SFB 649 Discussion Paper 2012-029

Statistical Modelling of Temperature Risk

Zografia Anastasiadou*
Brenda López-Cabrera*

* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

Statistical Modelling of Temperature Risk*

Zografia Anastasiadou, Brenda López-Cabrera†

April 23, 2012

Abstract

Recently the topic of global warming has become very popular. The literature has concentrated its attention on the evidence of such effect, either by detecting regime shifts or change points in time series. The majority of these methods are designed to find shifts in mean, but only few can do this for the variance. In this paper we attempt to investigate the statistical evidence of global warming by identifying shifts in seasonal mean of daily average temperatures over time and in seasonal variance of temperature residuals. We present a time series approach for modelling temperature dynamics. A seasonal mean Lasso-type technique based with a multiplicative structure of Fourier and GARCH terms in volatility is proposed. The model describes well the stylised facts of temperature: seasonality, intertemporal correlations and the heteroscedastic behaviour of residuals. The application to European temperature data indicates that the multiplicative model for the seasonal variance performs better in terms of out of sample forecast than other models proposed in the literature for modelling temperature dynamics. We study the dynamics of the seasonal variance by implementing quantile and expectile functions with confidence corridor to detrended and deseasonalized residuals. We show that shifts in seasonal mean and variance vary from location to location, indicating that all sources of trends other than mean and variance would rise trends over spatial scales. The local effects of temperature risk support the existence of global warming.

Keywords: Weather, temperature, seasonality, variance, global warming, expectile, quantile
JEL classification: G19, G29, G22, N23, N53, Q59

1 Introduction

Recently the topic of global warming has become very popular. The importance of this statement relies on the fact that as temperatures rise, the variability of climate will increase, leading to an increase in temperature extremes, which will translate into significant economic losses.

*The financial support from the Deutsche Forschungsgemeinschaft via SFB 649 "Ökonomisches Risiko", Humboldt-Universität zu Berlin is gratefully acknowledged.

†Institute for Statistics, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany. Email: lopezcab@wiwi.hu-berlin.de (corresponding author)

Part of the literature has concentrated its attention on the economic impact of global warming. Evidence concerning this issue has been emphasized in the works of Mendelsohn et al. (2000), Nordhaus and J. G. Boyer (2000), Horowitz (2001) and Nordhaus (2006) examining the output losses by sector once temperature increases by at least 2.5 or 3.0 degree Celsius. Other part of the literature, where we situate this paper, has concentrated its attention on the evidence of such effect, either by detecting regime shifts or change points in weather time series. See for example Alexander et al. (2006), who present trend analyses for the temperature indices followed by precipitation indices. The majority of these methods are designed to find shifts in mean, but only few can do this for the variance. See for instance the works of Smith and Sardeshmukh (2000), Beniston and Goyette (2007) or Michaels et al. (1998). Another drawback of these methods is that the estimation results rely on the time range of data.

The causes of trends in temperatures are believed to include natural causes or human impact (urbanisation and an increase in the concentration of greenhouse gases in the atmosphere), as well as other factors such as long term variability, the frequency of hazards or extreme events and the solar cycle. Moreover, IPCC (2007) points out that global warming can increase the intensity and frequency of extrem events.

Motivated with the recent findings of IPCC and considering that changes in variance might have greater impact than mean effects, we focus on gradual temperature changes and attempt to find evidence of shifts in seasonal variance of temperature residuals over time. Our reason for studying temperature shifts is simply to better understand the characteristics of observed trends and confirm whether global warming is changing predictability of weather. The majority of weather forecasting literature has based its results of evidence of global warming on structural atmospheric models, IPCC (2007). For such complex systems, an alternative modeling path is given by data-driven (statistical) techniques where the evolution of the system is studied by recording time series. Statistical models like in Campbell and Diebold (2005) have succeeded in the development of non-structural modelling and forecasting of times series trend. Besides density forecast does not necessary require a structural model, but it does require accurate approximations to stochastic dynamics.

