Climate change, trends in extremes, and model assessment for a long temperature time series from Sweden

Journal:	*Environmetrics*
Manuscript ID:	env-09-0137
Wiley - Manuscript type:	Special Issue Paper
Keywords:	climate change, shift function, GEV distribution

Climate change, trends in extremes, and model assessment for a long temperature time series from Sweden

Peter Guttorp[1] and Jia Xu[2]

[1]University of Washington and Norwegian Computing Center

[2]Simon Fraser University

ABSTRACT:

Many problems in current climate research deals with extreme events. Since by definition there are few observations of really extreme events, it is a statistical challenge to assess whether observed trends are significant. In this paper we illustrate one method to look for climate signals in extreme temperature data, and how to compare the data to a climate reconstruction based on a regional model.

Key words: Climate change, GEV distribution, shift function.

1. Stockholm temperature data

One of the longest temperature series in the world has been collected in a park in Stockholm since 1756. While the observing location has been moved twice during this time, it has always been at or near the north wall of the Observatory (Moberg et al., 2002, 2003).

[Figure 1 about here]

The series, shown in Figure 1, has been corrected for the heat island effect on mean temperature as well as for changes in the calculations of daily mean temperature, for one jump due to a miscalibrated thermometer, and for another jump that may possibly be due to the painting of the thermometer screen (Moberg et al., 2003). It is easier to see visually what happens with the extremes. While the maxima look relatively stationary there is some indication that the minima show a slight trend. From the point of view of climate change research, most general circulation models predict that at the latitude of Stockholm we should see an increase in annual minima, a decrease in annual range, and a slight increase in annual mean temperature (IPCC, 2007). We thus will look at low temperatures to see if we can visualize the expected climate signal.

[Figure 2 about here]

Figure 2 shows the minimum annual mean daily temperature for the Stockholm series. The smooth line is a lowess trend line, i.e. a locally weighted local polynomial fit (Cleveland, 1979). There is a clear indication of a continuing increase in minimum temperature through the entire series, although the rate slows down towards the end. Comparing to the corresponding series from nearby Uppsala (cf. Bergström and Moberg, 1997) a similar increase in the minimum annual temperature occurs there from around 1850 on. This increase coincides with the start of a substantial population increase in Uppsala (where population tripled between 1850 and 1890; http://www2.historia.su.se/urbanhistory/cybcity/index.htm). Hence it is possible that the increase in minimum temperature may be related to urbanization and the heat island effect (Akbari, 2001 points out that while Los Angeles mean annual temperature had increased by 2.5K, the annual daily minimum had increased by 4K). Stockholm's population was increasing similarly to Uppsala's, but with the substantial growth starting around 1800 instead of 1850.

[Figure 3 about here]

An alternative way of viewing the series is to look at the point process of extremely cold days. In Figure 3 we show the coldest 0.1% days from the entire series. Since cold weather is associated with high pressure systems that usually last a few days, one would expect the point process to be clustered, which is indeed the case.

The figure suggests that cold winters have become less common since about 1900, but that the number of cold days in a cold winter remains about the same (2–6). However, it would be premature to draw any firm conclusions about a decreasing trend in the rate of the point process.

2. Parametric modeling

We would expect that the negative of the annual minimum temperature should follow a GEV-distribution (Coles, 2001), which has cdf $G(x;\mu,\sigma,\xi) = \exp\left(-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right)$, for arguments x such that the quantity in square brackets is positive. The parameters μ, σ and ξ are measures of location, scale and shape, respectively. After fitting the cdf to the negative of the data it is straightforward to change back to the original scale, for which the location parameter will change sign.

[Figure 4 about here]

Figure 4 shows fitted distributions together with a hanging rootogram (Tukey 1972) of the difference between roughly the first and the second half of the data (before and after 1879). Jarušková and Rencová (2008) analyzed several long European temperature series, looking for change points, either of the jump type or a ramp of hockey stick form. Their analysis indicated that there may be either a jump (around 1901) or a ramp (starting around 1887). The asymptotic standard error of the rootogram, $(2n\Delta)^{-\frac{1}{2}}$, where Δ is the histogram bin width, is computed under the assumption of iid observations (dotted line in right panel of Figure 4). There is some relatively slight autocorrelation in the minimum series, but this should not affect the standard errors substantially. The rootogram plot indicates substantially higher probability of low temperatures in the first half of the data.

