JOINT TRANSPORTATION
RESEARCH PROGRAM
INDIANA DEPARTMENT OF TRANSPORTATION
AND PURDUE UNIVERSITY

Development of a Multiple Water Course Joint Probability Analysis Procedure for Indiana Watersheds

Ankit Ghanghas, Sayan Dey, Venkatesh Merwade

SPR-4430 • Report Number: FHWA/IN/JTRP-2023/09 • DOI: 10.5703/1288284317616

RECOMMENDED CITATION

Ghanghas, A., Dey, S., & Merwade, V. (2023). *Development of a multiple water course joint probability analysis procedure for Indiana watersheds* (Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2023/09). West Lafayette, IN: Purdue University. https://doi.org/10.5703/1288284317616

AUTHORS

Ankit Ghanghas
Graduate Research Assistant
Lyles School of Civil Engineering
(765) 426-2403
aghangha@purdue.edu
Corresponding Author

Sayan Dey, PhD
Post Doctoral Research Assistant
Lyles School of Civil Engineering
Purdue University

Venkatesh Merwade, PhD
Professor of Civil Engineering
Lyles School of Civil Engineering
Purdue University

JOINT TRANSPORTATION RESEARCH PROGRAM

The Joint Transportation Research Program serves as a vehicle for INDOT collaboration with higher education institutions and industry in Indiana to facilitate innovation that results in continuous improvement in the planning, design, construction, operation, management and economic efficiency of the Indiana transportation infrastructure. https://engineering.purdue.edu/JTRP/index_html

Published reports of the Joint Transportation Research Program are available at http://docs.lib.purdue.edu/jtrp/.

NOTICE

IMAGE CITATION

Boomsa, J. (2019, April 3). *Fort Wayne is a small city in Northern Indiana* [Photograph]. iStock. https://www.istockphoto.com/photo/fort-wayne-is-a-small-city-in-northern-indiana-gm1139979584-304897433

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.
FHWA/IN/JTRP-2023/09

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
Development of a Multiple Water Course Joint Probability Analysis Procedure for Indiana Watersheds

5. Report Date
January 2023

6. Performing Organization Code

7. Author(s)
Ankit Ghanghas, Sayan Dey, and Venkatesh Merwade

8. Performing Organization Report No.
FHWA/IN/JTRP-2023/09

9. Performing Organization Name and Address
Joint Transportation Research Program
Hall for Discovery and Learning Research (DLR), Suite 204
207 S. Martin Jischke Drive
West Lafayette, IN 47907

10. Work Unit No.

11. Contract or Grant No.
SPR-4430

12. Sponsoring Agency Name and Address
Indiana Department of Transportation (SPR)
State Office Building
100 North Senate Avenue
Indianapolis, IN 46204

13. Type of Report and Period Covered
Final Report

14. Sponsoring Agency Code

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

16. Abstract

The design of hydraulic structures located near a confluence of two streams must take into consideration the flows from both of the streams. A hydraulic structure located on a small tributary that drains into a large river immediately downstream is not just affected by the flow in the tributary, but also by the backwater flow from the downstream river. Currently INDOT uses a tabular summary (Table 1.1 and Table 7.3 in the HEC-22 manual) of joint probabilities of coincident flows in designing hydraulic structures at confluences. However, the source of the table is unknown, and the tabular summary provides coincidental flows for only 1% and 10% probabilities, and thus it cannot be used directly for other probabilities. This study analyzed the interdependence of flows in mainstream and tributary and then developed a Gumbel-Hougard Copula-based procedure for estimating joint probabilities for confluences in Indiana. The study found that the mainstream and tributary streamflow are significantly correlated with Kendall's Tau varying generally ranging from 0.5 to 0.8. Furthermore, the Kendall's Tau, which is the key parameter for Gumbel-Hougard Copula, was found to be significantly related to drainage area ratio (DAR). Regression-based equations between DAR and τ are used as a basis to relate DAR to joint probabilities at confluences. The study also found that the currently used tabular summary (Table 1.1 and Table 7.3 in HEC-22 manual) resulted in significantly conservative design estimates and therefore overdesigned structures.

17. Key Words
joint probability, copula, confluence design, coincident flows, Indiana, NWM, national water model

18. Distribution Statement
No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)
Unclassified

20. Security Classif. (of this page)
Unclassified

21. No. of Pages
25 including appendices

22. Price

Form DOT F 1700.7 (8-72)

Reproduction of completed page authorized

EXECUTIVE SUMMARY

Introduction

The design of hydraulic structures located near the confluence of two streams must take into consideration the flows from both the streams. A hydraulic structure located on a small tributary that drains into a large river immediately downstream is not just affected by the flow in the tributary, but also by the backwater flow from the downstream river. Currently, INDOT uses a tabular summary (Table 1.1) of joint probabilities of coincident flows when designing hydraulic structures at confluences. However, the source of the table is unknown and also the tabular summary provides coincidental flows for only 1% and 10% probabilities, and thus it cannot be directly used for other probabilities.

Findings

This study analyzed the interdependence of flows in the mainstream and the tributary and developed a Gumbel-Hougard Copula-based procedure for estimating joint probabilities for confluences in Indiana. The study found that the mainstream and tributary streamflow are significantly correlated with Kendall's Tau and generally vary from 0.5 to 0.8. Furthermore, the Kendall's Tau, which is the key parameter for Gumbel-Hougard Copula, was found to be significantly related to drainage area ratio (DAR). Regression-based equations between DAR and τ were used as a basis to relate DAR to joint probabilities at confluences. The study also found that the currently used tabular summary (Table 1.1 in the HEC-22 manual) resulted in significantly conservative design estimates, which can lead to over-designed hydraulic structures.

Implementation

The findings from this study translate into a code that can be implemented in MS Excel. This program was submitted to INDOT as part of the final project report. A brief description of the MS Excel file and how to use it is included in the final project report. This program will be used by the hydraulics division at INDOT for design projects and local federal aid projects.

CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Any hydraulic structure is designed to comply with specific performance objectives such as a 50-year flood or 100-year flood, which are estimated based on the flood frequency analysis for the stream on which the structure is located. However, a hydraulic structure located near a confluence can be influenced by high flows from both the confluent streams. The level of influence depends on the location of the structure relative to the confluence. For instance, consider the three structures near the confluence as presented in Figure 1.1. Structure H1 in Figure 1.1 is sufficiently close to the confluence and its design is based on the sum of discharges from both the mainstream and its tributary. Structure H1 is designed very similar to the ones designed downstream of the confluence. On the other hand, structure H3 is sufficiently far upstream of the confluence, and it is not influenced by the flow in the mainstream channel. Structure H3 is designed like any traditional structure using univariate flood frequency analysis of the tributary. If distance X_{max}, also referred as influence reach, in Figure 1.1 is considered as the maximum distance to get affected by the flows at the confluence, any structure lying within the influence reach will not just be affected by the discharge in the tributary but also affected by the backwater flowing from the mainstream. For the design of the structures lying within the influence reach, it is critical to understand the joint behavior of discharges from confluent streams. Thus, accurate joint probability estimates are key to designing these near confluence infrastructures.

Traditionally the Indiana Department of Transportation (INDOT), as well as most agencies across the US, use Table 7-3 from HEC-22, *Urban Drainage Design Manual* (Brown et al., 2009) for joint probability analysis at confluences (see Table 1.1). This joint probability table (JPT) is believed to be developed by the Norflok District of United States Army Core of Engineers (USACE) (Kilgore et al., 2010). However, no documentation related to the creation of the JPT, including the data used, methodology and location of the study is available. Moreover, the JPT provides joint probabilities for only 10-year and 100-year design periods, thus limiting its use for other design flows. Other common joint probability techniques for confluences include the use of bivariate or multivariate frequency analysis, total probability analysis and copulas. Some other techniques, including regression relations (Kirby et al., 2002), marginal analysis and synthetic storm cell/runoff modelling have also been proposed, but they are not widely used.

Bivariate or multivariate frequency analysis is a widely method in hydrology to study the joint impact of two or more variables. Bivariate normal distribution was first used by Sackl and Bergmann (1987) to describe flood events using both flood peak and volume. Shiau (2003) then used bivariate extreme value distribution to model extreme flood events using flood peak and flood runoff volume. With regard to confluence designs, Morris and Calise (1987) and Raynal and Salas (1987) used bivariate frequency analysis to analyze concurrent flooding near confluences. Although bivariate or multivariate frequency analysis is a good tool for joint probability analysis, it assumes same type of distribution

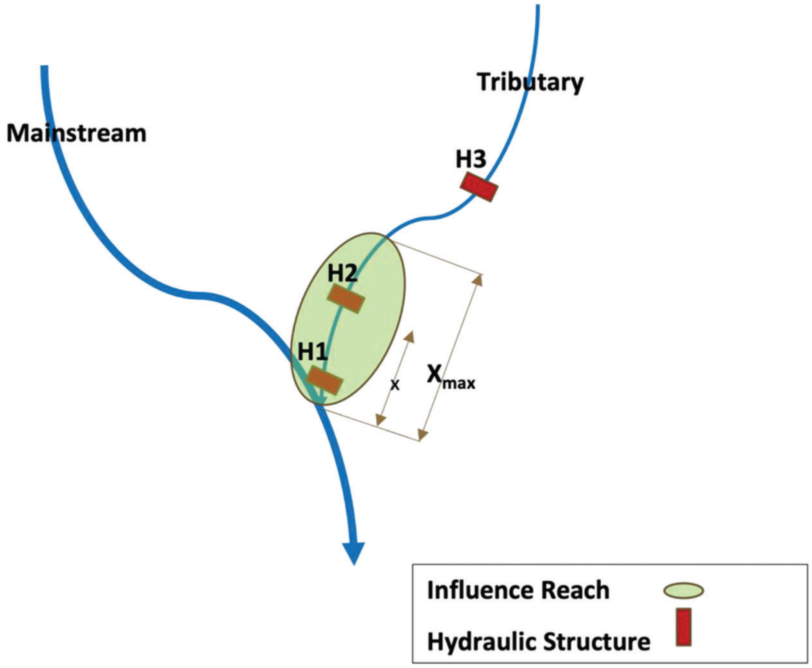


Figure 1.1 Problem statement schematic. The figure visualizes the influence region where structures are affected by backwater from mainstream.