Due to the local nature of weather, we follow a time series approach for modelling and forecasting temperature. Our contribution is twofold. First, we present a seasonal mean Lasso-type technique based with a multiplicative structure of Fourier and GARCH terms in volatility. This stochastic model for daily average temperatures identifies well the stylised facts of temperature (seasonality, intertemporal correlations and a variance describing the heteroscedastic residuals) and it captures weather extremes caused by long term climate variability, leading to normal residuals. Second, with help of novel statistical tools, we attempt to find statistical evidence of global warming by detecting upward trends in the seasonal mean of daily average temperatures and in the seasonal variance of temperature residuals over time. We implement quantiles and confidence corridor for expectile functions to the heteroscedastic residuals to show shifts in seasonal variation. The advantage of using such approaches is that are pasimounious and simple, flexible, inexpensive and it is purely data-driven. The quantification of temperature risk is of particular importance for hedgers and traders of weather risk because of the impact of tail events on market prices.

The application to temperature data of the industrial Blue Banana European area in-

dicates that the proposed model provides a better out of sample fit over other models proposed in the literature of modelling temperature dynamics which is of important relevance for the pricing of weather derivatives. The obtained results reveal shifts in seasonal mean and seasonal variation of daily average temperatures, which vary from location to location, indicating that all sources of trends other than mean and variance would rise trends over spatial scales. We conclude that the local effects of temperature risk (local global warming) support the existence of global warming.

This paper is structured as follows. Section 2 shows the tstochastic model for daily average temperature dynamics. Section 3 presents the methodology of quantile regression and expectile functions. The empirical analysis to real data and the performance of the multiplicative model over other models are done in Section 4. The quantile and expectile curves of the seasonal variation over the years are displayed in this section as well. Section 5 concludes the paper. All the computations were carried out in `R` and `Matlab`. The temperature data was obtained from Bloomberg.

2 Temperature Model

There exist many ways to measure climate change. Since our interest is on the statistical analysis of global warming, the most representative variable for the climate variable will be some measurement of long temperature. For our purpose, we equate global warming with a change in average temperature. Inspired by the work on nonstructural modelling and forecasting of times series and due to the local nature of weather, we follow a stochastic model for daily average temperature.

In order to estimate the evolution of temperature in time, the following discrete model for temperature dynamics as in Benth et al. (2007) and Härdle and López-Cabrera (2011) is constructed:

$$T_t = X_t + \Lambda_t \tag{1}$$

- where T_t is the average temperature in day $t, t = 1, ..., T$. T_t is computed as $T_t = \frac{T_{t,max}+T_{t,min}}{2}$.
- Λ_t is the seasonal function which is nonparametric approximated with a series of basis functions and a Lasso penalty estimator (see Tibshirani (1996)):

$$\arg\min_{\beta} \sum_{i=1}^{T} \|T_t - \beta\Psi(t)\|^2 + \lambda\|\beta\|_1 \tag{2}$$

  where T_t is a vector of daily averages temperatures, $\Psi(t) = (\varphi_1(t), \ldots, \varphi_K(t))^\top, \varphi_k : 1 \leq k \leq K)$ is a vector of known basis functions and λ is a penalty term which shrinks the unknown Fourier coefficients $\beta = (\beta_1, \ldots, \beta_K)$ to zero.

- X_t is a p-order autoregressive process $AR(p)$

$$X_{t+p} = \sum_{i=1}^{p} \alpha_i X_{t+p-i} + \varepsilon_t, \varepsilon_t = \sigma_t \eta_t \tag{3}$$

where η_t is white noise and σ_t is the smooth seasonal volatility. σ_t is assumed to follow a seasonal pattern as well.

