Detection Probability of Trends in Rare Events: Theory and Application to Heavy Precipitation in the Alpine Region

CHRISTOPH FREI AND CHRISTOPH SCHÄR

Climate Research, ETH, Zurich, Switzerland

(Manuscript received 13 September 1999, in final form 21 July 2000)

ABSTRACT

A statistical framework is presented for the assessment of climatological trends in the frequency of rare and extreme weather events. The methodology applies to long-term records of event counts and is based on the stochastic concept of binomial distributed counts. It embraces logistic regression for trend estimation and testing, and includes a quantification of the potential/limitation to discriminate a trend from the stochastic fluctuations in a record. This potential is expressed in terms of a detection probability, which is calculated from Monte Carlo–simulated surrogate records, and determined as a function of the record length, the magnitude of the trend and the average return period (i.e., the rarity) of events.

Calculations of the detection probability for daily events reveal a strong sensitivity upon the rarity of events: in a 100-yr record of seasonal counts, a frequency change by a factor of 1.5 can be detected with a probability of 0.6 for events with an average return period of 30 days; however, this value drops to 0.2 for events with a return period of 100 days. For moderately rare events the detection probability decreases rapidly with shorter record length, but it does not significantly increase with longer record length when very rare events are considered. The results demonstrate the difficulty to determine trends of very rare events, underpin the need for long period data for trend analyses, and point toward a careful interpretation of statistically nonsignificant trend results.

The statistical method is applied to examine seasonal trends of heavy daily precipitation at 113 rain gauge stations in the Alpine region of Switzerland (1901–94). For intense events (return period: 30 days) a statistically significant frequency increase was found in winter and autumn for a high number of stations. For strong precipitation events (return period larger than 100 days), trends are mostly statistically nonsignificant, which does not necessarily imply the absence of a trend.

1. Introduction

Analyses of trends in observational climate records provide insights into the characteristics and magnitude of global and regional climatic variations during the last decades to centuries. Beside the consideration of mean climate parameters such as seasonal mean temperature and precipitation, considerable interest has emerged in long-term trends of the occurrence and severity of rare and extreme weather phenomena. Recent trend analyses of extremes include tropical storms (e.g., Landsea et al. 1999), deep extratropical cyclones and wind storms (e.g., Schmidt and von Storch 1993; Schiesser et al. 1997; Bijl et al. 1999; Alexandersson et al. 2000), extreme temperature and frost (e.g., DeGaetano 1996; Stone et al. 1996), heavy precipitation (e.g., Auer and Böhm 1996; Karl and Knight 1998; Suppiah and Hennessy 1998), high intensity streamflow (e.g., Changnon and Demissie 1996), and droughts (e.g., Szinell et al. 1998; Rebetez 1999). Current interest in trends of extreme weather phenomena relates to their potential for severe and adverse impacts on human life, civil infrastructure, and natural ecosystems with far-reaching socioeconomic consequences. Attention into trends of extremes is also fostered by concerns that the climate may become regionally more variable or extreme as a result of global climate change (e.g., Fowler and Hennessy 1995; Kattenberg et al. 1996; Frei et al. 1998; Easterling et al. 2000).

The determination of long-term trends from observational records is confronted with the fundamental difficulty that stochastic variations in a climate record limit the accuracy with which a long-term trend can be estimated. Two types of misinterpretations can arise: first, the estimate is erroneously taken as the signal of a long-term trend, but its value is essentially controlled by the stochastic variations in the record (error of type I). Second, an existing long-term trend in the climate parameter under consideration is not identified because it is overshadowed by short-term, stochastic fluctuations (error of type II). Stochastic variations exert a limitation on the detectability of a trend from a finite record. While these errors pertain to trend analyses of all kinds of cli-

Corresponding author address: Christoph Frei, Climate Research ETH, Winterthurerstr. 190, CH-8057 Zürich, Switzerland.
E-mail: christoph.frei@geo.umnw.ethz.ch

mate parameters (including means), trend assessments for extremes and rare events are particularly exposed to these statistical complications.

Errors of type I are controlled by means of a statistical test, limiting the probability of erroneous trend detection at a prespecified level of significance. Conventional statistical tests are, however, often inappropriate for testing trends in rare weather events. For example the parametric Student's *t*-test assumes that trend residuals are normally distributed, whereas data from extreme events, such as event frequencies, means above thresholds or quantiles are usually skewed and deviate significantly from normality. Again, frequently used nonparametric methods based on ranks, such as Spearman rank correlation and Mann–Kendall tests (Mann 1945; Kendall 1955; Sneyers 1990), are not applicable when a large number of ties impedes the unequivocal assignment of ranks (see also Lettenmaier et al. 1994). This latter restriction is relevant for extreme event frequencies as this data has a discrete distribution. Various methods have been proposed to accommodate for the specific nature of rare event data in trend tests: following the concept of a stochastic Poisson process, Keim and Cruise (1998) avoid the problem of discrete event counts by considering interarrival times of heavy rainfall events. A stochastic simulation approach using a low-order autoregressive moving average model has been chosen by Karl et al. (1996) to analyze trends in indices of climate extremes. A related methodology using generalized extreme value distributions has been adopted by Zwiers and Kharin (1998) to assess changes in climate extremes between GCM control and $2 \times CO_2$ experiments. Again, DeGaetano (1996) applies a randomization procedure to test for trends in annual counts of temperature threshold exceedence.

While appropriate statistical tests are used to control type I errors, less attention is usually given to type II errors. Quantitative information on the ability to discriminate (by means of statistical inference) an eventual trend from stochastic variations is, however, important for the interpretation of the test results. The formal acceptance of the null hypothesis ''no trend'' does not necessarily imply the absence of a trend. Rather it is the confidence interval of the trend estimate that gives the range of statistically compatible trend magnitudes. For a linear regression trend analyses, the confidence interval increases with increasing residual variance and decreasing record length (see statistics textbooks, e.g., Rice 1995), which implies that weak trends are more difficult to discern (statistically significant) from short-term records with a large ''noise'' component. For rare weather events it is intuitive that the accuracy of trend estimates will also be influenced from the rarity of the events, with confidence intervals increasing with decreasing sample size. This would imply stronger limitations in trend detectability with increasing rarity of events. Therefore a quantitative consideration of type II errors appears particularly relevant for the assessment of trends in extreme events.

In this study we present a statistical framework for the assessment of trends in rare weather events. The framework embraces both a trend test (consideration of type I errors) as well as a quantification of the potential/limitation of trend detectability (type II errors). The methodology is based on the binomial process, which is adopted as a simple stochastic model for annual and seasonal counts of rare events. Trend estimation and testing is conducted using logistic linear regression. The potential of trend detectability, represented in the form of a detection probability, is quantified as a function of record length, trend magnitude, and event rarity. This is accomplished using the binomial process in Monte Carlo simulations of surrogate, ''trendy'' records. Advantages of this parametric concept are that: it is adapted to the statistical nature of event frequencies (counts); it is naturally linked to a regression model including an appropriate statistical test; and it is simple enough to allow for a quantification of type II errors. Yet this concept is restricted to the study of event count records and other methods are required for extreme intensities or sequences of extreme quantiles.

Our study is also motivated by GCM simulations of an increase in greenhouse gas concentrations. These have revealed an increase of mean precipitation at high and middle latitudes (Kattenberg et al. 1996). Results suggest that the simulated changes in means are associated with and mainly affected by shifts toward more frequent heavy precipitation events (e.g., Whetton et al. 1993; Hennessy et al. 1997; Zwiers and Kharin 1998). An increase of heavy precipitation events was also found in simulations with regional climate models for winter and autumn over Europe (Schär et al. 1996; Jones et al. 1997; Frei et al. 1998). The model results have fostered a number of trend studies, which report on increasing heavy precipitation occurrence/intensity for a number of regions on the globe (Iwashima and Yamamoto 1993; Rakhecha and Soman 1994; Karl et al 1995; Suppiah and Hennessy 1998; Forland 1998; Groisman et al. 1999; Osborn et al. 2000).

As an application of the trend assessment methodology developed in this study we will thus examine trends of heavy daily precipitation from centennial rain gauge records in the region of the European Alps (Switzerland). In central Europe an increasing trend of mean wintertime precipitation is observed during this century (Schönwiese et al. 1994). Regionally in the Alpine mountain range this increase amounts to 15%–20% and is statistically significant, but no significant trend is found for other seasons of the year (Widmann and Schär 1997). One purpose of this trend study is to investigate whether long-term changes in heavy events have contributed to the increase of the seasonal mean, in line with GCM-based expectations for global warming. Also it is interesting to learn about frequency changes, which do not reflect in the seasonal means. The database for the trend analysis is similar to that used by Widmann and Schär (1997). It comprises 113 Swiss rain gauge stations with records that cover the period 1901–94 with daily resolution.

Section 2 introduces the statistical framework of binomial event counts, logistic regression, and the notion of detection probability. Theoretical calculations of the detection probability are presented in section 3, and results from the application of the trend analysis to heavy precipitation in the Alpine region follow in section 4. The paper closes with a summary and pertinent conclusions in section 5.

2. Statistical method

The concept of a binomial random process can be used for statistical modeling of the temporal variations of weather events. This concept is adapted to deal with count data exclusively. Hence its adoption relies on a prior definition of ''events,'' for example, the exceedence of some threshold, and the subsequent counting of events over suitable pools (e.g., a season). Data most common for this methodology are time series of seasonal or annual counts of events, which hereafter are termed ''count records'' for brevity.

The binomial concept considers the count n of events at a particular time (e.g., the number of heavy daily precipitation in a particular summer) as the realization of a binomial distributed random variable. Recall that the binomial distribution describes a random process consisting of m independent trials (e.g., number of summer days) with probability π for a ''successful'' trial (e.g., threshold exceedence). The binomial distribution function (i.e., the probability for n events) is given by

$$B(n;\, \pi, m) = \binom{m}{n} \cdot \pi^n \cdot (1 - \pi)^{m-n} \tag{1a}$$

with

$$\binom{m}{n} \equiv \frac{m!}{n! \cdot (m-n)!}.$$

The expected value $\langle n \rangle$ and variance var(n) of the distribution are (see, e.g., Rice 1995):

$$\langle n \rangle = m \cdot \pi, \qquad \mathrm{var}(n) = m \cdot \pi \cdot (1 - \pi). \tag{1b}$$

Strictly, a binomial process consists of independent trials with constant probability π. This presumption may be in conflict with the reality of rare event occurrences for two related reasons. First, the day-to-day variation of weather conditions as well as the seasonal variation of extremes are associated with variations of the probability of rare events. Second, the persistence of weather systems can reflect in autocorrelation in the data, and hence a dependence of the event probability on the outcome of past trials. Both of these effects (or a combination thereof) do result in an increase of the variance in the data above the nominal binomial variance in (1b). [Overdispersion as a result of variations in π across trials is discussed in detail by Stigler (1986), and McCullagh and Nelder (1989); overdispersion as a result of autocorrelation is discussed in Liang and Zeger (1986) and Diggle et al. (1994).] The excess variance can be expressed with the help of an overdispersion factor σ^2, representing the ratio of effective variance in the data to binomial variance, that is,

$$\mathrm{var}(n) = \sigma^2 \cdot m \cdot \pi \cdot (1 - \pi). \tag{2}$$

Overdispersion arises in most practical applications of binomial models (McCullagh and Nelder 1989). For the statistical testing of trends in event records, we will take account of the overdispersion present in the observational data. However, this factor is not considered in the estimation of trend detection probabilities (see section 2b), implying an optimistic estimation of this quantity.

a. *Trend estimation and testing*

The statistical model of logistic linear regression is used in this study for estimating and testing long-term trends in count records. Logistic regression is a special case of a formal generalization of linear regression concepts commonly summarized under the term *generalized linear models.* These statistical models apply for a broad class of theoretical distributions, *the exponential distribution family,* which embraces the binomial distribution. In the following we briefly summarize logistic regression in its application for trend modeling of rare events. A detailed description of the theory of *generalized linear models* and the logistic case can be found in McCullagh and Nelder (1989) and Dobson (1990).

