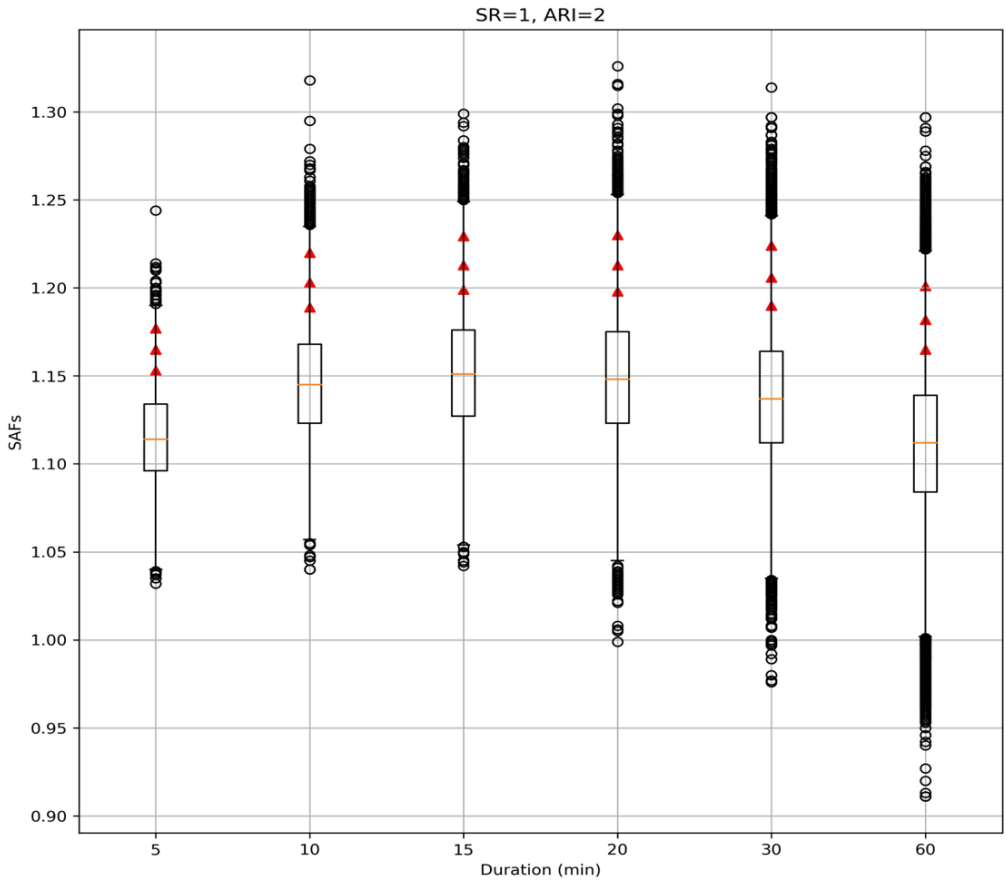


Figure 4.11. Individual SAFs for sampling ratio = 1, considering all stations, durations, and ways of totalizing. The boxplots show the statistics of individual SAFs, where the central red mark indicates the median, while the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers show the smallest and largest value in the sample, barring outliers, shown as black circles. The red triangles correspond to percentiles 90th, 95th, and 98th, in each case.

As expected, there is less bias when estimating the 30-min DDF values with 15-min data (SR = 2): in that case, the median SAF is 1.04, and the 75th percentile 1.05, while percentiles 90th and 95th are about 1.06. In the mean, these results compare well with the 1.05 value in Atlas 14, but again SAFs vary widely: many stations have SAFs above 1.07 or 1.08, with a maximum of 1.10.

In turn, Fig. 4.12 depicts the experimentally obtained mean at-a-station SAFs (i.e., averaged across all ways of totalizing, for each station) as a function of the sampling ratio; it can be seen that, on average, the effects of using totalized rainfall data become negligible (< 1%) for a sampling ratio larger than 6. For the U.S. context, this would mean that when we use the 15-min gauges to estimate the 90-min rainfall, the underestimation should be less than 1% on average. But of course, there will still be some stations with a larger negative bias for all durations under 90 minutes, due to the variability, as depicted by the red crosses (percentiles 95th, 98th, and 99th) in Fig. 4.12.

An interesting preliminary result (not shown here) is that mapping the mean SAFs for different durations reveals a clear spatial structure across Germany. We applied geostatistical tests which found a significant spatial correlation for about half of the durations.

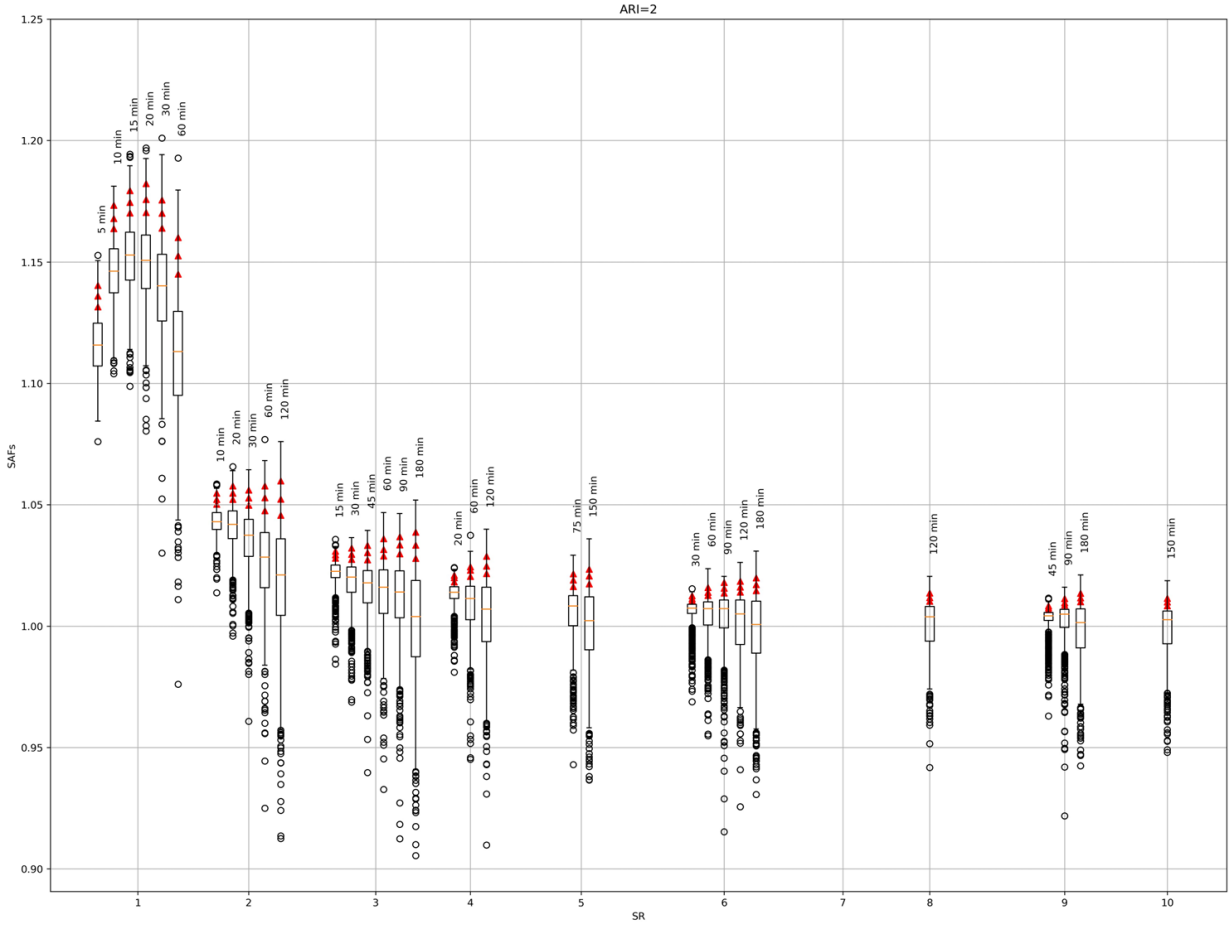


Figure 4.12. Distributions of at-a-station averaged SAFs, across all 862 stations and all possible ways of totalizing, for different durations, as a function of the sampling ratio SR (which ranges from 1 to 10). Boxplots show the statistics of these at-a-station mean SAFs, where the central red mark indicates the median, while the bottom and top edges of the box indicate the 25th & 75th percentiles, respectively, the whiskers show the smallest and largest value in the sample, barring outliers, shown as black circles, and the red triangles correspond to percentiles 90th, 95th, and 98th, in each case.

The conclusion from this detailed work is that mean SAFs are not independent with respect to the rainfall duration, within the range of shorter durations that we studied (up to 2 hours). The SAF values for 5 min are too low, as they are estimated from 1-min data which are not fine enough, but it is clear from the vast number of data incorporated in the distributions shown in Fig. 4.11 (e.g., for 30 minutes, there are 862 gauges times 30 different ways of totalizing for a total of 25,860 individual SAFs) that there are consistent trends in SAFs depending on duration. Again though, these effects are much less relevant, in engineering terms, than the problems due to the variability in the individual, at-a-station values of the SAFs. We suggest that the mean SAFs used in current DDF values (Atlas 14) are slightly biased, but most importantly, they do not reflect the actual variability across gauges as well as at gauges.

