Clim. Past Discuss., 7, 2893–2935, 2011
www.clim-past-discuss.net/7/2893/2011/
doi:10.5194/cpd-7-2893-2011

Climate of the Past Discussions

Inferences on weather extremes and weather-related disasters: a review of statistical methods

H. Visser[1] **and A. C. Petersen**[1,2,3]

[1]PBL Netherlands Environmental Assessment Agency, The Hague, The Netherlands
[2]Centre for the Analysis of Time Series, London School of Economics and Political Science (LSE), London, UK
[3]Institute for Environmental Studies (IVM), VU University, Amsterdam, The Netherlands

Received: 23 August 2011 – Accepted: 12 September 2011 – Published: 16 September 2011

Correspondence to: H. Visser (hans.visser@pbl.nl)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

The study of weather extremes and their impacts, such as weather-related disasters, plays an important role in climate-change research. Due to the great societal consequences of extremes – historically, now and in the future – the peer-reviewed literature on this theme has been growing enormously since the 1980s. Data sources have a wide origin, from century-long climate reconstructions from tree rings to short databases with disaster statistics and human impacts (30 to 60 yr).

In scanning the peer-reviewed literature on weather extremes and impacts thereof we noticed that many different methods are used to make inferences. However, discussions on methods are rare. Such discussions are important since a particular methodological choice might substantially influence the inferences made. A calculation of a return period of once in 500 yr, based on a normal distribution will deviate from that based on a Gumbel distribution. And the particular choice between a linear or a flexible trend model might influence inferences as well.

In this article we give a concise overview of statistical methods applied in the field of weather extremes and weather-related disasters. Methods have been evaluated as to stationarity assumptions, the choice for specific probability density functions (PDFs) and the availability of uncertainty information. As for stationarity we found that good testing is essential. Inferences on extremes may be wrong if data are assumed stationary while they are not. The same holds for the block-stationarity assumption. As for PDF choices we found that often more than one PDF shape fits to the same data. From a simulation study we conclude that both the generalized extreme value (GEV) distribution and the log-normal PDF fit very well to a variety of indicators. The application of the normal and Gumbel distributions is more limited. As for uncertainty it is advised to test conclusions on extremes for assumptions underlying the modeling approach. Finally, we conclude that the coupling of individual extremes or disasters to climate change should be avoided.

CPD
7, 2893–2935, 2011

Weather extremes and weather-related disasters

H. Visser and A. C. Petersen

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

1 Introduction

The study of weather extremes and impacts thereof plays an important role in climate-change research. Due to the great societal consequences of extremes – historically, now and in the future – the peer-reviewed literature on this theme has been growing enormously since the important findings of Mearns et al. (1984) and Wigley (1985). These authors showed that small shifts in the mean of a weather or climate variable might lead to a strong non-linear shift in the frequency of extreme values of that variable. Examples of recent publications on extremes are Trenberth and Jones (2007, Sect. 3.8), Gamble et al. (2008) and Karl et al. (2008). Furthermore, the literature shows that inferences on extremes can be based on all types of meteorological/climatological information: documentary evidence and paleo-climatological proxies (Battipaglia et al., 2010; Stoffel et al., 2010; Büntgen et al., 2011), instrumental data (Alexander et al., 2006; Klein Tank et al., 2006), disaster statistics (Pielke, 2006; Bouwer, 2011) and model-generated climate data (Kharin and Zwiers, 2005; Tebaldi et al., 2006; Orlowsky and Seneviratne, 2011).

In scanning the peer-reviewed literature on weather extremes and impacts we noticed that many different methods are used to make inferences on extremes. However, discussions on methods are rare. One exception we found, is that of Wigley (2009) and Cooley (2009), where the use of linear trends and normal distributions (Wigley) is opposed to the use of extreme value theory with time-varying parameters (Cooley). Clearly, the calculation of a return period of, say, once in 500 yr, based on a normal distribution will deviate from that based on a generalized extreme value (GEV) distribution. In other words, the specific choice of methods (here the shape of PDFs) might influence the inferences made on these extremes. Another example is the particular choice of a trend model to highlight temporal patterns in extreme-weather indicators. Conclusions based on an OLS straight line might differ from those made by more flexible trends. And the inclusion or exclusion of uncertainty information may influence inferences made either: a rising trend or increasing return periods are not necessarily statistically significant.

CPD
7, 2893–2935, 2011

In this article we will review the statistical methods used in the peer-reviewed literature. First, we will give a concise overview of methods applied. These methods deal with the computation of return periods of extremes, chances of crossing a pre-defined high (or low) threshold, or with the estimation of a trend in weather indicators which are defined as extreme by their definition (number of warm and cold days, annual maximum of 1-day/5-day precipitation, global number of floods, etc.).

Next to this overview we will discuss a number of methodological aspects. We will discuss (i) the assumption of a stationary climate when making inferences on extremes, (ii) the choice of (extreme value) probability distributions for the data at hand, (iii) the availability of uncertainty information and (iv) the coupling of weather or disaster statistics to climate change. As for point (iv) we will pay attention to methods in the peer-reviewed literature and to the way these results are assessed by the Intergovernmental Panel on Climate Change (IPCC).

There are two aspects of weather extremes and their impacts (disasters) which will not be dealt with in this methods review. The first aspect concerns the quality of the data, and more specifically, methods for testing the quality of data and correcting them, if necessary. For homogeneity issues the reader is referred to Aguilar et al. (2003) and Klein Tank et al. (2009). For the reliability of disaster statistics please refer to Gall et al. (2009).

The second methodological aspect not dealt with, is that of methods for detecting anthropogenic influences in climate or disaster data. For detection studies in relation to extremes please refer to Hegerl and Zwiers (2007), Zwiers et al. (2011) and Min et al. (2011). A critical view has been given by Stephens et al. (2010). For a review on detecting climate change influences in disaster trends, the reader is referred to Höppe and Pielke (2006) and Bouwer (2011). We further note that we will use the term “climate change” in the general sense, thus climate change both due to natural and anthropogenic influences (unless denoted otherwise).

CPD
7, 2893–2935, 2011

The contents of this article is as follows. In Sect. 2 we will give a concise description of how inferences on extremes are made in the peer-reviewed literature. Then, we will discuss these methods in Sects. 3 through 6 with respect to four aspects: the assumption of a stationary climate (Sect. 3), assumptions on probability distributions (Sect. 4), the use of uncertainty information (Sect. 5) and the coupling of extremes to climate change (Sect. 6). Conclusions are given in Sect. 7. A number of statements throughout this article will be illustrated by an analysis of annual maxima of daily maximum temperatures for station De Bilt in the Netherlands (TXX_t; Figs. 1, 4 and 6).

2 Methods for making inferences on extremes

2.1 Preliminaries

There is a diverse use of terminology of weather extremes. Terms used in the literature comprise weather or climate extremes, weather or climate extreme events, weather or climate indicators, weather or climate extreme indicators, indices of extremes, and weather-related disasters. Various definitions can be found in Alexander et al. (2006) or the ECA website http://eca.knmi.nl/indicesextremes/indicesdictionary.php. In the context of this article we need only one simple distinction. We will discern two types of indicators: weather indicators and extreme weather indicators. We will not make a conceptual distinction between "weather indicators" or "climate indicators". Weather indicators, as used here, comprise variables such as regional or global mean temperatures, annual total precipitation, annual averaged wind speeds, etc. Extreme weather indicators comprise a wide range of indicators which use the predicate "extreme" since they have or point to adverse, or sometimes benign, societal consequences. Extreme weather indicators can be constructed from weather indicators. Examples are the annual maximum value of daily maximum temperatures (TXX_t), the number of warm or cold days, the annual maximum value of one- or five-day precipitation totals ($RX1D_t$, $RX5D_t$), the annual number of flood disasters or the annual global economic damage

CPD
7, 2893–2935, 2011

H. Visser and A. C. Petersen

due to weather-related disasters. All indicators from disaster research will fall here in the category of "extreme weather indicators".

In case of weather indicators, inferences on extremes are made in terms of (i) lists of extreme values over some predefined sample period, (ii) the chance on exceeding a pre-defined high (or low) threshold, (iii) the corresponding return period, (iv) the indicator value that is crossed once in x yr (the "x-year return period") and (v) the comparison of probability density functions (PDFs), calculated over distinct sub-periods of interest. As for extreme weather indicators, inferences can be made as mentioned above, or by estimating a trend through the data.

An example illustrates a number of these inference possibilities. Suppose we are interested in the following extreme weather indicator: annual extreme temperatures $\mathrm{TXX_t}$, with t in years. For the Netherlands we constructed a time series of this indicator over the period 1901–2010 for station De Bilt. Homogeneity tests showed a large discontinuity in 1950, the year where the type of temperature screen changed. Therefore, we decided to limit analyses to the period 1951–2010. Other homogeneity tests were satisfactory (Visser, 2007). The $\mathrm{TXX_t}$ series is shown in Fig. 1.

The upper panel shows the data along with the Integrated Random Walk (IRW) trend model and 95 % confidence limits (Visser, 2004; Visser and Petersen, 2009). Tests showed the residuals (or in Kalman filter terms: innovations or on-step-ahead prediction errors) to be normally distributed. These normal distributions are shown in the lower panel for the years 1951, 1980 and 2010. The panel shows how chances (p_{t}^{35}) of crossing a certain threshold, here 35 °C, is changing for these three distributions (the yellow area). For this example we find $p_{1951}^{35} = 0.002$, $p_{1980}^{35} = 0.02$ and $p_{2010}^{35} = 0.18$. Average return periods are simply gained by taking the inverse of p_{t}^{35}, yielding return periods R_{t}^{35} of once in 420, 62 and 5.6 yr, respectively. For the calculation of annual 20-yr return periods ($\mathrm{TXX_t^{20}}$) we choose the yellow area such that it covers 5 % of right-hand tail of the normal distributions. For this example we find the temperature thresholds 32.8, 34.1 and 36.4 °C, respectively (cf. Fig. 6).

CPD
7, 2893–2935, 2011

In the following paragraphs we will categorize methods as to the stationarity assumptions researchers have made in their analyses. We discern four situations: no assumptions on stationarity (Sect. 2.2), assuming a stationary climate (Sect. 2.3), assuming a block-stationary climate (Sect. 2.4), and assuming a non-stationary climate (Sect. 2.5). Trend methods are reviewed in Sect. 2.6.

