Jane, R.; Wahl, T.; Malagon Santos, V.; Misra, S.K.; White, K.D. (2022). Assessing the potential for compound storm surge and extreme river discharge events at the catchment scale with statistical models: sensitivity analysis and recommendations for best practice. *J. Hydrol. Eng. 27(3)*.

Published version: https://dx.doi.org/10.1061/(asce)he.1943-5584.0002154

NIOZ Repository: http://imis.nioz.nl/imis.php?module=ref&refid=352877

Assessing the potential for compound storm surge and extreme river discharge events at the catchment scale with statistical models: sensitivity analysis and recommendations for best-practice

Robert Jane[1*]; Thomas Wahl[2]; Victor M. Santos[3]; Shubhra K. Misra[4]; Kathleen D. White[5]

[1]Postdoctoral Researcher, Department of Civil, Environmental and Construction Engineering, and National Center for Integrated Coastal Research, University of Central Florida, 12800 Pegasus Drive, Suite 211, Orlando, FL 32816-2450, USA. *Corresponding author r.jane@ucf.edu

[2]Assistant Professor, Department of Civil, Environmental and Construction Engineering, and National Center for Integrated Coastal Research, University of Central Florida, 12800 Pegasus Drive, Suite 211, Orlando, FL 32816-2450, USA.

[3]Ph.D. Student, Department of Civil, Environmental and Construction Engineering, and National Center for Integrated Coastal Research, University of Central Florida, 12800 Pegasus Drive, Suite 211, Orlando, FL 32816-2450, USA.

[4]Hydraulic Engineer, Coastal Engineering Section, United States Army Corps of Engineers Galveston District, 2000 Fort Point Road, Galveston, TX 77550, USA.

[5]Lead, Climate Preparedness and Resilience Community of Practice, United States Army Corps of Engineers, Headquarters, Engineering and Construction Directorate, 441 G Street NW, Washington, DC 20001, USA.

Abstract

Two-sided extreme conditional sampling is regularly coupled with copula theory to assess the dependence between flood-risk drivers such as extreme precipitation or river discharge, and storm surge. The approach involves many subjective choices. Choices include sampling techniques used to identify extreme events (Block maxima or Peaks-over-Threshold; POT), whether to account for the fit of marginal distributions, and time-lags considered between the two drivers. In this study, estimates of the potential for compound events at three sites along the Texan Gulf Coast, where the U.S. Army Corps of Engineers (USACE) is undertaking Coastal Storm Risk Management (CSRM) projects, are shown to be highly sensitivity to model set-up. A pragmatic approach accounting for marginal fit in a POT framework is proposed and shown to

provide stable estimates of the compounding potential for high discharge and storm surge events. -We also explore the effect of using precipitation as a proxy for discharge in the absence of sufficiently long discharge records.

Keywords (up to 8 alphabetical order): Compound flooding, Coastal flood risk, Copulas, Extreme value analysis, Gulf of Mexico, Multivariate statistical modelling, Sensitivity analysis.

1. Introduction

Flooding in coastal catchments is often a result of interacting meteorological, hydrological, and oceanographic phenomena such as heavy precipitation, a driver of increased river discharge, storm surge and waves. The co-occurrence of multiple phenomena – a compound event – may exacerbate the impact of a flood event compared to when individual phenomena occur in isolation. Many of the phenomena are driven by common meteorological forcing. For instance, low pressure systems can produce heavy rainfall and also strong onshore winds and an inverse barometric effect able to generate large storm surges. More recently, hotspots for coincident high wind and heavy rainfall at midlatitudes have been linked to regions with a high occurrence frequency of atmospheric rivers (Ridder et al., 2020). Neglecting any statistical dependence between the phenomena causing flooding in coastal catchments may lead to under design of flood defense structures and underestimation of flood risk (Moftakhari et al., 2019; Gori et al. 2020).

River discharge and storm surge often interact nonlinearly. Consequently, separate numerical modeling of the phenomena often results in the underestimation of flooding characteristics (e.g., Kumbier et al., 2018; Loveland et al., 2020). To account for the interactions, hydraulic models are often linked (one way) with hydrologic and ocean circulation models, often with time varying boundary condition (e.g., Gori et al., 2020). Similarly, linked hydrologic models (e.g., Joyce et al., 2018) and ocean circulation models (e.g., Valle-Levinson et al., 2020) are also prevalent in the literature. The latter neglect the non-linear interaction between surge and rainfall-runoff while linked hydraulic and linked ocean circulation models do not account for the direct effect of precipitation over the model domain potentially underestimating water levels (Santiago-Collazo et al., 2019). Wing et al. (2019) recently coupled a hydraulic model with NOAA forecasts of streamflow, rainfall, and coastal surge height to generate medium term forecast of flood inundation for Hurricane Harvey. The hydraulic model used rainfall and surge information to generate new event specific depth grids of pluvial and coastal flooding accounting for the dynamics of the surge as it interacts with the shoreline and moves inland. Despite more complex linking techniques of loosely coupled models (e.g., Blanton et al., 2020), where information between separate models is exchanged iteratively, and tightly coupled models (e.g., Lee et al., 2020) involving the integration of source code limitations remain. Nevertheless, continued advancements in computer technology and numerical modeling mean the development of comprehensive compound inundation models considering all pertinent physical processes and their interactions is becoming an increasingly realistic prospect (Santiago-Collazo et al., 2019; Bakhtyar et al., 2020).

Assessments conducted at the global (e.g., Ward et al., 2018; Couasnon et al., 2020) and national scale (e.g., Wahl et al., 2015) demonstrated that although regional patterns can be discerned, exposure to compound flooding depends on localized factors such as catchment size and morphology. A prudent first step in assessing and managing flood risk is to therefore assess the propensity for coincident high discharge and storm surge at the catchment scale to prioritize areas for future development of compound inundation models.

The Sabine Pass to Galveston Bay Coastal Storm Risk Management project (S2G-CSRM) involves three separate sub-projects in southeast Texas: one covers the lower Brazos River Basin (Freeport) and the other two cover the Sabine and Neches River Basins (Orange County and Port Arthur). Hurricane Harvey, which made landfall approx. $195\ km$ east of Freeport in August 2017, is one of several tropical storms that caused costly compound flood events in the eastern U.S. in recent years (NOAA, 2020). While Harvey stalled with its center over or near the Texas coast for four days, heavy-rain-producing convective systems including high precipitation efficiency supercells in its outer rain bands developed repeatedly, leading to record breaking rainfall totals and streamflow across southeastern Texas (Brauer et al., 2020; Watson et al., 2020). The meteorological situation has been associated with previous tropical cyclones in the region (Blake and Zelinsky, 2018). The synoptic conditions combined with the morphology of the Houston-Galveston Bay region, caused the phase-lagged rainfall-related land-derived discharge from two Bayous to coincide with storm surge exacerbating flooding around Houston (Valle-Levinson et al., 2020).

Two-sided conditional sampling has become an established technique for identifying bivariate extremes in compound flooding studies. The technique involves pairing the (detrended, if necessary) declustered excesses using a Peaks over Threshold (POT) approach or block maximum (BM) of one flooding driver with the maximum value of another within a lag of $\pm\ \Delta$ days. Wahl et al. (2015) applied the approach to hourly storm surge data (tidal and mean sea level (MSL) influence removed) from 30 tide gauges (TGs) around the U.S. and daily precipitation totals derived from stations within 25 *km* of these TGs. In the study, annual maximum storm surges are paired with the highest precipitation within a three-day window, i.e. within ±1 day from the timing of maximum surge, and vice versa. A similar study was undertaken by Ward et al. (2018) using observed sea levels (and skew surge) and river discharge data for 187 station combinations across the globe. To identify marginal extremes, Ward et al. (2018) used the annual maximum approach and time-lags between -5 and +5 days. Recently, Santos et al. (2020) assessed the potential for compound flooding around Sabine Lake by applying the conditional sampling approach to the coastal water level at Sabine Pass and discharge of Sabine and Neches rivers draining into the lake. Based on a correlation analysis over a range of lags and POT thresholds, the analysis was implemented without applying a lag when conditioning on surge (COS) while a lag of between -5 and +5 days was adopted when conditioning on discharge (COD). The lags applied in Santos et al. (2020) reflect an impact-based observation: while water level extremes (the impact) in Sabine Lake usually coincide with coastal surge maxima, discharge maxima often occur with significant lags.

To date, estimates of the degree of dependence between extreme discharge and storm surge along the U.S. Gulf coast vary primarily depending on whether discharge or precipitation (as a proxy for discharge) is used to represent the freshwater contribution. In Wahl et al. (2015) almost all sites in the region exhibited significant correlation regardless of the conditioning variable. Conversely, Ward et al. (2018) found the COS sample from the Sabine River to be only one of two out of seven discharge station - tide gauge combinations along the Gulf coast to exhibit significant correlation ($0.3<\tau<0.4$) between discharge and skew surge. The lack of significant correlation in the region was attributed to large catchment sizes and the highest surges typically occurring during hurricane season, when the discharge of the region's rivers are at their lowest. Santos et al. (2020) found significant correlation between combined discharge and surge over a range of lags for both conditional samples. The variation in Ward et al. (2018) and Santos et al. (2020) findings are mainly due to lags being accounted for differently.

Subjective choices are an inherent part of the model set up in any extreme value analysis. Choices include the detrending and sampling approaches used to identify marginal extremes and their associated parameters (Arns et al., 2013). Ward et al. (2018) explored the sensitivity of the dependence observed between discharge and skew surge to the sampling method by repeating the analysis using a POT approach. A 0.99 (quantile) threshold produced results consistent with using annual maxima, while a 0.95 threshold yielded lower correlation coefficients and fewer significant correlations. In contrast with the general trend for the 0.95 threshold, significant correlation was found for the COD sample at the station combinations involving both the Sabine River and Brazos River. Camus et al. (2020) adopted a two-sided conditional sampling approach to analyze the pairwise relationships between four flooding drivers – precipitation, discharge, significant wave height and a combined storm surge and wave component – around the coasts of the Eastern North Atlantic Ocean, Mediterranean Sea and Black Sea. Linear trends were observed in the correlation coefficients associated with conditional samples where POT approaches with thresholds generating an average of either three or six events per year and an annual maximum sampling approach were used to derive marginal extremes. The latter typically led to the largest coefficients. Using monthly maximum, Santos et al. (2020) reached similar conclusions regarding the potential for compound events as a POT approach where the threshold is set to approximately three events per year.

