Indices for daily temperature and precipitation extremes in Europe analyzed for the period 1901–2000

Anders Moberg,[1,2] Philip D. Jones,[3] David Lister,[3] Alexander Walther,[4] Manola Brunet,[5] Jucundus Jacobeit,[6] Lisa V. Alexander,[7] Paul M. Della-Marta,[8] Jürg Luterbacher,[8] Pascal Yiou,[9] Deliang Chen,[4] Albert M. G. Klein Tank,[10] Oscar Saladié,[5] Javier Sigró,[5] Enric Aguilar,[5] Hans Alexandersson,[11] Carlos Almarza,[12] Ingeborg Auer,[13] Mariano Barriendos,[14] Michael Begert,[15] Hans Bergström,[16] Reinhard Böhm,[13] C. J. Butler,[17] John Caesar,[7] Achim Drebs,[18] Dmitra Founda,[19] Friedrich-Wilhelm Gerstengarbe,[20] Giusi Micela,[21] Maurizio Maugeri,[22] Hermann Österle,[20] Kreso Pandzic,[23] Michael Petrakis,[19] Lidija Srnec,[23] Radim Tolasz,[24] Heikki Tuomenvirta,[18] Peter C. Werner,[20] Hans Linderholm,[4] Andreas Philipp,[6] Heinz Wanner,[8] and Elena Xoplaki[8]

[1] We analyze century-long daily temperature and precipitation records for stations in Europe west of 60°E. A set of climatic indices derived from the daily series, mainly focusing on extremes, is defined. Linear trends in these indices are assessed over the period 1901–2000. Average trends, for 75 stations mostly representing Europe west of 20°E, show a warming for all temperature indices. Winter has, on average, warmed more (~1.0°C/100 yr) than summer (~0.8°C), both for daily maximum (TX) and minimum (TN) temperatures. Overall, the warming of TX in winter was stronger in the warm tail than in the cold tail (1.6 and 1.5°C for 98th and 95th, but ~1.0°C for 2nd, 5th and 10th percentiles). There are, however, large regional differences in temperature trend patterns. For summer, there is a tendency for stronger warming, both for TX and TN, in the warm than in the cold tail only in parts of central Europe. Winter precipitation totals, averaged over 121 European stations north of 40°N, have increased significantly by ~12% per 100 years. Trends in 90th, 95th and 98th percentiles of daily winter precipitation have been similar. No overall long-term trend occurred in summer precipitation totals, but there is an overall weak (statistically insignificant and regionally dependent) tendency for summer precipitation to have become slightly more intense but less common. Data inhomogeneities and relative sparseness of station density in many parts of Europe preclude more robust conclusions. It is of importance that new methods are developed for homogenizing daily data.

[1]Department of Meteorology, Stockholm University, Stockholm, Sweden.

[2]Now at Department of Physical Geography and Quaternary Geology, Stockholm University, Stockholm, Sweden.

[3]Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, UK.

[4]Regional Climate Group, Earth Sciences Centre, Göteborg University, Göteborg, Sweden.

[5]Climate Change Research Group, Physical Geography, Universitat Rovira i Virgili, Tarragona, Spain.

[6]Institute of Geography, University of Augsburg, Augsburg, Germany.

[7]Hadley Centre, Met Office, Exeter, UK.

[8]Institute of Geography, Climatology, and Meteorology and NCCR Climate, University of Bern, Bern, Switzerland.

[9]Laboratoire des Sciences du Climat et de l'Environnement, UMR Commissariat à l'Energie Atomique–Centre Nationale de le Recherche Scientifique, Gif-sur-Yvette, France.

[10]Royal Netherlands Meteorological Institute, De Bilt, Netherlands.

[11]Swedish Meteorological and Hydrological Institute, Norrköping, Sweden.

[12]Instituto Nacional de Meteorología, Madrid, Spain.

[13]Central Institute for Meteorology and Geodynamics, Vienna, Austria.

[14]Department of Modern History, University of Barcelona, Barcelona, Spain.

[15]Climate Services, Federal Department of Home Affairs, Federal Office of Meteorology and Climatology (MeteoSwiss), Zürich, Switzerland.

[16]Department of Earth Sciences, Uppsala University, Uppsala, Sweden.

[17]Armagh Observatory, Armagh, UK.

[18]Finnish Meteorological Institute, Helsinki, Finland.

[19]Institute for Environmental Research and Sustainable Development, National Observatory of Athens, Athens, Greece.

[20]Potsdam Institute for Climate Impact Research, Potsdam, Germany.

[21]Osservatorio Astronomico di Palermo, Palermo, Italy.

[22]Istituto di Fisica Generale Applicata, Universitá di Milano, Milan, Italy.

[23]Meteorological and Hydrological Service, Zagreb, Croatia.

[24]Czech Hydrometeorological Institute, Prague, Czech Republic.

1. Introduction

[2] The importance of assessing trends in weather extremes is often emphasized. The principal reason is that extreme weather conditions related to temperature, precipitation, storms or other aspects of climate, can cause loss of life, severe damage and large economic and societal losses [*Nutter*, 1999; *Changnon et al.*, 1999]. Some types of weather extremes are expected (from climate models) to become more frequent in the future because of anthropogenic influences on climate [e.g., *Kharin and Zwiers*, 2000; *Semenov and Bengtsson*, 2002; *Hegerl et al.*, 2004; *Groisman et al.*, 2005]. It is thus of great interest to analyze the occurrence of past extremes, to see if changes are already apparent. During the last 8 years or so, a number of assessments have been made of changes and trends in climate extremes in different parts of the world. The most recent world-wide assessments of changes in observed daily temperature and precipitation extremes have been made by *Alexander et al.* [2006] and, for precipitation only, by *Groisman et al.* [2005].

[3] *Alexander et al.* [2006] report widespread significant changes in temperature extremes for the period 1951–2003, especially those related to daily minimum temperatures. Changes in daily maximum temperature are less marked, implying that our world in many places has become less cold rather than hotter. Precipitation changes have been much less coherent than temperature changes, but annual precipitation has shown a widespread significant increase. *Groisman et al.* [2005] found disproportionate changes during the past decades in heavy and very heavy precipitation compared to the change in the annual and/or seasonal precipitation. Their results indicate an increasing probability of intense precipitation events for many extratropical regions.

[4] The relative sparseness of long digitally available records of daily temperature and precipitation measurements hampers analyses of observed changes in climate extremes. For this reason, assessments of changes in extremes [e.g., *Alexander et al.*, 2006; *Groisman et al.*, 2005] predating the 1950s are not possible in many parts of the world. Even for Europe, which is one of the most data-rich regions, it is difficult to assess changes in temperature and precipitation extremes over the whole of the past century. *Moberg and Jones* [2005] estimated such changes for Europe, but they were only able to use about 40 stations for precipitation and 30 for temperature. In fact, their analysis was essentially restricted to only some parts of central and western Europe. One of their main conclusions was that more digitized observational data from various European subregions are needed to permit a more spatially extensive analysis of changes in climate extremes over the whole of the last century.

[5] Efforts to improve the density and spatial coverage over Europe, for stations with daily temperature and precipitation records going back at least to 1901, have recently been made in the project "European and North Atlantic daily to multidecadal climate variability" (EMULATE).

This project has developed a database containing more than 200 daily station records. The purpose of this paper is to give an overview of the EMULATE temperature and precipitation database and to present some overall results from analyses of these data, with emphasis on the occurrence of extremes, but we also compare changes in extremes with changes in the mean. Substantial efforts have been made to quality control the data. Homogeneity issues are covered as much as practically possible, although no general homogeneity testing of all data has been performed.

[6] Another difficulty of assessing trends in weather extremes is related to the fact that extremes are rare events. The detection probability decreases the rarer the event [*Frei and Schär*, 2001; *Klein Tank and Können*, 2003]. For this reason, assessments of trends are often based on indices for rather "moderate" extremes, defined by the 90th or 95th percentiles in the distribution of daily observations. EMULATE calculated indices for temperature and precipitation extremes with standard software similar to those used in several other recent analyses of climatic extremes [e.g., *Alexander et al.*, 2006; *Klein Tank and Können*, 2003; *Klein Tank et al.*, 2006]. Some new indices are also defined and the EMULATE catalogue contains 64 climate indices (available at http://www.cru.uea.ac.uk/cru/projects/emulate/).

[7] In this paper, temperature and precipitation trends over the twentieth century are studied using a selection of 19 indices. Trends in different percentiles of the daily data distribution are analyzed for both winter and summer and are compared with trends in the means. Results are presented in a consistent way so that different indices can easily be compared across Europe, for example by expressing precipitation indices as percentages of the 1961–1990 averages. We should also mention that several other aspects of changes in temperature and precipitation extremes, including their relations to changes in atmospheric circulation patterns and sea surface temperatures, will be addressed in other papers originating from EMULATE (in preparation; many results are available on the project website).

2. Data Set
2.1. Data Collection

[8] A database with 230 stations having daily temperature and/or precipitation series starting before 1901 has been established (only a few stations start a few years later). Most stations are located in Europe, but about 30 are located in the Asian part of the former Soviet Union. The data can be sorted into three categories depending on the source: (1) publicly available databases, (2) original data digitized within EMULATE, and (3) data obtained through personal contacts with various holders of digital data that were not publicly available. The data elements collected are daily maximum temperatures (TX), daily minimum temperatures (TN) and daily precipitation totals (PREC). Many station series also include daily mean temperatures (TMEAN).

[9] There were four principal publicly available data sources. One is the Web site of the European Climate
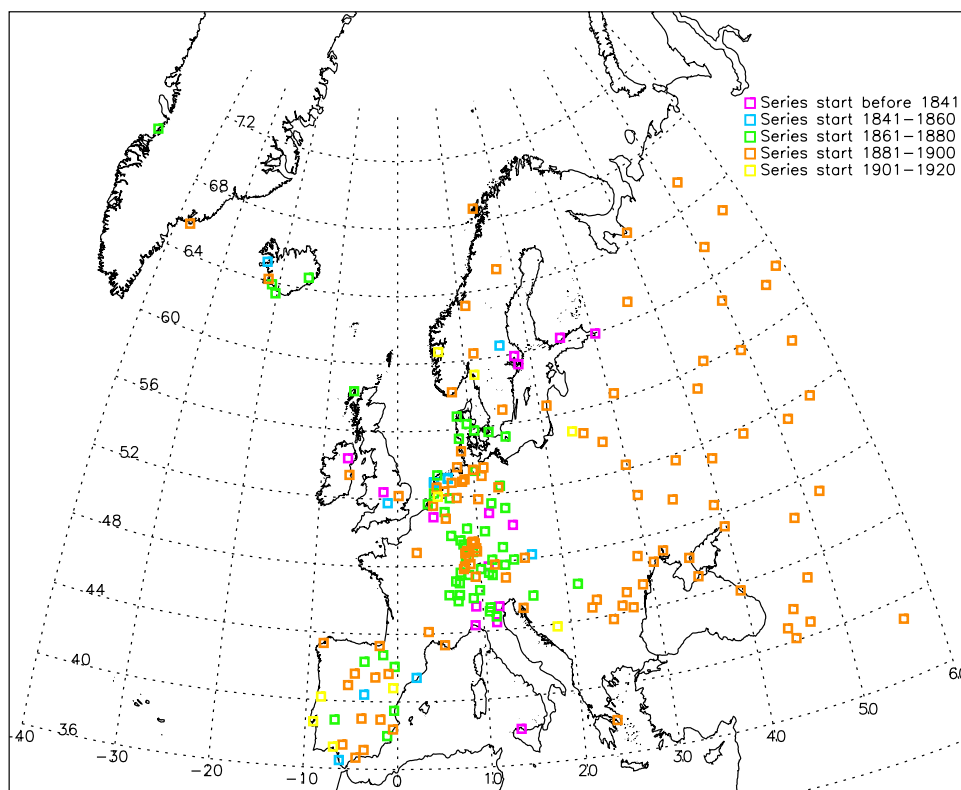
Figure 1. Locations of all 195 stations west of 60°E with daily temperature or precipitation series in the EMULATE database. The start period for each series is indicated by the colors.

Assessment & Data set (ECA&D; http://eca.knmi.nl) [*Klein Tank et al.*, 2002], with data from several European countries (67 stations used here). The second is the Carbon Dioxide Information Analysis Center (CDIAC) providing temperature data from countries in the former Soviet Union. These data [*Razuvaev et al.*, 1993] were updated for use by *Kiktev et al.* [2003] and again for the study by *Alexander et al.* [2006]. Stations compiled for *Alexander et al.* [2006] are used in this study. The third source is the National Climatic Data Center, from which we obtained precipitation data from countries in the former Soviet Union [*Groisman et al.*, 2005]. We used 64 stations from CDIAC and NCDC together. The fourth source is Deutscher Wetterdienst (25 German stations). In addition, eight station series were obtained from published CD-ROMs, including seven from the IMPROVE project [*Camuffo and Jones*, 2002] and one from the EARTHINFO CD (http://www.earthinfo.com). Within EMULATE, 22 long daily records from Spain have been developed directly from the original sources [*Brunet et al.*, 2006]. Finally, data for about 50 stations were provided by individual scientists or national meteorological services in the following countries: Austria, Croatia, Czech Republic, Finland, Germany, Greece, Iceland, Italy, Spain, Sweden, Switzerland and the U.K.

2.2. Station Network

[10] Among the station files, 223 have PREC, 178 have TX, and 169 have TN data. Figure 1 shows the geographical distribution and start period of all 195 stations located west of 60°E, which is the area of interest in this paper. Ten very long station records start before 1801.

[11] The network is fairly dense across parts of Europe, particularly over parts of central and western Europe (the Netherlands, Germany, Austria and Switzerland). Compared to the previous analyses of century-long European daily temperature and precipitation series [*Moberg and Jones*, 2005], EMULATE makes a substantial improvement to the geographical distribution and density of stations. Notable improvement is achieved over Spain, Russia and some other former Soviet Union countries. There are, however, still areas where station density could potentially be much improved (France, the Fennoscandian countries, the British Isles, some parts of eastern Europe and the central and eastern Mediterranean region), but this would require further digitizing of daily data that so far only exist in the form of printed year books and other kinds of written documents. The amalgamation of already digitized data from other projects would also help to extend the station network for future studies. For example, many long Italian precipitation records, not available here, have recently been digitized [*Brunetti et al.*, 2004].