The logistic trend model expresses a transformed form of the expected value of counts (or equivalently the event probability π) as a linear function of time:

$$\eta(\pi) = \alpha + \beta \cdot t. \tag{3}$$

Here t is time (the year of the period); α, β are the regression intercept and coefficient, respectively, to be estimated from the data; and η is a prescribed monotonic link function. The essence of the link function is to provide a transformation from the value range of $\pi \in [0, 1]$ onto the real axis, to ensure compatibility with the linear model on the rhs of Eq. (3). In principle various link functions can be appropriate. Here we have chosen the canonical link of the logistic regression model:

$$\eta(x) = \mathrm{logit}(x) \equiv \log\left(\frac{x}{1 - x}\right). \tag{4}$$

As a result, the temporal variations of the expected value of events take the form:

$$\pi(t;\, \alpha, \beta) = \exp(\alpha + \beta \cdot t)/[1 + \exp(\alpha + \beta \cdot t)]. \tag{5}$$

The magnitude of the trend, as given by model parameter β, is conveniently expressed as the odds ratio Θ, defined as

$$\Theta \equiv \frac{\pi(t_2)}{1 - \pi(t_2)} \bigg/ \frac{\pi(t_1)}{1 - \pi(t_1)} = \exp[\beta \cdot (t_2 - t_1)]. \quad (6)$$

The odds ratio represents the relative change in the ratio of the events against nonevents during the period (t_1, t_2) and is an exponential function of the period length. In the case of rare events ($\pi \ll 1$), Eq. (5) approximates an exponential trend and the odds ratio represents the fractional change of the rare event probability from the beginning to the end of the period [$\Theta \approx \pi(t_2)/\pi(t_1)$]. In our application of the logistic trend model in section 4, the trend magnitude, determined from the 94-yr time series, will be expressed as a centennial odds ratio Θ_{100} {i.e., the change over a 100-yr period [see Eq. (6)]}:

$$\Theta_{100} = \exp(\beta \cdot 100), \quad (7)$$

for better comparability with the theoretical results of section 3.

The parameters α and β of the logistic trend model are estimated according to the maximum likelihood principle by fitting the logistic regression model to observed event records. We use an iterative algorithm, known as Fisher scoring (see McCullagh and Nelder 1989), using the implementation in the S-plus statistics package (Venables and Ripley 1997). The statistical significance of the estimated trend parameter can be inferred from the *p*-value testing against the null hypothesis ($\beta = 0$). (The *p* values represent the probability of accepting the null hypothesis.) In our applications *p* values were obtained by using the deviance difference between the trend model and null hypothesis as the test statistic (McCullagh and Nelder 1989). Test results will be given for a two-tailed test with a significance level of 5%.

Following the procedure proposed by McCullagh and Nelder (1989) some additional scaling was made to the test statistic by the overdispersion σ^2, with the latter estimated from the residual deviance. The compensation for overdispersion ensures that some important deviations of the observed data from the strict binomial model can be accounted for in the trend tests. For instance, the occurrence of heavy precipitation during synoptically controlled seasons is likely to be subject to serial correlation from 1 day to the next. This contributes to overdispersion in the trend model, that is, larger residual variance than expected from strictly binomial data, and it broadens the distribution function of the test statistic under the null hypothesis. Hence the scaling with overdispersion intrinsically compensates for serial correlation in the event processes, without explicit statistical modeling. (An alternative approach for the compensation of autocorrelation-related overdispersion in confidence intervals was proposed by Liang and Zeger (1986) and Diggle et al. (1994) and has shown results very similar to the approach by McCullagh and Nelder (1989) for our application in section 4.)

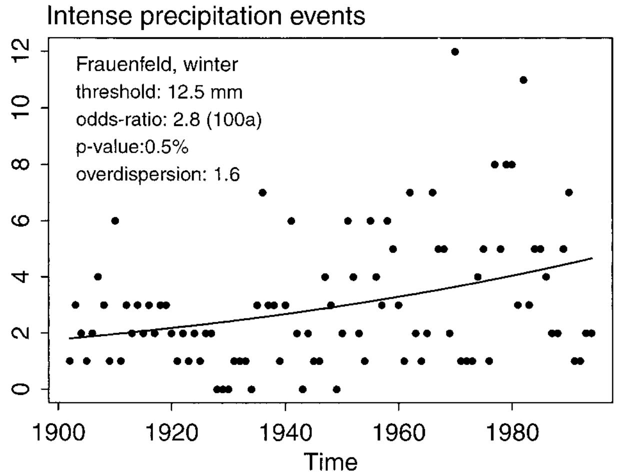

Fig. 1. Wintertime counts of intense daily precipitation events at station Frauenfeld in northeastern Switzerland. The full line depicts the fit of the logistic regression model (see section 2) to the data. Specific values of the fit and the trend test are listed in the inset (threshold for intense precipitation, trend magnitude as odds ratio Θ_{100}, *p* value of trend test, overdispersion σ^2).

As an example, Fig. 1 displays the counts of intense 24-h precipitation events for the winter seasons of 1901–94 at station Frauenfeld in northeastern Switzerland. Here and later in the study, an *intense* precipitation event is defined as the exceedence of a threshold that corresponds to the daily precipitation amount exceeded in the time series once per 30 days during the climatological period (counts above empirical quantile approach). At Frauenfeld this threshold is 12.5 mm day^{-1} in winter. The substantial increase in the occurrence of such events is evident from the record. The logistic regression model estimates a centennial increase by a factor 2.8 [Θ_{100}, see Eq. (7)], which is significant at a high level of confidence ($p = 0.5\%$). The overdispersion estimated with the model amounts to $\sigma^2 = 1.6$, indicating larger than binomial variance. Notice that the distribution of the data in the count record of Fig. 1 is substantially skewed and there is a tendency for increasing variance during the period. These characteristics stand in contrast to the assumptions for Gaussian linear regression, but they are a natural ingredient of the binomial distribution and logistic regression. Notice from Eq. (1b) that an increase of the expected value is coupled to an increase of the variance for $\pi < 0.5$ (i.e., in particular for rare events). We should mention that the trend analysis for station Frauenfeld yielded a trend magnitude at the upper range of results obtained for the Alpine station sample (see later section 4). This example was chosen here to visually illustrate the methodology.

In summary, logistic regression provides a suitable methodology for trend analyses of count records. The method has been applied in a wide range of scientific disciplines, but to our knowledge this is the first study

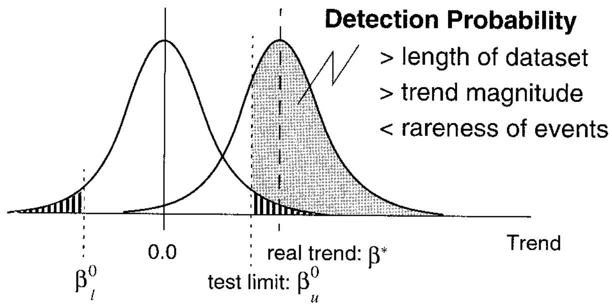

FIG. 2. Schematics of trend testing. Probability density functions of trend estimates under the null hypothesis and for an alternative with a presumed real trend. Trend estimates beyond test limits (striped) are statistically significant. The shaded area represents the detection probability under the alternative.

to apply it to rare weather events in climatology. Apart from providing a methodology for trend inference, the binomial concept also offers a simple scheme for quantifying type II errors.

b. Trend detection probability

Type II errors refer to the case where an existing real trend is not statistically detected by the trend test, due to the stochastic fluctuations in the record. While statistical testing aims at discriminating between signal and noise in a given data record, a quantification of type II errors addresses the problem of trend detectability, given a prescribed magnitude of the trend.

In principle the problems with type I and type II errors do not necessarily require separate statistical treatment. Both errors are due to the uncertainty of the trend estimate, which can be described in terms of a confidence interval. However it has not been customary to display confidence bounds in trend analyses, as the primary focus is to underline the evidence (i.e., statistical significance) of a trend, rather than the uncertainty of its estimate. Also, in practice it is more helpful to have information on estimation uncertainties prior to the exploration of the data series. We therefore pursue our quantification of type II errors largely decoupled from and prior to the analysis of data, by calculating a measure for trend detectability, the *detection probability.*

To illustrate the notion of *detection probability,* Fig. 2 alludes to the conventional procedure of a statistical test. Schematic probability density functions (pdf) depict the distribution of trend estimates for independent records in the absence (null hypothesis) and presence (alternative) of a given real trend. (The pdfs could equally represent the distribution of a suitable test statistic.) Their width characterizes the estimation uncertainty due to the stochastic fluctuations in the observational record. A trend estimate is considered as statistically significant if its value is to the tail-end of either of two thresholds (β_l^0 and β_u^0, two-tailed test, stippled areas) given by the pdf for the null hypothesis and the significance level. From a set of independent records with a real trend β^* (estimates represented by the right-hand pdf), some of the estimates will be considered as statistically significant (shaded area), while for some others the null hypothesis is accepted despite the trend. The fraction of significant trends obtained (shaded area) is a measure of detectability to be expected for a single record, and it will be termed the *detection probability.*

It is evident that the detection probability can become significantly smaller than 1 if the estimation uncertainty is comparable to the magnitude of the real trend. Hence the detection probability is a function of the magnitude of the trend and the length of the record under consideration. Moreover for rare weather events the detection probability for a given trend magnitude also depends on the rarity of the events considered, due to the decrease in the effective sample size.