TABLE 1.1
Joint Probability Table Used by INDOT Based on Table 7-3 of HEC-22

	Frequencies for Coincidental Occurrence			
	10-Year Design		100-Year Design	
Area Ratio	Mainstream	Tributary	Mainstream	Tributary
10,000 to 1	1	10	2	100
	10	1	100	2
1,000 to 1	2	10	10	100
	10	2	100	10
100 to 1	5	10	25	100
	10	5	100	25
10 to 1	10	10	50	100
	10	10	100	50
1 to 1	10	10	100	100
	10	10	100	100

for all the dependent variables thus limiting its application. Total Probability analysis is another well-established tool for joint probability analysis and is widely used by the United States Army Corps of Engineers for estimating flood magnitudes (USACE, 1993, 2018). Nathan and David (2004) used total probability to estimate annual damage caused by interior flooding. Total Probability is a useful method because it can produce complete probability distribution function for tributary stage, but it is computationally expensive when used for a large number of stations. The computational cost limits the use of this method for large scale implementation.

In the recent years, Copulas are widely being used to solve multivariate hydrological problems due to their low computational requirements and versatility (Bender et al., 2014, 2016; de Michele & Salvadori, 2003; Ganguli & Reddy, 2013; Gilja et al., 2018; Grimaldi & Serinaldi, 2006; Kao & Govindaraju, 2010; Karmakar & Simonovic, 2008, 2009; Kilgore et al., 2010; Peng et al., 2017; Wahl et al., 2012; Wang, 2016; Wang et al., 2009; Zhang & Singh, 2006). Copulas are distribution agnostic and allow the use of different marginal distributions for all the random variables used in joint probability analysis. When considering the joint probability of confluences, this property of Copulas enables using different probability density functions for mainstream and tributary flows. Favre et al. (2004) used Frank-Clayton and Farlie-Gumbel-Morgenstern Copulas for bivariate flood frequency analysis downstream of a reservoir and a watershed. Copulas have also been coupled with Monte-Carlo scheme to effectively estimate joint probability for gaged and ungaged confluences (Wang, 2016; Wang et al., 2009). The generalized approach for using Copulas for joint probability analysis at confluences was first presented in the National Cooperative Highway Research Program report (Kilgore et al., 2010). Kilgore et al. (2010) used 83 gaging station pairs across the United States and found the Gumbel-Hougard Copula to perform better than bivariate distribution and total probability for design of confluences. Using Gumbel-Hougard Copula for joint probability analysis at confluence was also recommended by Bender et al. (2016) and Gilja et al. (2018), which presented case studies for confluences in Germany and Croatia, respectively.

Given the large uncertainty in Table 7-3 of HEC-22 and complex procedures for joint probability analysis, this study aims to develop a simple procedure for estimating joint probabilities for confluent streams in Indiana. The study delivers Indiana specific JPT similar to the already familiar Table 7-3 of HEC-22, but it includes more design frequencies. This study is performed using the Gumbel-Hougard Copula across 4,500 confluences identified from the National Hydrography Dataset Medium Resolution (NHD-MR) to estimate joint probabilities across Indiana confluences. A regional method to estimate better joint probabilities in hydrologically similar regions across Indiana is also presented.

2. DATA

The study emphasizes on a comprehensive statewide analysis of confluences and focuses on analyzing confluences of two streams across Indiana. Although, the United States Geological Survey (USGS) maintains a network of nearly 275 streamflow gaging stations with continuous records across Indiana, only a handful of these stations are located on both mainstream and tributary and can be utilized for studying the joint behavior of streams at a confluence. Moreover, a robust and reliable flood frequency analysis is often based (or subject to) on long term availability of streamflow observations and among the handful of USGS gages across confluences, only a few have long term stream records (greater than 10 years). This limits the use of USGS stream gages to perform a robust and comprehensive statewide analysis. To overcome these limitations, this study uses data from the National Water Model retrospective run dataset version 2.1 (NWM_Rv2.1).

NWM_Rv2.1 runs WRF-Hydro (Salas et al., 2018), which is a process-based model that utilizes Noah-MP (Niu et al., 2011; Yang et al., 2011) for simulating land surface processes, kinematic routing for overland flow, and Muskingum-Cunge for channel routing. NWM_Rv2.1 provides hourly streamflow records for ~2.7 million streams as defined in the National Hydrography Dataset Medium Resolution (NHD-MR) across the Contiguous United States (CONUS) over a period of 42 years from February 1979 to December 2020. However, this study only uses hourly streamflow data from 1st January 1980 to 31st December 2008. Furthermore, the relationship between confluence behavior and stream characteristics is developed using drainage area, stream order and the results are clustered over similar hydrologic regions as defined by Rao (2004) for Indiana (Figure 3.2). The drainage area and stream order for all streams are obtained from the NHD-MR dataset.

3. METHODOLOGY

Confluence is a location where two or more streams (mainstream and its tributaries) meet. Among the merging streams, this study defines the stream with the largest drainage area as the mainstream and the other streams as tributaries. The overall methodology involves identifying confluences from a stream network, computing coincidental flow pairs, and finally describing the joint probability distribution for each confluence. A detailed description of each step is provided below.

3.1 Identification of Confluences

Confluences across Indiana are identified using the Medium Resolution National Hydrography Dataset (NHD MR). Confluence streams for the study are defined as a pair of discrete streams draining to the same node. The distinction to select discrete streams is essential to avoid confluences formed as a result of riverine islands. Confluences formed due to riverine islands can be studied by only using the mainstream discharges and do not need joint probability analysis. This study primarily focuses on two-stream confluences, so confluences with three or more streams converging at one location are excluded. The NWM streamflow contain relatively large uncertainty for smaller watersheds, so the study focuses only on confluences with mainstream order greater than or equal to 2, thus excluding first order stream confluences and a minimum tributary drainage area of 1 km^2. A total of nearly 4,500 confluences satisfying the above criteria are identified and used in this study.

3.2 Computation of Coincidental Flow Pairs

The most simplistic and widely used approach to perform joint probability analysis at a confluence involves identifying annual maximum series on mainstream and tributary to form a joint characteristic function (Wang, 2016; Wang et al., 2009). This approach assumes that the two annual maxima in a year (on mainstream and tributary) are completely correlated. However, this may not necessarily be true for all confluences (Figure 3.1) and a design based on such an approach will yield a conservative flow value, thus resulting in over designing of structures at the confluence.

An alternative approach to identify flows for joint probability analysis is to use coincidental flows. A coincidental flow pair includes simultaneous flows from both the mainstream and the tributary. In the case of instantaneous time series, coincidental flow pair includes values at a given instant, whereas for daily time series it will include values occurring on the same day. However, as evident from Figure 3.1, the highest flows due to any precipitation event may not occur simultaneously on both the mainstream and the tributary. The lag time/time to peak would vary based on watershed characteristics as well as the intensity, duration, and spatial extent of the precipitation event. Thus, defining coincidental flow pair using values at a given time instant or day may not result in truly correlated flows between the mainstream and tributary watersheds.

To address the issue related to lag times in finding coincidental flow pairs, this study first finds the annual maximum streamflow on the mainstream. Based on a time buffer or window (t days) centered around the time of annual maxima peak on the mainstream, maximum streamflow for the tributary falling within this window is picked. Using this criterion, two dataset pairs namely Peak on Mainstream pair (POM) and Peak on Tributary pair (POT) are formed.

The larger buffer window may sometimes result in more coincidental flow pairs, but some of these pairs may not be truly correlated. For example, the flow on the focus stream may be significantly lower than the true peak by the time the maximum flow within the buffer window is achieved on the tributary. To address this issue, this study tried three different coincidental flow criteria including a buffer time window of 1, 7, and 15 days. Subsequently the analysis present here uses seven-day buffer window to compute coincident flow pairs as using the 15-day buffer resulted in similar overall conclusions.

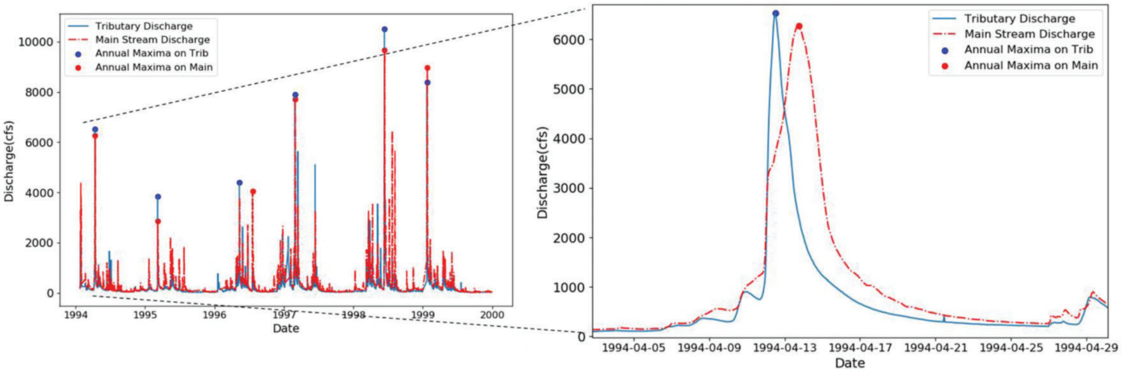


Figure 3.1 Comparison of mainstream and tributary discharge (USGS 3334000 and USGS 3334500 gage station, respectively). The image illustrates that peak discharge for mainstream and tributary do not occur simultaneously.

