

CLIMATOLOGY

Dependence of drivers affects risks associated with compound events

Jakob Zscheischler* and Sonia I. Seneviratne

Compound climate extremes are receiving increasing attention because of their disproportionate impacts on humans and ecosystems. However, risks assessments generally focus on univariate statistics. We analyze the co-occurrence of hot and dry summers and show that these are correlated, inducing a much higher frequency of concurrent hot and dry summers than what would be assumed from the independent combination of the univariate statistics. Our results demonstrate how the dependence structure between variables affects the occurrence frequency of multivariate extremes. Assessments based on univariate statistics can thus strongly underestimate risks associated with given extremes, if impacts depend on multiple (dependent) variables. We conclude that a multivariate perspective is necessary to appropriately assess changes in climate extremes and their impacts and to design adaptation strategies.

INTRODUCTION

Compound climate extremes are extreme events for which more than one variable is involved. These events often have disproportionate impacts on humans and ecosystems (1–4). Risk assessments so far usually focus on univariate statistics (5, 6), even when multiple stressors are considered (5, 7, 8). Different definitions for a compound event have been suggested in recent years (1, 5), and generally, extreme impacts are part of the definition. For instance, Leonard *et al.* (1) define a compound event as “an extreme impact that depends on multiple statistically dependent variables or events.” The risk of an extreme impact is thus related to the dependence structure of the driving variables. Stronger dependence between drivers can increase the risk of a compound event (9). Risk can be defined as follows (10): Risk = Hazard × Vulnerability × Exposure. Here, hazard comprises the probability of a climate extreme with a potentially large impact. Throughout the paper, we focus on the hazard part of the equation. However, a change in likelihood of the hazard directly affects risk.

Concurrent extreme droughts and heat waves have been observed to cause a suite of extreme impacts on natural and human systems alike. For example, they can substantially affect vegetation health (11–13), prompting tree mortality (14) and thereby facilitating insect outbreaks (15) and fires (16). In addition, hot droughts have the potential to trigger and intensify fires (17, 18) and can cause severe economic damage (19). By promoting disease spread, extremely hot and dry conditions also strongly affect human health (20–22).

Temperature and precipitation play a vital role for all living systems, and in particular, plants are sensitive to climatic variations during the growing season. It has long been known that during summer, temperature and precipitation are generally anticorrelated at interannual scales (23, 24). A global analysis reveals that precipitation and temperature averaged over the warmest 3 months (denoted by “warm season” in the remainder of the article) are strongly negatively correlated in many land regions of the world (Fig. 1). This negative correlation, prevalent in most models (fig. S1), is to a large extent driven by land surface feedbacks (24–26), associated with impacts of soil moisture limitation on surface temperature (25, 27). These feedbacks tend to be dominant in transitional climate regimes between dry and wet climates

(25, 28) and also include interactions with boundary layer processes (29). In addition, synoptic-scale correspondence between cloud cover/precipitation and incoming shortwave radiation can play a role (26). Regions with a particularly high negative correlation include the southeastern United States, the Amazon region, southern Africa, western Russia, large parts of India, and northern Australia. Over ocean regions, this correlation is often positive, in particular, for regions affected by El Niño, indicating that ocean conditions drive the atmosphere (24). The negative correlation between temperature and precipitation during the warm season over land should lead to an occurrence rate of hot and dry summers that is higher than if both variables were uncorrelated. Here, we focus on land only and aim at quantifying this effect. Furthermore, because of their wide-ranging impacts, detecting and quantifying changes in the co-occurrence of extremes in temperature and precipitation (30, 31) under a warming climate are important for making reliable risk projections. Hao *et al.* (30) quantified changes in concurrent monthly extremes in temperature and precipitation over the observational period and detected substantial increases in the occurrence of joint warm and dry months. Similarly, Mazdiyasi and AghaKouchak (31) demonstrated an increase in concurrent meteorological droughts and heat waves in the United States. However, these studies do not separate the overall warming trend from changes in the dependence between temperature and precipitation. Here, we quantify how the dependence structure between warm season temperature and precipitation changes under a strong greenhouse-gas forcing scenario and how this influences the likelihood of extremely hot and dry warm seasons.

RESULTS

A simple approach to investigate the occurrence rate of extremely hot and dry warm seasons consists of counting the number of years in which both variables exceed a quantile-based threshold (30). As an example, we pick a grid point in western Russia where precipitation and temperature are strongly negatively correlated in summer ($r = -0.63$; 1901–2013). We now count the number of summers in which temperature exceeds the 90th percentile and in which precipitation is at the same time below the 10th percentile (that is, negative precipitation exceeds the 90th percentile). Five summers fulfill this criterion (Fig. 2A). If precipitation and temperature were uncorrelated, this number would lie between 1 and 2 [the expected number is $(1 - 0.9) \times (1 - 0.9) \times 113 = 1.13$].

Copyright © 2017
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
NonCommercial
License 4.0 (CC BY-NC).

Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland.
*Corresponding author. Email: jakob.zscheischler@env.ethz.ch