In section 3 we will display quantitative estimates of the detection probability and its functional dependencies. The calculations are based on Monte Carlo simulated surrogate event records superimposed by a prespecified long-term trend. By reference to the binomial statistics, the number of events $n(t)$ in year t is simulated as a random variable with binomial distribution [see Eq. (1a)], taking the form:

$$n(t) \approx B[\cdot;\, \pi(t;\, \alpha,\, \beta),\, m] \tag{8}$$

($\approx$ meaning "distributed as"). Here the temporal dependence of the event probability π models the imposed trend and is specified in line with the logistic regression trend model given in Eq. (4). The parameters of the trend model α and β are calculated by specification of the odds ratio Θ [see Eq. (6)] and the time-average event probability $\overline{\pi}$. The latter is obtained from an integration of Eq. (5) over the time period under investigation:

$$\overline{\pi} = \frac{1}{T}\int_0^T \frac{\exp(\alpha + \beta t)}{1 + \exp(\alpha + \beta t)}\, dt = \frac{1}{\beta T}\log\left[\frac{1 + \exp(\alpha + \beta T)}{1 + \exp(\alpha)}\right], \tag{9}$$

where $T = t_2 - t_1$ is the length of the period. Notice that $\overline{\pi}^{-1}$ is a measure of the overall event rarity and corresponds to the expected value of the average return period over the climatological period (t_1, t_2). Expressions for α, and β in terms of T, Θ, and $\overline{\pi}$ can then be derived from Eqs. (6) and (9) and take the form:

$$\alpha = \log\left(\frac{\Theta^{\overline{\pi}} - 1}{\Theta - \Theta^{\overline{\pi}}}\right) \quad \text{for } \Theta \neq 1$$

$$\alpha = \log\left(\frac{\overline{\pi}}{1 - \overline{\pi}}\right) \quad \text{for } \Theta = 1 \quad \text{and} \tag{10a}$$

$$\beta = \frac{\log(\Theta)}{T}. \tag{10b}$$

The detection probability for a particular setting of observation period T, trend magnitude Θ, and the time-average event probability $\overline{\pi}$ is calculated from a sample of 2000 surrogate event records. For each of the series a logistic regression model is fitted and the trend tested for statistical significance as described in section 2a. The fraction of trend estimates in the sample, significant at the 5% level, is then associated with the detection probability. (It should be mentioned that the calculation of detection probabilities for the logistic trend model can in fact also be undertaken analytically, making use of the Fisher information matrix and accepting the asymptotic gaussianity of maximum likelihood estimates. The analytical derivations are extensive, and it was decided to perform the calculations by statistical simulation, which for the present situation is straightforward to implement.)

It is important to notice that the surrogate records of the Monte Carlo technique constitute idealizations of meteorological count records. Their construction assumes independence of events between successive days and a smoothly varying trend component. As a result the surrogate records tend to underestimate the variance of real count records, which are affected by overdispersion. Therefore the detection probabilities inferred from the Monte Carlo technique constitute optimal estimates of trend detectability, and overdispersion is likely to impose additional limitations. The consideration of overdispersion for detection probabilities would require the use of more complex stochastic models involving, for instance, Markov processes. This is not considered explicitly in this study as it would introduce additional parameters and ambiguities as regards the specification of the model. Yet for our application to heavy precipitation trends in the Alpine region, it will be interesting to consider variations of the overdispersion in the real count records (see section 4c) as this is an additional factor modulating trend detectability.

3. Limits of trend detection

This section presents quantitative estimates of the principal potential/limitation to detect trends of rare weather events in climate series. This potential is expressed by means of the *detection probability*, the expected chance to discern a given trend as statistically significant (see section 2b and Fig. 2). Detection probabilities are calculated as a function of trend magnitude [the odds ratio Θ, see Eq. (6)], the record length T in years, and the average occurrence of the events $\overline{\pi}$. The calculations are based on Monte Carlo–simulated surrogate trendy event records (see section 2b) and the logistic regression trend model (see section 2a), using a significance level of 5%. Results are presented for records of year-by-year counts of daily events. Two cases are distinguished where counts are either from one season [seasonal count records, number of independent trials $m = 90$, Eq. (1a)] or the whole year (annual count records, $m = 365$).

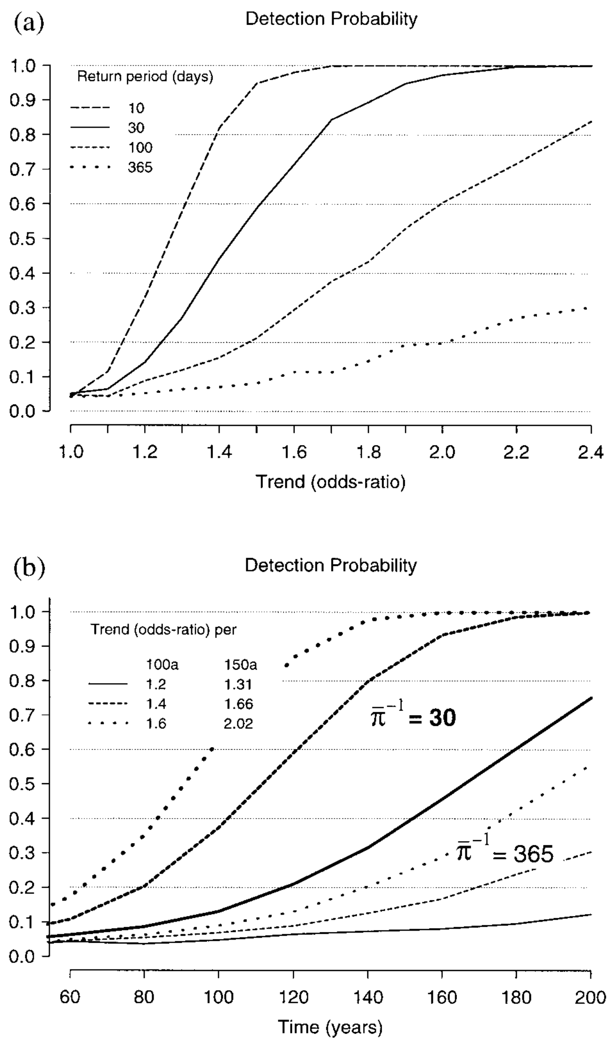

FIG. 3. Detection probability for trends in records of rare weather events. Results for seasonal count records ($m = 90$). (a) Detectability, for records of 100-yr length, are given as a function of trend magnitude (odds ratio Θ, horizontal axis) and for events with an average return period $\overline{\pi}^{-1}$ of 10, 30, 100, and 365 days. (b) Detectability for events with $\overline{\pi}^{-1} = 30$ days (thick lines) and $\overline{\pi}^{-1} = 365$ days (thin lines) is given as a function of record length and for three values of Θ, as listed in the inset. The odds ratio Θ approximates the fractional change of events over the period.

Figure 3a depicts the detection probability as a function of the trend magnitude Θ for seasonal count records of $T = 100$ yr. Results are shown for events with an average return period of 10, 30, 100, and 365 days, respectively (i.e., $\overline{\pi}^{-1} = 10, 30, 100, 365$). Notice that the odds ratio Θ (horizontal axis) corresponds approximately to the relative change of event probability over the 100-yr period (see section 2a). We have used increasing trends ($\Theta \geq 1$) for the horizontal axis in Fig. 3a, but detection probabilities are identical for Θ and Θ^{-1}.

The results show an obvious gradual increase of detection probability with trend magnitude. Moreover the detection probability is remarkably sensitive upon the rarity of the events: for moderate events with a return period of 10 days ($\overline{\pi}^{-1} = 10$) a centennial change of a few 10% can be detected with reasonable chance, but the detection probability rapidly drops with more rare events. For example a long-term change by a factor of 1.5 in the occurrence of intense events (return period 30 days) is identified with a probability of 0.6, and a similar trend yields a detection probability of 0.2 for events with a return of 100 days. Considering events with $\overline{\pi}^{-1} = 365$, even a doubling or halving of their occurrence ($\Theta = 2.0$) can be identified as statistically significant only with a probability of 0.2 in a 100-yr record.

The availability of daily climate records over a centennial period is limited in practice and trend analyses of extreme events are often restricted to shorter period records. With a shorter data period the probability for trend detection might suffer from a poor prominence of eventually ongoing long-term changes and the restricted sample size affecting the error of trend estimates. Figure 3b depicts the evolution of detection probability calculated for three selected scenarios of long-term trend. The scenarios pertain to the logistic trend model [see Eq. (5)] and a centennial trend magnitude of $\Theta = 1.2$, 1.4, and 1.6, respectively. (Odds ratios for 150 yr are also given in the inset.) The results demonstrate that the observation period is a critical factor for trend detection of intense weather events: while a 60-yr period is hardly sufficient to resolve rare event trends for any of the three scenarios, there is a substantial increase of detection probability with increasing record length in the case of $\overline{\pi}^{-1} = 30$ (thick lines). For very rare events ($\overline{\pi}^{-1} = 365$, thin lines) improved detectability is only obtained with a substantially longer and in practice unrealistic observation period (or much more dramatic trends). Notice that for such severe events even a doubling of occurrence over a 150-yr period is only detected with a chance of 0.25. Evidently the serious limitations for trend detection in very rare, severe weather events are not only critical for currently available climate records but even for those available in a few decades from now.

