if a max-stable copula (max-stable process)[10] is applied. This means that step 5 (Fig. 1) is not sufficient.

When computing the regional risk curve of a sub-basin (Fig. 1, step 7), Jongman *et al.*[1] describe its flood hazard by one discharge variable according to their dependence model. They validate this approximation by their Supplementary Fig. 1[1]. This is questionable (Supplementary Fig. 3). I count 13 peaks over the 70% threshold for the discharge of Hungary in 22 years. But only 54% of these peaks correspond to a flood and 53% of the floods do not exceed the 70% threshold. The Supplementary Fig. 1 of Jongman *et al.*[1] does not represent a sufficient validation of the approximation.

In step 9, the estimated hazard is linked with exposure and vulnerability to estimate the flood risk (step 10). Therein, the discrepancy between observations and the model should not be larger than a random discrepancy with an exceedance probability of 5%, which is the accepted significance level[1,11–13]. The model has to be rejected in the case of smaller exceedance probabilities, which correspond to larger discrepancies. The estimated loss for a 10-year return period is exceeded 6 times in 14 years (Fig. 2a in ref. 1), resulting in an exceedance probability of <1% (binomial distribution), and the Solidarity Fund claims (Fig. 2b in ref. 1) exceed a 20-year return period 4 times, which corresponds to an exceedance probability of 0.04%. These significant discrepancies demand rejection of the model. The discrepancies could have arisen at one or more of the ten steps discussed above. However, some issues could be improved, as explained in the Supplementary Information accompanying this Correspondence. ❐

References

1. Jongman, B. *et al. Nature Clim. Change* **4,** 264–268 (2014).
2. Lindsey, J. K. *Parametric Statistical Inference* (Clarendon Press, 1996).
3. Upton, S. & Cook, I. *A Dictionary of Statistics* 2nd edn (Oxford Univ. Press, 2008).
4. Genest, C. & Favre, A-C. *J. Hydrol. Eng.* **12,** 34–368 (2007).
5. Stephens, M. A. in *Goodness-of-Fit Techniques. Statistics: Textbooks and Monographs* Vol. 68 (eds D'Augustino, R. B. & Stephens, M. A.) 97–193 (Dekker, 1986).
6. Hosking, J. R. M., Wallis, J. R. & Wood, E. F. *Hydrol. Sci. J.* **30,** 85–109 (1985).
7. Genest, C., Rémillard, B. & Beaudoin, D. *Insur. Math. Econ.* **44,** 199–213 (2009).
8. Keef, C., Svensson, C. & Tawn, J. A. *J. Hydrol.* **378,** 240–252 (2009).
9. Keef, C., Tawn, J. A. & Svensson, C. *Appl. Statist.* **58,** 601–618 (2009).
10. Ribatet, M. & Sedki, M. *J. Soc. Française Statist.* **153,** 138–150 (2012).
11. Clark, R. T. *Wat. Resour. Res.* **38,** http://dx.doi.org/10.1029/2001WR000917 (2002).
12. van Gelder, P., Wamg, W. & Vrijling, J. K. *Statistical Estimation Methods for Extreme Hydrological Events. New Concepts for Security* 199–252 (Springer, 2007).
13. Wan Jaafar, W. Z., Liu, J. & Han, D. *Wat. Resour. Res.* **47,** W07503 (2011).

Additional information

Supplementary information is available in the online version of the paper.

Mathias Raschke
Independent Researcher, Stolze-Schrey-Strasse.1, Wiesbaden 65195, Germany.
e-mail: mathiasraschke@t-online.de

Reply to 'Statistics of flood risk'

Jongman *et al.* reply — Raschke[1] has a number of concerns about the modelling approach and assumptions applied in our Letter[2]. Here, we address each of the concerns raised and provide our view on the choice of the specific methodological aspects, and on their validity.

First, Raschke questions whether the Gumbel distribution is appropriate for estimating discharge, compared with the generalized extreme value (GEV) distribution. In a previous study, Dankers and Feyen[3] performed a likelihood ratio test for the GEV distribution against the Gumbel distribution in the same model domain as used in our Letter. They show that the three-parameter GEV distribution does not yield a significant improvement over the two-parameter Gumbel distribution in a large majority of river locations. They also found no evidence that either the GEV or Gumbel model produced consistently higher or lower estimates for the discharge return levels across Europe. Moreover, estimates of the GEV shape parameter become very unstable for return periods beyond the length of the fitting period, which is typically the case for flood events that exceed protection standards and largely determine expected annual flood damages. Based on this reasoning, we elected to use the Gumbel distribution and conclude that our analysis in that regard is statistically sound.