2.3. Completeness of Data Series

[12] Figures 2a–2c show how the number of stations varies with time, separately for TX, TN and PREC data. The black lines show the number of stations with nearly complete data in each year (<4 missing days per 3-month season). This number increases from just a few stations in 1840 to rather high values after around 1900, although the highest numbers are reached in the period 1950–1990. Thus, even though all stations have some data back to around 1901, the graphs show that many records are not
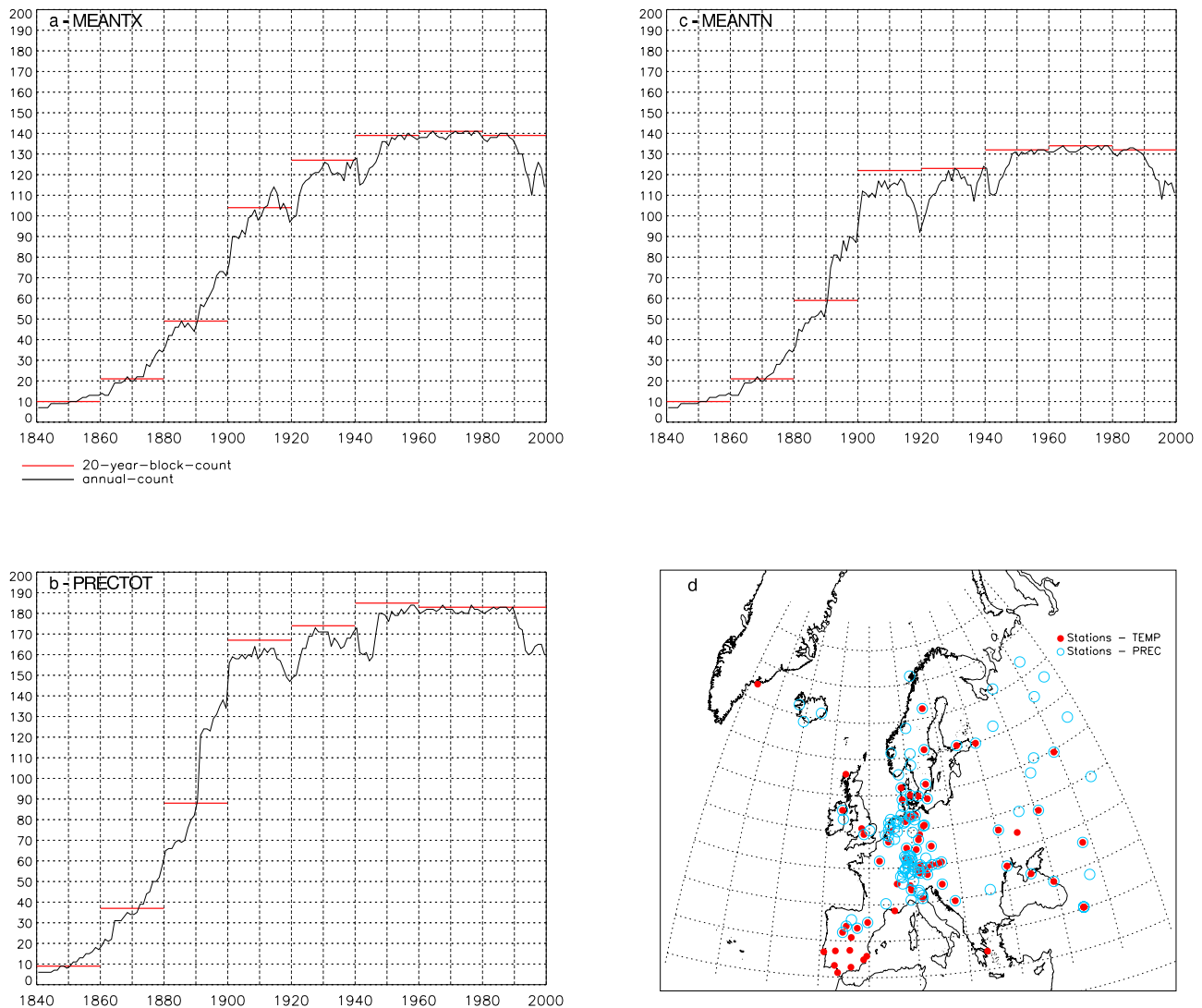
Figure 2. Overview of the completeness of daily station data. (a) Number of stations (black lines) in the EMULATE database with nearly complete daily data in each year (<4 missing days per 3-month season) for TX as a function of time. (b and c) Same for PREC and TN. The red lines in Figures 2a–2c show the number of stations having at least 10 years with nearly complete data per 20-year block. (d) Geographical distribution of stations west of 60°E that pass certain criteria required for calculation of linear trends both for temperature (75 stations) and precipitation (121 stations) over the period 1901–2000. These criteria are described in the methods section. Latitudes and longitudes are given in Figure 1.

complete in the first half of the twentieth century. The drop in station numbers after around 1990 is mainly because we were unable to update many records from the former Soviet Union. Figures 2a–2c also illustrate the completeness for blocks of 20 years, represented by the red lines. These show the number of stations with at least 10 years of complete or nearly complete data (using the same criterion as above) in each block.

[13] This analysis reveals a notable jump in station numbers between 1881–1900 and 1901–1920. After 1901, the number of stations in the 20-year blocks is always high for TN and PREC, but somewhat lower values for TX in the block 1901–1920 reveal that many stations have missing TX data in this period. These stations are mainly located in the former Soviet Union. Despite this, we consider the time period 1901–2000 appropriate for achiev-

ing a long study period and also both a wide geographical spread and high density of stations. All trend analyses in this paper are calculated over this period. Other papers originating from EMULATE will, however, also analyze data from before 1901.

[14] Figure 2d shows the distribution of 75 stations with temperature data and 121 stations with precipitation data that pass certain further criteria for data completeness and other data properties required for trend calculations over the period 1901–2000 (these criteria are described in the methods section). It is clear, from comparison with Figure 1, that the number of stations that can be used for trend analyses is reduced once these criteria are imposed. Two particular features concerning the distribution of stations in Figure 2d are worth pointing out. First, the reason why so few temperature stations appear over the former

Soviet Union is related to the incomplete TX data. Figure 2d only shows temperature stations where both TX and TN data pass the completeness criteria. Second, the few precipitation stations on the southern half of the Iberian peninsula are, on the contrary, not related to lack of data. It is instead the very dry climate in summer that causes difficulties in calculating trends for some indices chosen for this study. For our Europe-average trend analyses we required that each precipitation index must have an average exceeding 1 mm for the 1961–1990 period. (The reason for this is described later.) Many stations in the Mediterranean region do not fulfill this criterion in summer.

2.4. Quality and Homogeneity Controls

[15] Basic quality controls (QC) have been undertaken for all series. For most stations in countries in the former Soviet Union we mainly rely on the QC made by *Alexander et al.* [2006] for temperature data and by *Groisman et al.* [2005] for precipitation data. For all other stations, the QC undertaken during this study is described in Appendix A. Additional QC of Spanish data digitized within EMULATE have also been made by the Climate Change Research Group at the university in Tarragona [*Brunet et al.*, 2006].

[16] No overall homogeneity tests have been applied to the entire EMULATE database. To accomplish such an extensive task would require large resources in terms of personnel with detailed local knowledge of the station histories in all countries encompassed by the database. Another, more intrinsic, difficulty is the spatial heterogeneity of daily precipitation. If the decorrelation radius is ~40 km for daily precipitation data, which has been found for the Alpine region [*Auer et al. 2005*], then homogenization based on methods that require nearby reference stations is in principle impossible given the station density for large parts of our database.

[17] There is, furthermore, still a need to develop methods for homogeneity testing of daily climate data. So far, most homogeneity testing of long instrumental records has been made on monthly (sometimes seasonal or even annual) series. Methods have been developed for interpolating monthly adjustment factors to a daily resolution [*Moberg et al.*, 2002; *Vincent et al.*, 2002]. However, interpolating monthly adjustments to daily values does not guarantee that all aspects of data inhomogeneities are properly adjusted. A station relocation, for example, may very well influence the shape of data distribution and not simply involve a shift in the location of the mean. Considerable effort could be expended on developing homogenization tools that can account for such inhomogeneities before we can analyze reliably all aspects of climatic changes using daily data. In particular, in the context of analyzing changes in the occurrence of extremes, homogeneity testing methods that could be applied to the entire data distribution would be highly useful.

[18] Steps toward the development of such methods have been undertaken within EMULATE [*Della-Marta and Wanner*, 2006]. Their method and the one of *Trewin and Trevitt* [1996] are specifically designed to homogenize the second- and third-order moments (i.e., the variance and skewness) of daily temperature series. Currently, 26 long-term TX series for stations mainly in central western Europe have been homogenized using the *Della-Marta and Wanner* [2006]

technique and appear in the work of P. M. Della-Marta et al. (Summer heat waves over Europe since 1880, their changes and relationship to large scale forcings, submitted to *Climate Dynamics,* 2006). These homogenized series were, however, not included here because of the data not being available at the time of writing.

[19] The fact that no overall homogeneity tests have been made here does not mean that the homogeneity status of the data series are entirely unknown. In particular, the Spanish temperature series have been extensively quality controlled, adjusted for biases related to different properties of early and modern thermometer screens and have also been subject to statistical homogeneity tests and homogenization [*Brunet et al.*, 2006]. In this case, the *Vincent et al.* [2002] approach was adopted to interpolate monthly corrections to daily resolution. Precipitation series from the former Soviet Union had also been adjusted for changes in rain gauge types and observing conditions [*Groisman et al.*, 1991, 2005]. These adjustments are most substantial for winter precipitation. Homogeneity of the ECA data has been assessed and documented by *Wijngaard et al.* [2003]. However, ECA did not correct or reject any data, but rather flagged those stations thought to be suspect. Both TMEAN and PREC data have been homogenized for six of the Swiss stations used [*Begert et al.*, 2005], using interpolation of monthly adjustments to daily resolution, but no TX or TN data were homogenized. Also, two stations in Sweden have had their TMEAN series homogenized [*Bergström and Moberg*, 2002; *Moberg et al.*, 2002], but not TX, TN and PREC. The Spanish precipitation series have been quality controlled but not yet homogenized (work is in progress). Most records provided by data holders in various countries had been subject to routine quality controls before being provided to EMULATE, but generally no homogenization had been undertaken.

[20] The EMULATE database, even if far from fully homogeneous, is the best currently available collection of European daily temperature and precipitation records with long series starting in 1901 or earlier. The experience from numerous previous homogeneity assessments [e.g., *Groisman et al.*, 1991; *Moberg and Alexandersson*, 1997; *Nordli et al.*, 1997; *Böhm et al.*, 2001; *Tuomenvirta*, 2001; *Wijngaard et al.*, 2003; *Auer et al.*, 2005; *Begert et al.*, 2005], however, clearly reveals that unhomogenized data often contain both abrupt breaks and gradual artificial changes for various reasons. Therefore we stress that the exact numerical values of trends (or other measures of climatic change) estimated from the EMULATE data set (or any other not fully homogenized data set) should be treated with care.

3. Climate Indices

[21] EMULATE has defined 64 climate indices derived from the daily temperature and/or precipitation series. Most of these indices measure some aspect of climate extremes, while a few give information about mean conditions. Software has been developed and used to calculate, for each station, time series of each index for the four traditional climatological seasons December to February (DJF), March to May (MAM), June to August (JJA) and September to November (SON), and also for all twelve 2-month

Table 1. List of Climate Indices Used in This Study

Index Name	Explanation	Unit
Temperature Indices		
TN2P, TX2P	2nd percentile of daily TN or TX	°C
TN5P, TX5P	5th percentile of daily TN or TX	°C
TN10P, TX10P	10th percentile of daily TN or TX	°C
MEANTN, MEANTX	Mean of daily TN or TX	°C
TN90P, TX90P	90th percentile of daily TN or TX	°C
TN95P, TX95P	95th percentile of daily TN or TX	°C
TN98P, TX98P	98th percentile of daily TN or TX	°C
Precipitation Indices		
PRECTOT	precipitation total	mm
SDII	simple daily intensity index	mm
	(average precipitation per wet day, i.e.,	
	per day with precipitation >1 mm)	
PREC90P	90th percentile of daily PREC	mm
	(percentile defined on the basis of	
	all days in a season,	
	i.e., including both wet and dry days)	
PREC95P	95th percentile of daily PREC	mm
PREC98P	98th percentile of daily PREC	mm

seasons December to January, January to February, etc. In this paper, we confine our analyses to the JJA and DJF seasons. Data files containing all index time series for each station are available on the EMULATE web site (http://www.cru.uea.ac.uk/cru/projects/emulate/, with exception of a few stations with limited restrictions imposed by the data providers) together with software documentation, information about the file structure and definitions of all indices, and a "trend atlas" which shows results from calculations of long-term trends in all 64 indices.

[22] Different research groups have used somewhat different index definitions. The most commonly used index software at present is probably RClimDex (http://cccma.seos.uvic.ca/ETCCDMI/software.html), which was developed on behalf of the World Meteorological Organization Commission for Climatology (CCl)/World Climate Research Programme (WCRP) project on Climate Variability and Predictability (CLIVAR) Expert Team on Climate Change Detection, Monitoring and Indices (ETCCDMI). This software calculates a total of 27 indices derived from TX, TN and PREC. A dictionary of 40 indices is also provided on the ECA&D web site, giving an overview of which indices have been used by different research groups. EMULATE has strived to define indices that are comparable with those from other groups. However, because EMULATE has defined a larger total number of indices than used elsewhere, it has not always been possible to use index names and definitions identical to those used by others. Although this may confuse some readers who are already familiar with other index names, the index names are chosen to obtain an internally consistent terminology. Most of the indices involve estimations of percentiles. (We should mention that EMULATE has applied the bootstrapping method advised by *Zhang et al.* [2005], to avoid inhomogeneities at the boundaries of baseline periods for all indices that count the number of occurrences above (or below) a percentile-based threshold).

[23] Previous studies dealing with changes in climate extremes have often paid relatively little attention to comparing trends in different percentiles. Rather, investigators

have often studied changes in only the 10th/90th or only the 5th/95th percentiles. In this paper, we instead compare changes in several percentiles with those in the mean. To achieve this, a subset of fourteen temperature (seven each for TX and TN) and five precipitation indices is studied. These indices are listed with their names, brief explanations and units in Table 1. Most of these indices measure how a certain percentile in the empirical distribution of daily data in a given season varies with time. We use the percentiles 2, 5, 10, 90, 95 and 98 for temperatures, but only the upper percentiles 90, 95 and 98 for precipitation. Thus we can assess changes in both the cold and warm tail of the temperature distributions, whereas only the upper tail is examined for precipitation. The percentiles for precipitation data are estimated from samples containing all days in a season, i.e., including both dry and wet days. Names for the chosen percentile indices are of the style TX2P, TX5P, TX10P, etc. NB! The same index names have other meanings in the RClimDex and ECA&D terminology, where they denote percentile-based day count indices (expressed as percentages) rather than the temporal variation of percentiles. Our choice of temperature indices makes comparison of trends in the cold and warm tails with the trend in the mean more easily interpreted than the day-count indices, as the unit in all our indices is °C. Because we are analyzing 3-month seasons (~90-day seasons), the values of TX98P, TX95P and TX90P roughly correspond to the daily maximum temperature on the day with the second, fifth and ninth highest value, respectively, among all days in the season analyzed (with analogous relationships for all the other temperature and precipitation percentile indices).

[24] We also use some indices for average conditions. For temperatures, these are simply the arithmetic averages of the TX or TN data in each season (MEANTX and MEANTN). As a measure of average precipitation conditions we choose the commonly used "simple daily intensity index" (SDII), which gives the average precipitation per wet day during a season, where a wet day is defined as having a daily precipitation of at least 1 mm. In addition, we study the seasonal precipitation total (PRECTOT). If different trends are observed for PRECTOT and SDII, this would reflect changes in the character of precipitation. For example, a positive trend in SDII and a zero PRECTOT trend would imply that precipitation has become more intense, but also less frequent as the total amount is not changed. A positive PRECTOT trend and a zero SDII trend would instead imply that precipitation falls more often but with unchanged intensity.

4. Analysis Methods

[25] Temperature and precipitation changes over the 20th century are measured by estimating the linear trends over the period 1901–2000. The trend estimator used is the ordinary least squares (OLS) method. There are certainly various robust nonparametric trend estimators that also would suit well, but we decided to use the OLS method here. In the previous similar study by *Moberg and Jones* [2005], trends were estimated with both the OLS method and a robust nonparametric method. Their experience was that the size of the estimated trends was very similar with both methods. This is in line with findings by *Cohn and*

Lins [2005], who concluded that the trend magnitude can often be determined with little ambiguity, while the corresponding statistical significance is less certain. *Cohn and Lins* [2005], however, emphasize the importance of taking account of the long-term persistence when testing the significance of trends in hydroclimatological time series. *Kiktev et al.* [2003], in a study comparing modeled and observed trend in indices of daily climate extremes, used the OLS estimator in conjunction with a bootstrap technique for significance testing that accounts for serial correlation. They also used two alternative methods for assessing trend significance, a Mann-Kendall test and a linear model that allowed for red noise. They found that results with all three significance testing methods agree well. Moreover, *Moberg and Jones* [2005] used the same bootstrapping technique as *Kiktev et al.* [2003] together with both the OLS and the robust estimator. They found that sometimes more stations reach significance with one method and sometimes with the other. However, significance was generally reached more easily with the OLS method because of the smaller variance of this estimator.

[26] Although different methods for trend estimation and significance testing are in use, there is no universally accepted best technique. It seems that taking the autocorrelation into account in significance testing is more important for the overall results than the choice of trend estimator or method to assess significance of climatological time series. On the basis of these findings by *Cohn and Lins* [2005], *Kiktev et al.* [2003] and *Moberg and Jones* [2005], we decided to use the OLS estimator and to assess the significance of trends using a two-tailed t-test where the degrees of freedom are reduced to account for serial correlation (lag-1 autocorrelation). The 5% level is chosen to determine if a trend is significantly different from zero.

[27] Trend analysis is applied to the data in three different ways to each temperature and precipitation index. First, we are interested in an overall analysis, where we compare Europe-average trends in the means with trends in the percentiles. Second, we study the spatial distribution of statistically significant trends for different indices. Third, we compare the pattern of trends in the various indices in six selected subregions.