The results of Fig. 3 point to a substantial decrease of detection probability with the rarity of events when a similar trend magnitude is assumed. Yet natural trend signals might be highly variable across the frequency spectrum. Therefore additional insights on trend detectability can be gained by considering physically or statistically motivated trend scenarios describing long-term variations for the whole frequency spectrum. Here we consider such a scenario that has been suggested for daily precipitation and is motivated from simulations of global warming with global and regional climate models (Fowler and Hennessy 1995; Jones et al. 1997; Frei et al. 1998). It assumes a relative change of daily precipitation amounts independent of event intensity, that is, a change in the scale parameter of the probability density function. In terms of event frequencies the *scale-shift*

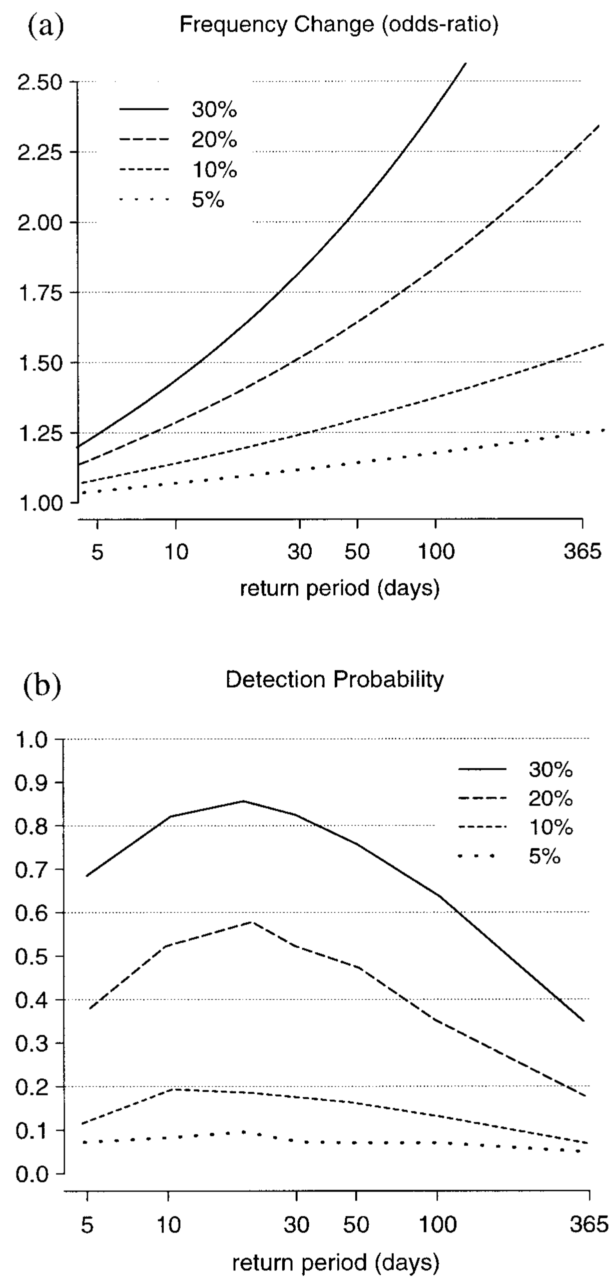

FIG. 4. (a) Frequency change of precipitation events under the "scale-shift" scenario (i.e., from the rescaling of a daily rainfall pdf typical for the Alpine region). Frequency change is displayed in terms of the odds ratio Θ as a function of event return period ($\overline{\pi}^{-1}$, in days) and for four scaling factors (listed in the inset). (b) Detection probability for the scale-shift scenarios depicted in panel a (i.e., intensity progressive trends). Probability is given as a function of return period (days, horizontal axis) and is valid for seasonal count records of 100-yr length.

scenario results in a progressively stronger change for rare high-intensity events (e.g., Fowler and Hennessy 1995). Figure 4a depicts several examples of this scenario based on a frequency distribution of daily precipitation characteristic for the Alpine region.

Detection probabilities for the scale-shift scenarios of Fig. 4a have been calculated assuming a gradual change over a 100-yr period according to the trend evolution of Eq. (5). The results (see Fig. 4b) exhibit maximum detection probability in a range of return periods between 10 and 30 days, that is, for moderately rare events. The decrease toward smaller return periods is the result of a decreasing trend amplitude as prescribed by the scenarios. On the other hand, toward longer return periods, the increasing trend amplitude cannot fully compensate for and is still dominated by the limitations of detectability with increasing event rarity. Hence even in the presence of an intensity progressive trend signal, the detection probability may be more limited for extreme than for intense events.

The detection probability calculated for Figs. 3 and 4 represent the situation for seasonal count records. Seasonal stratification in trend analyses is most appropriate for areas and parameters with a distinct annual cycle. Rare events often are prominent in particular seasons, and a nonstratified procedure might obscure compensating trends in different seasons. Yet in certain cases (e.g., in the Tropics) analyses of annual count records are also feasible and the corresponding detection probabilities are of interest. Figure 5 displays the results for annual count records (i.e., $m = 365$) considering similar return periods as those for the seasonal case. For similar settings of average return period, record length, and trend magnitude, the detection probability is higher for the annual ($m = 365$) compared to the seasonal case ($m = 90$, Fig. 3), which is a result of the increase in sample size. The functional behavior is similar to the seasonal case.

It is important to recall that the calculation of detection probability has neglected overdispersion for the construction of the surrogate count records and hence provides more ideal conditions for trend detection (i.e., lower variance) than can be expected in real data (see section 2). An account of overdispersion for the detection probability could be accomplished by statistical simulation using some explicitly autocorrelated time series models such as a first-order Markov chain. Alternatively, the effect of overdispersion could be considered by an appropriate reduction of the real sample size m to the effective sample size m/σ^2 of the overdispersion free situation (McCullagh and Nelder 1989). These calculations are not outperformed explicitly here, as overdispersion varies greatly between climate parameters and can also depend on event probability $\overline{\pi}$ (as was found for heavy precipitation in the Alps). Therefore the theoretically determined detection probabilities of this section should be considered as an upper bound of what can be expected with observed climate data.

This section has quantitatively illustrated the limitations of discriminating long-term trends from stochastic variations. This information is fundamental to trend analyses of rare weather events. Extreme events, which are most relevant from an impact point of view, are particularly exposed to the limitations. The use of more moderate thresholds for the definition of events, and efforts to obtain longer period observations appear important to obtain more conclusive results from a trend analysis. Moreover the results illustrate that the absence of a statistically significant trend is not synonymous to the absence of a trend. In fact the chance of detection might be marginal even for quite substantial trends present in the data. Quantitative knowledge about limits of trend detection constitutes an important element for a proper assessment of results from a trend analysis, especially when rare weather events are considered.

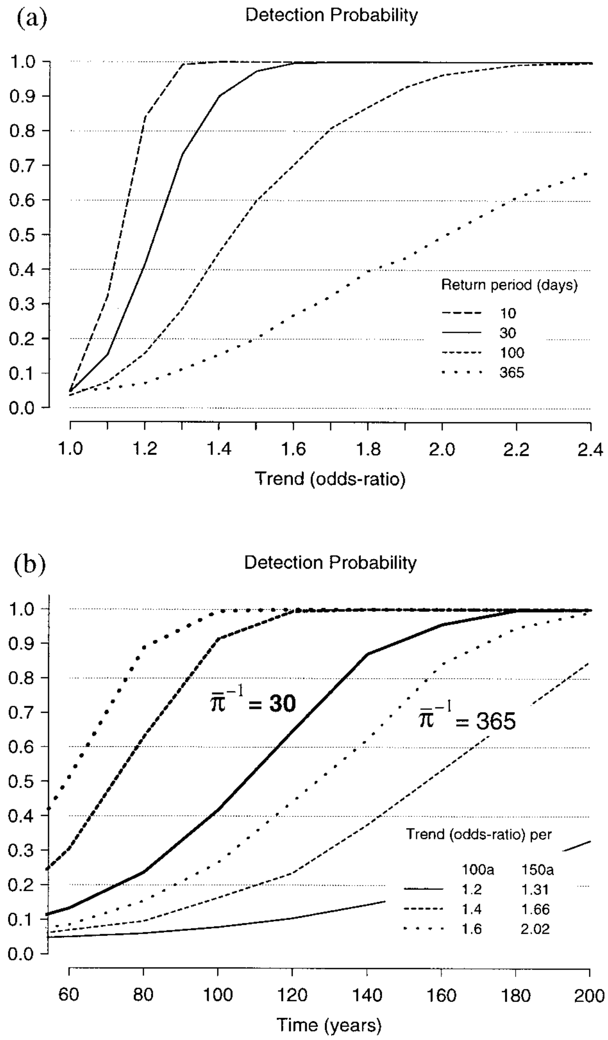

Fig. 5. As Fig. 3 but for annual count records ($m = 365$).

4. Heavy precipitation trends in the Alpine region

In this section the statistical framework of binomial distributed event counts will be applied to examine long-term trends of heavy precipitation occurrences in the Swiss Alps. Trend estimation and testing is based on

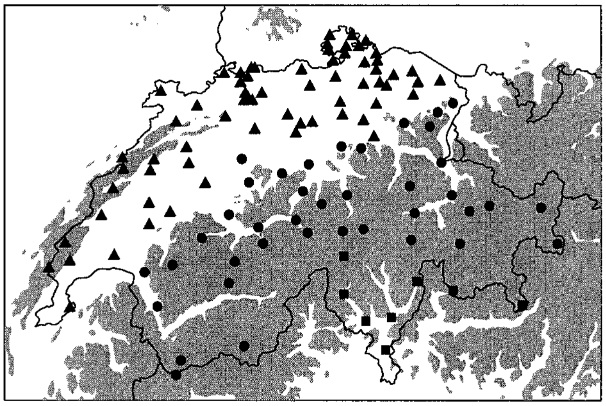

FIG. 6. Geographic outline of Switzerland with topography (areas above 1000 m MSL are shaded) and the distribution of the 113 long-term rain gauge stations used in the trend analyses. Station symbols refer the geographical regions used in Table 1 (triangles: *north,* circles: *Alps,* squares: *south*).

TABLE 1. Typical threshold values (mm day^{-1}) used for the definition of heavy precipitation events. Thresholds are defined individually for each station and month of the year, corresponding to events with an average return period of 10 days for *moderate,* 30 days for *intense,* 100 days for *strong,* and 365 days for *extreme* precipitation. Numbers in the table represent median values for stations to the *north* of, within the *Alps,* and to the *south* of the ridge (see station labels in Fig. 6).

Category (return period)		Winter mm day^{-1}	Spring mm day^{-1}	Summer mm day^{-1}	Autumn mm day^{-1}
Moderate	North	8	9	12	9
(10 days)	Alps	9	10	15	10
	South	6	15	18	17
Intense	North	15	16	23	18
(30 days)	Alps	17	18	25	20
	South	21	34	39	39
Strong	North	24	25	35	29
(100 days)	Alps	28	28	37	33
	South	39	52	70	63
Extreme	North	34	36	48	40
(365 days)	Alps	43	40	53	48
	South	58	72	99	97

the logistic regression analysis as described in section 2, and the results on trend detectability in section 3 will serve as background for the assessment of the trend results.

a. Data

The dataset for this trend study is composed of daily precipitation series at 113 rain gauge stations in Switzerland (see Fig. 6). The data have been provided by MeteoSwiss Zürich and embrace all Swiss rain gauges for which a continuous and complete daily record is available throughout the 94-yr period 1901–94. The station sample is similar to that used for a trend analysis of mean precipitation by Widmann and Schär (1997). With a typical interstation distance of 20 km it constitutes a long-term observation system with an exceptional density. Details of the Swiss rain gauge network can be found in Müller and Joss (1985) and Weingartner (1992).

Several categories of heavy daily precipitation events will be considered to span a range of intensities. These are defined in terms of threshold exceedence of the daily precipitation totals, and count records are formed of the seasonal occurrences in each year. The thresholds are defined as the upper 1/10, 1/30, 1/100, 1/365 quantiles of the empirical distribution function (counts above empirical quantile approach). Hence the count records correspond to daily rainfall events with an average return period of 10, 30, 100, and 365 days and will be referred to as *moderate, intense, strong,* and *extreme* events, respectively. The thresholds are determined individually for each station and month of the year. This procedure compensates for the pronounced regional variations between stations (e.g., the local exposure), and the seasonal cycle. It ensures comparability of the trend results across the station sample. Notice that the choice of return periods is similar to those chosen for the calculation of detection probabilities in section 3. The inclusion of moderate thresholds is motivated from these theoretical considerations, which suggest that trend analyses for very rare, extreme events are compromised by a low detection probability.