Effects of station density

To assess the potential effects of station density on regional DDF estimates, we computed at each station the percentual differences between DDF values simulated for a range of lower gauge densities and the "true" DDF values obtained for the full density case.

For each gauge's region of influence (ROI), we generated 100 simulated DDF values for each studied density, obtained by randomly resampling the gauges within the ROI. For this analysis, any difference exceeding ± 5% was deemed to be significant.

The plots in Fig. 4.13 show DDF values converging towards their "true" values, as gauge density increases. The panels on the left, for durations of 15 min and 2 h and ARI of 1 yr, depict many more cases of underestimation than those on the right, for ARI=25 yrs. Note that the ordinates represent the total number of simulations, i.e., they reflect the combined variability across stations and random realizations of lower density networks. We need to delve further into these results by analyzing what happens at individual rain gauges. Across both ARIs, the plots reveal a larger proportion of significant negative biases than significant positive ones. However, it is important to point out that the ratio of significant negative to significant positive biases decreases monotonically with increasing ARI. Interestingly, the plots across durations but for the same ARI are nearly identical, suggesting that duration only has negligible effects, at least for the shorter durations we are considering, relevant to urban hydrology.

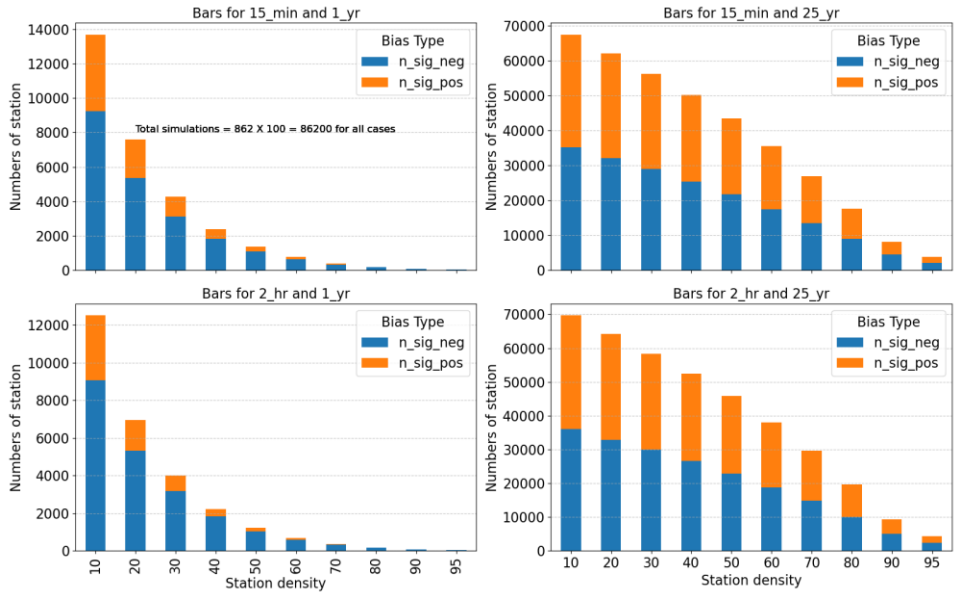


Figure 4.13. Bar plot for the number of significant positive differences (blue) and significant negative differences (orange) across all German stations, for durations of 15 min (above) and 2 h (below) and average recurrence intervals of 1 yr. (left) and 25 yrs. (right), as a function of station density in %.

Maybe more important than the possible biases introduced by the low density of stations that we happen to have in the Southeastern U.S. and Tennessee (which is about 6%, so even smaller that the worst case in Fig. 4.13), is the fact that the number of significant deviations (positive or

negative) quickly increases with decreasing density, tending to grow exponentially for the more frequent ARIs, but only linearly for the higher ARIs analyzed in this work.

In simpler words, DDFs estimated with regional frequency analyses from low-density networks are much more uncertain. We propose that this uncertainty is probably much larger than that reported in Atlas 14 by using the traditional methodology of resampling maxima at each gauge.

4.5 Underestimation of frequent floods in the United States

As outlined in Bulletin 17c (England et al., 2019), the current preferred approach for performing flood frequency analysis (FFA) in the U.S. is based on the use of annual maxima (AM) series, adopting the Log-Pearson III (LP3) distribution for modeling AM magnitudes.

However, AM frequency analysis (FA) is conceptually unsuitable for accurately predicting the average interarrival time (AIT) of frequent floods (FFs), i.e., those that occur on average once every five years or less (Ball et al. 2019), as discussed in Section 2.6. In other words, if one is interested in determining the $\boldsymbol{T}$-year event, AM-FA would likely underestimate the peak flow for $\boldsymbol{T}$ < 5 years (Langbein 1949; Dell'Aira et al., 2023). The methodology outlined in Section 3.6 can quantify the amount of underestimation of the $\boldsymbol{T}$-year event. For example, Fig. 4.14 shows the underestimation in the 1.5-year flood at 432 watersheds distributed over the entire CONUS, determined by comparing the estimates obtained from AM- and POT-FFA (Eq. 3.9).

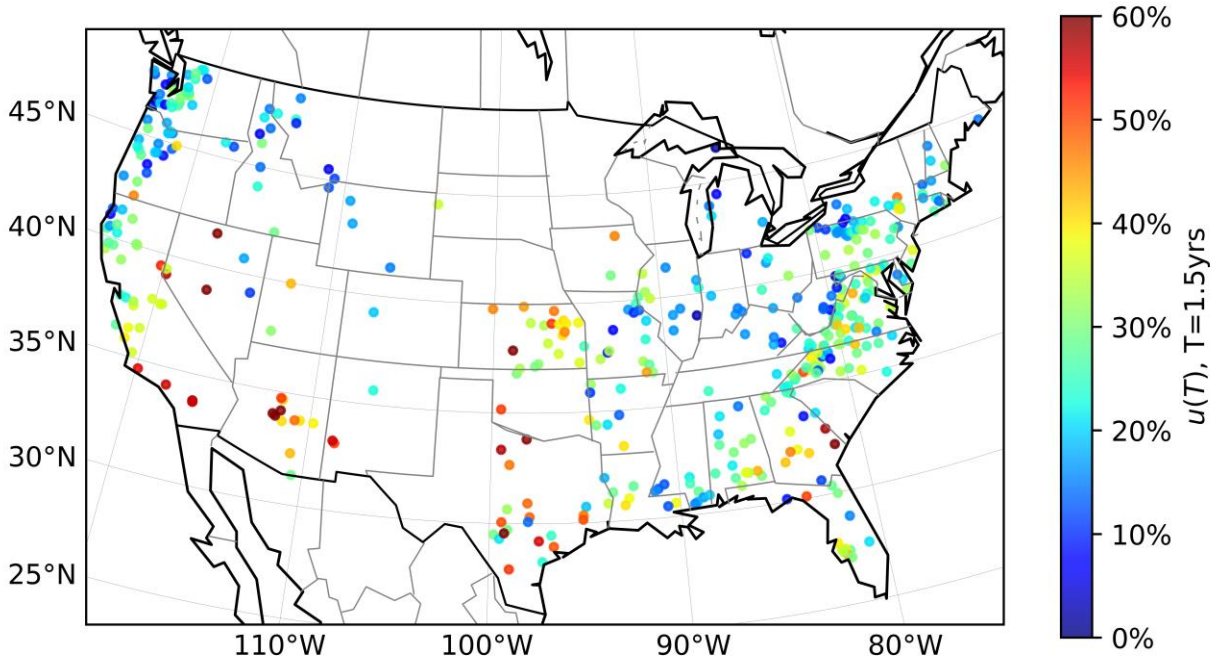


Figure 4.14. Spatial pattern of observed FF underestimation $\boldsymbol{u}(\boldsymbol{T})$ for the 1.5-year event.

The spatial patterns in flood underestimation revealed by Fig. 4.14 clearly suggest that there is a strong geographical dependence, while the values indicate that it can be severe in some regions, reaching magnitudes as high as 60%. Please note here that using the definitions in Equations 3.9 and 3.10, an underestimation of 60% actually means that the magnitude of the "true" 1.5-year flood, as obtained from partial duration analysis, will be 2.5 times larger than

that estimated using annual maxima. Specifically, areas with arid and semi-arid climates present clusters of watersheds with the highest underestimation rates across the nation, so that a preliminary suggestion would be that hydrologists should be particularly careful about their choice of Flood Frequency Analysis methodology (i.e., use Peaks over Thresholds instead of Annual Maxima) in these regions.

For the specific case of Tennessee, there are only two basins in the analysis, with moderate underestimation (in the order of 15 to 25%), but some watersheds in neighboring areas of surrounding states do show higher differences between AM- and POT-estimated frequent floods.