In bivariate and higher dimensional settings, time-lag(s) considered between variables represents an additional subjective choice. For instance, Ward et al. (2018) only detected significant correlation between annual maximum discharge and skew surge at the station combination involving the Sabine River when applying a lag, with the correlation peaking at a 5-day lag. Further afield, Camus et al. (2020) reported negligible differences in the highest pairwise correlations between four flooding drivers when extending the maximum lag from 3- to 10-days.

After implementing two-sided conditional sampling, copula theory is typically invoked to quantify the nature of the dependence structure within the two conditional samples. Ward et al. (2018) fitted three Archimedean copulas - Gumbel, Clayton, and Frank- which possess upper, lower, and no tail dependence, respectively, to station combinations exhibiting significant correlation between annual maximum discharge and total sea level. According to two goodness-of-fit criteria, the Clayton copula was the most suitable for modeling dependence at the station combination involving the Sabine River. In addition to the three Archimedean copulas used in Ward et al. (2018), Wahl et al. (2015) also applied the Galambos and Hüsler-Reiss extreme value copulas to the conditional samples. The Gumbel copula was deemed the best fitting copula for both samples at Galveston. Santos et al. (2020) fit a trivariate vine copula to capture the dependence between the surge data at Sabine Pass and discharge of the Sabine and Neches Rivers utilizing the VineCopula R package (Schepsmeier et al., 2015). The pair copula composing the vine were selected based on the Akaike information criterion (AIC), a goodness of fit measure which rewards the goodness of fit while penalizing each additional parameter thereby discouraging overfitting. Detailed analysis was presented for the sample derived when conditioning on the Sabine River discharge as it was the case that implied the strongest compounding effects. A vine copula composed of two Frank copulas and the Joe copula was deemed the best fit for both the POT and BM approaches.

In this paper, we aim to provide robust estimates of the potential for extreme river discharge-storm surge events in three watersheds. This is achieved by meeting three objectives. The first objective is to gauge the relative potential for compound events across the three case study sites. The estimates are likely to be influenced by the subjective choices required in the model set-up. The second objective is therefore to assess the sensitivity of the estimates to these choices in order to propose a robust best practice approach that attempts to mitigate the subjectivity. The absence of sufficiently long discharge time series often prevents a robust statistical analysis; hence in practice precipitation is commonly adopted as a proxy for discharge. The final objective is to investigate the effect of using precipitation as a proxy for discharge on estimates of the potential for compound events.

2. Case Study site & data

The S2G-CSRM project involves upgrades to existing Hurricane Flood Protection Systems (HFPS) at Freeport and Port Arthur and the creation of a new CSRM system for Orange county. The Orange county CSRM project requires the construction of two gate structures in Cow and Adams bayous, adjacent watersheds in the roughly crescent shaped Sabine Basin (Figure 1) with an area of approx. 515 and $96\ km^2$, respectively. The main channels of both bayous are slow moving streams which flow into the Sabine River a few kilometers upstream of Sabine Lake. The main streams in Cow and Adams bayous are approx. 37 and $28\ km$ long and tidally influenced for around 32 and 13 km, respectively, above their confluences with the Sabine River (Parsons, 2006). The lower reaches of both streams have been channelized, straightened, and dredged for navigational purposes, leading to oxbows replacing formerly more sinuous channels. The two bayous drain approximately half of Orange County where the National Flood Insurance Program's flood claims database totals over $\$125$ million in claims between 1979 and 2008 (OCDD & TWDB, 2015). The claims arise due to both prolonged or successive storms that produced heavy rainfall such as Hurricane Rita in Sept. 2006 and storm surge dominated events like Hurricane Ike in Sept. 2008.

The Brazos River extends from its headwaters at the foot of the south plains in eastern New Mexico, east and then southeast before draining into the Gulf of Mexico in the marshes south of Freeport. The Brazos River watershed encompasses approximately $109{,}000\ km^2$ and has been fully regulated since 1980, containing 16 major and 13 smaller regional water supply reservoirs (Brazos BBEST, 2012). Although there are no reservoirs on the mainstem of the middle and lower Brazos River, construction and operation of those in the basin likely affects the hydrology in this portion of the river (TIFP & BRA, 2010). The lower Brazos River is a large lowland meandering system with approximately $100{,}000$ residence and $51{,}000$ insurable structures worth $\$19.4$bn located in the 1% annual exceedance floodplain (BRA, 2019).

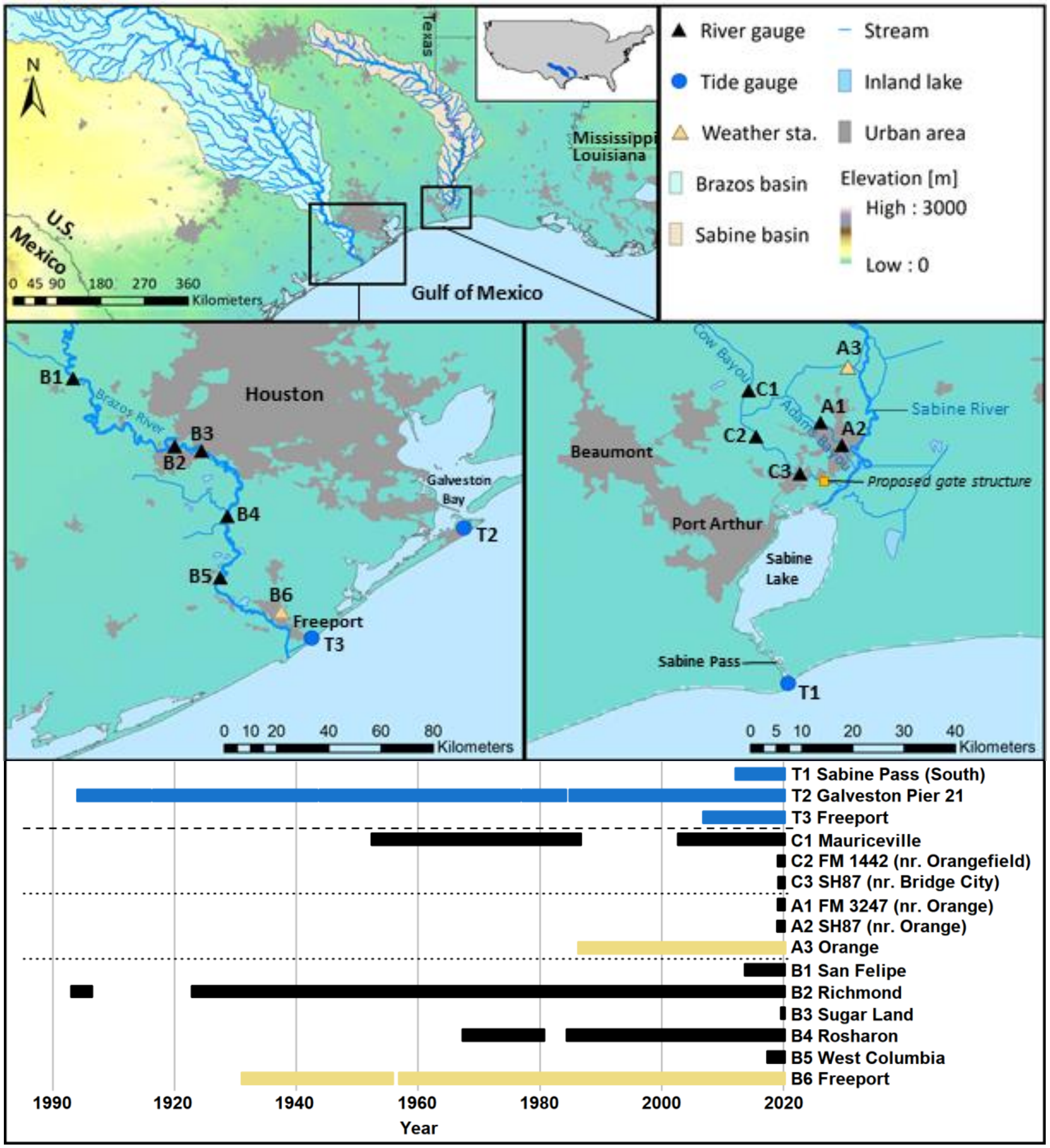


Figure 1: Location of study sites and data completeness. Top row: Location of the Sabine and Brazos River Basins in the western portion of the northern Gulf of Mexico coast and in terms of the entire U.S. in the inset. Middle row: Position of tide gauges, streamflow gauges, and weather stations in the Brazos River Basin (left) and Sabine River Basin (right). Bottom row: Completeness of the records at the gauges and stations.

The data utilized in this study consists solely of in-situ observations. Gauges measuring daily average river discharge are selected based on record duration, completeness, and proximity to the mouth of the river or bayou; see Figure 1. The discharge in Cow Bayou is represented by the gauge near Mauriceville (USGS 08031000). The streamflow records in Adams Bayou are too short to form part of a robust assessment. Daily precipitation totals from the weather station at Orange (USC 00416680) obtained from the National Oceanic and Atmospheric Administration (NOAA) National Climatic Data Center (https://www.ncdc.noaa.gov/) were adopted as a proxy for discharge. The agreement of the precipitation at Orange and nearest stage shown in Figure SM.1 in the Supplementary Material alludes to the varying temporal scales of the two processes. The streamflow gauge at Richmond (USGS 0811400) and another further upstream at Rosharon (USGS 08116650), shown in Figure SM.2, are utilized to investigate the influence of the coastal water level along the Brazos River. To test the effect of using precipitation as a proxy for discharge, the analysis undertaken in the Brazos River is repeated with daily precipitation totals replacing daily average river discharge. More specifically, daily precipitation totals from the NOAA National Climatic Data Center's weather station at Freeport (USC 00413340) were effectively adopted as a proxy for the discharge at both Richmond and Rosharon.

Coastal water levels are obtained from NOAA tide gauges (https://tidesandcurrents.noaa.gov). The water level is composed of the mean sea level, astronomical tide, and non-tidal residual (NTR). The NTR primarily comprises storm surge signal resulting from atmospheric pressure variability and wind setup as well as inter decadal/inter annual trends, seasonal cycles, and river discharge (Serafin and Ruggiero, 2014). It is thus the meteorologically driven component of the water level, i.e. that which will potentially exhibit dependence with hydrological processes and hence the component upon which our analysis is undertaken. The NTR is referred to simply as surge hereafter. The coastal water level is represented by the Sabine Pass (South) TG in the Sabine Basin and the Freeport TG in the Brazos Basin. The TG records are considerably shorter than the discharge/precipitation records but are extended and missing values filled-in by linear regression on the longer Galveston Pier 21 TG record (Figures SM.3-SM.4).

