Bivariate Design Storms Don't Work: A theory-driven investigation and stylized case study

Nathan Geldner[1]

[1]The Barbara Geldner Foundation (not yet incorporated)

Context

To model compound flood risk, researchers and practitioners are increasingly using bivariate design storms—design events derived from a bivariate distribution of flood drivers (e.g. rainfall intensity, storm tide) that are assumed to correspond to specific return periods (e.g. a 100-year storm) based only on the distribution of drivers and without regard for the actual resutling flood depths[1][2][3][4]. According to these methods, a 100-year storm is one where there is a 1/100 chance of seeing a storm with **both higher rainfall and higher storm tide** in a given year. Unfortunately, this has **almost nothing** to do with the 100-year storm that actually matters for flood risk analysis, which is one where there is a 1/100 chance of seeing **higher flood depths** in a given year. These methods are highly misleading, wildly mis-estimate flood risk, and typically involve plainly incorrect mathematics.

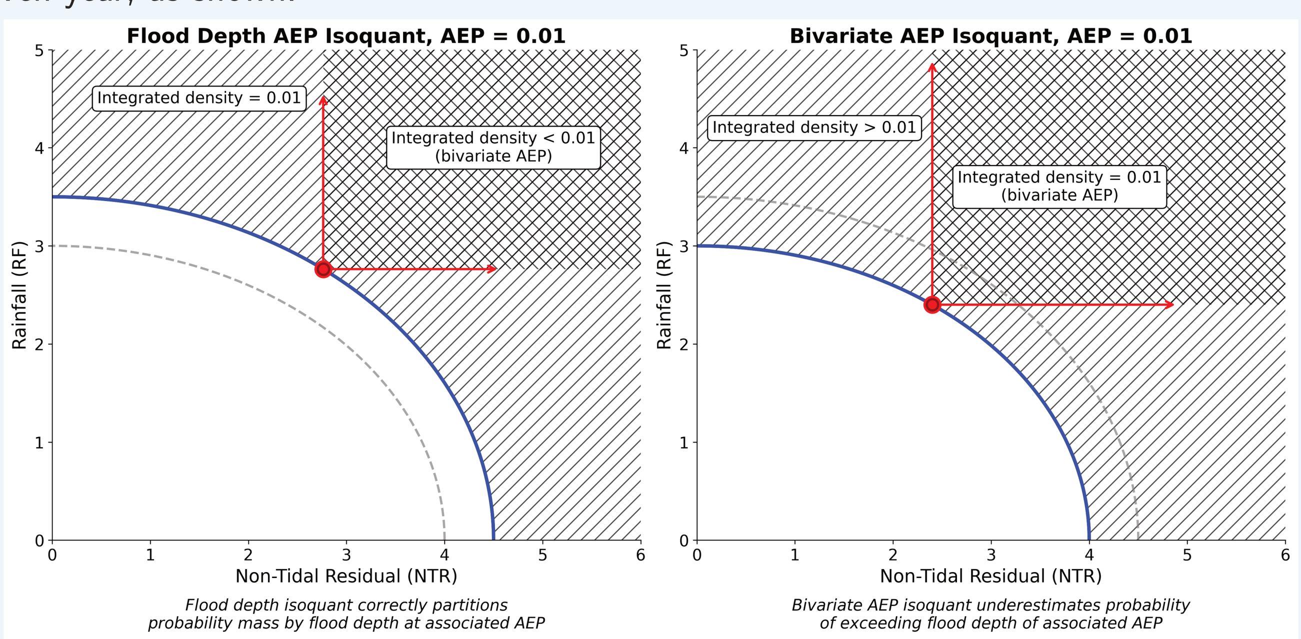
We address three core issues, which together are entirely fatal to the validity of any attempt to generate bivariate design storms for compound flood risk management, showing that response-based approaches accounting for more than two scalar factors are strictly necessary.

