

# Summarising changes in air temperature over Central Europe by quantile regression and clustering

S. M. Barbosa<sup>1</sup>, M. G. Scotto<sup>2</sup>, and A. M. Alonso<sup>3</sup>

<sup>1</sup>Instituto Dom Luiz, Universidade de Lisboa, Portugal

<sup>2</sup>Departamento de Matemática, Universidade de Aveiro, Portugal

<sup>3</sup>Departamento de Estadística, Instituto Flores de Lemus, Universidad Carlos III de Madrid, Spain

Received: 8 February 2011 – Revised: 5 May 2011 – Accepted: 29 September 2011 – Published: 9 December 2011

**Abstract.** The analysis of trends in air temperature observations is one of the most common activities in climate change studies. This work examines the changes in daily mean air temperature over Central Europe using quantile regression, which allows the estimation of trends, not only in the mean but in all parts of the data distribution. A bootstrap procedure is applied for assessing uncertainty on the derived slopes and the resulting distributions are summarised via clustering. The results show considerable spatial diversity over the central European region. A distinct behaviour is found for lower (5 %) and upper (95 %) quantiles, with higher trends around  $0.15\text{ }^{\circ}\text{C decade}^{-1}$  at the 5 % quantile and around  $0.20\text{ }^{\circ}\text{C decade}^{-1}$  at the 95 % quantile, the largest trends ( $>0.2\text{ }^{\circ}\text{C decade}^{-1}$ ) occurring in the Alps.

## 1 Introduction

Temperature changes can have profound impacts on socio-economic activities and on human health. Spells of excess heat, for example, are associated with increasing mortality rates (e.g. Patz et al., 2005), changes in water resources (e.g. Zappa and Kan, 2007) and higher occurrence of forest fires (Trigo et al., 2006). Over Europe, significant trends in extreme temperatures are expected due to changes in large-scale circulation and snow cover extent (e.g. Della-Marta et al., 2007; van den Besselaar et al., 2010)

Due to its considerable societal relevance, the temporal evolution of temperature extremes is the focus of intense scientific research. Temperature extremes are often analysed by fitting trends to pre-defined extreme indices (e.g. Frich et al., 2002) or data percentiles (e.g. Hertig et al., 2010), or by applying extreme value theory (e.g. Nogaj et al., 2006; Hugu

et al., 2008; Parey, 2008; Scotto et al., 2011). However, in a climate change context the temporal evolution of mean temperature is of relevance as well. Variability in the mean influences variability in extremes (although it does not completely determine), and is easier to quantify since the available sample size is naturally much larger than in the case of extreme temperature events. Furthermore, correct simulation of the temporal mean of climate variables is one of the first requirements for numerical climate models.

The long-term temporal evolution of a time series of averaged temperature observations (e.g. global or daily mean temperature) is often characterised by a single number, the slope of a linear trend model fitted to the data. In that case, the variability is assumed to remain constant in time – when fitting a linear trend by ordinary least squares, the mean is assumed to evolve in time but the data distribution (and specifically the variance) is assumed to remain unchanged over the considered period. However, this is not necessarily a realistic assumption. Of particular interest is the assessment of whether such a hypothesis is reasonable, or if the distribution itself changes in time, leading to different slope values for different parts of the data distribution.

Quantile regression (Koenker and Basset, 1978; Koenker and Hallock, 2001) provides a well defined statistical framework for estimating the rate of change not only in the mean as in ordinary regression, but in all parts of the data distribution. This work addresses the temporal evolution of daily mean temperature over Central Europe. Quantile regression is applied in order to thoroughly quantify the variability structure of daily mean temperature, and assess whether just the mean temperature evolves in time, or if there are changes in the data distribution itself, thereby obtaining a more complete picture of the temporal changes in daily mean temperature.

For regional studies, not only the temporal evolution but also the spatial distribution of temperature changes is of scientific interest. The regional variability of air temperature observations is often analysed by taking each



Correspondence to: S. Barbosa  
(sabarbosa@fc.ul.pt)

individual temperature time series and summarising the information for the region of interest in terms of maps of individual features. An alternative approach to spatially characterise regional variability is cluster analysis (e.g. Scotto et al., 2009; Mahlstein and Knutti, 2010). The aim of the present study is to describe regional temperature variability over Europe by combining quantile regression and a time series clustering procedure in the analysis of European daily mean temperature records. The data and methodological approach are described in the next section, results are presented in Sect. 3 and concluding remarks are given in Sect. 4.