3. Methodology

To assess compound flooding potential, an approach similar to the bivariate analysis in Jane et al. (2020) and summarized in Figure 2 is employed. The statistical models in the approach require Independent and Identically Distributed (IID) random variables, i.e. the statistical parameters such as mean and variance should remain constant over time and free of “trends, shifts, or periodicity” (Salas, 1993). The first step of the methodology is therefore to detrend the data. The surge time series are detrended using a linear fit and the seasonality removed by applying a 3-month moving average window. The behavior of hydrological processes may change abruptly due to distributional or structural changes, principally through human intervention (e.g., due to land use changes in the catchment), or more gradually, for instance due to climate change (Wu et al., 2020). To prevent change points being misinterpreted as trends and vice versa, it is essential that they are analyzed together (Sharma et al., 2016). The Pettitt test (Pettitt, 1979) for detecting change points is commonly used alongside the Mann–Kendall test for monotonic trends in hydrological studies (e.g., Rahmani et al., 2015; Tosunoglu et al., 2020). No change points or monotonic trends were detected in the annual maximum discharge/precipitation records in the catchments, see Figures SM.5-SM.9.

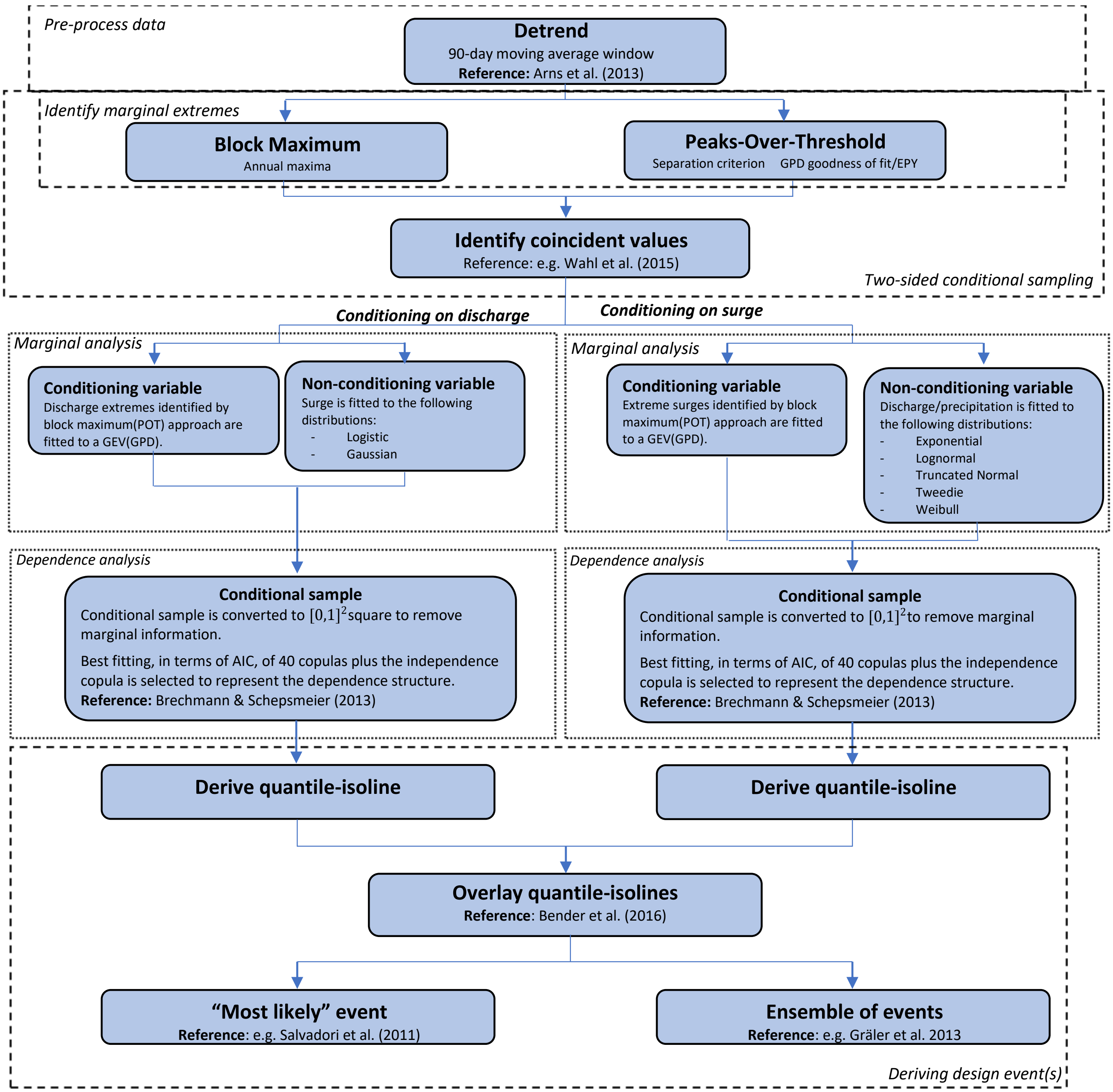


Figure 2: Flow chart of methodology in Section 2.2 of Jane et al. (2020).

The next step of the methodology is to identify marginal extremes. The two most commonly applied methods for identifying (univariate) extremes are BM and POT. For a sequence of IID random variables, if the linearly normalized BM converge in distribution, then the limit must be a Generalized Extreme Value (GEV) distribution (Gnedenko, 1943). Under similar conditions, if there is a non-degenerate limiting distribution for excesses above a high threshold, Picklands (1975) and Balkema and de Haan (1974) showed it will be a Generalized Pareto Distribution (GPD). To identify independent marginal excesses above a threshold using a POT approach, Ward et al. (2018) used a storm window approach pairing excesses within a specified time window. However, results were highly sensitive to whether a 3-, 5- or 7-day window was adopted. In this work a separation criterion of 3-days is used for lags of 2 days or less, otherwise the separation criterion is set to twice the lag. This ensures events remain independent once a lag is applied in the two-sided conditional sampling (e.g., each surge is only paired with one discharge/precipitation value and vice versa). Several truncated parametric non-extreme value marginal distributions are fitted to the discharge/precipitation observation in the COS samples, namely the Exponential, Log-normal, Truncated normal, and Tweedie distributions. Similarly, the Logistic and Gaussian, both non-truncated non-extreme value parametric distributions, are fitted to the surge observations in the COD samples.

Next, a conditional sample is derived by pairing the marginal extremes with the maximum observed value of the other variable within a predetermined lag. To ensure conservative estimates the lag must respect the temporal scale of extreme surge and extreme discharge events. Based on the results of the correlation analysis and results in Santos et al. (2020), lags of ± 5 days were applied to surge when COD and a lag of ± 1 day to discharge when COS.

When adopting a POT approach, the threshold above which (declustered) excesses form part of the conditional sample also needs to be selected. The choice of threshold represents a trade-off between ensuring the asymptotic assumptions of the statistical models are not violated whilst retaining enough observations to prevent biased parameter estimates. This study adopts a pragmatic approach where the goodness-of-fit of the marginal distribution is used to inform the threshold used to derive the conditional samples. More specifically, the normalized root mean squared error of the GPD residuals ($NRMSE_{GPD}$) for a (conditional) sample containing more than 30 events was required to be less than 0.3. This rather simplistic approach to threshold selection could be replaced by more sophisticated thresholds selection techniques in the future (Northrop et al., 2017; Solari et al., 2017). Thresholds used in the marginal and dependence analyses to derive the conditional samples can differ. A preliminary analysis found that using a higher marginal threshold after obtaining conditional samples only offered limited improvements in the marginal fit. Consequently, a GPD was fit to all the observations of the conditioning variable above the threshold used to derive the conditional sample. The threshold with a $NRMSE_{GPD}$ of less than 0.3 that maximizes the correlation between discharge/precipitation and surge is selected in the proposed approach (see Case 1 in Table 1).

Table 1: Cases propagated through the full analysis outlined in Figure 2.

Case	**Lag (COS)**	**Lag (COD)**	**Marginal Sampling**	**Identifying (quantile) threshold**	**Correlation**
1	± 7	± 1	POT	$<0.3 NRMSE_{GPD}$	Maximum
1+	± 7	± 1	POT	$<0.3 NRMSE_{GPD}$	Maximum with UTD*
2	± 7	± 1	POT		Maximum
3	± 5	± 1	POT	$<0.3 NRMSE_{GPD}$	Maximum
4	± 3	± 1	POT	$<0.3 NRMSE_{GPD}$	Maximum
5	± 7	± 1	POT	1 EPY	Maximum
6	± 7	± 1	POT	3 EPY	Maximum
7	0	0	POT	$<0.3 NRMSE_{GPD}$	Maximum
8	± 7	± 1	Block Maxima		

*UTD - Upper Tail Dependence

To gauge the sensitivity of the results to the subjective choices in the model setup several other test cases, summarized in Table 1, are propagated through the full analysis. Case 1 is designed to assess the effect of selecting the copula family that maximizes the correlation on occasions where copulas with and without UTD satisfy the marginal fit criterion. In Case 1+ the threshold which maximizes the correlation between the drivers and satisfies the marginal fit criterion but also leads to a copula with UTD being selected is adopted. Case 1+ is thus not run if Case 1 selects a copula with UTD, or if no copulas satisfying the marginal fit criterion possess UTD. Three cases similar to the proposed approach but that do not use the fit of the GPD to decide the appropriate (quantile) thresholds are also analyzed. The first of these cases chooses the threshold that maximizes the correlation (Case 2) while the other two adopt the threshold yielding on average 1 and 3 Events Per Year (EPY) (Cases 5 and 6). Other test cases involve reducing the lag when COS to ± 5 and ± 3 days (Cases 3 and 4) or not using any lag when conditioning on either variable (Case 7). In the final test case (Cases 8), a BM approach is implemented, eliminating the need to define a threshold. In Case 8, years with less than 80% complete records are omitted.

In terms of the dependence modeling, the best fitting of 40 copulas is subsequently identified for each sample from which the joint return period (RP) contours (given the RP definition) and design events can be obtained. In the univariate setting, return period and thus design events are uniquely defined. In contrast, in the multidimensional domain there are a plethora of equally mathematically valid definitions of RP each with perceived strengths and limitations (e.g. Salvadori et al., 2011). The “AND” definition was recently recommended for estimating the joint return period of river discharge and ocean levels in the FEMA (2015) procedure for assessing compound flood hazard in tidal channels and estuaries (Moftakhari et al., 2019). The "AND" RP definition is hence exclusively adopted in this study. Each multivariate RP definition relates to a specific underlying probability structure. Hence, comparisons between the results obtained using the different definitions are not meaningful (Serinaldi et al., 2015).