3.3 Fitting Copulas to Coincidental Flow Pairs

Modelling of hydrological phenomena often requires multivariate analysis given their dependence on several variables. Copulas are one such dependence functions which can be used to link and capture the dependence between two or more random variables. Copulas can be used to analyze the joint behavior of multiple (more than 3) independently distributed univariate random variables.

Let $X = (X_1, X_2, X_3,\ldots,X_n)$ be a random vector of size n with continuous marginal distributions functions (CDFs) F_1, F_2, F_3,…,F_n Following the Sklar's theorem (Sklar, 1959), the joint cumulative distribution function H(X) of the random vector X is represented as,

$$H(X) = C\{F_1(X_1), F_2(X_2), F_3(X_3), \ldots, F_n(X_n) : \theta\} \quad x \in R^n \qquad \text{(Eq. 3.1)}$$

where C: $[0,1]^n \rightarrow [0,1]$ is the n dimensional Copula function with a dependence parameter θ. Detailed theoretical background can be found in Nelsen (2006) and Sklar (1959).

Among the various families of Copula functions, the Archimedean Copulas have been widely used in hydrology due to their simplicity and effectiveness in capturing the symmetry, associativity, and wide range of other dependence characteristics among the input variables. Gumbel-Hougard Copula, Clayton Copula, Ali-Milkail-Haq Copula and Frank Copula are some of the most well-known bivariate Archimedean Copulas, each with its own individual characteristics. Clayton Copula has lower tail dependence, Frank Copula has no tail dependence, Gumbel-Hougard Copula has upper tail dependence and Ali-Mikhail-Haq Copula can be modulated to fit both upper and lower tail. Ali-Mikhail-Haq Copula is versatile however it's not suitable to represent joint probability when highly positive or negative correlation exists between the input variables (Gilja et al., 2018; Kumar, 2010). Therefore, this study uses Gumbel-Hougard Copula given its good upper tail dependence for fitting annual maxima flows.

Gumbel-Hougard Copula can be described by Equation 3.2 below.

$$C_\theta(F_1(X_1), F_2(X_2)) = \exp\{-[(-\ln(F_1(X_1)))^\theta + (-\ln(F_2(X_2)))^\theta]^{\frac{1}{\theta}}\} \qquad \text{(Eq. 3.2)}$$

where θ is the dependence parameter. The dependence parameter is a function of Kendall's Tau τ given by Equation 3.3

$$\theta = \frac{1}{1-\tau} \qquad \text{(Eq. 3.3)}$$

The Gumbel-Hougard Copula is a simple function, and the joint probability of the dependent variables can be defined using just the Kendall's Tau (τ). Kendall's Tau is a rank-correlation metric between two random variables and can take values from -1 to 1. The degree of correlation is lowest for 0 and highest as the absolute value of τ gets to 1. A value of $\tau = 1$ suggests a strong positive correlation, whereas $\tau = -1$ suggests strongest negative correlation between the variables. More details on calculating τ can be found in Genest and Favre (2007) and Kendall (1938).

3.4 Regression and Regionalization

The study focuses on understanding the variation in τ as it is the single dependent parameter which governs the Gumbel-Hougard Copula and hence governs the joint probability of confluences in this study. Additionally, to analyze the variation in joint behavior of confluent streams across different regions, τ is clustered in eight hydrologically similar regions (Figure 3.2) (Ramachandra Rao & Srinivas, 2006; Rao, 2004).

Traditional confluence design approaches use simple relationships between drainage area ratio (DAR) and design return periods as DAR is found to be closely related to the return periods on confluences (Brown et al., 2009; Kilgore et al., 2010). DAR refers to the ratio of mainstream drainage area to the tributary drainage area. To facilitate the adoption of new methodology and familiarity with the traditional joint probability table, this study links DAR to joint design return periods by correlating DAR to τ. The study performs a simple linear regression between the logarithm of DAR and correlation parameter Kendall's Tau τ (Equation 3.4).

$$\tau = m \ln R_a + c \qquad \text{(Eq. 3.4)}$$

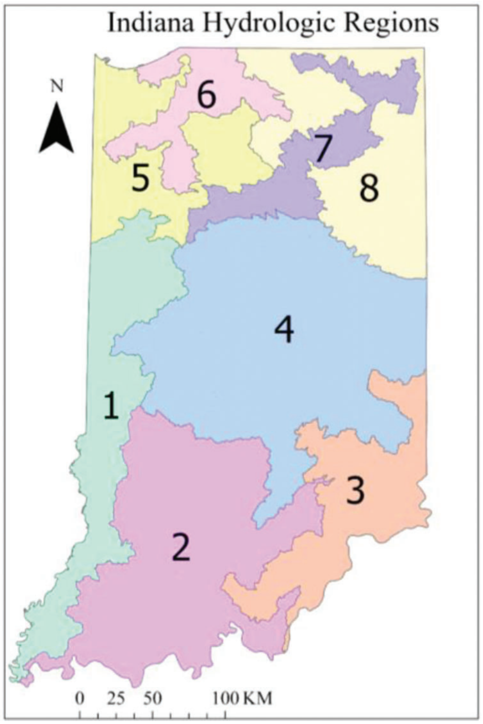


Figure 3.2 Hydrologically similar regions in Indiana (Rao, 2004).

where R_a represents DAR, m is the slope of linear regression and c is the intercept, and τ can then be used to find the joint probability using Gumbel-Hougard Copula using Equations 3.2 and 3.3. Regression equations for computing τ are developed for confluences across the entire state and for confluences in each region (Figure 3.2).

4. RESULTS

Figure 4.1 presents the summary of variation in τ across Indiana and hydrologically similar regions. Both POM and POT show similar variation in τ as its statewide median value for POM and POT datasets is 0.73 and 0.70, respectively. Both datasets have same upper quartile at 0.82. The lower quartile τ for POM and POT is slightly different at 0.60 and 0.62, respectively, but not statistically different (Figure 4.1(a)). Regionally, as shown in Figures 4.1(b) and 4.1(c), the variation in τ is significant in the interior part that includes with Regions 5–7. Regions closer to Lake Michigan have significantly lower τ compared to other regions. In a boarder sense Northern Indiana (Regions 5–8) has relatively lower τ compared to Southern Indiana (Regions 1–4). Even when analyzing τ across different regions, both POM and POT show similar inter-region variation and the overall difference in POM and POT is negligible.

The geographic variation of τ (Figure 4.2) show that the confluences formed by lower order streams entering into a higher stream order stream tend to have lower value of τ for the POM dataset. This implies that for the POM dataset, as the stream order difference between mainstream and tributary increases, the discharge in the smaller stream does not significantly impact the joint behavior. Similar behavior is also observed in the POT dataset, but it is not as prominent compared to the POM dataset, implying that for POT dataset, the discharge in smaller stream can impact the joint behavior of the confluence. Moreover, a larger difference in stream order is mostly accompanied by a larger DAR, implying that DAR has a significant correlation with τ.

To further analyze the relation between DAR and τ, confluences across Indiana are clustered into bins based on DAR as presented in Figure 4.3. τ tends to decrease exponentially with an increase in DAR. Moreover, the mean τ is lower compared to median τ for lower DAR while the mean τ is higher compared to median τ for higher DAR, implying a shift from negatively skewed distribution of τ for lower DAR to a positively skewed distribution of τ for higher DAR. This is in accordance with the findings of Kilgore et al. (2010) and Brown et al. (2009), which imply a lower correlation between streams for higher values of DAR.

Having established that Kendall's τ decreases exponentially with DAR, a simple linear regression is performed between the logarithm of DAR and τ. The statewide best fit equation and variation of τ with DAR is presented in Figure 4.4. POM shows a slightly better fit with an R^2 of 0.36 compared to POT, which has an R^2 of 0.35. The values for R^2 are consistent with the findings of Kilgore et al. (2010) which presented a relation between τ and DAR for select confluences across the United States. To account for the inter-region variability in τ (Figure 4.1), best fit logarithmic equations between τ and DAR are developed for different regions (Figure 4.5). Figure 4.5 presents the best fit R^2 values for the different regions across Indiana. While Region 1 shows best fit with an R^2 of 0.59 for POM and 0.56 for POT, Region 6 has the worst fit with an R^2 of 0.09 for POM and 0.07 for POT. The hydrology of both Region 5 and Region 6 is influenced by Lake Michigan, which may explain large variability and poor R^2 in the region. Relatively poor accuracy and inconsistent streamflow of NWM near the lakes and coastlines could also contribute to larger variability in τ in Region 5 and Region 6. All other regions show relatively better regional fit between τ and DAR compared to statewide results. The exact regional equations and plots are presented in Appendix A.

Practically, hydraulic design at confluences should be based on the maximum of the design flows obtained from the POM and POT datasets. Considering that the difference between POM and POT datasets for Indiana is not statistically significant, this study recommends using either POM or POT based design flows. Table 4.1 presents the best fit statewide and regional equations to estimate τ using DAR. Both Regions 5 and 6 have a poor regional fit so the use of statewide best fit equation is recommended for these regions.