## 2 Data and methods

### 2.1 Data

Time series of daily mean temperature are obtained from the blended European Climate Assessment (ECA) dataset (Tank et al., 2002; Klok and Tank, 2009). Only stations in western Europe with data from at least January 1901 to December 2007 and with less than 2 % of missing observations have been selected (Fig. 1 and Table 1). As a pre-processing step, seasonality is removed from each temperature record via sinusoidal regression at annual and semi-annual frequencies (e.g. Kedem and Fokianos, 2002). The ECA dataset is subject to quality control procedures but inhomogeneities that could influence the analysis of extreme temperatures can remain (Wijngaard et al., 2003). The deseasoned data are therefore inspected for potential outliers. The record from Wien was removed from the analysis due to the presence of some suspicious values (exceeding  $7\sigma$ ) in the last part of the record.

### 2.2 Statistical methods

#### 2.2.1 Quantile regression

Quantile regression is a well-defined statistical framework for regression on quantiles rather than regression on the mean. Although it was first introduced in econometrics by Koenker and Basset (1978), quantile regression is being applied in various geoscience contexts (e.g. Koenker and Schorfheide, 1994; Cade and Noon, 2003; Baur et al., 2004; Elsner et al., 2008; Barbosa, 2008).

Given a random variable  $Y$  with cumulative continuous distribution function  $F_Y(y)$ , the quantile  $\tau$  is defined as the value  $Q_Y(\tau)$  such that  $P[Y \leq Q_Y(\tau)] = \tau$ ,  $0 \leq \tau \leq 1$ . The quantile function  $Q_Y(\tau)$  is defined from the cumulative distribution function  $F_Y(y)$  as  $Q_Y(\tau) = F_Y^{-1}(\tau)$ . Then, considering the conditional distribution of  $Y$  given  $X = x$ , the conditional quantile function  $Q_{Y|X}(\tau; x)$  verifies  $P[Y \leq Q_{Y|X}(\tau; x) | X = x] = \tau$ . Whereas ordinary regression is based on the conditional mean function  $E[Y|X = x]$  and minimisation of the respective residuals, quantile regression

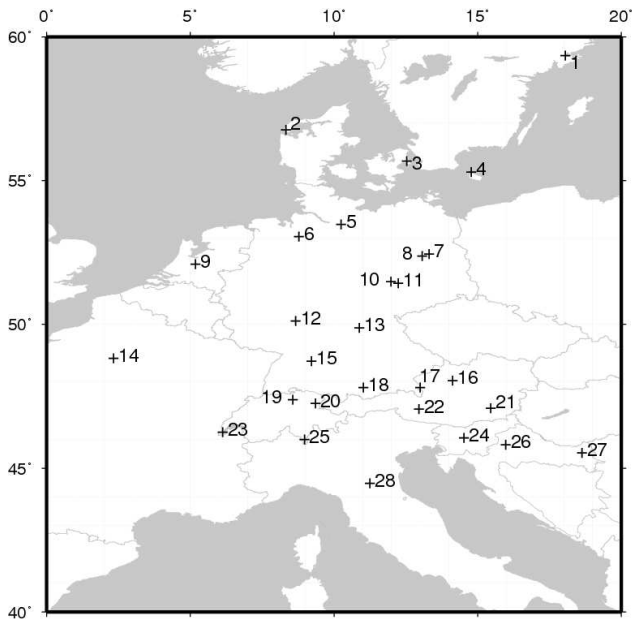
**Table 1.** Daily mean air temperature data summary.

	Station (Abbreviation)	% missing	Height (m)
1	STOCKHOLM (STO)	0	44
2	VESTERVIG (VES)	0.04	18
3	KOEBENHAVN (KOE)	0.06	9
4	HAMMER ODDE FYR (HOF)	0.58	11
5	HAMBURG (HAM)	0.75	35
6	BREMEN (BRE)	0.98	5
7	BERLIN (BER)	1.51	58
8	POTSDAM (POT)	1.51	81
9	DE BILT (DEB)	0.05	2
10	HALLE (HAL)	0.04	104
11	LEIPZIG (LEI)	0	141
12	FRANKFURT (FRE)	0.27	112
13	BAMBERG (BAM)	1.56	239
14	PARIS (PAR)	0.01	75
15	STUTTGART (STU)	0.33	401
16	KREMSMUNSTER (KRE)	0.12	383
17	SALZBURG (SAL)	1.89	437
18	HOHENPEISSENBERG (HOH)	1.68	977
19	ZUERICH (ZUE)	0.21	556
20	SAENTIS (SAE)	0.25	2490
21	GRAZ (GRA)	0.52	366
22	SONNBLICK (SON)	0.08	3106
23	GENEVE (GEN)	0.26	405
24	LJUBLJANA BEZIGRAD (LJU)	1.1	299
25	LUGANO (LUG)	0.25	273
26	ZAGREB-GRIC (ZAG)	0.07	156
27	OSIJEK (OSI)	0.82	88
28	BOLOGNA (BOL)	0.15	53