In the bivariate domain, isolines contain an infinite number of possible design events. Bender et al. (2016) proposed an approach for deriving the "AND" RP isoline when implementing two-sided conditional sampling. A single design event is typically obtained using the "most likely" strategy where the design event is defined as the event on the isoline with the maximum (relative) probability density given the observed data. Jane et al. (2020) found that estimated probability density along the isoline was highly sensitive to a few observed events, and therefore simulated a large sample of events from the fitted copulas, in proportion with the empirical data. The joint distribution was then estimated non-parametrically by applying a Kernel Density Estimator (KDE) to the simulated sample, using the *ks* R package (Duong, 2007), before the probabilities associated with the events on the isoline were extracted and scaled to lie in [0,1]. Salavadori et al. (2011) and Gräler et al. (2013) among others extol the benefits of considering an ensemble rather than a single design event, in conjunction with expert opinion on the interaction of the phenomenon and infrastructure. In addition to implementing the approach in Jane et al. (2020) to derive the single most likely "design event", ensembles of 1,000 design events are also generated here by sampling along the isoline.

4. Results

4.1 Correlation analysis and copula selection

Statistically significant correlation in terms of Kendall's rank correlation coefficient τ (i.e. testing the null hypothesis $H_0: \tau = 0$) is detected for all samples, see Figures SM10-SM.17. The magnitude of any positive correlation and proportion of thresholds/lag combinations yielding statistically significant correlation is greater in the Sabine region than in the Brazos Basin. Despite its location further downstream, due to the shorter discharge record at Rosharon, any positive correlation is generally smaller and less likely to be statistically significant than for the corresponding threshold-lag combination at Richmond. In both catchments when COS the correlation generally increases with lag and as the quantile threshold is lowered, while for higher thresholds (>0.99) correlation generally reduces as the lag is increased. In the COD sample at Cow Bayou and Richmond the likelihood of observing statistically significant positive correlation increases with lag and as threshold decreases. At Adams Bayou, significant correlation is detected in the sample conditioning on precipitation for thresholds less than 0.98 and around the 0.99 threshold independent of lag. Significant positive correlation is detected sporadically in the COD sample at Rosharon, typically at lower lags.

The Tweedie(Logistic) distribution was found to be the best fitting in terms of AIC for the non-conditioned variable in the COS(COD) samples. For plots demonstrating the goodness-of-fit of the candidate marginal distributions and GPD at all the sites under the proposed approach, see Figures SM.25-SM40.

Table 2: Properties of the conditional samples derived at the study sites for the nine test cases outlined in Table 1. Kendall's τ correlation coefficient, size of the sample (both raw (N) and in terms of events per year (EPY)), and the best fitting copula are provided. Bold font indicates significant Kendall's τ coefficients and copulas with UTD.

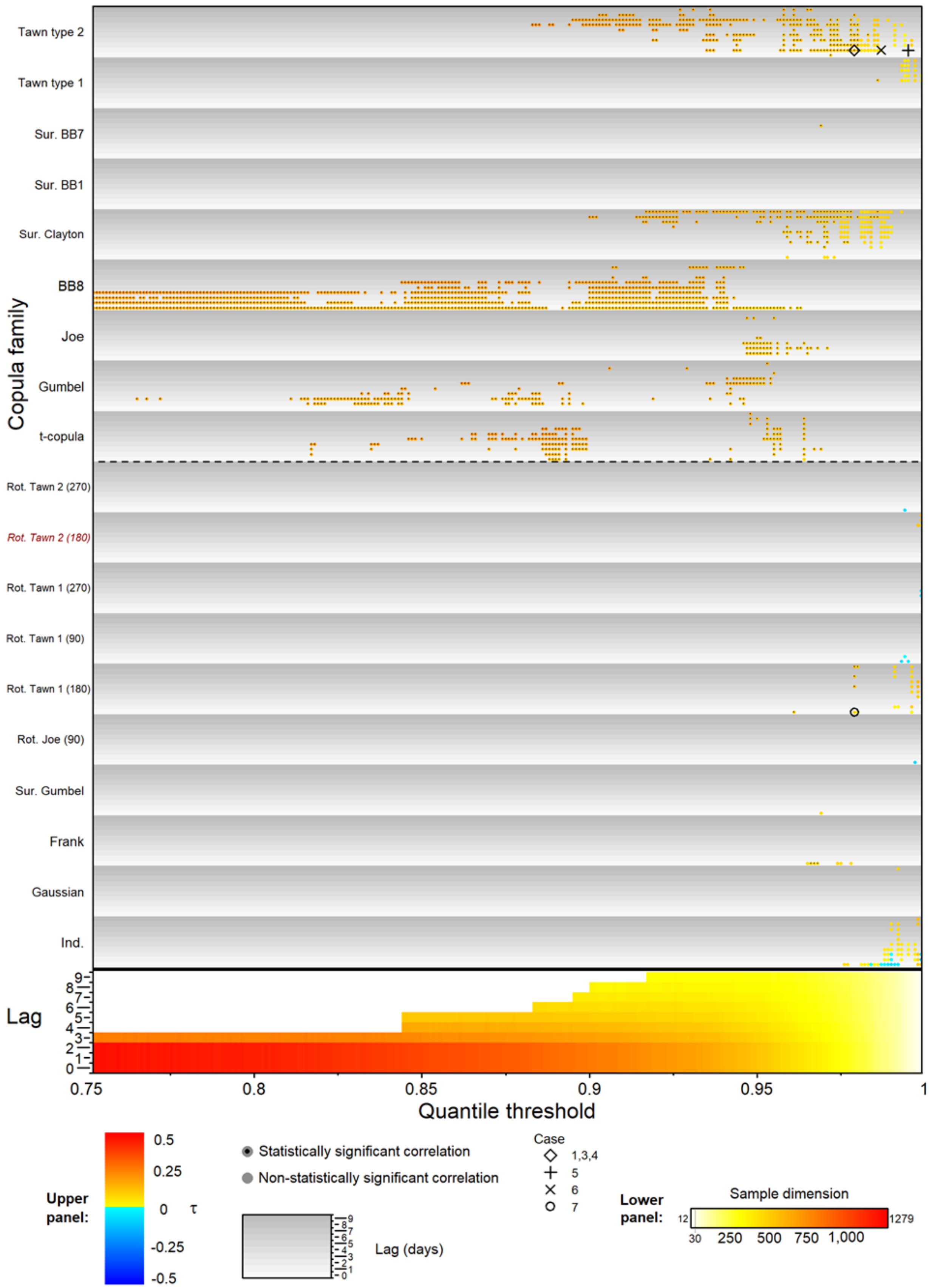
	Cow Bayou								Adams Bayou							
	COS				COD				COS				COD			
Case	τ	Copula	N	EPY	T	Copula	N	EPY	τ	Copula	N	EPY	τ	Copula	N	EPY
1	**0.112**	**Tawn 2**	265	4.81	**0.148**	Sur. Gum.	154	3.05	**0.223**	**t-cop**	128	3.44	**0.197**	**Gumbel**	77	2.44
1+	**0.112**	**Tawn 2**	265	4.81	**0.136**	**Gumbel**	209	4.13								
2	**0.121**	**BB8**	695	13.01	**0.202**	Gaus.	222	4.45	**0.260**	Frank	843	24.12	**0.197**	**Gumbel**	77	2.44
3	**0.112**	**Tawn 2**	265	4.81	**0.110**	Gaus.	157	3.41	**0.223**	**t-cop**	128	3.44	**0.192**	Gaus.	62	1.89
4	**0.112**	**Tawn 2**	265	4.81	0.113	Gaus.	114	2.22	**0.223**	**t-cop**	128	3.44	**0.213**	Gaus.	81	2.56
5	0.009	**Tawn 2**	66	1.00	-0.025	**t-cop**	50	1.00	0.111	Ind.	45	1.15	-0.023	Ind.	30	0.98
6	0.027	**Tawn 2**	167	3.00	**0.137**	Gaus.	153	3.00	**0.147**	**t-cop**	118	2.97	0.107	**Joe**	96	2.99
7	0.075	Tawn 1 (180)	265	4.81	0.147	**Joe**	31	0.61	**0.161**	Frank	158	4.35	**0.209**	Gaus	948	28.92
8	0.015	Tawn 2 (180)	49	1.00	-0.006	Ind.	49	1.00	0.149	**t-cop**	33	1.00	-0.222	Ind.	33	1.00

	Richmond								Rosharon							
	COS				COD				COS				COD			
Case	τ	Copula	N	EPY	T	Copula	N	EPY	τ	Copula	N	EPY	τ	Copula	N	EPY
1	0.035	Sur. Joe	257	2.56	**0.105**	**Tawn 1**	269	2.75	0.082	Frank	218	4.13	0.100	Sur. Joe	115	2.37
1+	**0.075**	**BB8**	2136	21.89	**0.105**	**Tawn 1**	269	2.75	0.074	**Tawn 2**	207	3.96	0.076	**BB7**	112	2.28
2	**0.079**	Frank	2027	20.84	**0.105**	**Tawn 1**	269	2.75	**0.112**	Frank	1112	22.00	-0.100	Joe (90)	74	1.52
3	0.035	Sur. Joe	257	2.56	**0.125**	Frank	295	2.98	0.082	Frank	218	4.13	**0.134**	Sur. Joe	125	2.57
4	0.035	Sur. Joe	257	2.56	0.111	**t-cop**	63	0.64	0.082	Frank	218	4.13	0.233	**Gumbel**	31	0.63
5	-0.011	Ind.	98	1.00	-0.029	**t-cop**	101	1.02	-0.073	Ind.	54	0.939	-0.074	Ind.	51	1.03
6	-0.001	Tawn 1 (180)	295	2.95	**0.094**	Frank	294	3.00								
7	0.026	Sur. Joe	257	2.56	**0.108**	Clayton	324	3.33	0.036	Ind.	218	4.13	-0.074	Sur. Gum.	31	0.63
8	0.064	Clayton	94	1.00	0.017	Ind.	97	1.00	-0.020	Clayton	47	1.00	0.195	Ind.	48	1.00

*1+ case at Richmond and Rosharon do not satisfy marginal goodness-of-fit criteria.

