Reviewer 1

The authors would like to thank the reviewers for their constructive comments that have improved this manuscript.

Reviewer #1's comments (or edits) have been numbered using the format Q.L where Q is the reviewer comment and L is the line number referred to in the reviewer uploaded file "reviewer's nhess-2021-241-RC1.pdf". The authors' response to each comment is given below each question/comment using the format R.L.

Q.L9 remove ','

R.L9 The comma has been removed.

Q.L12 I'm sure this will be explained later, but right now I don't understand what this means.

R.L12 The data pairs are shown in Table 2 and the authors have updated the text for clarity. The new text reads: Wet season coinciding water level and precipitation sampling benefits from a dramatic increase in data pairs, improved goodness of fit statistics, and provide a range of physically realistic pairs.

Q.L21 I find this a bit misleading. The fact that such a small SLR causes a doubling of the odds of the 50-year flood event only shows there is a relatively small difference between the 25-year and 50-year event (namely 5 cm), probably be because of relatively modest storm surges. That means Southern California is actually fortunate to not have very extreme high sea levels during extreme events. And that actually makes the area potentially less vulnerable.

R.L21 The author's agree that typical US West Coast storm surge magnitudes are small (~10 cm, Flick et al., 1998) when compared to multi-meter hurricane generated storm surges experienced in regions with wider continental shelfs. Along the US West Coast, tides dominate marine water levels. Urbanized regions have been built to accommodate the spring tides. Ironically, it is this modest storm surge that make the region highly sensitive to even minor changes in sea level. For example, a 10 cm sea level rise results in spring tides being identical to (or larger than) many historical storm event water levels. The impacts of sea level rise on coastal flooding and vulnerability in California have been demonstrated in the literature (e.g., Tebaldi et al., 2012, Vitousek et al., 2017, Taherkhani et al., 2020). Moftkahari et al., (2015) shows the impact of minor sea level perturbations on flooding (Figure 1). San Francisco, in particular, is highly sensitive to even small (3 cm) increases in sea level (red outline, Figure 1).

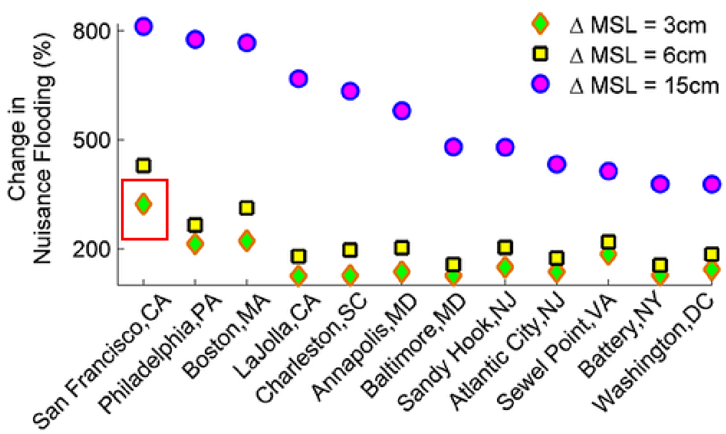

Figure 1. Relative vulnerability along U.S. coast to a unified MSL rise. Adapted from Moftakhari et al., (2015).

Q.L31 I don't understand the second part of this sentence. How can you have different outcomes for an event? An event is single (not plural) so the outcome is single as well. Perhaps you mean different events with the same return period? That would make more sense grammatically. However, in that (multivariate) case it is important to note that the return period loses its meaning (as known in the univariate sense).

R.L31 The authors would like to thank the reviewer for pointing out this inconsistency in the language. The authors have adopted the reviewer's suggestion and updated the text for clarification. The new text reads "Notably, compound events that share a common return period may produce vastly different flooding outcomes.".

Q.L35 remove 'potential'

R.L35 The word 'potential' has been removed.

Q.L36 The more severe of what? Which things are compared here in this univariate case? Or is this the multivariate case in which each variable is analyzed separately? Then please state this more explicitly.

R.L36 The authors have updated the text for clarification. The new text reads "For example, FEMA recommends characterizing compound events by developing univariate water level and discharge statistics, modeling each separately, and then adopting the more severe flooding result for transitional areas (FEMA 2011, 2016c)."

Q.L35 remove "Ironically"

R.L35 The word ironically has been removed.

Q.L81 Reduce?

R.L81 The authors suggestion has been adopted and the text has been updated.

Q.L87 Extended compared to what? I would like to know the period of observations for the three stations.

R.L87 Table 2 in the manuscript shows the observation windows used for the study. The full high/low and hourly OWL and precipitation at all sites up to 8/31/21 are provided for the reviewer in the table immediately below.

	Precipitation		High-Low Tide		Hourly Tide	
Site	Start	End	Start	End	Start	End
Santa Monica	7/1/1948	12/19/2013	8/1/1979	8/31/2021	11/22/1973	8/31/2021
Sunset	7/1/1948	12/1/2012	8/1/1979	8/31/2021	11/28/1923	8/31/2021
San Diego	7/1/1948	12/19/2013	7/1/1979	8/31/2021	8/1/1924	8/31/2021

Using hourly tide data adds over 50 years of additional observed water level records and 31 additional years overlapping precipitation observations for considering compound events at Sunset and San Diego and six years for Santa Monica. The text has been updated to specify the observation windows. The revised text in L84-89 is:

Observed water levels from the Los Angeles (Station ID: 9410660), La Jolla (Station ID: 9410230), and Santa Monica (Station ID: 9410840) tide gauges are available on NOAA's Tides and Currents for daily high-low, hourly, or six-minute intervals (NOAA, 2021 Accessed 2021d). Verified hourly water levels (m NAVD88) had the longest record length at all three stations and provided an additional 31-years of observations overlapping precipitation data for Los Angeles and La Jolla, and 6-years for Santa Monica. The resulting observations windows are November 22, 1973 to December 19, 2023 for Santa Monica, July 1, 1948 to December 1, 2012 for Sunset and July 1, 1948 to December 19, 2013 for San Diego (Table 2).

Q.L96 Are. How was this done?

R.L96 The authors thank the reviewer for catching this grammatical error. The text has been updated with "are". Precipitation measurements were converted to mm/hr by dividing the total event precipitation by the event time. This is reflected in the revised text which now reads "Precipitation measurements were converted to a mm/hr rate by dividing the total event precipitation by the event time to match the hourly OWL measurements."

Q.L108 remove 'also'

R.L108 The word 'also' has been removed.

Q.L110 replace its with their

R.L110 The sentence has been updated per the reviewer's suggestion and now reads "In the case of coinciding sampling, pairs that had three or more OWL measurements missing within the 24-hour window were manually reviewed and removed if their tidal peak was clearly missing. Specifically for WMM sampling, months with more than half their observations missing were also reviewed and removed if the tidal peak was missing."

Q.L132 Are

R.L132 The authors thank the reviewer for catching this grammatical error, is has been changed to are.

Q.L150 Kendall scenario.

R.L150 The text has been updated to include the word scenario.

Q.L170 But pdfs are very easy to construct from cdfs with numerical approximations.

R.L170 Uncertainties can be quantified and explored without PDFs, but establishing the most likely events associated to a specific return period requires continuous PDFs . The PDF is generated by taking the derivative of the CDF. Any CDF discontinuities (e.g., in piecewise functions) result in an undefined PDF. Several copulas (e.g., Cuadras-Auge, Shih-Louis, Marshal-Olkin, and Fischer-Hinzmann) employ "min" statements in their CDFs which causes their PDFs (i.e., the derivative of the CDF) to have undefined locations (e.g., Sadegh et al., 2018). Similarly, Raftery, Linear-Spearman, and Cube copulas are piecewise CDF functions with conditional statements, imparting undefined locations in the PDF. In other cases, (e.g., Gaussian, Student-t, and Husler-Reis) distribution functions embedded into the CDF results in a complex partial derivate for the conditional scenarios which are required to establish the most likely event values.