is based on the conditional quantile function and minimisation of the sum of asymmetrically weighted absolute residuals  $\sum_{i=1} \rho(\tau) |y_i - Q_{Y|X}(\tau; x = x_i)|$ , where  $\rho$  is the tilted absolute value function. Further details can be found in Koenker and Hallock (2001) and Koenker (2005).

#### 2.2.2 Clustering procedure

In this section, the time series clustering procedure proposed to classify the time series of daily mean temperature based on the corresponding distributions for quantile slopes at lower (0.05), middle (0.5) and upper (0.95) quantiles is introduced. The starting point is the panel of time series  $(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(T)})$ . The strategy for clustering the time series is carried out in three stages: firstly, for a fixed (but arbitrary) quantile, the algorithm starts with the estimation of the distribution corresponding to quantile slope estimates; second, the corresponding dissimilarity matrix is computed. Finally, a dendrogram based on the application of classical cluster techniques to the dissimilarity matrix is built and that provides the different clusters formed by the distributions of the quantile slopes. The agglomerative hierarchical method



**Fig. 1.** Map of the stations used in the analysis. The station numbers are as in Table 1.

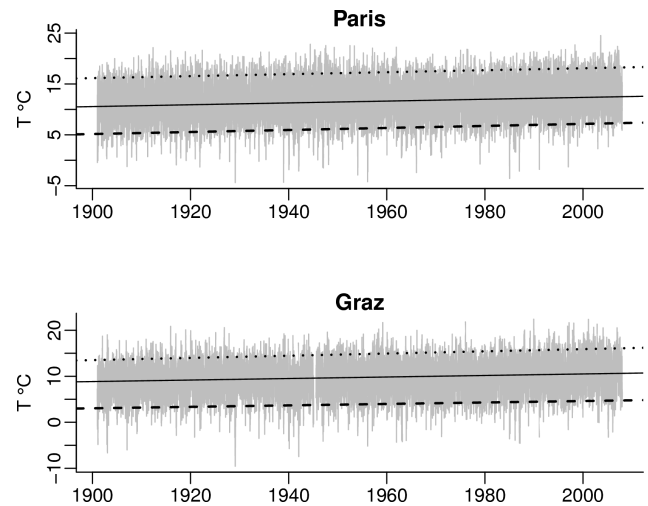
with unweighted average distance (average linkage) is used as grouping criteria.

The clustering procedure is based on the computation of distances among pairs of distribution functions. Therefore an adequate metric between univariate distribution functions is required. The simplest one is the absolute difference among the mean of the distributions, but two distributions can coincide in mean and be completely different in other aspects. Alonso et al. (2006) and Vilar et al. (2010) demonstrated that the choice of metric plays a key role and should reflect the final goal of the clustering procedure. Scotto et al. (2010, 2011), proposed to use the weighted  $L_2$ -Wasserstein distance as a metric since it could be approximated by a fast computational procedure and it has two nice interpretations: (1) as a weighted sum of quantiles squared differences, so it takes into account more than the mean/median behavior and (2) in the case of a Gaussian distribution, a weighted sum of mean differences and standard deviation differences. Since in the present study, the interest is not only in the mean behavior of the distribution but also in the uncertainty related to estimation, this weighted  $L_2$ -Wasserstein distance between two quantile slope distributions has been adopted.

### 3 Results and discussion

#### 3.1 Quantile slopes

Quantile regression has been applied to each temperature record in order to describe the temporal variability of