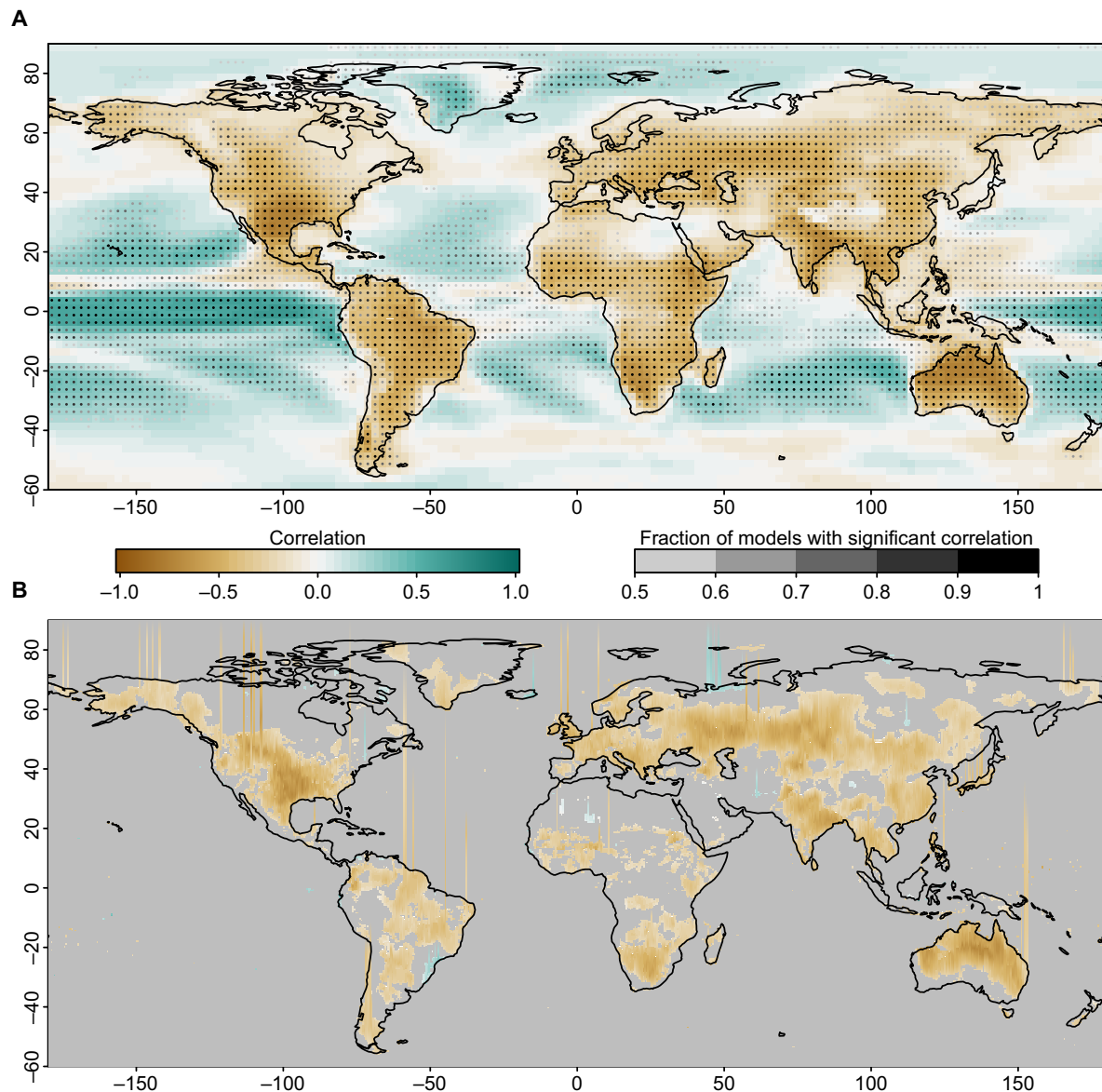


Fig. 1. Correlation between temperature and precipitation during the warm season. The warm season is determined as the hottest 3-month period in the temperature climatology. The correlation is computed as the interannual correlation of the yearly averaged values of temperature and precipitation over the considered 3-month period. **(A)** Model mean of correlations of all CMIP5 models (1870–1969). Stippling is shaded according to the fraction of models that show significant correlations at the 0.05 level if this fraction is larger than 0.5. **(B)** Mean of the correlations of the observation-based data sets CRU (1901–2013), Princeton (1901–2012), and Delaware (1901–2012). Oceans and areas, where less than two of the three data sets show significant correlations at the 0.05 level, are colored in gray.

Although this is a straightforward approach to investigate compound extremes, it has the disadvantage that, for very extreme events, these exceedances contain only very few samples, thus requiring very long time series. For instance, if temperature and precipitation are uncorrelated and we count occurrences for which both variables exceed their 90th percentile, this results in only one event, on average, in the case of a 100-year time series, rendering an investigation of changes in occurrence rates unfeasible. To overcome the shortcomings of counting exceedances, we model here the dependence structure of temperature and precipitation with copulas (32) and subsequently derive exceedance probabilities and return periods (Materials and Methods). Analogous to the simple approach introduced above, we define bivariate

extremes as the concurrent exceedance of some predefined quantile. For two random variables X (for example, temperature) and Y (for example, negative precipitation), we compute

$$p = \Pr(X > x \cap Y > y) \quad (1)$$

for some x and y representing the same quantile of X and Y , respectively. We model Eq. 1 with the help of copulas, which model the dependence between X and Y (see Materials and Methods). This dependence thus affects the probability p of a bivariate extreme. The return period RP in years of this bivariate extreme is then given as $RP = 1/p$ (note that we have one value per year). We quantify the