Q.L171 Therefore it was decided to remove …

R.L171 The text has been updated in accordance with the reviewer's suggestion and now reads "therefore it was decided to remove those copulas…"

Q.L172 Equation?

R.L172 The equation reference has been added after "Conditional 3" to reference the associated equation.

Q.L173 - That surprises me. This should not happen with well chosen values of dx and dy.

R.L173 –The probability space is divided into grid spacings of 0.0005 between 0 and 0.8 and 0.00005 from 0.8 to 1. This high resolution interval is designed to prevent any poor estimations caused by discretization, but in isolated instances negative probabilities occur when a partial derivative is calculated.

Q.L198 Which probability distributions were used to fit the marginals?

R.L198 Section 4.1 and Table 3 specify the selected probability distributions for marginal statistics by each site and sampling method. The reference to Figure 3 in L198 was removed to avoid confusion.

Q.L230 I wouldn't call this 'impacts' as that word has other meanings in flood risk management.

R.L230 The authors thank the reviewer for pointing this out. The text has been revised and now reads "San Diego WMC conditional CDFs display individual copulas effects (Fig. 5)".

Q.L233 Mutually consistent?

R.L233 The text has been updated and now reads "The Roch-Alegre and Fischer-Kock provide very similar results for both precipitation and water level (black and green lines, Fig. 5a, b, e, f)."

Q.L247 I would like some more elaboration on the number of observations ending up at the upper right side of the isolines. Taking the length of the observation period into account that could provide additional evidence in favor of or against some of the copulas.

R.L247 We will address this comment in two parts: the first addressing the observations surpassing the isolines and the second addressing the length of observations affecting copula selection.

Critical multivariate events may occur in different regions depending on the hazard type. For example, in multivariate drought studies where the axes are precipitation and soil moisture the critical area representing meteorological drought conditions (i.e., non-exceedance-non-exceedance extremes) lies below the isoline (i.e., Region I). Any data pair above and to the right of the isoline would, in this case be considered 'safe' or events of no concern.

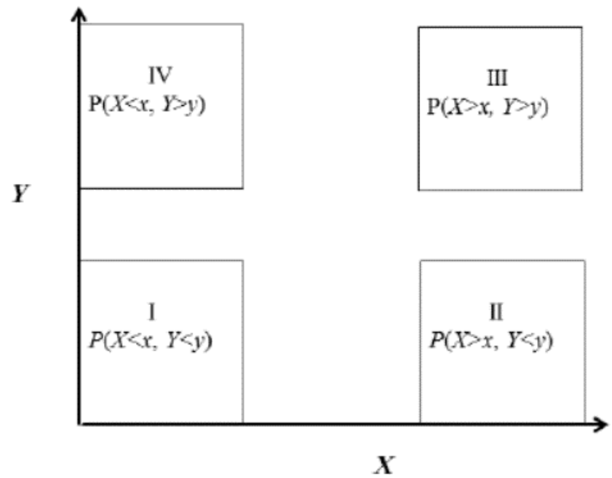

Figure 1, Compound event regions, adapted from Hao et al., (2018).

Conversely, in exceedance-exceedance applications like compound flooding events where, in this study, the axes are precipitation and observed water level the events of interest exist in Region III (Figure 1).

Events above and to the right of the isoline represent more extreme compound events. At the 10-year return period (Figure 2a) a number of exceedance-exceedance events would be expected given the near 70-year observation window. When the return period is more extreme (i.e., 100 year, Figure 2b) pairs on the upper right are minimal or zero for well fit copulas (i.e., Roche-Alegre, Fischer-Kock, Tawn).

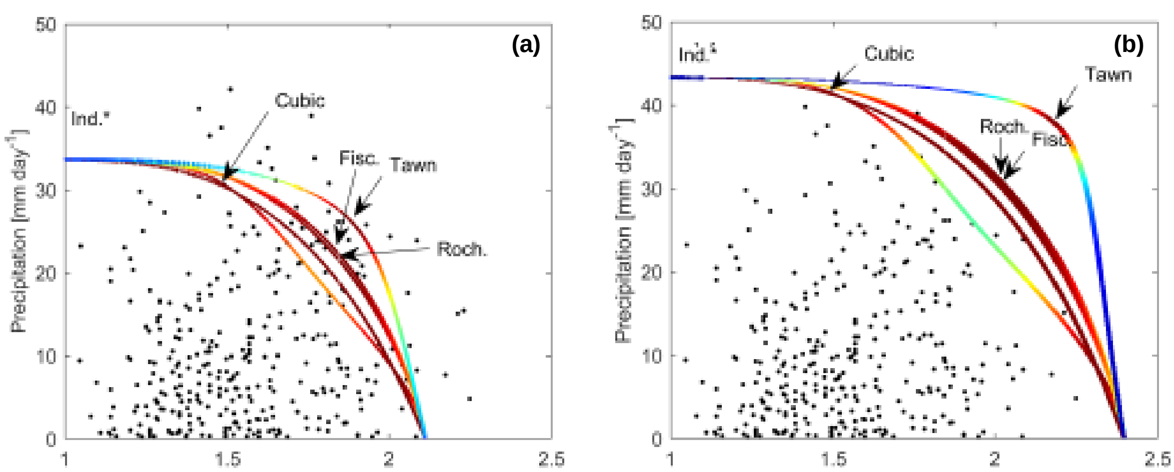

Figure 2 (a) 10 year and (b) 100 year return periods for various copulas using the AND scenario. Adapted from Figures 6a and 7a, Lucey and Gallien, in review.

Alternatively, sampling impacts can be considered in regard to observations above the isoline. Annual Maximum (AM, Figure 3 blue x) sampling pairs the single largest precipitation and OWL observations within a given year (without regard to co-occurrence), which are clearly shown as exceedance-exceedance. Similarly, Wet season monthly maximum (WMM) pairs the single largest precipitation and OWL observations within each wet season month. Maximum parings (annual or wet season) do not represent an observed compound event since the pairs did not co-occur, rather were developed from sampling the largest water level and precipitation event which occurred in a given time frame. This maximum sampling is recommended FEMA guidance as a “worse-case scenario" approach (FEMA, 2016). Unsurprisingly, this manifests as a number of pairs occurring in region III (blue x’s, black dots).

If historically observed, physically realistic events considered using Annual Coinciding (AC) or the Water Month Coinciding (WMC) data pairs, it becomes apparent in the 100-year return period (Figure 3b) that coinciding events are well described by the isoline. Only one event exceeds the isoline (red arrow).

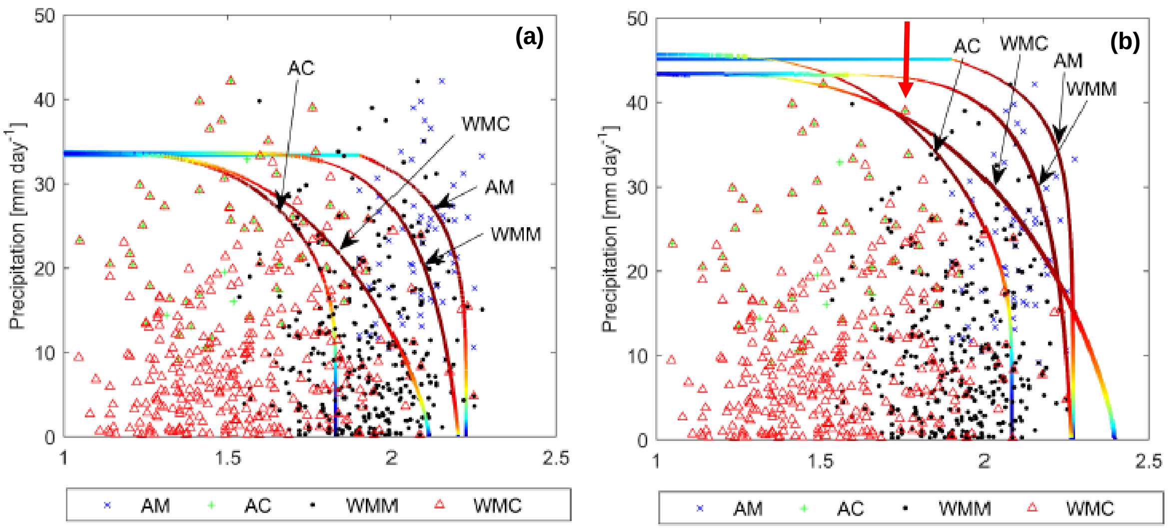

Figure 3, (a) 10 year return period isolines and (b) 100 year return period for the Fischer-Kock.