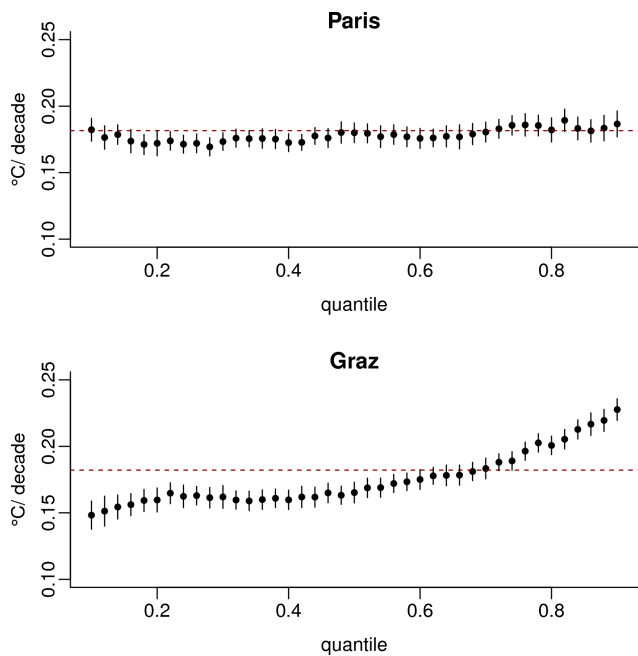


**Fig. 2.** Time series of daily mean air temperature and trends for quantiles 0.05 (dashed line), 0.5 (solid line) and 0.95 (dotted line).

different quantiles of the data distribution. The quantile slopes and corresponding standard errors have been derived using the algorithm of Koenker and D'Orey (1987). As an illustration, detailed results are presented for Paris (PAR) and Graz (GRA). The results for all the stations are presented in Sect. 3.2.

The Paris and Graz records are shown in Fig. 2, along with the quantile slopes at quantiles 0.05, 0.5 and 0.95, corresponding respectively to the lowest 5 %, 50 % (median) and 95 % of the ordered observations. A more complete description of the quantile regression results is given in Fig. 3, which displays for the Paris and Graz records the quantile slopes and the corresponding standard errors computed for quantiles 0.1 to 0.9 in steps of 0.02. Figure 3 clearly shows a distinct pattern for the two records. In the case of Paris, the derived slopes are very similar for all quantiles, and also similar to the usual ordinary least squares slope indicating that the distribution of temperature values is approximately symmetric and the rate of change is the same for all parts of the data distribution. However, in the case of Graz the lower, middle, and upper quantiles behave differently with the upper quantiles of the temperature data distribution increasing at a much faster rate than the middle and lower values.

In order to further assess uncertainty on the quantile slope estimates, a bootstrap procedure is applied. Bootstrap allows us to obtain a distribution of slope values instead of a single punctual estimate, but it is a delicate strategy in the case of non-independent data, since the temporal structure of the time series needs to be preserved in the bootstrap samples. Furthermore, bootstrap procedures often assume stationarity, an assumption which is not verified by most hydro-climatic series. In this work, maximum entropy bootstrap (Vinod and de Lacalle, 2009) is used since it preserves the temporal



**Fig. 3.** Quantile slopes ( $^{\circ}\text{C decade}^{-1}$ ) and corresponding standard errors for Paris and Graz records. The horizontal dashed line represents the usual ordinary least squares slope.

structure of the original series in the bootstrap replicates without assuming stationary behavior. For each time series of daily mean temperature, 200 replicates have been obtained by maximum entropy bootstrap – experiments (not shown) indicate that in this case there is no need of using a bigger number of ensembles. Quantile regression is then applied to each replicate, resulting in a sample of 200 quantile slopes instead of a single punctual estimate. The results are shown as histograms in Fig. 4 and allow the assessment of the dispersion around the punctual quantile slope estimates. Graz exhibits a clear, different distribution for the upper quantile slope relative to the the median and lower slopes. The asymmetry in the bootstrap distributions is a consequence of the asymmetry of the quantile slope statistic.