Fig. 4.15 depicts the underestimation for a range of average interarrival time (AIT) values from 1.5 to 5 years, considering both the observed (Eq. 3.9; shown on left panels) and the theoretical, duality-derived underestimation (Eq. 3.13; shown on right panels). The duality-derived underestimation $\hat{u}(\xi_g, \sigma_g, \mu_g, T)$ computed from the duality-derived quantile ratio $\hat{r}(\xi_g, \sigma_g, \mu_g, T)$ (Eq. 3.13) matches well the observed underestimation $u(T)$ (i.e., that obtained from Eq. (3.10) by computing the quantiles applying frequency analyses to both AM and POT series and then using the quantile ratio $r(T)$ from Eq. 3.9). It is clear that the spatial structures for the different AITs are fairly similar. This indicates that Eq. (3.14) provides an adequate tool for bias-correcting underestimated frequent flood magnitudes obtained using annual maxima. However, caution is advised when using Eq. (3.14) at locations with mixed flood populations, for the reasons discussed in Dell'Aira et al. (2023). Both the observed underestimation rate $u(T)$ and the duality-derived underestimation $\hat{u}(\xi_g, \sigma_g, \mu_g, T)$ decrease for increasingly larger return periods, as expected, since results from AM- and POT-frequency analysis methods tend to converge for larger return periods (Langbein, 1949).

As a concluding remark, it should be noted that the problem of underestimation due to using annual maxima in frequency analysis does not affect the estimation of large-return-period floods, typically applied in the design of major water structures and infrastructure. However, many environmental applications related, e.g., to river restoration (Wohl et al., 2015), do require accurate predictions of flood events that occur with relatively high frequency, due to their major impacts on stream geomorphology (Wolman & Miller, 1960) and ecology (Johnson et al., 1995; Meier, 2008; Wohl et al. 2015).

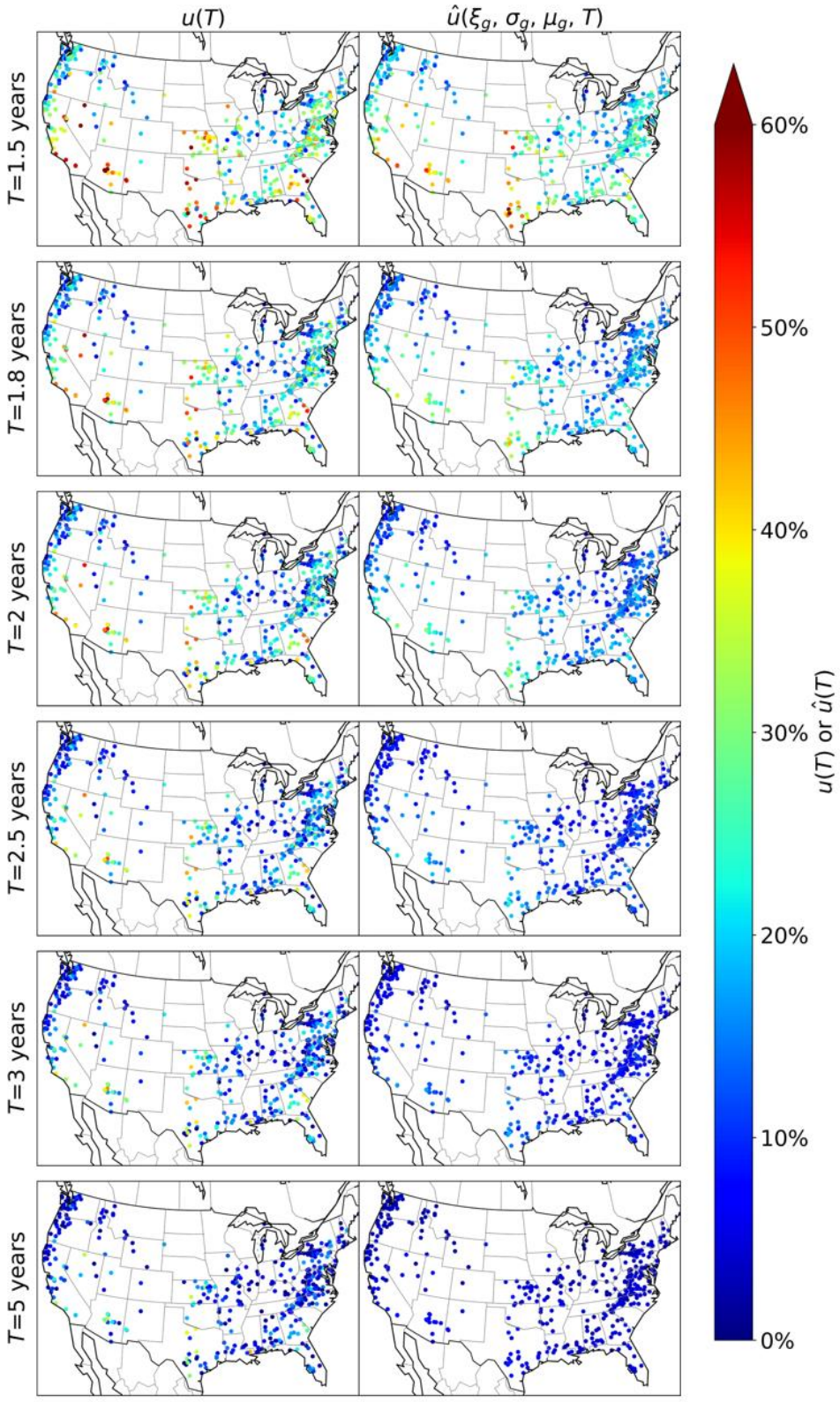


Figure 4.15. Spatial distribution of the observed underestimation $u(T)$, at left, and the duality-derived underestimation $\hat{u}(\xi_g, \sigma_g, \mu_g, T)$, at right, for a range of Ts.

Chapter 5 Conclusions

In this concluding chapter, we describe how our findings can benefit TDOT and, more generally, the practice of hydrologic engineering in the U.S. and beyond. We present these by subjectively ranking them in decreasing order of perceived importance and applicability. In each case we discuss potential applications of our findings, outline future steps for their implementation, and consider recommendations and promising future research lines. The last section provides a summary of our key recommendations.

5.1 Applications and future implementation of the connectivity-based urbanization index

As outlined in Section 4.2, the state of Tennessee has experienced significant growth in land development in those watersheds located within, or close to, major metropolitan areas, with strong local heterogeneities in the average annual increases in the percentage of impervious area. As discussed before, this is one of the reasons why current regional regression equations for peak-flow predictions in urban basins in Tennessee (and surrounding areas) necessitate major updates.

Our connectivity-based urbanization index ***UI***, described in Section 3.3 and benchmarked against an established regression model in Section 4.3, represents a better alternative than the percentage of impervious area (***IA***) as a descriptor for the level of urbanization at the basin scale, as it is able to capture information on the spatial distribution of urbanized areas within the basin and how this interacts with the hydrologic response. This feature of our proposed ***UI*** should help derive more robust regional models for peak-flow predictions, that can better adapt to changes in the level of urbanization over time.

We are currently communicating with the USGS StreamStats team towards the future implementation of the connectivity-based ***UI*** in regional peak-flow equations. To ensure applicability at the scale of entire states, computationally less-expensive formulations of the connectivity index are required. Our observations on the low sensitivity of the performance of the connectivity-based urbanization index suggests that this can be achieved by adopting weighting functions that require minimal number of iterative steps for computing ***UI***. Our preliminary results (not published in this report) indicate that this is a viable strategy.

The applicability of this novel procedure is not limited to urban watersheds in Tennessee; we are currently piloting its expansion to other candidate case studies, considering urbanized regions in Virginia (Austin, 2014), some major urban centers in Arizona (Kennedy & Paretti, 2014), and three large homogenous regions (based on physiographic considerations) across the states of Georgia and the Carolinas (Feaster et al., 2014). All the case-study regions contain watersheds with sizes spanning across three orders of magnitude, with the largest ones in the order of a few thousands of square kilometers. If our preliminary findings presented herein are substantiated, this research on urbanization indices would not only benefit TDOT, but all agencies charged with drainage design in urban areas, at all levels (federal, state, county, municipal).

In view of the future implementation of the connectivity-based ***UI*** in StreamStats, at the scale of whole states, in collaboration with the USGS, further parallelization of the program is advised from a computational standpoint. This is because the implementation in StreamStats will require the computation of connectivity raster maps for one or more states. We have identified Nvidia CUDA (Halfhill, 2008) as a good candidate programming platform to GPU-accelerate our program. Other authors (e.g., Qin & Zhan, 2012) have already successfully tested computationally demanding algorithms for raster operations in CUDA. An alternative (or synergic) approach for reducing the computational burden could be searching for specific formulations of the HCI that display lower computational demand, considering specific weighting functions that allow to neglect some of the otherwise required calculations. Our results suggest that finding weighting functions that require a convenient number of iterations, from a computational perspective, without losing too-much information and predictive power, should be a viable course of action.

As a final, future projection about the potential applications of the urbanization index, we suggest that subsequent improvements of the ***UI*** should also account for the presence of stormwater infrastructure in urban watersheds, as this represents an additional source of hydrologic connectivity. We recommend that further research efforts be performed in collaboration with TDOT and the USGS, and maybe other agencies, to facilitate acquiring detailed data on local drainage infrastructure, thus paving the way for such expansion of the connectivity-based ***UI***. Preliminary data that would be required involve the stormwater infrastructure for at least a few dozen urbanized basins. Such maps of stormwater infrastructure should preferably be in GIS format, for ensuring compatibility with the software that we have developed in the framework of the present project.

