Testing for Increasing Weather Risk

Weining Wang, Ihtiyor Bobojonov, Wolfgang Karl Härdle, Martin Odening

W. Wang, Institute for Statistics and Econometrics, Humboldt-Universität zu Berlin, Germany

I. Bobojonov, Department of Agricultural Economics, Humboldt-Universität zu Berlin, Germany

W. K. Härdle, Humboldt-Universität zu Berlin and C.A.S.E. - Center for Applied Statistics and Economics, Berlin, Germany

M. Odening, Department of Agricultural Economics, Humboldt-Universität zu Berlin, Germany

Phone: ++49 30 2093 6487

Fax: ++49 30 2093 6465

e-mail: m.odening@agrar.hu-berlin.de

Abstract It is an undisputed fact that weather risk increases over time due to climate change. However, qualification of this statement with regard to the type of weather risk and geographical location is needed. In this paper we use local *t*-tests, change point tests and Mann-Kendall tests to analyze the trends of weather risk indices that are relevant from an agricultural viewpoint. Local test procedures offer more information about the timing and the kind of change in weather risk than global tests do. We also use quantile regression to analyze changes in the tails of weather index distributions. These methods are applied to temperature and rainfall based weather indices in three different climatic zones. Our results show that weather risk follows different patterns depending on the type of risk and the location. We also find differences in the sensitivity of the statistical test procedures.

Keywords: *weather extremes, agricultural risk, change point test, quantile regression*

1 Introduction

Weather is an extremely important production factor for many sectors in an economy. Agriculture, tourism, the energy and the insurance sector are only few examples of weather sensitive businesses. Companies in these sectors are naturally concerned about unfavorable weather conditions and much attention is paid to the development of risk management tools that allow coping with weather perils. The appropriateness and the effectiveness of weather risk management tools, however, crucially depend on the answer to the question: Is weather risk increasing over time or not? The answer to this question seems obvious. It is a more or less undisputed fact that weather risk increases over time due to climate change. Meteorologists and mass media report the occurrence of extremes droughts, floods, heat waves and forest fires in short time intervals. Also a wide range of scientific analyses provide empirical evidence for increasing weather risk (e.g. Beniston and Stephenson 2004; Vasiliades et al. 2009). Despite of the vast empirical literature that already exists in this field we believe that the statistical measurement of weather risk and weather extremes deserves further attention. The main message of this paper is that the increasing-weather-risk hypothesis needs to be qualified in several directions.

First, it is not obvious what weather risk actually means, because there is no single clear definition. The different sectors in an economy are affected in a very specific, complex and often nonlinear manner by extreme weather events. For example, fruit farmers are concerned about the occurrence of frost during the blooming period, but it is of minor importance how cold it is, given that a frost day occurs. That means, even if the temperature becomes more risky in a sense that the temperature distribution has fatter tails, this does not necessarily imply an increased frost risk exposure of the farmer. Moreover, economic risk does not only result from catastrophic or extreme weather events. For example, even a moderate deficit in rainfall during the vegetation period may cause damages to crops and result in severe income losses.

Second, the finding that weather risk in a well-defined sense has increased over time holds only for a specific location where weather data have been recorded and it is not trivial to draw conclusions about the change of weather risk in other locations. In fact, several studies show differences in the trends of weather extremes between geographical regions (e.g. Alexander et al. 2006; Tebaldi et al. 2006).

Third, the time dimension of increasing weather risk needs further investigation. From the viewpoint of predicting future risk exposures it makes an obvious difference, if weather risk increases steadily over time or if the increase is characterized by jumps and discontinuities (Fischer et al. 2012).

Finally, improving the methodology of assessing changes in weather variability is considered as one of the important topics related to risk and uncertainty assessment studies (Hegerl et al. 2007; Tebaldi et al. 2006).

In this paper a battery of statistical test procedures is utilized, among them a local change point test, a local *t*-test and a local Mann-Kendall test. Moreover, upper and lower quantiles of weather indices and their confidence bands are estimated using quantile regression. These techniques are applied to four weather indices, two of them are temperature based (Growing Degree Days (GDD), Frost Days Index (FDI)) and two are rainfall based (Cumulative Rainfall Index (CRI) and Potential Flood Index (PFI)). These four indicators have been used in earlier studies on weather risk exposure of agricultural systems (e.g. World Bank 2005). The objective of this paper is the statistical analysis of weather risk indicators, particularly testing for significant trends and changes over time. Given the limited number of weather indices and locations, that we analyze here, our results are rather illustrative than comprehensive. However, the suggested procedures can be easily applied to other weather events and locations. Our contribution to the existing literature is twofold: First, we apply local instead of global tests. Estimation is done nonparametrically allowing for a flexible estimation of the trends, both in means and in quantiles. Second, we compare alternative tests for trend detection and discuss their sensitivity. The paper is organized as follows: Section 2 provides a brief overview about the relevant literature. Section 3 explains the test procedures followed by a description of the data and the weather indices (section 4). Section 5 presents the results of the various test procedures. The paper ends with conclusions on the statistical measurement of changing weather risk.