Second, considering the length of observations (i.e., the number of events) will influence copula selection (Tong et al., 2015, Sadegh et al., 2017).

Tong et al., (2015) explicitly considers data record length on copula selection, distribution characteristics (mean, standard deviation, skewness, autocorrelation), entropy, goodness of fit (Akaike information criterion, AIC), parameter estimator methods, and return period uncertainty. This study utilizes annual maximum flood peak data between 1893 to 2004 along the Yangtze River in China. Only three single parameter copulas are considered: Clayton, Frank, and Gumbel. In circumstances with minimal data availability (<40-years) the best fitting copula varied between the Frank and Gumbel but evolved to a Frank when record lengths were extended (i.e., 40- to 80-years). Copula fittings were insensitive to time period windows (e.g. period between 1910-1992 vs. 1917-1999 were both well fit by the Frank copula). When the data availability was reduced, distribution characteristics varied. For example, entropy, a measurement of disorder (higher entropy meaning a likelier state) decreases with a shorter data length and longer data records improved AIC values (i.e., minimalized the AIC).

In Sadegh et al., (2017), the Tawn (3-parameter), BB1 (2-parameter), and Burr (1-parameter) copulas are fit to precipitation and soil moisture data given 68-years of monthly (816 pairs), 68-years of annual (68 pairs), and 34-years (34 pairs) of annual observations. The Tawn best described the most data dense observations (68 years of monthly data), followed by the BB1 for the 68 years of annual data, and then the Burr for the 34 years of annual data. In this case, the three parameter Tawn was the only copula able to identify an asymmetric dependence between precipitation and (biased towards) soil moisture apparent in the longer, denser data. Additionally, longer records (i.e., increased data availability) reduced uncertainty along the isolines. Although observational record length implications are beyond the scope of this specific study, it is of great interest and we anticipate future work in this area.

Q.L251 One word.

R.L251 The text has been changed to whereas.

Q.L276 remove ‘unique’

R.L276 The word ‘unique’ has been removed

Q.L280 remove ‘also’

R.L280 The word ‘also’ has been removed.

Q.L283 Equal to or greater than.

R.L283 The text has been updated and now reads “A water level equal to or greater than 1.68 m NAVD88 forces valve closures…”

Q.L297 In our study.

R.L297 The text has been revised and now reads “Gaussian and Student t copulas were excluded from this study due to their lack of a computationally simple derivative or integral”

Q.L308 remove ‘likely’

R.L308 ‘likely’ has been removed

Q.L349 – 119, 42

R.L349– All values in the manuscript are provided with two decimal places to maintain consistency.

Q.L364 remove ‘the’

R.L364 ‘the’ has been removed

Q.L366 Particularly

R.L366 The reviewers suggestion has been adopted and the revised text reads “…they are fundamental to coastal flooding, particularly in regions…”.

Q.L373 Records are not quadrupled, you just sample more data from the record

R.L373 The authors agree with the reviewer and the text has been revised for accuracy. The new sentence is “Wet season sampling quadruples data pairs (Table 2), providing additional historical joint event information”.

References

- FEMA: Guidance for Flood Risk Analysis and Mapping: Coastal Flood Frequency and Extreme Value Analysis, https://www.fema.gov/sites/default/files/2020-02/Coastal_Flood_Frequency_and_Extreme_Value_Analysis_Guidance_Nov_2016.pdf, 2016.
- Flick, R. E.: A comparison of California tides, storm surges, and mean sea level during the El Niño winters of 1982-83 and 1997-98, Shore& Beach, 66, 7–11, 1998.
- Hao, Z., Singh, V. P., and Hao, F.: Compound extremes in hydroclimatology: a review. Water, 10(6), 718, 2018.
- NOAA: Tides & Currents, Online, https://tidesandcurrents.noaa.gov, Accessed 2021.
- Sadegh, M., Ragno, E., and AghaKouchak, A.: Multivariate Copula Analysis Toolbox (MvCAT): describing dependence and underlying uncertainty using a Bayesian framework, Water Resources Research, 53, 5166–5183, 2017.

- Sadegh, M., Moftakhari, H., Gupta, H. V., Ragno, E., Mazdiyasni, O., Sanders, B., Matthew, R., and AghaKouchak, A.: Multihazard scenarios for analysis of compound extreme events, Geophysical Research Letters, 45, 5470–5480, 2018.
- Tong, X., Wang, D., Singh, V. P., Wu, J. C., Chen, X., and Chen, Y. F.: Impact of data length on the uncertainty of hydrological copula modeling, Journal of Hydrologic Engineering, 20(4), 05014019, 2015.

Reviewer 2

The authors would like to thank the reviewer for their comprehensive review and constructive comments that have improved this manuscript.

Reviewer #2's comments have been numbered using the format QX.Y where Q is the reviewer query, X is the reviewer comment number and Y represents a subdivision of the comment if there are multiple items requiring attention. Responses follow the format of reviewer's "nhess-2021-241-RC2_supplement.pdf". The authors' response to each comment is given below each question/comment using the format RX.Y.

Q1.1 Their definition of compound events are not correct. The sampled AM extremes do not represent the compound event. Since unlike drought, which is a slowly developing phenomena, the occurrence of floods is faster; varies from few hours (e.g., flash flood) up to a week of time scale (considering inundation effects). The authors define AM event to be when the maximum sampling pairs of single largest precipitation and Observed Water Level (OWL) happens within *a year or month*. However, the paired events can only be qualify as 'compound coastal-pluvial or riverine floods' when the two drivers occurs coincidently if not successively within a *limited time window*.

R1.1. The authors agree that the annual maximum method does not represent an observed compound event. However, FEMA, a primary flood regulatory agency in the United States specifically suggests characterizing compound events using the annual maximum pairing (FEMA, 2016). The text has been substantially updated to reflect this key distinction that the reviewer points out. The new text in L103-114 of the revised manuscript is:

Compound flood probabilities are determined with combinations of sampling methods: Annual Maximum (AM), Annual Coinciding (AC), Wet Season Monthly Maximum (WMM), and Wet Season Monthly Coinciding (WMC). AM sampling pairs the single largest precipitation and OWL observations within a given year (without regard to co-occurrence), where AC sampling pairs the single largest precipitation observation within a given year to the largest OWL observation within its 24-hour accumulation period. A summary of each sites' associated gauges, observation windows, and number of pairs is provided in Table 2. Southern California's wet season is defined between October to March and provides a majority of the total annual rainfall (Cayan and Roads, 1984; Conil and Hall, 2006). It is likely for extreme compound events to occur during this period.

Wet season monthly maximum pairs the single largest precipitation and OWL observations within each wet season month. Although strictly speaking maximum parings (annual or wet season) do not technically represent an observed compound event since the co-occurrence of precipitation and water levels follows the FEMA guidance for considering a "worst case scenario" approach (FEMA, 2016c). Wet season monthly coinciding sampling pairs the single largest precipitation observation within each wet season month to the largest OWL observation within its 24-hour accumulation period, providing more realistic pairs compared to maximum sampling.

Q1.2 Based on large-scale climatic pattern this time window is often taken as a within a week of occurrence of the first event. This is because it may take a few days to inundate from rivers as well as coast to know the combined impact, which may not be possible to detect within ± 1-day of occurrence of the event. Secondly, a large watershed may respond within few days of occurrence of the storm event – it may take a few days' time to reach water to flow to the outlet, when it meet with coastal storms. Therefore, a lag-days to be considered to sample such events, however, it would be based on the time of concentration of the watershed and not the user defined input.