The proposed Lasso estimator in (2) captures the global trend in time with orthogonal Legendre polynomial basis:

$$\varphi_1(t) = 1, \varphi_2(t) = t, \varphi_3(t) = 3t^2 - 1 \tag{4}$$

and periodic variations with Fourier series:

$$\begin{aligned} \varphi_4(t) &= sin(2\pi t/p), & \varphi_5(t) &= cos(2\pi t/p), ... \\ \varphi_6(t) &= sin(2\pi t/(p/2)), & \varphi_7(t) &= cos(2\pi t/(p/2)), ... \end{aligned} \tag{5}$$

where $p = 365$ (seasonal effects) or $p = 365 \cdot 11$ (large period effects). The period of 11 years indicates the 11-year solar activity cycle, according to meteorologists, see Racskoa et al. (1991) or Parton and Logan (1981). The number of basis functions K is usually region/climate specific. Other approaches to estimate (2) are given in Benth et al. (2007) model, who uses Fourier series, or in Benth et al. (2011), who considers Local Linear Regression (LLR) for the seasonal mean:

$$\arg \min_{e,f} \sum_{i=1}^{365} \{\bar{T}_t - e_s - f_s(t-s)\}^2 K\left(\frac{t-s}{h}\right) \tag{6}$$

where $\bar{T}_t$ is the mean over years of daily averages temperatures, h is the bandwidth, $K(\cdot)$ is a Kernel. The advantage of LLR estimator is that no global function is required for the model fitting.

Before checking inter-temporal correlations with an autoregressive process $AR(p)$, defined in (3), the process X_t has to be tested for stationarity. For that reason two tests are applied, Augmented Dickey-Fuller (ADF) test for a unit root and KPSS test for stationarity. If H_0 of ADF test is rejected and the H_0 of KPSS test cannot be rejected then X_t is a stationary process and can be modeled with an $AR(p)$. The order p is suggested by plotting the Partial Autocorrelation Function (PACF) of X_t and confirmed by the Bayesian Information (BIC) Criterion see Hurvich and Tsai (1989).

Empirical work has shown that the process in (3) is a heteroscedastic process with periodic pattern. Therefore the empirical variance σ_t^2 is estimated as follows: divide the residuals into 365 groups, so that each group corresponds to each day over all year, then for each group estimate the average of squared residuals and finally smooth the curve. A one step smoothing model for σ_t^2 has been proposed in Benth et al. (2007) as:

$$\hat{\sigma}^2_{t,FTS} = c_1 + \sum_{i=1}^{L} \left\{ c_{2i} \cos\left(\frac{2i\pi t}{365}\right) + c_{2i+1} \sin\left(\frac{2i\pi t}{365}\right) \right\} \tag{7}$$

Since the optimal choice of L in (7) depends on the region/climate to be considered, Benth et al. (2011) proposed to smooth the data with a Local Linear Regression (LLR), $\hat{\sigma}^2_{t,LLR}$ estimator, thus reducing the number of parameters to be estimated:

$$\hat{\sigma}^2_{t,LLR} = \arg \min_{a,b} \sum_{i=1}^{365} \left\{\hat{\varepsilon}_t^2 - a(t) - b(t)(t-t_0)\right\}^2 K\left(\frac{t-t_0}{h}\right) \tag{8}$$

where $\hat{\varepsilon}_t^2$ is the average of squared residuals on each day over all years, h is a bandwidth and $K(\cdot)$ is a Kernel.

Alternatively, Campbell and Diebold (2005) consider an additive model of a truncated Fourier function and a GARCH process of the form:

$$\hat{\sigma}^2_{t,FTSG} = c_1 + \sum_{i=1}^{L} \left\{ c_{2i} \cos\left(\frac{2i\pi t}{365}\right) + c_{2i+1} \sin\left(\frac{2i\pi t}{365}\right) \right\} + \alpha_1 \varepsilon^2_{t-1} + \beta_1 \sigma^2_{t-1} \tag{9}$$