[28] The number of stations was allowed to vary somewhat between the three trend analyses. In the first case, we applied the most strict criterion of data completeness. For this overall analysis, we required that all stations must pass the data completeness criteria for all indices analyzed and for both winter and summer (although separately for temperature and precipitation data). The reason for using the most strict criterion for the overall analysis is that we wanted to use the same set of stations for all different indices. This ensures that average trends in different indices can be directly compared.

[29] Using the completeness criteria described below, the overall analysis could be undertaken for the 75 temperature and 121 precipitation stations indicated in Figure 2d. For these stations, we calculated the average trends for each index to obtain Europe-wide averages. We also provide estimates of the uncertainties in these averages, by calculating 95% confidence intervals. These are obtained in the classical way using the standard deviations (of the individual trend estimates at each station and for each index) and the t-distribution.

[30] For the other two trend analyses, we could loosen the data completeness criteria somewhat to allow more stations to be used. Therefore maps showing trends in indices based on TN data include more stations in countries in the former Soviet Union than maps showing TX data, because of the earlier mentioned incompleteness of former Soviet Union TX series. Similarly, maps with winter precipitation trends include more stations on the Iberian peninsula than those shown in Figure 2d.

[31] Here we describe the criteria for data completeness: First, we required that a station had sufficient daily values during the season-in-question to produce seasonal index values (<4 missing days per season). Then, we went further by checking that each 20-year block (1901–1920, 1921–1940,.., 1981–2000) is complete enough. We used a stricter threshold for the first and last block (at least 15 years complete according to the 4-day per season criterion) than for the intermediate blocks (which must have at least 10 years complete).

[32] Comparison of trends for different stations and indices is straightforward for the temperature indices, as the numerical values (unit is °C/100 yr) of each index series are always of about the same size. For precipitation, however, trends expressed in mm vary considerably among different sites and different indices. Therefore precipitation trends are expressed as percentages of the 1961–1990 climatological average for each respective index. To account for the problem with very rare precipitation events at some stations, which can sometimes lead to zero (or very near zero) climatological averages of the precipitation percentile indices, we additionally require that the 1961–1990 seasonal mean values of the precipitation indices must exceed 1 mm for inclusion in the trend analyses. If we had not used this additional criterion, it would have been impossible for some stations (and quite meaningless for other stations) to compare the precipitation trends for different indices, when these are expressed as percentages of the mean for the 1961–1990 period.

[33] An additional analysis is undertaken to see how similar, or how different, the index time series for mean conditions are compared with those for the various percentiles. To obtain a simple measure, we used the linear (Pearson) correlation coefficient. Correlations have been calculated, for each station, between MEANTX and all the TX percentile indices, and similarly between MEANTN and all TN percentile indices. For both PRECTOT and SDII, separately, correlations have been calculated with the three (upper) percentile precipitation indices.

5. Results

[34] The overall results of trend estimates, for the stations indicated in Figure 2d, are summarized in Figures 3–5. It should be noted that these overall results are regionally biased toward those parts of Europe where the station density is highest. The temperature stations rather well represent most areas west of 20°E except northern Fennoscandia. There are only a few temperature stations east of 20°E, mainly located south of 52°N. The precipitation stations better represent Europe as a whole, except for the
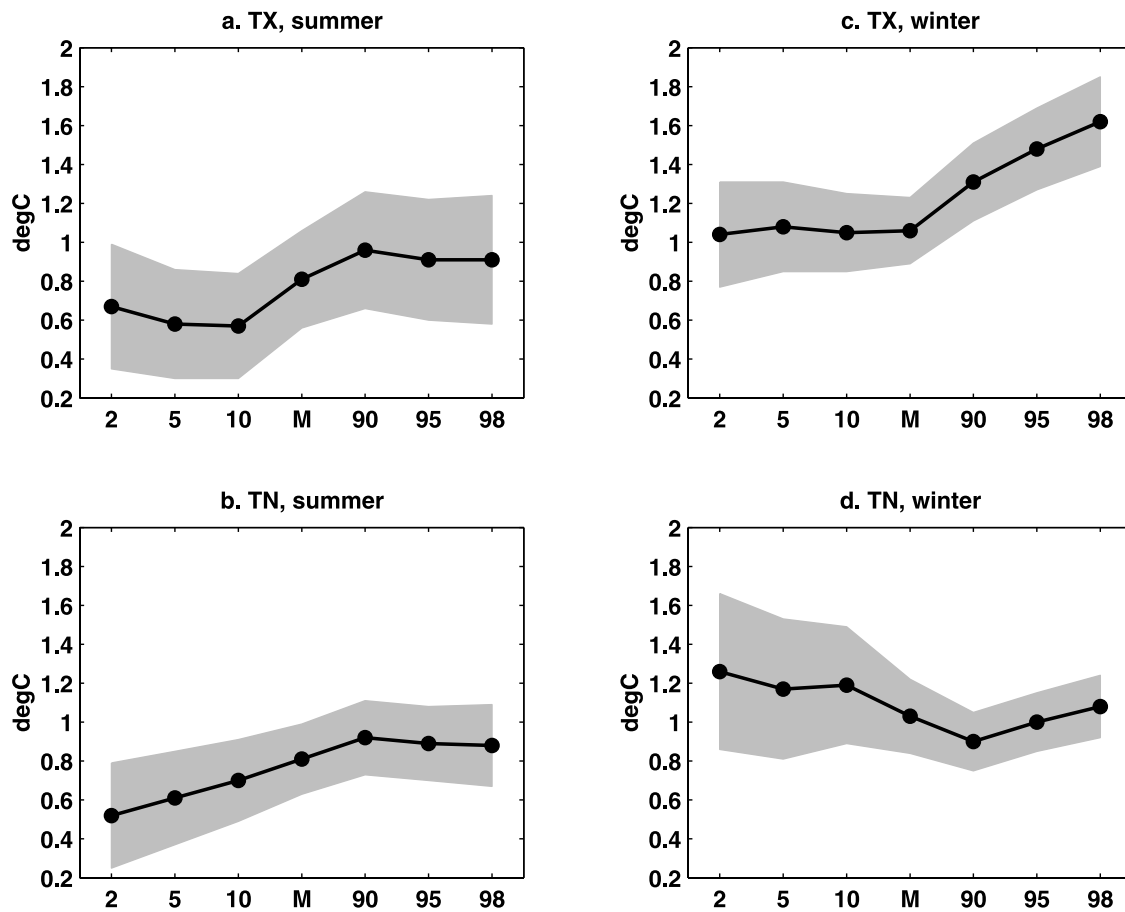
Figure 3. Average trends(black dots) in temperature indices over the 1901–2000 period (unit is °C/ 100 yr). The averaging is made over individual trends at 75 stations located west of 60°E (see Figure 2d). 95% confidence intervals for the average trends are shown with grey shading. Trends are shown for (a) TX in summer, (b) TN in summer, (c) TX in winter, and (d) TN in winter. Summer (winter) is the June–August (December–February) period. The labels on the horizontal axes refer to TX2P, TX5P, TX10P, MEANTX, TX90P, TX95P and TX98P in Figures 3a and 3c, and similarly to TN2P, etc., in Figures 3b and 3d. Definitions of these indices are given in the main text and Table 1.

part south of 40°N. Analyses of correlations between the percentile indices and the indices for mean conditions are shown in Figures 6 and 7. Maps showing the size of trends at all stations (passing the data completeness criteria) where trends are significant have been produced and inspected for each index. A selection of these maps is shown to give an overall view of the main findings regarding the spatial pattern of trends (Figures 8–10). Furthermore, locations of the stations used to define six regional series are indicated in Figure 11. Time series from these six regions are plotted for a few selected indices to illustrate their temporal evolution (Figures 12–14). Finally, Figure 15 compares the pattern of average trends in different indices for the six regions.

5.1. Overall Temperature Trends

[35] Figure 3 shows the average trends over the period 1901–2000 in all temperature indices for the 75 stations. The average trends are positive (i.e., warming) for all indices. Winter has warmed more than summer; MEANTN and MEANTX warmed on average about 1.0 to 1.1°C/100 yr in winter and about 0.8°C/100 yr in summer.

[36] Given that no overall homogeneity tests have been applied to the entire data set, it is relevant to ask whether these observed average trends are reliable before inspecting the results in more detail. Therefore we compared with European summer and winter mean temperature data extracted from the widely used 5° × 5° (latitude, longitude) gridded monthly temperature data set CRUTEM2v [*Jones and Moberg*, 2003]. The rationale for this comparison is that the European subset of CRUTEM2v contains a much larger number of series than those analyzed here, and that *Jones and Moberg* [2003] included as many station series as possible that had been homogenized by national meteorological services in different countries. It is therefore expected that the European subset of CRUTEM2v shows reliable trends in mean temperatures. Hence a similar average trend in the EMULATE series would indicate an overall degree of homogeneity of the data.

[37] CRUTEM2v data from the region 35–60°N, 10°W–50°E were used, which corresponds quite well to the area where a large majority of the EMULATE stations analyzed are located. The warming trend during 1901–2000 in this subset of CRUTEM2v is 0.9°C for winter and
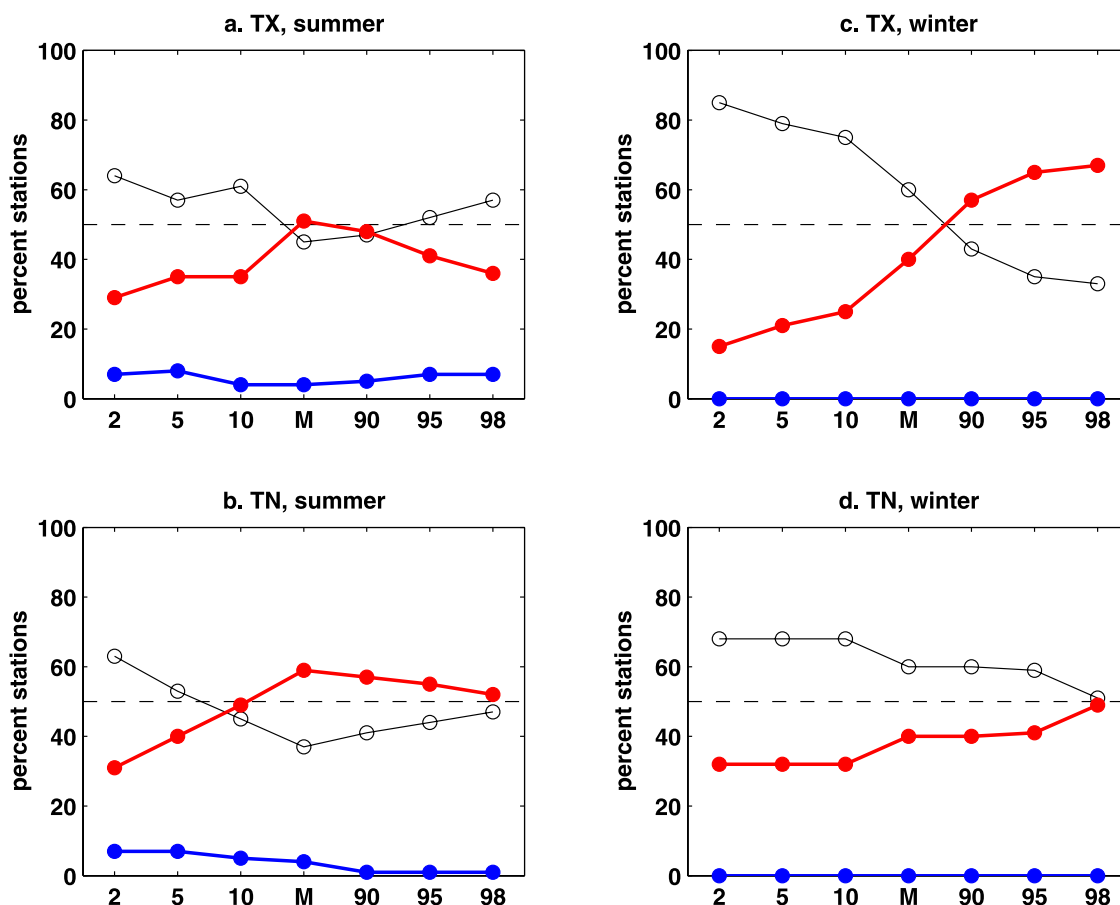
Figure 4. Fractions (in percent) of the 75 stations analyzed in Figure 3 that have significant (at the 5% level) and insignificant trends in temperature indices. Red indicates significant warming trends. Blue indicates significant cooling trends. Black indicates insignificant trends. The order of subplots and labels on the horizontal axes are as in Figure 3.

0.6°C for summer. These numbers are within the 95% confidence intervals (although close to the lower part of the range) for the estimated trends in EMULATE MEANTX and MEANTN. Hence we judge the difference between the two data sets to be sufficiently small (in particular given the difficulties of comparing unweighted station data with gridded data) to say that trends in the EMULATE temperature data are on average reliable for the purposes of this study. This does, however, not imply that trends at each individual station are correct.

[38] A notable finding (in Figure 3) is the particularly strong warming trend for TX98P in winter (1.62°C). The lower limit of the confidence interval for this index is even above the upper limits for the confidence intervals both for MEANTX and for all three lower TX percentile indices (which all have mean trends between 1.04 and 1.08°C). This implies an asymmetric warming of winter TX data, with a stronger warming in the warm tail than in the cold tail. The warming of TX95P is only slightly smaller (1.48°C) than that for TX98P.

[39] For summer, both TX and TN show a similar tendency for asymmetric warming (stronger in the warm than in the cold tail). However, the confidence intervals for high and low percentiles overlap. Hence we cannot conclude that an asymmetric warming in summer has been

significantly detected. TN trends in winter suggest a tendency of asymmetric warming of the opposite kind (stronger in the cold than in the warm tail). Confidence intervals for the cold-tail percentiles, however, are very wide. This implies large differences among stations. This behavior is further illustrated later on, when we present results for the subregions.

[40] Figure 4 reveals that there are no stations with significant cooling trends for any index in winter, and that only a few percent of the stations have significant cooling trends in summer (for all indices). Significant warming station trends are, on the other hand, quite common for all indices. This complements the information from Figure 3 that warming has occurred on average for all percentiles. Furthermore, the fraction of stations with significant warming trends for TX indices in winter is strongly dependent on the percentile. This fraction increases monotonically from 15% for TX2P to 65% for TX98P. Thus the strong average warming of TX98P in winter (seen in Figure 3c) is associated with the property that a majority of stations have significant warming trends for this index.

5.2. Overall Precipitation Trends

[41] The average trends in precipitation indices for the 121 precipitation stations (Figure 2d) are illustrated in
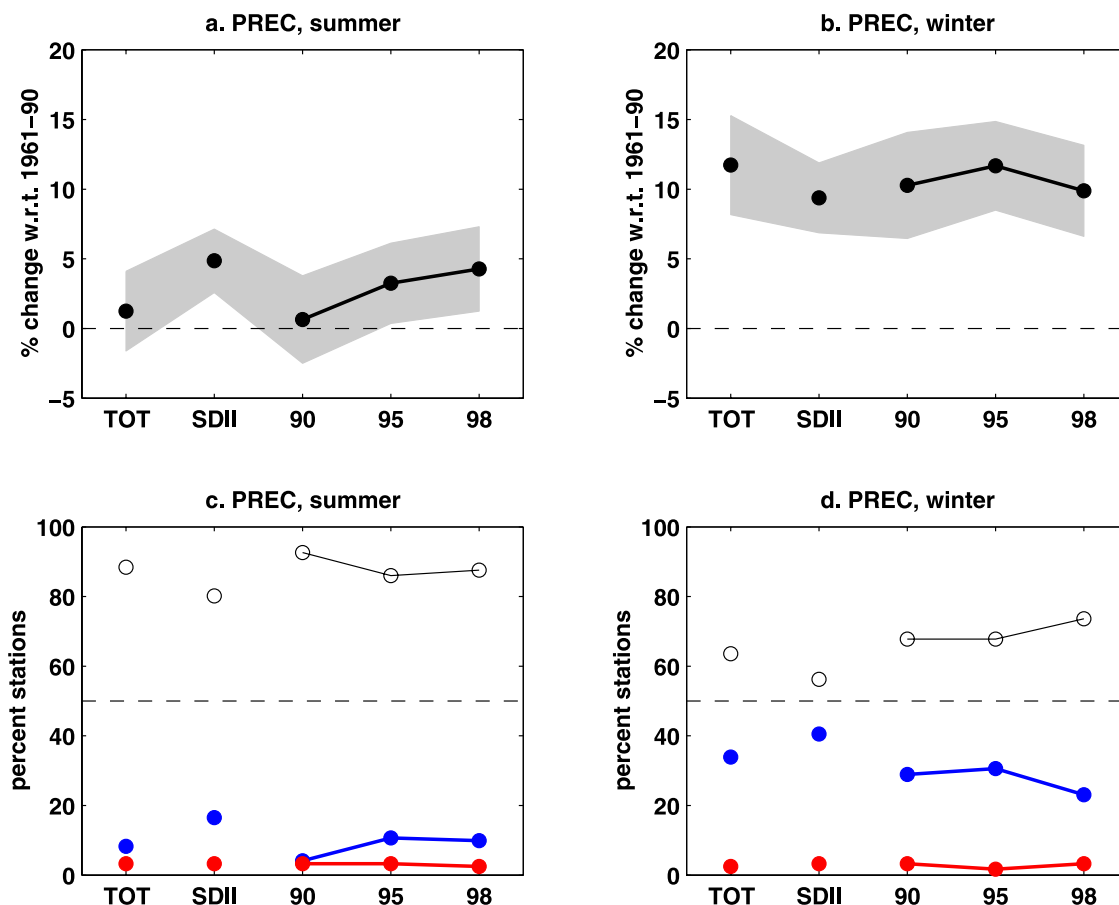
Figure 5. Overview of trends in precipitation indices over the 1901–2000 period at 121 stations west of 60°E (see Figure 2d). The trends averaged over the 121 stations are shown (black dots) for (a) summer and (b) winter (unit is % change with respect to the 1961–1990 average for each respective index), together with their 95% confidence intervals (grey shading). The fractions (in percent) of the 121 individual stations that have significant (at the 5% level) and insignificant trends are given for (c) summer and (d) winter. Red indicates significant drying trends. Blue indicates significant wetting trends. Black indicates insignificant trends. The labels on the horizontal axes refer to PRECTOT, SDII, PREC90P, PREC95P and PREC98P. Definitions of these indices are given in the main text and Table 1.