Core Issues:

- The bivariate annual exceedance probability (AEP) used by these methods is wildly different from the flood depth AEP of interest. If bivariate models were appropriate in the first place, resulting design storms would always (deterministically) be less intense than they should be.
- Published methods for generating bivariate design storms have clear mathematical errors on the order of 2+2=5. Particularly, these methods fit one bivariate distribution for storms with one extreme forcing (e.g. rainfall), another bivariate distribution for storms with another extreme forcing (e.g. storm tide), and use a plainly incorrect method to combine them.
- Bivariate models simplify storm attributes to two scalar driver intensities. This is simply not enough information. Flood outcomes are highly sensitive to factors which are excluded from bivariate analysis, meaning that any bivariate model will be largely unhelpful in understanding flood risk

Bivariate AEPs inherently underestimate flood depth AEPs

Bivariate design storms are sampled from bivariate AEP isoquants. In the context of rainfall and tidally driven events, this method defines a 100-year storm as one with a 0.01 probability of having both its rainfall and storm tide exceeded in a given year, as shown.

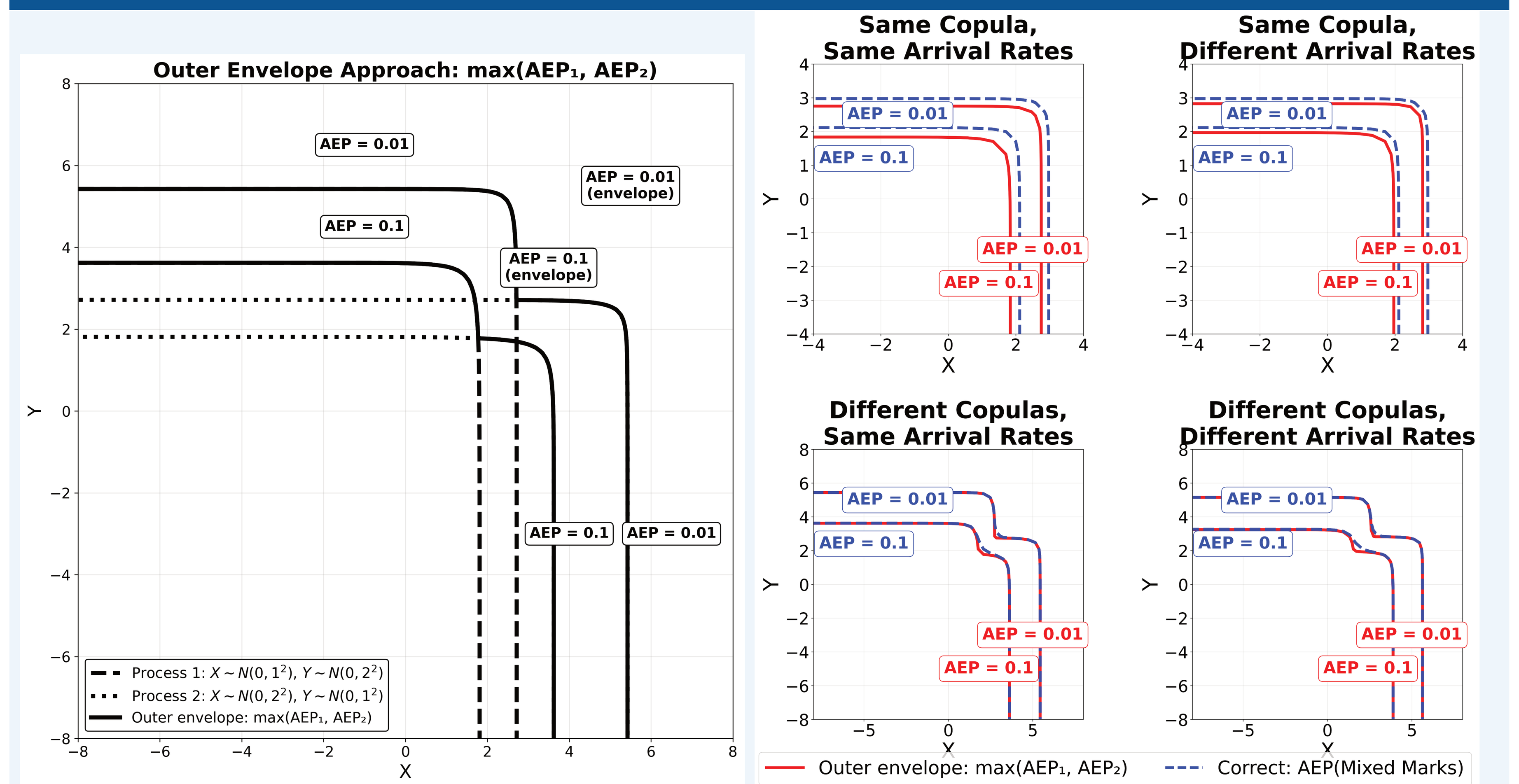


As a consequence, there is more than a 0.01 probability of seeing a storm exterior to the bivariate AEP isoquant in a given year. In contrast, the set of storms which generate 100-year flood depths by definition are such that there is exactly a 0.01 probability of seeing a storm with greater flood depths (i.e. a storm exterior to the flood depth isoquant). Therefore, if a bivariate model could fully characterize flood risk, the e.g. 100-year bivariate design storm would be guaranteed to produce less than 100-year flood depths.