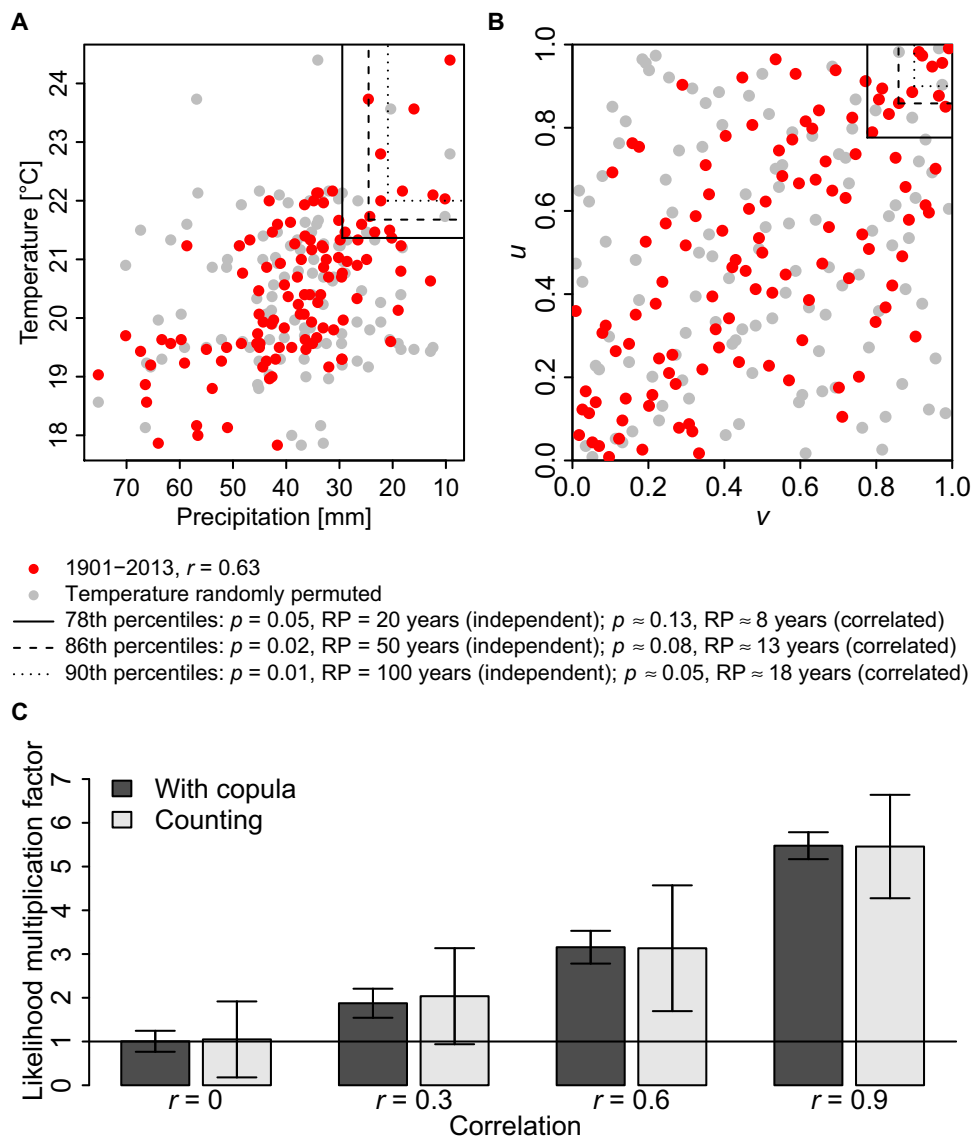


Fig. 2. Dependence affects the likelihood of bivariate extremes. (A) Temperature and negative precipitation averaged across June, July, and August (warm season) at 56.25°E, 51.25°N in CRU for the time period 1901–2013 (red). Data where the values of temperature are randomly permuted are shown in gray. (B) Same data as in (A) transformed into normalized ranks. The regions where both variables concurrently exceeded both 78th (solid), 86th (dashed), and 90th percentiles (dotted), corresponding to 20-year ($p = 0.05$), 50-year ($p = 0.02$), and 100-year ($p = 0.01$) bivariate return periods for independent data under the condition that temperature and negative precipitation exceeded the same quantile ($u = v$), are depicted. Return periods of the original, correlated data based on the same thresholds correspond to approximately 8 years ($p \approx 0.13$), 13 years ($p \approx 0.08$), and 18 years ($p \approx 0.05$). (C) Comparison between estimating changes in the likelihood of bivariate extremes by counting extremes (light bars) and modeling the extremes with copulas (dark bars) for different coefficients of correlation. The increase in likelihood due to the correlation is shown, taking an event in which both variables are independent and exceed their 90th percentile as reference (that is, a 100-year event). Whiskers represent 1 SD over 83 repetitions.

impact of the correlation between temperature and precipitation on the likelihood of extremely hot and dry warm seasons (see Fig. 2 and Materials and Methods). For instance, for the grid point in western Russia, a 20-year event ($p = 0.05$) based on independent (permuted) temperature and precipitation becomes an 8-year event ($p \approx 0.13$) if the correlation between temperature and precipitation is taken into account (Fig. 2B). Similarly, a 50-year event ($p = 0.02$) becomes a 1-year event ($p \approx 0.08$). Dependence (measured here as correlation) between climate variables thus directly influences the likelihood of compound extremes (Fig. 2C). Both approaches, the

counting approach and the approach based on copulas, lead to similar results on artificial data (Fig. 2C), although with higher uncertainty associated with the counting approach. The comparison is shown for bivariate extremes based on 90th percentiles (corresponding to a 100-year event for independent data, that is, $p = 0.01$). In principle, it is also possible that variables are strongly correlated, but their extremes do not co-occur [so-called tail independence (32); see Materials and Methods]. However, below, we show that the correlation between temperature and precipitation is a good indicator for the likelihood of an extremely hot and dry warm season. This may be related to

the fact that the analyzed events are not extreme enough to see the effect of tail (in)dependence.

The return period of an extremely hot and dry warm season that has a 100-year return period if temperature and precipitation were not correlated is reduced to merely 16 years in some regions due to the negative correlation between temperature and precipitation (fig. S2). This corresponds to a sixfold increase in likelihood (Fig. 3, A and B). The spatial distribution of this change in likelihood of extremely hot and dry warm seasons explains why some regions experience these compound events more often than others do. For example, averaged over central North America, the likelihood of a 100-year summer based on independent temperature and precipitation is increased by a factor 4.0 ± 0.9 because of their correlation. The likelihood of a 100-year event is increased by a factor of 3.6 ± 0.6 and 3.4 ± 0.5 in the Amazon and South Africa, respectively; by a factor of 3.2 ± 1.0 and 3.2 ± 0.5 in central Europe and South Asia, respectively; and by a factor of 3.0 ± 0.7 in East Asia (Fig. 3C). These numbers are based on state-of-the-art climate models (see Materials and Methods) and are smaller for observation-based data sets (Fig. 3D), probably due to noise and incomplete coverage by weather stations (see Materials and Methods).