Figures 5a and 5b. Precipitation has increased significantly in winter (Figure 5b), with confidence intervals for the average trends being well above zero for all five indices. Average precipitation trends for summer are small (Figure 5a); confidence intervals for two of the indices (PRECTOT and PREC90P) even cross the zero level. When expressed as % change in relation to the 1961–1990 normals, the average wetting trends in winter vary between 9 and 12% among the indices (12% for PRECTOT). In summer the average wetting trends vary between 1 and 5% (1% for PRECTOT).

[42] To make a rough estimate of the overall reliability of these precipitation trends, in a similar manner to that for temperatures, we compare with trends estimated from published gridded precipitation totals. We used the 5° × 5° gridded monthly precipitation data "g55wld0098.dat" [*Hulme*, 1994, 1996] (updated to 1998 on http://www.cru.uea.ac.uk/~mikeh/datasets/global/). Because of the weak spatial coherence of precipitation, some experimentation was made to select and weight precipitation grid point data such that they geographically correspond to the

locations of the EMULATE stations. This exercise returned a positive trend of 8% for winter and 2% for summer. The summer trend is well within the 95% confidence interval for PRECTOT in the EMULATE data, whereas the winter trend in the gridded data is just at the lower end of the range. We conclude that the EMULATE data and the gridded data agree quite well with respect to the overall summer and winter precipitation trends during the 20th century.

[43] The larger average wetting trends in winter than summer are associated with substantially more stations having significant wetting trends in winter. Between 23 and 41% of the stations have significant wetting trends in winter (Figure 5d). Only between 4 and 17% of the stations have significant wetting trends in summer (Figure 5c). Very few significant drying trends (2–3% of the stations) are found both in winter and summer (which is about the number expected by chance if the true trend is zero). As many as 80–93% of the stations have insignificant trends in summer, which further emphasizes the absence of long-term trends in European summer precipitation over the past century.
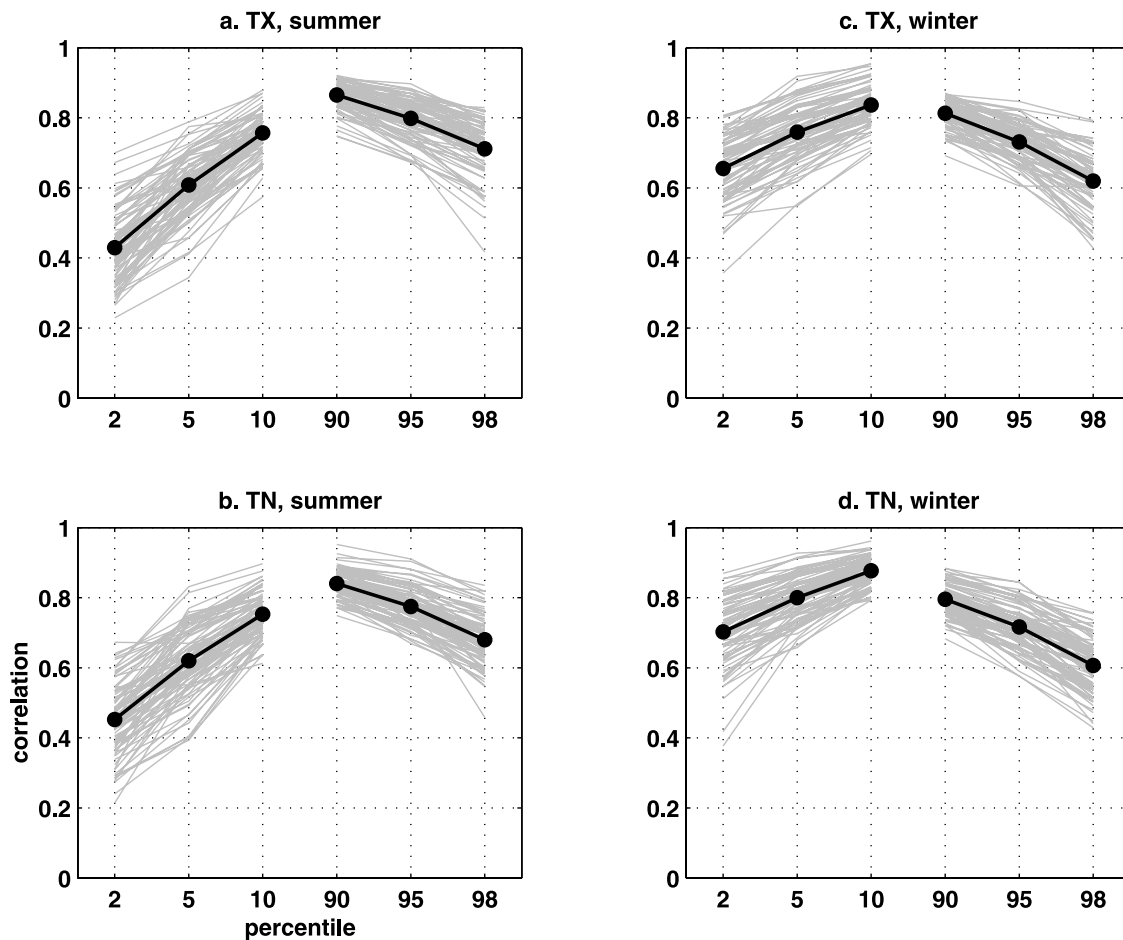
Figure 6. Correlations between temperature percentile indices and TMEAN for (a) TX in summer, (b) TN in summer, (c) TX in winter, and (d) TN in winter. The shaded lines indicate correlations for individual stations. Solid lines show the average correlation for all stations (essentially the same stations as in Figures 3 and 4). The labels on the horizontal axes refer to TX2P, TX5P, TX10P, TX90P, TX95P and TX98P in Figures 6a and 6c, and similarly to TN2P, etc., in Figures 6b and 6d.

[44] Despite the small and mostly insignificant summer precipitation trends, one further issue relating to Figure 5a is worth pointing out. The average summer trends in SDII, PREC98P and PREC95P (+5%, +4% and +3% respectively) have confidence intervals that do not cross the zero level. Note, in particular, also that SDII has the largest trend. Hence there is a tendency for summer precipitation to have become slightly more intense on average, but also less common because PRECTOT did not increase. On the other hand, the confidence intervals for trends in all five indices overlap, so it would be a step too far to say that a trend toward more intense and rarer summer precipitation during the past century has been significant.

5.3. Correlations Between Temperature Indices

[45] For temperature data, the correlations are, on average, strongest between the mean and the 10th/90th percentile indices, weaker between the mean and the 5th/95th percentiles and weaker still between the mean and the 2nd/98th percentiles (Figure 6). This rather unsurprising result holds for both TX and TN in winter and summer. On a more detailed level we observe that, for winter, the mean shows

slightly higher correlations with the lower percentiles than with the upper percentiles. In summer the situation is the opposite; that is, the mean is more strongly correlated with the upper percentiles than with the lower ones. Figure 6 further reveals that temporal variability (at individual sites) in the cold tail of the temperature distribution in summer can be notably different from variability in the mean (correlations down to 0.2 for some stations, and ~0.4 on average, both for TX and TN). Variations in the warm tail in summer and in both tails in winter are more similar to variations in the corresponding means (average correlations are between 0.6 and 0.7 for the 98th percentile in both winter and summer, and also for the 2nd percentile in winter).

5.4. Correlations Between Precipitation Indices

[46] For PRECTOT, the correlations with the three percentile indices are mostly quite strong (~0.7–0.9 on average) both in winter and summer (Figures 7a and 7c). In summer, however, the spread of correlation values among stations is notably larger than for winter. For PREC90P some stations actually have very weak (down to ~0.1)
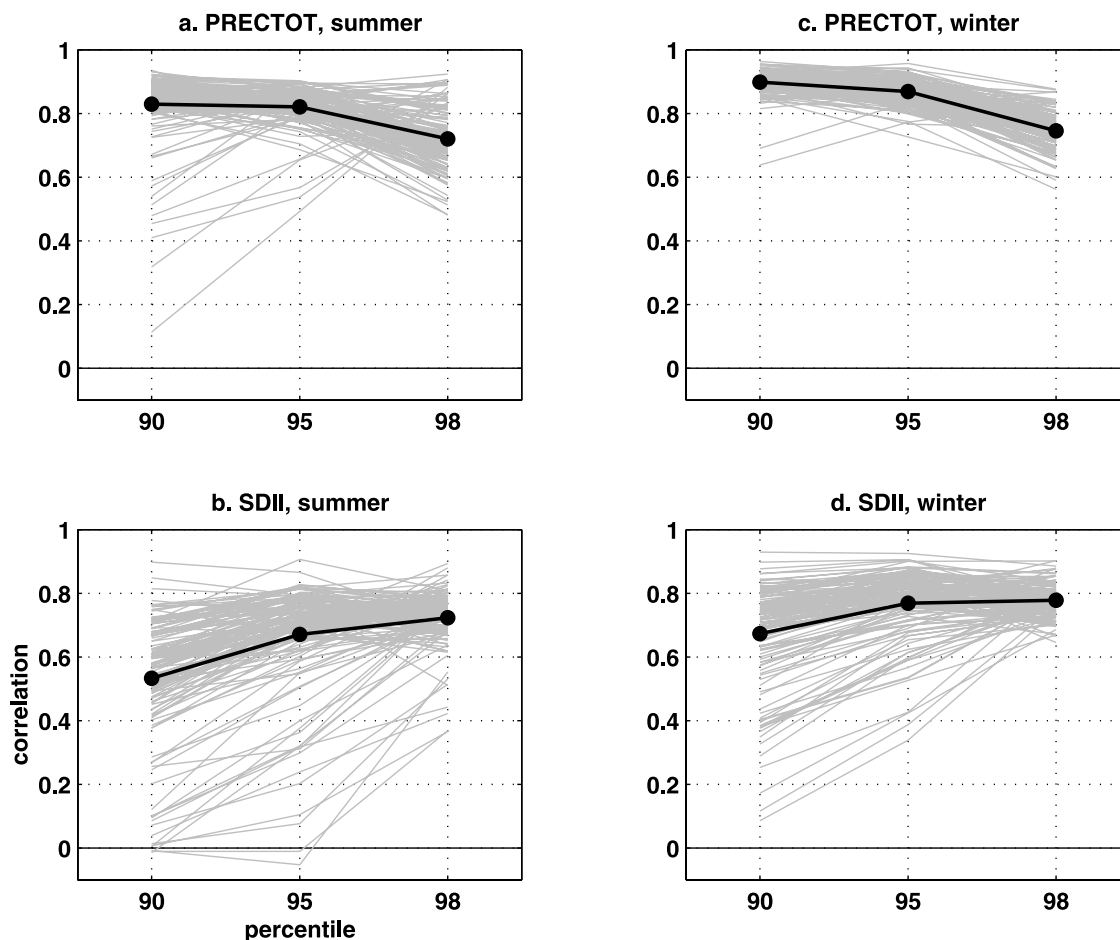
Figure 7. Correlations between precipitation percentile indices and two indices for average conditions. (a) Correlations with PRECTOT in summer, (b) correlations with SDII in summer, (c) correlations with PRECTOT in winter, and (d) correlations with SDII in winter. The shaded lines indicate correlations for individual stations. Solid lines show the average correlation for all stations (essentially the same stations as in Figure 5). The labels on the horizontal axes refer to PREC90P, PREC95P and PREC98P.

correlations with PRECTOT in summer. All nine stations with PRECTOT versus PREC90P correlations below 0.6 (in Figure 7a) are located in the Mediterranean region. These weak correlations are certainly related to the character of climate in this region with few wet days per summer (which cause PREC90P to be equal to zero in many years while PRECTOT varies from year to year). A similar effect, but to a much lesser extent, is seen for some stations also in winter. Two stations in Figure 7c have PRECTOT versus PREC90P correlations below 0.7, whereas all other stations have values above 0.8. The two stations with the lower correlations are located in Spain.

[47] Correlations between SDII and the percentile indices exhibit a much larger spread among stations compared to PRECTOT, both for summer and winter (Figures 7b and 7d). For some stations, correlations between SDII and both PREC90P and PREC95P are even close to zero in summer. All twelve stations with the weakest SDII versus PREC90P correlations ($r < 0.11$) in summer (Figure 7b) are from the Mediterranean region. For winter data (Figure 7d), eight of the nine stations with the weakest correlations between SDII and PREC90P are also Mediterranean. It is, as above, the

few wet days per season that cause the notably different behavior of SDII and PREC90P. We conclude that PREC90P (and to a lesser extent PREC95P) is difficult to interpret and maybe not so useful in dry climates (because of its property of often having zero-values). The indices PRECTOT, SDII and PREC98P are more "nicely" behaved and thus more easily allow comparison of trends in different types of climates. However, very dry summers with no wet days would make the calculation of SDII impossible.