Another approach is given in Härdle et al. (2011), who show that the GARCH effects in (9) are small and therefore propose a local adaptive modelling approach to find at each point, an optimal smoothing parameter to locally estimate the volatility. In a recent paper, Benth and Saltyte Benth (2010) suggest a multiplicative model of Fourier and GARCH terms in volatility:

$$\hat{\sigma}^2_{t,MFTSG} = \left[c_1 + \sum_{i=1}^{L} \left\{ c_{2i} \cos\left(\frac{2i\pi t}{365}\right) + c_{2i+1} \sin\left(\frac{2i\pi t}{365}\right) \right\} \right] * \left(\alpha_1 \varepsilon^2_{t-1} + \beta_1 \sigma^2_{t-1} \right) \tag{10}$$

This multiplicative approach is motivated by the fact that variance is positive and GARCH effects are still found as an exponentially decaying autocorrelation function of squared standardized residuals $(\frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,FTS}})^2$, where $\hat{\varepsilon}_t$ are the residuals estimated by the $AR(p)$ process.

3 Methodology

3.1 Local Quantile Regression

Shifts of seasonal variance of temperature residuals can be detected with the application of local quantile regression. We follow Härdle et al. (2011) proposal, of an adaptive local quantile regression algorithm. It was shown, that quantile curves are good indicators for finding shifts in variance of local temperature residuals.

The τth quantile curve is given by the following formula:

$$Y_i = l(X_i) + \epsilon_t \tag{11}$$

with $P(\epsilon_t > 0) = \tau$ and $l(x)$, the conditional quantile function $F^{-1}_{Y|x}(\tau)$ which can be approximated by a polynomial. Y_i and X_i, with $i = 1, \ldots, n$, are independent random variables and $\tau \in (0, 1)$. The adaptive part comes from the bandwidth selection.

3.2 Expectile Function

The seasonal variation from the fitted temperature model of equation 3 can also be analyzed by expectile curves (EC). The τ-conditional expectile $v_\tau(x)$, $0 < \tau < 1$, given x, is defined as

$$v(x) = \arg\min_{\theta} \mathrm{E}\left\{ \rho_\tau(y - \theta) | X = x \right\}, \tag{12}$$

where $\rho_\tau(u) = |\mathbf{1}(u \leq 0) - \tau| u^2$ is the loss function. Note that $\rho^*_\tau(u) = |\mathbf{1}(u \leq 0) - \tau| u$ leads to quantile regression framework. Guo and Härdle (2012) introduced the localized

nonlinear smoother $v_n(x)$ of the expectile regression curve and constructed confidence corridor around the estimated expectile function of the conditional distribution of Y given x. The advantage of expectiles over quantiles is that they capture the extreme events reported in the data - the special behavior of non-average observations.

4 Empirical Analysis

In this section we present the results of the empirical analysis to real data, the model validation and the shifts of the seasonal variance of temperature residuals via confidence corridor of expectile curves and quantiles.

4.1 Data

IPCC (2007) points out that global warming has been detecting in two periods: the pre-1946 and the post-1976 period. We limit our study to the second period. The modeling of temperature dynamics is implemented on four cities of the so called "Blue Banana" European area: Amsterdam, London, Paris and Rome. We are interested on these cities since the "Blue Banana" area covers one of the highest concentrations of people, money and industry in the world and the aim of our analysis is to detect evidence of global warming, as shifts in mean and variance, in this specific indusrial territory of Europe.

The temperature data for each city contain daily average temperatures T_t, measured in °C and are defined as $T_t = \frac{T_{t,max}+T_{t,min}}{2}$. The observations are from January 1, 1973 to October 10, 2009. There were 0.03% missing values in the data which were handled by computing the mean of the time neighbouring observations. The leap years were removed in order the estimation procedure to be more consistent. Figure 1 summarises the monthly average temperatures and monthly temperature means of the available datasets. The data for each city is splitted in two datasets. The first data set consists of 13140 observations from January 1, 1973 to December 31, 2008 (in-sample data) which is used for model estimation. The second data set consists of 283 observations from January 1, 2009 to October 10, 2009 (out-of-sample data) and is used for model validation.