Considering a single largest precipitation observation within a year or month to the largest OWL observation within its 24-hour accumulation period, is the one, which could be categorized as a compound event, per se. However, here also the question lies whether the watershed is large/medium sized and the land use pattern of the watershed. For example, in case of a large rural (or agricultural intensive) watershed the time of concentration would depend upon the catchment area and flow path length to travel water from the remote point to the catchment outlet. Therefore, the choice of 24-hr/a day may not be adequate to model dependency between two drivers.

Based on the above two comments I find the definition of compound event sampling adopted in the manuscript is erroneous. Rather, their definition could be categorized to solely a multivariate interpretation of extremes.

R1.2 This study considers the co-occurrence of precipitation and high marine water in a semi-arid, highly urbanized coastal watersheds. The key distinction is that the precipitation component causes pluvial (i.e., surface water flooding from intense rainfall) rather than fluvial (i.e., riverine) flooding and occurs on much shorter time scales. The United State Geological Survey (USGS) and the United States Department of Agriculture (USDA) classify individual watersheds with hydrologic unit codes (HUC). HUC10 officially represent watersheds. The time of concentration (T_c) for each HUC watershed were estimated using digital terrain data and the Kirpich (Kirpich, 1940), FAA (1970), and SCS Lag (1975) methods that have been previously applied in southern California coastal watersheds (e.g., Kang et al., 2008, Parker et al., 2019). HUC10 watershed T_c range is ~1-10 hours depending on individual watershed characteristics. All locations are less than 24-hour T_c. From literature perspective, storm surge and precipitation 24 hour windows are commonly used (e.g., Wahl et al. 2015, Xu et al., 2014, Xu et al., 2019, Yang et al., 2020). Given both the rapid T_c of these particular watersheds and the literature, the authors believe the 24-hour time window is appropriate.

Q1.3 Further, for wet season monthly maximum and wet season monthly coinciding method of samples, the sampling methods are not properly described.

R1.3 The authors have provided additional detail and substantially updated the text to more expansively describe the sampling methods. The new text in L110-114 of the revised manuscript is:

Wet season monthly maximum pairs the single largest precipitation and OWL observations within each wet season month. Although strictly speaking maximum parings (annual or wet season) do not technically represent an observed compound event since the co-occurrence of precipitation and water levels follows the FEMA guidance for considering a “worst case scenario" approach (FEMA, 2016c). Wet season monthly coinciding sampling pairs the single largest precipitation observation within each wet season month to the largest OWL observation within its 24-hour accumulation period, providing more realistic pairs compared to maximum sampling.

Q1.4 It is not clear whether these two samples follow the iid behavior. This is because as the sampling was performed only based on wet seasons, if they are sampling large number of events then the method fails to preserve iid assumption for frequency analysis. This would calls for nonstationary method, instead of stationary method adopted here.

For example, in Page 24, Line 309: By sampling entire wet season, you would not be able to sample iid events; also it is not the true representative of rare events.

R1.4 Our assumption for iid behavior is consistent with current literature (Mendoza and Jiménez, 2006, Li et al., 2014, Kapelonis et al., 2015, Shao et al., 2020) considering fast time scale events. These studies ensure iid using a minimum separation window between events on the scale from hours to 5-days.

Similarly, in our study the separation window between events is sufficiently large to ensure iid given we sample a single event per year, when using annual sampling, and once per month during the wet season (October to March), when using wet season sampling. Additionally, the events occurring in the watersheds within this study are fast responding and will occur within a 24 hour window (please see response R1.2).

Q2.1. The authors have used the vast array of copulas, based on their simulation they infer that the three families are describing the paired event characteristics sufficiently well. However, they have not shown any formal goodness-of-fit that suggests credibility of selected copulas to fit the multivariate extremes. Neither, they discuss out of the three which copula family performs the best for modelling multivariate extremes.

R2.1 The Bayesian Information Criterion (BIC) and Maximum Likelihood (ML) goodness of fit metrics are provided in section 3.5, equations 16-19. Figure 4 presents the ML values for each tested copula and sampling strategy. ML values describe how well the parameters of an assumed distribution represent a given sample (Chapter 7.5 and 7.6 from DeGroot et. al., 2014), and have been previously used as a goodness of fit metric (e.g., Sadegh et al., 2017, 2018). The authors refrain from explicitly recommending a particular copula since a single 'best' copula did not emerge for given samplings across all three sites. However, the authors point out the most well fit (and poorly fit) copulas and their impacts on return periods (e.g., Tables 5, 6). Akaike Information Criterion (AIC, Figure A3) and BIC (Figure A4) goodness of fit metrics have been added to the Appendix and below.

Q3.1 Although they have coined a term "Structural" in the return period concept, but in the discussion or results section, I could not find any dedicated section that distinguishes structural or failure probability concepts.

R3.1 Section 3.2.5 specifically introduces the "Structural" hazard scenario and associated literature. The text has been substantially updated to provide additional context. The revised text (L171-175) now reads:

The "Structural" scenario considers the probability of an output from a structural function, Ψ(x) exceeding a design load or capacity (z) (Salvadori et al., 2016). For example, De Michele et al. (2005) and Volpi and Fiori (2014) used a structural function to evaluate a dam spillway while Salvadori et al. (2015) considers the preliminary design of rubble mound breakwater. In this work, the structural failure function focuses on the question ``what is the probability of a water level forcing tide valve closure and subsequent flooding during a precipitation event?".

Q4.1 15. Line 327: 'OR' scenario may not provide an accurate estimate of compounding condition since in this case, only one of the variable is assumed to exceed over the other. The simultaneous or joint exceedance of variables are considered in 'AND' scenario case only.

R4.1 The authors agree with the reviewer's comment that the 'AND' scenario provides only the joint exceedance of both variables. The 'OR' scenario probability space encompasses three regions: events dominated by the primary variable, events dominated by both variables, and events dominated by the secondary variable (Figure 1, Salvadori et al., 2016). Literature presents multiple studies that have considered the "OR" hazard scenarios (Table 1, Salvadori et al. 2016).

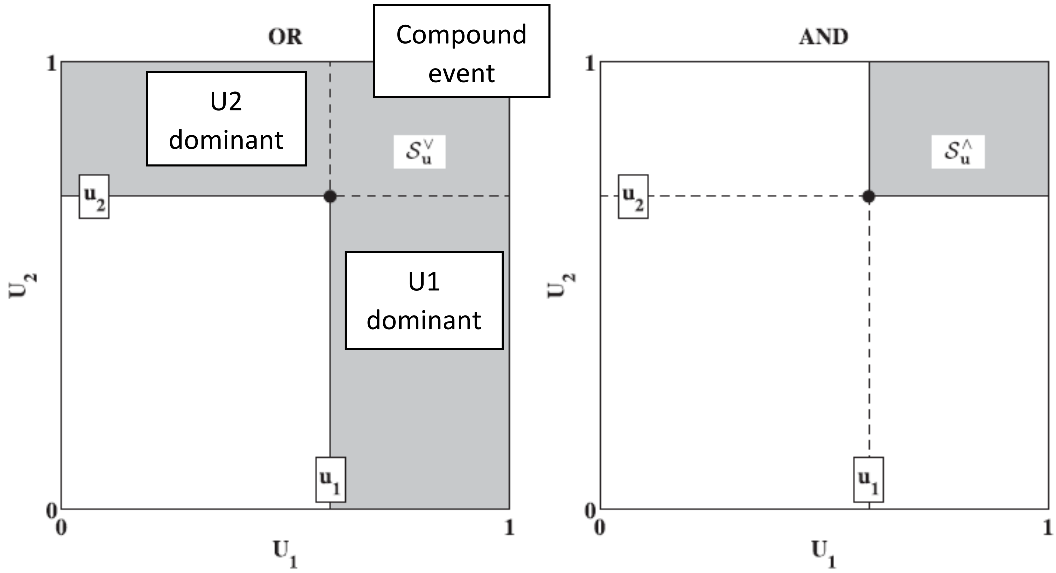

Figure 1 - 'OR' and 'And' Probability spaces. Adapted from Salvadori et al., (2016). All definitions are from Salvadori et al., (2016).