In Tennessee, not only are gauged urban watersheds scarce, but they are highly clustered in space, and in many cases do not coincide with the location of rain gauges. It will be difficult to improve our predictive tools for engineering hydrology if we do not increase the density and representativity of the hydrometeorological network (the number of stream gauging stations and rain gauges) in the state. This must be done in a concerted way, so as to ensure adequate spatial coverage across different urban areas in Tennessee, a range of variability in the properties of the gauged watersheds, and a concurrent measurement of the rainfall input, as recommended in Section 5.2, below.

5.2 Uncertainties, biases, and trends in extreme precipitation

Our findings demonstrate that it is highly probable that using the currently available rainfall data in Tennessee (and the Southeastern U.S. as a whole), with a sparse (low-density) network consisting mostly of 15-minute Fisher and Porter rain gauges, introduces negative biases as well as a large variability in our estimates of extreme precipitation; in other words, IDF-DDF values are underestimated and uncertain. It is clear that these issues should warrant further research.

We recommend that our methodological approaches, derived on German precipitation data, be applied to rain gauges in Tennessee and surrounding states. We see potential benefits in reproducing our study for those 1-min ASOS stations with longer records, to locally quantify the variability as well as the bias introduced by the use of 15-min rainfall data.

Given the nature of precipitation events in this part of the world, we hypothesize that the biases may be even larger than those observed in Germany.

An immediate practical recommendation from our findings is that given their inherent variability both across gauges and at a gauge, we should not use mean or median rainfall sampling adjustment factors in order to uniformly correct DDF values derived from fixed maxima. Instead, we should think about moving to a more conservative approach, maybe using the 75^{th} or the 90^{th} percentile of the SAFs derived at those fewer gauges with higher (1-min) resolution. A potential future research line would be to attempt to understand the spatial structure of SAFs, so that gauge location can be accounted for when correcting DDF values obtained from fixed, totalized data.

This research proposes a new urbanization index ***UI*** to better represent the effects of urban development on the hydrologic response of basins, thus helping tackle the non-stationarity issue in urban hydrology due to changing land-cover conditions. However, as explained before, this is not the sole source of non-stationarity in urban floods: trends in precipitation can also change the frequency and magnitude of large floods. Regarding the trends in extreme precipitation in the Southeast U.S., our preliminary findings of an increase in their frequency corroborate previous studies. In the case of Tennessee, we find that for durations below 9 hours, 30 to 45% of the rain gauges (depending on duration) show increasing or significantly increasing numbers of extreme events. These trends are slightly more evident in central and eastern Tennessee.

We did not study the trends in the magnitude of the events, because of lingering questions about the quality of the data from the 15-minute gauges, which make up all of the stations with longer records, that are needed to test for temporal trends. Future funded research could possibly attempt to "clean" and merge the COOP data from different sources, jointly working with NOAA, in order to analyze rainfall trends with different techniques, either more statistical (trend tests) or else more "event-based," using the information contained in the individual storms at each gauge location, as was initially proposed but finally not pursued within this project.

It will be difficult to locally document the effects of station density on DDF-IDF value estimation, given the low density of weather stations across Tennessee (and the U.S.). Still, we recommend further research into this topic, aiming at incorporating gauge density into our confidence estimates for DDF-IDF values. It should be clear that a higher density of stations results in many more closely located gauges, all sampling the same "extreme rainfall climate," which in turn should decrease the uncertainty around estimated extreme rainfall quantiles.

A tentative recommendation based on all of the above ideas, is that the DDF/IDF values in Atlas 14 need to be updated. Moreover, we recommend that state agencies increase the current low density of rain gauges as well as the temporal resolution of the current COOP gauges, considering 1-minute resolution equipment, so that in the future these topics can be pursued with more adequate data in terms of density and temporal resolution. If more stream gauging stations are installed in urban watersheds across the state, it will be important to concurrently measure precipitation in those basins so as to maximize benefits to designers and planners.

5.3 Revisiting current techniques for frequent flood estimation

Because performing frequency analyses with annual maxima underestimates peak floods for return periods below 5 years, we recommend that Bulletin 17c incorporates adjustments for frequent-flood prediction. Our proposed method is a valid candidate both for assessing the amount of potential underestimation at a given location, and correct biased frequent flood predictions. This is of concern in a range of river-related disciplines; for example, a stream restored considering a bankfull discharge that was underestimated because of this issue could end up flooding as much as three times more often than per-design conditions (Dell'Aira et al., 2023).

5.4 New approach to study urbanization effects on floods

Using an LSTM artificial neural network model to derive data-driven rainfall-runoff models, Dell'Aira et al. (2022) showed that given that specific parts within a given basin affect its hydrologic response more than others, accurate flow predictions can be still obtained considering the precipitation input only at those locations, neglecting rainfall over the remaining, less relevant portions of the basin.

We suggest that a future methodology to investigate the effects of urbanization changes can be developed based on this type of models, by searching for spatial correlations between areas with high rates of urbanization-induced connectivity and those portions of a watershed whose rainfall inputs have the highest predictive power. The experimental procedure should involve a range of watersheds with different sizes, to ensure generalization of the findings across a variety of hydrologic and geomorphic conditions.

The main advantage of the proposed procedure is that it implicitly excludes the effects of trends in precipitation, ultimately disentangling land-cover changes and precipitation trends, i.e., the two main sources of non-stationarity in a basin's hydrologic response (see Section 2.1).

5.4 Summary of key recommendations

- Further pursue the development of the urbanization index ***UI***, also including the effects of stormwater drainage infrastructure on hydrologic connectivity. The current ***UI*** described in this report only considers surface hydrologic connectivity, driven by relief and the nature of surface patches (their slope, roughness, infiltrability, etc.), but most urbanized and urbanizing areas contain stormwater sewerage systems, with their own, separate effects on hydrologic connectivity. Including information on both types of connectivity should enhance the explanatory power of ***UI*** in peak flow equations for urbanizing areas, further reducing estimation uncertainty.
- Simultaneously increase the number of gauged basins as well as the density of rain gauges across urban areas of Tennessee. The design for such enhanced hydrometeorological monitoring networks (sometimes known as "mesonets") should consider both a more distributed coverage across the different urban areas of the state, as well as a wider range of watershed characteristics. Rain gauging equipment for the monitored watersheds should be state-of-the-art, with 1-minute resolution, and a spatial density that ensures coverage of

extreme, small-scale convective events (i.e., thunderstorms). This would allow for future development of locally tailored peak flow equations, better estimation of IDF-DDF values, and improvements in methods for engineering design, that would increase urban resilience to extreme hydrometeorological events across Tennessee.

- Perform further, more detailed studies of: (i) how totalized, 15-min data underestimate actual maxima, (ii) what uncertainty or bias is introduced due to the current, low density of weather stations, (iii) trends in both frequency and magnitude of extreme precipitation, and (iv) how the development of novel, event-based rainfall analysis techniques can help extract more information from the sparse, short precipitation records available in Tennessee. Collectively, results from these studies would allow for updating IDF-DDF values in Tennessee, so that the present, higher values of rainfall are reflected in engineering designs. In this respect, it should be noted that there is increasing evidence, including results from this research, that the currently accepted extreme rainfall values for design, from NOAA's Atlas 14, are underestimated.
- Convey to practitioners that commonly performed flood frequency analyses based on annual maxima can severely underestimate frequent floods (say, those with return period ≤ 5 years), and thus should not be used in engineering applications requiring such estimates, like the determination of bankfull floods, or for river restoration projects. This would ensure that designs are not based on biased estimations, which increases their risk of failure.
- Pursue the study of machine learning models for explaining the hydrologic effects of urbanization, disentangling trends in land cover from those in precipitation. We think that this is the next frontier for meaningful improvement in hydrologic estimates needed for design.

References

Abrahart, R., See, L., and Kneale, P. 1999. Applying saliency analysis to neural network rainfall-runoff modelling. Proceedings of the *4th International Conference on Geocomputation,* Fredericksburg, Virginia, USA, 25-28 July.

Aqil, M., Kita, I., Yano, A., and Nishiyama, S. 2007. Neural networks for real time catchment flow modeling and prediction. *Water Resources Management*, 21(10), 1781–1796.

Austin, S.H. 2014. Methods and equations for estimating peak streamflow per square mile in Virginia's urban basins. *U.S. Geological Survey Scientific Investigations Report* 2014–5090. http://dx.doi.org/10.3133/sir20145090.

Barbero, R., Fowler, H.J., Blenkinsop, S., Westra, S., et al., 2019: A synthesis of hourly and daily precipitation extremes in different climatic regions. *Weather and Climate Extremes*, 26, 100219.

Ball, J., Babister, M., Nathan, R., Weeks,W., Weinmann, E., Retallick, M., Testoni, I. 2019. *Australian Rainfall and Runoff: A Guide to Flood Estimation*; Commonwealth of Australia (Geoscience Australia).