Second, Raschke claims that the Clayton copula that we used to convolute loss distributions in individual basins into a single continental one was applied inappropriately because seasonality was not incorporated, an incorrect tail dependency was assumed and no other copulas were considered. In the Letter, we indeed argue that seasonality affects discharge patterns (Supplementary Fig. 2 in ref. 2), and we agree that an extended dependency model could include seasonality as an additional informative covariate. In our Letter, we established one single dependency structure that represents the entire co-occurrence of peak discharges between basins over the entire time series; future research could further contribute to this methodology by developing separate copula parameters for different seasons.

It is true that some basins may show increasing tail dependency whereas others do not. We therefore chose a family of copulas that, on the one hand, can model independency but, on the other hand, can model possible strong increasing tail dependencies. The Clayton model shows the highest increase in tail dependence compared with the other models mentioned and was therefore selected. The Clayton copula was extensively (but not exclusively) tested for Romania (see Supplementary Information in ref. 2) and showed very good fit there. Also, the Clayton copula has the required flexibility to model different forms of dependencies. In situations with no or decreasing tail dependence, as mentioned by Raschke for the United Kingdom, the resulting Clayton copula would represent full independence. As additional information, we included the no-dependence scenario as well. We therefore provided a wide bandwidth of possible large-scale losses rather than a sensitivity analysis showing the effects of using different copula models and parameters. Note that our model is a special case of the so-called 'vine copulas'[4] but the huge number of components (1,007 sub-basins) did not allow us to estimate the vine structure.

Third, Raschke argues that the dependence model for monthly peaks cannot be applied to annual extremes because the Clayton copula is not a max-stable copula. We used monthly maximum discharges of the entire time series to establish the copula parameters; we used annual maximum values to compute extreme (1 in 2 year to 1 in 500 year) flows. In doing so, our model indeed assumes similar relative basin dependencies for low- and high-return-period flows. In many cases, the correlation between the monthly peaks will probably be similar to those between the annual peaks, and the annual peaks are indeed a subset of the monthly peaks. However, this relationship is likely to be more complex, and we encourage further research on the comparative application of max-stable and non max-stable copulas.

Fourth, Raschke states that peak discharges aggregated at the national level do not always fully explain the reported damaging floods, and that river discharges are therefore not a good proxy for floods. Raschke is correct to state that the relationship between discharges and reported events does not always hold, and we can identify two reasons behind this. First, there are some months in which simulated total peak discharges (at the aggregated national scale) were relatively low, although flood losses were reported. If we examine the damage database used in our study[5], it becomes clear that these specific flood events are very small. For both the May 1991 and July 1994 events in Austria, the total reported losses do not exceed US$100,000. Second, there are months with high simulated river discharge (at the aggregated national scale) but without reported flood losses. This effect is likely to occur because the high discharges do not always happen in populated areas where they cause losses. Following the terminology of the Intergovernmental Panel on Climate Change[6], the peak discharge deviation may have coincided with a flood hazard, but the lack of exposure results in no flood risk. In those cases, our model would therefore simulate floods, but with low or zero economic losses. Modelled discharge peaks versus observed gauge discharge at 554 stations across Europe have been fully validated[7]. We emphasize that the analysis of discharge correlations is conducted at the level of 1,007 individual sub-basins and that the economic risk modelling is performed at the grid-cell level (100 m × 100 m) rather than the national scale.

Raschke's final argument relates to the overestimation of relatively frequent losses, specifically for the 1-in-10-year return period. We acknowledge that our model outcomes do not perfectly represent reported losses, as can be expected. There are a substantial number of uncertain elements in our modelling chain, some of which can be validated while others cannot. These model elements include the grid-cell-based damage modelling, the assessment of discharge correlations, the dependency modelling and the protection standard estimation. In addition to uncertainties surrounding tail dependency in different basins, we acknowledge that uncertainties surrounding the newly developed protection standard database can lead to overestimation of high-frequency losses, as Raschke points out. Whereas validation of the modelled protection levels was performed with the data available, the number of empirical data points is very limited (Supplementary Table 2 in ref. 2). For the same reason, we necessarily assumed homogeneous protection levels within each basin, while this is often not the case in reality. Hence, for a basin with a protection level of 100 years, we assume that no inundation (and therefore no damages) would occur below this frequency anywhere in the basin, whereas some regions (for example, peripheral urban or semi-urban areas) may not have the same level of protection as more densely populated areas.