Q5.1 Page 3, line # 60: Although authors have pointed annual maxima sampling generates a worst case scenario; this has been followed earlier. Moftakhari et al. (2017) modeled failure probability for the extreme scenario aiding disaster response considering concurrence of the largest annual freshwater inflow to the lower estuary and the corresponding largest observed hourly water level within ±1 day.

R5.1 Moftakhari et al. (2017) explores compound coastal-fluvial (marine-riverine) events. In contrast, our study explores a coastal-pluvial type event where high marine water levels and precipitation co-occur. In both cases indeed, the annual maximum presents a 'worst-case'. However, the results are distinct since the events considered fundamentally differ.

Q6.1 Page 3, line # 65: cold & wet fall season (Ganguli et al., 2019a.b), where authors compare coastal compound floods relative to winter seasonal peak discharge and found larger amplifications upon considering compounding effects than solely accounting for high winter (November-March) discharge over northwestern Europe.

R6.1 The authors would like to thank the reviewer for pointing out this work and have added the suggested citations to the text. The definition of "winter" used in Ganguli et. al., 2019a, 2019b does not strictly align with the definition of 'wet season' given the differences in climatology, hydrology, and coastal dynamics between Southern California and Northwestern Europe.

Q7.1 Page 3, line # 89: The authors have claimed that in current compound flood literature, the terms tides and storm surges are interchanged; this is not true. Check Devlin et al., (2017); Ganguli and Merz, (2019); Ganguli et al. (2020) as a reference. While in Ganguli and Merz (2019) observed high coastal water level was used that includes tides, storm surges and wave setup, for Ganguli et al. (2020) only meteorologically driven skew surge was considered.

R7.1 The authors thank the reviewer for pointing this error out. The term 'storm surge' should have been 'water level'. The text has been updated to reflect this change and now reads. "It is worth noting, that within the body of compound flooding literature, the terms tide and water level may be interchanged (e.g., Lian et al., 2013; Xu et al., 2014; Tu et al., 2018; Xu et al., 2019, Yang et al., 2020)."

Q8.1 Table 1: several references are missing:

a. River discharge and water level: (Ganguli and Merz, 2019a, b).

b. River discharge and storm surge: (Ganguli et al., 2020).

c. River discharge and volume (Reddy and Ganguli, 2012).

d. Rainfall and tide: (Bevacqua et al., 2020).

e. Combination of river discharge, volume, and duration (Ganguli and Reddy, 2013).

R8.1 The authors thank the reviewer for providing these studies. Ganguli and Merz, (2019a, b) was added to "River discharge and water level". Ganguli et al., (2020) was added to "River discharge and storm surge". Reddy and Ganguli, (2012) and Ganguli and Reddy, (2013) were added to "Combination of river discharge, volume, and duration". Bevacqua et al., (2020) was added to "Rainfall and tide".

Q9.1 Fig. 2 Captions: Expand each of the terms, AM, AC, WMM, WMC, SM, S, and SD.

R9.1 The reviewer's suggestion has been adopted and the caption updated accordingly.

Q10.1 Page 8: Line #148: separates supercritical vs sub-critical region.

R10.1 `The text has been updated to include the reviewer's suggestion and now reads: The "Kendall" (K) scenario highlights an infinite set of OR events that separate the subcritical (i.e, "safe") and supercritical (i.e., "dangerous") statistical regions.

Q11.1 In Eqn. 12: it was not shown how individual marginal CDFs were included in the expression to get the Kendall 'AND' return period.

R11.1 Equation 12 is related to the "Survival Kendall". The Survival Kendall scenario (Equation 11) utilizes a "survival copula" which utilizes "survival CDFs" detailed in Section 3.2.2. Additionally, the reader is referred to Salvadori et al. (2013), Salvadori and De Michele (2004), and Serinaldi (2015) for additional information.

Q12.1 In Eqn. 13: Are you considering Kendall's return period to define this case? If it is, it should be properly expressed.

R12.1 Equation 13 described the "Structural" scenario (e.g., De Michele et al., 2005, Salvadori et al., 2016). Section 3.2.5 specifically introduces the "Structural" hazard scenario and associated literature. Section 4.4 details the application of the "Structural" scenario which, in this case, follows a Conditional 1 type event. Please also see **R3.1** for additional details.

Q13.1 Page 10: Line 210, the choice of marginal distribution depends on the tail property: if the shape of density function show fast decaying pattern; exponential /Gumbel (GEV-I) distribution would be good; however, for long upper tail, heavier tail distribution is being preferred. Therefore, I would suggest to summarize the basic summary statistics of driver variables, including their skewness and excess kurtosis. Based on that they can make inferences, why certain univariate marginal fits the best.

R13.1 Marginal BIC values, which MvCAT uses to select optimal distributions, for OWL (Figure A1) and precipitation (Figure A2) are now provided in the Appendix and, for the reviewer's convenience, below. Additionally, marginal distributions were independently tested using a visual chi square test (e.g., Requena et al., 2013).

Q14.1 Page 12: line #245, It is not clear if the goodness-of-fit is performed for the choice of copulas? Different copulas behaves differently and one has to select a copula class, which can represent the sample on the basis of their ability to simulate complete vs upper tail dependences. This concern I have also raised as a major comment.

R14.1 Please see responses R2.1 for a comprehensive explanation of goodness-of-fit metrics.

Q15.1 Page 18: Line 269, We only consider compounding aspects if they occur within a limited or close time intervals, for example, within a week of occurrence; because of large sized of a catchment, it is physically may not be possible that both events co-occur simultaneously; a lag effect to be considered during event sampling. The response of medium to large sized watershed to a rain event is proportional to the time of concentration of the watershed. This can be utilized to estimate the lag effect.

R15.1 Precipitation drives a pluvial flooding event on the relatively small urbanized coastal watersheds which exhibit a rapid time of concentration on the order of 1-10 hours. The reviewer is referred to **R1.2** for additional information regarding the watershed time of concentration.