Greenhouse gas-induced climate change is projected to lead to a strong increase in temperature in many regions of the world, accompanied by differing trends in precipitation (33). Under the business-as-usual climate change scenario, the state-of-the-art climate models

collected in the Coupled Model Intercomparison Project Phase 5 (CMIP5) project very large increases in the concurrent exceedance of the historical 90th percentiles of temperature and dryness during the warm season (Fig. 4A). In many regions, the occurrence rate of extremely hot and dry warm seasons increases by a factor of 10 between the historical time periods (1870–1969) and the 21st century. These exceptional changes are largely driven by strong long-term trends in temperature and precipitation. If we subtract these trends (see Materials and Methods), the CMIP5 models project and intensification of the predominant negative interannual correlation between temperature and precipitation in many regions of the world (Fig. 4B). That is, this negative interannual correlation between temperature and precipitation (Fig. 1) intensifies under future climate change, in addition to the change in mean climate. Particularly in the northern extratropics, but also in the Amazon region and in Indonesia, the change in negative correlation can be up to -0.2 in the model mean. Little change or a slight increase in correlation (that is, decrease in negative correlation) is projected in the Mediterranean, Central America, the Sahel, and northern and eastern Australia. As demonstrated below, an increase in negative correlation translates into an increase of the occurrence rate of extremely dry and hot warm seasons (see also Fig. 2C).

Although important for assessments of future risks associated with climate extremes, the investigation of changes in the occurrence of compound events has received only limited attention so far (5, 6).

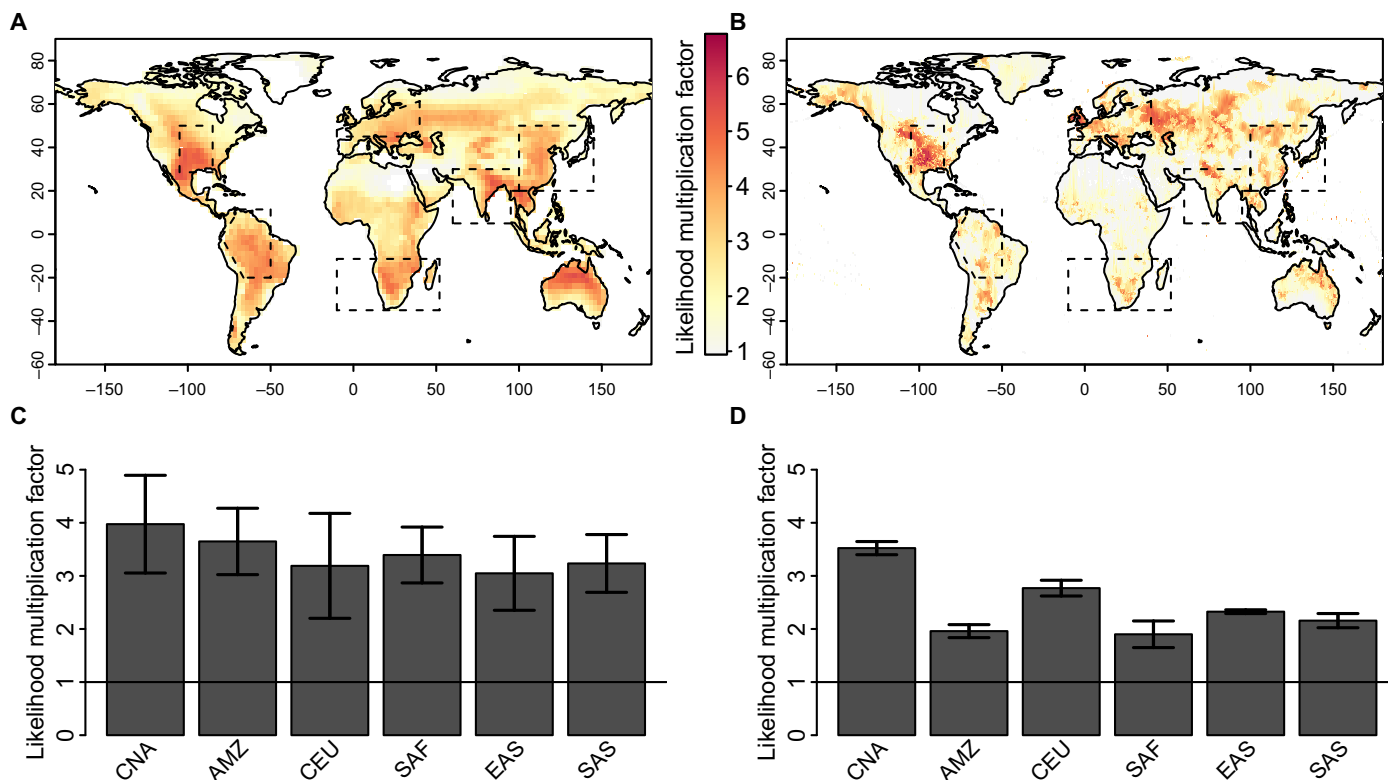


Fig. 3. Increase in likelihood of extremely hot and dry warm seasons due to dependence. Starting from a 100-year event with independent temperature and negative precipitation (that is, both exceed their 90th percentile), the increase in likelihood of these events due to the dependence between temperature and correlation is shown. (A) Average of the increases in likelihood across all CMIP5 models. (B) Average of the increases in likelihood in the data sets CRU, Princeton, and Delaware. (C) Increases in likelihood were averaged over CMIP5 models across the regions central North America (CNA), Amazon (AMZ), central Europe (CEU), South Africa (SAF), East Asia (EAS), and South Asia (SAS). Whiskers represent 1 SD over all models. (D) As in (C) but averaged over observation-based data sets CRU, Princeton, and Delaware. Whiskers represent 1 SD over all three data sets.