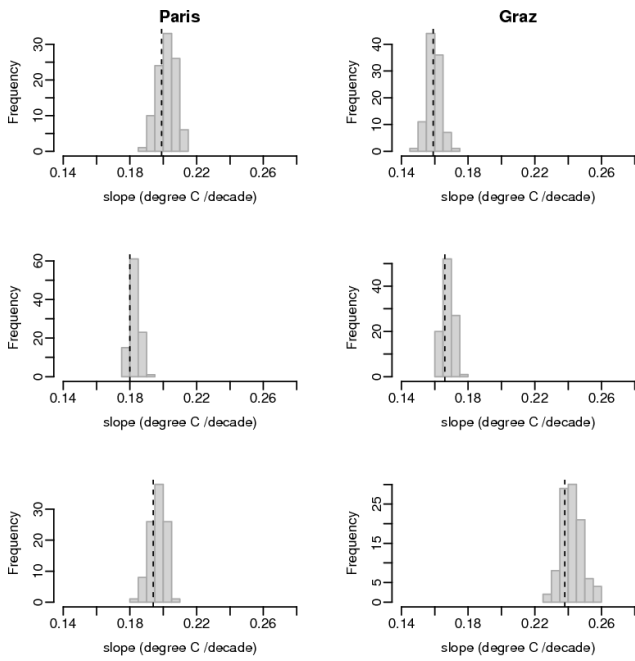
### 3.2 Clustering

The analysis outlined in Sect. 3.1 is repeated for each one of the 28 stations and the resulting punctual estimates are displayed in Table 2. Maximum entropy bootstrap is applied to assess uncertainty, yielding a set of distributions of quantile slope estimates as shown in Fig. 4 for the two selected records. In order to summarise those distributions, the average linkage procedure is applied to obtain dendrograms of slopes for quantiles 0.05 (Fig. 5), 0.5 (Fig. 6) and 0.95 (Fig. 7).

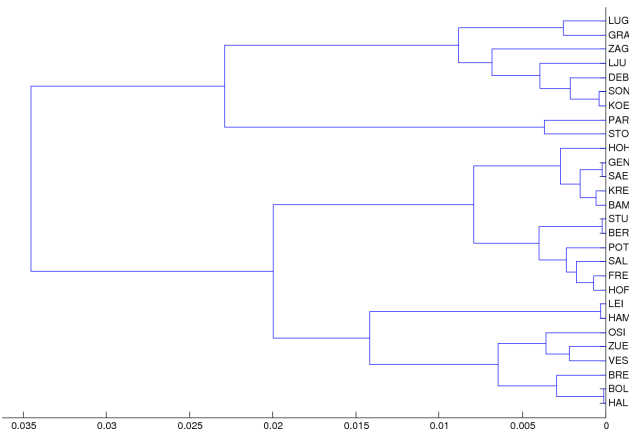
**Table 2.** Quantile slopes ( $^{\circ}\text{C decade}^{-1}$ ) and corresponding standard errors (in parenthesis). The sign \* denotes values that are not statistically significant (for a 5 % significance level).

		$\tau = 0.05$	$\tau = 0.5$	$\tau = 0.95$
1	STO	0.190 (0.016)	0.065 (0.007)	0.029 (0.011)
2	VES	0.051 (0.013)	0.094 (0.005)	0.147 (0.009)
3	KOE	0.135 (0.011)	0.117 (0.006)	0.166 (0.010)
4	HOF	0.102 (0.009)	0.089 (0.005)	0.134 (0.009)
5	HAM	0.016* (0.014)	0.039 (0.007)	0.106 (0.011)
6	BRE	0.037 (0.013)	0.042 (0.007)	0.064 (0.012)
7	BER	0.091 (0.014)	0.068 (0.008)	0.090 (0.013)
8	POT	0.105 (0.014)	0.099 (0.008)	0.127 (0.013)
9	DEB	0.131 (0.013)	0.121 (0.006)	0.153 (0.012)
10	HAL	0.044 (0.015)	0.034 (0.008)	0.049 (0.012)
11	LEI	0.017* (0.015)	0.020 (0.008)	0.034 (0.013)
12	FRE	0.100 (0.012)	0.096 (0.008)	0.124 (0.011)
13	BAM	0.076 (0.013)	0.081 (0.008)	0.100 (0.012)
14	PAR	0.199 (0.012)	0.180 (0.007)	0.194 (0.010)
15	STU	0.091 (0.012)	0.098 (0.008)	0.111 (0.011)
16	KRE	0.077 (0.014)	0.089 (0.007)	0.114 (0.011)
17	SAL	0.096 (0.016)	0.092 (0.009)	0.085 (0.012)
18	HOH	0.073 (0.013)	0.118 (0.010)	0.190 (0.013)
19	ZUE	0.057 (0.011)	0.066 (0.008)	0.115 (0.012)
20	SAE	0.081 (0.015)	0.149 (0.009)	0.242 (0.011)
21	GRA	0.159 (0.013)	0.166 (0.007)	0.238 (0.009)
22	SON	0.136 (0.019)	0.177 (0.008)	0.205 (0.010)
23	GEN	0.081 (0.010)	0.084 (0.007)	0.080 (0.011)
24	LJU	0.143 (0.013)	0.109 (0.007)	0.182 (0.011)
25	LUG	0.152 (0.010)	0.158 (0.005)	0.113 (0.009)
26	ZAG	0.120 (0.013)	0.075 (0.008)	0.148 (0.013)
27	OSI	0.062 (0.016)	0.033 (0.008)	0.069 (0.011)
28	BOL	0.044 (0.011)	0.098 (0.006)	0.148 (0.009)