5.5. Spatial Distribution of Temperature Trends

[48] For temperature data, we show maps for 1901–2000 trends in the 2nd and 98th percentiles (TN2P, TN98P, TX2P, TX98P) for both summer and winter (Figures 8 and 9). This provides an overview of trends in the lowest and highest temperature percentiles analyzed here. Overall, the main findings for these most extreme percentiles, in terms of spatial patterns, hold also for the other low and high percentiles respectively (maps not shown), although the size of trends in the less extreme percentiles may differ somewhat (compare Figure 3). The locations of all stations analyzed are indicated with triangles on the maps, whereas
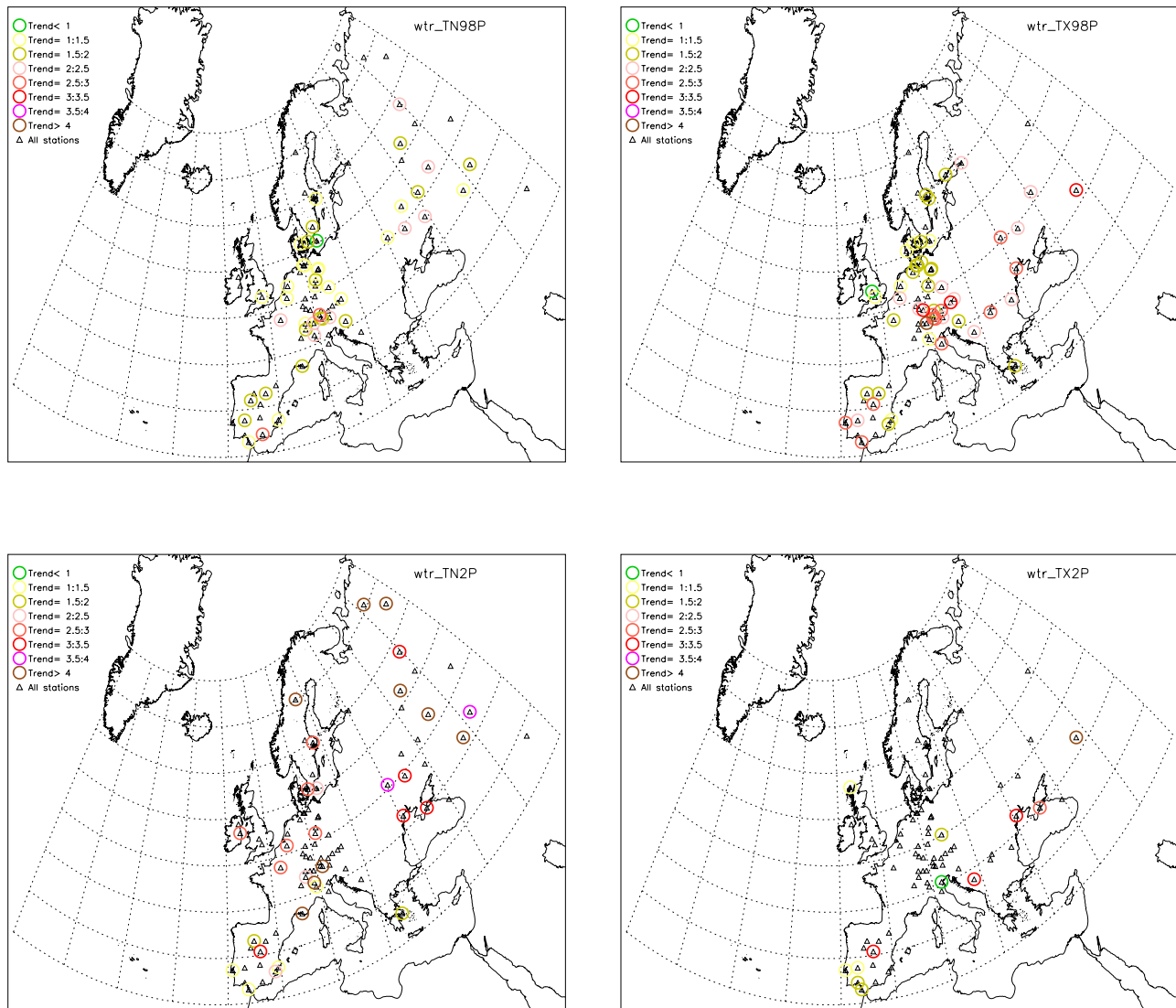
Figure 8. Locations of stations with significant (at the 5% level) trends over the period 1901–2000 in temperature indices for winter, for each of the four indices TN2P, TN98P, TX2P and TX98P. The colors indicate the size of trends (unit is °C/100 yr) as identified in the legends. All stations that pass a data completeness criterion are indicated with black triangles. Triangles with no associated circles denote stations where trends are statistically insignificant. Latitudes and longitudes are given in Figure 1.

statistically significant linear trends over the period 1901–2000 (unit is °C/100 yr) account for the numbers represented by colored circles.

[49] Starting with the TX98P index for winter (Figure 8), we find that significant warming trends occur across the entire region covered by the station network. The largest warming trends (often 2.5–4°C) occur in the southern parts (south of ∼45–50°N), whereas trends in the northern parts are generally smaller (1–2°C). The map for TN98P in winter also shows widespread warming trends, with most significant trends being between 1–2.5°C. This index, however, does not show clear signs of larger trends in the southern parts. For TN2P, strong significant winter warming trends (many >2.5°C, a few even >4°C) occur all over the study region (but mostly 1–2°C on the Iberian peninsula). The map for TX2P has too few significant trends for any clear spatial pattern to emerge.

[50] The maps for summer data (Figure 9), reveal that significant warming trends are frequent both for TN98P and TX98P over central and western Europe. On the Iberian peninsula, significant summer warming trends are frequent for TX98P but not for TN98P. The significant warming trends are often larger for TX98P (often 2–4°C) than TN98P (mostly 1–2°C). A few scattered stations with cooling summer trends are also found, but some of these occur in data-sparse regions so little can be concluded as to whether they reflect regional cooling or stations with inhomogeneous records. The overall impression, however, is a dominance of warming summer trends in the two 98th percentile indices. The spatial pattern of summer trends for the 2nd percentile indices is quite different. For both TX2P and TN2P, central and western Europe has many insignificant or weak trends (a mixture of warming and cooling). The map for TN2P is dominated by insignificant trends east
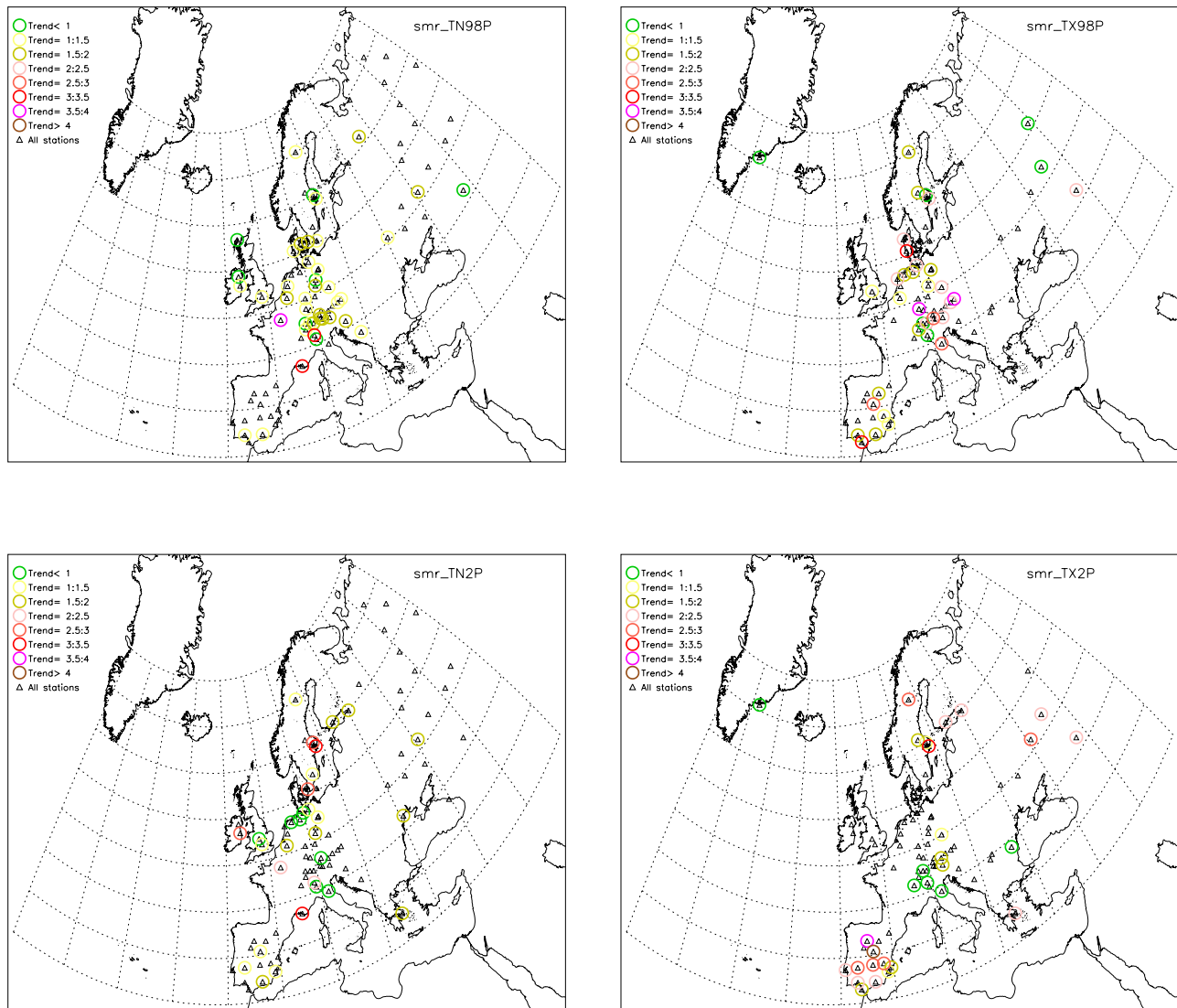
Figure 9. Same as in Figure 8 but for summer.

of 30°E. The Iberian peninsula has either insignificant or rather weak warming trends (1–2°C). On the map for TX2P, larger significant summer warming trends (often 2–3°C) occur over the Iberian peninsula, Fennoscandia and southwestern parts of former Soviet Union.

### 5.6. Spatial Distribution of Precipitation Trends

[51] For winter precipitation, maps with trends for four of the five indices analyzed (PRECTOT, SDII, PREC90P, PREC98P) are shown in Figure 10. The trend values indicated by colors refer to changes over the 100 years in % relative to the 1961–1990 average. We do not show any maps with precipitation trends in summer, as the number of stations with significant trends is low in this season.

[52] The main impression is that significant wetting trends in winter are common in the data-rich central and western Europe for all indices. There is also slight evidence for larger significant wetting trends in PRECTOT compared to SDII (in agreement with Figure 5). In central and western Europe significant wetting trends for PRECTOT are mostly between 20 and 40%, whereas SDII has about equal

amounts of significant 10–20% and 20–40% wetting trends.

[53] A dominance of insignificant winter precipitation trends is seen on the Iberian peninsula, on Iceland and the British Isles. East of 30°E, most stations have insignificant trends for SDII and PREC98P, whereas there is a mixture of insignificant trends and significant wetting and drying trends for PRECTOT and PREC90P. All four maps in Figure 10 show a few stations with significant drying trends. Some of these are very large (in the class 40–80% drying). Most of the few drying stations are located in data sparse regions, however, so it is not easy to assess whether the drying trends are real or due to inhomogeneous data. These outliers among trends certainly has an effect, both on the overall average trends and the width of the associated confidence intervals, shown in Figure 5.

[54] Overall, the analysis suggests that winter precipitation in Europe has increased significantly over the 20th century in central and western Europe. Furthermore, the observations indicate that winters in this part of Europe have experienced an increase both in the frequency of
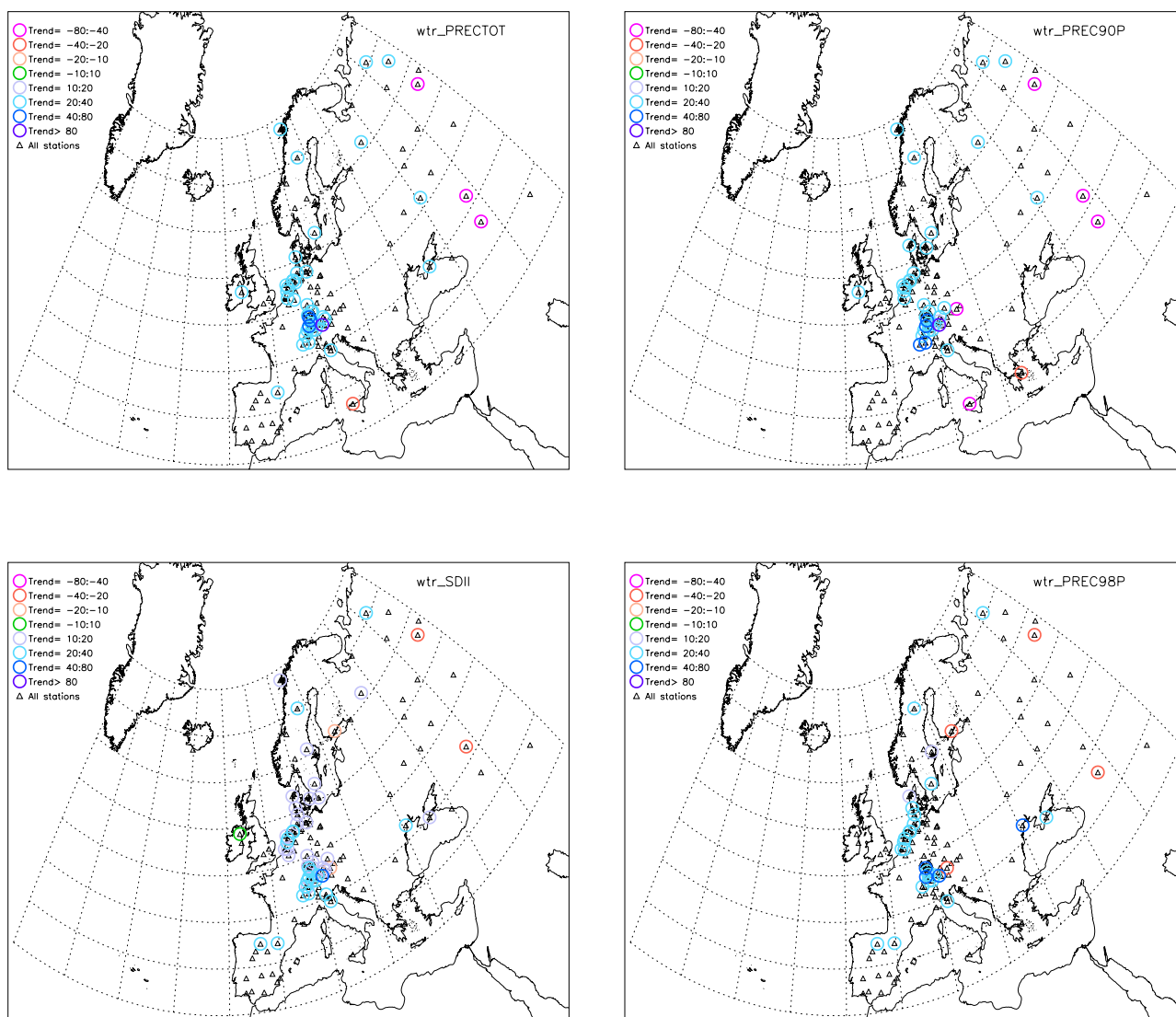
Figure 10. Locations of stations with significant (at the 5% level) trends over the period 1901–2000 in precipitation indices for winter, for each of the four indices PRECTOT, SDII, PREC90P and PREC98P. The colors indicate the size of trends (unit is % change with respect to the 1961–1990 average for each respective index) as identified in the legends. Graphical presentation is otherwise as in Figure 8.

precipitation events (because PRECTOT often increased more than SDII) and at the same time an increase in the average precipitation per wet day (because SDII increased). There is no clear evidence, however, for any disproportionately large increase in unusually heavy precipitation events in winter (because trends in PREC98P are about the same as for PREC90P and PRECTOT). This observation is more easily seen in the average trends shown in Figure 5 than on the maps in Figure 10.

5.7. Selection of Six Subregions

[55] Six European subregions have been selected (Figure 11). These regions, and their associated stations, were selected on the basis of visual inspection of the geographical distribution of stations, in combination with visual comparisons of time series plots of indices for neighboring stations. The regions are identified by the following names: IBERIA, GTALPINE, GERMANY, SSCAND,

NEEUROPE and SEEUROPE. Together, these regions cover much of the area represented by the entire station network. We should also mention that preliminary efforts were made to identify regions objectively using clustering algorithms, but this idea was abandoned because different results could be obtained for different climate indices.