Appendix

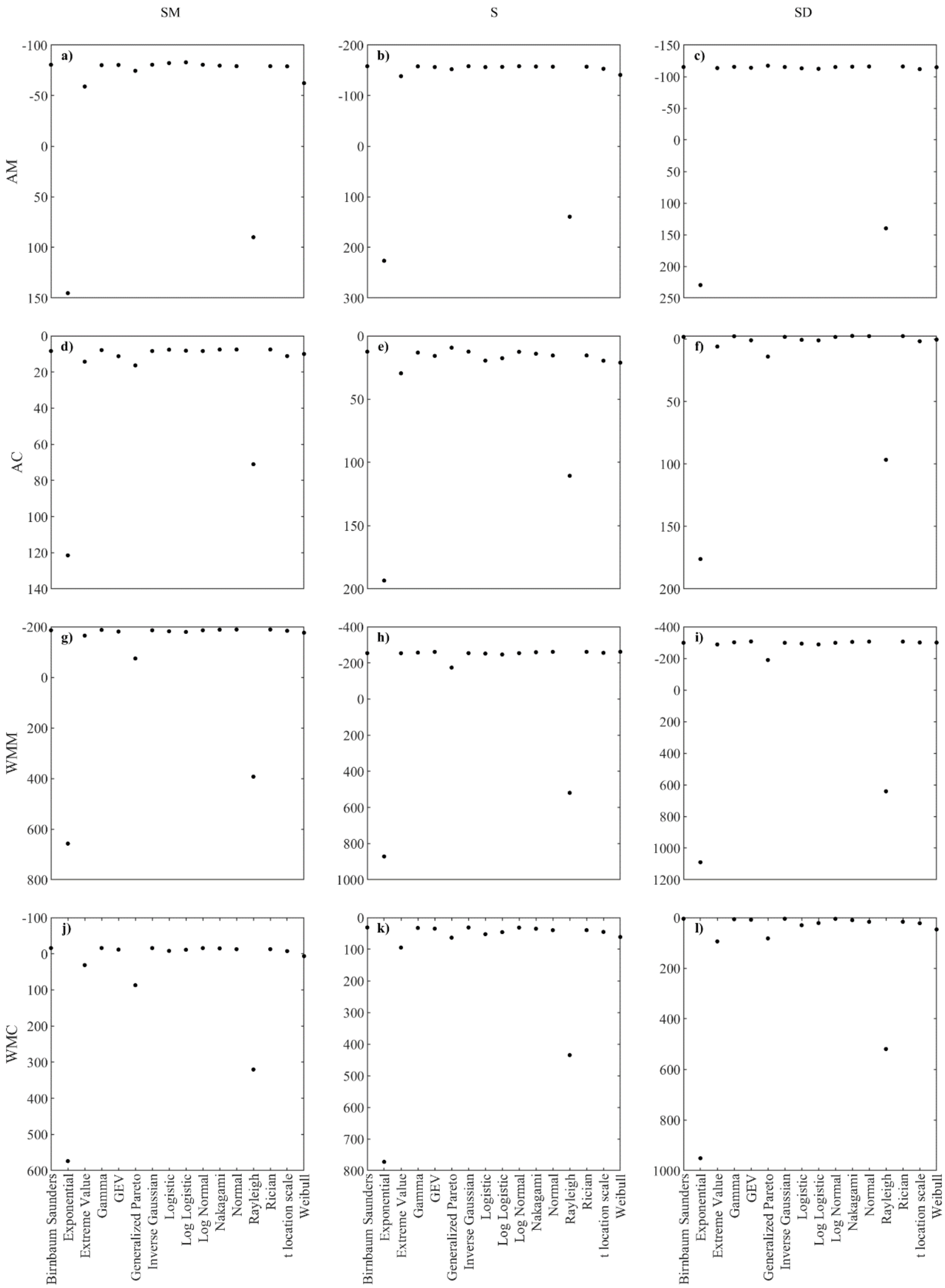

Figure A1 – Marginal OWL BIC values per fitted copula for SM (left column), S (middle column), and SD (right column) (a), (b), (c) AM, (d), (e), (f) AC, (g), (h), (i) WMM, and (j), (k), (l) WMC. Y-axis is orientated to display best BIC (top) to worst BIC (bottom).

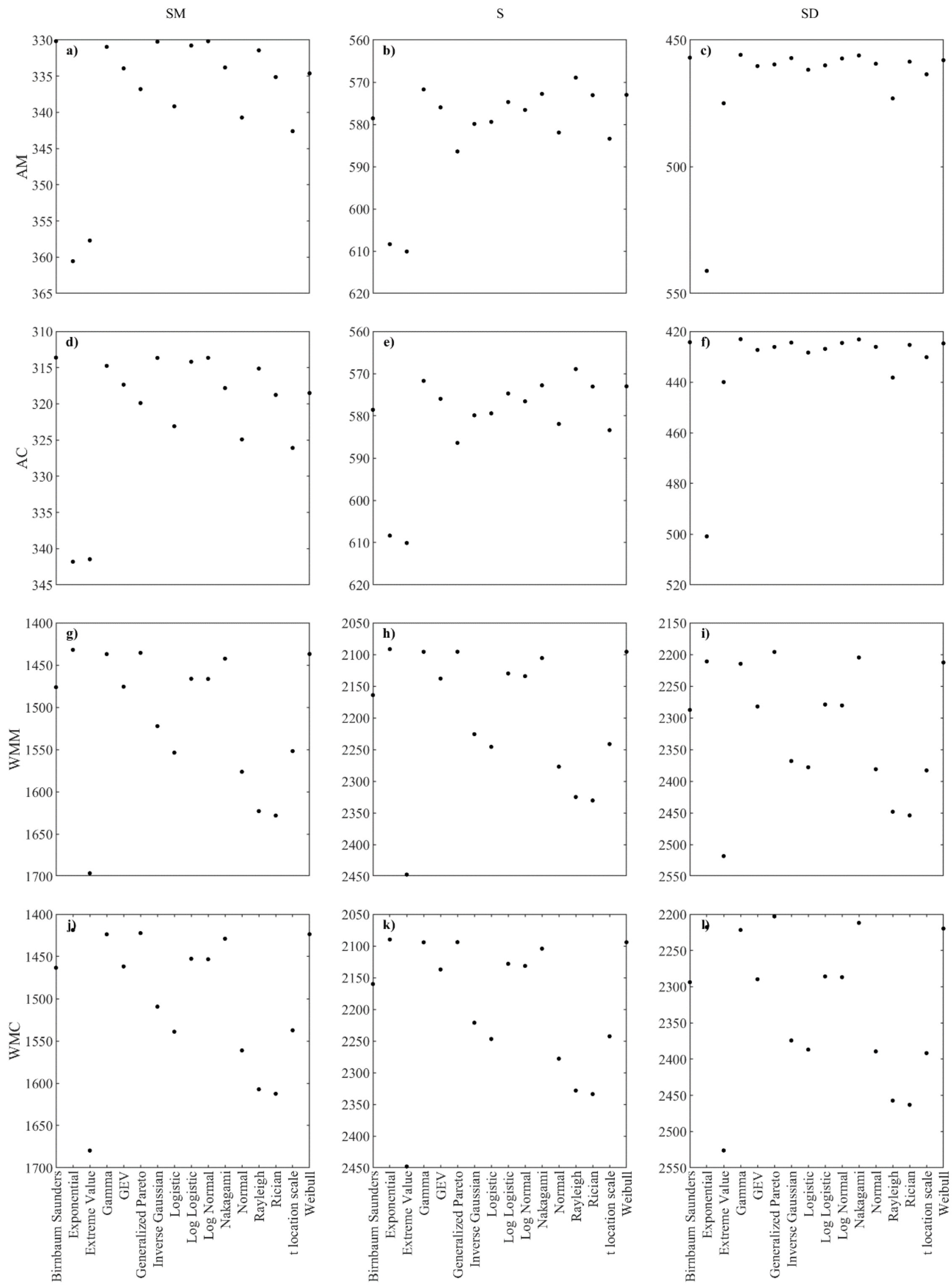

Figure A2 – Marginal precipitation BIC values per fitted copula for SM (left column), S (middle column), and SD (right column) (a), (b), (c) AM, (d), (e), (f) AC, (g), (h), (i) WMM, and (j), (k), (l) WMC. Y-axis is orientated to display best BIC (top) to worst BIC (bottom).