Table 1 lists typical threshold values for station samples in three geographically different regions of Switzerland (notice the station labels in Fig. 6). For the station samples to the north of and in the Alps, summer is the main season of heavy precipitation (high threshold values), due to summertime convection and heavy thunderstorms. More frequent heavy events occur to the south of the ridge where local convection and topographic precipitation linked to moist southerly airflows contribute to peak activities in summer and autumn. Climatological analyses of heavy precipitation in the Alps can be found, for example, in Courvoisier (1981), Geiger et al. (1991), and Frei and Schär (1998).

b. Trend results

Results of the trend analyses obtained for *intense* and *extreme* precipitation events over the period 1901–94 are reproduced in Figs. 7–10. The seasonal maps (Figs. 7 and 9) depict the sign and statistical significance of the trend estimates (circles: increasing, triangles: decreasing, filled symbols: $p < 5\%$) for each of the 113 long-term rain gauge series. Results for all four categories of heavy precipitation are summarized in Table 2.

For *intense* precipitation events, Fig. 7 reveals a seasonally distinct trend signal: In spring and summer the trend results in the station sample are roughly balanced between increasing and decreasing estimates and there is only a small number of stations for which the trends are statistically significant. In contrast for winter and autumn the station charts suggest an increasing trend

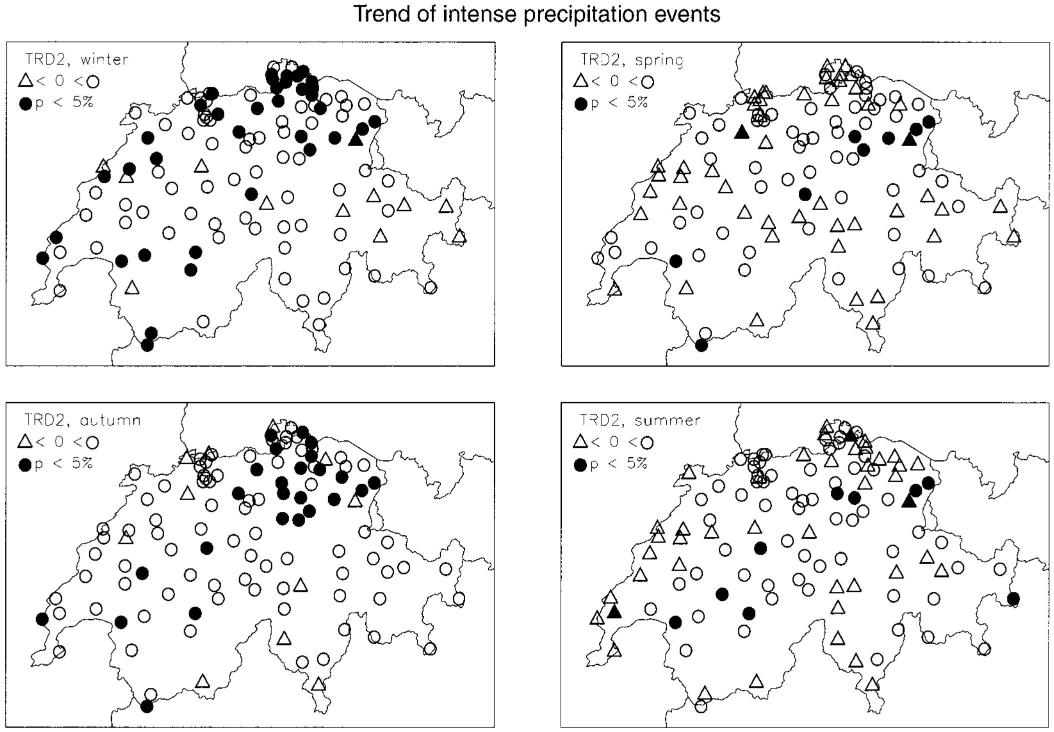

FIG. 7. Seasonal charts showing trend results in the frequency of *intense* daily precipitation (i.e., events occurring once per 30 days on average over the climatological period). Symbols represent the sign of the trend estimate (circle: increasing, triangle: decreasing) and its statistical significance (filled symbols: $p < 5\%$). Trends were estimated and tested from seasonal event counts over the period 1901–94 using logistic linear regression.

signal with a clear prominence of positive estimates, and a high number of sites with significant test results (36 for winter, 25 for autumn, see also Table 2). The northeastern flatland region of the country exhibits a particularly high density of stations with significant increases, but there is also a considerable number of significant trend results over western Switzerland for winter. Notice that there is one site with a negative trend signal in all seasons (significant, except in autumn). The series of this station, located at 2500 m mean sea level (MSL) on mount Säntis, is known for its homogeneity problems (P. Hachler, MeteoSwiss 1999, personal communication).

Histograms of the trend estimates for intense precipitation [expressed as centennial values $\beta \cdot 100$ and Θ_{100}, see section 2a, Eq. (7)] are depicted in Fig. 8 and reveal considerable interstation variation during all seasons. Nevertheless the histograms compactly illustrate the seasonal variation of the trend signal in Switzerland, with symmetrical distributions for spring and summer and a prominent offset in the histograms for winter and autumn. The median values of the trend estimates Θ_{100} (fractional change over 100 yr) amount to 1.54 and 1.35 for winter and autumn, respectively.

Although considerable limitations for the detection of centennial trends in *extreme* precipitation (return period 365 days) must be expected, trend results for this event category are briefly illustrated. The seasonal trend maps

TABLE 2. Number of Swiss long-term rain gauge series with positive/negative and significant ($p < 5\%$) trends in the occurrence of *moderate, intense, strong,* and *extreme* daily precipitation events. Total number of stations is 113.

Category (return period)		Winter		Spring		Summer		Autumn	
			$p < 5\%$		$p < 5\%$		$p < 5\%$		$p < 5\%$
Moderate	+	102	40	67	12	43	1	84	5
(10 days)	−	11	0	46	0	70	3	29	0
Intense	+	101	36	63	8	70	9	102	25
(30 days)	−	12	1	50	2	43	3	11	0
Strong	+	98	15	59	5	75	12	96	28
(100 days)	−	15	0	54	1	38	0	17	1
Extreme	+	89	3	66	5	87	13	87	19
(365 days)	−	24	1	47	2	26	0	26	1

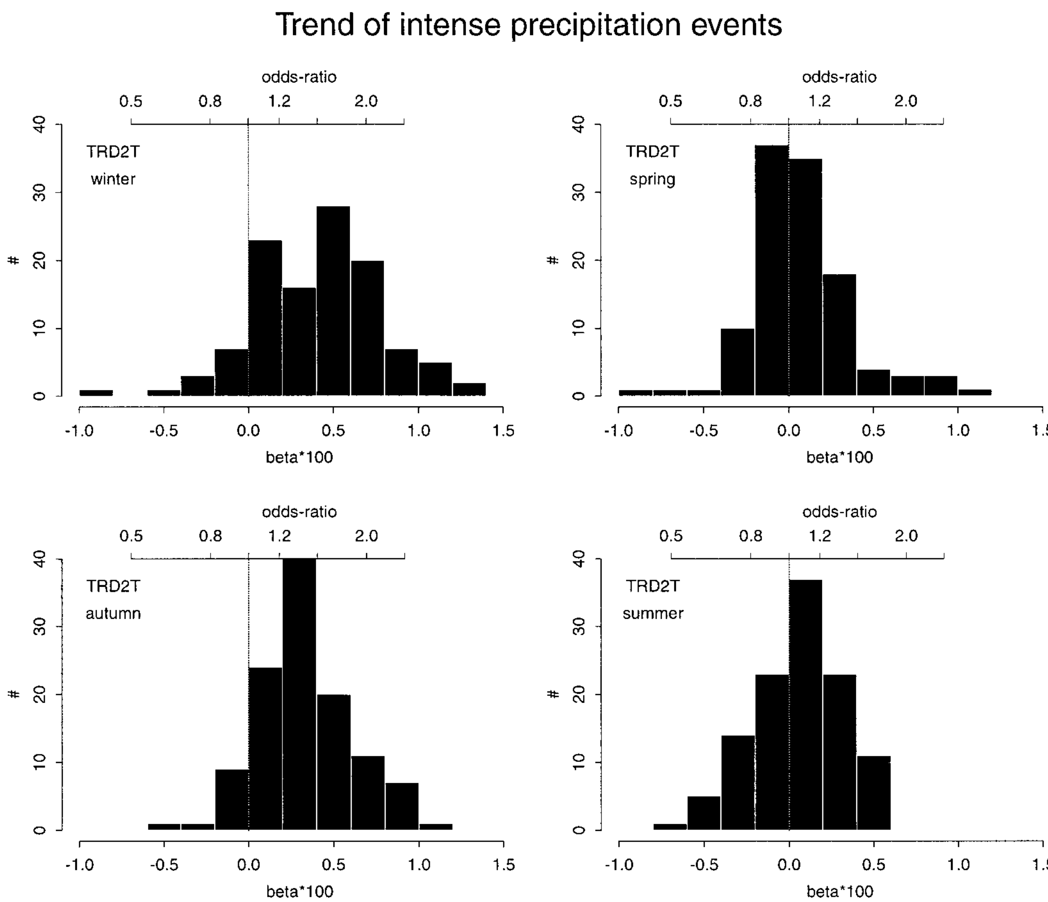

FIG. 8. Histograms of trend estimates for *intense* daily precipitation (average return period 30 days) obtained for 113 Swiss long-term rain gauge stations. Trend estimates (horizontal axes) are given in centennial values: $\beta \cdot 100$ (bottom axes) and the odds ratio Θ_{100} (top axes). The latter represents approximately the relative frequency change between the beginning and the end of the century (see also section 2a).

(Fig. 9) show only few series for which trends are statistically significant. Perhaps the moderate number of significant increases in summer (13) and autumn (19) could be taken as a weak indication. Still the absence of statistical evidence for trends should not be mistaken as the necessary absence of long-term changes. The histograms of the trend estimates (Fig. 10) reveal substantial variations and hence point to the uncertainty of the trend estimates. (Notice the difference in scale between Figs. 8 and 10.) For example, the trend estimate for wintertime at station Frauenfeld is substantial ($\Theta_{100} = 1.75$), however, it is not statistically significant ($p = 15\%$).