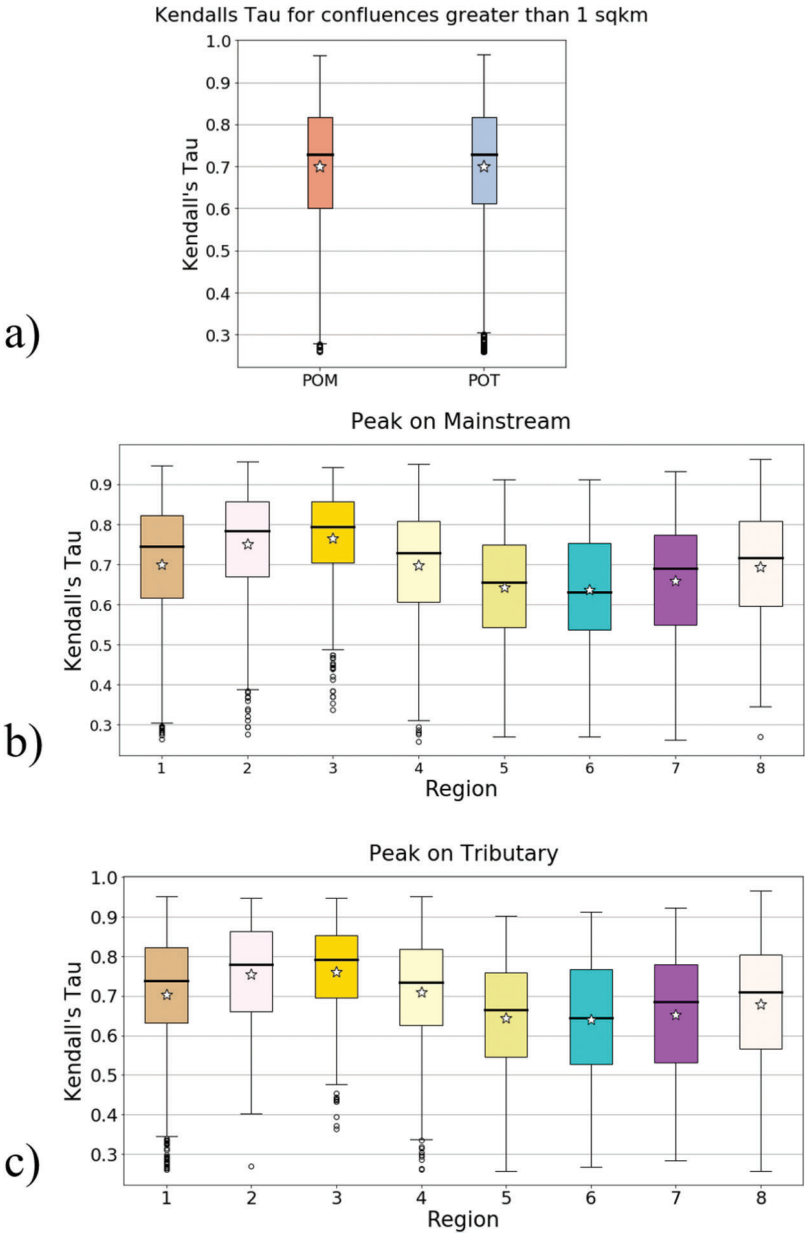


Figure 4.1 Box plot of Kendall's Tau variation across confluences. (a) Kendall's Tau variation across Indiana for both POM and POT. (b) Box plot of Kendall's Tau variation for POM dataset across different hydrologically similar regions. (c) Box plot of Kendall's Tau variation for POT dataset across different hydrologically similar regions.

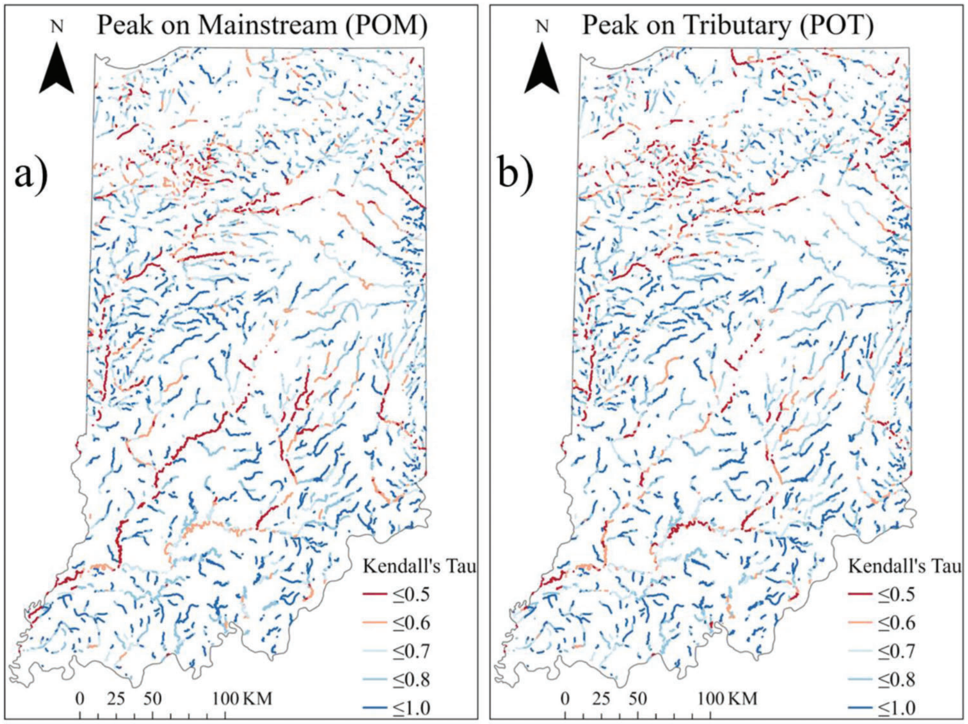


Figure 4.2 Geographic variation of Kendalls' Tau across Indiana. (a) Variation of Kendall's Tau for POM dataset. (b) Variation of Kendall's Tau for POT dataset.

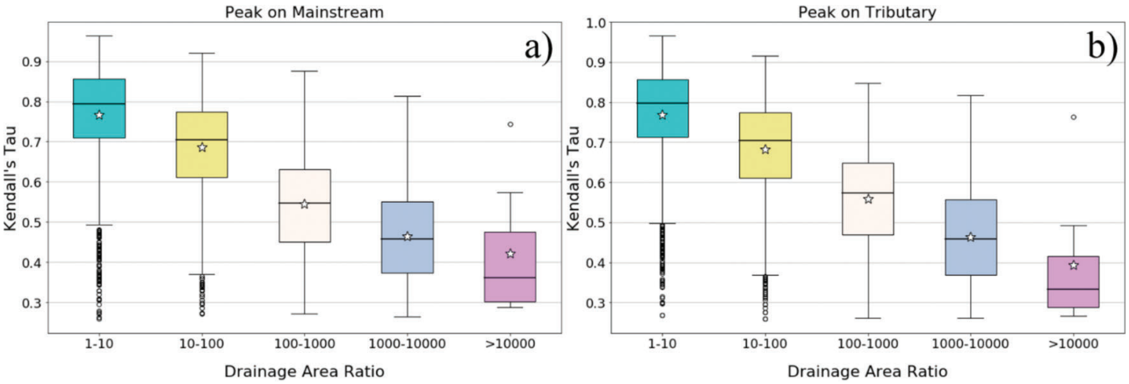


Figure 4.3 Effect of drainage area ratio on Kendall's Tau. Figure presents the variation in τ for confluences clustered in 1–10, 10–100, 100–1,000, 1,000–10,000, >10,000 DAR bins. (a) Variation for POM dataset. (b) Variation for POT dataset.

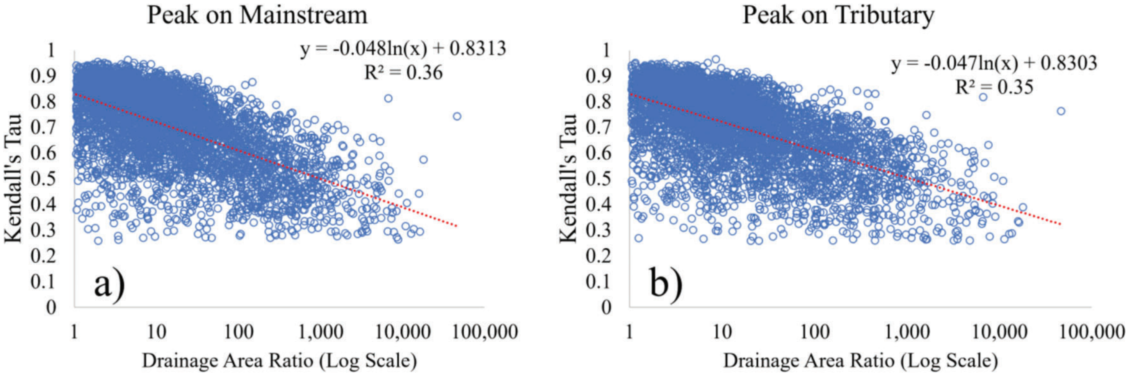


Figure 4.4 Statewide best fit regression equation and variation of Kendall's Tau with drainage area ratio. (a) Variation for POM dataset. (b) Variation for POT dataset.