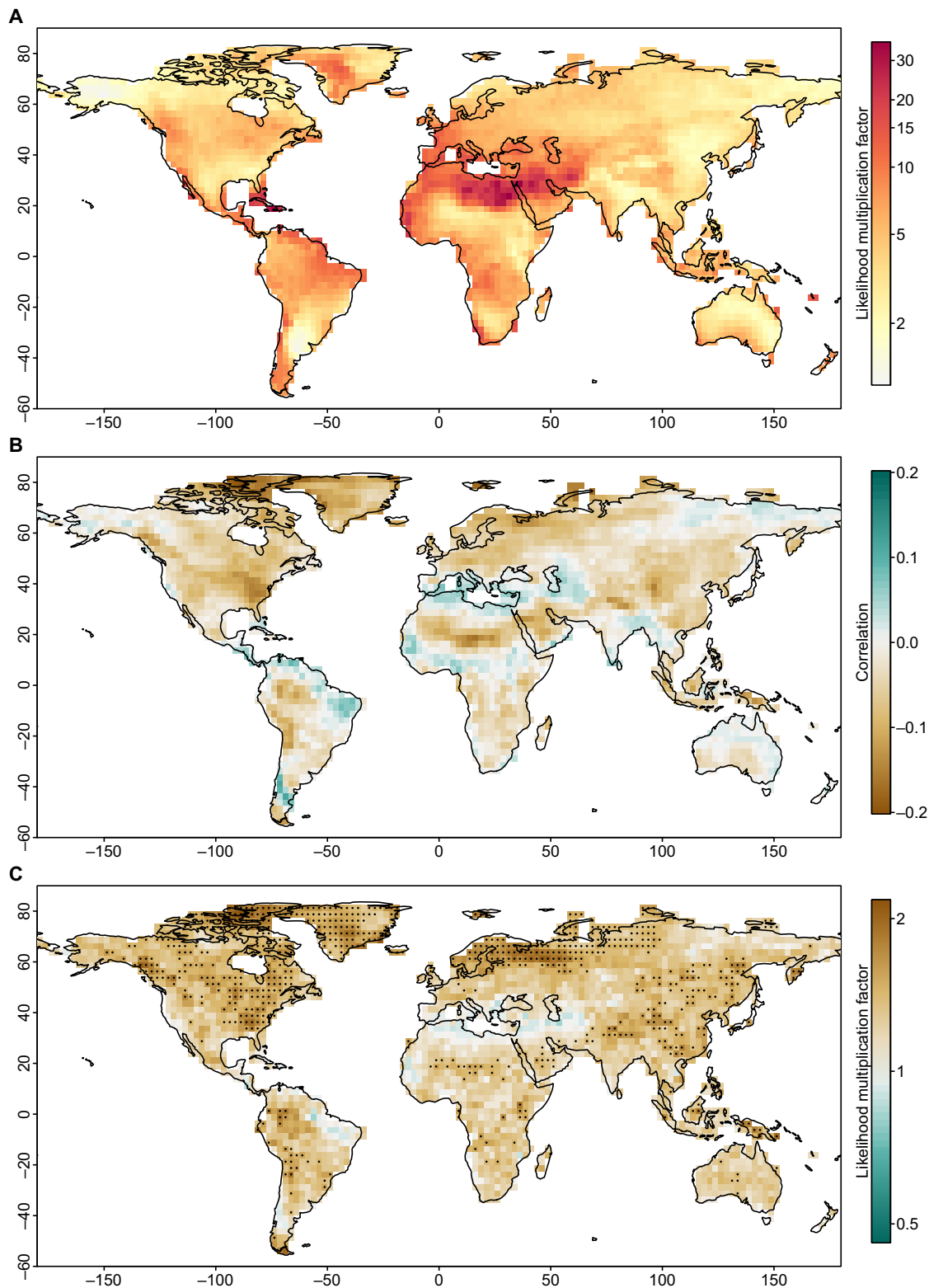


Fig. 4. Future projections. (A) Increase in likelihood of concurrently exceeding the historical 90th percentiles of temperature and negative precipitation averaged over the warm season during the 21st century. The average over all CMIP5 models is shown. (B) Change in interannual correlation between temperature and precipitation averaged over the warm season between 1870–1969 and the 21st century. The average over all CMIP5 models is shown. (C) Change in likelihood that an extremely hot and dry warm season with a return period of 100 years during 1870–1969 will occur during the 21st century. The average across all CMIP5 models is shown. Stippling highlights locations where models show a significant increase in likelihood in the 21st century ($P < 0.1$). For (B) and (C), temperature and precipitation during the warm season were linearly detrended in both time periods before further analysis.

Here, we translate the detected changes in correlation between detrended temperature and precipitation into changes in likelihood of experiencing concurrent extremes. Counting the occurrence of events lying above the 90th percentile for temperature and below the 10th percentile of precipitation slightly increases between the two periods 1871–1969 and 2001–2100 (fig. S3), with spatial patterns approximately resembling the change in correlation (Fig. 4B). However, uncertainties are spatially high.

Using copulas to model the dependence of temperature and precipitation allows an assessment of the change in likelihood of a 100-year event between the historical time period and the 21st century, ignoring the climate change signal on the trend. This likelihood is increased by a factor of up to 2 in the model average in some regions (spatial mean, 1.32), leading to a doubling in occurrence rate (Fig. 4C). Spatial patterns for changes in likelihood of 20- and 50-year events look similar (fig. S4). Corresponding to the change in negative correlation between temperature and precipitation (Fig. 4B), the increase in likelihood of extremely hot and dry warm seasons is largest in the northern extratropics, eastern Asia, and some parts of western South America. Despite large model uncertainties, in about 19% of the land area, this increase in likelihood is significant ($P < 0.1$; see Materials and Methods), including in eastern North America, eastern Asia, northwestern Russia, and some regions in the Amazon.

DISCUSSION

Warmer temperatures naturally lead to an increase in the co-occurrence of hot temperatures and meteorological droughts, if precipitation does not change (30, 31, 34). Our results demonstrate that, in addition to the trend induced by warming, the strengthening of the dependence between temperature and precipitation further exacerbates the increase in co-occurrence of very hot and dry warm seasons in many regions. This increase can approach a doubling of probability in the 21st century for 100-year events of the historical time period. These results suggest that even if systems adapt to mean climate change, they will be hit by extremely hot and dry warm seasons more frequently. Quantifying this effect is highly relevant for future projections of impact assessments because the co-occurrence of extremely hot and dry conditions causes disproportionate impacts. Statistical projections of future impacts for which temperature and precipitation are relevant may thus not be very reliable. For making reliable projections, Earth system models need to be validated to represent the correct covariability between climatic variables. For many regions, this is a challenge because observational data sets are not well constrained (see fig. S5 and Materials and Methods).