2 Review on weather risk measurement

With regard to the measurement of weather risk, two approaches can be distinguished. The first approach captures weather extremes by indices such as hottest and coldest days of the year, warm nights, consecutive dry days and consecutive wet days. The Expert Team on Climate Change Detection and Indices

(ETCCDI) (Klein Tank et al. 2009) has proposed a list of 27 indicators, that is widely accepted and many studies refer to this catalogue (e.g. Alexander et al. 2006; Frich et al. 2002; Tebaldi et al. 2006). A change of weather risk is acknowledged, if the mean values of these indices show a significant increase or decrease over time. There are numerous approaches available for statistical trend detection. The *t*-test and the Mann-Kendall test are among the most commonly used methods to examine if the value of the time series data increases or decreases over time (Liu et al. 2008; Hu et al. 2012). The latter is nonparametric and able to test for a nonzero slope of the mean function as well as for jumps in the mean. In the second approach, one can directly analyze variances or quantiles of the probability distribution of more basic weather variables like daily temperature and daily rainfall and test if the tails of the distributions change over time. For that purpose quantile regression or extreme value theory can be applied. Examples of this kind of analysis are Beniston and Stephenson (2004) and Siliverstovs et al. (2010).

Almost all aforementioned studies test for changes in weather extremes in a fixed time window whose length is determined by the availability of data, i.e. global tests are often conducted. Global tests, however, assume that the error distributions are independent and identically or even normally distributed which is unlikely to be the case for long time series. In fact, global *t*-tests and Mann-Kendall tests are plagued by false detection of trends due to the existence of serial autocorrelations in weather data. This problem as well as solution strategies have been addressed by Noguchi et al. (2011) and Bayazit and Önöz (2009). A further drawback of global tests is that local variations of weather risk may not be detected or precisely located. In other words, global tests are silent about the exact time when a change of weather risk occurred. Such information, however, may be valuable from an economic perspective, for example for the calculation of insurance premia.

In order to cope with the aforementioned problems local test procedures have been developed (Mercurio and Spokoiny 2004; Chen et al. 2010). In particular, Andriyashin et al. (2006) introduce a local test procedure for trend tests, e.g. a *t*-test, a change point test and a Mann-Kendall test. This procedure allows one to decide if there exist trends or jumps based on a sequence of tests in smaller sub-periods. In contrast to global tests such a procedure provides information about the local variations of weather risk, i.e. the type and the timing of risk changes. Another advantage of local tests is the fact that normality and independence are more likely to exist in local windows rather than in long time series. The next section describes the idea of local tests in greater detail.

3 Methodology

3.1 Local tests for the trend in weather variables

In figure 1 we illustrate the motivation of local tests for rainfall risk (measured as monthly rainfall sum in May). The slope of the black line represents the trend over the entire observed period (1948-2008). Visual inspection shows that there is no significant trend. In contrast, dividing the observation period into 4 time windows with a window size of 15 years reveals that there are positive and negative trends in the subintervals. Testing just for a trend over the whole period would disregard this information.

Starting from a small sample size, we increase one observation in each subsequence regression until it covers the whole sample period (Spokoiny 2009; Bertranda 2000; Härdle et al. 2011). This moving window estimation framework, which is often used in financial analysis, adapts to the local stationary feature of many time series data. It is recognized that a nonstationary time series in long term may be locally stationary, without autocorrelations and normally distributed. As an example for this statistical procedure, we demonstrate the localization of the *t*-test. The Mann-Kendall test and the change point test can be localized in a similar fashion. The main idea of the *t*-test is to identify whether the slope of the linear model is significantly different from zero. The linear model has the form:

$$m(j) = \alpha + \beta_j \quad j = 1,\ldots,n, \quad (1)$$

where *n* is the number of observations, $m(j)$ denotes the mean function $E(Y_j)$. Y_j is the temperature or rainfall index for the year *j*, and *α* and *β* are the intercept and slope parameters. These parameters are estimated by minimizing the sum of squares:

$$(\hat{\alpha}, \hat{\beta}) = argmin_{\alpha,\beta} \sum_j \{Y_j - \alpha - \beta_j\}^2 \quad (2)$$

The statistical significance of the slope parameter *β*, i.e. the trend, is tested by the null hypothesis $\beta = 0$, and $\beta \neq 0$ otherwise. Test statistics t^* are defined as:

$$t^* = \frac{\hat{\beta} - 0}{\tilde{\sigma}}, \quad (3)$$

where $\tilde{\sigma}$ is from the residuals of the linear model in Eq. (2), and t^* asymptotically converges to a *t* distribution with n-2 degree of freedom. Instead of assuming linearity of $m(j)$ in Eq. (1) for the whole sample period one can divide the data set into small subintervals (windows) *L* and estimate the parameters α_L and β_L for each window *L*. Under this condition, parameters *α* and *β* may also change with respect to subintervals (windows). The sequence of test results is interpreted by *p*-values, i.e. the probability of test statistic to fall in the rejection region of the null hypothesis. We identify change points only if a sufficient number of consecutive significant signals in the *p*-values sequence occur. How consecutive they should be is basically controlled by aggregation parameters γ and *κ*. γ indicates the minimum number of subsequent *p*-values eligible to create a summation measure and *κ* is the maximum number of insignificant *p*-values to drop the summation to zero. To decide when exactly a change is considered to be significant a threshold value η for the cumulative signals (1-*p*-values) needs to be specified. There is no common agreement on how to select the aggregation parameters γ and *κ* and the threshold value η, although these are very important hyper parameters that affect the outcome of the tests. Following Andriyashin et al. (2006) we set the value of γ equal to 4 and *κ* equal to 2. The threshold value η is set to the half of the maximum accumulated signals (1-*p*-values). From a broader perspective our local test procedure can be viewed as a special type of multiple testing. A multiple testing procedure provides more convincing (robust) test results as it accumulates information from individual tests, but the trade-off is that it induces more hyper parameters. These hyper parameters have a direct link to the conservativeness of the test, but usually there is no fixed rule to select them. It is rather suggested to use scenario-based decisions, namely to choose the hyper parameters that work well for the data set. Our choice of hyper parameters has not returned too many alerts and these alerts can be traced by inspecting the original time series.