[56] The inspection of time series for various indices revealed homogeneity problems for some stations. This led us to exclude some station series (both from the regional analysis and from the previously described Europe-average analyses). Nevertheless, as will be demonstrated, it appears that there are still homogeneity problems for some of the stations that were included. The set of stations finally selected for the regional analysis include some stations that were not used in the trend analyses (Figures 3–5 and 8–10). This is because here we could increase the number of contributing stations by also allowing some stations that did not pass the completeness criteria used when calculating
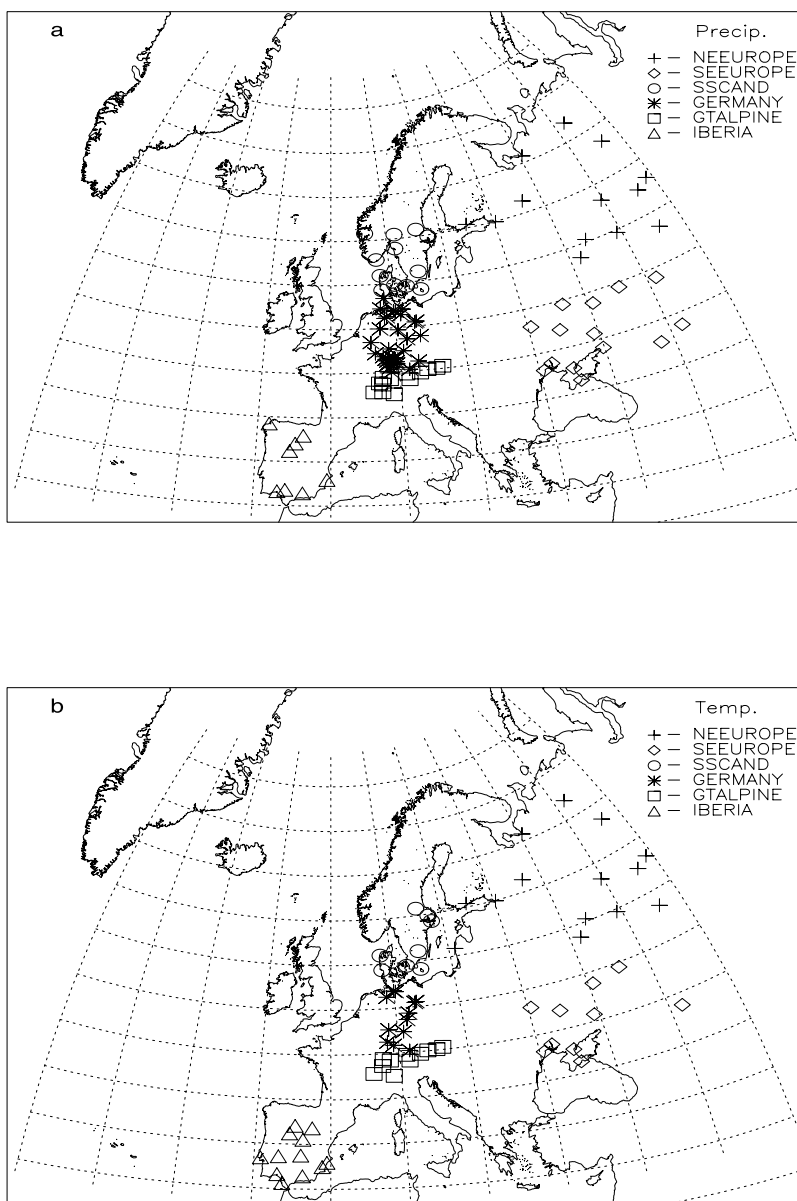
Figure 11. Locations of the stations used for the regional analyses in Figures 12–15. (a) Stations with precipitation data. (b) Stations with temperature data. The numbers of stations in each region are IBERIA (12 TEMP/8 PREC), GTALPINE (11/13), GERMANY (12/33), SSCAND (8/12), SEEUROPE (10/14), and NEEUROPE (14/13). Latitudes and longitudes are given in Figure 1.

trends at individual stations. For each region, the same set of stations is used for all temperature indices and another set for all precipitation indices. The number of stations used for each region is given in the caption for Figure 11.

### 5.8. Temperature Index Time Series for Subregions

[57] For each region and each temperature index, time series for all stations have been plotted together with the time series obtained by averaging the index series for all stations within the same region. Examples of these plots are shown for two chosen indices (TN98P and TX98P) for winter (Figure 12) and summer (Figure 13). The regional average series are plotted in black and all individual station series in grey. Smoothing of the regional average

series is made to highlight changes on decadal and longer timescales.

[58] As mentioned before, TX98P for winter is the index that shows the largest overall warming trends and is also the index for which the largest number of stations have significant trends. The detailed time evolution of this index varies among regions, but the overall warming trend can be observed in all six regions (Figure 12). The most steady linear trend is seen for IBERIA. Five of the regions (all but NEEUROPE) also show relatively high levels of winter TX98P in the late 1950s followed by a temporary drop to around 1970.

[59] The time series plots for TX98P in summer (Figure 13) show rather variable between-region details in their temporal evolution of the warm tail of the summer
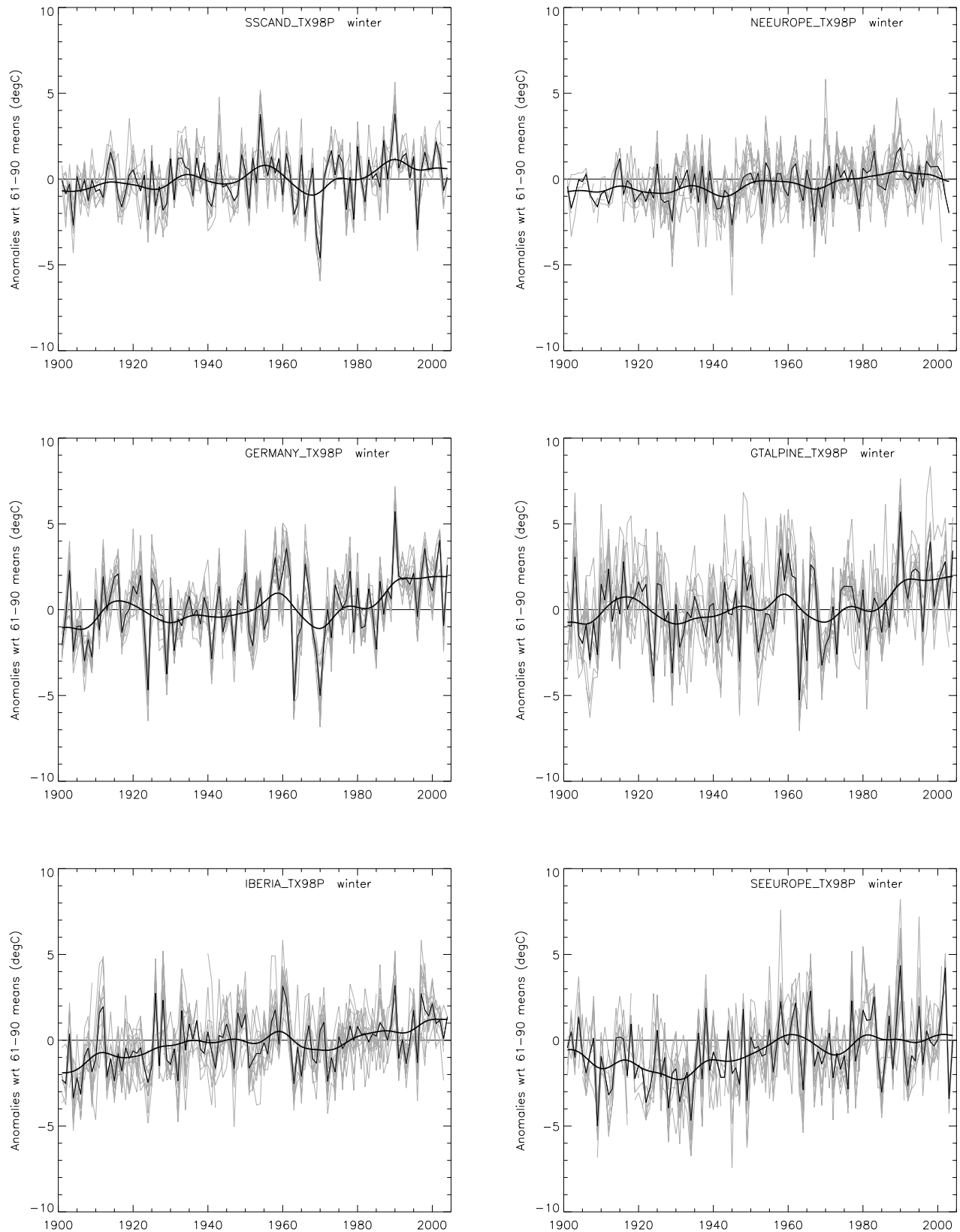
Figure 12. Time series for the TX98P index in winter (DJF) for stations in six different regions (IBERIA, GTALPINE, GERMANY, SSCAND, SEEUROPE and NEEUROPE). Unit is anomalies (°C) with respect to the 1961–1990 average. Shaded lines indicate individual stations. Solid lines indicate the average anomalies for the individual stations. Smooth solid lines indicate Gaussian 20-year smoothing of the regional average annual anomalies. The locations of stations used for the six regions are indicated in Figure 11. All winters are dates by the January.
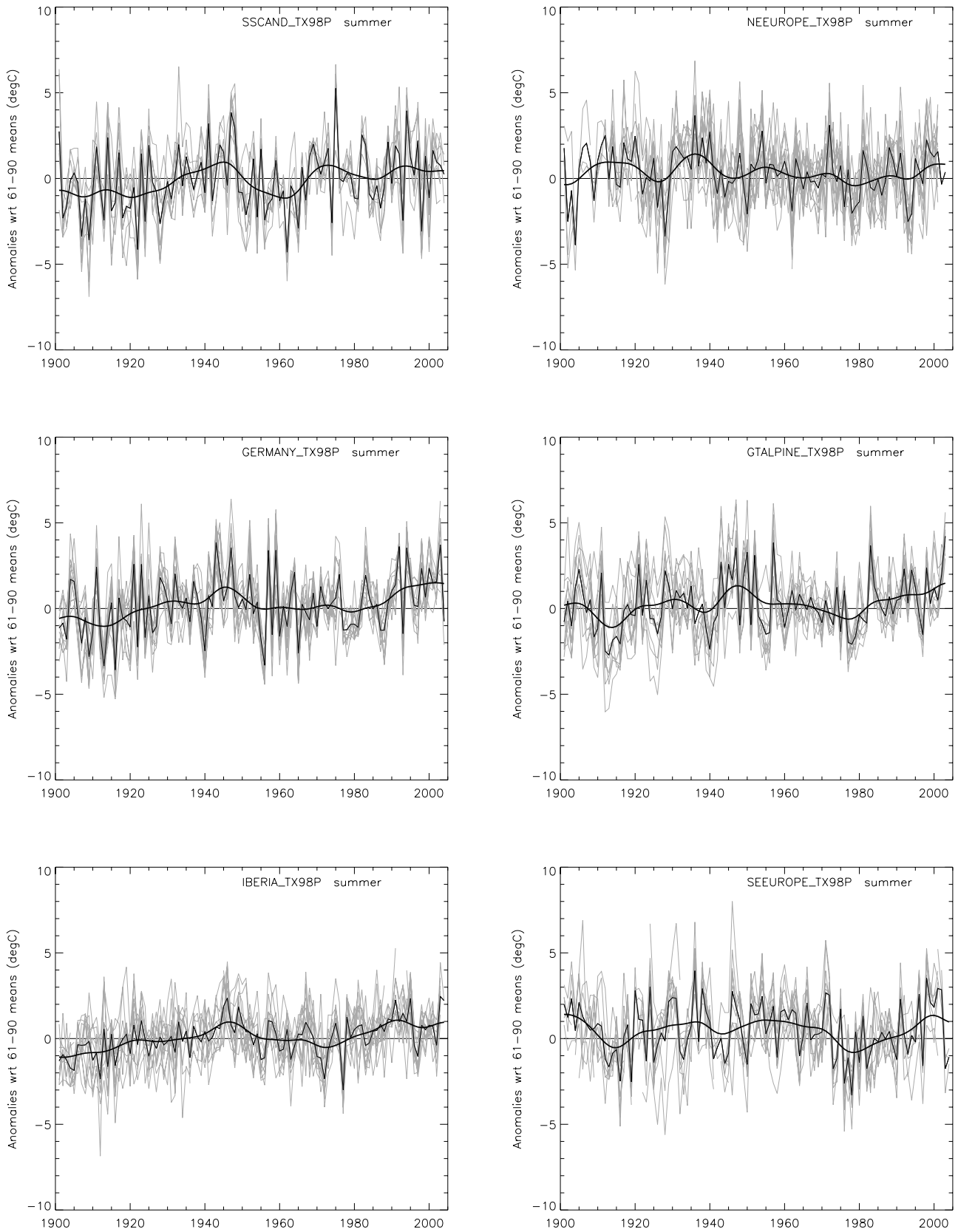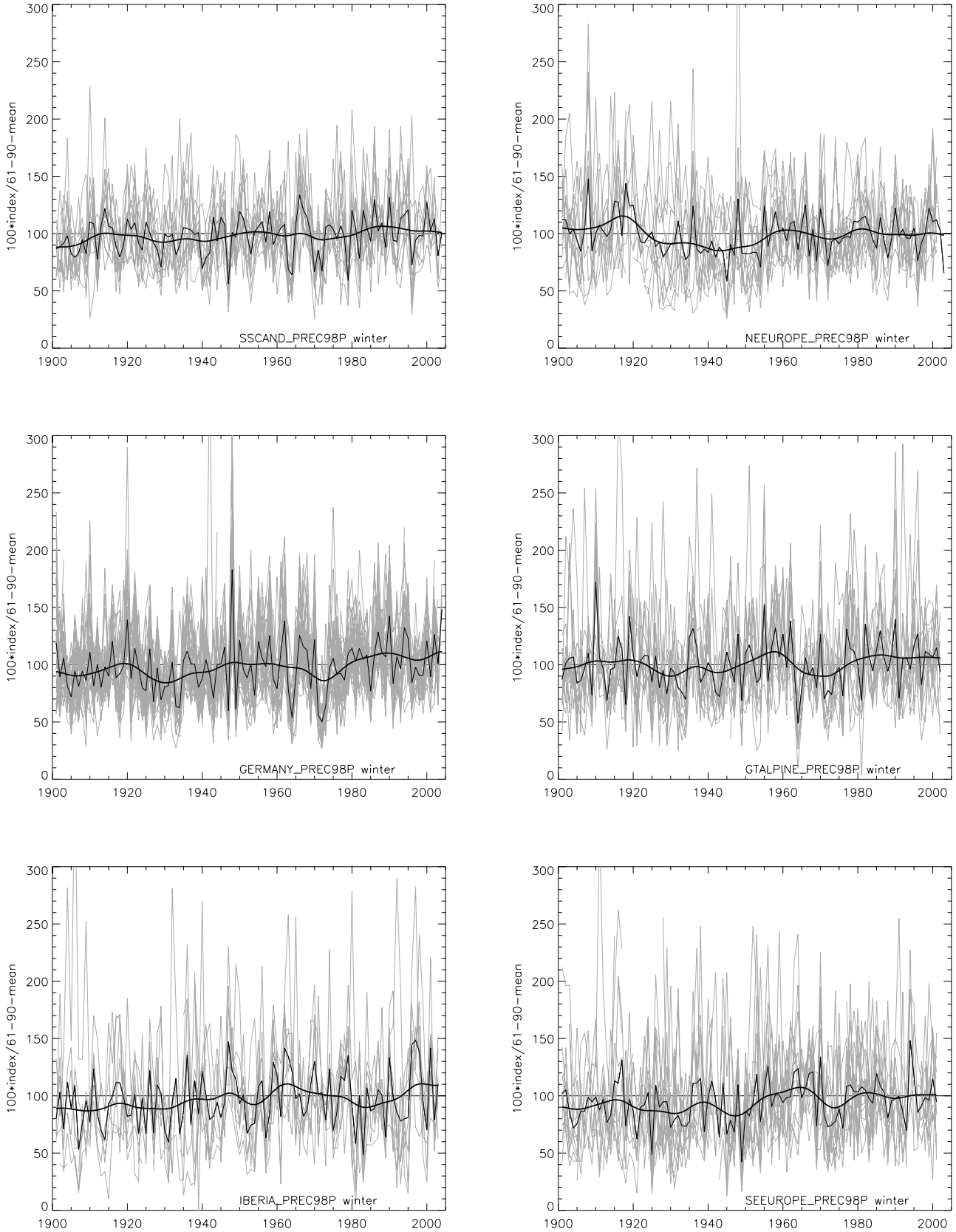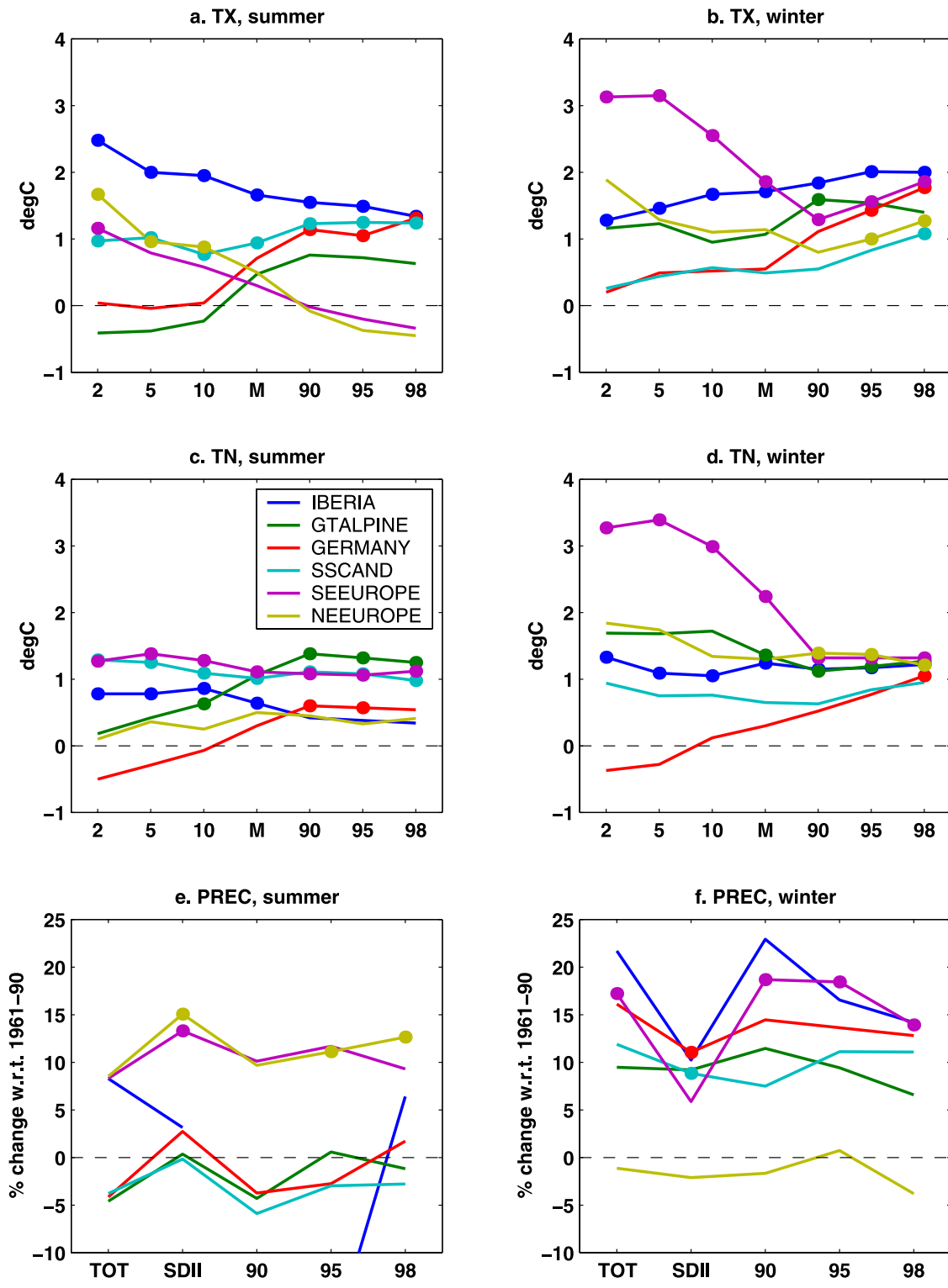
Figure 13. Same as in Figure 12 but for summer (JJA).