2.2 Enumerating extreme events

The first method of treating extremes is simply by enumerating a number of record-breaking values. These records can be discussed with respect to their spreading over time. If x of the highest values occurred in the past decade, this might give an indication of a shifting climate. The method of enumeration is often applied in communication to the media. An example is the annually returning discussions on the extremity of global mean temperatures. E.g. see the NOAA and NASA GISS websites http://www.noaanews.noaa.gov/stories2011/20110112_globalstats.html and http://www.giss.nasa.gov/research/news/20110113/, discussing the extremity of the 2010 value.

In the peer-reviewed literature enumeration is found only incidentally. For instance, Prior and Kendon (2011) studied the UK winter of 2009/2010 in relation to the severity of winters over the last 100 yr. They give an overview of coldness rankings for monthly and seasonal average temperatures, as well as rankings for the number of days with snow. Furthermore, Battipaglia et al. (2010) study temperature extremes in Central Europe reconstructed from tree-ring density measurements and documentary evidence. Their tables and graphs show a list of warm and cold summers over the past five centuries.

In the grey literature (reports) many examples of enumeration can be found. Buisman (2011) gives a detailed description of weather extremes and disasters, for a large part based on documentary information in the area of the Netherlands. His enumeration covers the period from the Middle Ages up to the present. Enumerations of disasters in recent decades are found in, e.g. WHO (2011) and Munich Re (2011).

2.3 Assuming a stationary climate

There is a wide range of studies which assume the data at hand to be stationary. Stationarity means in practice that the data are stable over time (no trends, breaks, shocks, ramps or changes in variance over time). Using this assumption, a number of statistical techniques have been applied. We give a number of examples from the literature. Theoretical background of extreme- value methods can be found in Coles (2001) or the literature mentioned, and is not clarified here.

Zorita et al. (2008) consider the likelihood that the observed recent clustering of warm record-breaking mean temperatures at global, regional and local scales may occur by chance in a stationary climate. They conclude this probability to be very low (under two different hypotheses).

Schär et al. (2004) estimate a normal distribution through monthly and summer temperatures in Switzerland, 1864–2003, to characterize the 2003 European heat wave (their Figs. 1 and 3). Barriopedro et al. (2011) estimate a normal distribution for European summer temperatures for 1500–2010. See their Fig. 2. The five coldest and highest values are highlighted. The 2010 summer temperature appears to be the highest by far.

Wehner (2010) estimates GEV distributions to pre-industrial control runs from 15 climate models in the CMIP3 dataset. These control runs are assumed to be stationary. 20-yr return periods are estimated for annual maximum daily mean surface air temperatures along with uncertainties in these return periods. Min et al. (2011) also estimate the GEV distribution. They analyze 49-yr time series of the largest one-day and five-day precipitation accumulations annually ($RX1D_t$ and $RX5D_t$). Afterwards, these distributions are used to transform precipitation data to a "probability-based index" (PI), yielding a new 49-yr time series with values between 0 and 1. Time-dependent behavior of the PI_t series is shown by estimating trends (their Fig. 1).

Della-Marta et al. (2009) apply the Peaks Over Threshold (POT) approach with declustering, and apply it to extreme wind speed indices (EWIs). The POT parameters are regarded to be time-independent.

2.4 Assuming a block-stationary climate

The block-stationary assumption differs from overall stationarity in that researchers define a period or "block" of a certain length, typically between 20 to 30 yr, where climate is assumed to be stationary. For this block PDF shapes can be assumed or fitted. Then, statistics such as average return periods can be calculated for differing blocks over time. We give four examples from the literature:

- Kharin and Zwiers (2007) evaluate temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. For that purpose they assume climate to be stationary over 20-yr periods. For selected blocks GEV distributions are estimated and 20-yr return periods are calculated. They argue that longer return periods (≥50 yr) are less advisable given the short block length of 20 yr.

- Beniston and Diaz (2004) use a block length of 30 yr to analyze the rarity of the 2003 heat wave in Europe. They estimate a normal distribution through mean summer maximum temperature data at Basel, Switzerland, for the 1961–1990 period. They argue that what may be regarded as an extreme beyond the 90th percentile under current (= stationary) climate, becomes the median by the second half of the 21st century. Their results are repeated in Trenberth and Jones (2007, p. 312, Fig. 2, lower panel).

- Barriopedro et al. (2011) analyze multi-model projections of future mega-heatwaves (their Fig. 4). To this end they choose blocks of 30 yr and base their return-period calculations on these 30-yr blocks. Uncertainties in return periods are gained through 1000 times resampling of block data. Zwiers et al. (2011) choose 10-yr blocks for the location parameter of the GEV distribution. The other two GEV parameters are kept constant in their approach.

CPD
7, 2893–2935, 2011
Weather extremes and weather-related disasters
H. Visser and A. C. Petersen

– Brown et al. (2010) analyze temporal changes in PDFs in their Figs. 5 and 6. Data are seasonal minimum and maximum temperatures over the period 1893–2005, taken from northeastern US stations. Their block size is around 28 yr. No specific PDF shape is assumed.

2.5 Assuming a non-stationary climate

Opposed to the approaches in the preceding sections there are techniques which are designed for non-stationary situations. Such techniques are time-varying extreme value approaches based on GEV and POT theory with time-varying coefficients (e.g. Coles, 2001, Sect. 6), or methods where trends and the behavior of residuals are studied. These latter methods fall apart into two groups: (i) methods which analyze trends and residuals in two stages by detrending the data first and analyzing residuals afterwards, and (ii) methods where these two stages are combined into one analysis. Examples of applying non-stationary extreme value theory are the following:

– Clarke (2002) estimates time trends in Gumbel distributed data by means of generalized linear models (GLMs). He discusses estimation issues and the detection/existence of small time trends. Trömel and Schönwiese (2005, 2007) analyze monthly total precipitation data from a German station network of 132 time series, covering the period 1901–2000. They use a decomposition technique which results in estimations of Gumbel distributions with a time-dependent location and scale parameter.

– Kharin and Zwiers (2005) estimate extremes in transient climate-change simulations. Their sample period is 1990–2100. They assume annual extremes of temperature and precipitation to be distributed according to a GEV distribution with all three parameters time-varying (linear trends). In doing so their GEV model has six unknown parameters to be estimated. Brown et al. (2008) essentially follow the same approach for extreme daily temperatures over the period 1950–2004.

CPD
7, 2893–2935, 2011

- Fowler et al. (2010) estimate GEV distributions with linear changing location parameters and apply this technique to UK extreme precipitation simulations over the period 2000–2080. Their approach deviates from that of Kharim and Zwiers (2005) and Brown et al. (2008) in that they do not assume this approach to be the only approach possible. They estimate eight different modeling approaches and evaluate the best fitting model using Akaike's AIC criterion.

- Parey et al. (2007) assume a POT model with time-varying parameters and analyze 47 temperature stations in France over the 1950–2003 period. As in Fowler et al. (2010) they consider a suit of models such as situations where station data are assumed to be stationary versus those where they are assumed to be non-stationary.

The examples thus far integrated the estimation of trends and extreme value distributions in one model. Other approaches are directed to trend estimation as such and analyze the behavior of residuals afterwards. Two examples of this two-stage approach are:

- In dendroclimatology tree-ring data are often detrended first and residuals are analysed afterwards, e.g. Stoffel et al. (2010), Visser et al. (2010) and Büntgen et al. (2011).

- Charpentier (2011) analyzes temperature data from Paris over the period 1900–2004 to evaluate the extremity of the 2003 heatwave. First, he estimates two trend models for detrending purposes: a spline nonlinear regression model and a LOWESS regression model. Then he applies a suit of ARMA, SARIMA, GARMA and fractional processes to analyze the residuals.

Examples of the integrated treatment of trends and residuals are the following:

- Wigley (2009) analyzes changes in return periods using OLS trend fitting plus a normal distribution for the residuals. He gives an example for monthly mean

CPD
7, 2893–2935, 2011
Weather extremes and weather-related disasters
H. Visser and A. C. Petersen

summer temperatures in England (the CET database). We come to this approach into more detail in Sect. 4.1.

- Visser and Petersen (2009) apply a trend model from the group of structural time series models, the so-called Integrate Random Walk (IRW) model. This IRW model has the advantage of being flexible where the flexibility can be chosen by maximum likelihood optimization. The OLS straight line is a special case of the IRW model. They apply this trend model to an indicator for extreme cold conditions in the Netherlands for the period 1901–2008. Return periods are generated along with uncertainty information on temporal changes in these return periods (cf. the TXX_t example shown in the Figs. 1, 4 and 6).

2.6 Trends in extreme weather indicators

If a particular weather or climate indicator has an extreme character by definition, important inferences can be gained by estimating trends through the time series available. If we scan the climate literature on trend methods, an enormous amount of models arises. We found the following trend models or groups of models (without being complete): low pass filters (various binomial weights; with or without end point estimates), ARIMA models and variations (SARIMA, GARMA, ARFIMA), linear trend with OLS, kernel smoothers, splines, the resistant (RES) method, Restricted Maximum Likelihood AR(1) based linear trends, trends in rare events by logistic regression, Bayesian trend models, simple Moving Averages, neural networks, Structural Time-series Models (STMs), smooth transition models, Multiple Regression models with higher order polynomials, exponential smoothing, Mann-Kendall tests for monotonic trends (with or without correction for serial correlations), trend tests against long-memory time series, robust regression trend lines (MM or LTS regression), Seidel-Lanzante trends incorporating abrupt changes, wavelets, Singular Spectrum Analysis (SSA), LOESS and LOWESS smoothing, Shiyatov corridor methods, Holmes double-detrending methods, piecewise linear fitting, Students t-test on sub-periods in time, extreme value theory with a time-varying location parameter, and last but not least, some form of expert judgment (drawing a trend "by hand"). See Mills (2010) and references therein for a discussion.

This long list of trend approaches holds for trends in climate data in general. However, the number of trend models applied to extreme weather indicators, appears to be much more limited. The trend model almost exclusively applied, is the OLS straight line. This model has the advantage of being simple and generating uncertainty information for any trend difference $[\mu_t - \mu_s]$ (indices t and s are arbitrary time points within the sample period)[1]. Uncertainty estimates are available since the slope of the trend is estimated along with its uncertainty. Examples of OLS trend fitting are given by Brown et al. (2010). They estimate trends in 17 temperature and 10 precipitation indices (all for extremes) at 40 stations. Their sample period is 1870–2005. Furthermore, Brown et al. (2010) analyze the sensitivity of their results with respect to the linearity assumption. To do so, they splitted the sample period in two parts of equal length and estimated the OLS trends on these two sub-periods.