We apply the same idea of localization to the change point test and the Mann-Kendall test (cf. Andriyashin et al. 2006). The tests differ in their assumptions. The *t*-test and the change point test require normality (i.e., symmetric and unimodal). In contrast, the Mann-Kendall test does not presume normality, but as a nonparametric test it requires a large sample size. Also, all the three tests assume the independence of the data, they are inappropriate when a serial correlation is present (Noguchi et al. 2011).

3.2 Quantile regression and confidence bands

Quantile regression is one of the widely used techniques in modern statistics (Koenker and Bassett 1978). In contrast to traditional regression, quantile regression looks at different quantiles of the conditional distribution instead of looking at the conditional mean curve. This approach is particularly useful for analyzing the tail behavior of the conditional distribution of climate variables. Considering independent random variables $\{\varepsilon_t\}_{j=1}^{n}$, the model can be presented as:

$$Y_j = l(j) + \varepsilon_j, (4)$$

where the q[th] quantile of the distribuiton of ε_j is 0, and $l(j)$ is the q[th] quantile of Y_j. The nonparametric estimation for a fixed point $i(1 \le i \le n)$ is calculated as:

$$\begin{aligned}\hat{l}(i) = & argmin_{l(i)}(1-q)\sum_{j=1}\left(l(i)-Y_j\right)I\left(l(i)>Y_j\right)w_j \\ & -q\sum_{j=1}\left(Y_j-l(i)\right)I\left(l(i)\le Y_j\right)w_j,\end{aligned} \qquad (5)$$

where $I(.)$ is the indicator function, $w_j = K((j-i)/h)$ with $K(.)$ as a kernel function (e.g. Gaussian Kernel), and h is a bandwidth.

The confidence band estimation follows Härdle and Song (2010) and can be presented as following:

$$\hat{l}(j) \pm (nh)^{-1/2}\{d_n + c(\rho)(2\delta \log n)^{-1/2}\}\hat{\lambda}(j), (6)$$

where ρ is the significance level. d_n, δ, $\hat{\lambda}(j)$ and $c(\rho)$ are constants related to the kernel and the conditional distribution. The confidence band around a nonparametric estimator can be used for testing the functional form.

4 Description of data and specification of weather indices

We apply the tests to weather data at three locations: Taipei (Taiwan), Berlin (Germany) and Mason (Iowa, USA).[1] These weather stations have been selected for two reasons. First, they are located in different agro-ecological environments

[1] Taipei weather data was obtained from Central Weather Bureau of Taiwan (http://www.cwb.gov.tw), Berlin data from German Weather Service (http://www.wetterdienst.de) and Mason data from Wilson et al. (2007). The coordinates of the three weather stations are Lat. 52.466, Long. 13.4 in Berlin, Lat. 25.033, Long. 121.517 in Taipei and Lat. 43.15, Long. -93.199 in Mason.

and thus it can be expected to find different weather patterns over time. According to the Köppen climate classification Berlin has temperate and humid continental climate, Taipei has warm oceanic and humid subtropical climate and Mason has warm and humid continental climate. Second, the quality of the weather records is good in terms of length and completeness of the data. Daily temperature and precipitation records are available for the years 1910-2008 in Taipei, 1948-2010 in Berlin and 1905-2010 in Mason. The data have been checked for missing values, outliers and discontinuities following usual guidelines (e.g. Aguilar et al. 2003). The number of missing values is low for all three stations. 0.9% (2.7%) of temperature data (precipitation data) are missing in Mason. The respective figures for Taipei are 0.09% (9.5%) and the data for Berlin are complete. Linear interpolation and mean substitution have been applied in case of missing temperature data and precipitation data, respectively.

From these data we calculate four indices that can reasonably reflect weather risk exposure in agriculture. The indices are closely related to those proposed by the ETCCDI (Klein Tank et al. 2009), but some modifications were necessary to make the risk indicators more suitable to agriculture.[2] The first index is the Growing Degree Days (GDD) index. The GDD is frequently used to measure the risk of insufficient temperatures during the vegetation season (World Bank 2005):

$$GDD = \sum_{j=b}^{d} \max\{(T_{\max,j} + T_{\min,j})/2 - T_{\text{base}}, 0\}\,, \quad (7)$$

where $T_{\max,j}$ and $T_{\min,j}$ denote maximum and minimum daily temperature and T_{base} represents the temperature that triggers plant growth. We set this parameter to 5°C. *b* and *d* denote the beginning (March 1st) and the end (October 31st) of the vegetation season, respectively. The rationale of the GDD is that plant growth (and hence yields) is proportional to temperature above the threshold T_{base}.

Another temperature related index is the Frost Days Index (FDI):

$$FDI = \sum_{j=b}^{d} \mathbf{1}\{T_j < 0\} \quad (8)$$

where $\mathbf{1}\{.\}$ is the indicator function, which counts the number of days when temperature is below zero. The FDI takes into account the risk associated with frost on the beginning of the vegetation period, e.g. frost during the flowering period for fruit trees or low temperature risk during the germination period of crops.