Figure 14. Same as in Figure 12 but for the PREC98P index in winter. Unit is % of the 1961–1990 average.

Figure 15. Trends in (a–d) temperature and (e and f) precipitation indices over the 1901–2000 period (units are as in Figures 3, 5a, and 5b) for regional average index series for six regions, defined in Figure 11. Trends that are significant at the 5% level are indicated with circles. The labeling of the horizontal axes are as in Figures 3 and 5.

temperature distribution. Although all six regions show anomalies in the last decade or so that are above the 1961–1990 average, there is evidence also for earlier periods with similar equally high values in all six regions. Four of the regions (SSCAND, GERMANY, GTALPINE, IBERIA) have high summer TX98P values in the 1940s. NEEU-ROPE has high values in the 1930s that are even slightly above the most recent values, and rather high values are

seen also in the 1910s. SEEUROPE differs from the other five regions in that the smoothed regional average is above the 1961–1990 mean for most of the record. The beginning of the 20th century appears to have been about as warm as the end of the 20th century for this region. Only two rather short intervals show anomalies below the 1961–1990 mean; the 1910s (quite opposite to the high levels for NEEUROPE in this period) and around 1975–1985. There is, however, rather substantial intraregion variability in SEEUROPE, in particular between 1920 and 1940, which makes it difficult to interpret the temporal evolution of TX98P in this region.

5.9. Precipitation Index Time Series for Subregions

[60] To give regional examples of precipitation time series, we have chosen to show PREC98P series for winter (Figure 14). Intraregion variability for this precipitation index is large. The same holds for all other precipitation indices too (not shown). This is an expected result, given that spatial correlations for precipitation are small. Overall, the large intraregion variability in the six regions illustrates quite well the general difficulties of assessing long-term changes in precipitation extremes within a region. These difficulties can arise both because of inhomogeneous data and/or because station density is too low. A notable observation here concerns NEEUROPE. It appears that the thirteen stations used for this region can be grouped into two main categories according to their behavior before around 1960. In that period, some stations have systematically rather high values compared to the other ones. This is most clearly seen for the period around 1920–1950. This finding suggests that not all homogeneity issues regarding precipitation series from former Soviet Union countries have been taken into account. The contrast to the behavior of the GERMANY series is striking, where all contributing stations seem to agree broadly about the temporal variations. In GERMANY, however, the average distance between stations is much smaller than in NEEUROPE, so a smaller intraregion variability should naturally be expected.

5.10. Regional Trends for Temperature and Precipitation Indices

[61] We conclude the result section with an analysis of 1901–2000 trends calculated for each of the six regions. Regional average series have been calculated for each temperature and precipitation index (as exemplified by the black series in Figures 12–14). The outcome of this analysis is presented in Figure 15, with colors representing different regions. The graphical presentation is designed to enable direct comparison with the overall trend results in Figures 3 and 5. For a given region, the trend estimates for the different indices are plotted and connected with lines. This makes it easy to see the pattern in how trends vary among indices for the various regions. Trends that are significantly different from zero are indicated with circles. Note that, here, significance is assessed for the trend in the regional-average time series for each index. (This differs from Figures 4, 5c, and 5d, where significance was tested for trends at each individual station, and then the fractions of stations having significant trends were plotted).

[62] A first impression for temperature indices (Figures 15a–15d) is the substantial interregional differences as regards trend patterns. It becomes immediately

clear that the overall pattern of Europe-wide mean trends illustrated in Figure 3 is not valid for all European sub-regions. In particular, the notable behavior of trends in TX in winter in Figure 3 (i.e., increasing trends with increasing percentile) is definitely not occurring in the SEEUROPE region. This region rather has its largest warming trends (>3°C) in the two lowest percentiles, TX2P and TX5P. Similar large trends are also found for TN2P and TN5P in SEEUROPE. Together, this reveals a reduction of the occurrence of very cold winter temperatures in this region. One reason why the behavior of TX indices for SEEUROPE was not more reflected in Figure 3, is likely that some former Soviet Union stations were excluded from the earlier analyses because of incomplete TX data. This implies that one needs to be careful with interpreting the results in Figure 3 as representing Europe-wide average conditions. Another notable observation for winter data is that all six regions have nearly the same warming trend (∼1°C) for TN98P, whereas the spread of trends in the cold tail (TN2P) is large for TN data (between −0.4°C for GERMANY and +3.3°C for SEEUROPE).

[63] For summer temperature data we observe that, although most indices for most regions show warming trends, there are also some regions for which small cooling trends are seen for some indices. No cooling trends, however, are statistically significant whereas many warming trends are significant. IBERIA stands out as having the largest warming trends among regions for all TX indices, with the largest trend occurring for TX2P (+2.5°C). This region, together with SEEUROPE and NEEUROPE, shows a tendency for a decreased range of daily maximum temperatures in summer. This is at odds with the overall pattern for summer TX trends in Figure 3, which rather suggests a tendency toward an increased range of daily maximum temperatures in summer. This latter behavior seems to be represented only in GERMANY and GTALPINE. This observation suggests that the high station density in the German and Alpine regions has a disproportional effect on the overall trends in Figure 3.

[64] A notable observation for summer precipitation data (Figure 15e) is the different character of trends in both NEEUROPE and SEEUROPE compared to SSCAND, GERMANY and GTALPINE. The two easternmost regions have positive trends (∼10 to 15% increases) for all five indices, whereas the three other mentioned regions have trends that mostly are either close to zero or slightly negative. No trend could be calculated for summer PREC90P in IBERIA because its average for the period 1961–1990 was less than 1 mm. The corresponding value for PREC95P was too small to provide a useful measure of the trend when expressed as a percentage. This explains the strange behavior of IBERIA summer precipitation trends in Figure 15e.

[65] In winter, positive trends are seen for five of the six regions (Figure 15f). The exception is NEEUROPE where trends are close to zero or slightly negative. Recall that the PREC98P time series for individual stations in NEEUROPE (Figure 14) were grouped in two categories before around 1960. If only the stations belonging to the "lower" category had been used, the overall trend in winter precipitation would likely have been positive, as for all the other regions. This adds further substance to our suspicion that the former

Soviet Union precipitation series may not be fully adjusted for all problems in the entire period back to 1901.

6. Discussion and Conclusions

[66] We have presented a new collection of long daily temperature and precipitation records, starting before 1901, for 230 stations spread across much of Europe and the Asian part of the former Soviet Union. This database was built within the EU-project EMULATE: "European and North Atlantic daily to multidecadal climate variability." On the basis of the daily series, a set of 64 climatic indices have been defined. Most of these indices measure some aspects of climate extremes, while some measure the mean climatic conditions. Index series for most stations are publicly available on the EMULATE website (the original station data are not available on the Website). In this paper, we have analyzed climatic changes over the period 1901–2000 for stations west of 60°E, using a subset of 19 indices. Most of these chosen indices are defined as different low and high percentiles (2nd, 5th, 10th, 90th, 95th, 98th), while a few measure changes in the mean. In this section, we will first highlight some data quality and homogeneity aspects we consider to be of importance for future studies. Then we briefly discuss the main results obtained from analyses of the climatic indices, before we give some recommendations for future work.

6.1. Data Quality and Homogeneity Aspects

[67] As mentioned in the data section, no overall homogeneity assessment of the data has been made. The main reason is that homogenization of the entire database, containing more than 200 daily temperature and precipitation series with more than 100-yearlong records from 26 countries would be a huge task, requiring a separate project devoted specifically to data homogenization. This is far beyond the scope and resources of EMULATE. Consequently, we have to rely on the efforts that were made on basic data quality checks for all station series, and also on the homogenization of a subset of the series that was made either within this project or before the project started (by some of those institutes and individual scientists from whom we obtained the data). The data quality control was in itself a large task that led to correction of a number of incorrect daily temperature or precipitation values. Despite all these controls on data errors, our experience is that it is virtually impossible to completely avoid all errors in databases like this.

[68] A good approach, when analyzing large observational climate databases, is to look at the data in many different ways and always be suspicious about any strange behavior of data from a particular station. Examples from our analyses are the plots of correlations between precipitation percentiles and both precipitation totals and precipitation per wet day (like Figure 7). It turned out that this kind of plot helped to identify a few stations with anomalous behavior. A closer inspection of these station data showed that a few individual highly erroneous daily values could corrupt the statistics for the entire series. When we discovered such errors, the erroneous data were corrected or removed before the station series was used in the subsequent analyses, and all earlier analyses were repeated after

the correction had been made. Similarly, inspection of time series plots of index series from several neighboring stations (like Figures 12–14) also revealed a number of data quality problems.

[69] We can summarize our experience concerning data quality issues and translate it into a rule: "The more the observational climate data series are analyzed, the larger number of problems will be found, and hence more errors can be corrected. It is vital, before any climate analyses are undertaken, to always inspect the observational data closely." This rule certainly holds for most types of large climatic databases, whether daily or monthly and regardless of which variable is studied. Furthermore, our experience also illustrates the value of exchange of data and the creation of multicountry/regional data sets.

6.2. Observed Temperature and Precipitation Changes

[70] As mentioned in the introduction, a recent worldwide analysis of changes in climatic extremes by *Alexander et al.* [2006] report widespread significant changes in temperature extremes, especially those related to daily minimum temperatures. According to their global study (for data since 1951), changes in daily maximum temperatures are less marked, implying that our world in many places has become less cold rather than hotter. The results from our European study based on the EMULATE data, however, reveals that the overall global findings by *Alexander et al.* [2006] for the post-1950 period are not representative for Europe if the entire twentieth century is considered. An overall warming is observed also in our study, but we find only a small difference, or no difference at all, between average trends in daily minimum and maximum temperatures when we average trends for 75 stations across Europe.

[71] It is worth noting that our results indicate that there were no significant trends in the daily temperature range (DTR) over the 20th century as a whole, when looking at the average of the analyzed European series. This gives a European century-scale perspective on previous findings by *Easterling et al.* [1997] and *Vose et al.* [2005], who reported decreasing DTR trends in most parts of the globe during the second half of the 20th century. However, when analyzed over the last 25 years, there is no global trend in DTR [*Vose et al.*, 2005].

[72] The difference between our findings and those of *Alexander et al.* [2006] are likely due both to different spatial domains (global versus Europe) and to the different periods analyzed. We have analyzed the rather long period of a whole century, and this might include different trends within different subperiods. For example, the dominance in daily minimum temperature warming (compared to daily maximum warming) is often seen as a particular impact of enhanced man-made greenhouse warming, which has accelerated during the last 50 years and was not as strong during the first half of the 20th century. Thus long periods for trend calculations often are not able to identify important changes that are working on a smaller timescale. Another example would be the recent observed decrease in Mediterranean winter precipitation which does not emerge in a comparable manner on a centennial timescale. An advantage of calculating centennial trends, is the longer time

perspective in which the most recent changes can be viewed.

[73] A similar issue related to the length of the time period has previously been discussed by *Klein Tank and Können* [2003]. They analyzed indices for daily temperature extremes across Europe for the latter half of the 20th century. They found a "symmetric" warming of the warm and cold tails for their entire study period, but when analyzing the second half of it they found an "asymmetric" warming, with the warm tail warming faster than the cold tail. We find some tendencies for such an asymmetry in the European warming, also when the entire century is analyzed. This can be seen in our Europe-wide average, both for daily maximum and minimum temperatures in summer and for daily maximum temperatures in winter. However, when the data are separated into subregions, the pattern breaks down as regional differences are quite large.

[74] As regards precipitation, the global study of *Alexander et al.* [2006] reveals that annual precipitation and the number of heavy precipitation days have shown widespread significant increases during the second half of the twentieth century although with some regional variation. Recall also that *Groisman et al.* [2005] found disproportionate changes, during the past decades, in heavy and very heavy precipitation compared to the change in the annual and/or seasonal precipitation. They found an increasing occurrence of intense precipitation events for many extratropical regions. Our analyses, for the entire 20th century, provide some more information about European conditions. We do see an average increase of winter precipitation over the century, but we do not see any evidence for any disproportionately large increase of heavy precipitation events. Rather, the trends in the upper percentiles of the distribution of daily precipitation totals in winter change roughly in proportion to the trend in seasonal precipitation totals. Moreover, the winter trend for the precipitation per wet day (the intensity index SDII) is also about the same.

[75] Given some previous discussions of winter precipitation trends [e.g., *Førland and Hanssen-Bauer*, 2000; *Schmidli et al.*, 2002; *Xoplaki et al.*, 2004], we draw the attention to the hypothesized problem that a climate warming may cause measured precipitation in winters to increase even if the true precipitation is unchanged. This is because the relative proportions of precipitation falling as snow and rain would change toward more rain and less snow. This in turn implies that the well-known bias due to undercatch of precipitation that falls as snow [e.g., *Groisman et al.*, 1991] would decrease in a warming climate. It is relevant to ask how this hypothesized effect has influenced the increased winter precipitation in Europe observed here. There are, however, very few studies made that can give any quantitative answers. *Schmidli et al.* [2002] looked at this problem for Alpine stations, but found the effect to be marginal, whereas it was quantifiable in the Arctic climate on Svalbard [*Førland and Hanssen-Bauer*, 2000]. It is not possible to extrapolate results from the Alps or Svalbard to the rest of Europe. However, given that most EMULATE stations are located at low elevations, where the snow component is relatively small, and that the effect was marginal in the Alps where there is more snow, it seems likely that the effect on the overall results found here are small.