[Figure 5 about here]

Given the smooth apparent change in the minimum temperatures, it seems reasonable to try to model a trend in the GEV parameters. A running window estimation (see Figure 5) of the parameters indicates an increasing location parameter, but relatively constant scale and shape parameters. A likelihood ratio test decides in favor of a GEV model having constant shape and scale, and a linear increase in the location parameter of the model (see Gilleland et al. 2009 for details of the model and the fitting procedure). Table 1 shows the various models and the likelihood values. If the linear increase in the location parameter is continued, we would extrapolate the minimum annual temperature to average -9.2°C by the year 2100.

[Table 1 about here]

3. Comparison to a regional climate model

Climate can be thought of as the distribution of weather. A climate model, therefore, does not produce output that is directly comparable (on a day-by-day or even year-by-year basis) to weather data. Rather, one needs to look at the distribution over a number of years of the two outputs. Were climate stationary, the comparison would be more powerful the longer stretch of data we compared. However, since we are looking for indications of changing climate, we need to compare relatively short stretches of data, thus reducing the power of the comparison.

The weather at a given station cannot reasonably be compared to the output of a global climate model, typically operating at a spatial resolution of 3–5 degrees. Instead, regional climate models are used. A regional model is intended to predict local consequences of various climate scenarios (describing greenhouse gas emissions, policy alternatives etc.). It is typically constrained by the output of a global model, which is the only way we can forecast the large-scale consequences of the scenarios. However, when we want to compare regional model output to data, it is useful to constrain the regional model with observed weather data (reanalysis of actual observations using the latest weather forecasting technology). This is the closest a regional model can come to data.

[Figure 6 about here]

For the Stockholm station, we are using the Rossby Center Coupled Regional Climate Model RCA3 (Kjellström et al., 2005), constrained by the European Centre for Medium-Range Weather Forecasts ERA40 reanalysis data (Uppala et al. 2005). The model operates on a time scale of 3 hrs, and a spatial scale of 50 km. In Figure 6 we show a QQ-plot of the observed and model output data. It is clear that the model output is shifted towards higher temperatures. In other words, the regional model is oversmoothing the extremes. The model is tuned to match averages, so expecting it to reproduce the distribution of extremes is perhaps unfair. However, we expect the model to be used to describe probabilities of extreme weather events, and this misfit casts some doubt over its usefulness for this purpose. Doksum and Sievers (1976) introduced the shift function $\Delta(x)$, defined by $F(x+\Delta(x)) \overset{d}{=} G(x)$, to compare two distribution functions F and G. Using the empirical distribution functions F_n and G_n, the natural nonparametric shift function estimate is $\hat{\Delta}(x) = F_n^{-1}\left(G_n(x)\right) - x$. Simultaneous confidence bounds are obtained from the Kolmogorov-Smirnov test statistic (Doksum and Sievers, 1976).

[Figure 7 about here]

In Figure 7 we see that the distribution of the model output apparently is shifted 2–6°C upwards compared to the data. Also, since the horizontal line at height 0 falls outside of the simultaneous confidence band, we are able to reject the hypothesis of no difference between the two distributions at the 95% level. Of course, the regional model is calculating a spatial average of the temperature, while the observations are in a single location. In fact, the regional model averages over separate calculations for water, forested land and open land. The observation is, of course, on open land, which would tend to have the lowest temperature on a cold day. Using several series from the region one could estimate the average minimum temperature over a grid square, as in Meiring et al. (1998) (see also the recent work by Mannshardt-Shamseldin et al. 2009). The discussion in Kjellström et al. (2005) indicates that the model bias may be related to the representation of moisture in clouds and consequent downward longwave radiation.

4. Discussion

The minimum temperature calculated from the data in this paper are the annual minima of daily mean temperatures, calculated as indicated in Moberg et al. (2002). After 1859 there are observations of the actual minimum daily temperature (and these are used in the calculation of the daily mean temperature for these years). Comparing the annual daily minima to the annual minimum daily means (for the years for which we have both series available) it turns out that the actual minima are well approximated by the minimum daily means, shifted down by 3.5 °C. The correlation between the two series is high (0.94). Hence no essential differences would be obtained were one to analyze the shorter series of observed daily minima.