Rainfall is an important production factor in agriculture. Here we consider the Cumulative Rainfall Index (CRI), which captures drought risk and is defined as:

$$CRI = \sum_{j=k}^{m} R_j\,, \quad (9)$$

Herein R_j denotes daily rainfall, and *k* and *m* are the beginning and end of the accumulation period. We focus on the rainfall in May since precipitation shortfalls are most harmful in this month.

Bad yields, however, may also result from excessive rainfall. In this study we use the Potential Flood Indicator (PFI) to measure this kind of weather risk:

[2] Note that we want to quantify the yield risk of a diversified crop production in general. The analysis of the weather risk exposure of a particular crop would require a more narrow definition of the weather indices.

$$PFI = \max_{\tau \in \{1,\ldots,365-s+1\}} \left(\sum_{j=\tau}^{s+\tau-1} R_j \right), \quad (10)$$

The PFI measures rainfall in the wettest *s*-day-period of the year. Following usual convention we set *s* to 5 days (e.g. Frich et al. 2002; Xu et al. 2010).
Descriptive statistics of these indices for the three locations are presented in Supplement A. To check the precondition of the local *t*-test and the local change point test we carry out normality tests (Anderson Darling and Kolmogorov Sminov) for the indices in each time window. Supplement B displays averaged *p*-values of these tests for all subperiods which suggest that the observations are mostly locally normal with the exception of the FDI in Berlin. Moreover, we tested for serial autocorrelation. ACF tests show that no serial autocorrelation is present during small window sizes which makes the application of local tests suitable for our data[3].

5 Results

In what follows we present the results of the local trend tests (subsection 5.1) and the quantile regression (subsection 5.2). The rationale behind these two views is that the considered weather indices themselves already capture agricultural production risk. That means changes in the mean level of these indices virtually reflect a change in the production risk. This view is adopted by the local trend tests. The quantile regression, however, takes a closer look at the extremes of the indices, i.e. risk in a statistical sense.

5.1 Local trend tests

We demonstrate the application of local *t*-tests, change point tests and Mann-Kendall tests for the case of GDD in Mason (Iowa). Instead of testing for a trend in the whole period, we divide the data set into overlapping windows with a size of 20 years each. The years on the x-axis in figure 2 indicate the beginning of a 20-years-window. The last window, for example, starts in 1990 and covers the time period until 2010. Next, beginning with the first window we calculate test statistics and the according *p*-values for a trend (or a jump) in the sub-periods.
The upper panel in figure 2 depicts the plot of signals (1-*p*-values) of the local tests for consecutive overlapping data windows. A signal value close to one indicates that there exists a trend or a jump inside the window. However, existence of trend is not decided by only looking at one sample window but rather at a group of windows in the neighborhood. We try to make decisions on trends or jumps based on the accumulated significant signals. The cumulated signals are depicted in the lower panel of figure 2.
One starts to sum up signals only if there are more than $\gamma = 4$ consecutive windows with significant signals (the significance level is set to 0.05, which means that significant signals need to be greater than 0.95). The summation stops (the sum is set to zero) if one observes more than $\kappa = 2$ insignificant signals in the subsequent periods.
A trend (or change point) is considered as significant if the cumulated *p*-values in the lower panel of figure 2 cross the threshold values of the respective test, which are represented by horizontal lines. This procedure is conducted in a similar way for all three tests. The change point test, for example, indicates significant jumps

[3] Autocorrelation test results are available from the authors upon request.

in the level of the GDD in Mason during the period 1932-1941, in the year 1965 and from 1978 until 2009. The direction of the change can be identified by means of the sign of the slope parameter β in case of *t*-test and the sign of $\hat{\mu}_1 - \hat{\mu}_2$ in case of the change point test.

The tests are conducted for all indices at all three weather stations with the exception of the FDI in Taipei, since no frost occurs in this region. Table 1 summarizes the results. The column “From” reports the beginning and the column “To” shows the end of the time period, in which a change of a weather risk indicator appears to be significant. To be precise, “From” indicates the beginning of the first 20-years-window showing a significant change in the mean of the risk indicator and “To” means the beginning of the last 20-years-window, until which this change continues. Column “Sign” displays the direction of the change, i.e. an increase or decrease, and column “Test” reports the type test which has indicated the change.

Apparently, the detected changes differ in their duration, as well as in their direction. In some cases there are only a few consecutive windows where a change appears to be significant while other changes are more persistent. In particular, we find a clear positive trend for the GDD in Berlin detected by all tests during different periods. This finding coincides with the results of other studies which also indicate increasing trends of temperature indices in this region (e.g. Kürbis et al. 2009). Taipei has also experienced an increasing level of GDD except one short period showing a negative trend. Similar results are also discussed in EPA (2009).

Summarizing the trend in Mason is more complex since positive and negative changes occur during the observation period. In contrast to Taipei, the phases with negative trend dominate during the observed period indicating a declining plant growth potential in this area. Declining temperatures in summer periods in Iowa were also reported by the ICCIC (2011). Similar to the ICCIC study we find a positive trend of the FDI in Mason that indicates increasing frost risk in this area during the spring period. No such tendency can be found for Berlin.

According to the change point test there is a clear positive jump in the CRI in Berlin between the windows starting 1956 and 1963 which shows a reduced risk of water shortage in the vegetation period compared with the first half of the past century. A significant positive jump of the CRI is also pointed up for Taipei. In contrast, several positive and one negative changes of the CRI occur in Mason. It is difficult to make a clear statement on the risk of excessive rainfall when there are positive and negative trends during the whole observed period. However, we can summarize the general trend as positive in Mason since most of the trends are positive except the time window starting in 1939. Thus the results of the local tests are similar to the finding of the ICCIC (2011) stating that precipitation in Iowa increased.