Other examples of OLS linear trend fits can be found in Klein Tank et al. (2006) and Alexander (2006), be it that the significance of the trend slope is estimated differently. Klein Tank et al. apply the Student's t-test, while Alexander et al. apply Kendall's tau-based slope estimator along with a correction for serial correlation according to a study of Wang and Swail. Karl et al. (2008, Appendix A) choose linear trend estimation in combination with ARIMA models for the residuals. This is another way of correcting for serial correlation.

In the field of disaster studies OLS trends are the dominant method, be it that the original data are log-transformed in most cases. See Pielke (2006, Figs. 2 and 3) or Munich Re (2011, p. 47) for examples. Another trend method in this field is the moving average trend model where the flexibility is influenced by the length of the averaging window chosen. See Pielke (2006, Fig. 5) for an example.

[1]The OLS regression model reads as $y_t = \mu_t + \varepsilon_t = a + b \cdot t + \varepsilon_t$, with "$a$" the intercept, "$b$" the slope of the regression line and ε_t a noise process. Now, the variance of any trend differential $[\mu_t - \mu_s]$ follows from $\mathrm{var}(\mu_t - \mu_s) = \mathrm{var}(\hat{b} \cdot (t-s)) = (t-s)^2 \cdot \mathrm{var}(\hat{b})$.

Occasionally, other trend approaches for extreme weather indicators are reported. E.g. Klein Tank et al. (2006) use the LOWESS smoother to highlight trend patterns in extreme weather indicators (their Figs. 3, 4, 6 and 7). Tebaldi et al. (2006) do not apply any specific trend model but show increases or decreases over two distant 20-yr periods: indicator differences between 2080–2099 and 1980–1999, and between 1980–1999 and 1900–1919 (their Figs. 3 and 4). Visser (2005) applies sub-models from the class of STMs to estimate trends and uncertainty in weather indicators where trends may be flexible. The measure of flexibility is estimated by ML optimization. Frei and Schär (2001) apply logistic regression to time series of very rare precipitation events in the Alpine region of Switzerland. They include a quantification of the potential/limitation to discriminate a trend from the stochastic fluctuations in these records. Hu et al. (2011) apply Mann-Kendall tests with correction for serial correlation (no actual trend estimated in this approach).

Finally, some authors acknowledge that the use of a specific trend model, along with uncertainty analysis, may lead to deviating inferences on (significant) trend changes. Therefore, they chose to evaluate trends using more than one trend model. E.g. Moberg and Jones apply two different trend models to the same data: the OLS trend model and the resistant (RES) model. Subsequently, they evaluate all their results with respect to these two trend models. Even more methods are evaluated by Young et al. (2011). They estimate five different trend models to 23-yr wind speed and wave height data and evaluate uncertainty information for each model (their supporting material). We note that the application of more than one trend model to the same data has been published more often (not specifically for the evaluation of extremes). The reader is referred to Harvey and Mills (2003), Mills (2010) and references therein.

CPD
7, 2893–2935, 2011

3 Stationarity assumptions

3.1 Stationarity

We have seen in Sect. 2 that methods fall apart with respect to their assumption of stationarity (Sects. 2.3, 2.4 and 2.5). At first glance one may judge this choice as a matter of taste. As long as one makes his or her assumptions clear, all seems okay at this point. Of course, there is no problem as long as the data at hand are truly stationary, such as in the study of Wehner (2010) who estimates GEV distributions to pre-industrial control runs from 15 climate models, part of the CMIP3 dataset. The same holds for Villarini et al. (2011) who apply GEV distributions for extreme flooding stations with stationary data over time only.

However, inferences might go wrong if data are assumed to be stationary while they are not. Figure 2 gives an illustration of this point by simulation. Suppose that a specific weather index shows an increasing trend pattern over time. However, the year-to-year variability slowly decreases over time (heteroscedastic residuals). Now, if we would assume these data to be stationary, we would conclude that the frequency of high extremes is decreasing over time. This conclusion could be easily interpreted as an absence of climate change. However, the increasing trend in these data is contradictory to this conclusion. The example shows that conclusions on the influence of climate change should not be done on the behavior of extremes alone. Proper methods for stationary checks should be applied.

A second danger of assuming stationarity while data are in fact non-stationary, occurs if GEV distributions are applied. GEV distributions are very well suited to fit data which are stable at first and start to rise at the end. See the simulation example in Fig. 3, upper panel. This example is composed of an exponential curve where normally distributed noise is added. Now, if we regard this hundred-year long record as stationary and estimate for example the Gumbel distribution to these data, a perfect fit is found, as illustrated in the lower panel.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

CPD

7, 2893–2935, 2011

Weather extremes and weather-related disasters

H. Visser and A. C. Petersen

This result might seem surprising, but it is not. The residuals of the simulated series are normally distributed, having symmetric tails. Due to the higher values at the end of the series the right-hand tail of the distribution will become "thicker" than the tail of the normal distribution if we discard the non-stationarity at the end of the series. And this is exactly the shape of the right-hand tail of the Gumbel distribution, and more generally the GEV distribution. In practice the GEV distribution will give a good fit in many such occasions since it has three fit parameters instead of the two of the Gumbel distribution.

Our conclusion is that care should be taken if climate is assumed to be stationary. If data are assumed to be stationary while they are not, inferences might become misleading.

Thus, proper testing for stationarity versus non-stationarity is a prerequisite. For examples of stationarity tests please refer to Feng et al. (2007), Fowler et al. (2010), Furió and Meneu (2011), Villarini et al. (2011) and Rea et al. (2011).

3.2 Block stationarity

As we have shown in Sect. 2.4, a number of authors assume their data to be stationary over short periods of time, typically periods of 20 to 30 yr. Such assumptions are often made in climatology and is clearly reflected in the definition of "climate" (IPCC, 2007, WG I, Annex I): Climate in a narrow sense is usually defined as the average weather, or more rigorously, as the statistical description in terms of the mean and variability of relevant quantities over a period of time ranging from months to thousands or millions of years. The classical period for averaging these variables is 30 yr, as defined by the World Meteorological Organization [...].

Of course, if the (extreme) weather indicator at hand shows a stable behavior over the block period chosen, the choice for stationarity satisfies. However, due to rapid climate change, the stationarity assumption may be invalid, even for very short periods. Young et al. (2011) give such examples for 23-yr extreme wind speed and wave height data. They find many significant rising trends (their Table 1 and Fig. 3).

Another example is the TXX_t series shown in the upper panel of Fig. 1. The Figure shows an almost linear increase of these annual maximum temperatures. To analyze the local behavior of this trend more closely, we estimated the trend differences $[\mu_{2010} - \mu_{t-1}]$ and $[\mu_t - \mu_{t-1}]$ along with 95 % confidence limits (statistical approach explained in Visser, 2004). See Fig. 4. The figure shows that the trend value μ_{2010} in the final year 2010 is significantly larger than any trend value μ_s in the period 1951–2009 ($\alpha = 0.05$). The lower panel shows an even stronger result: all trend differences $[\mu_t - \mu_{t-1}]$ over the period 1967–2010 are significantly positive ($\alpha = 0.05$).

Again, our conclusion is that care should be taken in assuming stationarity, even for such short periods of time (20 to 30 yr). Changes in extreme weather variables may be highly significant even over these short periods.

4 Choice of probability distribution assumptions

4.1 PDF shapes: normal or GEV?

As described in Sects. 2.2, 2.3 and 2.4 different types of probability distributions have been applied to both stationary and non-stationary data. E.g. Beniston and Diaz (2004) applied the normal distribution, Visser and Petersen (2009) applied the log-normal distribution, Trömel and Schönwiese (2006) applied the Gumbel distribution and Brown et al. (2008) applied the GEV distribution. This leads to the question which distribution is preferable in which situation? Or would it be possible that different PDFs fit equally well to the same data? If the latter were true, it would still be worthwhile to choose the PDF with care if extrapolations are made far beyond the sample record length (return periods of once in 500 to 1000 yr, as in Della-Marta et al. (2009) or Lucio et al. (2010)).

In this context a discussion between Wigley (2009) and Cooley (2009) is relevant. Wigley estimated linear trends and normal distributions to monthly mean temperatures in England (the CET database, Parker and Horton (2005)). Cooley estimated GEV distributions with time-varying parameters to annual maxima of daily maximum

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

CPD

7, 2893–2935, 2011

Weather extremes and weather-related disasters

H. Visser and A. C. Petersen

temperatures, also taken from the CET database. He finds a linear fit for the GEV location (mean) parameter, and constants for the variance and shape parameter. Cooley discusses the advantages of taking the GEV distribution rather than the normal distribution. Who is right, or are both right?

We re-estimated the CET TXX_t data[2] with the IRW trend model (cf. Fig. 1), and checked the distribution of the residuals. The IRW flexibility is estimated by ML optimization and appears to be a straight line, mathematically equal to the OLS linear trend. The innovations (= one-step-ahead prediction errors) show perfect normal behavior and we conclude that a straight line, along with normally distributed residuals, gives feasible results for these TXX_t data. Compared to the trend of Cooley, our trend appears to have a slightly steeper slope: 0.0155 ± 0.005 $(1-\sigma)$ against their slope estimate 0.0142. This result implies that (i) more than one PDF may be applied to the same data and (ii) the choice of the PDF shape (slightly) influences the trend slope estimate (cf. the simulation example shown in Fig. 3).

4.2 Comparing four PDF shapes

To get a better grip on this "PDF shape discussion" we have tested four PDF shapes frequently encountered in the literature, on the same data. PDF shapes are (i) the normal distribution, (ii) the log-normal distribution, (iii) the Gumbel distribution and (iv) the GEV distribution (of which the Gumbel distribution is a special case). For such a test, we performed two groups of simulations yielding a number of TXX_t and $RX1D_t$ "look alikes". We varied the time series length N (65, 130 and 1300 yr) and the number of effective days N_{eff} (1, 60, 180 and 365 days). The latter parameter mimics the effective number of independent daily data within a year for a certain weather variable. Details are given in Appendix A.

[2] The CET TXX_t temperatures can most easily be downloaded from the Climate Explorer website: http://climexp.knmi.nl/. Choose "Daily climate indices", "UK temperatures", "maximum", and in the bottom panel: "annual (Jan–Dec)" and new variable: "max".

An example from these simulations is given in Fig. 5. Here we have plotted four PDFs for the same TXX_t simulation ($N_{eff} = 60$ days; $N = 130$ yr). This simulation resembles the Wigley-Cooley case with daily CET temperatures since 1880. The four panels show the Kolmogorov-Smirnov goodness of fit test, along with three graphic presentations (as in the lower panel of Fig. 3). The panels show that the only distribution which fits not very well, is the Gumbel distribution (right tail deviates in the QQ plot, panel lower left).