While the annual minimum daily mean temperatures do not exhibit much serial correlation, the daily mean temperature series itself clearly indicates long term memory (Smith 1993). This has important consequences for homogenization techniques. The traditional work by Alexandersson (1986) assumes independent normally distributed values. Simulation studies indicate that the significance level of the hypothesis test for step changes (relative to comparison series that are reasonably well correlated with the series being studied) are quite a bit higher than the nominal size. We intend to pursue this in a later paper (cf. also Lund et al., 2007).

One would perhaps expect the regional model to be better tuned to annual average temperatures than to annual minima. A comparison similar to that in the previous section, indicates that the regional model climate mean annual temperature is shifted up by 1.7° C compared to the observations. This is a known artifact of RCA3, and is likely also due to the cloud representation (Kjellström et al., 2005).

Acknowledgements: The authors are grateful to Daniela Jarušková, Erik Kjellström and Paul Whitfield for helpful comments on the manuscript.

References

Akbari, H. (2005). *Energy Saving Potentials and Air Quality Benefits of Urban Heat Island Mitigation* (PDF) (19 pp, 251K). Lawrence Berkeley National Laboratory. Available at http://www.osti.gov/bridge/servlets/purl/860475-UlHWIq/860475.PDF.

Alexandersson, H., 1986: A homogeneity test applied to precipitation data. *J. Climatol.* **6**, 661–675.

Cleveland, W. S. (1979) Robust locally weighted regression and smoothing scatterplots. *J. Amer. Statist. Assoc.* **74**, 829–836.

Coles, Stuart (2001). *An Introduction to Statistical Modeling of Extreme Values.* Berlin: Springer-Verlag.

Doksum, K. A. (1974) Empirical probability plots and statistical inference for nonlinear models in the two-sample case. *Ann. Statist.* **2**, 267–277.

Doksum, K. A. and Sievers, G. L. (1976) Plotting with confidence: Graphical comparisons of two populations. *Biometrika* **63**, 421–434.

Gilleland, E., Katz R. and Young G. (2009) extRemes: Extreme value toolkit. R package version 1.60. http://CRAN.R-project.org/package=extRemes

IPCC (2007) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Solomon, S., Qin, D., Manning, M. et al. (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jarušková, D. and Rencová, M. (2008) Analysis of annual maximal and minimal temperatures for some European cities by change point methods. *Environmetrics* **19**, 221–330.

Kjellström, E., Bärring, L., Gollvik, S., et al. (2005) *A 140-year simulation of European climate with the new version of the Rossby Centre regional atmospheric climate model (RCA3)*. Reports Meteorology and Climatology 108, SMHI, SE-60176 Norrköping, Sweden, 54 pp.

Lund, R., Wang, X. L., Lu, Q et al. (2007) Changepoint Detection in Periodic and Autocorrelated Time Series. *J. Climate* **20,** 5178–5190.

Mannshardt-Shamseldin, E. C., Smith, R. L., Sain, S. R., Mearns L. D. and Cooley, D. (2009) Downscaling extremes: A comparison of extreme value distributions in point-source and gridded precipitation data. Accepted for *Annals of Applied Statistics.* Available as preprint at http://www.unc.edu/depts/statistics/postscript/rs/ MannshardtShamseldinSmithSain.pdf

Moberg, A., Bergström, H., Ruiz Krisman, J. and Svanerud, O. (2002) Daily Air Temperature And Pressure Series For Stockholm (1756–1998) *Clim. Change* **53,** 171–212.

Moberg, A., Alexandersson, H., Bergström, H. and Jones, P.D. (2003) Were Southern Swedish temperatures before 1860 as warm as measured? *Int. J. Climatology* **23**, 1495–1521.

R Development Core Team (2009) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

R.L. Smith (1993), Long-range dependence and global warming. In *Statistics for the Environmen*t, edited by V. Barnett and F. Turkman, John Wiley, Chichester, 141–161.

Tukey, J. W. (1972). Some graphic and semigraphic displays. In *Statistical Papers in Honor of George W. Snedecor*, ed. T. A. Bancroft. Iowa State University Press.