The dendrogram for the lower quantile clearly discriminates two groups: stations with larger slopes, typically  $>0.1^{\circ}\text{C decade}^{-1}$  (LUG, GRA, ZAG, LJU, DEB, SON, KOE, PAR, STO) and the remaining stations with smaller slopes. The cluster of large slopes further distinguishes the stations with the highest slopes,  $>0.15^{\circ}\text{C decade}^{-1}$  (PAR and STO) from the other stations. The cluster of small slopes further subdivides into stations with moderate slopes and stations with very small or non-significant trends (LEI, HAM, OSI, ZUE, VES, BRE, BOL, HAL). The dendrogram for the median quantile (Fig. 6) first distinguishes the stations with highest slopes,  $>0.15^{\circ}\text{C decade}^{-1}$  (LUG, GRA, SAE, SON, PAR). Within the remaining stations, the major subdivision discriminates the stations with the lowest slopes,  $<0.04^{\circ}\text{C decade}^{-1}$  (LEI, OSI, HAL, BRE, HAM). A similar pattern is found in the dendrogram for the upper quantile (Fig. 7). The first major subdivision clusters the stations with highest slopes (GRA, SAE, SON, LJU, HOH, PAR). The remaining stations are subdivided into clusters of moderate slopes and a cluster of stations with the lowest slopes, typically  $<0.1^{\circ}\text{C decade}^{-1}$  (GEN, SAL, BER, OSI, BRE, HAL,

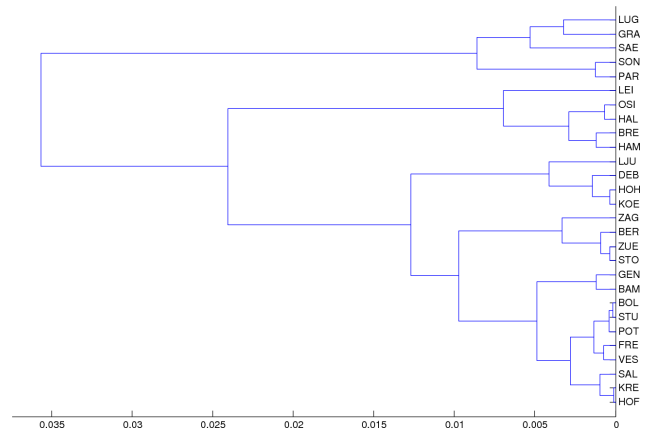


**Fig. 4.** Bootstrap distribution of quantile slopes for Paris (left) and Graz (right) records. From top to bottom: quantile  $\tau = 0.05$ ,  $\tau = 0.5$  and  $\tau = 0.95$ . The dashed vertical line indicates the punctual estimate (Table 2).

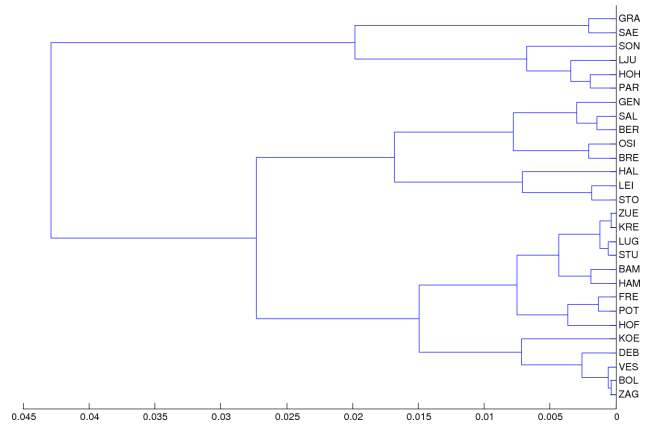


**Fig. 5.** Dendrogram for 5% quantile slopes.

LEI, STO). Although the clustering results mainly reflect the quantile slope values for each station, the advantage of the adopted clustering procedure is that it classifies the distribution of slopes resulting from the bootstrap analysis, instead of a single punctual value.



**Fig. 6.** Dendrogram for 50% quantile (median) slopes.



**Fig. 7.** Dendrogram for 95% quantile slopes.

#### 4 Discussion and conclusions

In the present study, quantile regression and clustering are applied to the study of changes in daily mean air temperature over Central Europe. Quantile regression allows us to compute trends at different quantiles of the data distribution within a well-defined statistical framework. Maximum entropy bootstrap allows the assessment of the uncertainty on the computed slopes while taking into account data serial dependence and non-stationary behavior. Finally, a classical clustering procedure allows us to summarise the resulting distributions of sample quantile slopes.