Although the simulation excercise described in Appendix A, is certainly not exhaustive, the following inferences can be made:

- both log-normal and GEV distribution fit very well for the vast majority of simulations, (both TXX_t and $RX1D_t$ simulations). This result is in line with the many examples of these PDFs in the literature, applied to real data.

- the Gumbel distribution fits only moderately to the TXX_t simulations. Much better fits are found for data which are skewed in nature, such as in case of the $RX1D_t$ simulations. This result is in line with the findings of Trömel and Schönwiese (2007) who find Gumbel distributions for 132 precipitation series in Germany (1901–2000). No Gumbel distributions have been reported in the literature for temperature data, which is in line with our TXX_t simulation results.

- The normal distribution fits well for the TXX_t simulations as long the number of years is rather small (sample periods shorter than ~130 yr). This result is in line with the Wigley-Cooley discussions for CET data since 1880. For skewed data, as in the second group of simulations, the normal distribution is not a good choice.

One might conclude from the inferences above that the GEV distribution would be the ideal PDF choice in general: (i) it fits in almost all cases and (ii) it has an interpretational background in relation to extremes. However, we note that the estimation of time-varying GEVs in combination with linearity assumptions on the three parameters,

CPD

7, 2893–2935, 2011

Weather extremes and weather-related disasters

H. Visser and A. C. Petersen

demands the estimation of six parameters (Kharin and Zwiers, 2005). And the linearity assumption for GEV parameters might be limiting in some cases. In contrast, the estimation of flexible trends and normal distributions (as in the TXX_t examples for CET and De Bilt) (i) does not fit for skewed data and (ii) lacks interpretation. However, it demands the estimation of only one parameter. Also uncertainty information on extremes is gained more easily (cf. Fig. 6). The same advantage is gained after taking logarithms of the indicator at hand, as in Visser and Petersen (2009, their Fig. 5 and Appendix). And the simulations in Appendix A show that log-normal distributions fit very well.

5 Uncertainty information

5.1 Available statistical techniques do not suffice in all cases

Uncertainty information is an important source of additional information pertaining to inferences on extremes. Within climate science, and particularly within the Intergovernmental Panel on Climate Change (IPCC), there has been increased attention for dealing with uncertainties over the last decade or so (see e.g. Moss and Schneider, 2000; Petersen, 2000, 2006; IPCC, 2005; Risbey and Kandlikar, 2007; Swart et al., 2009; Hulme and Mahony, 2010; Mastrandrea et al., 2010). Uncertainty plays a role in Sects. 2.3 through 2.6.

We scanned the literature as for their treatment of statistical uncertainties. In doing so, we discerned three levels of statistical uncertainty information:

- Class 0: research giving no statistical uncertainty information.

- Class 1: research giving point-estimate uncertainty for extreme statistics. Here, we mean uncertainty statistics at one specific point in time, such as confidence limits for a return period R_t or confidence limits for a trend estimate μ_t. An example for extremes has been given in the three panels of Fig. 6. An example for trends has been given in Fig. 1, upper panel.

CPD
7, 2893–2935, 2011

– Class 2: research giving uncertainty information both for point estimates and for differential estimates. Here, we mean "Class 1" uncertainty information along with uncertainty information on differential statistics such as the return-period differential $[R_t - R_s]$, or trend differentials $[\mu_t - \mu_s]$ (times t and s lie in the sample period with $t > s$)[3]. An example has been given in Fig. 4. This graph shows the trend differentials $[\mu_{2010} - \mu_t]$ and $[\mu_t - \mu_{t-1}]$, along with 95 % confidence limits.

With respect to return periods or the chance for crossing pre-defined thresholds we found only rarely examples of "Class 0". In most cases "Class 1" uncertainty information is given: Feng and Nadarajah (2007), Della-Marta et al. (2009), Fowler et al. (2010), Wehner (2010) and Lucio et al. (2010). However, we found that "Class 2" uncertainty information is lacking almost completely. The only example we found, was in a previous paper of ours (Visser and Petersen, 2009). There we give approximate uncertainty estimates for return period differentials in an Appendix A.

As for trends we only rarely found examples of "Class 0" uncertainty. Examples lacking uncertainty information are mostly found in the estimation of trends in disaster data: although OLS linear trends have been applied (and thus uncertainty information is easily available), no uncertainty information is given in publications. Other examples are those where moving averages of other digital filters have been applied. These trend models are not statistical in nature and, thus, do not give uncertainty information.

Since most articles apply OLS linear trend fits to their data, both "Class 1" and "Class 2" uncertainty information are covered at the same time (cf. footnote 1). Examples are Klein Tank and Können (2003), Klein Tank et al. (2006), Alexander et al. (2006), Brown et al. (2010), Min et al. (2011) and Charpentier (2011). Brown et al. (2008) give full statistical uncertainty information for the time-varying location parameter of the GEV

[3]We note that some researchers apply the Mann-Kendall test for monotonic trends (e.g. Nasri and Modarres, 2009; Young et al., 2011). Here, the significance is tested for the whole sample period only, without specifying the trend shape. Thus, this approach does not fall within the category of "Class 2". The same holds for trends based on ARIMA models: trends are filtered from the data but an actual trend is not given.

CPD
7, 2893–2935, 2011

distribution. Trends from the class of structural time series models (STMs), as shown here in Figs. 1 and 4, give a generalization of the OLS linear trend: they also give full statistical uncertainty information (Visser, 2004; Visser et al., 2010).

Our “uncertainty scan” shows that full uncertainty information (“Class 2”) is missing for statistics such as return periods or the chance for crossing thresholds. And the reason for that is simple: the statistical literature on extremes, such as Coles (2001), does not report methods to compute these differential uncertainties. Therefore, our conclusion is a simple one: such methods should be developed. For trend estimation we conclude that full uncertainty information is available as long as OLS linear trends or trend models from the class of STMs are chosen.

5.2 Best modeling practices and uncertainty

As described at the end of Sect. 2.6 some authors have chosen to apply more than one trend model to analyze their data. This type of sensitivity analysis does not evaluate uncertainties in estimators only, but also tries to find the influence of under-lying model assumptions – thus often moving beyond the realm of statistical uncertainty into scenario (what-if) uncertainty. See Mills (2010) and Charpentier (2011, Sect. 2). Other examples are:

- Moberg and Jones (2005) evaluate trends in extreme weather indicators using two trend models: the OLS linear trend and the RES method. The latter method is more appropriate if the data contain outliers and behave non-normal.
- A variation is given by Young et al. (2011, Table S1). They present five different significance tests for their trends.

In fact, the evaluation of different trend models, and corresponding uncertainty inferences, is a way of evaluating structural uncertainty, i.e. evaluating the potential influence of specific model assumptions.

An illustration of the importance of considering more than one trend model, is given in Fig. 7. The upper panel shows the economic losses due to global weather-related disasters, as published by Munich Re (2010). We note that this graph is published in many variations such as losses due great natural disasters, etc. (Pielke, 2006, Fig. 2; Munich Re, 2011, p. 47). The trend is estimated by fitting the OLS linear trend model, after taking logarithms of the event data. The result is an exponential increasing trend. If an IRW trend is estimated, where the flexibility is optimized by ML (Visser, 2004), a different trend pattern arises (lower panel): an increase up to 1995 and a stabilization afterwards. The trend value in 2009 is significant higher than trend values before 1987 (tested for $\alpha = 0.05$, graph not shown here). This example illustrates that the interpretation of trend patterns in extreme weather indicators might be influenced by the trend method chosen.

Another approach to assess structural uncertainty is the evaluation of the stationarity/non-stationarity of the data at hand (cf. discussion in Sect. 3). Examples are:

- Feng and Nadarajah (2007) estimate both stationary and non-stationary GEVs, and calculate return periods for both approaches.
- Fowler et al. (2010) evaluate 8 GEV models, both stationary and non-stationary. For choosing the most appropriate model they use the AIC criterion.

We found two other sensitivity approaches which could be categorized under the term "best modeling practices". In the field of future extremes it might be of importance to evaluate extreme statistics on the basis of more than one GCM or RCM. Examples are:

- Kharin et al. (2007) give multi-model uncertainty limits for 20-yr return periods in their Figs. 3, 5, 6 and 7, based on 14 IPCC AR4 models.
- Wehner (2010, Fig. 1) calculates the inter-model uncertainty for return periods, based on daily data from 15 different CMIP3 models.

CPD
7, 2893–2935, 2011

– Barriopedro et al. (2011, Figs. 4 and S12) evaluate return periods for mega-heatwaves on the basis of 11 RCMs and one reanalysis run.

A second sensitivity approach deals with the sensitivity of trend estimates and corresponding uncertainties in relation to the sample period length. Examples are:

– Moberg and Jones (2005, Table III) show significant trends for four periods: 1901–1999, 1921–1999, 1901–1950 and 1946–1999.

– Klein Tank et al. (2006, Table 2) show trend decadal increments with uncertainties for the periods 1961–2000 and 1901–2000.

We note that an analoguous sensitivity example for linear trends has been given by Trenberth and Jones (2007, FAQ 3.1, Fig. 1) for global mean temperatures.

In our judgment, some form of sensitivity analysis is important to assess the reliability of results. This conclusion of course pertains more generally to environmental research.

6 Coupling extremes or disasters to climate change

There are several ways to couple trends in extremes or disaster to (anthropogenic) climate change (see, e.g. Hegerl and Zwiers (2007), Zwiers et al. (2011) and Min et al. (2011) for spatio-temporal approaches). One has to be careful, however, in coupling individual extremes to climate change. In fact, statistical inferences are about chances for groups of events and not about individual events.

Even though most publications do not strictly couple single extremes to climate change, that is, with 100 % certainty, many are suggestive about the connection while they focus actually on the changed chances. A recent example on flooding is Pall et al. (2011) and an example of suggestive information on the Pakistan floodings in 2010 is given in Fig. 8. An earlier example is constituted by publications on the 2003 heatwave in Europe.

CPD
7, 2893–2935, 2011

PBL (2010) has analyzed the presenting of this 2003 heat wave as a consequence of climate change in IPCC (2007). The Working Group II Summary for Policymakers states, for Europe, on p. 14: "For the first time, wide-ranging impacts of changes in current climate have been documented: retreating glaciers, longer growing seasons, shift of species ranges, and health impacts due to a heat wave of unprecedented magnitude. The observed changes described above are consistent with those projected for future climate change." This text, as well as its counterparts in Table TS 4.2 of the Technical Summary (p. 51) and in the Executive Summary (p. 543), present the health impacts from the 2003 heat wave as an example of "wide-ranging impact of changes in current climate". Thus, the text implicitly suggests that the 2003 heat wave is the result of recent climate change.