Bell, C.D., S.K. McMillan, S.M. Clinton, and A.J. Jefferson 2016. Hydrologic response to stormwater control measures in urban watershed. *Journal of Hydrology*, 541, 1488-1500.

Beven, K. J. 2012. *Rainfall-Runoff Modelling: The Primer*. John Wiley & Sons.

Bezak, N., Brilly, M., & Šraj, M. 2014. Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal*, 59(5), 959-977.

Boning, C.W. 1977. Preliminary Evaluation of Flood Frequency Relations in the Urban Areas of Memphis, Tennessee. *U.S. Geological Survey Water-Resources Investigations* 77-132, 58 pp.

Borselli, L., Cassi, P., Torri, D. 2008. Prolegomena to sediment and flow connectivity in the landscape: A GIS and field numerical assessment. *Catena*, 75(3), 268-277.

Boyd, M.J., Bufill, M.C., Knee, R.M., 1993. Pervious and impervious runoff in urban catchments. *Hydrological Sciences Journal*, 38 (6), 463–478. doi.org/10.1080/ 02626669309492699.

Braud, I., P. Breil, F. Thollet, M. Lagouy, F. Branger, C. Jacqueminet, S. Kermadi, and K. Michel. 2013. Evidence of the impact of urbanization on the hydrologic regime of a medium-sized periurban catchment in France. *Journal of Hydrology*, 485, 5-13.

Berndt, Christian, Ehsan Rabiei, and Uwe Haberlandt. 2014. Geostatistical merging of rain gauge and radar data for high temporal resolutions and various station density scenarios. *Journal of Hydrology* 508:88–101. doi: 10.1016/j.jhydrol.2013.10.028.

Brown, V. M., Keim, B. D., & Black, A. W. 2020. Trend analysis of multiple extreme hourly precipitation time series in the Southeastern United States. *Journal of Applied Meteorology and Climatology*, *59*(3), 427–442. https://doi.org/10.1175/JAMC-D-19-0119.1

Burn, D. H. 1990. Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resources Research*, 26(10), 2257-2265.

Campolo, M., Andreussi, P., and Soldati, A. 1999. River flood forecasting with a neural network model. *Water Resources Research*, 35(4), 1191–1197.

Cavalli, M., & Marchi, L. 2008. Characterisation of the surface morphology of an alpine alluvial fan using airborne LiDAR. *Natural Hazards and Earth System Sciences*, 8(2), 323-333.

Cavalli, M., Trevisani, S., Comiti, F., & Marchi, L. 2013. Geomorphometric assessment of spatial sediment connectivity in small Alpine catchments. *Geomorphology*, 188, 31-41.

Cohn, T. A., England, J. F., Berenbrock, C. E., Mason, R. R., Stedinger, J. R., & Lamontagne, J. R. 2013. A generalized Grubbs-Beck test statistic for detecting multiple potentially influential low outliers in flood series. *Water Resources Research*, 49(8), 5047-5058.

Couta, D., Zhang, Y.-K., and Li, Y.-M. 2019. River flow forecasting using long short-term memory. *DEStech Transactions on Computer Science and Engineering* (2019 International Conference on Artificial Intelligence and Computing Science (ICAICS 2019). ISBN: 978-1-60595-615-2).

Dalrymple, T. 1960. *Flood-Frequency Analyses* (No. 1543). US Government Printing Office.

Dawson, C. W. and Wilby, R. L. 1999. A comparison of artificial neural networks used for river forecasting. *Hydrology and Earth System Sciences Discussions, European Geosciences Union*, 3(4), 529–540.

Dawson, C. W. and Wilby, R. L. 2001. Hydrologic modelling using artificial neural networks. *Progress in Physical Geography*, 25(1), 80–108.

Dell'Aira, F., Chy, T. J., Goebel, T. H., & Meier, C. I. 2022. Inferring hydrologic properties of the rainfall-runoff conversion process through artificial neural network modeling. In *World Environmental and Water Resources Congress* 2022, American Society of Civil Engineers (pp. 1264-1278).

Dell'Aira, F., Cancelliere, A., & Meier, C. I. 2023. Underestimation of Frequent Floods when Using Annual Maxima for Frequency Analysis: Drivers, Spatial Variability, and Possible Solutions. *ESS Open Archive*.

DePriest, T.A. 2018: Urbanization in Jackson, Madison County, TN and the Environmental Impact of Urbanization on the Region. PhD Dissertation, University of Memphis, 84 pp.

Douglas, E.M., and C.A. Fairbank. 2011. Is precipitation in Northern New England becoming more extreme? Statistical analysis of extremerainfall in Massachusetts, New Hampshire, and Maine and updated estimates of the 100-year storm. *Journal of Hydrologic Engineering*, 16(3), 203-217.

Eagleson, P.S. 1978: Climate, soil, and vegetation. 2. The distribution of annual precipitation derived from observed storm sequences. *Water Resources Research*, 14, 713-721

Ebrahimian, A., B.N. Wilson, and J.S. Gulliver. 2016. Improved methods to estimate the effective impervious area in urban catchments using rainfall-runoff data. *Journal of Hydrology*, 536, 109-118.

England, J.F., Jr., Cohn, T.A., Faber, B.A., Stedinger, J.R., Thomas, W.O., Jr., Veilleux, A.G., Kiang, J.E., and Mason, R.R., Jr. 2019, Guidelines for determining flood flow frequency—Bulletin 17C (ver. 1.1, May 2019): *U.S. Geological Survey Techniques and Methods*, book 4, chap. B5, 148 p., https://doi.org/10.3133/tm4B5.

Epps, T.H., and J.M. Hathaway. 2018. Establishing a framework for the spatial identification of effective impervious areas in gauged basins: Review and case study. *Journal of Sustainable Water in the Built Environment*, 4(2)

Feaster, T.D., Gotvald, A.J., & Weaver, J.C. 2014. *Methods for estimating the magnitude and frequency of floods for urban and small, rural streams in Georgia, South Carolina, and North Carolina, 2011 (ver. 1.1, March 2014)*. U.S. Geological Survey Scientific Investigations Report 2014–5030, 104 p., http://dx.doi.org/10.3133/sir20145030

First Street Foundation. 2023. *The 8th National Risk Assessment: The Precipitation Problem*. New York.

Gamble, C.R., and J.G. Lewis. 1977. Technique for Estimating Depth of 100-Year Floods in Tennessee. *US Geological Survey Open-File Report* 77-668, Nashville, Tennessee.

Gamble, C.R. 1983. Technique for Estimating Depth of Floods in Tennessee. *US Geological Survey Water Resources Investigations* 83-4050, Nashville, Tennessee

Grubbs, F., and G. Beck. 1972. Extension of sample sizes and percentage points for significance tests of outlying observations. *Technometrics*, 11(1), 1-21.

Gupta, R.S. 2017. *Hydrology and Hydraulic Systems*. 4th Edition. Waveland Press, Long Grove, Illinois.

Halfhill, T. R., 2008. Parallel processing with CUDA: Nvidia's high-performance computing platform uses massive multithreading. *Microprocessor Report* 22, 1–8.

Herrera, Sixto, Sven Kotlarski, Pedro M. M. Soares, Rita M. Cardoso, Adam Jaczewski, José M. Gutiérrez, and Douglas Maraun. 2019. Uncertainty in gridded precipitation products: Influence of station density, interpolation method and grid resolution." *International Journal of Climatology* 39(9):3717–29. doi: 10.1002/joc.5878.

Hollis, G.E. 1975. The effect of urbanization on floods of different recurrence interval. *Water Resources Research*, 11(3), 431-435.

Homer, C., Dewitz, J., Jin, S., Xian, G., Costello, C., Danielson, P., Gass, L., Funk, M., Wickham, J., Stehman, S. and Auch, R. 2020. Conterminous United States land cover change patterns 2001–2016 from the 2016 national land cover database. *ISPRS Journal of Photogrammetry and Remote Sensing*, 162, pp.184-199.

Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., and Megown, K. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing*, v. 81, no. 5, p. 345-354.

Hooke, J., J., Souza, & M., Marchamalo. 2021. Evaluation of connectivity indices applied to a Mediterranean agricultural catchment. *Catena*, 207, 105713.

Hornik, K., Stinchcombe, M., and White, H. 1989. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366.

Hosking, J. R. M., & Wallis, J. R. 1997. *Regional Frequency Analysis: An Approach based on L-moments*. Cambridge University Press

Hu, Y., Yan, L., Hang, T., and Feng, J. 2020. Stream-flow forecasting of small rivers based on LSTM. *arXiv preprint arXiv:2001.05681*.

Husic, A., & Michalek, A. 2022. Structural hillslope connectivity is driven by tectonics more than climate and modulates hydrologic extremes and benefits. *Geophysical Research Letters*.

Intergovernmental Panel on Climate Change (IPCC). 2023. *Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the*

Intergovernmental Panel on Climate Change (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781009325844

Jacobson, C.R. 2011. Identification and quantification of the hydrologic impacts of imperviousness in urban catchments: A review. *Journal of Environmental Management*, 92, 1438-1448.

Jenkins, C.T. 1960. *Floods in Tennessee, Magnitude and Frequency*. Tennessee Department of Highways, 68 pp.