Published methods combine processes incorrectly

Published methods on bivariate design storms use peak-over-threshold sampling to train two separate bivariate distributions—in the case of rainfall and storm tide, one distribution is trained on events with extreme rainfall, and another is trained on events with extreme storm tide. Bivariate AEP isoquants are derived for select return periods for each distribution, and the two resulting AEP isoquants for a given return period are combined to what is described by the authors as an "outer envelope", which is equivalent to saying that the annual exceedance probability of a given storm in the combined process is equal to the maximum of its annual exceedance probabilities between the two distributions.

Published methods combine processes incorrectly (continued)



The authors of these methods assert that this is a conservative heuristic. It is not conservative, and no heuristic is necessary. Under the standard formulation of storm arrivals as a marked Poisson process in which the mark distribution is the rate-weighted mixture of the two individual sub-processes, the correct combined annual exceedance probability can be expressed simply as

$$AEP_{combined} = 1 - (1 - AEP_1)(1 - AEP_2)$$

Comparing the published approach with the correct approach shows that the "outer envelope" heuristic is not conservative, and produces less extreme design storms than the correct approach in many cases as shown above.

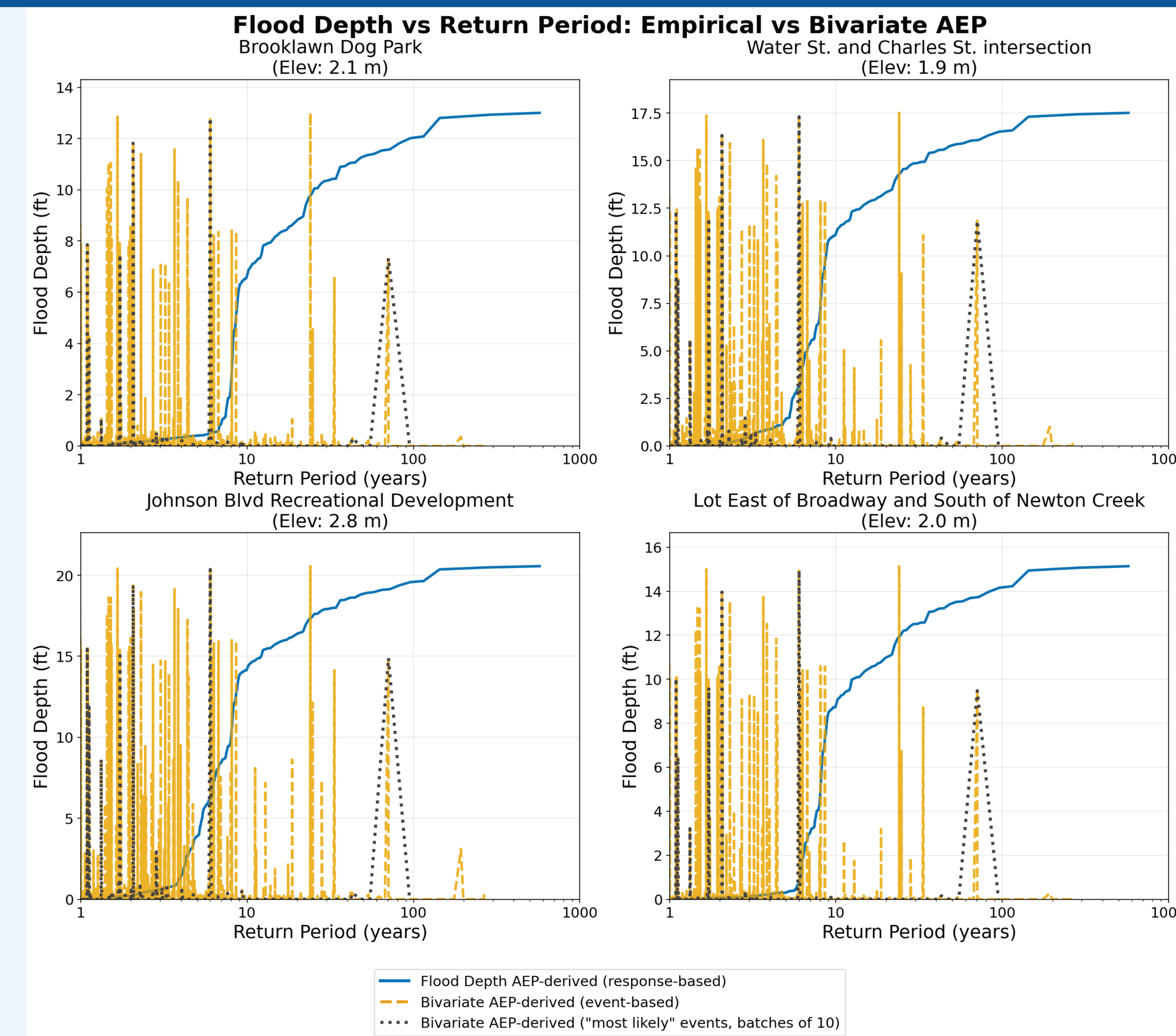
Bivariate models leave too much residual variance to be useful