However, one can never attribute a specific extreme weather event of the past – such as that particular heat wave – to changes in current climate. In fact, we agree with Schär and Jendritzky (2004) who stated the following: "The European heatwave of 2003: was it merely a rare meteorological event or a first glimpse of climate change to come? Probably both." Stott et al. (2004) come to a comparable conclusion: "It is an ill-posed question whether the 2003 heatwave was caused, in a simple deterministic sense, by a modification of the external influence on climate – for example, increasing concentration of greenhouse gases in the atmosphere – because almost any such weather event might have occurred by chance in an unmodified climate."

7 Conclusions

In this article we have given a concise overview of methods applied in the peer-review literature. Furthermore, we have evaluated these methods as for specific choices that researchers have made. These choices are (i) the choice for a specific type of stationarity, (ii) the choice for a specific PDF shape for the data (or residuals) at hand, (iii) the treatment of uncertainties and (iv) the coupling of extremes or disasters to climate change. We draw the following conclusions:

CPD
7, 2893–2935, 2011
Weather extremes and weather-related disasters
H. Visser and A. C. Petersen

- In making a choice for treating data as stationary or non-stationary, good testing is essential. Inferences on extremes may be wrong if data are assumed stationary while they are not (cf. Figs. 2 and 3). Some researchers choose for block-stationarity (blocks of 20 to 30 yr). However, climate may be non-stationary even for such short periods (cf. Figs. 1 and 4). Thus, such an assumption needs testing too.

- In calculating statistics such as average return periods, a certain PDF shape is assumed. We found that often more than one PDF shape fits to the same data (cf. the Cooley-Wigley example, and Fig. 5). From a simulation study we conclude that both the GEV and the log-normal PDF fit very well to a variety of indicators (both symmetric and skewed data/residuals). The normal PDF performs well for data which are (i) essentially symmetrical in nature (such as extremes for temperature data) and (ii) have relatively short sample periods (∼130 yr). The Gumbel PDF fits well for data which are skewed in nature (such as extreme indicators for precipitation). For symmetrical situations the Gumbel PDF does not perform very well.

- Statistical techniques are not available for all cases of interest. We found that theory is lacking for uncertainties for differential statistics of return periods, i.e. uncertainties for a particular differences $[R_t - R_s]$. For trends these statistics are available as long as OLS trends or structural time series models (STMs) are chosen (cf. Figs. 1 and 4).

- It is advised to test conclusions on extremes with respect to assumptions underlying the modeling approach chosen (structural uncertainty). Examples are given as for (i) the application of different trend models to the same data, (ii) stationary versus non-stationary GEV models, (iii) evaluation of extremes for a suite of GCMs or RCMS to evaluate statistics in the future, and (iv) the role of the sample period length. An example has been given where the choice of a specific trend model influences the inferences made (Fig. 7).

- The coupling of extremes to climate change should be performed by spatio-temporal detection methods. However, in the communication of extremes to the media it occurs that researchers couple one specific exceptional extreme event or disaster to climate change. This (suggestive) coupling should be avoided (Fig. 8). Statistical inferences are always directed to chances for groups of data. They do not allow to give a clue for one specific occurrence within that group.

Appendix A

Simulation and PDF shapes

As described in Sect. 4.2 we have tested four PDF shapes frequently encountered in the literature, on the same data. PDF shapes are: the normal, the log-normal, the Gumbel and the GEV distribution (of which the Gumbel distribution is a special case). For such a test, we performed two groups of simulations, one yielding TXX_t "look alikes" and one yielding $RX1D_t$ "look alikes". The first set is totally based on random drawings from a normal distribution for daily values; the second set is based on real daily precipitation totals over the period 1906–2005 (De Bilt, the Netherlands). We varied the time series length N (65, 130 and 1300 yr) and the number of effective days N_{eff} (1, 60, 180 and 365 days). The latter parameter mimics the effective number of independent daily data within a year for a certain weather variable. The judgment of distributional fit has been done with two criteria: visual inspection of the QQ plot and the p value from the Kolmogorov-Smirnov goodness of fit test ($p < 0.05$: bad result; $p > 0.80$: very good result). See Fig. 5 for an example. Each judgment was repeated three times to rule out the influence of incidental deviating simulation results.

Table 1 shows that the log-normal and the GEV distribution give good fits for all simulations (all judgments are "+/++" or "++"). This result is independent of the specific choices made for N_{eff} or N. The normal distribution fits well for the TXX_t "look alikes" as long as time series are shorter than ~130 yr of length and N_{eff} shorter than

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |
CPD
7, 2893–2935, 2011

Weather extremes and weather-related disasters

H. Visser and A. C. Petersen

180 days. The fit for the precipitation simulations are moderate to bad throughout. For the Gumbel distribution, the situation is the other way around: a moderate result for the temperature simulations and a good result for the precipitation simulations. Time series with 1300 yr of length are the only exception here.

References

Aguilar, E., Auer, I., Brunet, M., Peterson, T. C., and Wieringa, J.: Guidelines on climate metadata and homogenization, WMO report, WMO/TD No. 1186, 2003.

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Rupa Kumar, K., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D. B., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M., and Vazquez-Aguirre, J. L.: Global observed changes in daily climate extremes of temperature and precipitation, J. Geophys. Res., 111, D05109, doi:10.1029/2005JD006290, 2006.

Barriopedro, D., Fischer, E. M., Luterbacher, J., Triog, R. M., and Garcia-Herrera, R.: The hot summer of 2010: redrawing the temperature record map of Europe, Science, 332, 220–224, 2011.

Beniston, M. and Diaz, H. F.: The 2003 heat wave as an example of summers in a greenhouse climate? Observations and climate model simulations for Basel, Switzerland, Global Planet. Change, 44, 73–81, 2004.

Bouwer, L. M.: Have disaster losses increased due to anthropogenic climate change?, B. Am. Meteorol. Soc., 92, 39–45, 2011.

Brandsma, T., Können, G. P., and Wessels, H. R. A.: Empirical estimation of the effect of urban heat advection on the temperature series of De Bilt (The Netherlands), Int. J. Climatol., 23, 829–845, 2002.

Brown, S. J., Caesar, J., and Ferro, C. A. T.: Global change in extreme daily temperature since 1950, J. Geophys. Res., 113, D05115, doi:10.1029/2006JD008091, 2008.

Brown, P. J., Bradley, R. S., and Keimig, F. T.: Changes in extreme climate indices for the Northeastern United States, 1870–2005, J. Climate, 23, 6555–6572, 2010.

Buisman, J.: Extreme Weather! A summary of cold winters and hot summers, hail and tornados, storms and floodings, Van Wijnen, Franeker, 2011 (in Dutch).

Büntgen, U., Tegel, W., Nicolussi, K., McCormick, M., Frank, D., Trouet, V., Kaplan, J. O., Herzig, F., Heussner, K. U., Wanner, H., Luterbacher, J., and Esper, J. : 2500 Years of European climate variability and human susceptibility, Science, 331, 578–582, 2011.
Charpentier, A.: On the return period of the 2003 heat wave, Climatic Change, online version, 2011.
Clarke, R. T.: Estimating time trends in Gumbel-distributed data by means of generalized linear models, Water Resour. Res., 38, 16/1–16/11, 2002.
Coles, S.: An introduction to statistical modeling of extreme values, Springer Series in Statistics, Springer-Verlag, London, 2001.
Cooley, D.: Extreme value analysis and the study of climate change. A commentary on Wigley 1988, Climatic Change, 96, 77–83, 2009.
Della-Marta, P. M., Mathis, H., Frei, C., Liniger, M. A., Kleinn, J., and Appenzeller, C.: The return period of wind storms over Europe, Int. J. Climatol., 29, 437–459, 2009.
Feng, S., Nadarajah, S., and Hu, Q.: Modeling annual extreme precipitation in China using the generalized extreme value distribution, J. Meteorol. Soc. Jpn., 85, 599–613, 2007.
Fowler, H. J., Cooley, D., Sain, S. R., and Thurston, M.: Detecting change in UK extreme precipitation using results from the climateprediction.net BBC climate change experiment, Extremes, 13, 241–267, 2010.
Furió, D. and Meneu, V.: Analysis of extreme temperatures for four sites across Peninsular Spain, Theor. Appl. Climatol., 104, 83–99, 2011.
Gall, M., Borden, K. A., and Cutter, S. L.: When do losses count? Six fallacies of natural hazards loss data, B. Am. Meteorol. Soc., 90, 799–809, 2009.
Gamble, J. L., Ebi, K. L., Grambsch, A. E., Sussman, F. G., and Wilbanks, T. J.: Analyses of the effects of global change on human health and welfare and human systems, CCSP report, Synthesis and Assessment Product, 4.6, 2008.
Guha-Sapir, D., Vos, F., Below, R., and Ponserre, S.: Annual disaster statistical review 2010. The numbers and trends, CRED report, available at: http://www.cred.be/sites/default/files/ADSR_2010.pdf, 2011.
Harvey, A. C.: Forecasting, structural time series models and the Kalman filter, Cambridge University Press, Cambridge, 1989.
Harvey, D. I. and Mills, T.: Modelling trends in Central England temperatures, J. Forecasting, 22, 35–47, 2003.
Hegerl, G. C. and Zwiers, F. W.: Understanding and attributing climate change, in: Climate

CPD
7, 2893–2935, 2011
Weather extremes and weather-related disasters
H. Visser and A. C. Petersen