The nature of the UTD of the best fitting copulas for tested threshold-lag combinations (Figures 3, SM.18-SM.24) suggests compound flooding potential is higher in the Sabine Basin than the Brazos Basin. In common with previous studies (Ward et al. 2018; Santos et al. 2020), UTD is detected in the COS samples in the Sabine Basin. In contrast with Ward et al. (2018), when COD in Adams Bayou there appears to be UTD for thresholds >0.9, particularly for lags >2 days. The best fitting copulas possess UTD for approximately half of the COD samples in Cow Bayou once the threshold exceeds 0.9. The independence copula is frequently the best fitting for the COD samples in Cow Bayou when using zero or a one-day lag, or thresholds >0.95. In the Brazos Basin, the best fitting copulas for both conditional samples tend not to possess UTD.

366



Tawn type 2
Tawn type 1
Sur. BB7
Sur. BB1
Sur. Clayton
BB8
Joe
Gumbel
t-copula
Rot. Tawn 2 (270)
Rot. Tawn 2 (180)
Rot. Tawn 1 (270)
Rot. Tawn 1 (90)
Rot. Tawn 1 (180)
Rot. Joe (90)
Sur. Gumbel
Frank
Gaussian
Ind.
Copula family
Lag
0 1 2 3 4 5 6 7 8 9
0.75 0.8 0.85 0.9 0.95 1
Quantile threshold
Upper panel:
0.5
0.25
0 τ
-0.25
-0.5
Statistically significant correlation
Non-statistically significant correlation
Lag (days)
Case
1,3,4
5
6
7
Lower panel:
Sample dimension
12
1279
30 250 500 750 1,000

Figure 3: Upper panel: Best fitting, according to AIC, of the 40 copula families in the VineCopula R package (Schepsmeier et al., 2015) for the COS sample at Cow Bayou. Colored markers denote the value of τ and the best fitting copula for combinations of quantile threshold (0.75 to 1 at a 0.001 interval) and lag (0-9 days) for which a conditional sample could be derived, i.e., each threshold-lag combination has only one colored marker across the different copulas (black dots indicate that the correlation is significant at the 95% confidence level). Only copula families above the dashed horizonal line possess UTD. Larger markers denote threshold-lag combinations and copula families chosen for the test cases outlined in Table 1. The BM approach is not associated with any threshold, consequently the name of the copula family (along the y-axis) selected in test case 8 is highlighted in red italics. Lower Panel: Sample dimension of the (conditional) sample for each threshold-lag combination.

The sensitivity of the best fitting copula to the sampling approach including the choice of POT threshold and lag resulted in substantial variation in the copulas associated with test case samples (Table 2). The least variation in the best fitting copulas among the test cases was found at Adams Bayou where three(four) distinct copulas were associated with the six(eight) COS(COD) samples. The greatest disparity was at Richmond where a different copula was the best fitting for each of the COS samples and five distinct copulas were the best fitting for the eight COD samples. In general, the proposed approach (case 1) and the 3 EPY approach (case 6) most often resulted in the selection of copulas with the prevailing class of UTD. On the other hand, the 1 EPY (case 5), no lag (case 7), and BM approaches (case 8) resulted in copulas with the class of UTD found for fewer threshold-lag combinations being selected.

4.2 Case 1

Isolines associated with several RPs derived according to the proposed approach (case 1) are shown in Figure 4. Except for the 5-year RP at Cow Bayou, isolines in the Sabine region are concave indicating drivers can combine to produce compound events with far lower RPs compared to the marginal RPs. The isolines become increasingly concave as RP increases, a consequence of the decrease in coincident extremes in the conditional samples; this is particularly evident at Cow Bayou. Furthermore, the small proportion of events in both conditional samples at Cow Bayou demonstrates that to obtain coincident extremes, a lag needs to be applied. Nevertheless, design events associated with each of the four RPs are all found at or close to the intersection of the two quantile isolines, where neither flooding driver is particularly dominant. The COD sample at Cow Bayou yields the flattest part of any isoline in the region, as expected given its lack of significant correlation and UTD in its best fitting copula.

The weaker relationship between the drivers signified by the general lack of curvature in the isolines for the Brazos River sites compared with those derived for the sites in the Sabine Basin is evident in Figure 4. The very small correlations ($\tau<0.01$) among the COS samples in the Brazos Basin coupled with the subsequent selection of copulas with no UTD leads to isolines which are partially convex. On the other hand, the COD sample at Richmond exhibits significant correlation and is responsible for the most pronounced part of any of the isolines calculated for the sites along the Brazos River. The design event associated with the 5-year RP at Richmond is the only design event for any RP at any site to lie outside the region where both drivers exceed the quantile thresholds used to derive the conditional samples. The associated discharge is small and thus the 5-year RP design event at Richmond can be considered a “surge only” event (i.e. not a compound event).

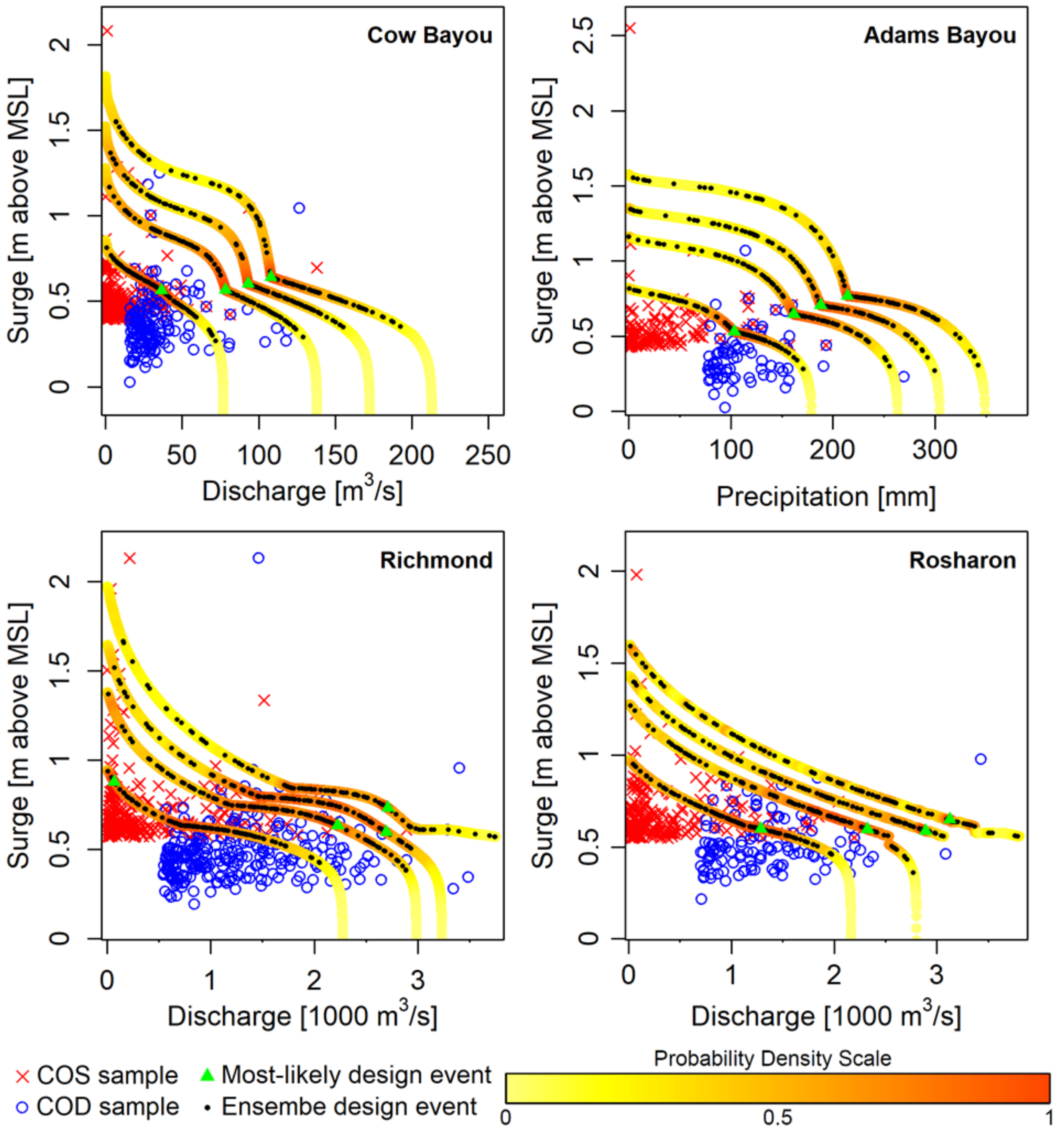


Figure 3: Isolines associated with the (from origin) 5-, 25-, 50- and 100-year RPs derived according to the proposed approach (case 1) superimposed on the conditional samples used in their generation (COS: red; COD: blue). Colored contours denote the relative likelihood (see color bar) of events on an isoline given the data. Black dots represent the ensemble of 1,000 events sampled along the isolines. Green triangles represent the most likely design points.

4.3 Comparison among test cases and with the independence assumption

The sensitivity of the isolines to the subjective choices in the model set up is demonstrated in Figure 5. It is immediately apparent that the isolines derived without considering any lag (case 7) show the greatest disparity with those associated with the other test cases. The RPs obtained at the margins for the no-lag and BM approaches (case 7 and 8) are the least homogenous with the other approaches. As opposed to the isolines derived by the no-lag approach, those derived by the BM approach generally have a similar shape to those given by most other approaches including the 3 EPY and case 1. Despite this, the BM approach consistently underestimates the potential for compounding effects at Cow Bayou. Similarly, the approach identical to case 1 but disregarding the marginal fit (case 2) and the 1 EPY on average approach (case 5) fail to detect compound effects in Adams Bayou and to a lesser extent at Cow Bayou. Case 2 also resulted in large overestimation of the RPs at Rosharon and Richmond as compared to most other approaches. Isolines derived using the 3 EPY on average approach (case 6) are closest to those obtained using case 1 at the three sites at which they could be derived. The approach using a 5-day lag when COD (case 3) at Cow Bayou results in isolines similar to case 1; reducing the lag further to 3-days (case 4) yields a more conservative isoline for the COD part of the isoline. At Adams Bayou and to a lesser extent the sites along the Brazos River, both lag-reduction approaches (cases 3 and 4) underestimate compounding effects relative to the other approaches.