Jones, K. A., Niknami, L. S., Buto, S. G., & Decker, D. 2022. Federal Standards and Procedures for the National Watershed Boundary Dataset (WBD): Chapter 3 of Section A, *Federal Standards, Book 11, Collection and Delineation of Spatial Data* (No. 11-A3). US Geological Survey.

Jovanovic, T., F. Sun, T. Mahjabin, and A.I. Mejía. 2018. Disentangling the effects of climate and urban growth on streamflow for sustainable urban development: A stochastic approach to flow regime attribution. *Landscape and Urban Planning*, 177, 160-170.

Karim, F., Hasan, M., & Marvanek, S. 2017. Evaluating annual maximum and partial duration series for estimating frequency of small magnitude floods. *Water*, 9(7), 481.

Keast, D., & Ellison, J. 2013. Magnitude frequency analysis of small floods using the annual and partial series. *Water*, 5(4), 1816-1829.

Kennedy, J.R., & Paretti, N.V. 2014. *Evaluation of the magnitude and frequency of floods in urban watersheds in Phoenix and Tucson, Arizona.* U.S. Geological Survey Scientific Investigations Report 2014–5121, 29 p. http://dx.doi.org/10.3133/sir20145121.

Kermadi, S., I. Braud, C. Jacqueminet, F. Branger, F. Renard, and K. Michel. 2012. Relationships between climatology, land cover, and flooding in the Yzeron catchment (western Lyon, France): Contribution of a combined analysis using hydro-climatological data and land-cover mapping using very high resolution imagery. *Climatologie*, 9, 83-127 (in French).

Kim, J., E. Shu, K. Lai, M. Amodeo, J. Porter, and E. Kearns. 2022. Assessment of the standard precipitation frequency estimates in the United States. *Journal of Hydrology: Regional Studies*, 44: 101276

Kiran, K. G., & Srinivas, V. V. 2021. A Mahalanobis distance-based automatic threshold selection method for peaks over threshold model. *Water Resources Research*, 57, e2020WR027534.

Konrad, C. P. 2003. Effects of Urban Development on Floods. *Report 076-03 USGS Publication Warehouse* (https://doi.org/10.3133/fs07603)

Kratzert, F., Herrnegger, M., Klotz, D., Hochreiter, S., & Klambauer, G. 2019. *NeuralHydrology–Interpreting LSTMs in Hydrology*. Explainable AI: Interpreting, Explaining and Visualizing Deep Learning, Springer, p. 347-362.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., and Herrnegger, M. 2018. Rainfall–runoff modelling using long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences*, 22(11), 6005–6022.

Kumar, D. N., Raju, K. S., and Sathish, T. 2004. River flow forecasting using recurrent neural networks. *Water Resources Management*, 18(2), 143–161.

Ladd, D.E., and Law, G.S. 2007, *Tennessee StreamStats: A web-enabled geographic information system application for automating the retrieval and calculation of streamflow statistics*: U.S. Geological Survey Fact Sheet 2007-3081, 2 p.

Langbein, W.B., 1949. Annual floods and the partial-duration flood series. *Transactions, American Geophysical Union*, 30 (6), 879–881. (doi:10.1029/TR030i006p00879).

Law, G.S, and G.D. Tasker. 2003. *Flood-Frequency Prediction Methods for Unregulated Streams of Tennessee, 2000.* U.S. Geological Survey Water-Resources Investigations Report 03-4176, Nashville, Tennessee.

Law, G.S., Tasker, G.D., and Ladd, D.E. 2009, *Streamflow-characteristic estimation methods for unregulated streams of Tennessee*: U.S. Geological Survey Scientific Investigations Report 2009-5159, 212 p., 1 pl.

Leopold, L.B. 1968. Hydrology for Urban Land Planning – A Guidebook on the Hydrologic Effects of Urban Land Use. In: *Geological Survey Circular* 554, U.S. Geological Survey, Washington D.C., 18 pp.

Lengfeld, Katharina, Pierre-Emmanuel Kirstetter, Hayley J. Fowler, Jingjing Yu, Andreas Becker, Zachary Flamig, and Jonathan Gourley. 2020. Use of radar data for characterizing extreme precipitation at fine scales and short durations. *Environmental Research Letters,* 15(8):085003. doi: 10.1088/1748-9326/ab98b4.

Liu, Y. B. and De Smedt, F. 2004. *WetSpa Extension, A GIS-based Hydrologic Model for Flood Prediction and Watershed Management*. Documentation and User Manual.

Llabrés-Brustenga, A., Rius, A., Rodríguez-Solà, R., & Casas-Castillo, M. C. 2020. Influence of regional and seasonal rainfall patterns on the ratio between fixed and unrestricted measured intervals of rainfall amounts. *Theoretical and Applied Climatology*, 140: 389-399.

Mallakpour, I., & G. Villarini. 2017. Analysis of changes in the magnitude, frequency, and seasonality of heavy precipitation over the contiguous USA. *Theoretical and Applied Climatology*, *130*, 345-363.

Marasini, A., P. Muñoz-Proboste, C.I. Meier, and N. Kafle. 2023. Hershfield factors for extreme precipitation: Variability and a proposal for a unified definition. Oral presentation at the *Joint Federal Interagency Conference on Sedimentation and Hydrologic Modeling*, SEDHYD 2023, May 2023, St. Louis, MO

Meier, C.I., J.S. Moraga, G. Pranzini, and P. Molnar. 2016: Describing interannual variability of precipitation with the derived distribution approach: effects of record length and resolution. *Hydrology and Earth Systems Sciences*, 20, 4177-419

Meier, C.I., P. Muñoz-Proboste, and P. Molnár. 2019. On the variability of Hershfield-type rainfall sampling adjustment factors. *Geophysical Research Abstract* 2019-18583, European Geosciences Union.

Meier, C.I., P.I. Muñoz-Proboste, and A. Marasini. 2022. By how much do we really underestimate extreme precipitation due to rain gauge totalization? A systematic analysis of variability in rainfall sampling adjustment factors. *AGU Fall Meeting Abstracts* 2022, H42B-1249.

Meier, C.I., P.I. Muñoz-Proboste, A. Marasini, N. Kafle, and F. Dell'Aira. 2023. Hershfield rainfall sampling adjustment factors in urban hydrology: Variability, current practice, and a proposal.

UrbanRain2023: International Workshop on Precipitation in Urban Areas, ETH-Zürich, Pontresina, Switzerland, November 2023.

Mejía, A.I., and G.E. Moglen. 2010a. Spatial distribution of imperviousness and the space-time variability of rainfall, runoff generation, and routing. *Water Resources Research*, 46, W07509.

Mejía, A.I., and G.E. Moglen. 2010b. Impact of the spatial distribution of imperviousness on the hydrologic response of an urbanizing basin. *Hydrologic Processes*, 24, 3359-3373.

Miller, J.D., H. Kim, T.R. Kjeldsen, J. Packman, S. Grebby and R. Dearden. 2014. Assessing the impact of urbanization on storm runoff in a peri-urban catchment using historical change in impervious cover. *Journal of Hydrology*, 515, 59-70.

Miller, J.D., and T. Brewer. 2018. Refining flood estimation in urbanized catchments using landscape metrics. *Landscape and Urban Planning*, 174, 34-49.

Moglen, G.E., and D.E Shivers. 1983. *Methods for Adjusting US Geological Survey Rural Regression Peak Discharges in an Urban Setting*. U.S. Geological Survey Scientific Investigations Report 2006-5270.

Morbidelli, R., Saltalippi, C., Flammini, A., Picciafuoco, T., Dari, J., & Corradini, C. 2018. Characteristics of the underestimation error of annual maximum rainfall depth due to coarse temporal aggregation. *Atmosphere*, 9(8), 303.

Muñoz, P.I., C.I. Meier, and P. Molnár. 2018. Effects of location, season, and storm type on Hershfield-type rainfall sampling adjustment factors. *AGU Fall Meeting Abstracts* (Vol. 2018, pp. H41L-2266).

Neely, B.L., Jr. 1984. *Flood Frequency and Storm Runoff of Urban Areas of Memphis and Shelby County, Tennessee*. U.S. Geological Survey Water-Resources Investigations Report 84-4110, 56 pp.

Over, T.M., R.J. Saito, A.G. Veilleux, J.B. Sharpe, D.T. Soong, and A.L. Ishii. 2017. *Estimation of Peak Discharge Quantiles for Selected Annual Exceedence Probabilities in Northeastern Illinois*. USGS Scientific Investigations Report 2016-5050.

Pan, X., Rahman, A., Haddad, K., & Ouarda, T. B. 2022. Peaks-over-threshold model in flood frequency analysis: a scoping review. *Stochastic Environmental Research and Risk Assessment*, 1-17.

Papalexiou, S. M., Dialynas, Y. G., & Grimaldi, S. 2016. Hershfield factor revisited: Correcting annual maximum precipitation. *Journal of Hydrology*, 542, 884-895.

Patterson, J.L. 1964. *Magnitude and Frequency of Floods in the United States, Part 7, Lower Mississippi River Basin*. U.S. Geological Survey Water-Supply Paper 1681, 636 pp.