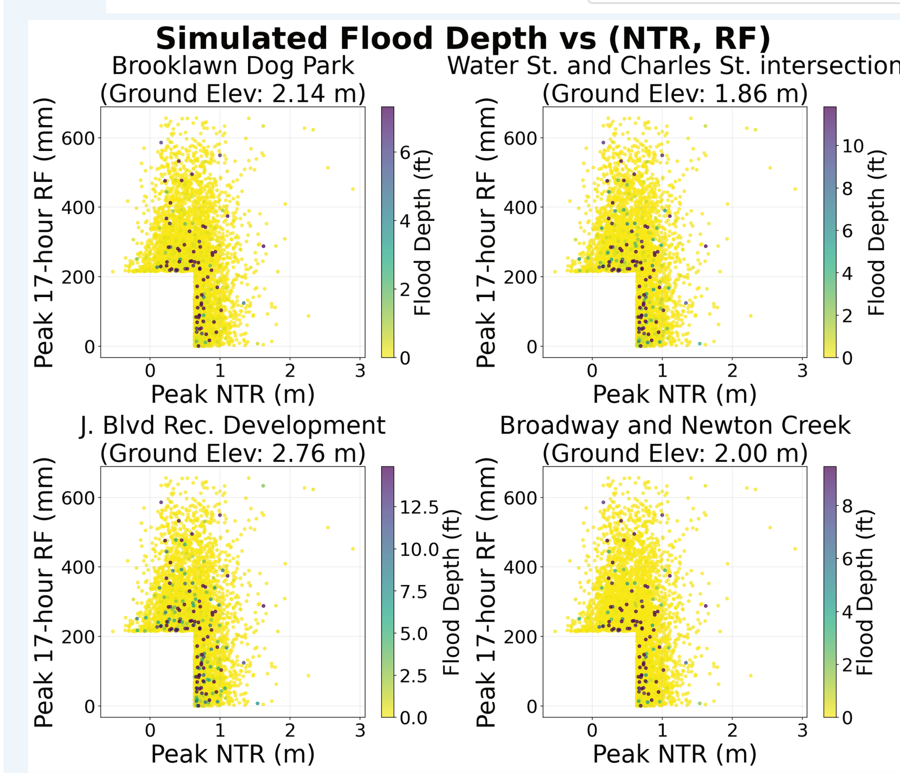
We reproduce the work of [4], which randomly samples non-tidal residual (NTR) and rainfall intensity from the bivariate copulas used to generate bivariate design storms, and constructs realistic events by using sampled values to scale up time series from randomly selected historical events with randomly selected tidal signals and lag times between peak non-tidal residual and peak rainfall intensity. Choice of rainfall accumulation time to scalarize rainfall intensity (in this case, 17 hours) is selected to maximize correlation with peak NTR, in line with previous work on bivairate design storms, though we note that this is an arbitrary and highly problematic choice. Events are simulated here with an uncalibrated SFINCS model of Gloucester City, NJ. We observe that the bivariate annual exceedance probabilities derived by this model have little to no relation to flood depth annual exceedance probabilities. We further see that the bivariate distribution of scalarized forcing intensity does not come close to fully explaining the variability of peak flood depths. This is not resolved by plotting peak flood depths as a function of more explanatory pairs of rainfall and tidal predictors—in this case 3-hour peak rainfall, which has the highest correlation with peak flood depth (.275, while 17-hour accumulation is negatively correlated), and peak tide. This strongly suggests that no bivariate model is adequate for flood risk management, and any adequate model must explicitly account for dependencies among all salient parameters including lag time between forcings and rainfall intensity at multiple accumulation times, potentially at multiple spatial scales.