As in ordinary regression, the slopes for a fixed quantile are not the same across Europe, reflecting the spatial dependence of air temperature trends. For example, in the central part of the data distribution (median quantile), trends vary from  $0.02\text{ }^{\circ}\text{C decade}^{-1}$  to  $0.18\text{ }^{\circ}\text{C decade}^{-1}$ . Furthermore, quantile regression allows us to assess trends at different quantiles of the data distribution. While for some

stations (e.g. PAR, GEN) the trend is the same for all quantiles, the results show that most stations exhibit different slopes for the 5%, 50% and 95% quantiles. This is also reflected in the difference between clusters for different quantiles. Thus, the rate of temporal change is not the same for all parts of the data distribution, as implicitly assumed in ordinary regression. At the 5% quantile, the largest trends are around  $0.15\text{ }^{\circ}\text{C decade}^{-1}$ , while at the 95% quantile the largest trends are around  $0.20\text{ }^{\circ}\text{C decade}^{-1}$ . This indicates a tendency towards larger increases in the upper part of the data distribution (large positive temperature anomalies).

The largest gradient of spatial variability occurs at the 95 % quantile, with slopes ranging from  $0.03\text{ }^{\circ}\text{C decade}^{-1}$  to  $0.24\text{ }^{\circ}\text{C decade}^{-1}$ . The largest trends ( $>0.2\text{ }^{\circ}\text{C decade}^{-1}$ ) are found in the Alps (GRA, SAE, SON). Most stations show a much larger trend in the upper part (95% quantile) than in the central and lower part of the data distribution, consistent with the projected warming for Central Europe (Kjellstrom et al., 2007). The northernmost station, STO, exhibits a contrasting behavior with highest trend ( $0.19\text{ }^{\circ}\text{C decade}^{-1}$ ) at the lowest quantile, decreasing to  $0.03\text{ }^{\circ}\text{C decade}^{-1}$  at the 95 % quantile.

The quantile regression results are presented here in terms of slopes at fixed quantiles for all the stations. An alternative approach would be to perform clustering on the distribution of slopes for all quantiles of each station (as displayed in Fig. 3) in order to better compare changes in slopes across different quantiles. This is computationally very intensive, and further methodological work is required on how to incorporate the uncertainty on the slope estimate at each quantile in such a clustering procedure. However, clustering of slopes across quantiles remains a promising perspective for further work, along with the application of the methodology to outputs from regional climate models.

*Acknowledgements.* This work is supported by FCT (contract under Programme Ciencia2008) and by Acções Integradas Luso-Espanholas (grants E-83/09, HP2008-008 and SEJ2007-64500).