Change, 2007: the Physical Science Basis, edited by: Solomon, S., Qin, D., Manning, M., Marquis, M., Averyt, K., Tignor, M. M. B., Leroy Miller Jr., H., and Chen, Z., Cambridge University Press, Cambridge, 2007.
Höppe, P. and Pielke Jr., R. A. (Eds.): Climate change and disaster losses. Understanding and attributing trends and projections, Hohenkammer workshop, available at: http://cstpr.colorado.edu/sparc/research/projects/extreme_events/munich_workshop/workshop_report.html, 2006.
Hu, Y., Maskey, S., and Uhlenbrook, S.: Trends in temperature and rainfall extremes in the yellow river source region, China, Climatic Change, online version, 2011.
Hulme, M. and Mahony, M.: Climate change: What do we know about the IPCC?, Prog. Phys. Geog., 34, 705–718, 2010.
IPCC: Guidance Notes for Lead Authors of the IPCC Fourth Assessment Report on Addressing Uncertainties, Intergovernmental Panel on Climate Change (IPCC), Geneva, Switzerland, 2005.
IPCC: Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, England, 2007.
Karl, T. R., Meehl, G. A., Miller, C. D., Hassol, S. J., Waple, A. M., and Murray, W. L.: Weather and climate extremes in a changing climate, CCSP report, Synthesis and Assessment Product, 3.3, 2008.
Kharin, V. V. and Zwiers, F. W.: Estimating extremes in transient climate change simulations, J. Climate, 18, 1156–1173, 2005.
Kharin, V. V., Zwiers, F. W., Zhang, X., and Hegerl, G. C.: Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations, J. Climate, 20, 1419–1444, 2007.
Klein Tank, A. M. G. and Können, G. P.: Trends in indices of daily temperature and precipitation extremes in Europe, 1946–99, J. Climate, 16, 3665–3680, 2003.
Klein Tank, A. M. G., Peterson, T. C., Quadir, D. A., Dorji, S., Zou, X., Tang, H., Santhosh, K., Joshi, U. R., Jaswal, A. K., Kolli, R. K., Sikder, A. B., Deshpande, N. R., Revadekar, J. V., Yeleuova, K., Vandasheva, S., Faleyeva, M., Gomboluudev, P., Budhathoki, K. P., Hussain, A., Afzaal, M., Chandrapala, L., Anvar, H., Amanmurad, D., Asanova, V. S., Jones, P. D., New, M. G. and, Spektorman, T.: Changes in daily temperature and precipitation extremes in central and south Asia, J. Geophys. Res., 111, D16105, doi:10.1029/2005JD006316, 2006.

Klein Tank, A. M. G., Zwiers, F. W., and Zhang, X.: Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation, WMO report, WCDMP-No. 72, http://www.wmo.int/pages/prog/wcp/wcdmp/wcdmp_series/documents/WCDMP_72_TD_1500_en_1.pdf, 2009.
Mastrandrea, M. D., Field, C. B., Stocker, T. F., Edenhofer, O., Ebi, K. L., Frame, D. J., Held, H., Kriegler, E., Mach, K. J., Matschoss, P. R., Plattner, G.-K., Yohe, G. W., and Zwiers, F. W.: Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties, Intergovernmental Panel on Climate Change (IPCC), Geneva, Switzerland, 2010.
Mearns, L. O., Katz, R. W., and Schneider, S. H.: Extreme high-temperature events: changes in their probabilities with changes in mean temperature, J. Clim. Appl. Meteorol., 23, 1601–1613, 1984.
Min, S. K., Zhang, X., Zwiers, F. W., and Hegerl, G. C.: Human contribution to more-intense precipitation extremes, Plus Supplementary Information, Nature, 470, 378–381, 2011.
Mills, T. C.: 'Skinning a cat': alternative models of representing temperature trends. An editorial comment, Climatic Change, 101, 415–426, 2010.
Moberg, A. and Jones, P. D.: Trends in indices for extremes in daily temperature and precipitation in Central and Western Europe, 1901–99, Int. J. Climatol., 25, 1149–1171, 2005.
Moss, R. and Schneider, S.: Uncertainties. Guidance Papers on the Cross Cutting Issues of the Third Assessment Report of the IPCC, edited by: Pachauri, R., Taniguchi, T., and Tanaka, K., Intergovernmental Panel on Climate Change (IPCC), Geneva, Switzerland, 2000.
Munich Re: Weather extremes, climate change, Cancún, 2010.
Munich Re: Topics Geo. Natural catastrophes 2010. Analyses, assessments, positions, Annual Munich Re report, 2011.
Nasri, M. and Modarres, R.: Dry spell trend analysis of Isfahan Province, Iran, Int. J. Climatol., 29, 1430–1438, 2009.
Neumayer, E. and Barthel, F.: Normalizing economic loss from natural disaster: a global analysis, Global Environ. Chang., 21, 13–24, 2011.
Orlowsky, B. and Seneviratne, S. I.: Global changes in extreme events: regional and seasonal dimension, Climatic Change, online version, 2011.
Pall, P., Aina, T., Stone, D. A., Stott, P. A., Nozawa, T., Hilberts, A. G. J., Lohmann, D., and Allen, M. R.: Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000, Nature, 470, 382–386, 2011.

CPD
7, 2893–2935, 2011

Weather extremes and weather-related disasters

H. Visser and A. C. Petersen

Parey, S., Malek, F., Laurent, C., and Dacunha-Castelle, D.: Trends and climate evolution: statistical approach for very high temperatures in France, Climate Change, 81, 331–352, 2007.

Parker, D. E. and Horton, E. B.: Uncertainties in central England temperature 1878–2003 and some improvements to the maximum and minimum series, Int. J. Climatol., 25, 317–342, 2005.

PBL: Assessing an IPCC Assessment. An Analysis of Statements on Projected Regional Impacts in the 2007 Report, PBL Netherlands Environmental Assessment Agency, The Hague, The Netherlands, 2010.

Petersen, A. C.: Philosophy of climate science, B. Am. Meteorol. Soc., 81, 265–271, 2000.

Petersen, A. C.: Simulating Nature: A Philosophical Study of Computer-Simulation Uncertainties and Their Role in Climate Science and Policy Advice, Het Spinhuis Publishers, available at: http://hdl.handle.net/1871/11385, 2006.

Pielke Jr., R. A.: Disasters, death and destruction. Making sense of recent calamities, Oceanography, 19, 138–147, 2006.

Pielke Jr., R. A., Gratz, J., Landsea, C. W., Collins, D., Saunders, M. A., and Musulin, R.: Normalized hurricane damage in the United States: 1900–2005, Natural Hazards Review, 9, 29–42, 2008.

Prior, J. and Kendon, M.: The UK winter of 2009/2010 compared with severe winters of the last 100 years, Weather, 66, 4–10, doi:10.1002/wea.735, 2011.

Rea, W., Reale, M., and Brown, J.: Long memory in temperature reconstructions, Climatic Change, online version, 2011.

Risbey, J. S. and Kandlikar, M.: Expressions of likelihood and confidence in the IPCC uncertainty assessment process, Climatic Change, 85, 19–31, 2007.

Schär, C. and Jendritzky,G.: Climate change: hot news from summer 2003, Nature, 432, 559–560, 2004.

Schär, C., Vidale, P. L., Lüthi, D., Frei, C., Häaberli, C., Liniger, M. A., and Appenzeller, C.: The role of increasing temperature variability in European summer heatwaves, Nature, 427, 332–336, 2004.

Shen, S. S. P., Gurung, A. B., Oh, H. S., Shu, T., and Easterling, D. R.: The twentieth century contiguous US temperature changes indicated by daily data and higher statistical moments, Climatic Change, online version, 2011.

Stern, D. I. and Kaufmann, R. K.: Detecting a global warming signal in hemispheric temperature

CPD

7, 2893–2935, 2011

Weather extremes and weather-related disasters

H. Visser and A. C. Petersen

series: a structural time series analysis, Climatic Change, 47, 411–438, 2000.

Stephens, G. L., L'Ecuyer, T., Forbes, R., Gettlemen, A., Golaz, J. C., Bodas-Salcedo, A., Suzuki, K., Gabriel, P., and Haynes, J.: Dreary state of precipitation in global models, J. Geophys. Res., 115, D24211, doi:10.1029/2010JD014532, 2010.

Stoffel, M., Bollschweiler, M., Butler, D. R., and Luckman, B. H.: Tree rings and natural hazards: a state-of art, Springer Verlag, Dordrecht, 2010.

Stott, P. A., Stone, D. A., and Allen, M. R.: Human contribution to the European heatwave of 2003, Nature, 432, 610–614, 2004.

Swart, R., Bernstein, L., Ha-Duong, M., and Petersen, A. C.: Agreeing to disagree: Uncertainty management in assessing climate change, impacts and responses by the IPCC, Climatic Change, 92, 1–29, 2009.

Tebaldi, C., Hayhoe, K., Arblaster, J. M., and Meehl, G. A.: Going to the extremes – an intercomparison of model-simulated historical and future changes in extreme events, Climatic Change, 79, 185–211, 2006.

Trenberth, K. A., Jones, P. D., Ambenje, P., Bojariu, R., Easterling, D., Klein Tank, A. M. G., Parker, D., Rahimzadeh, F., Renwick, J. A., Rusticucci, M., Soden, B., and Zhai, P.: Observations: surface and atmospheric climate change, in: Climate Change 2007: the Physical Science Basis, edited by: Solomon, S., Qin, D., Manning, M., Marquis, M., Averyt, K., Tignor, M. M. B., Leroy Miller Jr., H., and Chen, Z., Cambridge University Press, Cambridge, 2007.

Trömel, S. and Schönwiese, C. D.: A generalized method of time series decomposition into significant components including probability assessments of extreme events and application to observational German precipitation data, Meteorol. Z., 14, 417–427, 2005.

Trömel, S. and Schönwiese, C. D.: Robust trend estimation of observed German precipitation, Theor. Appl. Climatol., 93, 107–115, 2007.

Villarini, G., Smith, J. A., Serinaldi, F., Ntelekos, A. A., and Schwarz, U.: Analyses of extreme flooding in Austria over the period 1951–2006, Int. J. Climatol., doi:10.1002/joc.2331, online version, 2011.

Visser, H.: Estimation and detection of flexible trends, Atmos. Environ., 38, 4135–4145, 2004.

Visser, H.: The significance of climate change in the Netherlands. An analysis of historical and future trends (1901–2020), MNP report 550002007/2005, Bilthoven, http://www.rivm.nl/bibliotheek/rapporten/550002007.pdf, 2005.

Visser, H.: The chance for extreme warm days in the Netherlands, MNP report 550032010, Bilthoven, http://www.rivm.nl/bibliotheek/rapporten/550032010.pdf, 2007 (in Dutch).

CPD

7, 2893–2935, 2011

Visser, H. and Molenaar, J.: Trend estimation and regression analysis in climatological time series: an application of structural time series models and the Kalman filter, J. Climate, 8, 969–979, 1995.

Visser, H. and Petersen, A. C.: The likelihood of holding outdoor skating marathons in the Netherlands as a policy-relevant indicator of climate change, Climatic Change, 93, 39–54, 2009.

Visser, H., Büntgen, U., D'Arrigo, R., and Petersen, A. C.: Detecting instabilities in tree-ring proxy calibration, Clim. Past, 6, 367–377, doi:10.5194/cp-6-367-2010, 2010.