Persichillo, M. G., Bordoni, M., Cavalli, M., Crema, S., & Meisina, C. 2018. The role of human activities on sediment connectivity of shallow landslides. *Catena*, 160, 261-274.

Plavšić, J., Mihailović, V., & Blagojević, B. 2014. Assessment of methods for outlier detection and treatment in flood frequency analysis. In Proceedings of the Mediterranean Meeting on Monitoring, modelling and early warning of extreme events triggered by heavy rainfalls. PON 01_01503-MED-FRIEND project University of Calabria, Cosenza (Italy), June 26th-28th, 2014 (pp. 181-192).

Ponce, V. M. 2014. *Engineering Hydrology: Principles and Practices.* Online Edition, http://ponce.sdsu.edu/enghydro/index.html.

Pöschmann, Judith Marie, Dongkyun Kim, Rico Kronenberg, and Christian Bernhofer. 2021. An analysis of temporal scaling behaviour of extreme rainfall in Germany based on radar precipitation QPE data. *Natural Hazards and Earth System Sciences,* 21(4):1195–1207. doi: 10.5194/nhess-21-1195-2021.

Powell, E.J., and B.D. Keim. 2015. Trends in daily temperature and precipitation extremes for the Southeastern United States: 1948-2012. *Journal of Climate*, 28: 1592-1612.

Pryor, S.C., J.A. Howe, and K.E. Kunkel. 2009. How spatially coherent and statistically robust are temporal changes in extreme precipitation in the contiguous USA? *International Journal of Climatology*, 29(1), 31-45.

Qin, C. Z., & Zhan, L. (2012). Parallelizing flow-accumulation calculations on graphics processing units—From iterative DEM preprocessing algorithm to recursive multiple-flow-direction algorithm. *Computers & Geosciences*, *43*, 7-16.

Rahman, M.S., J.C. Senkbeil, and D.J. Keellings. 2023. Spatial and temporal variability of extreme precipitation events in the Southeastern United States. *Atmosphere,* 14(8): 1301.

Randolph, W.J., and C.R. Gamble. 1976. *Technique for Estimating Magnitude and Frequency of Floods in Tennessee*. Tennessee Department of Transportation, 52 pp.

Renard, K., Foster, G.R., Weessies, G.A., Mc Cool, D.K., Yodler, D.C. 1997. *Predicting soil erosion by water—a guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE)*. U.S.D.A.- A.R.S., Handbook No 703. 384 pp.

Robbins, C.H. 1984. *Synthesized Flood Frequency for Small Urban Streams in Tennessee*, US Geological Survey, Water-Resources Investigations Report 84-4182

Robson, A. J. and Reed, D. W. (1999) Flood Estimation Handbook: Volume 3. *Statistical Procedures for Flood Frequency Estimation*. Centre for Ecology & Hydrology.

Roy, A.H., and W.D. Shuster (2009). Assessing impervious surface connectivity and applications for watershed management. *Journal of the American Water Resources Association*, 45(1), 198-209.

Sauer, V.B., Thomas, W.O., Jr., Stricker, V.A., and Wilson, K.V. 1983. *Flood characteristics of urban watersheds in the United States*. U.S. Geological Survey Water-Supply Paper 2207, 63 p.

Schäfer, A. M. and Zimmermann, H. G. 2006. Recurrent neural networks are universal approximators. *International Conference on Artificial Neural Networks, Springer*, 632–640.

Schueler, T.R., L. Fraley-McNeal, and K., Cappiella, K. 2009. Is impervious cover still important? Review of recent research. *Journal of Hydrologic Engineering*, 14, 309–315.

Seo, Y., Choi, N.-J., Schmidt, A.R., 2013. Contribution of directly connected and isolated impervious areas to urban drainage network hydrographs. *Hydrology and Earth System Science,* 17 (9), 3473–3483. https://doi.org/10.5194/hess-17-3473-2013

Shamseldin, A.Y. 2010. Artificial neural network model for river flow forecasting in a developing country. *Journal of Hydroinformatics*, 12(1), 22–35.

Shuster, W.D, J. Bonta, H. Thurston, E. Warnemuende, and D.R. Smith 2005. Impacts of impervious surface on watershed hydrology: A review. *Urban Water Journal*, 2(4), 263-275.

Sillanpää, N., and H. Koivasulo. 2015. Impacts of urban development on runoff event characteristics and unit hydrographs across warm and cold seasons in high latitudes. *Journal of Hydrology*, 521, 328-340.

Skeeter, W.J., J.C. Senkbeil, & D.J. Keellings. 2019. Spatial and temporal changes in the frequency and magnitude of intense precipitation events in the Southeastern United States. *International Journal of Climatology*, *39*(2), 768-782.

Smith, B.K., J.A. Smith, M.L. Baeck, G. Villarini, and D.B. Wright. 2013. Spectrum of storm event hydrologic response in urban watersheds. *Water Resources Research*, 49, 2649-2663.

Soong, D.T., A.L. Ishii, J.B. Sharpe, and C.F. Avery. 2004. *Estimating flood-peak discharge magnitudes and frequencies for rural streams in Illinois*. USGS Scientific Investigations Report 2004-5103, 158 pp.

Southard, R.E. 2010. Estimating the magnitude and frequency of floods in urban basins in Missouri. *Scientific Investigations Report 2010-5073 USGS Publication Warehouse (https://doi.org/10.3133/sir20105073).*

Speer, P.R., and C.R. Gamble. 1964. *Magnitude and Frequency of Floods in the United States, Part 3B, Cumberland and Tennessee River Basins*, U.S. Geological Survey Water-Supply Paper 1676, 340 pp.

Sytsma, A., Bell, C., Eisenstein, W., Hogue, T., & Kondolf, G. M. 2020. A geospatial approach for estimating hydrologic connectivity of impervious surfaces. *Journal of Hydrology*, 591, 125545.

Tarboton, D. G. 1997. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, 33(2), 309-319.

Tavares, L. V., & Da Silva, J. E. 1983. Partial duration series method revisited. *Journal of Hydrology*, 64(1-4), 1-14.

U.S. Department of Agriculture. 1986. *Urban Hydrology for Small Watersheds.* Technical Report 55.

USGS. 1982. *Guidelines for Determining Flood Flow Frequency*. Bulletin 17B of the Hydrology Subcommittee, Interagency Advisory Committee on Water Data, US Geological Survey, Reston, VA.

U.S. Weather Bureau. 1953. *Rainfall Intensities for Local Drainage Design in the United States for Durations of 5 to 240 Minutes and 2-, 5-and 10-year Return Periods. Part I: West of the 115th Meridian*. Technical Paper N. 24, U.S. Dept. of Commerce, Washington, D.C.

Viessman, W., Jr., and G.L. Lewis. 1995. *Introduction to Hydrology,* 4th Edition. AddisonWesleyLongman.

Wang, C. H., & Holmes, J. D. (2020). Exceedance rate, exceedance probability, and the duality of GEV and GPD for extreme hazard analysis. *Natural Hazards*, 102, 1305-1321.

Weaver, J.D., and C.R. Gamble. 1993. *Flood Frequency of Streams in Rural Basins of Tennessee*. U.S. Geological Survey Water-Resources Investigations Report 76-121, 33 pp.

Wibben, H.C. 1976. *Effects of Urbanization on Flood Characteristics in Nashville-Davidson County, Tennessee*. U.S. Geological Survey Water-Resources Investigation 76-121, 40 pp.

Wohl, E., Lane, S.N. and Wilcox, A.C. 2015. The science and practice of river restoration. *Water Resources Research*, 51(8), pp.5974-5997.

Wright, D.B., C. Samaras, and T. Lopez-Cantu. 2021. Resilience to extreme rainfall starts with science. *Bulletin of the American Meteorological Society*, **102**, E808-E813

Wuebbles, D. J., and Coauthors. 2017. *Climate Science Special Report. Vol. I, Fourth National Climate Assessment*, U.S. Global Change Research Program, 470 pp., https://doi.org/10.7930/J0J964J6.

Wyżga, B. (1995). Evaluating the occurrence of low magnitude floods: a study of the reliability of the annual maximum series method. *Geografiska Annaler. Series A. Physical geography*, 23-33.

Xu, W., Jiang, Y., Zhang, X., Li, Y., Zhang, R., and Fu, G. 2020. Using long short-term memory networks for river flow prediction. *Hydrology Research*.

Yan, H. 2012. *Magnitude and Frequency of Floods for Rural, Unregulated Streams of Tennessee by L-Moments Method*. Master Thesis, University of Arkansas, Fayetteville.

Yang, G., L.C. Bowling, K.A. Cherkauer, and B.C Pijanowski. 2011. The impact of urban development on hydrologic regime from catchment to basin scales. *Landscape and Urban Planning*, 103, 237-247.

Yao, L., L. Chen, and W. Wei. 2016. Assessing the effectiveness of imperviousness on stormwater runoff in micro urban catchments by model simulation. *Hydrologic Processes*, 30, 1836-1848.

Zanandrea, F., Michel, G. P., Kobiyama, M., & Cardozo, G. L. 2019. Evaluation of different DTMs in sediment connectivity determination in the Mascarada River Watershed, southern Brazil. *Geomorphology*, 332, 80-87.