Uppala, S.M., Kållberg, P.W., Simmons, A.J., et al. (2005) The ERA-40 re-analysis. *Quart. J. R. Meteorol. Soc.* **131**, 2961–3012. doi:10.1256/qj.04.176

Wilk, M. B. and Gnanadesikan, R. (1968) Probability Plotting Methods for the Analysis of Data. *Biometrika* **55**, 1-17.

Table 1. Log likelihood values for GEV fits of the Stockholm annual minima.

Model	Estimated μ	-Log likelihood	Number of parameters
All years, fixed μ	-13.6	705.2	3
Early years, fixed μ	-15.3	350.7	3
Late years, fixed μ	-12.3	335.7	3
Early and late combined		686.4	6
Linear model in μ	-16.3 – -11.2	687.2	4

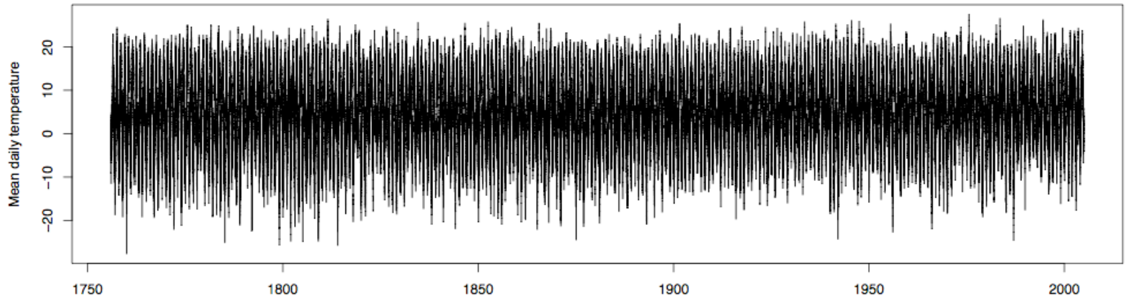

Figure 1: Mean daily temperature readings (°C) from Stockholm Observatory 1756-2004.

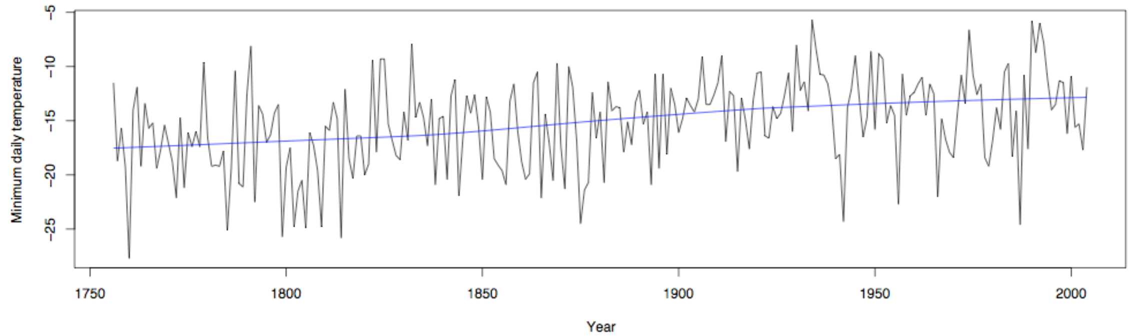

Figure 2. Annual minimum mean daily temperature for Stockholm. Trend curve is a locally weighted polynomial fit, obtained using a default lowess smoother (Cleveland, 1979) in R (R Development Core Team, 2009).

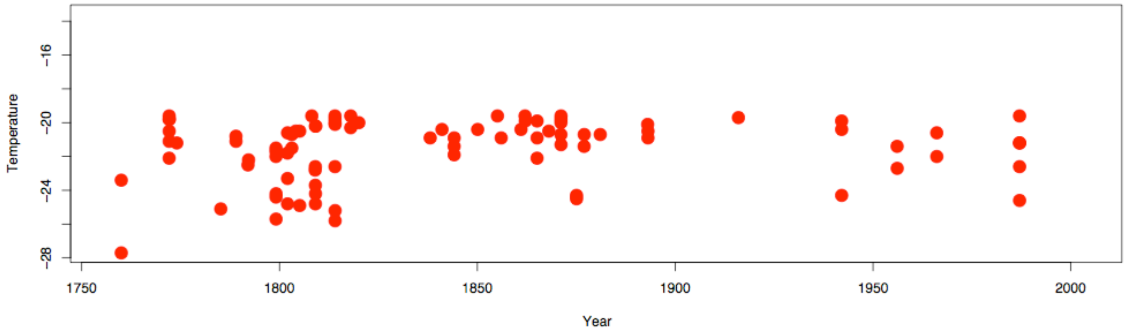

Figure 3. Times of the 0.1% lowest Stockholm temperatures.