The only way to reduce this specific uncertainty in future large-scale risk modelling studies would be to develop a detailed geo-referenced dataset of actual flood protection levels and observed losses. We emphasize that the method still represents the most sophisticated approach at the continental scale to date, as most large-scale models simply assume that no protection measures are in place, leading to large overestimations of risk[8].

While uncertainties persist and may propagate, especially in the lower ranges of modelled risk estimates, we reject Raschke's claim that this would falsify the risk model[1]. We do emphasize that we present a first approach to a continental-scale disaster risk assessment that includes basin dependencies, and that the results should therefore not be considered as a final answer. Although a full sensitivity analysis focusing on each individual part of the risk modelling was not possible in this study, the quantification of uncertainties and further validation of model elements on lower spatial levels should be a research priority. ❐

References
1. Raschke, M. *Nature Clim. Change* **4,** 843–844 (2014).
2. Jongman, B. *et al. Nature Clim. Change* **4,** 264–268 (2014).
3. Dankers, R. & Feyen, L. J. *Geophys. Res. Atmos.* **113,** D19 (2008).
4. Kurowicka, D. & Joe, H. (eds) *Dependence Modeling: Vine Copula Handbook* (World Scientific, 2011).
5. Munich Re *Munich Reinsurance Company Geo Risks Research* (NatCatSERVICE Database, 2013).
6. IPCC *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation* (eds Field, C. B. *et al.*) (Cambridge Univ. Press, 2012).
7. Rojas, R., Feyen, L., Bianchi, A. & Dosio, A. J. *Geophys. Res. Atmos.* **117,** D17109 (2012).
8. Ward, P. J. *et al. Environ. Res. Lett.* **8,** 044019 (2013).

Brenden Jongman[1]*, Stefan Hochrainer-Stigler[2], Luc Feyen[3], Jeroen C. J. H. Aerts[1], Reinhard Mechler[2], W. J. Wouter Botzen[1], Laurens M. Bouwer[4], Georg Pflug[2], Rodrigo Rojas[3,5] and Philip J. Ward[1]
[1]Institute for Environmental Studies, VU University Amsterdam, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands, [2]IIASA — International Institute for Applied System Analysis, Schlossplatz 1, Laxenburg 2361, Austria, [3]European Commission — Institute for Environment and Sustainability, Joint Research Centre, via E. Fermi 2749, I-21027 Ispra (VA), Italy, [4]Deltares, Rotterdamseweg 185, 2629 HD Delft, The Netherlands, [5]Present address: CSIRO, Land and Water, Private Bag Nr 5, PO Wembley, Perth, Western Australia 6913, Australia.
*e-mail: brenden.jongman@vu.nl

CORRESPONDENCE:

Spatiotemporal patterns of warming

To the Editor — Ji *et al.*[1] present a methodology to analyse global (excluding Antarctica) spatiotemporal patterns of temperature change, using mean monthly temperatures obtained from the updated Climate Research Unit (CRU) high-resolution gridded climate database[2,3]. Their analysis fails to take into account several key characteristics of the CRU database, seriously compromising the conclusions regarding the spatiotemporal patterns of global warming during the twentieth century.

Climatic data comes from thousands of stations scattered non-randomly across Earth, with much higher densities at mid-latitudes than in the tropics or the Arctic, creating spatial bias. A distance-weighted interpolation from available meteorological stations was implemented to fill spatial gaps in the CRU database[2–4]. Land pixels outside a search radius of 1,200 km from the closest meteorological station were given the corresponding CRU 0.5° 1961–1990 mean monthly climatology[4,5] (Supplementary Fig. 1; other search radii apply to other variables in the CRU database).

In terms of temporal bias, the CRU dataset logically contains many fewer observations in the early part of its record. This is particularly prevalent in remote tropical and Arctic regions, where temperature records abound with long-term climatological averages. Consequently, the temporal autocorrelation of such time series is artificially high, and the climatic variability they portray for the early decades of the record is meaningless (Fig. 1).

Ji *et al.*[1] fail to address these spatial and temporal biases. Supplementary Fig. 2 strongly suggests that the absence of a trend