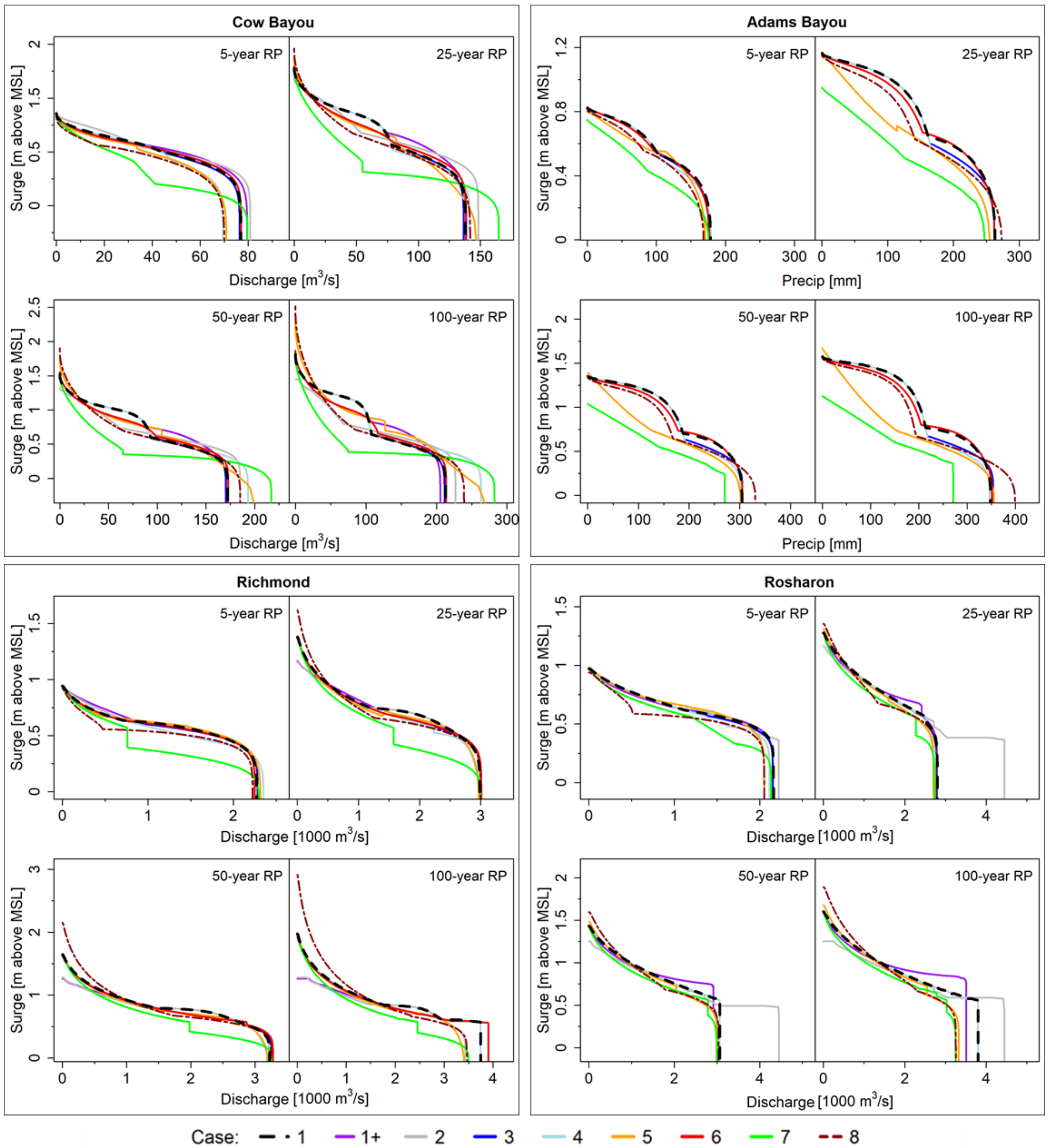


Figure 4: Isolines associated with various RPs according to the test cases (see color coding) summarized in Table 1.

To further assess the effect of ignoring dependence, RPs of the ensemble of 1,000 design events associated with the 100-year RP are recalculated assuming discharge and surge are independent. Univariate RPs are estimated using the distributions fitted to the conditional samples derived when implementing the proposed approach. More specifically, for each event, two independence RPs are calculated from the pairs of GPD and parametric distributions fitted to the two conditional samples. The independence RP is taken as minimum of these two RPs. Below the GPD threshold, RPs are estimated via empirical distributions. Diagnostic plots of the marginal fits can be found in Figures SM.25-SM.40.

Histograms displaying the RPs of the 100-year design events under the independence assumption (Figures 6-7) are representative of the common theme that compound extremes are more prevalent in the Sabine Basin than the Brazos Basin. Joint return periods greater than, equal to, or less than 100 years indicate that compound events are more, equally, or less likely than if the two drivers were independent. Adams Bayou possesses the highest percentage of events with return periods longer than 100 years under the independence assumption at 89%; this value drops to 79% at Cow Bayou. According to the proposed approach, under the assumption of independence, no events in the ensemble exceed their joint (100-year) RP at either site in the Brazos Basin indicating that extreme river discharge and surge events tend not to coincide in this area.

The test cases yielding independence RPs which best match those given by the proposed approach varies by site. Cases 3 and 4 show the best agreement with the proposed approach in the Brazos Basin while cases 1+ and 3 show the best agreement at Cow and Adams Bayou, respectively. The histograms demonstrate that the approaches ignoring the fit of the marginal extreme distribution (case 2), not considering a lag (case 7), and using a block maximum approach (case 8) are most likely to underestimate the propensity for compound events.

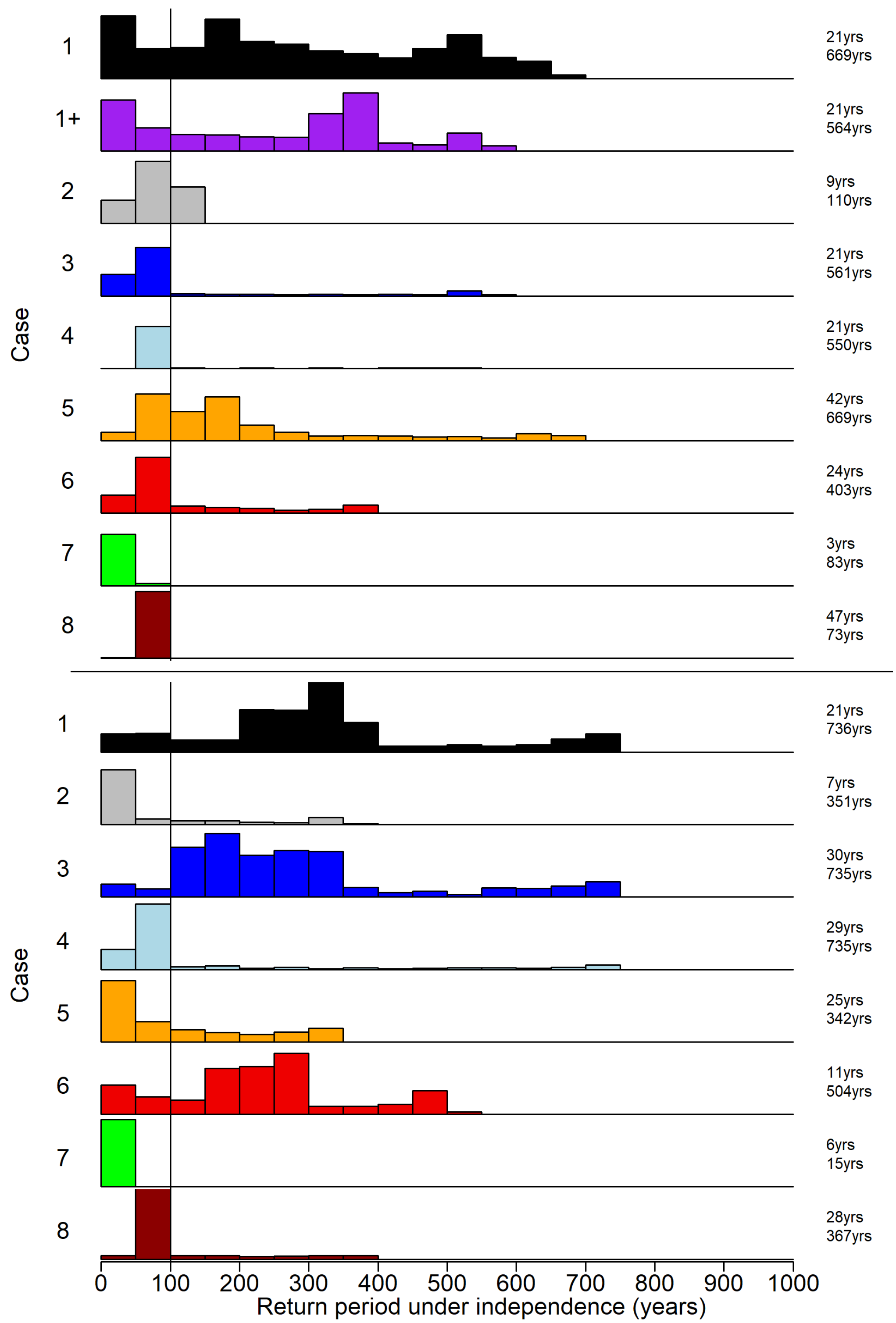


Figure 5: Histograms of the joint RPs of the ensembles of 100-year events under the independence assumption at Cow (upper) and Adams Bayou (lower). Min. and max. RPs are given on the right-hand side; color coding refers to the different test cases as outlined in the legend of Figure 5.