With regard to the PFI we find several positive jumps between 1928 and 1971 in Mason. This increase of risk, however, is partially compensated by a negative change of the PFI in the late seventies. No change in the risk of excessive rainfall was detected for Berlin.

The results show that the four risk indicators follow different trends. Moreover, their development is region specific. From table 1 it is also apparent that the sensitivity of the test procedures differs. The change point test detects changes more often than the *t*-test and the Mann-Kendall test. Actually, all changes of the CRI and the PFI are solely reported by the change point test and not confirmed by any other test.

It is mentioned in Andriyashin et al. (2006) that the change point tests showed the best performance on average in their case. In our analysis, the change point tests revealed most alerts of jumps, especially for rainfall indices. This may be partly due to the nature of the rainfall time series, in which one observes more sudden jumps instead of smooth trends. Another reason may be that the local normality of our time series gives more power to the change point tests than the Mann-Kendal tests.

5.2 Quantile regressions

We now turn to the results of the quantile regression, which complement the information provided in the previous sections. Instead of considering changes in the mean of weather (risk) indicators we now focus on changes in the tails of their distribution function. Figure 3a illustrates the results for the CRI in Taipei and figure 3b shows the results for the CRI in Mason and figure 3c for Berlin. We analyze the lower quantiles of the CRI in order to examine the weather risk exposure in agriculture associated with droughts.

The confidence bands can be employed for testing the significance of changes of the quantile over time. This can be done by checking whether a linear line with positive or negative slope fits into the corridor. For example, inspection of figure 3b shows that the 10% quantile (i.e. the downside risk of rainfall) of the CRI exhibits a positive trend in Mason since 1990, which indicates a declining probability of rainfall shortfalls in May. In contrast, we observe a decline of the 10% quantile of the CRI in Berlin during the seventies and eighties as well as in the last decade. A negative trend for Taipei can be detected in the first half of the previous century. The difference in trends between the cities is not surprising as many existing studies already found opposite trends in rainfall indices in different areas of the world (e.g. Alexander et al. 2006; Frich et al. 2002). However, it is interesting to see that the 10% quantile of the CRI in Berlin declines while a positive change in the mean value of this indicator has been detected by local tests in the previous section. One should recall, however, that the subjects of the two statistical procedures are not the same. The slope tests consider the mean values of CRI whereas the quantile regression looks at lower left tail of the CRI distribution. Both results can be interpreted such that the rainfall availability in May increased in Berlin, but the probability of having severe drought in some years increased at the same time.

All in all we find mixed evidence for the increasing-weather-risk-hypothesis. The 10% quantile of the GDD exhibits a positive trend in Taipei and no change can be observed in the GDD in Mason. We neither found evidence for increasing frost risk.

There is a significant increase in the 90% quantile of the PFI in Mason, which confirms the results of the local mean test. This trend, however, does not exist in the other two cities. Similar to the case of Berlin, the 10% quantile of the CRI tends to decline in Taipei when local tests detected some positive trend. This can be interpreted as an increasing exposure to drought risk at least in Berlin, since the level of rainfall is rather low in that region.

6 Conclusions

We have used a battery of statistical test procedures to analyze the presence of trends in weather indices, which are relevant to agriculture and capture parts of the production risk in this sector. Local versions of the Mann-Kendall test, *t*-test as well as change point test have been employed to analyze the trends of weather indices. Moreover, we used quantile regression techniques to analyze changes in extreme values of these weather indices. Three weather stations located in Europe, Asia and North America have been selected to investigate the dynamics of weather risk indicators in different climatic zones.

The results confirm our conjecture that a general trend of increasing weather risk cannot be identified. The analysis reveals that different weather risk indicators show different pattern over time. Moreover, the changes of the weather indices are also location specific. For example, we reveal declining temperature related risk (e.g. GDD) in Taipei and Berlin but increasing temperature risk in Mason. Another feature of weather risk is that indices rarely follow long lasting trends but exhibit local jumps, sometimes in opposite direction. We also found that the tests used in the local form have different sensitivity in detecting the changes. To be specific, the change point test and the *t*-test turned out to be more sensitive than the Mann-Kendall test. This may be explained by the finding of Yue et al. (2002) who report that *t*-test and the change point test are likely to outperform the Mann-Kendall test for small samples of normal data, while the Mann-Kendall test is better for nonnormal data.

One advantages of using tests in a local form is the possibility of examining when exactly changes in a weather risk indicator occurred. This is particularly advantageous for sectors which are interested in short term trends, e.g. insurance. This informational gain, however, comes at the cost of a more complicated test procedure. Additional parameters that have to be specified are the sample window size, the minimum number of consecutive significant windows and a threshold level for cumulated *p*-values. The determination of these parameters introduces subjective elements into local tests since there are only rules of thumb for their specification. Furthermore, applications of local tests can help to circumvent the non-normality in long time series by looking at a reasonably smaller window.

The detected changes in local tests may have similar directions to those of the quantile regression, but not necessarily. In some cases, trends detected by local tests and changes in the quantiles have opposite directions, for example in case of the CRI in Berlin. Which perspective on risk is more reasonable depends on the specific context. An agricultural producer is likely affected by changes in the mean of the weather indices, while an insurer is more concerned about changes in their extremes.