Table 2 summarizes the trend results for all categories of heavy precipitation considered and compactly reproduces the visual impression obtained from an inspection of seasonal trend maps for *moderate* and *strong* events (not shown): For the winter season there is a strong bias toward increasing trend estimates and a high portion of statistically significant increases in *moderate* and *intense* precipitation events. The spatial distribution of the trend is similar between the two event categories, that is, there are significant increases at a prominent cluster of stations over the northeast and also for numerous sites in the northwestern regions of the country. The wintertime trend signal is then gradually less apparent for the more rare *strong* and *extreme* events. In the autumn season significant trends are found at a considerable portion of the sites for *intense* and *strong* events. In this case the trend analysis did not reveal a trend signal in the occurrence of *moderate* precipitation events, in fact a high number of negative trend estimates (rarely significant though) was seen in the occurrence of events with even shorter return periods and also for the frequency of rainy days. Finally for spring and summer the results of the trend analyses do not exhibit conclusive trend signals at any of the event categories.

c. Discussion

The trend analyses for the Swiss long-term records has revealed a large number of stations with significant trends of *intense* precipitation events for winter and autumn and over the eastern pre-Alpine flatland regions.

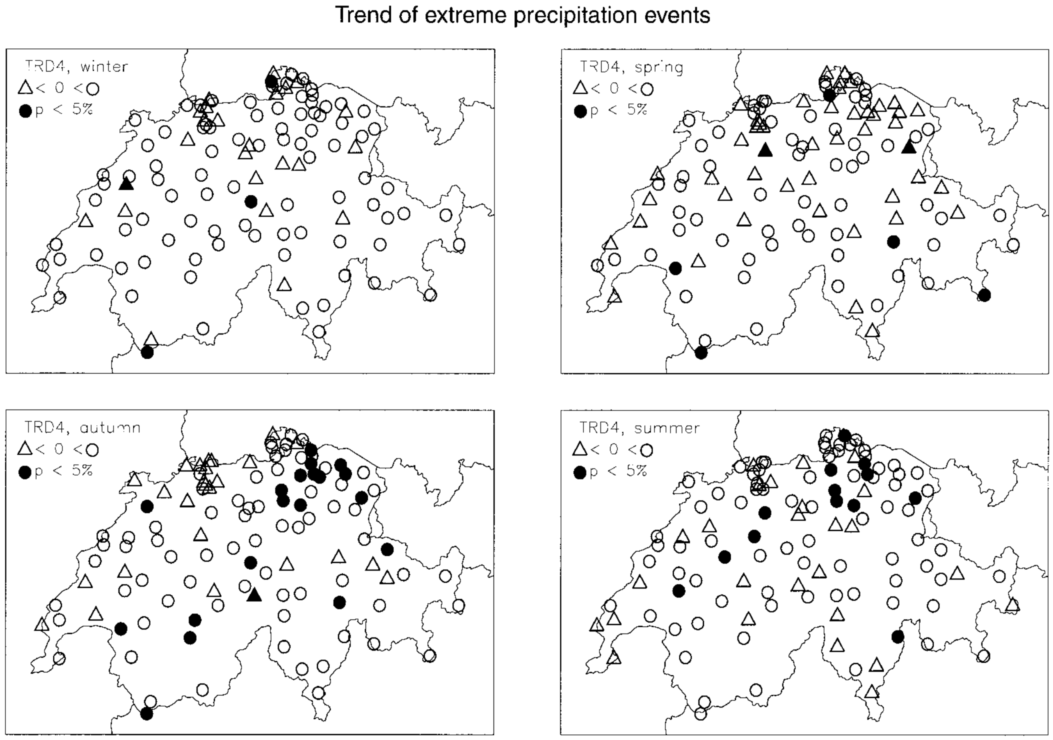

FIG. 9. As in Fig. 7 but for *extreme* daily precipitation (occurring once per 365 days on average over the climatological period).

As illustrated in section 3, the fundamental statistical limitations constitute a critical factor for trend detection, and the observed seasonal and spatial variation of the trend could, in principle, be an artefact of seasonal and spatial variations in trend detectability. By adapting the exceedence thresholds to the distribution function of each record, gross variations in detectability have been avoided. (With a fixed threshold value for all stations, the spatial variations of event frequency would have resulted in variations of the detection probability.) Apart from the mean event frequency, however, the overdispersion σ^2 is an additional factor contributing to seasonal and spatial variations of the detection probability. Recall that σ^2 is a measure for the residual variance in the logistic trend model (see section 2). A large value of σ^2 can result, for example, from the serial correlation of heavy precipitation events or nonstationary large-scale flow conditions. It increases the uncertainty of the trend estimate and constrains the detectability of long-term trends below the values given in section 3.

Figure 11 displays maps of the overdispersion σ^2 as estimated with the logistic trend model for the case of *intense* precipitation. Symbols represent values above, between, and below 0.8 standard deviations from the mean (see symbol legend for discriminating values). There are pronounced seasonal variations of σ^2 with high values in winter, only weak overdispersion in summer, and intermediate conditions in spring and autumn. These can be interpreted in terms of the characteristic weather systems responsible for heavy precipitation events (see, e.g., Schär et al. 1998): In winter large-scale persistent depressions and their frontal systems impose a high probability of heavy precipitation on consecutive days (serial correlation), whereas in summer small-scale convective fronts and locally triggered thunderstorms exhibit weaker persistence. The differences in overdispersion suggest that trend detectability for spring and summer is not excessively compromised by overdispersion, and that the results of the trend analysis reflect statistically sound seasonal variations in trend magnitude. As regards the spatial variations, higher than normal values for σ^2 are found for the few stations along the southern rim of the ridge (especially in winter and spring), and this might point toward more serious limitations in trend detection from excessive variance in these southern records.

While the logistic regression model provides a suitable statistical method to test for trends in individual count records, this method does not provide a means to assess the overall significance (field significance) of the station-by-station test results. As in many other trend analyses, we have taken an empirical approach, considering a large portion of significant trends as indicative of a regional trend (e.g., Table 2). A formal caveat of this interpretation is that trend estimates are not strictly statistically independent between stations, due to the spatial correlation within the network. For example, the high number of statistically significant test results in the

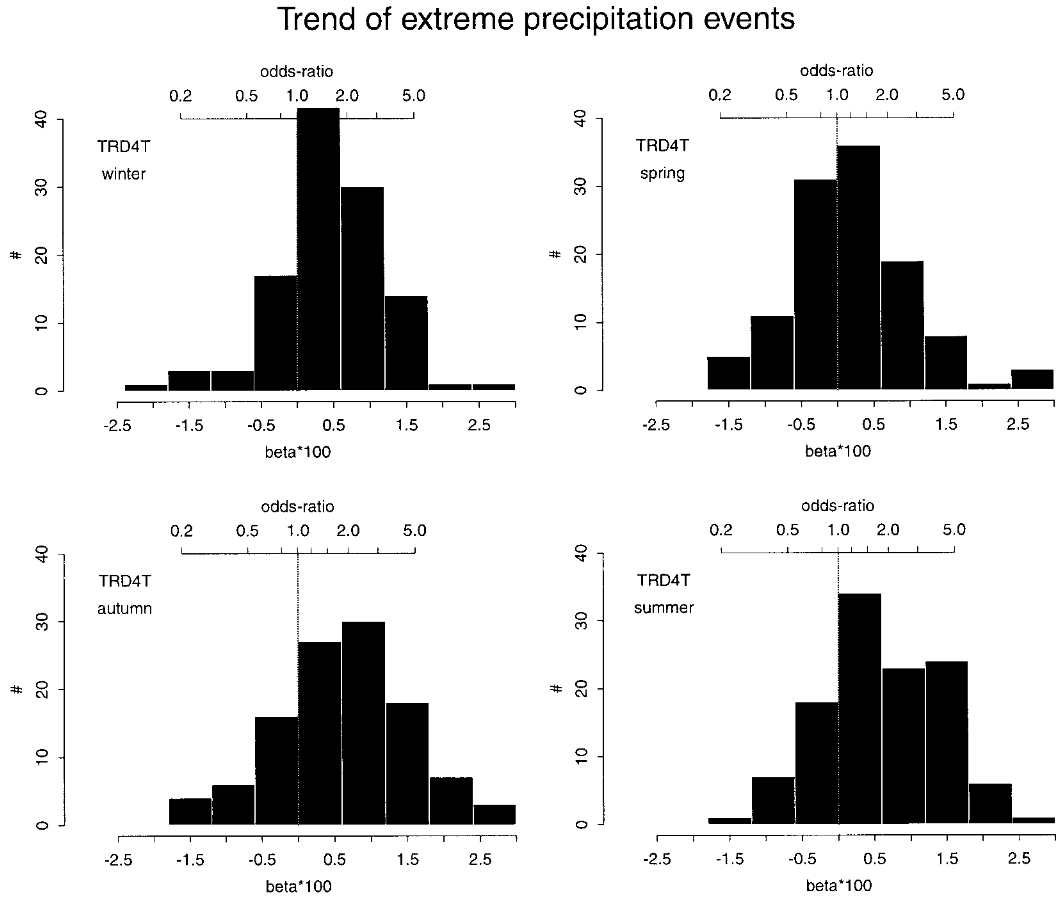

FIG. 10. As in Fig. 8 but for *extreme* daily precipitation (average return period 365 days). Note that the scale of the trend estimates (horizontal axes) is different from Fig. 8.

northeastern sector of the country could be influenced by correlation among the series. A final conclusion about field significance and regional variations of the trend signal would involve a spatiotemporal statistical model, which incorporates correlation between the count records of the station network. One possible option for this purpose would be a formal extension/adaptation of multivariate statistical methods (such as Principal Component Analyses; Preisendorfer 1988) for application with binomial count data.

Despite the special care taken by the data provider to keep observation practices stable through time, the Swiss long-term records could be affected by inhomogeneity, from observer changes, gauge displacements, instrumental renewals, or changes in the station environment. Again, in combination with systematic measurement biases, changes in the snowfall fraction or wind conditions could have affected larger portions of the station sample more gradually (e.g., Sevruk 1994; Groisman and Legates 1994). Although spurious effects on the trend results cannot be excluded, it is unlikely that the trends are solely an artifact of inhomogeneity. The main trend signals appear to be consistent among large station clusters. Moreover networkwide changes would have likely resulted in similar effects for all seasons. Again, the snowfall fraction is very small even in wintertime at the low elevations where most significant trends are found. A profound estimate of the role of inhomogeniety would require a correction of reported changes at individual stations (e.g., Easterling et al. 1996; Peterson et al. 1998), and also a quantification of more gradual and networkwide changes.

5. Conclusions

In this study we have presented a statistical framework for the assessment of long-term trends in rare weather events. The methodology operates on records of event counts, defined, for example, by threshold exceedence in daily climate records. It is based on binomial statistics and involves logistic regression for trend estimation and testing. Conceptually this regression is a natural model to deal with count records and it is also amenable to deviations from the binomial distribution related to excessive variance (overdispersion) in the data.