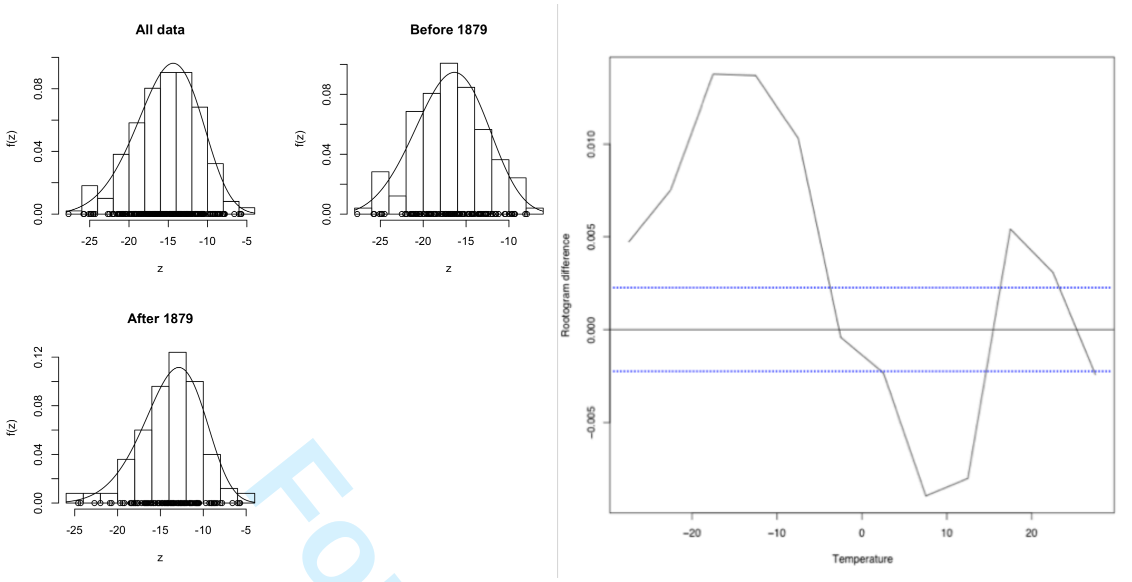

Figure 4. GEV fit using ExtRemes (Gilleland et al. 2009) to all data, first and second half (leftmost three panels). Parameter values are given in Table 1. Hanging rootogram of the difference between first and second half (rightmost panel). The dotted lines are two asymptotic standard errors above and below the x-axis (Tukey, 1972).