As a general conclusion we recommend to be specific when stating that weather risk increases under climate change. This refers to the type of risk, geographical location, time and duration of risk changes as well as the test procedure.

We consider the tests presented in this paper as a complement to the existing statistical toolbox that targets at the detection of changes in weather risk. However, we do not claim that the presented test procedures are superior to others and the analysis presented here can be extended in several directions. It would be worth to apply the test procedures to other regions and other weather events. This would allow for a generalization of the results reported here. Moreover, the power and the robustness of the statistical test procedures deserve further attention. Simulation experiments could be helpful to find out, which kind of test is most appropriate for identifying changes of specific weather risks.

References

Adamowski K, Bougadis J (2003) Detection of trends in annual extreme rainfall. Hydrological Process 17:3547–3560

Aguilar E, Auer I, Brunet M, Peterson TC, Wieringa J (2003) Guidelines on Climate Metadata and Homogenization, WCDMP 53, WMO-TD 1186, 55 pp World Meteorol Org Geneva

Alexander LV, Zhang X, Peterson TC, Caesar J, Gleason B, Klein Tank A, Haylock M, Collins D, Trewin B, Rahimzadeh F, Tagipour A, Ambenje P, Kumar KR, Revadekar J, Griffiths G, Vincent L, Stephenson DB, Burn J, Aguliar E, Brunet M, Taylor M, New M, Zhai P, Rusticucci M, Vazquez-Aguirre JL (2006) Global observed changes in daily climate extremes of temperature and precipitation. J Geophysical Res 111

Andriyashin A, Benko M, Härdle W, Timofeev R, Ziegenhagen U (2006) Color harmonization in car manufacturing processes, Applied Stoch. Models in Business and Industry., 22:519–532

Bayazit M, Önöz B (2007) To prewhiten or not to prewhiten in trend analysis? Hydrological Sciences J. 52(4)

Beniston M, Stephenson DB (2004) Extreme climatic events and their evolution under changing climatic conditions. Glob. Planet. Change 44:1-9

Bertranda P (2000) A local method for estimating change points: the “Hat-function”. Statistics: J. Theoretical and Appl. Statistics 34(3)

Chen Y, Härdle WK, Pigorsch U (2010) Localized Realized Volatility Modeling. J. Am. Stat. Association 105(492):1376-1393

Enviromental Protection Administration (EPA) 2009 Extreme events and disasters are biggest threat to Taiwan. Typhoon Morakot. Executive Yuan, R.O.C Taiwan.. http://www.epa.gov.tw

Fischer T, Gemmer·M, Liu L, Su B (2012) Change-points in climate extremes in the Zhujiang River Basin, South China, 1961–2007. Climatic Change 110:783–799

Frich P, Alexander L, Della-Marta P, Gleason B, Haylock M, Klein Tank A, Peterson T (2002) Observed coherent changes in climatic extremes during the second half of the twentieth century, Climate Research 19:193–212

Härdle W, Song S (2010) Confidence bands in quantile regression, Econometric Theory 26:1180–1200

Härdle W. (1990) Applied Nonparametric Regression. Camb. University Press

Härdle WK, Cabrera BL, Okhrin O, Wang W (2011) Localising temperature Risk. SFB 649 Discussion Paper 2011-001

Hegerl GC, Zwiers F, Kharin S, Stott P (2007) Detectability of anthropogenic changes in temperature and precipitation extremes, J. Climate. 17:3683–3700

Hu·Y, Maskey S, Uhlenbrook S (2012) Trends in temperature and rainfall extremes in the Yellow River source region, China. Climatic Change 110:403–429

Iowa Climate Change Impacts Committee (ICCIC) (2011 Climate change impacts on Iowa 2010. Report to the Governor and the Iowa General Assembly. www.energy.iowa.gov

Koenker R, Bassett G (1978) Regression quantiles, Econometrica 46(1):33–50

Kürbis K, Mudelsee M, Tetzlaff G, Brázdil R (2009) Trends in extremes of temperature, dew point, and precipitation from long instrumental series from central Europe. Theory Applied Climatology 98:187–195

Liu X, Xu Z, Yu R (2011) Trend of climate variability in China during the past decades. Climatic Change 109:503–516. DOI: 10.1007/s10584-011-0097-6

Mercurio D, Spokoiny V (2004) Statistical inference for time-inhomogeneous volatility models. Annals of Statistics 32:577–602

Noguchi K, Gel YR, Duguay CR (2011) Bootstrap-based tests for trends in hydrological time series, with application to ice phenology data. J.Hydrology 410(3-4):150-161

Siliverstovs B, Rainald Ö, Kemfert C, Jaeger C, Armin H, Kremers H (2010) Climate change and modelling of extreme temperatures in Switzerland, Stochastic Environ. Res Risk Assessment 24:311–326

Spokoiny V (2009) Multiscale local change point detection with applications to value at risk. The Annals of Statistics 37(3):1405-1436

Klein Tank AMG, Zwiers FW, Zhang X (2009) Guidelines on Analysis of extremes in a changing climate in support of informed decisions for adaptation. Climate Data and Monitoring WCDMP 72. World Meteorological Organization

Tebaldi C, Hayhoe K, Arblaster JM, Meehl GA (2006) Going to the extremes: An intercomparison of model-simulated historical and future changes in extreme events, Climatic Change 79(3–4):185–211