The concept of binomial counts has been employed to quantify the probability with which one can expect

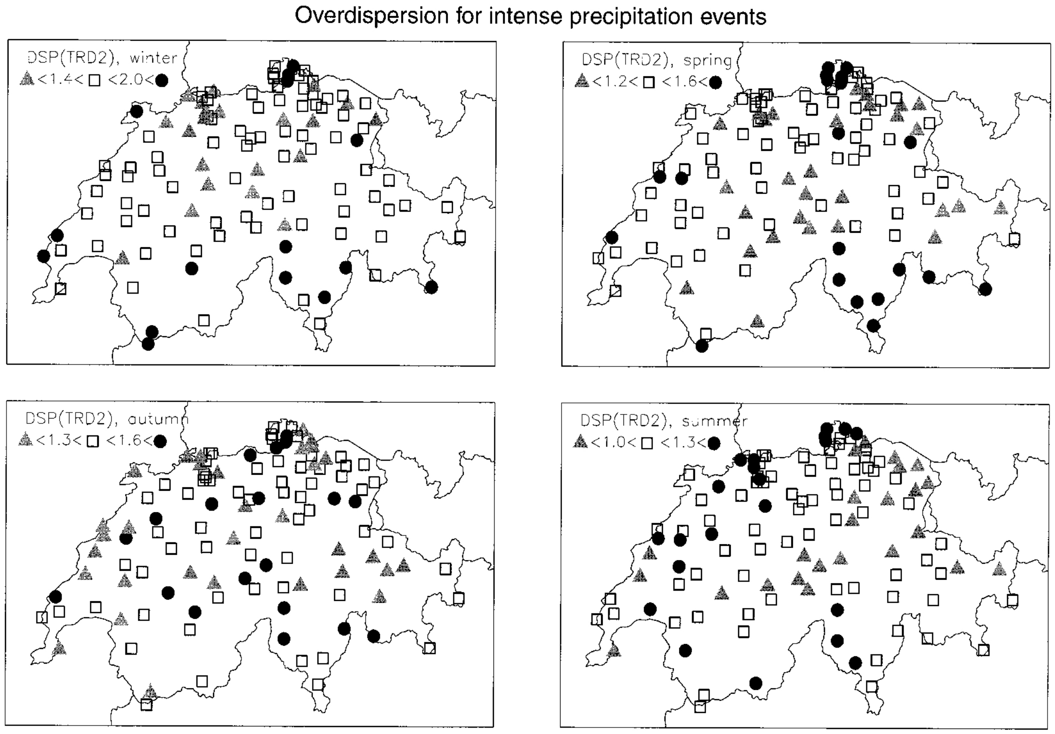

FIG. 11. Seasonal charts showing overdispersion σ^2 of the logistic trend model for *intense* precipitation (return period 30 days). Symbols at station locations represent anomalously low (triangles), typical (squares), and anomalously high (circles) values of σ^2 in the station sample of the respective season. The discriminating values, defined as the ± 0.8 standard deviations from the mean of the sample, are indicated in each panel. Notice the difference of discriminating values between seasons.

to discriminate a long-term trend from the stochastic variations in a count record. The results have exposed serious limitations for trend detection when rare extremes are considered. In a centennial record of seasonal counts, a change by a factor of 1.5 in the frequency of events with an average return period of 100 days is detected with a probability of 0.2 only. For shorter record periods and more seldom extremes, detectability becomes rapidly even worse.

The results have practical implications for trend analyses of rare events: first, long-term changes in the frequency distribution of a climate parameter might more readily be resolved by considering moderately intense rather than overtly extreme events. This is even the case when an intensity progressive change in the frequency distribution is assumed (see Fig. 4). Second, the length of the climate record is critical for trend detection and efforts for long-period records are beneficial for trend analyses in rare events. Third, to allow for a comparison of the trend results between regions and seasons, a comparable effective sample size should be assured between the records. The definition of event classes should be based upon their frequency, that is, by determining a station-dependent threshold with identical return period. Finally, a proper interpretation of trend results should be based on quantitative knowledge of the discriminating power of the trend test (or with the consideration of confidence bounds in trend estimates). The absence of statistical significance is not synonymous to the absence of trends in the data. Substantial long-term changes can be masked by the stochastic fluctuations associated with the small sample size in rare event records.

Our quantification of trend detectability strictly holds for count data and the logistic trend model. However similar limitations may also be expected from other statistical models, for example, the use of extreme value distributions for seasonal and annual extremes. The uncertainty of trend estimates is inherent to the small samples of rare events and it results from both the application of thresholds (as pursued in this study) or the restriction to extreme values.

It should be noted that the theoretical results of trend detectability derived in section 3 are valid for single station records. A more optimistic chance for trend detection may be expected from the combination of data from a station network. The associated increase in sample size compared to single station records could mitigate the uncertainty of the trend estimates and could help to extend trend detectability toward more severe event categories. Such an undertaking would require the development of a spatiotemporal statistical model, which takes into account the correlation structure between the stations of the network, can provide objective guidance on the regional subdivision of the data, and hence enables us to examine significance of regional trend patterns (field significance). Spatiotemporal trend models have been readily applied to deal with mean climate parameters (e.g., Widmann and Schär 1997) of a station network and are ideal for parameters with a near-Gaussian distribution. The extension of multivariate methods to rare event data with a non-Gaussian distribution requires more theoretical work but would be highly desirable for the expected increase in trend detectability.

The application of the logistic trend model to heavy precipitation events in the Alpine region has demonstrated practically the limitations of trend detection. For extreme events (return period 365 days), the number of stations with a significant trend was low in all four seasons. Yet the results must be considered poorly conclusive, as trends of a large amplitude where estimated without being statistically significant. The lack of statistical power is particularly regrettable from an impact point of view: *extreme* events as considered in this study are still not primarily representing situations responsible for serious flooding and damage. The Alpine region is well known for suffering from serious impacts of heavy precipitation, yet damage-causing events often involve precipitation amounts well beyond the 100 mm day^{-1} threshold (e.g., Röthlisberger 1998).

For intense precipitation in Switzerland (return period 30 days), the trend analysis has yielded substantial evidence of increasing trends in the course of this century. The increase was found for autumn and winter, seasons that are characterized by high synoptic weather activity. While the wintertime increase conforms to an increase of mean wintertime precipitation (e.g., Schönwiese et al. 1994; SMI 1996; Widmann and Schär 1997), the trend signal for autumn reflects mutually compensating long-term variations in the frequency distribution. The findings of this study confirm an earlier report on increasing heavy precipitation frequency in Switzerland (Courvoisier 1998), which was based on a subset of the records used in this study and conventional least squares regression.

The trend analysis for Switzerland conforms to the findings of increasing heavy precipitation during this century for other regions of the world, such as Scandinavia (Forland 1998), the continental United States (Karl and Knight 1998), the United Kingdom (Osborn et al. 2000), India (Rakhecha and Soman 1994), Japan (Iwashima and Yamamoto 1993), and Australia (Suppiah and Hennessy 1998). From experiments with global climate models it has been suggested that increasing precipitation intensities can result from an intensification of the atmospheric water cycle related to the anticipated anthropogenic warming in future (Fowler and Hennessy 1995; Hennessy et al. 1997; Zwiers and Kharin 1998; Meehl et al. 2000). This effect is expected primarily in middle and high latitudes where absolute humidity is primarily controlled by the ambient air temperature through the Clausius-Clapeyron relation (Mitchell and Ingram 1992; Gaffen et al. 1992; Bony et al. 1995; Trenberth 1999). Regional scenario simulations for $2 \times CO_2$ conditions (Jones et al. 1997) and idealized experiments (Schär et al. 1996; Frei et al. 1998) with limited-area climate models point toward the potential role of these processes for the occurrence of heavy winter and autumn precipitation over Europe. Again from a statistical analysis of observational data, Widmann and Schär (1997) conclude that the observed increase of mean wintertime precipitation is dominated by an increase of precipitation activity over trends in the frequency of weather types.

The observed increase of intense precipitation found in this study conforms with the ideas of an intensified water cycle and the observed long-term warming (e.g., Jones 1994; Beniston et al. 1994; Weber et al. 1997). But the role of long-term variations in the large-scale circulation pattern, such as those observed over the North Atlantic (Hurrell 1995; Hurrell and van Loon 1997), has yet to be assessed. For this purpose an extension of the binomial statistical framework for linking regional rare events to circulation or humidity variables would be highly desirable and could also serve the design of statistical techniques for downscaling rare extreme events from general circulation model output.

Acknowledgments. We are grateful to the Seminar for Statistics ETH (Dr. Verena Gelpke and Dr. Werner Stahel) for statistical advice and to MeteoSwiss Zürich for providing access to the long-term rain gauge records in Switzerland. We are also indepted to two anonymous reviewers for their valuable comments, which contributed to the improvement of this paper. The research was supported by the Swiss National Science Foundation (Contract SPP-U 5001-044602).

REFERENCES

Alexandersson, H., H. Tuomenvirta, T. Smith, and K. Iden, 2000: Trends of storms in NW Europe derived from an updated pressure data set. *Climate Res.,* **14,** 71–73.

Auer, I., and R. Böhm, 1996: Ein Beitrag zur Frage über die Zunahme extremer Niederschlagsereignisse und Ausweitung von Trockenperioden in einer wärmeren Atmosphäre anhand der Wiener Messreihe. *Wetter Leben,* **48,** 13–24.

Beniston, M., M. Rebetez, F. Giorgi, and M. R. Marinucci, 1994: An analysis of regional climate change in Switzerland. *Theor. Appl. Climatol.,* **49,** 135–159.

Bijl, W., R. Flather, J. G. de Ronde, and T. Schmith, 1999: Changing storminess? An analysis of long-term sea level data sets. *Climate Res.,* **11,** 161–172.

Bony, S., P. Duvel, and H. Le Treut, 1995: Observed dependence of the water vapor and clear-sky greenhouse effect on sea surface temperature; Comparison with climate warming experiments. *Climate Dyn.,* **11,** 307–320.

Changnon, S. A., and M. Demissie, 1996: Detection of changes in streamflow and floods resulting from climate fluctuations and land-use drainage. *Climatic Change,* **32,** 411–421.

Courvoisier, H. W., 1981: Starkniederschläge in der Schweiz in Abhängigkeit vom Druck-, Temperatur-, und Feuchtefeld. SMI Rep. 42, 59 pp. [Available from MeteoSwiss, Krähbühlstrasse 58, CH-8044 Zürich, Switzerland.]

——, 1998: Statistik der 24-stündigen Starkniederschläge in der Schweiz 1901–1996. SMI Rep. 194, 20 pp. [Available from MeteoSwiss, Krähbühlstrasse 58, CH-8044 Zürich, Switzerland.]