The intensification of the interannual correlation between temperature and precipitation during the warm season (Fig. 4B) is possibly caused by an increased land-atmosphere feedback in a warmer climate (25). Warmer temperatures and stronger radiative forcing generate higher evaporation rates, potentially drying out the soil earlier in the season and therefore reducing evaporative cooling during summer (25). In particular, in higher latitudes, an overall greener land surface in combination with a longer growing season may also lead to higher transpiration rates in spring (13), decreasing soil moisture and thus potentially further increasing temperatures by increasing sensible heat (35). The importance of the land surface for this change in correlation is also underlined by the fact that, over oceans, the negative correlation between temperature and precipitation is absent or reversed (Fig. 1A). In addition to land surface feedbacks, changes in dynam-

ical forcing may explain some of the observed trends in interannual correlation. In particular, changes in atmospheric circulation patterns related to changes in planetary waves may be partly responsible for the detected change in correlation in some regions (36). For instance, certain planetary waves have been linked to extreme conditions in precipitation and temperature in midlatitudes (37–39) and may be amplified in response to anthropogenic climate change (38, 39). However, model projections of circulation patterns are highly uncertain (40). Unraveling the drivers of the change in correlation is important for assessing the robustness of the model results and estimating future risks related to compound events.

Our analysis suggests that univariate assessments of extremes may fall short in communicating risks related to impacts of climate extremes, because often several variables are responsible for causing extreme impacts (1, 9). In addition, as we have shown, the multivariate structure may change over time. Thus, including the multivariate structure of relevant driver variables is crucial to realistically assess potential impacts related to compound events.

MATERIALS AND METHODS

Data

We used temperature (T) and precipitation (P) data from observation-based data sets and models from the CMIP5 archive (41). For observation-based data sets, we included CRU (V3.22, 1901–2013) (42), Princeton (1901–2012) (43), and Delaware (V3.01, 1900–2010) (44). From CMIP5, we used runs from 40 models covering the historic time period (1870–2005) and climate projections with the strongest greenhouse-gas forcing for the future (RCP8.5, 2006–2100). The names of the models, including the number of individual runs performed, are listed in table S1. In total, there are 83 runs available for the RCP8.5 scenario.

All our analyses are based on seasonal averages over the warm season (that is, one value per year). We defined the warm season as the hottest 3 months of the climatology of T . For the CMIP5 models, the warm season was defined based on 1870–1969. A change in time periods has little impact on the definition of warmest season or the interannual correlation between T and P (see below).

For CMIP5 data, before all calculations, all data were bilinearly interpolated to an equal 2.5° spatial grid. To compute changes in bivariate 100-year return periods of extremely hot and dry warm seasons, we chose the two time periods 1870–1969 and 2001–2100. These time periods were chosen as a trade-off between maximizing the climate change signal and, at the same time, maximizing the number of samples for computing bivariate extremes. Choosing 3 months as the length of an event is a compromise between potentially longer-lasting droughts and heat waves, which occur on shorter time scales. The combination of summer means in temperature and precipitation is a good indicator for the extremeness of the summer and its impacts (45). The computation of bivariate return periods and changes therein are described below.

Statistical analysis

Interannual correlations and model-data comparison

We computed interannual correlations between T and P averaged over the warm season at each location and model/data set. We investigated correlations in CMIP5 for the whole globe (land and ocean), but we used only land data for observation-based data sets. To eliminate the climate change signal on long-term trends, we linearly detrended both variables before computing correlations. Detrending has little

impact on the correlation for the period 1870–1969 (fig. S6A) but strongly affects the correlation during the 21st century (fig. S6B). This is because the strong warming trend in T in the 21st century overrides correlations at the interannual scale.

In comparison to observation-based data sets, CMIP5 models capture the negative correlation between T and P quite well (fig. S5). Observation-based data sets lie within the 10th-to-90th percentile range of CMIP5 models in 75% of the land area. In the remaining areas, the negative correlations between T and P were generally much stronger in CMIP5, for instance, in the Amazon, Mexico, large parts of Africa and western Australia, and northern Canada and Greenland. Although this might suggest that CMIP5 models generally overestimate the strength of correlation between T and P during the warm season in these areas, this may partly also be related to noise in the observation-based data sets and the spatial coverage of actual climate stations (46). Correlations in CMIP5 were stronger especially in the tropics and subtropics, where the station cover for deriving observation-based gridded climate data is sparse (42), thus leaving the covariation between T and P less well constrained.

To incorporate potential impacts of serial correlation, we tested whether seasonally averaged T and P are significantly serially correlated in time. Figure S7 shows that for most regions on land and most models, T and P are not significantly correlated at lag 1 ($P < 0.05$). For some tropical areas, including the Amazon, Congo, and Indonesia, most models show significant serial correlation at lag 1 for T . However, this is not the case for P . Hence, we concluded that serial correlation should not have a significant impact on the assessment of bivariate extremes. All the remaining analyses were based on land data only.

Compound extremes

We used two different approaches to investigate compound extremes of extremely hot and dry warm seasons: (i) We counted concurrent exceedances of T over the 90th percentile (“hot”) and P below the 10th percentile (“dry”), and (ii) we modeled the dependence of T and $-P$ with copulas and computed the bivariate return period from the fitted copula. Modeling dependence with a copula allowed us to investigate the effect of the dependence of T and P on bivariate return periods (see below).

Bivariate return periods with copulas

We analyzed bivariate return period using the concept of copulas, which are often used to describe the dependence between random variables (32). Here, we computed bivariate return periods of hot and dry warm seasons; accordingly, our two variables are T and $-P$, averaged over the warm season (47). For two random variables X (for example, T) and Y (for example, $-P$) with cumulative distribution functions $F_X(x) = Pr(X \leq x)$ and $F_Y(y) = Pr(Y \leq y)$, the joint distribution function of X and Y can be written as

$$F(x, y) = Pr(X \leq x, Y \leq y) = C(F_X(x), F_Y(y)) \quad (2)$$

with a copula C (48). C is a joint distribution function of the transformed random variables $U = F_X(X)$ and $V = F_Y(Y)$. Because of this transformation, the marginals U and V have uniform distribution. The probability of an event, where both variables exceed a given threshold, is given by (49, 50)

$$p = Pr(U > u \cap V > v) = 1 - u - v + C(u, v) \quad (3)$$

Bivariate extremes can be defined in other ways (49). However, this

definition, using the AND operator (51), is consistent with the approach of counting concurrent exceedances (see above). We defined bivariate extremes as the area where both variables exceed the same quantile-based threshold; hence, we always set $u = v$. The return period in years associated with the exceedance probability p is given by

$$RP = 1/p \quad (4)$$

Note that our analysis is based on one value per year.