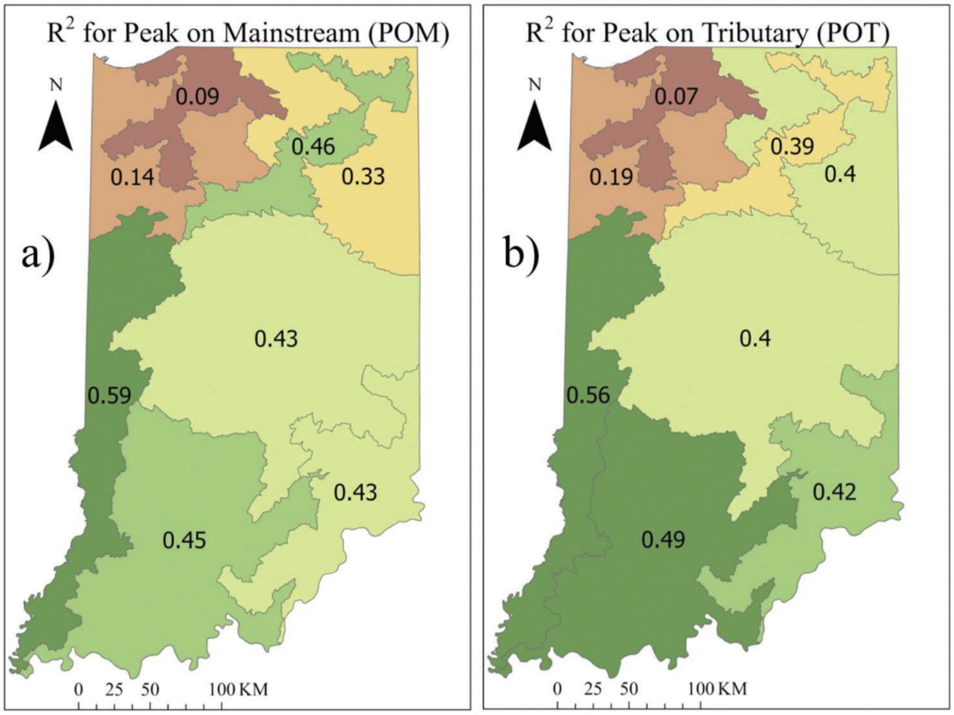


Figure 4.5 Best fit R^2 for different hydrologically similar regions of Indiana. (a) Best fit R^2 for POM dataset. (b) Best fit R^2 for POT dataset.

TABLE 4.1
Statewide and Regional Best Fit Equations for Kendall's Tau and Drainage Area Ratio

INDIANA: $\tau = -0048 \log(DAR) + 0.8313$

Region 1: $\tau = -0.057 \ln(DAR) + 0.8765$
Region 2: $\tau = -0.045 \ln(DAR) + 0.8826$
Region 3: $\tau = -0.055 \ln(DAR) + 0.8951$
Region 4: $\tau = -0.056 \ln(DAR) + 0.8563$
Region 5: $\tau = -0.048 \log(DAR) + 0.8313$
Region 6: $\tau = -0.048 \log(DAR) + 0.8313$
Region 7: $\tau = -0.048 \log(DAR) + 0.7952$
Region 8: $\tau = -0.055 \ln(DAR) + 0.8261$

5. IMPLEMENTATION

The potential combinations for different joint design frequency events can be computed using the Gumbel Hougard Copula (Equation 3.2) and the dependence parameter equations (τ) (Equation 3.3). Table 5.1–Table 5.5 here present the coincident flow frequencies associated with different τ for 10-year, 25-year, 50-year, 100-year and 500-year return period floods, respectively.

This report presents two methods for estimating joint flood frequency for confluences. Method 1 presents more accurate method based on regional regression equations, while Method 2 presents a summarized, simpler, and quicker method based on statewide regression equation. To simplify the implementation, users are recommended to round off the return period to nearest bigger and more familiar return period such as 1, 2, 5, 10, 25, 50, and 100-year).

5.1 Method 1

Implementation of joint return period estimation using Method 1 is summarized in Figure 5.1 and can be better understood by the following illustration. Suppose a structure in Region 4 is to be designed across a confluence with a DAR of 100 for 100-year joint return period. Assuming that the designer uses regional regression equations (Table 4.1), τ is estimated to be 0.59. Rounding off τ to 0.6 and using Table 5.4 for 100-year design, the recommended design flows would be a combination of 100-year and 5-year flow. The old method based on HEC 22 design manual (Table 1.1 or Table 7-3 of HEC-22) instead recommends a 100-year and 25-year flow combination which is significantly more conservative and leads to overdesigned structures (also see Appendix B).

TABLE 5.1
10-Year Joint Return Period Table

	10-Year Joint Return Period			
Kendall's	Return Period on One Stream			
Tau	1.25	2	5	Equal
0.2	9	7	3	4
0.3	10	8	5	5
0.4	–	9	6	6
0.5	–	10	8	6
0.6	–	–	9	7
0.7	–	–	10	8
0.8	–	–	–	9
0.9	–	–	–	9

TABLE 5.2
25-Year Joint Return Period Table

	25-Year Joint Return Period				
Kendall's	Return Period on One Stream				
Tau	1.25	2	5	10	Equal
0.2	23	20	13	7	8
0.3	24	22	17	11	11
0.4	25	24	21	16	13
0.5	–	25	23	20	15
0.6	–	–	24	22	17
0.7	–	–	25	24	19
0.8	–	–	–	25	21
0.9	–	–	–	–	23

TABLE 5.3
50-Year Joint Return Period Table

	50-Year Joint Return Period					
Kendall's	Return Period on One Stream					
Tau	1.25	2	5	10	25	Equal
0.2	47	41	30	21	7	15
0.3	49	46	39	31	15	20
0.4	50	49	45	40	25	25
0.5	–	50	48	45	34	30
0.6	–	–	50	48	42	34
0.7	–	–	–	50	47	39
0.8	–	–	–	–	49	43
0.9	–	–	–	–	50	46

TABLE 5.4
100-Year Joint Return Period Table

	100-Year Joint Return Period						
Kendall's	Return Period on One Stream						
Tau	1.25	2	5	10	25	50	Equal
0.2	95	86	67	53	32	12	28
0.3	98	94	84	74	54	28	39
0.4	100	98	94	88	73	48	49
0.5	–	100	98	95	87	68	59
0.6	–	–	100	99	95	83	68
0.7	–	–	–	100	99	93	77
0.8	–	–	–	–	100	99	85
0.9	–	–	–	–	–	100	93

TABLE 5.5
500-Year Joint Return Period Table

	500-Year Joint Return Period							
Kendall's	Return Period on One Stream							
Tau	1.25	2	5	10	25	50	100	Equal
0.2	485	452	395	354	294	239	167	132
0.3	496	486	462	439	398	353	286	189
0.4	499	497	489	480	459	431	381	243
0.5	–	500	498	496	488	474	445	293
0.6	–	–	–	500	498	494	481	341
0.7	–	–	–	–	500	499	497	385
0.8	–	–	–	–	–	–	500	426
0.9	–	–	–	–	–	–	500	464

5.2 Method 2

Assuming the familiarity of designers with the current joint probability table (Table 1.1), the Indiana specific frequency for coincidental occurrence is summarized and presented in Table 5.6. Table 5.6 presents the summary of the findings of the report, but Table 5.7 is aimed for direct practical implementation and is a conservative round up of Table 5.6 results. The joint flood frequency summarized in Table 5.6 and 5.7 are computed using Indiana statewide equation (Table 4.1) and approximating the frequency obtained from the τ to the nearest larger year (for example an equivalent frequency of 8 years is rounded up to 10 years). Table 5.7 is very similar to Table 1.1 (current method) and provides a quick way to determine the joint return periods. Considering the same illustration as above, Table 5.7 also recommends a combination of 100-year and 5-year flow for design of a 100-year structure near the confluence.

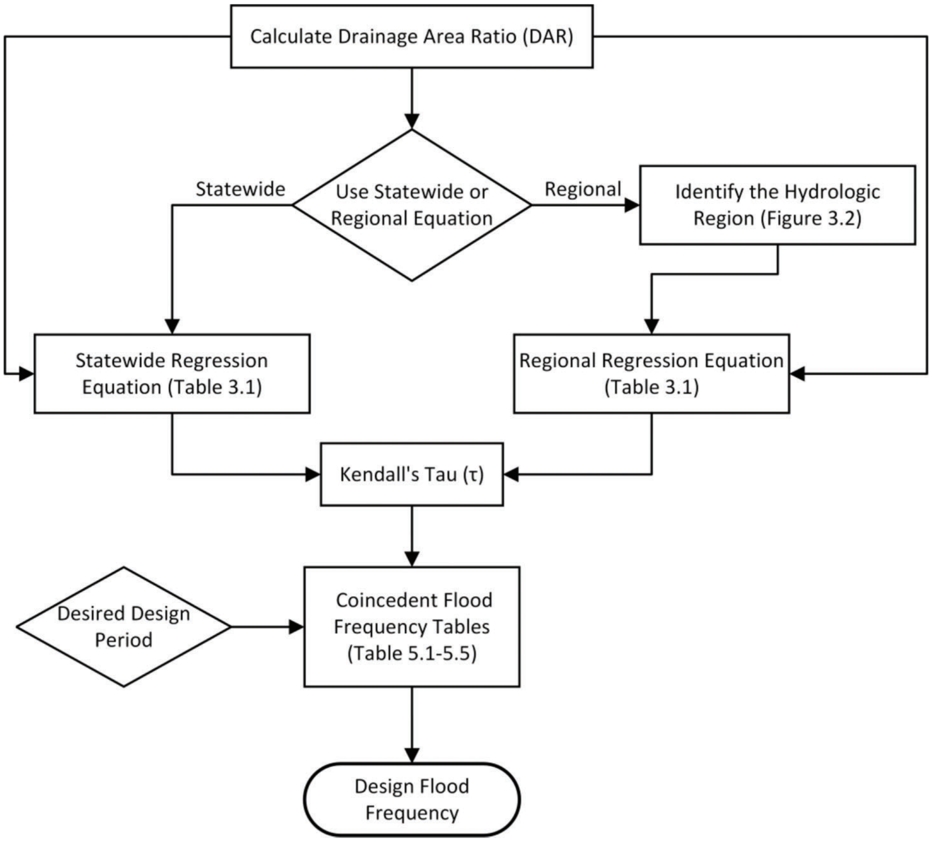


Figure 5.1 Flowchart for using copula-based joint probability analysis using Method 1.

