**TIME-DEPENDENT PEAKS OVER THRESHOLD METHODOLOGY.** To produce Fig. 4, we used data for 1948–2010 from the Global Historical Climate Network–daily dataset for stations in the contiguous United States, including only stations providing data for at least two-thirds of the days in that period. At each station, we found the station-specific 97th percentile of daily precipitation based on the entire period, using only days with at least 1 mm of precipitation. We then fit a station-specific time-varying statistical extreme value model (Coles 2001) to daily exceedances of the 97th percentile. We used only the maximum daily value when consecutive days exceeded the threshold to avoid temporal dependence from multiday storms (i.e., runs declustering with parameter \( r = 1 \); Coles 2001). We used a point process model for exceedances over a high threshold (or peaks over threshold), as in Tomassini and Jacob (2009) and Cooley and Sain (2010). The model is equivalent to a generalized Pareto distribution for excesses over a threshold combined with a Poisson process for the occurrence of threshold exceedances and is consistent with a generalized extreme value (GEV) distribution for block maxima. The basic parameters of the point process model can be expressed in terms of those of a GEV, namely, location, scale, and shape. The shape parameter determines the heaviness of the tail of the distribution, encompassing the Weibull (bounded tail), Fréchet (heavy tail), and Gumbel (light tail) distributions. We allowed the location parameter to vary linearly in time, while assuming the shape and scale parameters were constant over time. To minimize complexity, any seasonality in these parameters was ignored. As a result of this parameterization, the change over time in the return level (for any return
period) is linear with the same slope as that for the location parameter (Coles 2001). An additional consequence is that the change is not a function of the return period considered—that is, the 1948–2010 change in the 20-yr return level is the same as the 1948–2010 change in the X-yr return level for any X. Note that by fitting a separate shape parameter value at each location, we allowed for the possibility that the heaviness of the tail differs by location. Uncertainty estimates were based on the Hessian of the point process likelihood according to standard maximum likelihood theory, with the standard error for the return level depending not only on the standard error for the linear trend parameter, but also on the standard errors of the other parameters of the GEV as well. Standard diagnostics for extreme value distributions (Coles 2001) indicated no obvious lack of fit, and analysis with thresholds based on percentiles other than the 97th percentile (90th, 95th, 98th, 99th, 99.5th) indicated the results did not change substantially apart from the expected bias–variance trade-off as the percentile increased. The station-specific results are noisy because of the uncertainty in estimating the behavior of extremes from short time series. Statistical approaches that smooth over the noise are feasible, but standard techniques have not been developed, so we show the station-specific results without smoothing. The results are not sensitive to the available data criterion. We repeated the analysis for stations with 90% and 95% available data. We found that the stations excluded by these criterion levels exhibited the same spatial patterns as the stations with more complete data.

To account for multiple testing, we carried out a field significance analysis. Each of the 1,000 simulations consisted of 63 years of synthetic data resampled with replacement from the 63 years of observations comprising 1948–2010. Each resampled year included all the data from all locations for that year, thereby preserving the spatial dependence and within-year temporal structure but breaking the between-year dependence. This produced simulated datasets under the null hypothesis of no temporal trend across years. For each of the 1,000 simulated datasets, we carried out the point process model analysis, calculating the field significance $P$ value based on the number of locations with a $z$ score (change in return level divided by its standard error) exceeding 1, 1.64, and 1.96. In all three cases, none of the simulations had as high a proportion of stations with $z$ scores exceeding the value as the proportion of stations in the original analysis, giving $P < 0.001$.

**EXTREME PRECIPITATION WATER VAPOR ANALYSIS.** A set of extreme precipitation events (daily, 1-in-5-yr recurrence) used in Kunkel et al. (2012) was the basis for this analysis. For each station event, radiosonde data from the Integrated Global Radiosonde Archive were used to find the highest precipitable water value occurring within 3° latitude and longitude and on the day before or the day of the event. This was assumed to be the best representation of the water vapor environment available to the precipitation-producing system.

For each of seven regions similar to those used in the 2009 National Climate Assessment Report (Karl et al. 2009), we averaged these precipitable water values for two periods: 1971–89 and 1990–2009. We also averaged the values of the extreme precipitation index. The statistical significance of the differences was tested using the two-sample $t$ test. These two periods were compared because they span a period of sizeable changes in extreme precipitation occurrences and the data from the Integrated Global Radiosonde Archive (Durre et al. 2006) are most complete after 1970.

**REFERENCES**