Vasiliades L, Loukas A, Patsonas G (2009) Evaluation of a statistical downscaling procedure for the estimation of climate change impacts on droughts. Nat. Hazards Earth Syst. Science 9:879–894

Wang XL, Swail VR (2001) Changes of extreme wave heights in northern hemisphere oceans and related atmospheric circulation regimes, J Climate 14:2204–2220

Wilson LT, Yang Y, Lu P, Wang J, Nielsen-Gammon JW, Smith N, Fernandez CJ (2007) Integrated Agricultural Information and Management System (iAIMS): World Climatic Data.

http://beaumont.tamu.edu/ClimaticData/
World Bank (2005) Managing agricultural production risk: Innovations in developing countries, www.worldbank.org/rural
Xu W, Filler G, Odening M, Okhrin O (2010) On the systemic nature of weather risk. Agricultural Finance Rev. 70(2):267-284
Yue S, Pilon P, Caradias G (2002) Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. J. Hydrology. 259:254–271

Figure

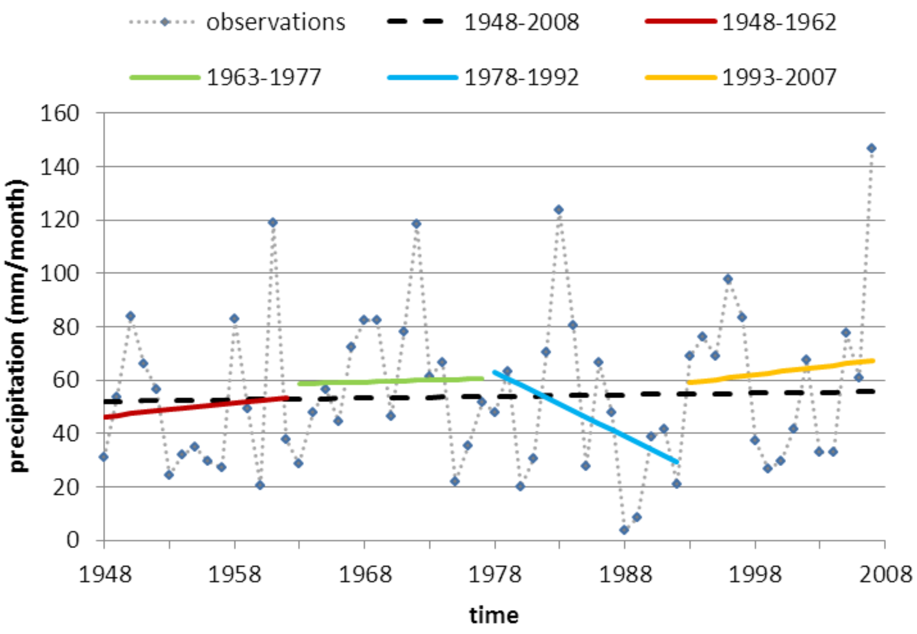


Figure 1: Local versus global test of changes in monthly rainfall (Berlin)