Appendices

Table A1. Urbanization trends (%IA) in developed HUC-12 basins in Tennessee from 2001 to 2021.

HUC-12 basin	2001	2004	2006	2008	2011	2013	2016	2019	2021
051301070204	8.93	8.93	9.06	9.06	9.14	9.14	9.33	11.45	11.46
051301080702	10.61	10.61	10.84	10.84	11.06	11.06	11.11	12.14	12.17
051302010306	8.51	8.51	9.96	9.96	10.66	10.66	11.19	14.99	15.24
051302010501	5.78	5.78	6.79	6.79	7.29	7.29	7.54	10.35	10.84
051302010602	3.95	3.95	4.89	4.89	5.91	5.91	6.79	9.35	10.65
051302010603	5.84	5.84	6.92	6.92	7.52	7.52	7.96	10.93	11.19
051302010605	6.97	6.97	7.85	7.85	8.05	8.05	8.26	11.69	11.88
051302020101	4.47	4.47	6.98	6.98	7.73	7.73	8.91	12.93	13.89
051302020102	25.45	25.45	26.66	26.66	27.26	27.26	27.46	33.88	33.98
051302020302	15.59	15.59	16.39	16.39	16.80	16.80	16.92	21.18	21.35
051302020304	17.60	17.60	18.14	18.14	18.47	18.47	18.81	23.28	23.57
051302020305	35.10	35.10	35.88	35.88	36.39	36.39	36.72	44.07	44.37
051302030106	5.84	5.84	7.78	7.78	8.36	8.36	8.73	12.31	12.35
051302030203	10.79	10.79	11.64	11.64	12.46	12.46	12.83	15.98	16.28
051302030204	3.97	3.97	6.01	6.01	7.08	7.08	8.09	11.84	11.98
051302030206	15.31	15.31	18.11	18.11	19.47	19.47	20.78	26.65	26.94
051302030301	7.34	7.34	8.87	8.87	9.37	9.37	9.85	13.45	13.83
051302030304	19.40	19.40	21.76	21.76	23.11	23.11	23.57	30.72	32.51
051302030305	6.30	6.30	7.49	7.49	7.84	7.84	8.11	10.93	11.24
051302030307	6.67	6.67	7.39	7.39	7.84	7.84	8.23	10.90	11.39
051302030308	11.67	11.67	12.93	12.93	14.42	14.42	15.35	20.97	21.58
051302030309	15.16	15.16	15.82	15.82	16.78	16.78	17.82	23.85	24.89
051302040105	14.64	14.64	16.39	16.39	17.43	17.43	18.23	22.96	23.18
051302040601	7.82	7.82	8.77	8.77	9.16	9.16	9.63	12.46	12.63
051302060401	7.52	7.52	8.16	8.16	8.99	8.99	9.06	10.36	10.36
051302060404	12.73	12.73	14.16	14.16	15.85	15.85	16.79	19.67	19.84
051302060406	11.34	11.34	12.45	12.45	13.43	13.43	14.08	17.82	18.29
051302060503	7.48	7.48	8.01	8.01	8.35	8.35	8.43	10.00	10.31
051302060603	2.74	2.74	5.43	5.43	6.92	6.92	8.53	10.93	11.51
051302060604	6.87	6.87	10.01	10.01	11.52	11.52	12.87	15.97	16.87
051302060707	5.24	5.24	6.67	6.67	7.90	7.90	8.76	11.38	11.79
051302060708	10.00	10.00	11.47	11.47	12.25	12.25	12.59	15.72	16.03
060101020404	5.46	5.46	5.46	5.46	5.46	5.46	7.96	12.40	12.40
060101020501	43.90	43.90	44.38	44.38	44.50	44.50	44.76	48.62	48.62
060101020502	8.32	8.32	9.11	9.11	9.56	9.56	9.85	11.76	11.78
060101020703	11.13	11.13	11.51	11.51	11.76	11.76	11.88	14.64	14.65

Table A1 (cont.) Urbanization trends in developed HUC-12 basins in Tennessee from 2001 to 2021.

060101020704	12.29	12.29	12.70	12.70	12.92	12.92	13.20	15.96	16.00
060101030503	8.52	8.52	8.92	8.92	9.03	9.03	9.13	10.99	10.99
060101030504	22.60	22.60	23.22	23.22	23.67	23.67	23.79	27.62	27.63
060101030506	16.90	16.90	18.32	18.32	19.38	19.38	19.63	22.58	22.66
060101040207	13.37	13.37	14.11	14.11	14.36	14.36	14.71	18.33	18.37
060101040308	12.51	12.51	13.79	13.79	14.18	14.18	14.33	16.79	16.90
060101070207	11.64	11.64	14.08	14.08	15.07	15.07	15.43	18.64	18.94
060101070307	5.09	5.09	8.21	8.21	9.09	9.09	9.37	12.84	13.10
060102010108	15.74	15.74	18.45	18.45	19.83	19.83	20.34	26.88	27.10
060102010110	8.26	8.26	9.30	9.30	9.71	9.71	10.01	13.25	13.42
060102010201	23.49	23.49	24.73	24.73	24.96	24.96	25.17	28.59	28.61
060102010202	34.47	34.47	35.30	35.30	35.64	35.64	35.84	39.74	39.77
060102010203	45.84	45.84	46.33	46.33	46.54	46.54	46.63	50.95	50.97
060102010204	16.73	16.73	17.46	17.46	17.91	17.91	18.12	21.48	21.58
060102010205	5.70	5.70	6.71	6.71	7.07	7.07	7.21	11.22	11.24
060102010206	8.86	8.86	9.70	9.70	10.04	10.04	10.30	12.57	12.61
060102010207	30.65	30.65	31.61	31.61	32.03	32.03	32.32	36.70	36.76
060102010208	22.79	22.79	25.33	25.33	26.37	26.37	26.85	32.10	32.32
060102010210	8.25	8.25	9.25	9.25	9.62	9.62	9.87	12.42	12.58
060102010302	8.43	8.43	9.18	9.18	9.54	9.54	9.93	12.06	12.13
060102070201	11.17	11.17	12.90	12.90	13.33	13.33	13.52	16.53	16.64
060102070202	14.93	14.93	17.28	17.28	17.80	17.80	18.33	22.00	22.29
060102070302	12.33	12.33	12.75	12.75	12.90	12.90	13.16	14.97	14.97
060102070403	10.70	10.70	11.20	11.20	11.41	11.41	11.51	13.40	13.42
060102080301	14.43	14.43	14.84	14.84	15.15	15.15	15.57	17.53	17.65
060200010502	8.16	8.16	8.83	8.83	9.17	9.17	9.40	11.86	11.93
060200010902	6.27	6.27	7.53	7.53	8.45	8.45	8.98	12.91	13.13
060200010903	27.59	27.59	27.90	27.90	28.13	28.13	28.26	32.54	32.61
060200010904	22.29	22.29	23.20	23.20	24.07	24.07	24.67	32.01	33.22
060200010905	19.22	19.22	21.41	21.41	24.22	24.22	24.78	29.94	30.20
060200011003	31.47	31.47	31.98	31.98	32.30	32.30	32.46	36.86	36.88
060200011105	7.35	7.35	8.34	8.34	8.52	8.52	8.66	10.37	10.38
060200011202	17.03	17.03	17.62	17.62	18.05	18.05	18.19	21.23	21.28
060200021404	15.41	15.41	16.19	16.19	17.07	17.07	18.25	21.05	21.14
060300020801	6.79	6.79	7.88	7.88	8.29	8.29	8.30	10.56	10.64
060400030201	3.26	3.26	5.44	5.44	6.33	6.33	7.22	10.00	10.93
060400030506	7.27	7.27	7.76	7.76	8.00	8.00	8.07	10.22	10.47
080101000703	17.60	17.60	18.56	18.56	19.27	19.27	19.45	22.02	22.03
080102050302	8.80	8.80	9.28	9.28	9.55	9.55	9.65	11.47	11.49
080102090406	9.83	9.83	10.19	10.19	10.40	10.40	10.61	11.74	11.95

Table A1 (cont.) Urbanization trends in developed HUC-12 basins in Tennessee from 2001 to 2021.

HUC-12 basin	2001	2004	2006	2008	2011	2013	2016	2019	2021
080102100305	16.78	16.78	18.12	18.12	19.17	19.17	19.85	23.91	24.02
080102100306	19.08	19.08	19.83	19.83	20.47	20.47	21.05	25.37	25.40
080102100307	32.66	32.66	35.38	35.38	36.75	36.75	37.10	41.39	41.48
080102100308	30.40	30.40	30.92	30.92	31.19	31.19	31.41	35.26	35.29
080102110101	7.60	7.60	10.22	10.22	11.52	11.52	12.92	15.97	16.41
080102110102	30.95	30.95	33.46	33.46	34.69	34.69	35.33	40.27	40.42
080102110103	37.88	37.88	38.85	38.85	40.00	40.00	40.31	44.79	45.25
080102110201	19.97	19.97	20.30	20.30	21.56	21.56	21.76	24.82	24.84
080102110301	10.03	10.03	10.69	10.69	10.86	10.86	10.96	13.01	13.72