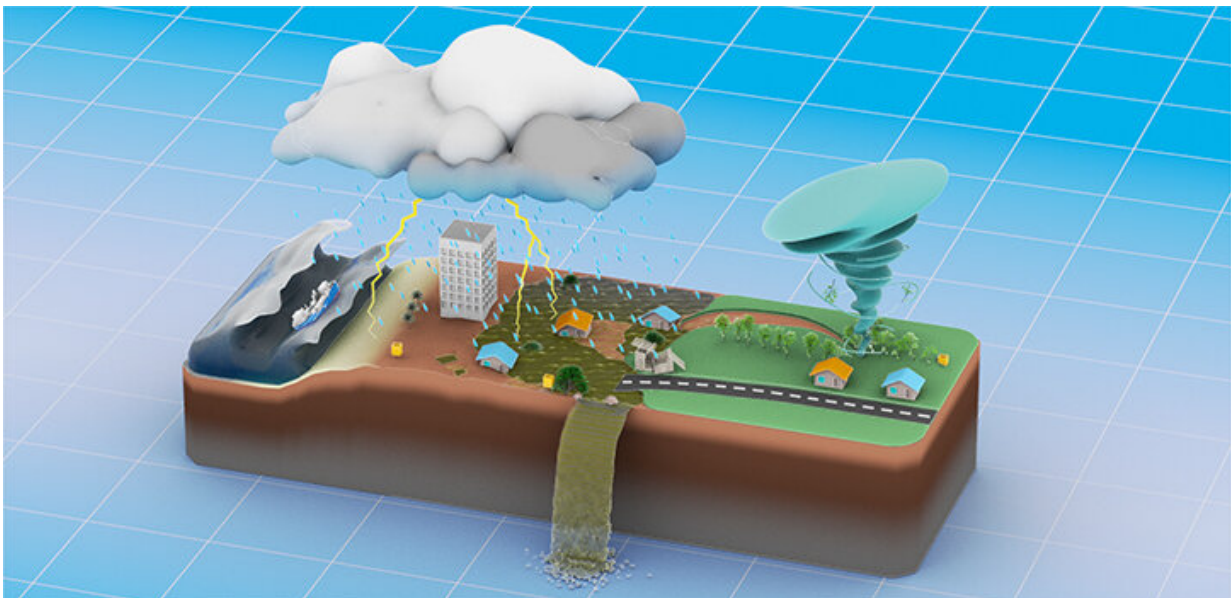


A better statistical model for environmental data

February 14 2022



KAUST statisticians have developed an improved statistical model for analyzing environmental data of extreme events, such as heavy rainfall or strong wind data. Reproduced with permission from . Credit: Wiley-VCH, KAUST, artwork by Heno Hwang

By clarifying inconsistencies in published theories and devising a flexible statistical model, KAUST researchers have established a more informed and reliable basis for selecting the most suitable statistical model for environmental data.

Despite a long history of development, the [statistical methods](#) used to analyze, process and make sense of data continue to evolve as new applications emerge. The analysis of very large environmental datasets has tested the limits of existing statistics and revealed niches where the available statistical methods fall over or could lead to erroneous results. One such area is in the analysis of extreme events, such as [heavy rainfall](#), strong winds or sea level changes.

"As the extent of events becomes more extreme, the dependence among spatial locations might decrease and eventually vanish," explains Zhongwei Zhang, Ph.D. student from KAUST's Extreme Statistics Group (extSTAT). "For example, as heavy rain becomes more extreme, the event tends to be more localized and the dependence tends to decrease between different sites. This is a typical feature of many types of [environmental data](#), and so models that correctly describe this 'asymptotic independence' are important for environmental applications."

While there are already numerous well-proven models for data characterized by asymptotic dependence, there are fewer for the independent case, particularly in the scientific literature. A [model](#) used commonly in financial analysis—the generalized hyperbolic distribution—has potential to be used for modeling asymptotic independence. However, the reported results for this model have been contradictory, with different researchers claiming the model can capture both asymptotic independence and dependence.

"The major contribution of this current work is a detailed theoretical investigation of the tail dependence properties of the multivariate generalized hyperbolic distribution model while clarifying the contradictory results in the literature," says Zhang.

Having been widely developed for financial applications, the generalized

hyperbolic distribution has been used to model financial crashes and other such extreme financial events, where the data are not necessarily asymptotically independent—financial contagion can cause many assets to fall simultaneously.

Zhang, with Raphael Huser's extSTAT team, corrected the tail description for the generalized hyperbolic distribution for asymptotic independence and, on that basis, developed a new flexible "copula" approach that models the dependence structure of a process at different locations.

"Our study shows that it is important that researchers are aware that all models have both advantages and disadvantages," notes Zhang. "If you plan to use a certain model, make sure you know its properties and limitations, especially when it comes to extrapolating outside the range of the observed data."

More information: Zhongwei Zhang et al, Modeling spatial extremes using normal mean-variance mixtures, *Extremes* (2022). [DOI: 10.1007/s10687-021-00434-2](https://doi.org/10.1007/s10687-021-00434-2)

Provided by King Abdullah University of Science and Technology

Citation: A better statistical model for environmental data (2022, February 14) retrieved 25 April 2024 from <https://techxplore.com/news/2022-02-statistical-environmental.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.