Von Storch, H. and Zwiers, F. W.: Statistical analysis in climate research, Cambridge University Press, Cambridge, 1999.

Wehner, M.: Sources of uncertainty in the extreme value statistics of climate data, Extremes, 13, 205–217, 2010.

WHO: Weather extremes in a changing climate: hindsight on foresight, WHO brochure, http://www.wmo.int/pages/mediacentre/news/documents/1075_en.pdf, 2011.

Wigley, T. M. L.: Impact of extreme events, Nature, 316, 106–107, 1985.

Wigley, T. M. L.: The effect of changing climate on the frequency of absolute extreme events, Climatic Change, 97, 67–76 (reprinted from Climate Monitor, 17, 44–55, 1988), 2009.

Young, I. R., Zieger, S., and Babanin, A. V.: Global trends in wind speed and wave height, Science, 332, 451–455 (with supporting online material), 2011.

Zorita, E., Stocker, T. F., and von Storch, H.: How unusual is the recent series of warm years?, Geophys. Res. Lett., 35, L24706, doi:10.1029/2008GL036228, 2008.

Zwiers, F. W., Zhang, X., and Feng, Y.: Anthropogenic influence on long return period daily temperature extremes at regional scales, J. Climate, 24, 881–892, 2011.

CPD

7, 2893–2935, 2011

Table 1. Judgments of distributional fits for (i) simulated meteorological data (cf. Fig. 5) and (ii) daily precipitation data in the Netherlands. Meaning of codes: – – stands for a very bad fit; – stands for a bad fit; + stands for a good fit; ++ stands for a very good fit. These judgments are based on visual inspection of the actual fits and on p-values of the Kolmogorov-Smirnov goodness of fit tests. All judgments are based on three repeated simulations (using different seeds in random number generation).

	Simulations based on normally distributed daily data				Simulations based on 100 yr of daily precipitation data in the Netherlands			
	Normal	Log-normal	Gumbel	GEV	Normal	Log-normal	Gumbel	GEV
$N = 65$ yr $N_{eff} = 1$ day	++	NA	– –	+	+	++	+	++
$N = 65$ yr $N_{eff} = 60$ days	+	+/++	–/+	+	–/+	+	+/++	+/++
$N = 65$ yr $N_{eff} = 180$ days	+	++	+	++	–/+	+/++	+	+/++
$N = 65$ yr $N_{eff} = 365$ days	–/+	+/++	+/++	++	–/+	+/++	+	+/++
$N = 130$ yr $N_{eff} = 1$ day	++	NA	– –	++	–/+	NA	+	++
$N = 130$ yr $N_{eff} = 60$ days	+	++	–/+	+/++	–/+	+/++	+	+/++
$N = 130$ yr $N_{eff} = 180$ days	–/+	+	+	++	–	+	+	+
$N = 130$ yr $N_{eff} = 365$ days	–	+	+	++	–	++	++	++
$N = 1300$ yr $N_{eff} = 1$ day	++	NA	– –	+	– –	NA	++	++
$N = 1300$ yr $N_{eff} = 60$ days	– –	+/++	– –	+/++	– –	+	–	–
$N = 1300$ yr $N_{eff} = 180$ days	– –	+	–/– –	++	– –	+	– –	– –
$N = 1300$ yr $N_{eff} = 365$ days	– –	+	– –	++	– –	–/+	– –	– –

CPD
7, 2893–2935, 2011

Weather extremes and weather-related disasters

H. Visser and A. C. Petersen

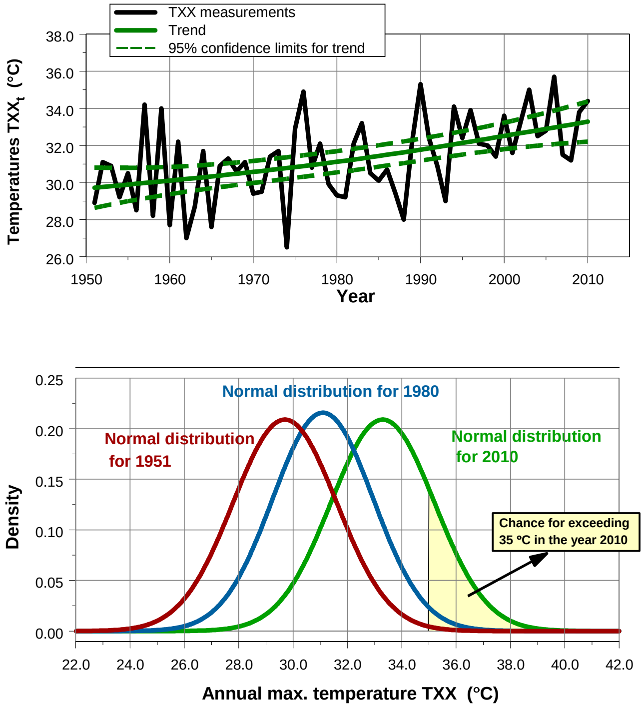

Fig. 1. Example of an extreme weather indicator, the TXX$_t$ series for station De Bilt in the Netherlands. The upper panel shows the annual data, along with an IRW trend fit (μ_t) and 95 % confidence limits. The lower panel shows three normal distributions corresponding to the years 1951, 1980 and 2010. The yellow area illustrates the chance of crossing the 35 °C threshold. Clearly, the area for the years 1951 and 1980 is much smaller, a phenomenon first shown by Mearns et al. (1984) and Wigley (1985). A return period is calculated as the inverse of these chances. For each of the three normal distributions one could calculate the temperature which is exceeded once in x years, the x-year return periods. For statistical details see Von Storch and Zwiers (1999, Sect. 2). Chances and return periods are further illustrated in Fig. 6.

CPD
7, 2893–2935, 2011

H. Visser and A. C. Petersen

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

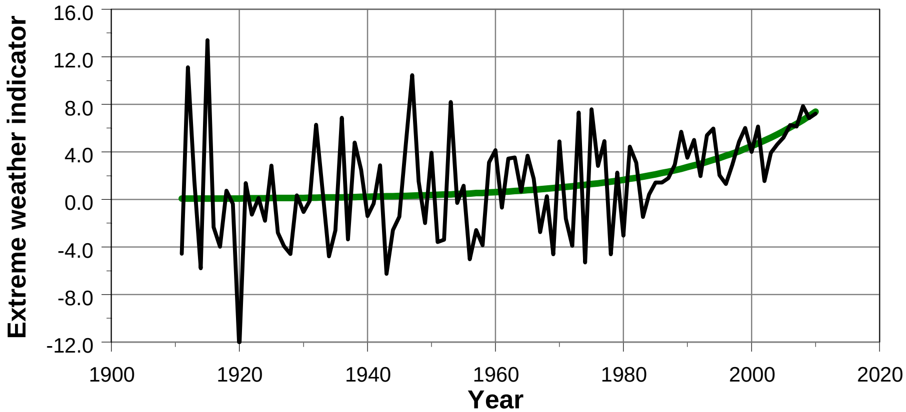

Fig. 2. Simulated extreme weather indicator for the sample period 1911–2010. The "measurements" are gained by choosing an exponential as a "trend" (green line) and adding a normally distributed white noise process to this trend. The variance of the noise is linearly decreasing over time.

CPD
7, 2893–2935, 2011

Weather extremes and weather-related disasters

H. Visser and A. C. Petersen

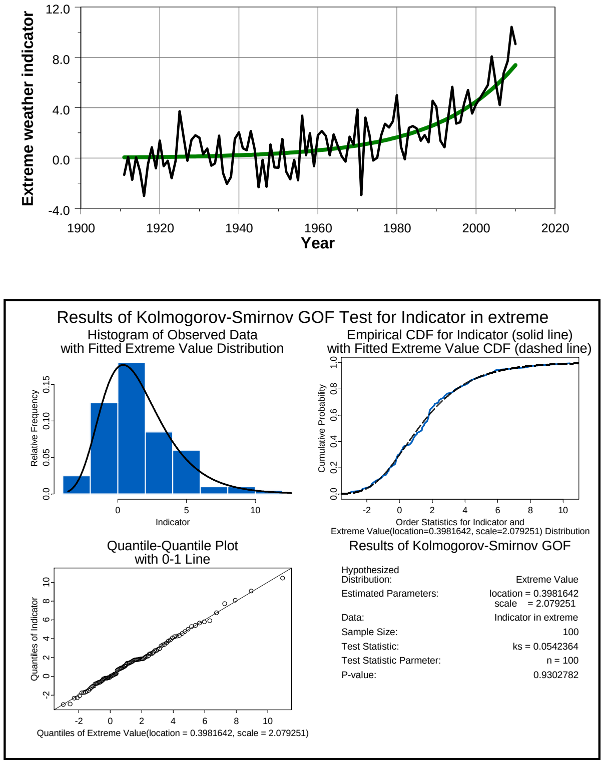

Fig. 3. The upper panel shows a simulated extreme weather indicator over the sample period 1911–2010. The "measurements" are gained by choosing an exponential as a "trend" and adding normal distributed white noise to this trend (constant variance) . If it is assumed that the measurements follow a stationary process, the data appear to follow an extreme value (Gumbel) distribution. This is shown in the lower four panels which are generated by the S-PLUS Envstats module. Shown is the Kolmogorov-Smirnov test, where the data are compared to a Gumbel distribution. The Gumbel distribution appears to fit very well (QQ-plot shows all data on the 0–1 line; p-value of the KS test is 0.93).