Archimedean copulas used in this study

We used four Archimedean copulas in this study: Frank, Clayton, Gumbel, and Joe. Archimedean copulas can be written as (32)

$$C : [0, 1]^2 \rightarrow [0, 1], C(u, v) := \varphi^{[-1]}(\varphi(u) + \varphi(v)) \quad (5)$$

with the generator φ and

$$\varphi^{[-1]} = \begin{cases} \varphi^{-1}(t), & 0 \leq t \leq \varphi \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

the pseudo inverse of φ . The generator functions for the four copulas are given by

$$\text{Frank copula: } \varphi(t) = -\ln\left(\frac{e^{-\vartheta t} - 1}{e^{-t} - 1}\right), \vartheta > 0 \quad (7)$$

$$\text{Clayton copula: } \varphi(t) = \frac{1}{\vartheta} (t^{-\vartheta} - 1), \vartheta > 0 \quad (8)$$

$$\text{Gumbel copula: } \varphi(t) = (-\ln t)^\vartheta, \vartheta > 1 \quad (9)$$

$$\text{Joe copula: } \varphi(t) = -\ln\left(1 - (1 - t)^\vartheta\right), \vartheta > 1 \quad (10)$$

To illustrate how these four copulas look like, we plotted 1000 random samples from the Frank, Clayton, Gumbel, and Joe copulas in fig. S8. Some of these copulas are able to model tail dependence, that is, the property that extremes are correlated. As illustrated in fig. S8, the Frank copula has no tail dependence, the Clayton copula has lower tail dependence, and the Gumbel and Joe copulas have upper tail dependence.

We further made use of the independent copula, which is given by

$$C(u, v) = uv \quad (11)$$

Model fitting

To obtain uniform distributions in the marginals, we transformed marginal distributions into normalized ranks, which is a common procedure when working with copulas (52, 53) and is the only reasonable choice if goodness of fit is to be tested appropriately (54). We fit four different Archimedean copulas (Clayton, Frank, Gumbel, and Joe; see above and fig. S8) and selected the one with the best fit relying on the Bayesian Information Criterion implemented in the R package VineCopula (55). Goodness of fit was tested based on the Cramér-von Mises statistic (54) implemented in the R package copula (56). P values below 0.05 (rejecting the

hypothesis that the copula is a good fit) were obtained for about 3% of the fits, which is an acceptable rate.

Although our analysis focused on bivariate extremes, because we used a copula-based approach, goodness of fit was tested on the whole distribution. Hence, goodness-of-fit statistics may be largely driven by the less extreme values. However, using copulas allowed us to switch between different return periods without the need to fit a new model for each of the different return periods (for example, see Figs. 2 and 4C and fig. S4). In addition, all results were averaged across >80 model runs (table S1), which reduced some of the uncertainty related to the fitting procedure (57). Figure 2C provides some confidence that our approach based on copulas captures the desired change in likelihood due to a change in dependence well.

Dependence affects risk of multivariate extremes

We analyzed the impact of varying correlation on bivariate return periods using simulated data. We simulated four times 100 samples from a Frank copula with $\theta = 0, 1.9, 4.4,$ and 12 with uniform marginals. This resulted in correlation coefficients of approximately $0, 0.3, 0.6,$ and $0.9,$ respectively (note that $\theta = 0$ corresponds to the independent copula). We computed changes in return periods due to dependence as follows: Let C_0 be the independent copula (Eq. 11), representing no correlation, and C_1 be one of the Frank copulas. For a given return period RP_0 and the corresponding probability of exceedance $p_0 = 1/RP_0,$ we can use Eqs. 3, 4, and 11 and compute

$$u = v = 1 - \sqrt{p_0} \quad (12)$$

We now look for the return period of $C_1,$ whose exceedance thresholds are defined by the same u and $v.$ The new exceedance probability p_1 is given by (Eq. 2; setting $u = v$)

$$p_1 = 1 - 2(1 - \sqrt{p_0}) + C(1 - \sqrt{p_0}, 1 - \sqrt{p_0}) \quad (13)$$

and hence, $RP_1 = 1/p_1.$ For an example, see Fig. 2. We further defined the likelihood multiplication factor as $p_1/p_0.$ Figure 2C shows the influence of dependence (measured as correlation) on the likelihood multiplication factor of a 100-year event ($p_0 = 0.01$) using the copula approach and simple counting. Uncertainty estimates are based on 83 repetitions, which is equivalent to the number of model runs used in this study. The counting approach is associated with much larger uncertainties as compared to the copula-based approach. This is related to the relatively small sample size (100) compared to the extremeness of the considered event (100-year event) and the fact that, in the counting approach, only discrete numbers are possible. The uncertainty for the copula approach at $r = 0$ reflects the uncertainty related to the fitting of the copula.

For modeled and observation-based T and P averaged over the warm season, we computed the likelihood multiplication factor starting with $p_0 = 0.01,$ corresponding to a 100-year return period for independent variables (Fig. 3). Figure S2 shows the effective return period if the thresholds to compute return periods had been defined under the assumption that T and P were independent (that is, showing RP_1).

Correlation is only one indicator for dependence between two variables and does not capture the dependence in the extremes well because it is based on the full range of the data. However, as our analysis shows, for seasonal T and $P,$ the correlation coefficient is generally a good indicator for the influence of dependence on the likelihood of bivariate extremes.