TABLE 5.6
Summary of Findings of Joint Probability Analysis for Indiana Specific Confluences

	Frequencies for Coincidental Occurrence							
	10-Year		25-Year		50-Year		100-Year	
Area Ratio	**Mainstream**	**Tributary**	**Mainstream**	**Tributary**	**Mainstream**	**Tributary**	**Mainstream**	**Tributary**
10,000 to 1	1	10	2	25	1.25	50	1.25	100
	10	1	24	1.25	49	1.25	98	1.25
1,000 to 1	2	10	2	25	2	50	2	100
	9	2	25	1.25	50	1.25	100	1.25
100 to 1	5	9	5	25	5	50	5	100
	9	5	25	5	50	5	100	5
10 to 1	9	9	10	25	15	50	25	100
	10	5	25	5	50	10	100	10
1 to 1	9	9	23	23	25	50	50	100
	9	9	23	23	50	25	100	50

TABLE 5.7
Tabular Summary of Joint Flood Frequency for Indiana Confluences (Implementation)

	Frequencies for Coincidental Occurrence							
	10-Year		25-Year		50-Year		100-Year	
Area Ratio	Mainstream	Tributary	Mainstream	Tributary	Mainstream	Tributary	Mainstream	Tributary
10,000 to 1	1	10	2	25	2	50	2	100
	10	1	25	2	50	2	100	2
1,000 to 1	2	10	2	25	2	50	2	100
	10	2	25	2	50	2	100	2
100 to 1	5	10	5	25	5	50	5	100
	10	5	25	5	50	5	100	5
10 to 1	10	10	10	25	10	50	25	100
	10	10	25	10	50	10	100	25
1 to 1	10	10	25	25	25	50	50	100
	10	10	25	25	50	25	100	50

6. CONCLUSION

Accurate calculation of discharges for coincident flooding on confluent streams is key to designing structures near confluences. Given the limitations of the currently used joint probability table (Table 7-3 HEC-22), this study comprehensively evaluates the current methodology, analyses the interdependence of the confluent streams in Indiana and subsequently proposes a new Copula-based joint probability table for implementation in Indiana. The following conclusions are drawn from this study.

1. Kendall's Tau τ (interdependence between mainstream and tributary discharges) can be used to estimate joint probability at confluences, using Gumbel-Hougard Copula.
2. The interdependence (Kendall's Tau τ) of mainstream and tributary discharges largely varies from 0.5 to 0.8. τ decreases with an increase in drainage area ratio (DAR) and stream order difference. This results in lower design discharges for large drainage area ratio confluences.
3. Drainage area ratio is significantly related to τ, but the correlation between DAR and τ can vary across regions within Indiana, with regions near Lake Michigan having the poorest correlation. For most locations in Indiana DAR can be effectively used in the regression equations to determine the interdependence between mainstream and tributary discharges. DAR can then be used to estimate joint probabilities. DAR is easy to estimate and has been previously widely used in joint probability analysis of confluences.
4. The current method using Table 7-3 of HEC-22 results in significantly conservative and overdesigned structures as compared to the Copula-based joint probability table for Indiana.

Two methods are proposed for implementing the findings of this study using the Copula-based joint probability analysis. Method 1 involves using regional regression equations to estimate τ and results in more accurate return period estimation. Method 2, on the other hand, involves using statewide regression equation and provides a summarized joint probability table which is easy and simple to use. It must be noted that Method 2 largely provides similar estimates as Method 1, but the simplifying assumptions and average may sometimes result in slight underestimation of return period compared to Method 1 primarily in Region 2 and 3 (Figure 3.2). For ease of implementation and familiarity among the contractors, the report recommends using Table 5.7. Table 5.7 provides an implementation oriented tabular summary of the Joint Flood Frequency for Indiana specific watersheds.

Although the Copula-based joint probability table improves the accuracy of calculating design discharges for confluent streams, it is important to note that the Copula-based method presented here is based on only using the drainage area ratio to estimate τ, hence the joint probability analysis. For confluent streams, the joint probability of flooding is also based on several hydrometeorological parameters including intensity, duration, spatial extent, location of rainfall, antecedent soil moisture, and changes in land use and land cover in the watershed. This study focused only on statistically analyzing the streamflow and relation with drainage area ratio and did not consider the above-mentioned parameters. A future study involving a regressive analysis of hydrometeorological, and other watershed parameters could enhance the understanding and accuracy of joint probability analysis at confluences.

REFERENCES

Bender, J., Wahl, T., & Jensen, J. (2014). Multivariate design in the presence of non-stationarity. *Journal of Hydrology*, *514*, 123–130. https://doi.org/10.1016/j.jhydrol.2014.04.017

Bender, J., Wahl, T., Müller, A., & Jensen, J. (2016). A multivariate design framework for river confluences. *Hydrological Sciences Journal*, *61*(3), 471–482. https://doi.org/10.1080/02626667.2015.1052816

Brown, S. A., Schall, J. D., Morris, J. L., Stein, S. M., & Warner, J. C. (2009, September). *Urban drainage design manual: Hydraulic engineering circular 22, Third Edition* (Publication No. FHWA-NHI-10-009). Federal Highway Administration. https://www.fhwa.dot.gov/engineering/hydraulics/pubs/10009/10009.pdf de Michele, C., & Salvadori, G. (2003). A generalized Pareto intensity-duration model of storm rainfall exploiting 2-Copulas. *Journal of Geophysical Research: Atmospheres*, *108*(D2). https://doi.org/10.1029/2002JD002534

Favre, A.-C., el Adlouni, S., Perreault, L., Thiémonge, N., & Bobée, B. (2004). Multivariate hydrological frequency analysis using copulas. *Water Resources Research*, *40*(1). https://doi.org/10.1029/2003WR002456

Ganguli, P., & Reddy, M. J. (2013). Probabilistic assessment of flood risks using trivariate copulas. *Theoretical and Applied Climatology*, *111*(1–2), 341–360. https://doi.org/10.1007/s00704-012-0664-4

Genest, C., & Favre, A.-C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, *12*(4), 347–368. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347)

Gilja, G., Ocvirk, E., & Kuspilić, N. (2018). Joint probability analysis of flood hazard at river confluences using bivariate copulas. *Gradevinar*, *70*(4), 267–275. https://doi.org/10.14256/jce.2173.2017

Grimaldi, S., & Serinaldi, F. (2006). Design hyetograph analysis with 3-copula function. *Hydrological Sciences Journal*, *51*(2), 223–238. https://doi.org/10.1623/hysj.51.2.223

Kao, S.-C., & Govindaraju, R. S. (2010). A copula-based joint deficit index for droughts. *Journal of Hydrology*, *380*(1–2), 121–134. https://doi.org/10.1016/j.jhydrol.2009.10.029

Karmakar, S., & Simonovic, S. P. (2008). Bivariate flood frequency analysis: Part 1. Determination of marginals by parametric and nonparametric techniques. *Journal of Flood Risk Management*, *1*(4), 190–200. https://doi.org/10.1111/j.1753-318X.2008.00022.x

Karmakar, S., & Simonovic, S. P. (2009, March). Bivariate flood frequency analysis. Part 2: A copula-based approach with mixed marginal distributions. *Journal of Flood Risk Management*, *2*(1), 32–44. https://doi.org/10.1111/j.1753-318X.2009.01020.x

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, *30*(1–2), 81–93. https://doi.org/10.1093/BIOMET/30.1-2.81

Kilgore, R. T., Thompson, D. B., & Ford, D. T. (2010, September). *Estimating joint probabilities of design coincident flows at stream confluences* (NCHRP REPORT 15-36). National Cooperative Highway Research Program.

Kirby, W., Riggs, H., Sauer, V., & Thomas, W., Jr., (2002). *The national flood frequency program, version 3: A computer program for estimating magnitude and frequency of floods for ungaged sites* (Water-Resource Investigations Report 02-4168). https://pubs.usgs.gov/wri/wri024168/pdf/entirereport.pdf

Kumar, P. (2010). Probability distributions and estimation of Ali-Mikhail-Haq copula. *Applied Mathematical Sciences*, *4*(14), 657–666.

Morris, C. D., & Calise, S. J. (1987). Bivariate analysis of concurrent flooding. In V. P. Singh (Ed.), *Hydrologic Frequency Modeling* (pp. 615–632). Springer Netherlands.

Nathan, P., & David, F. (2004). Interior floodplain flood-damage reduction study. *Journal of Water Resources Planning and Management*, *130*(2), 123–130. https://doi.org/10.1061/(ASCE)0733-9496(2004)130:2(123)

Nelsen, R. B. (2006). *An introduction to copulas* (2nd ed.). Springer. https://doi.org/10.1007/0-387-28678-0

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., & Xia, Y. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research: Atmospheres*, *116*(D12), 12109. https://doi.org/10.1029/2010JD015139

Peng, Y., Chen, K., Yan, H., & Yu, X. (2017). Improving flood-risk analysis for confluence flooding control downstream using copula Monte Carlo method. *Journal of Hydrologic Engineering*, *22*(8), 04017018. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001526

Ramachandra Rao, A., & Srinivas, V. V. (2006, March). Regionalization of watersheds by hybrid-cluster analysis. *Journal of Hydrology*, *318*(1–4), 37–56. https://doi.org/10.1016/J.JHYDROL.2005.06.003

Rao, A. R. (2004). *Regionalization of Indiana watersheds for flood flow predictions phase I: Studies in regionalization of Indiana watersheds* (Joint Transportation Research Program Publication FHWA/IN/JTRP-2002/02). West Lafayette, IN: Purdue University. https://doi.org/10.5703/1288284313303

Raynal, J. A., & Salas, J. D. (1987). A probabilistic model for flooding downstream of the junction of two rivers. In V. P. Singh (Ed.), *Hydrologic frequency modeling: Proceedings of the international symposium on flood frequency and risk analyses* (pp. 595–602).