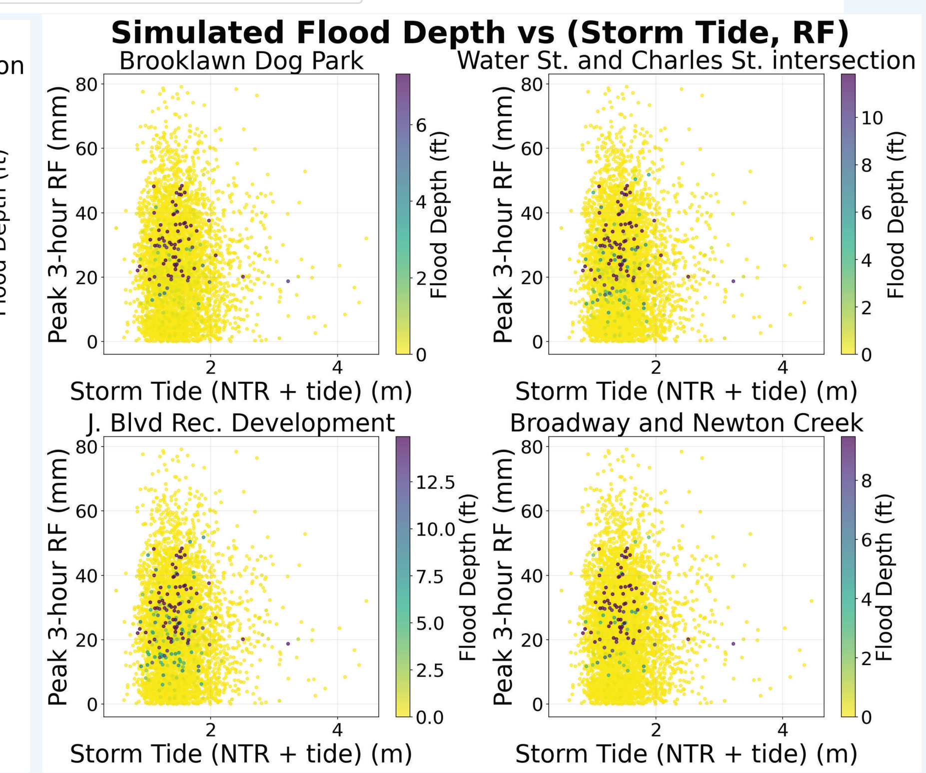
References

- Renato Amorim, Gabriele Villarini, Hanbeen Kim, Robert A Jane, and Thomas Wahl. A practitioner's approach to process-driven modeling of compound rainfall and storm surge extremes for coastal texas. *Journal of Hydrologic Engineering*, 30(5):04025025, 2025.
- Jens Bender, Thomas Wahl, Alfred Müller, and Jürgen Jensen. A multivariate design framework for river confluences. *Hydrological Sciences Journal*, 61(3):471–482, 2016.
- Hanbeen Kim, Gabriele Villarini, Robert Jane, Thomas Wahl, Shubhra Misra, and Alexander Michalek. On the generation of high-resolution probabilistic design events capturing the joint occurrence of rainfall and storm surge in coastal basins. *International Journal of climatology*, 43(2):761–771, 2023.
- Pravin Maduwantha, Thomas Wahl, Sara Santamaria-Aguilar, Robert Jane, Sönke Dangendorf, Hanbeen Kim, and Gabriele Villarini. Generating boundary conditions for compound flood modeling in a probabilistic framework. *EGUsphere*, 2025:1–29, 2025.

Bivariate models leave too much residual variance to be useful (continued)







Acknowledgments

This work was not supported by any grant or external funding.

Contact: ngeldner@thebarbarageldnerfoundation.com