[76] For summer, when averaging over all European stations with sufficiently complete data, we find no trend in the precipitation totals. There are, however, notable regional differences concerning the trends. It appears that summer precipitation trends have been positive in the eastern parts of the study domain, but zero or slightly negative in the western parts. For regions north of 40°N there is a slight tendency for a disproportionately large change in heavy precipitation events in summer, reflected by overall (small) positive trends both in the 98th percentile of daily precipitation and in the intensity index SDII. We do, however, not consider this change in precipitation character to be statistically significant. For the Mediterranean climate, it is difficult to assess trends in the precipitation percentiles in summer because the number of wet days can often be so small that trends become difficult to interpret. Data inhomogeneities and relative sparseness of the station density in many parts of Europe preclude more definitive conclusions from being drawn.

6.3. Future Prospects

[77] The EMULATE database makes a substantial improvement to the geographical distribution and density of stations with century-long daily temperature and precipitation records across Europe, as compared to previous investigations [see *Moberg and Jones*, 2005]. Nevertheless, it is still difficult to draw definitive and detailed conclusions with respect to trends in extremes in these climate variables over the entire twentieth century. The station density could certainly still be much improved in many European countries, but this would require further digitizing of daily data. Some activities are currently going on in this context, e.g., in the ENSEMBLES project (http://www.knmi.nl/samenw/ensembles_rt5/WP51.html), but even more efforts would be needed. It is also important that new methods be developed for homogenizing daily data, e.g., by following the path of *Della-Marta and Wanner* [2006]. We recommend the community of national meteorological institutes and other relevant organizations to cooperate on such efforts, to ensure that more accurate and more detailed assessments of long-term changes in climate extremes over Europe can be undertaken in the future.

Appendix A: Details About the Quality Checks (QC)

[78] The QC consisted of seven steps: (1) a check for logical errors (TX < TMEAN < TN, PREC < 0); (2) potential erroneous data defined as values exceeding a certain threshold were flagged; (3) for each outlier flagged, an estimated reasonably correct value was calculated from neighboring stations (for TX (TN) an estimate was also derived from TN (TX) for the same station) (step 3 could not be applied to precipitation data); (4) printed lists with outliers and their corresponding estimated reasonably correct values were produced; (5) these lists were inspected and individual values that were obviously wrong or suspected as being erroneous were identified; (6) reduced lists with only obviously wrong and suspect values were sent to the various data providers who compared these with the original sources; and (7) a correct value was inserted in our database whenever this action had been confirmed. During the subsequent climatic

analyses made in connection with this paper (e.g., the correlation analyses and time series plots described in the main text), a few more erroneous values that had not been detected during the basic QC were also found and corrected. The resulting data files are considered to be essentially free of serious errors, although an entirely error-free data set can never be guaranteed. Below, some additional information are provided on some of the QC steps.

[79] In step two, the threshold for defining an outlier (i.e., a potential erroneous value) for the daily temperature data (TX, TN, TMEAN separately) was chosen to be four times the standard deviation (4*SD) of daily data (calculated individually for all January, February, March, etc., days in the baseline period 1961–1990). For precipitation data, a threshold of 10*SD was chosen. These thresholds were selected (after having tried several other thresholds) as a compromise to ensure that the severe errors were captured without including too many correct values in the output list files.

[80] The third QC step, i.e., the estimation of reasonably correct values, could only be performed for the temperature data, but not for precipitation. For TX, TN and TMEAN separately, we estimated approximate values by using adjusted and weighted averages of data from stations with highly correlated series. These approximate reference values were calculated for the calendar day in question, in a fashion analogous to that used on a seasonal basis when creating reference series in homogeneity testing [see *Alexandersson and Moberg*, 1997], i.e., weighting the reference series temperature anomalies (with respect to the baseline period 1961–1990) with their squared correlation coefficient with the candidate series and then adding the climatological average for the candidate station to ensure that the absolute value is representative for this station. The reference stations selected were the six having the largest correlation coefficients (calculated using monthly mean series with the annual cycle removed) with the candidate series, with the additional requirement that the squared correlation must exceed 0.7. Because of this latter criterion, less than six reference stations had sometimes to be used. We also found, after some trials, that this kind of approach was not at all applicable for the daily precipitation data because of the weaker spatial correlations compared to temperatures.

[81] Furthermore (still QC step three), for outliers in TX (TN) series, an independent set of estimated values was derived from the corresponding TN (TX) values in the candidate series, by adding (subtracting) to these values the climatological average difference (TX-TN). These climatological differences were calculated for 5-day windows, centered on the calendar day in question, during the reference period 1961–1990.

[82] The lists produced in the fourth QC step contained, in addition to the outliers themselves, the following information: (1) for TN(TX) series, year, month, day, actual value (X) in database, climatological mean for that day of the year, climatological standard deviation (SD), (X-mean)/SD, estimate from reference stations, estimate from TX (TN), station name, and reference station names; (2) for TMEAN series, year, month, day, X, mean, SD, (X-mean)/SD, estimate from reference stations, station name, and reference station names; and (3) for PREC series, year, month, day, X, mean, SD, (X-mean)/SD, and station name.

[83] Additional lists with all correlations (for temperature data) between the candidate station and their reference stations were also produced.

[84] The automated part (steps 1–4) was followed by time-consuming manual work (steps 5–7), where the lists were visually inspected to identify the obvious errors and suspected wrong values. Reduced outlier lists were then sent to the individual data providers, who were asked to compare them with original data in their archives in order to determine which outliers could be confirmed or removed from the data set. On the basis of their answers, we corrected our data files whenever necessary and possible. This approach turned out to be an affordable way leading to a largely satisfying database for subsequent analyses.

References

Alexander, L. V., et al. (2006), Global observed changes in daily climate extremes of temperature and precipitation, *J. Geophys. Res.*, *111*, D05109, doi:10.1029/2005JD006290.

Alexandersson, H., and A. Moberg (1997), Homogenisation of Swedish temperature data, part I: Homogeneity test for linear trends, *Int. J. Climatol.*, *17*, 25–34.

Auer, I., et al. (2005), A new instrumental precipitation dataset for the Greater Alpine Region for the period 1800–2002, *Int. J. Climatol.*, *25*, 139–166.

Begert, M., T. Schlegel, and W. Kirchhofer (2005), Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000, *Int. J. Climatol.*, *25*, 65–80.

Bergström, H., and A. Moberg (2002), Daily air temperature and pressure series for Uppsala (1722–1998), *Clim. Change*, *53*, 213–252.

Böhm, R., I. Auer, M. Brunetti, M. Maugeri, T. Nanni, and W. Schöner (2001), Regional temperature variability in the European Alps: 1760–1998 from homogenized instrumental time series, *Int. J. Climatol.*, *12*, 1779–1801.

Brunet, M., O. Saladié, P. Jones, J. Sigró, A. Moberg, E. Aguilar, A. Walther, D. Lister, and D. López (2006), The development of a new daily adjusted temperature dataset for Spain (1850–2003), *Int. J. Climatol.*, doi:10.1002/joc.1338, in press.

Brunetti, M., M. Maugeri, F. Monti, and T. Nanni (2004), Changes in daily precipitation frequency in Italy over the last 120 years, *J. Geophys. Res.*, *109*, D05102, doi:10.1029/2003JD004296.

Camuffo, D., and P. Jones (Eds.) (2002), Improved understanding of past climatic variability from early daily European instrumental sources, *Clim. Change*, *53*, 1–392.

Changnon, S. A., E. R. Fosse, and E. L. Lecomte (1999), Interactions between the atmospheric sciences and insurers in the United States, *Clim. Change*, *42*, 51–67.

Cohn, T. A., and H. F. Lins (2005), Nature's style: Naturally trendy, *Geophys. Res. Lett.*, *32*, L23402, doi:10.1029/2005GL024476.

Della-Marta, P. M., and H. Wanner (2006), A method of homgenizing the extremes and mean of daily temperature measurements, *J. Clim.*, *19*(17), 4179–4197.

Easterling, D. R., et al. (1997), Maximum and minimum temperature trends for the globe, *Science*, *277*, 364–367.

Førland, E. J., and I. Hanssen-Bauer (2000), Increased precipitation in the Norwegian Arctic: True or false?, *Clim. Change*, *46*, 485–509.

Frei, C., and C. Schär (2001), Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region, *J. Clim.*, *14*, 1568–1584.

Groisman, P. Y., V. V. Koknaeva, T. A. Belokrylova, and T. R. Karl (1991), Overcoming biases of precipitation measurement: A history of the USSR Experience, *Bull. Am. Meteorol. Soc.*, 72, 1725–1733.

Groisman, P. Y., R. W. Knight, D. R. Easterling, T. R. Karl, T. C. Hegerl, and V. N. Razuvaev (2005), Trends in intense precipitation in the climate record, *J. Clim.*, 18, 1326–1350.

Hegerl, G. C., F. W. Zwiers, P. A. Stott, and V. V. Kharin (2004), Detectability of anthropogenic changes in annual temperature and precipitation extremes, *J. Clim.*, 17, 3683–3700.

Hulme, M. (1994), Validation of large-scale precipitation fields in General Circulation Models, in *Global Precipitations and Climate Change*, edited by M. Desbois and F. Désalmond, pp. 387–405, Springer, New York.

Hulme, M. (1996), Recent climatic change in the world's drylands, *Geophys. Res. Lett.*, 23, 61–64.

Jones, P. D., and A. Moberg (2003), Hemispheric and large-scale surface air temperature variations: Sn extensive revision and update to 2001, *J. Clim.*, 16, 206–223.

Kharin, V. V., and F. W. Zwiers (2000), Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM, *J. Clim.*, 13, 3760–3788.

Kiktev, D., D. M. H. Sexton, L. Alexander, and C. K. Folland (2003), Comparison of modeled and observed trends in indices of daily climate extremes, *J. Clim.*, 16, 3560–3571.

Klein Tank, A. M. G., and G. P. Können (2003), Trends in indices of daily temperature and precipitation extremes in Europe, 1946–1999, *J. Clim.*, 16, 3665–3680.

Klein Tank, A. M. G., et al. (2002), Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment, *Int. J. Climatol.*, 22, 1441–1453.

Klein Tank, A. M. G., et al. (2006), Changes in daily temperature and precipitation extremes in central and south Asia, *J. Geophys. Res.*, 111, D16105, doi:10.1029/2005JD006316.

Moberg, A., and H. Alexandersson (1997), Homogenization of Swedish temperature data. Part II: Homogenized gridded air temperature compared with a subset of global gridded air temperature since 1861, *Int. J. Climatol.*, 17, 35–54.

Moberg, A., and P. D. Jones (2005), Trends in indices for extremes in daily temperature and precipitation in central and western Europe 1901–1999, *Int. J. Climatol.*, 25, 1149–1171.

Moberg, A., H. Bergström, J. Ruiz Krigsman, and O. Svanered (2002), Daily air temperature and pressure series for Stockholm (1756–1998), *Clim. Change*, 53, 171–212.

Nordli, P. Ø., H. Alexandersson, P. Frich, E. J. Førland, R. Heino, T. Jónsson, H. Tuomenvirta, and O. E. Tveito (1997), The effect of radiation screens on Nordic time series of mean temperature, *Int. J. Climatol.*, 17, 1667–1681.

Nutter, F. W. (1999), Global climate change: Why U.S. insurers care, *Clim. Change*, 42, 45–49.

Razuvaev, V. N., E. G. Apasova, and R. A. Martuganov (1993), Daily temperature and precipitation data for 223 USSR stations, *ORNL/CDIAC-56, NDP-040, ESD Publ. 4194*, 134 pp., Carbon Dioxide Inf. Anal. Data Cent., Oak Ridge Natl. Lab., Oak Ridge, Tenn.

Schmidli, J., C. Schmutz, C. Frei, H. Wanner, and C. Schär (2002), Mesoscale precipitation variability in the region of the European Alps during the 20th century, *Int. J. Climatol.*, 22, 1049–1074.

Semenov, V. A., and L. Bengtsson (2002), Secular trends in daily precipitation characteristics: greenhouse gas simulation with a coupled AOGCM, *Clim. Dyn.*, 19, 123–140.

Trewin, B. C., and A. C. F. Trevitt (1996), The development of composite temperature records, *Int. J. Climatol.*, 16, 1227–1242.

Tuomenvirta, H. (2001), Homogeneity adjustments of temperature and precipitation series - Finnish and Nordic data, *Int. J. Climatol.*, 21, 495–506.

Vincent, L. A., X. Zhang, B. R. Bonsal, and W. D. Hogg (2002), Homogenization of daily temperatures over Canada, *J. Clim.*, 15, 1322–1334.

Vose, R. S., D. R. Easterling, and B. Gleason (2005), Maximum and minimum temperature trends for the globe: An update through 2004, *Geophys. Res. Lett.*, 32, L23822, doi:10.1029/2005GL024379.

Wijngaard, J. B., A. M. G. Klein Tank, and G. P. Können (2003), Homogeneity of 20th century European daily temperature and precipitation series, *Int. J. Climatol.*, 23, 679–692.

Xoplaki, E., J. F. González-Rouco, J. Luterbacher, and H. Wanner (2004), Wet season Mediterranean precipitation variability: influence of large-scale dynamics and trends, *Clim. Dyn.*, 23, 63–78.

Zhang, X., G. Hegerl, F. W. Zwiers, and J. Kenyon (2005), Avoiding inhomogeneity in percentile-based indices of temperature extremes, *J. Clim.*, 18, 1641–1651.

E. Aguilar, M. Brunet, O. Saladié, and J. Sigró, Climate Change Research Group, Physical Geography, Universitat Rovira i Virgili, E-43005 Tarragona, Spain.

L. V. Alexander and J. Caesar, Hadley Centre, Met Office, Exeter EX1 3PB, UK.

H. Alexandersson, Swedish Meteorological and Hydrological Institute, SE-60176 Norrköping, Sweden.

C. Almarza, Instituto Nacional de Meteorología, Leonardo Prieto Castro 8, E-28040 Madrid, Spain.

I. Auer and R. Böhm, Central Institute for Meteorology and Geodynamics, A-1190 Vienna, Austria.

M. Barriendos, Department of Modern History, University of Barcelona, c/Montalegre 6, E-08001 Barcelona, Spain.

M. Begert, Climate Services, Federal Department of Home Affairs, Federal Office of Meteorology and Climatology (MeteoSwiss), Kraehbuehlstrasse 58, P.O. Box 514, CH-8044 Zürich, Switzerland.

H. Bergström, Department of Earth Sciences, Uppsala University, SE-75105 Uppsala, Sweden.

C. J. Butler, Armagh Observatory, Armagh BT61 9DG, UK.

D. Chen, H. Linderholm, and A. Walther, Regional Climate Group, Earth Sciences Centre, Göteborg University, SE-40530 Göteborg, Sweden.

P. M. Della-Marta, J. Luterbacher, H. Wanner, and E. Xoplaki, Institute of Geography, Climatology, and Meteorology and NCCR Climate, University of Bern, CH-3012 Bern, Switzerland.

A. Drebs and H. Tuomenvirta, Finnish Meteorological Institute, FIN-00101 Helsinki, Finland.

D. Founda and M. Petrakis, Institute for Environmental Research and Sustainable Development, National Observatory of Athens, GR-11810 Athens, Greece.

F.-W. Gerstengarbe, H. Österle, and P. C. Werner, Potsdam Institute for Climate Impact Research, D-14412 Potsdam, Germany.

J. Jacobeit and A. Philipp, Institute of Geography, University of Augsburg, D-86159 Augsburg, Germany.

P. D. Jones and D. Lister, Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

A. M. G. Klein Tank, Royal Netherlands Meteorological Institute, NL-3730 AE De Bilt, Netherlands.

M. Maugeri, Istituto di Fisica Generale Applicata, Universitá di Milano, I-20122 Milan, Italy.

G. Micela, Osservatorio Astronomico di Palermo, I-90134 Palermo, Italy.

A. Moberg, Department of Physical Geography and Quaternary Geology, Stockholm University, SE-10691 Stockholm, Sweden. (anders.moberg@natgeo.su.se)

K. Pandzic and L. Srnec, Meteorological and Hydrological Service, 10000 Zagreb, Croatia.

R. Tolasz, Czech Hydrometeorological Institute, 14306 Prague, Czech Republic.

P. Yiou, Laboratoire des Sciences du Climat et de l'Environnement, UMR Commissariat à l'Energie Atomique–Centre Nationale de le Recherche Scientifique, F-91191 Gif-sur-Yvette, France.