Sackl, B., & Bergmann, H. (1987). A bivariate flood model and its application. In V. P. Singh (Ed.), *Hydrologic Frequency Modeling* (pp. 571–582). https://doi.org/10.1007/978-94-009-3953-0_40

Salas, F. R., Somos-Valenzuela, M. A., Dugger, A., Maidment, D. R., Gochis, D. J., David, C. H., Yu, W., Ding, D., Clark, E. P., & Noman, N. (2018). Towards real-time continental scale streamflow simulation in continuous and discrete space. *JAWRA: Journal of the American Water Resources Association*, *54*(1), 7–27. https://doi.org/10.1111/1752-1688.12586

Shiau, J. T. (2003). Return period of bivariate distributed extreme hydrological events. *Stochastic Environmental Research and Risk Assessment*, *17*(1), 42–57. https://doi.org/10.1007/s00477-003-0125-9

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut Statistique de l'Université de Paris*, *8*, 229–231.

USACE. (1993, March 5). *Hydrologic frequency analysis* (Engineer Manual 1110-2-1415). U.S Army Corps of Engineers. https://www.publications.usace.army.mil/portals/76/publications/engineermanuals/em_1110-2-1415.pdf

USACE. (2018). *Hydrologic design of interior areas* (Engineer Manual No. 1110-2-1413). U.S. Army Corps of Engineers. https://www.publications.usace.army.mil/Portals/76/Users/227/19/2019/EM_1110-2-1413.pdf?ver=2018-09-13-125346-687

Wahl, T., Mudersbach, C., & Jensen, J. (2012). Assessing the hydrodynamic boundary conditions for risk analyses in coastal areas: a multivariate statistical approach based on Copula functions. *Natural Hazards and Earth System Sciences*, *12*(2), 495–510. https://doi.org/10.5194/nhess-12-495-2012

Wang, C. (2016). A joint probability approach for coincidental flood frequency analysis at ungauged basin confluences. *Natural Hazards*, *82*(3), 1727–1741. https://doi.org/10.1007/s11069-016-2265-5

Wang, C., Chang, N.-B., & Yeh, G.-T. (2009). Copula-based flood frequency (COFF) analysis at the confluences of river systems. *Hydrological Processes*, *23*(10), 1471–1486. https://doi.org/10.1002/HYP.7273

Yang, Z.-L., Niu, G.-Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., Tewari, M., & Xia, Y. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. *Journal of Geophysical Research Atmospheres*, *116*(D12). https://doi.org/10.1029/2010JD015140

Zhang, L., & Singh, V. P. (2006). Bivariate flood frequency analysis using the Copula method. *Journal of Hydrologic Engineering*, *11*(2), 150–164. https://doi.org/10.1061/(ASCE)1084-0699(2006)11:2(150)

APPENDICES

Appendix A. Best Fit Regional Regression Equations for POM and POT

Appendix B. A Quick Summarized Implementation of Method 1 Based on MS Excel

APPENDIX A. BEST FIT REGIONAL REGRESSION EQUATIONS FOR POM AND POT

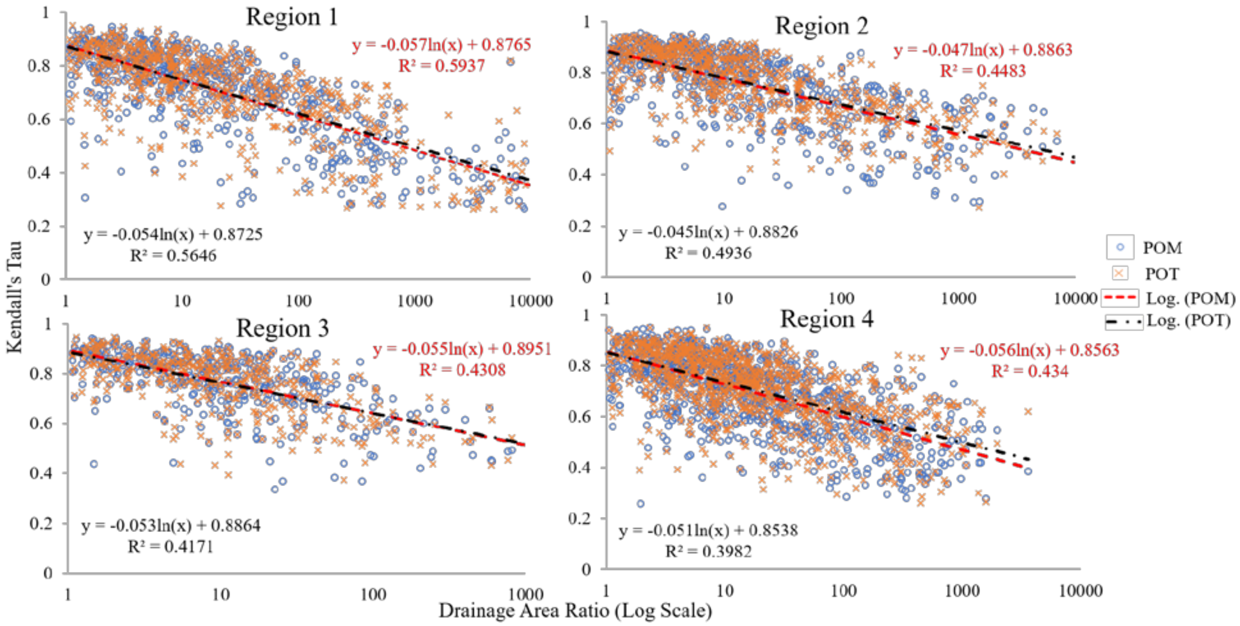


Figure A.1 Best fit regression equation and variation of Kendall's Tau with Drainage Area Ratio for Region 1, 2, 3 and 4.

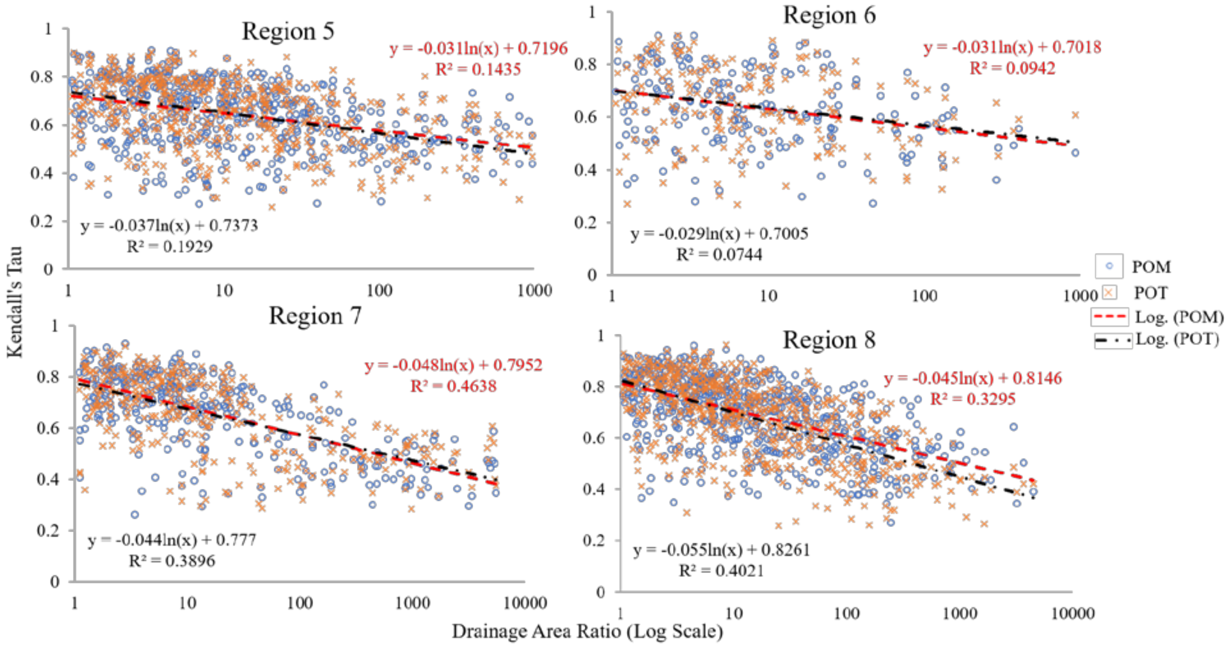


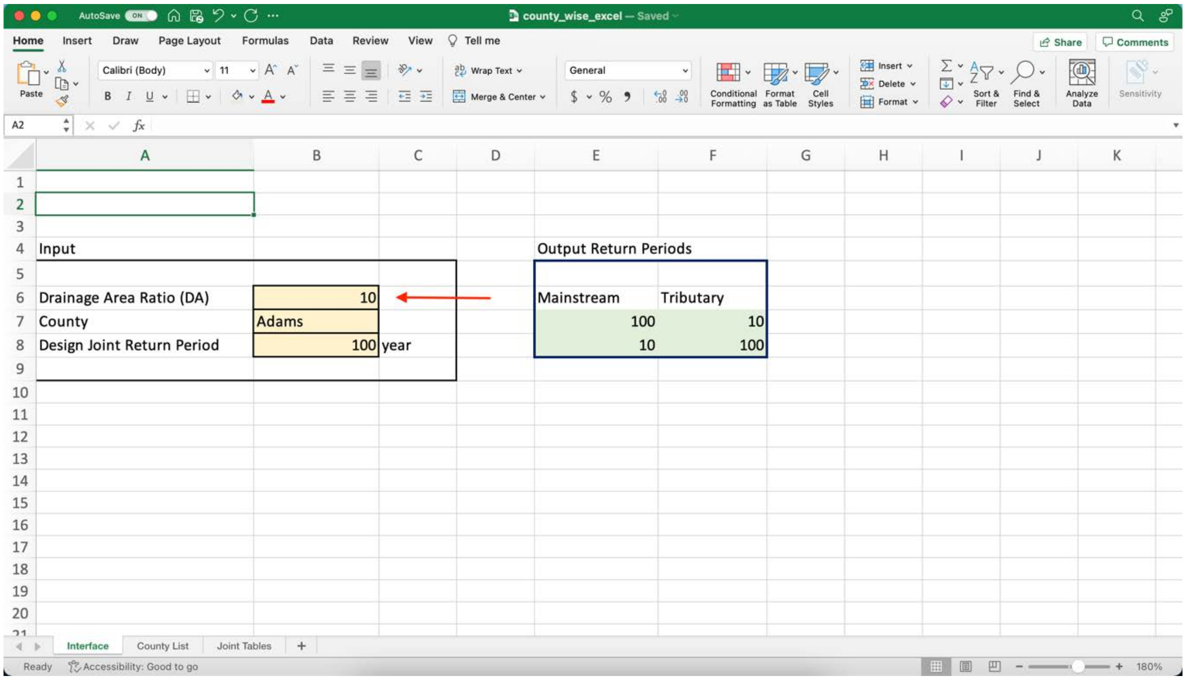
Figure A.1 Best fit regression equation and variation of Kendall's Tau with Drainage Area Ratio for Region 5, 6, 7 and 8.