Changes in bivariate return periods

The likelihood of a bivariate extreme may change if the dependence structure of the two underlying variables changes. A simple approach of assessing changes in bivariate extremes is to count concurrent exceedances of 90th percentiles in detrended T and $-P$ in 1870–1969 and 2001–2100. Figure S3 shows the differences in these counts. However, in this approach, the historical exceedance probabilities vary according to their dependence, and we cannot assess changes in, for example, 100-year events.

We investigated changes in the likelihood of bivariate extremes in CMIP5 by comparing their occurrence probabilities, as estimated by Eq. 3 between the two time periods. Let C_1 and C_2 be two copulas capturing the dependence between detrended T and $-P$ for the two time periods 1870–1969 and 2001–2100, respectively. For a given probability of occurrence $p_1,$ we obtained the thresholds $u = v = u^*$ by solving Eq. 3 for $u.$ We then computed the new probability of occurrence p_2 as (Eq. 3)

$$p_2 = 1 - 2u^* + C_2(u^*, u^*) \quad (14)$$

As above, we computed the likelihood multiplication factor as $p_2/p_1.$ We report changes in likelihood of a historic 100-year event (Fig. 4C) and 20- and 50-year events (fig. S4).

We tested whether the ensemble median of the models shows an increase in likelihood with the one-sided sign test (58). We corrected for multiple testing by controlling the false discovery rate, as suggested by Wilks (59), and highlighted regions for which the model ensemble median of the likelihood multiplication factor is significantly larger than 1 with stippling (adjusted $P < 0.1$).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/3/6/e1700263/DC1>

table S1. Global climate models used in this study, with number of runs in brackets.

fig. S1. Fraction of CMIP5 models with negative correlation between temperature and precipitation.

fig. S2. Effective return period of extremely hot and dry warm seasons.

fig. S3. Difference in the number of concurrently extreme hot and dry warm seasons between 1870–1969 and 2001–2100.

fig. S4. Change in likelihood that hot and dry warm seasons with return periods of 20 and 50 years during 1870–1969 will occur during the 21st century.

fig. S5. Comparison of correlations of temperature and precipitation between CMIP5 and observation-based data sets.

fig. S6. Difference in correlation of temperature and precipitation between original and detrended data.

fig. S7. Serial correlation in seasonal temperature and precipitation averaged over the warm season.

fig. S8. Random samples of the four Archimedean copulas used in this study.

REFERENCES AND NOTES

1. M. Leonard, S. Westra, A. Phatak, M. Lambert, B. van den Hurk, K. McInnes, J. Risbey, S. Schuster, D. Jakob, M. Stafford-Smith, A compound event framework for understanding extreme impacts. *WIREs Clim. Change* **5**, 113–128 (2014).
2. J. Zscheischler, A. M. Michalak, C. Schwalm, M. D. Mahecha, D. N. Huntzinger, M. Reichstein, G. Berthier, P. Ciais, R. B. Cook, B. El-Masri, M. Huang, A. Ito, A. Jain, A. King, H. Lei, C. Lu, J. Mao, S. Peng, B. Poulter, D. Ricciuto, X. Shi, B. Tao, H. Tian, N. Viovy, W. Wang, Y. Wei, J. Yang, N. Zeng, Impact of large-scale climate extremes on biospheric carbon fluxes: An intercomparison based on MsTMIP data. *Global Biogeochem. Cycles* **28**, 585–600 (2014).
3. O. Martius, S. Pfah, C. Chevalier, A global quantification of compound precipitation and wind extremes. *Geophys. Res. Lett.* **43**, 7709–7717 (2016).
4. M. Poumadère, C. Mays, S. Le Mer, R. Blong, The 2003 heat wave in France: Dangerous climate change here and now. *Risk Anal.* **25**, 1483–1494 (2005).

51. G. Salvadori, C. De Michele, Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. *Water Resour. Res.* **40**, W12511 (2004).
52. G. Salvadori, F. Durante, C. De Michele, On the return period and design in a multivariate framework. *Hydrol. Earth Syst. Sci.* **15**, 3293–3305 (2011).
53. F. Serinaldi, Can we tell more than we can know? The limits of bivariate drought analyses in the United States. *Stoch. Environ. Res. Risk Assess.* **30**, 1691–1704 (2016).
54. C. Genest, B. Rémillard, D. Beaudoin, Goodness-of-fit tests for copulas: A review and a power study. *Insur. Math. Econ.* **44**, 199–213 (2009).
55. U. Schepsmeier, J. Stoeber, E. Brechmann, B. Graeler, T. Nagler, T. Erhardt, VineCopula: Statistical inference of vine copulas, R package version (2012).
56. M. Hofert, I. Kojadinovic, M. Maechler, J. Yan, Copula: Multivariate dependence with copulas, R package version 0.999-9 (2014); <https://cran.r-project.org/package=copula>.
57. F. Serinaldi, An uncertain journey around the tails of multivariate hydrological distributions. *Water Resour. Res.* **49**, 6527–6547 (2013).
58. P. Sprent, *Applied Nonparametric Statistical Methods* (Chapman & Hall, 1989).
59. D. S. Wilks, “The stippling shows statistically significant gridpoints”: How research results are routinely overstated and over-interpreted, and what to do about it. *Bull. Am. Meteorol. Soc.* **97**, 2263–2273 (2016).

Acknowledgments: We thank the World Climate Research Programme’s Working Group on Coupled Modeling, which is responsible for CMIP, and we thank the climate modeling

groups (listed in table S1 of this paper) for producing and making available their model output. For CMIP, the U.S. Department of Energy’s Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. **Funding:** We acknowledge partial funding from the European Research Council DROUGHT-HEAT project funded by the European Community’s Seventh Framework Programme (grant agreement FP7-IDEAS-ERC-617518). **Author contributions:** Both authors conceived the study. J.Z. performed the analysis and prepared all figures. Both authors wrote the paper. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper are available from the respective websites hosting the data sets.

Submitted 25 January 2017

Accepted 15 May 2017

Published 28 June 2017

10.1126/sciadv.1700263

Citation: J. Zscheischler, S. I. Seneviratne, Dependence of drivers affects risks associated with compound events. *Sci. Adv.* **3**, e1700263 (2017).