DeGaetano, A. T., 1996: Recent trends in maximum and minimum temperature threshold exceedences in the northeastern United States. *J. Climate,* **9,** 1646–1660.

Diggle, P. J., K.-Y. Liang, and S. L. Zeger, 1994: *Analysis of Longitudinal Data.* Oxford University Press, 253 pp.

Dobson, A. J., 1990: *Introduction to Generalized Linear Models.* Chapman and Hall, 174 pp.

Easterling, D. R., R. C. Peterson, and T. R. Karl, 1996: On the development and use of homogenized climate datasets. *J. Climate,* **9,** 1429–1434.

——, G. A. Meehl, C. Parmesan, S. A. Changnon, T. R. Karl, and L. O. Mearns, 2000: Climate extremes: Observations, modeling and impacts. *Science,* **289,** 2068–2074.

Forland, E. J., 1998: Long-term variations in extreme 1-day rainfall in the nordic countries. *Proc. European Conf. on Applied Climatology,* Vienna, Austria, Central Institute for Meteorology and Geodynamics, CD-ROM.

Fowler, A. M., and K. J. Hennessy, 1995: Potential impacts of global warming on the frequency and magnitude of heavy precipitation. *Nat. Hazards,* **11,** 282–303.

Frei, C., and C. Schär, 1998: A precipitation climatology of the Alps from high-resolution rain-gauge observations. *Int. J. Climatol.,* **18,** 873–900.

——, ——, D. Lüthi, and H. C. Davies, 1998: Heavy precipitation processes in a warmer climate. *Geophys. Res. Lett.,* **25,** 1431–1434.

Gaffen, D. J., W. P. Elliott, and A. Tobock, 1992: Relationships between tropospheric water vapor and surface temperature as observed by radiosondes. *Geoph. Res. Lett.,* **19,** 1839–1842.

Geiger, H., J. Zeller, and G. Röthlisberger, 1991: Starkniederschläge des schweizereischen Alpen- und Alpenrandgebietes. WSL Special Report, 334 pp. [Available from Swiss Federal Institute for Forest, Snow and Landscape Research, Zürcherstr., CH-8903 Birmensdorf, Switzerland.]

Groisman, P. Y., and D. R. Legates, 1994: The accuracy of United States precipitation data. *Bull. Amer. Meteor. Soc.,* **75,** 215–227.

——, and Coauthors, 1999: Changes in the probability of heavy precipitation: Important indicators of climatic change. *Climatic Change,* **42,** 243–283.

Hennessy, K. J., J. M. Gregory, and J. F. B. Mitchell, 1997: Changes in daily precipitation under enhanced greenhouse conditions. *Climate Dyn.,* **13,** 667–680.

Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation: Regional temperature and precipitation. *Science,* **269,** 676–679.

——, and H. Van Loon, 1997: Decadal variations in climate associated with the North Atlantic Oscillation. *Climatic Change,* **36,** 301–326.

Iwashima, T., and R. Yamamoto, 1993: A statistical analysis of the extreme events: Long-term trend of heavy daily precipitation. *J. Meteor. Soc. Japan,* **71,** 637–640.

Jones, P. D., M. New, D. E. Parker, and I. G. Rigor, 1999: Surface air temperature and its changes over the past 150 years. *Rev. Geophys.,* **37,** 173–199.

Jones, R. G., J. M. Murphy, M. Noguer, and A. B. Keen, 1997: Simulation of climate change over Europe using a nested regional-climate model. II: Comparison of driving and regional model responses to a doubling of carbon dioxide. *Quart. J. Roy. Meteor. Soc.,* **123,** 265–292.

Karl, T. R., and R. W. Knight, 1998: Secular trends of precipitation amount, frequency and intensity in the United States. *Bull. Amer. Meteor. Soc.,* **79,** 231–241.

——, ——, and N. Plummer, 1995: Trends in high-frequency climate variability in the twentieth century. *Nature,* **377,** 217–220.

——, ——, D. R. Easterling, and R. G. Quayle, 1996: Indices of climate change for the United States. *Bull. Amer. Meteor. Soc.,* **77,** 279–292.

Kattenberg, A., F. Giorgi, H. Grassl, G. A. Meehl, J. F. B. Mitchell, R. J. Stouffer, T. Tokioka, A. J. Weaver, and T. M. L. Wigley, 1996: Climate models—Projections of future climate. *Climate Change 1995,* J. T. Houghton et al., Eds., Cambridge University Press, 285–357.

Keim, B. D., and J. F. Cruise, 1998: A technique to measure trends in the frequency of discrete random events. *J. Climate,* **11,** 848–855.

Kendall, M. G., 1955: *Rank Correlation Methods.* 2d ed. Charles Griffin, 196 pp.

Landsea, C. W., R. A. Pielke, A. M. Mestas-Nuñes, and J. A. Knaff, 1999: Atlantic basin hurricanes: Indices of climate change. *Climatic Change,* **42,** 89–129.

Lettenmaier, D. P., E. F. Wood, and J. R. Wallis, 1994: Hydro-climatological trends in the Continental United States, 1948–88. *J. Climate,* **7,** 586–607.

Liang, K.-Y., and S. L. Zeger, 1986: Longitudinal data analysis using generalized linear models. *Biometrika,* **73,** 13–22.

Mann, H. B., 1945: Nonparametric tests against trend. *Econometrica,* **13,** 245–259.

McCullagh, P., and J. A. Nelder, 1989: *Generalized Linear Models.* 2d ed. *Monogr. on Statistics and Appl. Probability,* No. 37, Chapman and Hall, 511 pp.

Meehl, G. A., F. Zwiers, J. Evans, T. Knutson, L. O. Mearns, and P. Whetton, 2000: Trends in extreme weather and climate events: Issues related to modeling extremes in projections of future climatic change. *Bull. Amer. Meteor. Soc.,* **81,** 427–436.

Mitchell, J. F. B., and W. J. Ingram, 1992: Carbon dioxide and climate: Mechanisms of changes in cloud. *J. Climate,* **5,** 5–21.

Müller, G., and J. Joss, 1985: Messnetze und Instrumente. *Der Niederschlag in der Schweiz.* Beitr. Geol. Schweiz - Hydrologie, Vol. 31, 17–47.

Osborn, T. J., M. Hulme, P. D. Jones, and T. A. Basnett, 2000: Observed trends in the daily intensity of United Kingdom precipitation. *Int. J. Climatol.,* **20,** 347–364.

Peterson, T. C., and Coauthors, 1998: Homogeniety adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol.,* **18,** 1493–1517.

Preisendorfer, R. W., 1988: *Principal Component Analysis in Meteorology and Oceanography.* Elsevier, 425 pp.

Rakhecha, P. R., and M. K. Soman, 1994: Trends in the annual extreme rainfall events of 1 to 3 days duration over India. *Theor. Appl. Climatol.,* **48,** 227–237.

Rebetez, M., 1999: Twentieth century trends in droughts in southern Switzerland. *Geophys. Res. Lett.,* **26,** 755–758.

Rice, J. A., 1995: *Mathematical Statistics and Data Analysis.* 2d ed. Duxburry Press, 602 pp.

Röthlisberger, G., 1998: Unwetterschäden in der Schweiz. WSL Rep. 346, 51 pp. [Available from Swiss Federal Institute for Forest, Snow and Landscape Res., Zürcherstro., CH-8903 Birmensdorf, Switzerland.]

Schär, C., C. Frei, D. Lüthi, and H. C. Davies, 1996: Surrogate climate-change scenarios for regional climate models. *Geophys. Res. Lett.,* **23,** 669–672.

——, T. D. Davies, C. Frei, H. Wanner, M. Widmann, M. Wild, and H. C. Davies, 1998: Current Alpine climate. *A View from the Alps: Regional Perspectives on Climate Change,* P. Cebon, et al., Eds., The MIT Press, 21–72.

Schiesser, H. H., C. Pfister, and J. Bader, 1997: Winter storms in Switzerland north of the Alps 1864/1865–1993/1994. *Theor. Appl. Climatol.,* **58,** 1–19.

Schmidt, H., and H. von Storch, 1993: German Bight storms analysed. *Nature,* **365,** 791.

Schönwiese, C.-D., J. Rapp, T. Fuchs, and M. Denhard, 1994: Observed climate trends in Europe 1891–1990. *Meteor. Z.,* **3,** 22–28.

Sevruk, B., 1994: Spatial and temporal inhomogeneity of global precipitation data. *Global Precipitations and Climate Change,* M. Desbois and F. Désalmand, Eds., NATO ASI Series I, Vol. 26, Springer, 219–230.

SMI, 1996: *Klima-90 Schlussbericht.* 228 pp. [Available from MeteoSwiss, Krähbühlstrasse 58, CH-8044 Zürich, Switzerland.]

Sneyers, R., 1990: On the statistical analysis of series of observations. WMO Tech. Note 143, WMO, 192 pp.

Stigler, S. M., 1986: *The History of Statistics.* Belknap Press, 410 pp.

Stone, R., N. Nicholls, and G. Hammer, 1996: Frost in Northeast Australia: Trends and influences of phases of the Southern Oscillation. *J. Climate,* **9,** 1896–1909.

Suppiah, R., and K. J. Hennessy, 1998: Trends in total rainfall, heavy rain events and number of dry days in Australia, 1910–1990. *Int. J. Climatol.,* **10,** 1141–1164.

Szinell, C. S., A. Bussay, and T. Szentimrey, 1998: Drought tendencies in Hungary. *Int. J. Climatol.,* **18,** 1479–1491.

Trenberth, K. E., 1999: Conceptual framework for changes of extremes of the hydrological cycle with climate change. *Climatic Change,* **42,** 327–339.

Venables, W. N., and B. D. Ripley, 1997: *Modern Applied Statistics with S-PLUS* 2d ed. Springer-Verlag, 462 pp.

Weber, R. O., P. Talkner, I. Auer, R. Böhm, M. Gajic-Capka, K. Zaninovic, R. Brazdil, and P. Fasko, 1997: 20th-century changes of temperature in the mountain regions of central Europe. *Climate Res.,* **36,** 327–344.

Weingartner, R., 1992: Niederschlagsmessnetze. *Hydrologischer Atlas der Schweiz,* Landeshydrologie und -geologie, EDMZ, Plate 2.1.

Whetton, P. H., A. M. Fowler, M. R. Haylock, and A. B. Pittock, 1993: Implications of climate change due to the enhanced greenhouse effect on floods and draughts in Australia. *Climatic Change,* **25,** 289–317.

Widmann, M., and C. Schär, 1997: A principal component and long-term trend analysis of daily precipitation in Switzerland. *Int. J. Climatol.,* **17,** 1333–1356.

Zwiers, F. W., and V. V. Kharin, 1998: Changes in the extremes of the climate simulated by CC GCM2 under CO_2 doubling. *J. Climate,* **11,** 2200–2222.