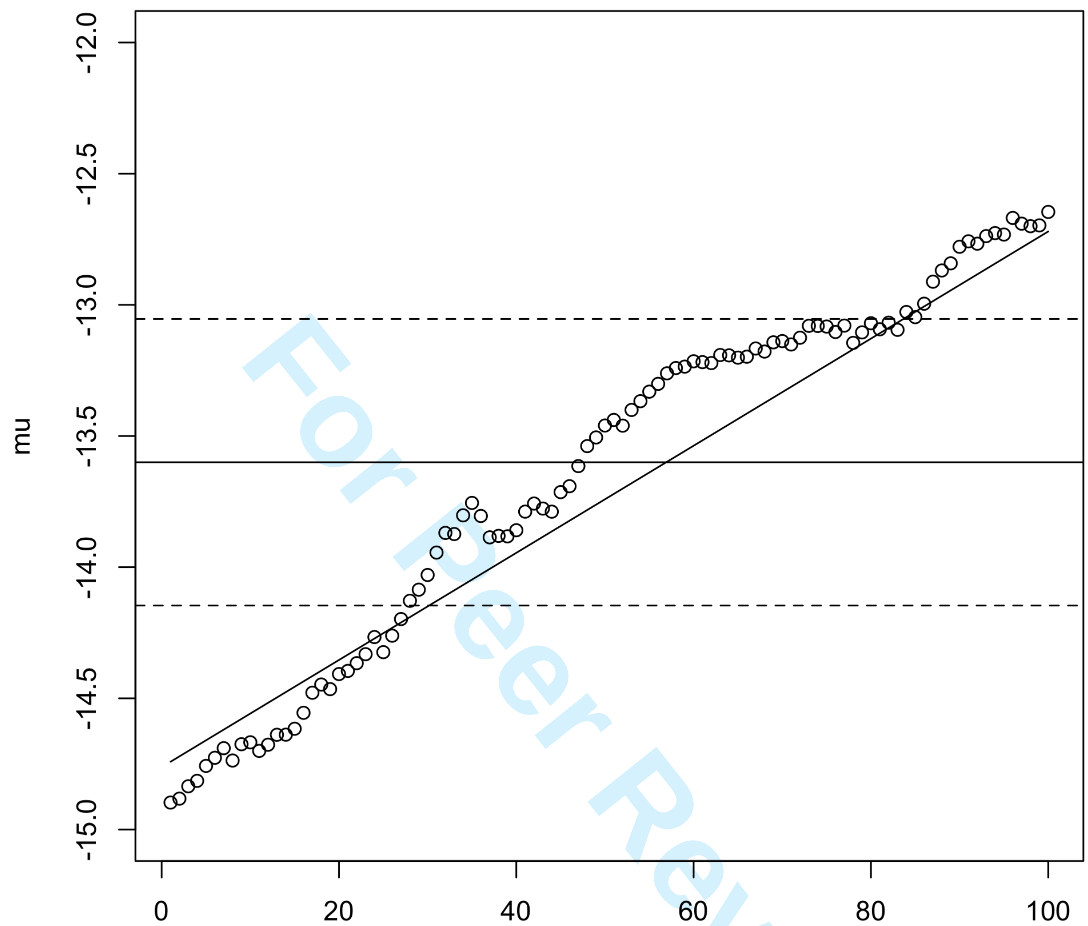

Figure 5. Running estimate of μ with a window size of 149 years. X-axis label indicates the ordinal of the window. The sloped line is the estimated linear mean from extRemes (Gilleland et al., 2009). The horizontal lines are the estimates assuming constant μ (solid) and two standard errors up and down (dashed).