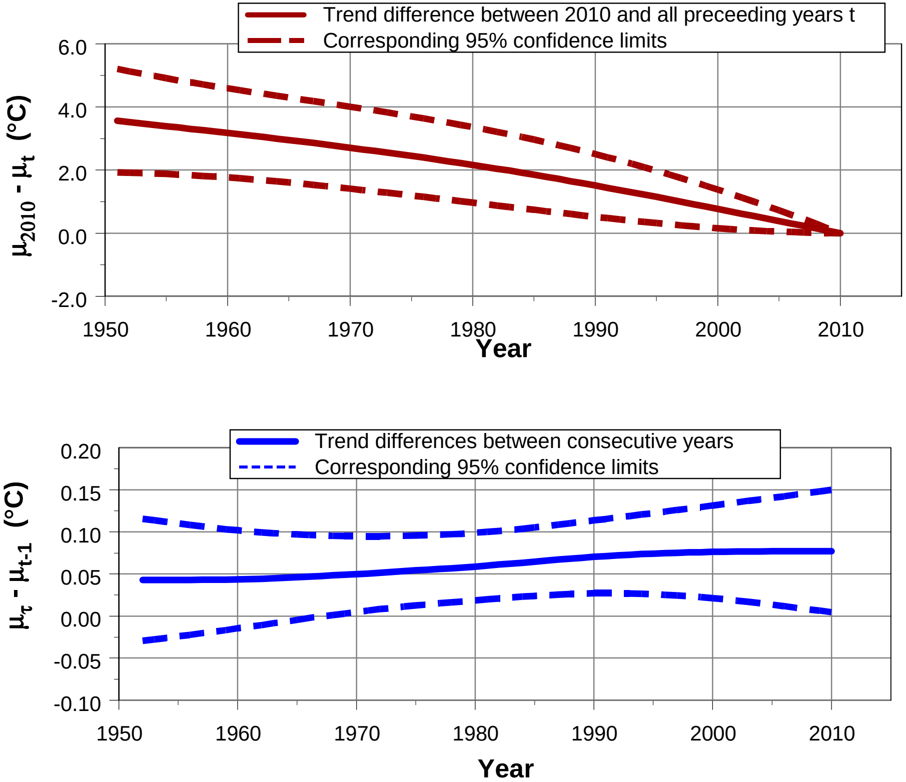

Fig. 4. Uncertainties for the IRW trend estimates shown in Fig. 1. The upper panel shows the trend difference $[\mu_{2010} - \mu_t]$ along with 95 % confidence limits, the lower panel the trend differences $[\mu_t - \mu_{t-1}]$ along with 95 % confidence limits.

CPD
7, 2893–2935, 2011

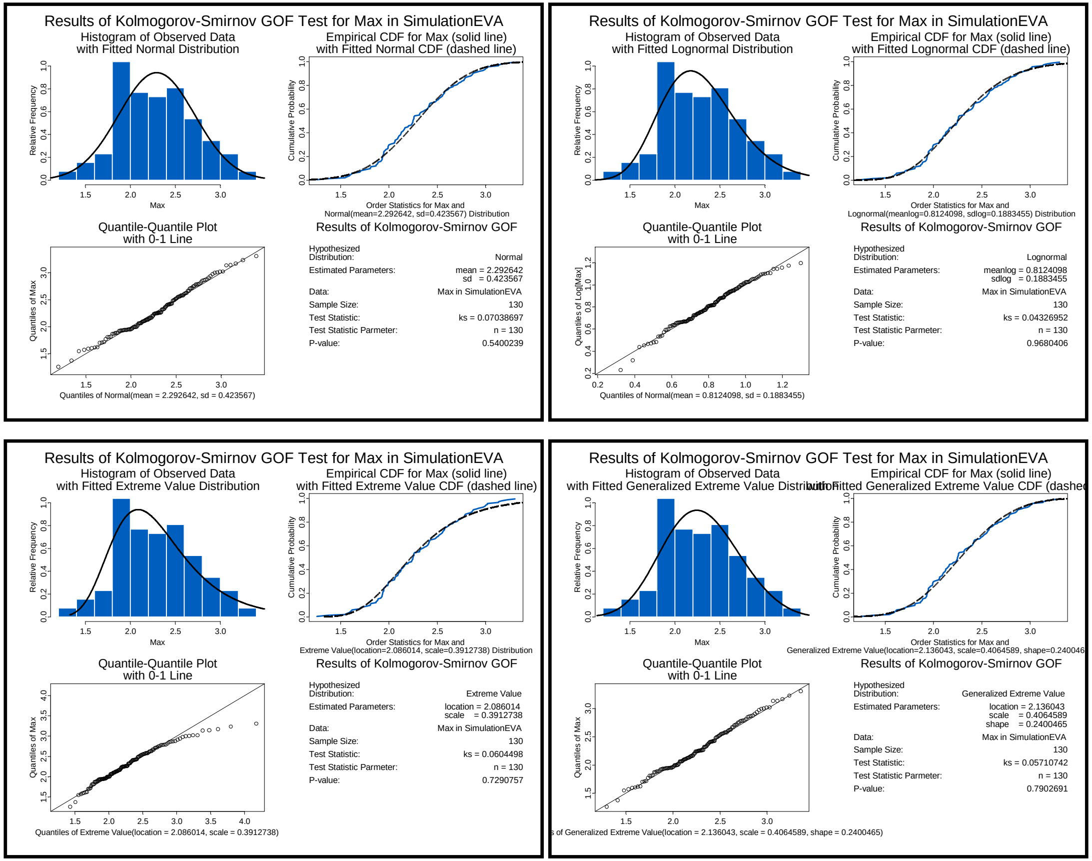

Fig. 5. Kolmogorov-Smirnov graphs and statistics for a simulation example shown in Table 1 of Appendix A (maximum of 60 effective data in a year, and a sample length of 130 yr). All four PDFs have been computed to the same data. Upper left panel: normal distribution; upper right panel: log-normal distribution; lower left panel: Gumbel distribution; lower right panel: GEV distribution.

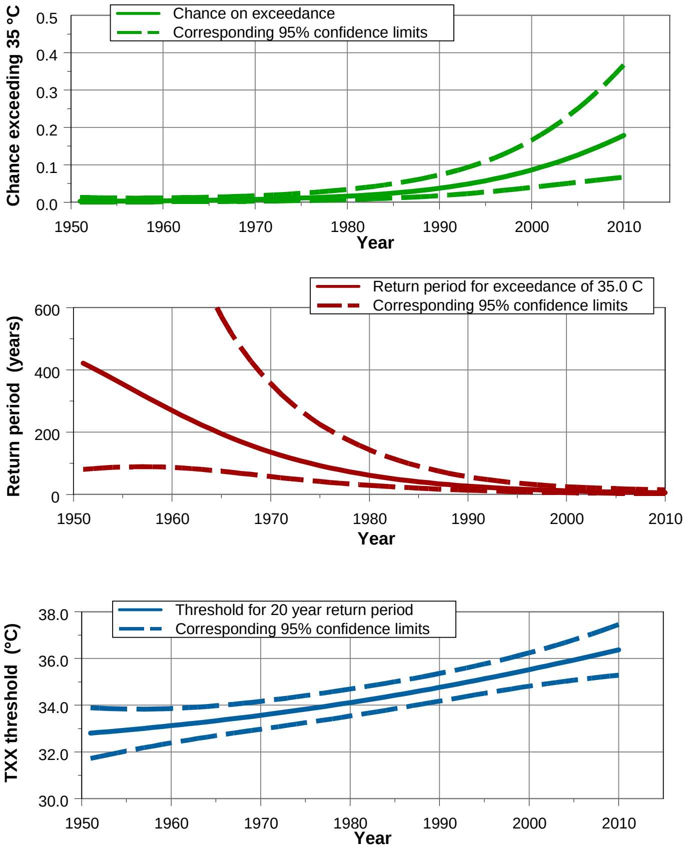

Fig. 6. Three characterizations of extremes with uncertainties corresponding to the TXX_t example from Fig. 1. The first panel shows the annual chance of crossing the 35 °C threshold (p_t^{35}); the second panel shows the corresponding return period for crossing that threshold (R_t^{35}, expressed in years); the third panel shows the temperature threshold that will be exceeded once every 20 yr (TXX_t^{20}, expressed in °C). In all three panels 95 % confidence limits are shown.

CPD
7, 2893–2935, 2011

Weather extremes and weather-related disasters

H. Visser and
A. C. Petersen

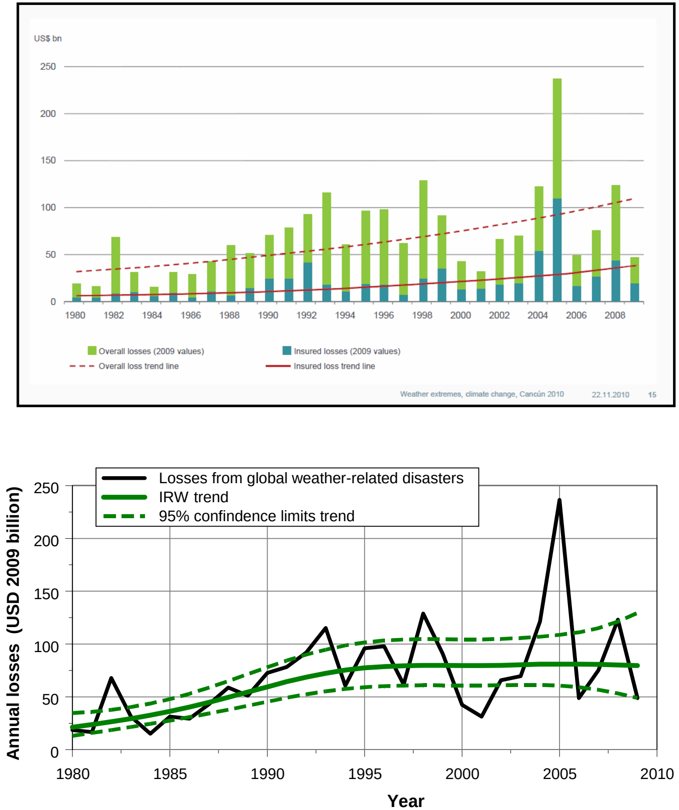

Fig. 7. Economic losses due to weather-related disasters in the period 1980–2009. The data and trend in the upper panel are taken from Munich Re (2010). The trend has been estimated by the OLS straight line fit after taking logarithms. The lower panel shows the IRW trend fit on logarithms of the same data. Flexibility of the trend has been optimized by ML estimation (Visser, 2004).

Is the Flooding in Pakistan a Climate Change Disaster?

Devastating flooding in Pakistan may foreshadow extreme weather to come as a result of global warming

By Nathanial Gronewold and Climatewire | August 18, 2010 | 21

Share Email Print 1 2 3 Next >

UNITED NATIONS -- Devastating flooding that has swamped one-fifth of Pakistan and left millions homeless is likely the worst natural disaster to date attributable to climate change, U.N. officials and climatologists are now openly saying.

Most experts are still cautioning against tying any specific event directly to emissions of greenhouse gases. But scientists at the World Meteorological Organization (WMO) in Geneva say there's no doubt that higher Atlantic Ocean temperatures contributed to the disaster begun late last month.

Image: IMAGE COURTESY OF WIKIMEDIA COMMONS

Fig. 8. Example of coupling climate change to one particular disaster: the 2010 flooding in Pakistan. Text taken from the Scientific American website: http://www.scientificamerican.com/article.cfm?id=is-the-flooding-in-pakist.

CPD
7, 2893–2935, 2011