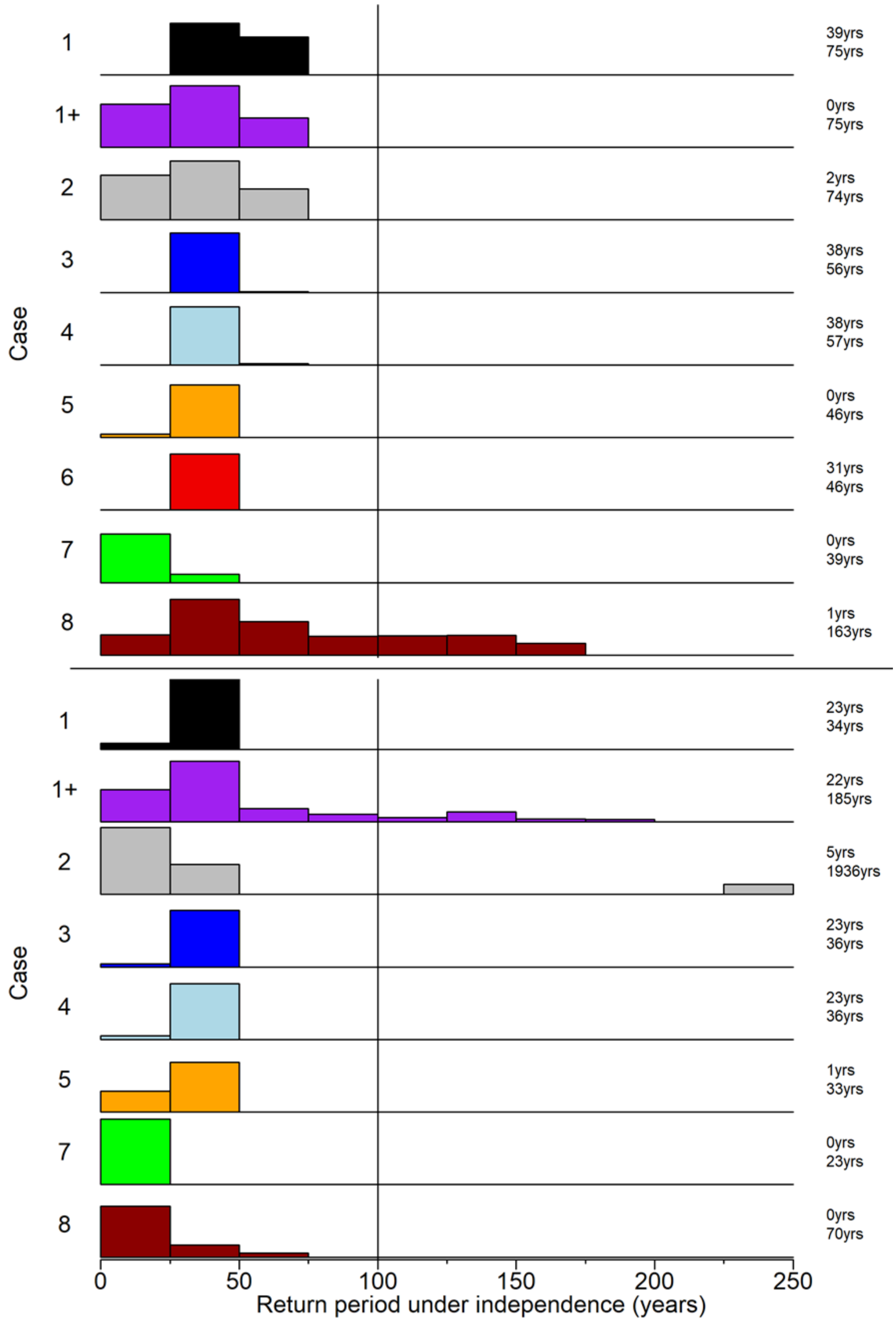


Figure 7: Histograms of the joint RPs of the ensembles of 100-year events under the independence assumption at Richmond (upper) and Rosharon (lower). Min. and max. RPs are given on the right-hand side; color coding refers to the different test cases as outlined in the legend of Figure 5.

4.3 Precipitation as a proxy for discharge

The analysis undertaken so far suggests Adams Bayou has a higher propensity for compound events than neighboring Cow Bayou. Of the National Climatic Data Center's weather stations, the weather station at Orange is the closest to the streamflow gauge at Mauriceville. The analysis undertaken at Adams Bayou is the same that would be carried out at Cow Bayou if precipitation was to be used as a proxy for discharge. The difference in compound event potential in the two bayous may arise due to differences in local catchment characteristics influencing run-off times, e.g. size, topography or land-use, or human activity such as coastal engineering works or channel modifications or because precipitation was adopted as a proxy for discharge in the case of Adams Bayou. The latter may result in the correlation between the fresh and saltwater contributions being decreased when using discharge compared to precipitation due to run-off processes and influence of antecedent groundwater level.

To further investigate the effect of adopting precipitation as a proxy for discharge, the Brazos Basin analysis is implemented with the precipitation measured at Freeport replacing the streamflow records. Isolines associated with several RPs and RPs of a thousand, 100-year precipitation-surge design events recalculated under the assumption of independence using same methodology as in the previous section is shown in Figure 8. The isolines are most reminiscent of those given by the proposed approach at Rosharon (see Figure 4), the closest site to the Freeport weather station. In common with the proposed approach when applied at Rosharon, neither copula possessed UTD and the Frank copula is the best fitting for the COS sample. The Gaussian copula was the best fitting when conditioning on precipitation. The lack of UTD resulted in the isolines becoming increasingly convex as the RP increased. Over 65% of the 100-year design events possessed RPs greater than 100 years under the independence assumption compared to zero when using discharge (Figure 8). Hence, in common with the results from the Sabine River Basin, substituting precipitation for discharge increased the estimated potential for compound events in the Brazos River.

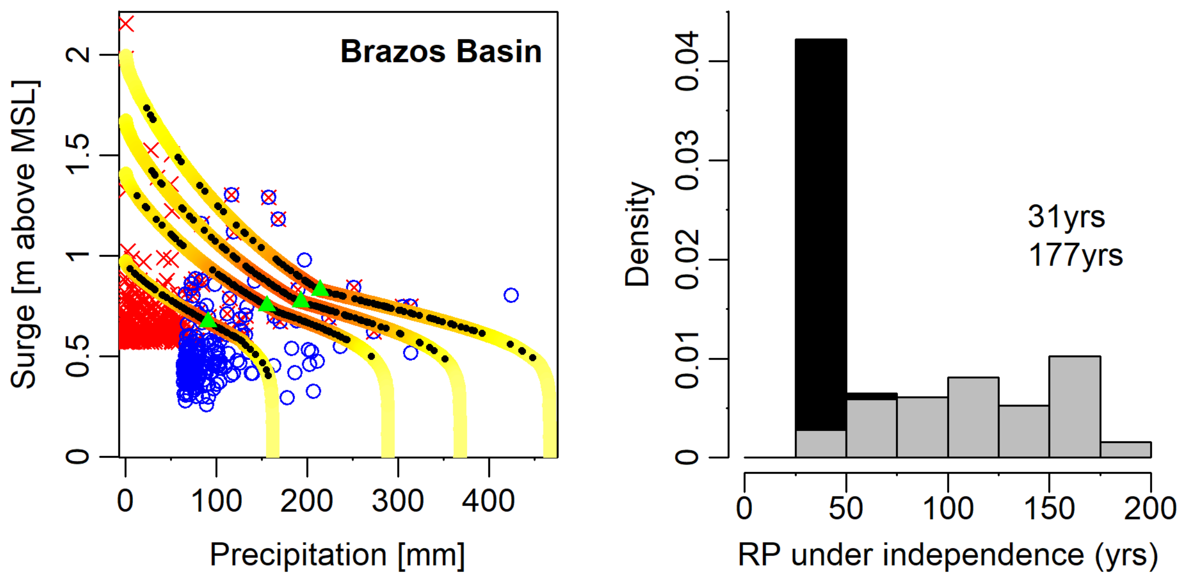


Figure 8: Left: Isolines associated with the 5-, 25-, 50- and 100-year RPs derived according to the proposed approach (case 1) superimposed on the conditional samples used in their generation. Colored contours denote the relative likelihood of events on an isoline given the data. Right: Histogram of the RPs of 1000, 100-year design events recalculated under the assumption of independence using Freeport precipitation (grey) and Richmond discharge (black) records to represent the freshwater contribution. Min. and max. RPs are given on the right-hand side.

5. Conclusion

The overarching aim of this study was to assess the propensity for high river discharge and storm surge events at three locations in southeast Texas: Cow and Adams Bayous (Sabine River Basin) and the Brazos River (Brazos River Basin). According to the methodology adopted in the study, i.e., combining two-sided conditional sampling with a copula analysis, compound events are a feature of both regions but more prevalent in the Sabine River Basin than the Brazos River Basin. The sensitivity of the estimated compound event potential to subjective choices in the model set up, namely the sampling method and the choices of threshold and lag were also explored to ensure robust results. Criteria which limit the size of the conditional samples, such as choosing on average one event per year, or using annual maximum, as well as not considering any time lags, underestimate compound flooding potential compared to the other approaches tested. A pragmatic best practice approach is proposed using the fit of the marginal extreme distribution to guide the choice of threshold for the conditional samples and including time lags. The approach provides stable estimates of the potential for compound flooding, while results tend to be on the conservative side when compared to the other methods.

As discharge time series are often too short for robust statistical analysis, we also assessed the sensitivity of results when using precipitation as a proxy. The greatest correlation is observed in Adams Bayou where precipitation is used throughout due to the lack of suitable discharge data. Repeating the analysis in the Brazos Basin but using precipitation as a proxy for discharge leads to inflated estimates of the potential for compound discharge-surge events.

A pertinent limitation of the study is that the length of the time series presently prevents a robust assessment of RPs greater than 100 years. Including more historic observational data or numerical model hindcasts (or synthetic event simulations) would allow the analysis of events associated with longer RPs.

Acknowledgments

RAJ and TW acknowledge financial support from the USACE Climate Preparedness and Resilience Community of Practice and Programs. This material is based in part on work supported by the National Science Foundation under Grant AGS-192938

Data Availability Statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

References

Arns, A., T. Wahl, I. Haigh, J. Jensen, and C. Pattiaratchi. 2013. "Estimating extreme water level probabilities: A comparison of the direct methods and recommendations for best practice." *Coast. Eng.* 81, 51-66. https://doi.org/10.1016/j.coastaleng.2013.07.003

Balkema, A. and L. De Haan. 1974. "Residual life time at great age." *Ann. Probab.* 792-804. https://doi.org/10.2307/2959306

Bakhtyar, R., Clark, E. P., Flowers, T., Maitaria, K., Velissariou, P., Trimble, B., et al. 2020. "A new 1D/2D coupled modeling approach for a riverine-estuarine system under storm events: Application to Delaware River Basin." *J. Geophys. Res. Oceans*, 125, e2019JC015822. https://doi.org/10.1029/2019JC015822

Blake, E., and D. Zelinsky. 2018, Hurricane Harvey (National Hurricane Center Tropical Cyclone Report).

Bender, J., T. Wahl, A. Müller, and J. Jensen. 2016. "A multivariate design framework for river confluences." *Hydrolog. Sci. J.*, 61(3), 3471-482. https://doi.org/10.1080/02626667.2015.1052816

[BRA] Brazos River Authority. 2019. Lower Brazos Flood Protection Planning Study. HALFF.

Blanton, B., K. Dresback, B. Colle, R. Kolar, H. Vergara, Y. Hong, N. Leonardo, R. Davidson, L. Nozick, T. Wachtendorf. 2020. "An integrated scenario ensemble based framework for hurricane evacuation modeling: Part 2—hazard modeling." *Risk Anal.,* 40, 117-133. https://doi.org/10.1111/risa.12990

Brauer, N., J. Basara, C. Homeyer, G. McFarquhar, and P. Kirstetter. 2020. Quantifying precipitation efficiency and drivers of excessive precipitation in post-landfall Hurricane Harvey. J. Hydrometeor., 21, 433–452. https://doi.org/10.1175/JHM-D-19-0192.1

[Brazos BBEST] Brazos River Basin and Bay Expert Science Team. 2012. Brazos River basin and bay expert science team environmental flow regime recommendations report. Final submission to the Brazos River Basin and Bay Area Stakeholder Committee. Austin (TX): Environmental Flows Advisory Group and the Texas Commission on Environmental Quality.