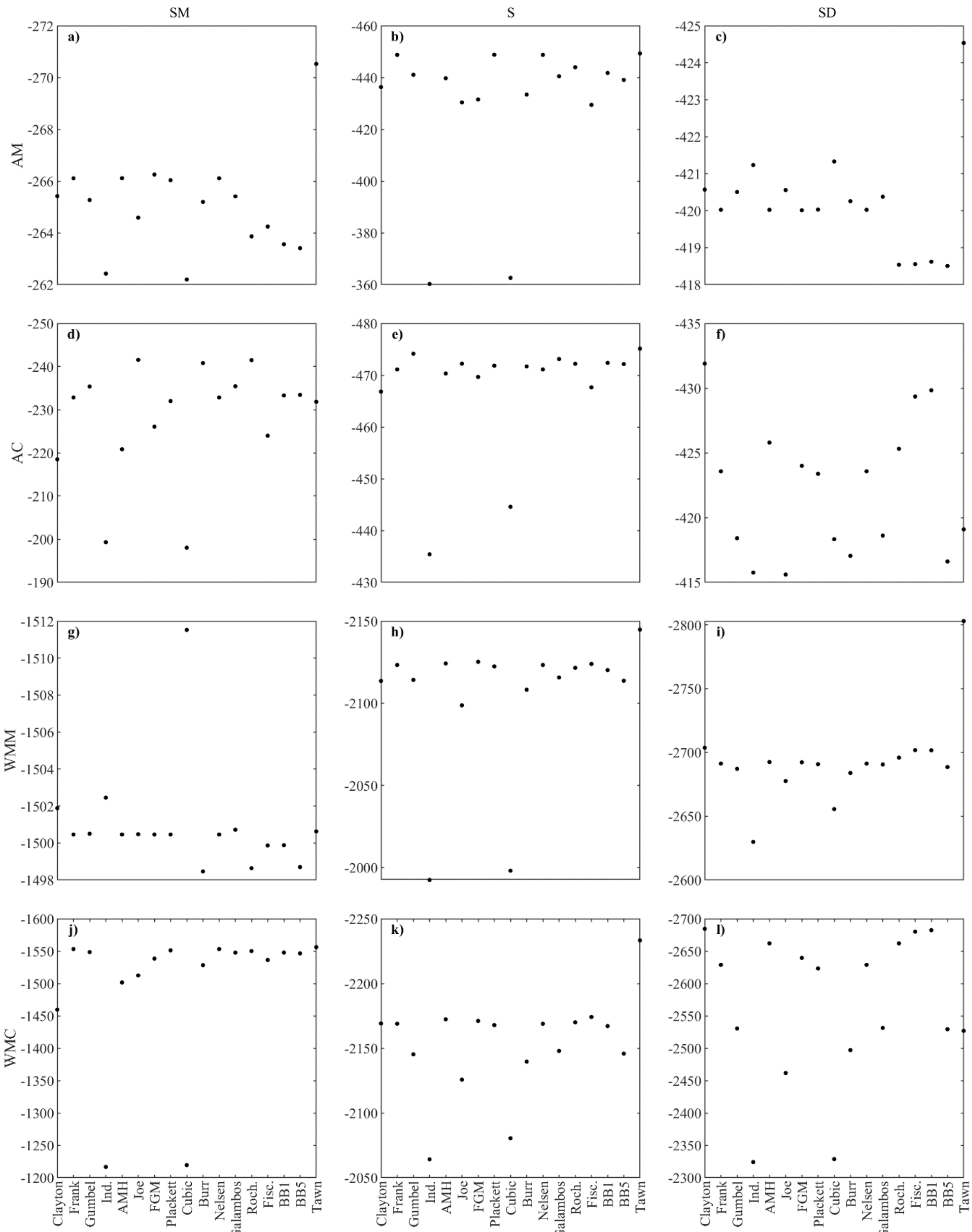

Figure A3 – Copula AIC values per fitted copula for SM (left column), S (middle column), and SD (right column) (a), (b), (c) AM, (d), (e), (f) AC, (g), (h), (i) WMM, and (j), (k), (l) WMC. Y-axis is orientated to display best AIC (top) to worst AIC (bottom).

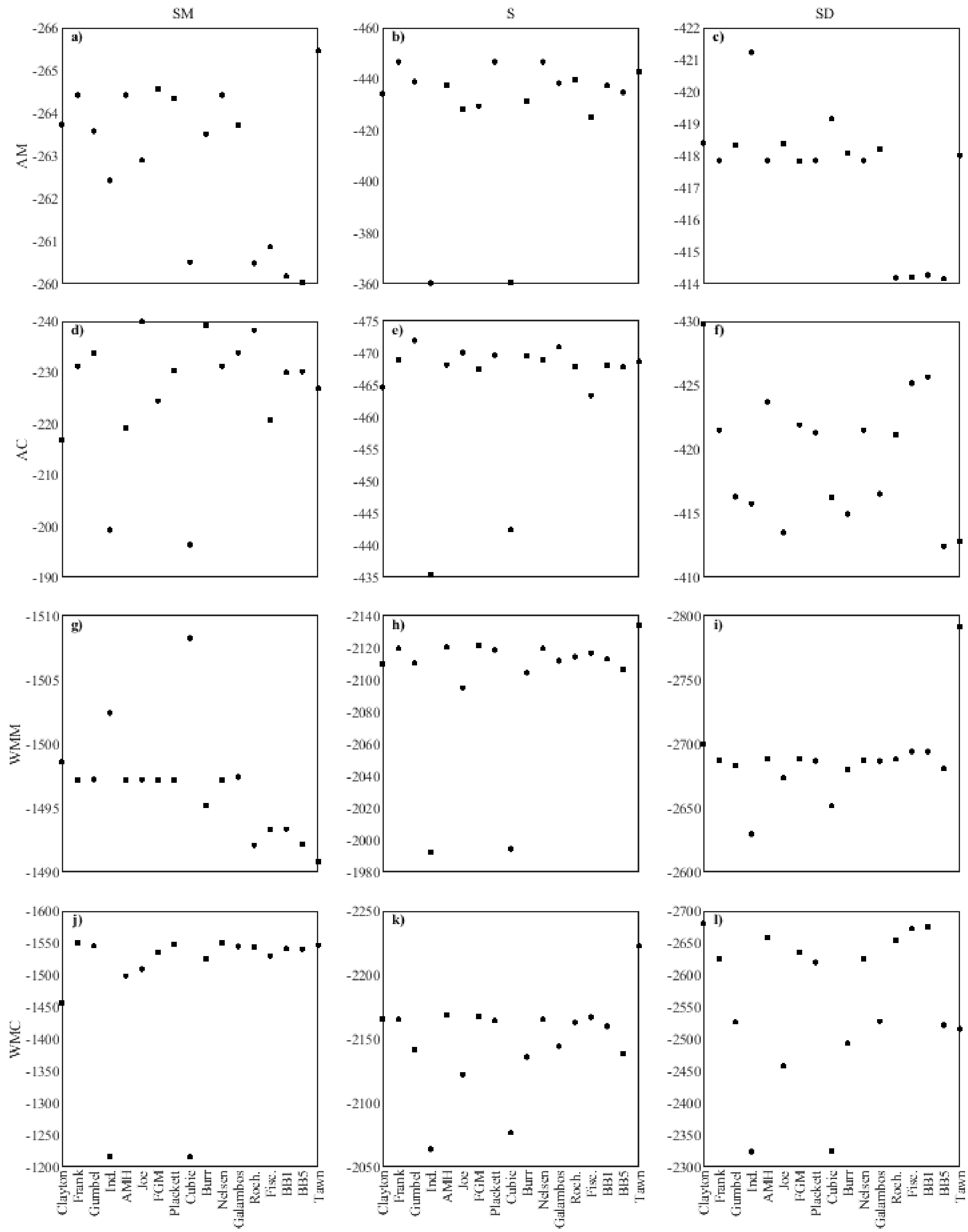

Figure A4 – Copula BIC values per fitted copula for SM (left column), S (middle column), and SD (right column) (a), (b), (c) AM, (d), (e), (f) AC, (g), (h), (i) WMM, and (j), (k), (l) WMC. Y-axis is orientated to display best BIC (top) to worst BIC (bottom).