APPENDIX B. A QUICK SUMMARIZED IMPLEMENTATION OF METHOD 1 BASED ON MS EXCEL

This document is associated with the report “*Multiple Water Course Joint Probability Analysis Procedure Development for Indiana Specific Watersheds*” and presents a simple implementation of Method 1 present in the report using MS Excel. To simplify the application process, the Excel identifies the hydrologically similar region (Figure 3.2 of the report) for all the counties in Indiana. The interface of the excel takes County Name, Drainage Area Ratio, and desired Design return period as input, then estimates the Kendall’s tau τ, rounds it off to the nearest 1 decimal number. The excel then looks up the coincident flow frequencies associated with the rounded of τ and the Design return period and outputs coincident flow frequencies approximated to the nearest 5-year multiple for that county.

The excel file has three spreadsheets named Interface, County List and Joint Tables. The excel sheet is designed in a way that the designer needs to work only with the Interface spreadsheets. The County list spreadsheet provide additional information like the hydrologic region associated with each county, the equation used, the exact τ and approximated τ for the given DAR and county. Joint Tables spreadsheet contains the joint probability tables for different τ and design return period used by the excel.

Step 1. Enter the Drainage Area Ratio in the cell marked by red arrow



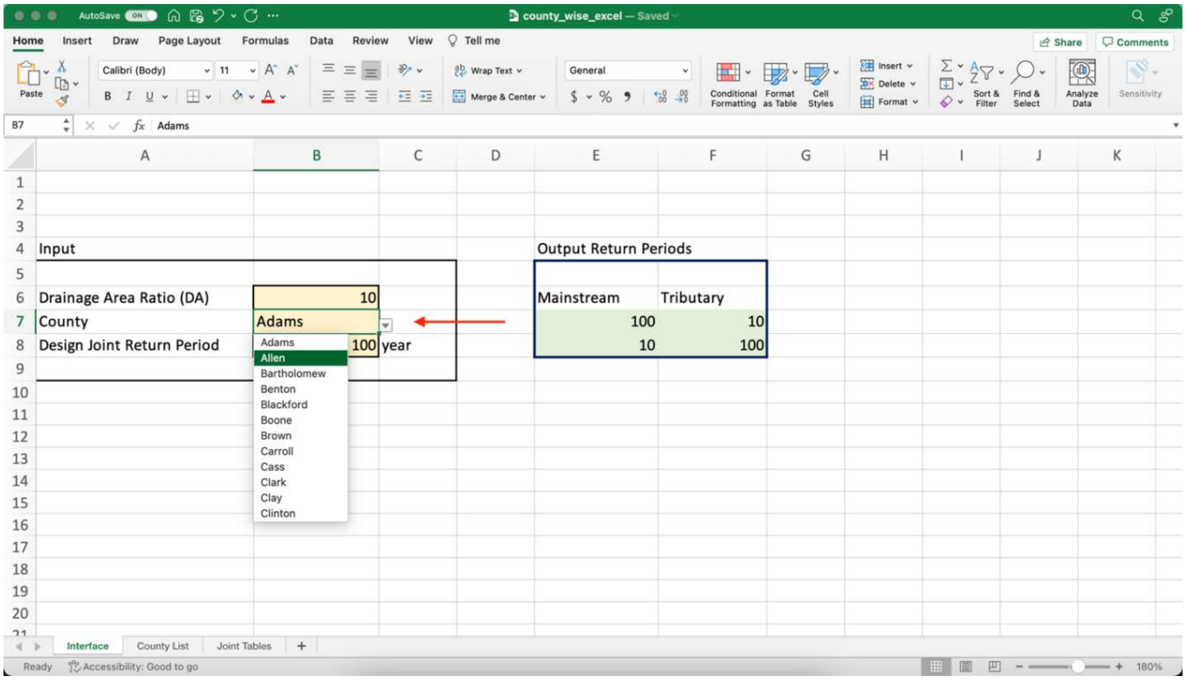
Input

Drainage Area Ratio (DA)	10	
County	Adams	
Design Joint Return Period	100	year

Output Return Periods

Mainstream	Tributary
100	10
10	100

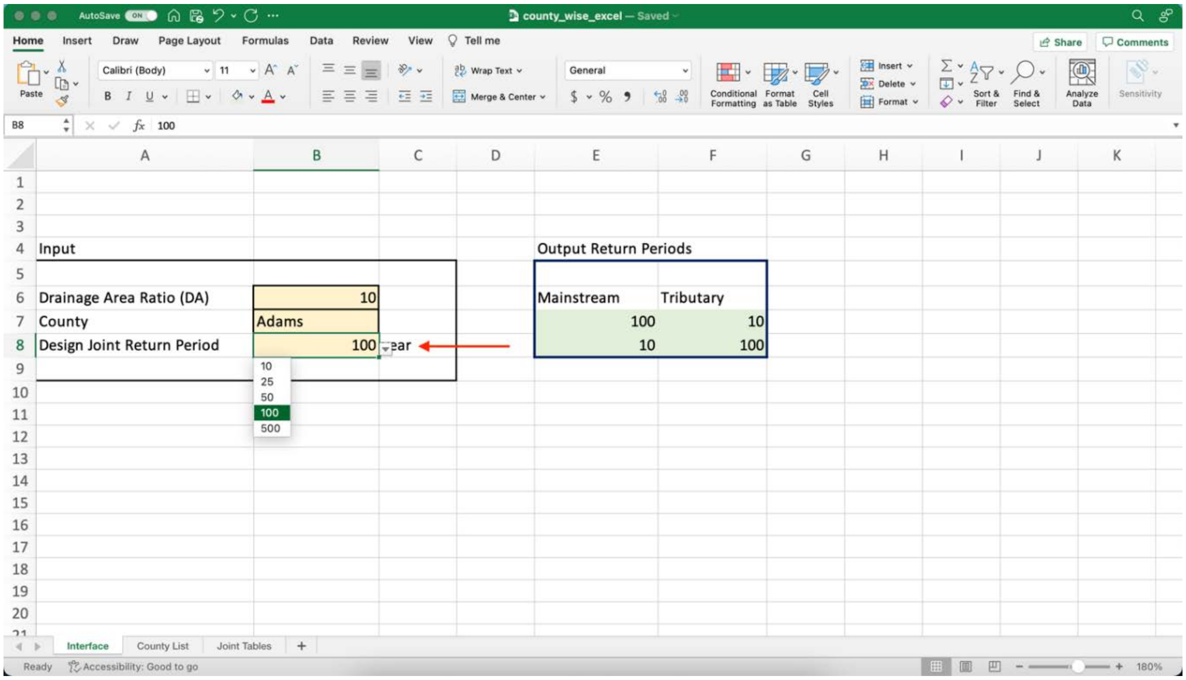
Step 2. Choose the county from the County List (cell marked by red arrow)



B7 Adams

	A	B	C	D	E	F
4	Input				Output Return Periods	
6	Drainage Area Ratio (DA)	10			Mainstream	Tributary
7	County	Adams			100	10
8	Design Joint Return Period	100	year		10	100

Step 3. Choose the Design Joint Return Period from the List (cell marked by red arrow)



B8 100

	A	B	C	D	E	F
4	Input				Output Return Periods	
6	Drainage Area Ratio (DA)	10			Mainstream	Tributary
7	County	Adams			100	10
8	Design Joint Return Period	100	ear		10	100

About the Joint Transportation Research Program (JTRP)

On March 11, 1937, the Indiana Legislature passed an act which authorized the Indiana State Highway Commission to cooperate with and assist Purdue University in developing the best methods of improving and maintaining the highways of the state and the respective counties thereof. That collaborative effort was called the Joint Highway Research Project (JHRP). In 1997 the collaborative venture was renamed as the Joint Transportation Research Program (JTRP) to reflect the state and national efforts to integrate the management and operation of various transportation modes.

The first studies of JHRP were concerned with Test Road No. 1—evaluation of the weathering characteristics of stabilized materials. After World War II, the JHRP program grew substantially and was regularly producing technical reports. Over 1,600 technical reports are now available, published as part of the JHRP and subsequently JTRP collaborative venture between Purdue University and what is now the Indiana Department of Transportation.

Free online access to all reports is provided through a unique collaboration between JTRP and Purdue Libraries. These are available at http://docs.lib.purdue.edu/jtrp.

Further information about JTRP and its current research program is available at http://www.purdue.edu/jtrp.

About This Report

An open access version of this publication is available online. See the URL in the citation below.

Ghanghas, A., Dey, S., & Merwade, V. (2023). *Development of a multiple water course joint probability analysis procedure for Indiana watersheds* (Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2023/09). West Lafayette, IN: Purdue University. https://doi.org/10.5703/1288284317616