Camus, P., I. Haigh, T. Wahl, and Nasr. 2020. "Compound flooding potential due to pluvial, fluvial and coastal drivers along the coast of Eastern North Atlantic Ocean, Mediterranean Sea and Black Sea: sensitivity analysis and spatial patterns." *Nat. Hazards and Earth Syst. Sci.,* in prep.

Couasnon, A., D. Eilander, S. Muis, T. Veldkamp, I. Haigh, T. Wahl, H. Winsemius, and P. Ward. 2020. "Measuring compound flood potential from river discharge and storm surge extremes at the global scale." *Nat. Hazards and Earth Syst. Sci.* 20 (2), 489-504. https://doi.org/10.5194/nhess-20-489-2020

Duong, T. 2007. "ks: Kernel density estimation and kernel discriminant analysis for multivariate data in R." *J. Stat. Softw*., 21 (7), 1-16. https://doi.org/10.18637/JSS.V021.I07

FEMA. 2015. *Guidance for flood risk analysis and mapping; combined coastal and riverine floodplain*, (No. Guidance Document 32), Washington, DC: FEMA.

Gnedenko, B.V. 1943. "Sur la distribution limite du terme maximum d'une serie aleatoire." *Ann. Math.* 44 (3), 423–453. https://doi.org/10.2307/1968974

Gori, A., N., Lin, and J. Smith. 2020. "Assessing compound flooding from landfalling tropical cyclones on the North Carolina coast." *Water Resour. Res.*, 56, e2019WR026788. https://doi.org/10.1029/2019WR026788

Gräler, B., M. van den Berg, S. Vandenberghe, A. Petroselli, S. Grimaldi, B. De Baets, and N. Verhoest. 2013. "Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation." *Hydrol. Earth Syst. Sci.* 17, 1281-1296. https://doi.org//10.5194/hess-17-1281-2013

Jane, R., L. Cadavid, J. Obeysekera, and T. Wahl. 2020. Multivariate statistical modelling of the drivers of compound flood events in South Florida." *Nat. Hazards and Earth Syst. Sci. Discuss.* 1-30. https://doi.org/10.5194/nhess-2020-82

Joyce, J., N. Chang, R. Bin, R. Harji, T. Ruppert and P. Singhofen. 2018. "Cascade impact of hurricane movement, storm tidal surge, sea level rise and precipitation variability on flood assessment in a coastal urban watershed." *Clim. Dyn.* 51, 1–27. https://doi.org/10.1007/s00382-017-3930-4

Kumbier, K., R. Cabral Carvalho, A. Vafeidis, and C. Woodroffe. 2018. "Investigating compound flooding in an estuary using hydrodynamic modelling: a case study from the Shoalhaven River, Australia." *Nat. Hazards and Earth Syst. Sci.* 18, 463–477. https://doi.org/10.5194/nhess-18-463-2018

Lee, S., T. Kang, D. Sun, J. Park, 2020. "Enhancing an Analysis Method of Compound Flooding in Coastal Areas by Linking Flow Simulation Models of Coasts and Watershed." *Sustainability*, 12, 6572. https://doi.org/10.3390/su12166572

Loveland, M., Kiaghadi, A., Dawson, C. N., Rifai H. S., Misra, S., Mosser, H. and Parola, A. 2020. "Developing a Modeling Framework to Simulate Compound Flooding: When Storm Surge Interacts with Riverine Flow. " *Front. Clim.*, 2:609610. https://doi.org/ 10.3389/fclim.2020.609610

Moftakhari, H., J. Schubert, A. AghaKouchak, R. Matthew, and B. Sanders. 2019. "Linking statistical and hydrodynamic modeling for compound flood hazard assessment in tidal channels and estuaries." *Adv. Water Resour.* 128, 28-38. https://doi.org/10.1016/j.advwatres.2019.04.009

[NOAA] National Oceanic and Atmospheric Administration. 2020. Billion-dollar weather and climate disasters: Table of events.

Northrop, P. J., N. Attalides, and P. Jonathan. 2017. “Cross-Validatory Extreme Value Threshold Selection and Uncertainty with Application to Ocean Storm Severity*.” J. R. Stat. Soc. Ser. C Appl. Stat.,* 66 (1), 93–120. https://doi.org/10.1111/rssc.12159

[OCDD & TWDB] Orange County Drainage District and Texas Water Development Board. 2015. Flood Protection Planning Study, Cow and Adams Bayou, Orange County Texas. Final Report. Beaumont, TX: Carroll & Blackman Inc.

Parsons. 2006. Nonpoint source modeling of the watersheds of Adams and Cow Bayous. Orange County Total Maximum Daily Load Project. Submitted to the Texas Commission on Environmental Quality, Austin, TX.

Pettitt, A. 1979. “A non-parametric approach to the change-point problem.” *J. R. Stat. Soc. C* 28 (2), 126–135. https://doi.org/10.2307/2346729

Rahmani, V., S. Hutchinson, J. Harrington Jr, J. Hutchinson, and A. Anandhi. 2015. “Analysis of temporal and spatial distribution and change-points for annual precipitation in Kansas, USA.” *Int. J. Climatol.* 35, 3879–3887. https://doi.org/0.1002/joc.4252

Ridder, N., A. Pitman, S. Westra, A. Ukkola, X. Hong, M. Bador, A. Hirsch, J. Evans, A. Di Luca and J. Zscheischler. 2020. “Global hotspots for the occurrence of compound events.” *Nat Commun* 11, 5956. https://doi.org/10.1038/s41467-020-19639-3

Sadegh M., E. Ragno, and A. AghaKouchak. 2017. “Multivariate Copula Analysis Toolbox (MvCAT): Describing Dependence and Underlying Uncertainty Using a Bayesian Framework” *Water Resour. Res.* 53 (6), 5166-5183. https://doi.org/10.1002/2016WR020242

Salas, J. D. 1993. “Analysis and modeling of hydrologic time series, in Handbook of Hydrology.” edited by D. Maidment, pp. 19.1–19.72, New York: McGraw-Hill.

Salvadori, G., C. De Michele, and F. Durante. 2011. “On the return period and design in a multivariate framework.” *Hydrol. Earth Syst. Sci.* 15 (11), 3293–3305. https://doi.org/10.5194/hess-15-3293-2011.

Santiago-Collazo, F.L., M. Bilskie, and S. Hagen. 2019. “A comprehensive review of compound inundation models in low-gradient coastal watersheds.” *Environ. Model. Softw.* 119, 166-181. https://doi.org/10.1016/j.envsoft.2019.06.002

Santos, V., T. Wahl, R. Jane, S. Misra, and K. White. 2020. “Assessing Compound Flooding Potential with Multivariate Statistical Models in a Complex Estuarine System under Data Constraints.” *Journal of Flood Risk Management*. Submitted.

Schepsmeier, U., J. Stoeber, E. Brechmann, B. Graeler, T. Nagler, T. Erhardt, C. Almeida, A. Min, C. Czado, and M. Hofmann. 2015. Package ‘VineCopula’. R package version, 2(5).

Sharma, S., D. Swayne, and C. Obimbo. 2016. “Trend analysis and change point techniques: a survey.” *Energ. Ecol. Environ.* 1 (3), 123–130. https://doi.org/10.1007/s40974-016-0011-1

Smith, R.L., and I. Weissman. 1994. “Estimating the extremal index.” *J. R. Stat. Soc. Ser. B* 56 (3), 515-528. https://doi.org/10.2307/2346124.

Solari, S., M. Egüen, M. Polo, and M. Losada. 2017. “Peaks Over Threshold (POT): A methodology for automatic threshold estimation using goodness of fit p-value” *Water Resour. Res.*, 53, 2833– 2849. https://doi.org/10.1002/2016WR019426

[TIFP & BRA] Texas Instream Flow Program and Brazos River Authority. 2010. Instream Flow Study of the Middle and Lower Brazos River: Draft study design. Submitted to the Middle and Lower Brazos River Sub-Basin Study Design Workgroup, Texas Water Development Board, Austin TX.

Tosunoglu, F., F. Gürbüz, and M. Ispirli. 2020. “Multivariate modeling of flood characteristics using Vine copulas. *Environ. Earth Sci.*, 79 (19), 459-479. https://doi.org/10.1007/s12665-020-09199-6

Valle-Levinson, A., M. Olabarrieta, and L. Heilman. 2020. “Compound flooding in Houston-Galveston Bay during Hurricane Harvey.” *Sci. Total Environ.* 747: 141272. https://doi.org/10.1016/j.scitotenv.2020.141272

Wahl, T., S. Jain, J. Bender, S. Meyers, and M. Luther. 2015. “Increasing risk of compound flooding from storm surge and rainfall for major US cities.” *Nat. Clim. Change* 5 (12), 1093–1097. https://doi.org/10.1038/nclimate2736.

Ward, P., A. Couasnon, D. Eilander, I. Haigh, A. Hendry, S. Muis, T. Veldkamp, H. Winsemius, and T. Wahl. 2018. “Dependence between high sea-level and high river discharge increases flood hazard in global deltas and estuaries” *Environ. Res. Lett.* 13: 084012. https://doi.org/10.1088/1748-9326/aad400

Watson, K., G. Harwell, D. Wallace, T. Welborn, V. Stengel, and J. McDowell. 2018. Characterization of peak streamflows and flood inundation of selected areas in southeastern Texas and southwestern Louisiana from the August and September 2017 flood resulting from Hurricane Harvey: U.S. Geological Survey Scientific Investigations Report 2018 – 5070. https://doi.org/10.3133/sir20185070

Wing, O. E., C. Sampson, P. Bates, N. Quinn, A. Smith, and J. Neal. 2019. “A flood inundation forecast of Hurricane Harvey using a continental-scale 2D hydrodynamic model.” *J. Hydrol. X* 4: 100039. https://doi.org/10.1016/j.hydroa.2019.100039

Wu, W., S. Westra, and M. Leonard. 2020. “Estimating the Probability of Compound Floods in Estuarine Regions.” *Hydrol. Earth Syst. Sci. Discuss.*, in review. https://doi.org/10.5194/hess-2020-456