References

- Bevacqua, E., Vousdoukas, M. I., Zappa, G., Hodges, K., Shepherd, T. G., Maraun, D., Mentaschi, L., and Feyen, L.: More meteorological events that drive compound coastal flooding are projected under climate change, Communications earth & environment, 1, 1–11, 2020.
- De Michele, C., Salvadori, G., Canossi, M., Petaccia, A., and Rosso, R.: Bivariate statistical approach to check adequacy of dam spillway, Journal of Hydrologic Engineering, 10, 50–57, 2005.
- DeGroot, M. H. and Schervish, M. J.: Probability and Statistics, Pearson Education, 4 edn., 2014.
- Diez, D. M., Barr, C. D., & Cetinskaya-Rundel, M.: Open Intro Statistics (3rd edition)., 2015.
- FEMA: Guidance for Flood Risk Analysis and Mapping: Coastal Flood Frequency and Extreme Value Analysis, https://www.fema.gov/sites/default/files/2020-02/Coastal_Flood_Frequency_and_Extreme_Value_Analysis_Guidance_Nov_2016.pdf, 2016.
- Ganguli, P. and Reddy, M. J.: Probabilistic assessment of flood risks using trivariate copulas, Theoretical and applied climatology, 111, 341–360, 2013.
- Ganguli, P. and Merz, B.: Trends in compound flooding in northwestern Europe during 1901–2014, Geophysical Research Letters, 46, 10 810–10 820, 2019a.
- Ganguli, P. and Merz, B.: Extreme coastal water levels exacerbate fluvial flood hazards in Northwestern Europe, Scientific reports, 9, 1–14, 2019b.
- Ganguli, P., Paprotny, D., Hasan, M., Güntner, A., and Merz, B.: Projected changes in compound flood hazard from riverine and coastal floods in northwestern Europe, Earth's Future, 8, e2020EF001 752, 2020.
- Guirguis, K. J., & Avissar, R.: A precipitation climatology and dataset intercomparison for the western United States. Journal of Hydrometeorology, 9(5), 825-841, 2008.
- Kang, J.H., Kayhanian, M. and Stenstrom, M.K.: Predicting the existence of stormwater first flush from the time of concentration, Water research, 42(1-2), pp.220-228, 2008.
- Kapelonis, Z. G., Gavriliadis, P. N., and Athanassoulis, G. A.: Extreme value analysis of dynamical wave climate projections in the Mediterranean Sea, Procedia Computer Science, 66, 210-219, 2015.
- Kirpich, Z. P.: Time of concentration of small agricultural watersheds, Civil engineering, 10(6), 362, 1940.
- Lian, J., Xu, K., and Ma, C.: Joint impact of rainfall and tidal level on flood risk in a coastal city with a complex river network: a case study of Fuzhou City, China, Hydrology and Earth System Sciences, 17, 679, 2013.
- Li, F., Van Gelder, P. H. A. J. M., Ranasinghe, R. W. M. R. J. B., Callaghan, D. P., and Jongejan, R. B.: Probabilistic modelling of extreme storms along the Dutch coast, Coastal Engineering, 86, 1-13, 2014.
- Leonard, M., Westra, S., Phatak, A., Lambert, M., van den Hurk, B., McInnes, K., Risbey, J., Schuster, S., Jakob, D., and Stafford-Smith, M.: A compound event framework for understanding extreme impacts, Wiley Interdisciplinary Reviews: Climate Change, 5(1), 113-128, 2014.
- Mendoza, E. T., and Jiménez, J. A.: A storm classification based on the beach erosion potential in the Catalonian Coast, In Coastal Dynamics 2005: State of the Practice, pp.1-11, 2006.
- Moftakhari, H. R., Salvadori, G., AghaKouchak, A., Sanders, B. F., and Matthew, R. A.: Compounding effects of sea level rise and fluvial flooding, Proceedings of the National Academy of Sciences, 114, 9785–9790, 2017.

- Parker, S.R., Adams, S.K., Lammers, R.W., Stein, E.D. and Bledsoe, B.P.: Targeted hydrologic model calibration to improve prediction of ecologically-relevant flow metrics, Journal of Hydrology, 573, pp.546-556, 2019.
- Reddy, M. J. and Ganguli, P.: Bivariate flood frequency analysis of upper Godavari River flows using Archimedean copulas, Water Resources Management, 26, 3995–4018, 2012.
- Requena, A. I., Mediero, L., and Garrote, L.: A bivariate return period based on copulas for hydrologic dam design: accounting for reservoir routing in risk estimation, Hydrology and Earth System Sciences, 17(8), 3023-3038, 2013.
- Sadegh, M., Ragno, E., and AghaKouchak, A.: Multivariate Copula Analysis Toolbox (MvCAT): describing dependence and underlying uncertainty using a Bayesian framework, Water Resources Research, 53, 5166–5183, 2017.
- Sadegh, M., Moftakhari, H., Gupta, H. V., Ragno, E., Mazdiyasni, O., Sanders, B., Matthew, R., and AghaKouchak, A.: Multihazard scenarios for analysis of compound extreme events, Geophysical Research Letters, 45, 5470–5480, 2018.
- Salvadori, G. and De Michele, C.: Frequency analysis via copulas: Theoretical aspects and applications to hydrological events, Water re-sources research, 40, 2004.
- Salvadori, G., Durante, F., and De Michele, C.: On the return period and design in a multivariate framework, Hydrology and Earth System Sciences, 15, 3293–3305, 2011.
- Salvadori, G., Durante, F., and De Michele, C.: Multivariate return period calculation via survival functions, Water Resources Research, 49, 2308–2311, 2013.
- Salvadori, G., Durante, F., Tomasicchio, G. R., and D'Alessandro, F.: Practical guidelines for the multivariate assessment of the structural risk in coastal and off-shore engineering, Coastal Engineering, 95, 77-83, 2015.
- Salvadori, G., Durante, F., De Michele, C., Bernardi, M., and Petrella, L.: A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities, Water Resources Research, 52, 3701–3721, 2016.
- Seneviratne, S., Nicholls, N., Easterling, D., Goodess, C., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C., and Zhang, X.: Changes in climate extremes and their impacts on the natural physical environment, in: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change, edited by Field, C., Barros, V., Stocker, T., Qin, D., Dokken, D., Ebi, K., Mastrandrea, M., Mach, K., Plattner, G.-K., Allen, S., Tignor, M., and Midgley, P., pp. 109–230, Cambridge University Press, Cambridge, UK, and New York, NY, USA, 2012.
- Serinaldi, F.: Dismissing return periods!, Stochastic Environmental Research and Risk Assessment, 29, 1179–1189, 2015.
- Shao, Z., Liang, B., and Gao, H.: Extracting independent and identically distributed samples from time series significant wave heights in the Yellow Sea, Coastal Engineering, 158, 103693, 2020.
- Shumway, R. H. and Stoffer, D. S.: Time series analysis and its applications, 3, New York: springer, 2000.
- Tu, X., Du, Y., Singh, V. P., and Chen, X.: Joint distribution of design precipitation and tide and impact of sampling in a coastal area, International Journal of Climatology, 38, e290–e302, 2018.
- U.S. Geological Survey (USGS) and U.S. Department of Agriculture (USDA), Natural Resources Conservation Service, Federal Standards and Procedures for the National Watershed Boundary

Dataset (WBD) (4 ed.): Techniques and Methods 11–A3, 63 p., https://pubs.usgs.gov/tm/11/a3/., 2013.

- Volpi, E. and Fiori, A.: Hydraulic structures subject to bivariate hydrological loads: Return period, design, and risk assessment, Water Resources Research, 50(2), 885-897, 2014.
- Wahl, T., Jain, S., Bender, J., Meyers, S. D., and Luther, M. E.: Increasing risk of compound flooding from storm surge and rainfall for major US cities, Nature Climate Change, 5, 1093, 2015.
- Xu, H., Xu, K., Lian, J., and Ma, C.: Compound effects of rainfall and storm tides on coastal flooding risk, Stochastic Environmental Research and Risk Assessment, 33, 1249–1261, 2019.
- Xu, K., Ma, C., Lian, J., and Bin, L.: Joint probability analysis of extreme precipitation and storm tide in a coastal city under changing environment, PLoS One, 9, e109 341, 2014.
- Yang, X., Wang, J., and Weng, S.: Joint Probability Study of Destructive Factors Related to the “Triad” Phenomenon during Typhoon Events in the Coastal Regions: Taking Jiangsu Province as an Example, Journal of Hydrologic Engineering, 25, 05020 038, 2020.
- Zscheischler, J., Westra, S., Van Den Hurk, B.J., Seneviratne, S.I., Ward, P.J., Pitman, A., AghaKouchak, A., Bresch, D.N., Leonard, M., Wahl, T., and Zhang, X.: Future climate risk from compound events, Nature Climate Change, 8(6), 469-477, 2018.