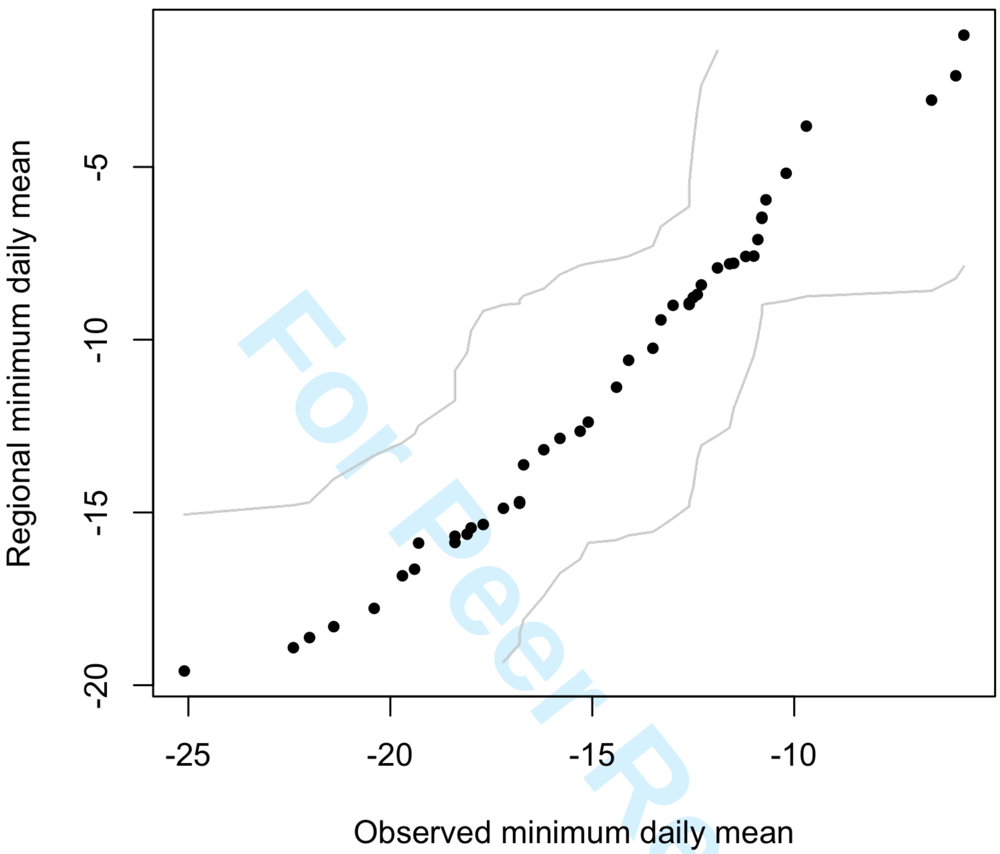

Figure 6. Q-Q plot (Wilks and Gnanadesikan 1968) of observed annual minima in the Stockholm temperature series 1960–2004 and regional model output forced by reanalysis data 1961-2005. The light lines are asymptotic 95% simultaneous confidence bands (Doksum, 1974).

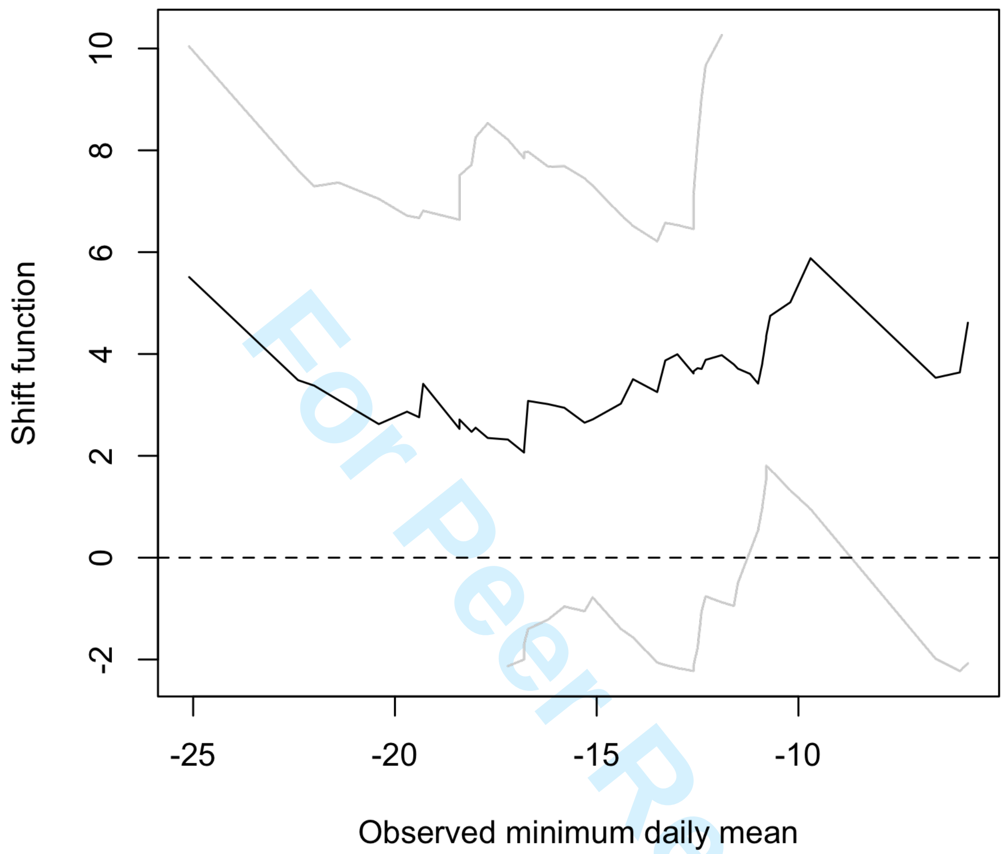

Figure 7. Shift function estimate (solid curve) for the relation between simulated and observed data. The horizontal dashed line corresponds to identical distributions. The light curves are asymptotic 95% simultaneous confidence bands (Doksum and Sievers, 1976).