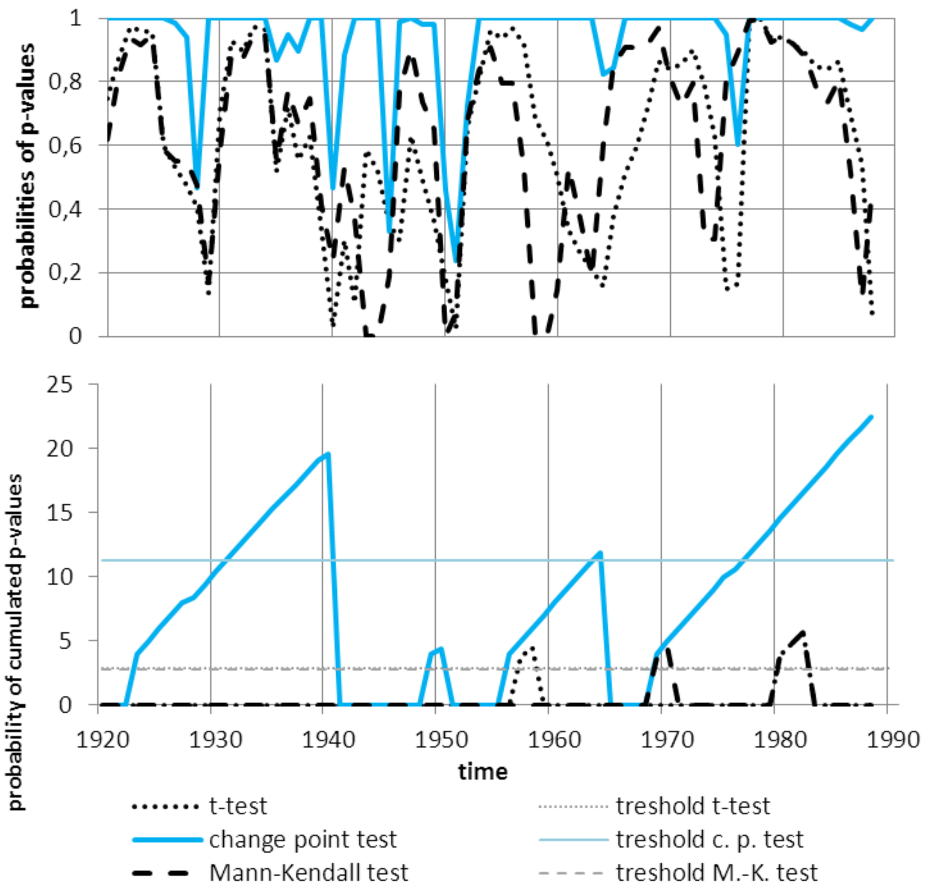
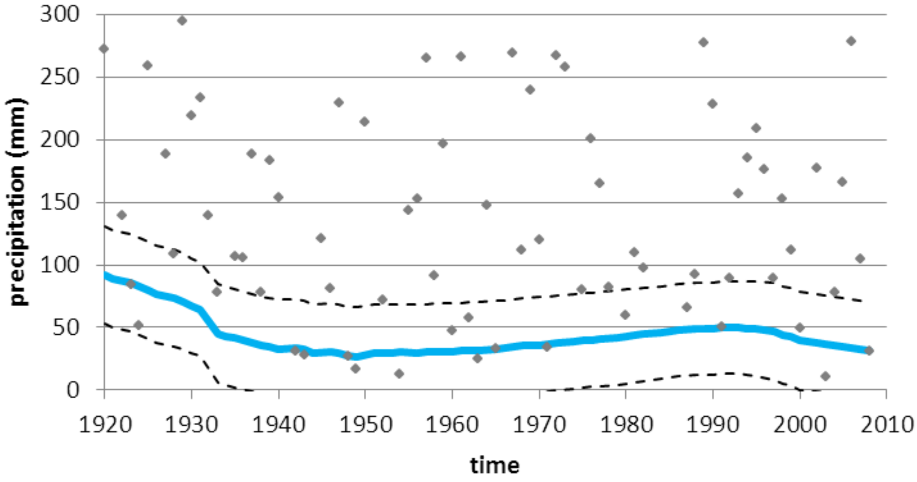
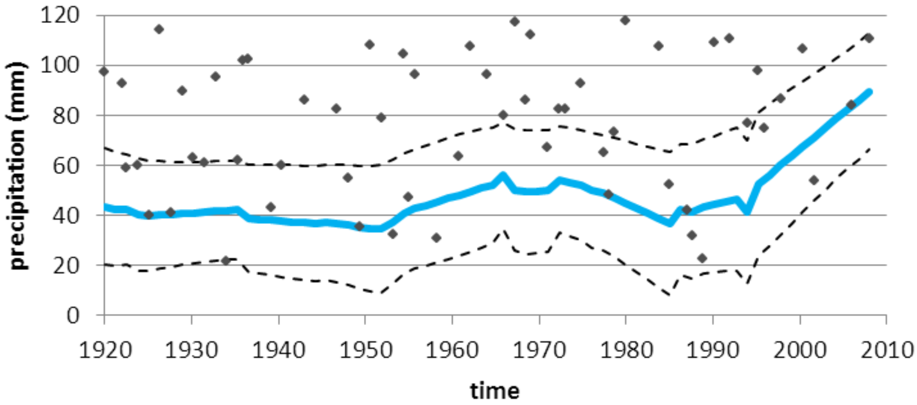


Figure 2: Local test results for GDD in Mason

a) Taipei



b) Mason



c) Berlin

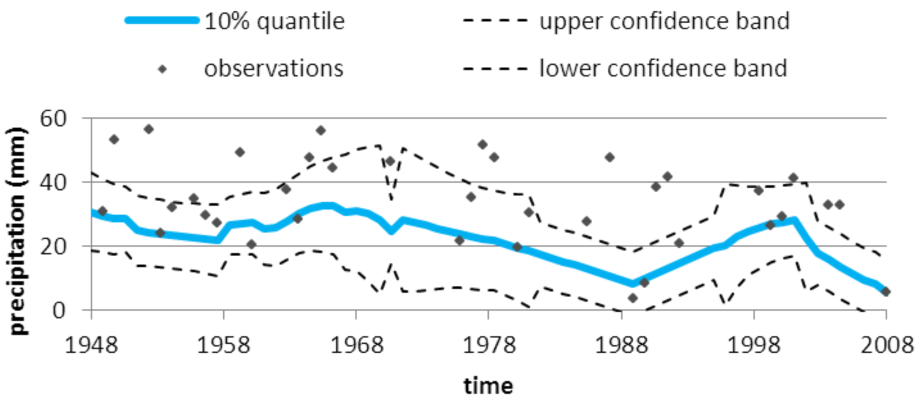


Figure 3: 10% quantiles of CRI with confidence bands

Table 1: Trends of indices

Indices	GDD				FDI				CRI				PFI			
Cities	From	To	Sign	Test	From	To	Sign	Test	From	To	Sign	Test	From	To	Sign	Test
Taipei	1939	1949	+	C,T,M	n.a.	n.a.	n.a.	n.a.	1953	1966	+	C	1949	1959	-	C
	1959	1960	-	C									1960	1965	+	C
	1981	2008	+	C												
Mason	1932	1941	-	C	1986	1987	+	T	1928	1933	+	C	1928	1929	+	C
	1958	1959	-	T					1939	1939	-	C	1961	1964	+	C
	1965	1965	+	C					1948	1950	+	C	1968	1971	+	C
	1970	1971	+	M					1988	1989	+	C	1977	1980	-	C
	1978	2009	-	C												
	1981	1983	-	M,T												
Berlin	1977	1982	+	T	none	none	none	none	1956	1958	+	C	none	none	none	none
	1979	2009	+	C					1963	1966	+	C				
	1988	1989	+	M,T												

C = change point test, M = Mann-Kendall Test, T = *t*-test

Supplementary Material

Click here to download Supplementary Material: supplement.pdf