Edited by: R. Lasaponara

Reviewed by: three anonymous referees

## References

- Alonso, A. M., Berrendero, J. R., Hernandez, A., and Justel, A.: Time series clustering based on forecast densities, *Comput. Stat. Data An.*, 51, 762–766, 2006.
- Barbosa, S. M.: Quantile trends in Baltic sea-level, *Geophys. Res. Lett.*, 35, L22704, doi:10.1029/2008GL035182, 2008.
- Baur, D., Saisana, M., and Schulze, N.: Modelling the effects of meteorological variables on ozone concentration – a quantile regression approach, *Atmos. Environ.*, 38, 4689–4699, 2004.
- Cade, B. and Noon, B.: A Gentle introduction to quantile regression for ecologists, *Front. Ecol. Environ.*, 1, 412–420, 2003.
- Della-Marta, P. M., Haylock, M., Luterbacher, J., and Wanner, H.: Doubled length of western European summer heat waves since 1880, *J. Geophys. Res.*, 112, D15103, doi:10.1029/2007JD008510, 2007.
- Elsner, J. B., Kossin, J. P., and Jagger, T. H.: The increasing intensity of the strongest tropical cyclones, *Nature*, 455, 92–95, 2008.
- Frich, P., Alexander, L., Della-Marta, P., Gleason, B., Haylock, M., Tank, A., and Peterson, T.: Observed coherent changes in climatic extremes during the second half of the twentieth century, *Clim. Res.*, 19, 193–212, 2002.
- Hertig, E., Seubert, S., and Jacobeit, J.: Temperature extremes in the Mediterranean area: trends in the past and assessments for the future, *Nat. Hazards Earth Syst. Sci.*, 10, 2039–2050, doi:10.5194/nhess-10-2039-2010, 2010.
- Huguet, F., Parey, S., Dacunha-Castelle, D., and Malek, F.: Is there a trend in extremely high river temperature for the next decades? A case study for France, *Nat. Hazards Earth Syst. Sci.*, 8, 67–79, doi:10.5194/nhess-8-67-2008, 2008.
- Kedem, B. and Fokianos, K.: Regression models for time series analysis, Wiley series in probability and statistics, Wiley, New Jersey, 2002.
- Kjellstrom, E., Barring, L., Jacob, D., Jones, R., Lenderink, G., and Schaer, C.: Modelling daily temperature extremes: recent climate and future changes over Europe, *Climatic Change*, 81, 249–265, 2007.
- Klein Tank, A. M. G., Wijngaard, J. B., Können, G. P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Müller-Westermeier, G., Tzanakou, M., Szalai, S., Pálsdóttir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Bukantis, A., Aberfeld, R., van Engelen, A. F. V., Forland, E., Miletus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., Antonio López, J., Dahlström, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L. V., and Petrovic, P.: Daily dataset of 20th century surface air temperature and precipitation series for the European Climate Assessment, *Int. J. Climatol.*, 22, 1441–1453, 2002.
- Klok, E. J. and Tank, A. M. G. K.: Updated and extended European dataset of daily climate observations, *Int. J. Climatol.*, 29, 1182–1191, 2009.
- Koenker, R.: Quantile regression, Cambridge University Press, New York, 2005.
- Koenker, R. and Basset, G.: Regression quantiles, *Econometrica*, 46, 33–50, 1978.
- Koenker, R. and D'Orey, V.: Computing regression quantiles, *Appl. Stat.*, 36, 383–393, 1987.
- Koenker, R. and Hallock, K.: Quantile Regression, *J. Economic Perspect.*, 15, 143–156, 2001.
- Koenker, R. and Schorfheide, F.: Quantile spline models for global temperature change, *Climatic Change*, 28, 395–404, 1994.
- Mahlstein, I. and Knutti, R.: Regional climate change patterns identified by cluster analysis, *Clim. Dynam.*, 35, 587–600, 2010.
- Nogaj, M., Yiou, P., Parey, S., Malek, F., and Naveau, P.: Amplitude and frequency of temperature extremes over the North Atlantic region, *Geophys. Res. Lett.*, 33, L10801, doi:10.1029/2005GL024251, 2006.
- Parey, S.: Extremely high temperatures in France at the end of the century, *Clim. Dynam.*, 30, 99–112, 2008.
- Patz, J. A., Campbell-Lendrum, D., Holloway, T., and Foley, J. A.: Impact of regional climate change on human health, *Nature*, 438, 310–317, 2005.

- Scotto, M., Barbosa, S. M., and Alonso, A. M.: Model-based clustering of Baltic sea-level, *Appl. Ocean Res.*, 31, 4–11, 2009.
- Scotto, M. G., Alonso, A. M., and Barbosa, S. M.: Clustering time series of sea levels: extreme value approach, *J. Waterw. Port C.-ASCE*, 136, 215–225, 2010.
- Scotto, M. G., Barbosa, S. M., and Alonso, A. M.: Extreme value and cluster analysis of European daily temperature series, *J. Appl. Stat.*, 38, 2793–2804, 2011.
- Trigo, R. M., Pereira, J. M. C., Pereira, M. G., Mota, B., Calado, T. J., Dacamara, C. C., and Santo, F. E.: Atmospheric conditions associated with the exceptional fire season of 2003 in Portugal, *Int. J. Climatol.*, 26, 1741–1757, 2006.
- van den Besselaar, E. J. M., Tank, A. M. G. K., and van der Schrier, G.: Influence of circulation types on temperature extremes in Europe, *Theor. Appl. Climatol.*, 99, 431–439, 2010.
- Vilar, J. A., Alonso, A. M., and Vilar, J. M.: Nonlinear time series clustering based on nonparametric forecast densities, *Comput. Stat. Data An.*, 54, 2850–2865, 2010.
- Vinod, H. D. and de Lacalle, J. L.: Maximum entropy bootstrap for time series: The meboot R Package, *J. Stat. Softw.*, 29, 1–19, 2009.
- Wijngaard, J. B., Tank, A. M. G. K., and Können, G. P.: Homogeneity of 20th century European daily temperature and precipitation series, *Int. J. Climatol.*, 23, 679–692, 2003.
- Zappa, M. and Kan, C.: Extreme heat and runoff extremes in the Swiss Alps, *Nat. Hazards Earth Syst. Sci.*, 7, 375–389, doi:10.5194/nhess-7-375-2007, 2007.