4.2 Temperature Dynamics modelling

We now proceed with the proposed methodolgy. We first correct the seasonality in mean. Figure 2 shows the LLR estimator in (6), which is fitted to the data using the Epanechnikov Kernel and a bandwidth proposed by Bowman and Azzalini (1997). We note, that a clear evidence of an upward trend of the temperature time series is not visible with this estimator. We next fit the seasonal Lasso type in (2), with basis functions given in (4) and (5). We observe in Figure 3 how the penalty term λ shrinks the coefficients β's of the basis functions $\Psi(t)$ for the four European cities.

Therefore the resulted seasonal models for London ($\Lambda_{t,L}$), Rome ($\Lambda_{t,R}$), Paris ($\Lambda_{t,P}$) and Amsterdam ($\Lambda_{t,A}$) consist of linear and quadratic terms (trend) and a seasonal part (mix-

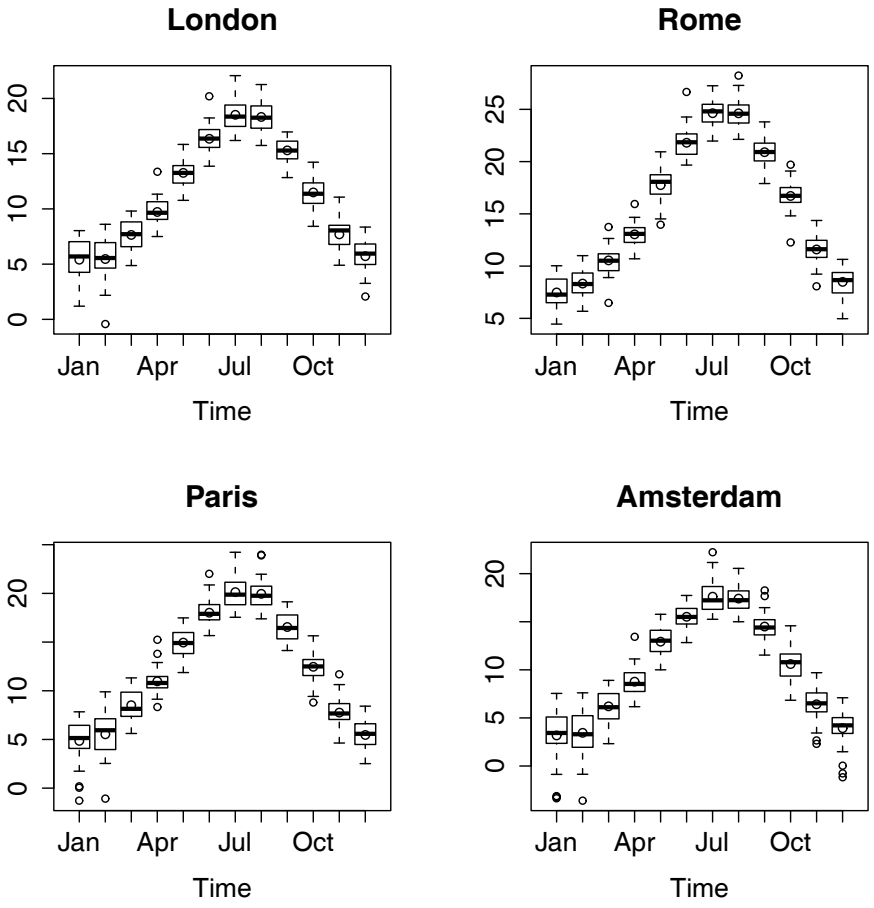

Figure 1: Monthly average temperatures and monthly temperature means (circles) for London, Rome, Paris, Amsterdam from 01.01.1973 - 31.12.2008.

ture of cos, sin):

$$\begin{aligned}
\Lambda_{t,L} &= \beta_1 + \beta_2 t + \beta_5 cos\{2\pi(t-\zeta_1)/365\} + \beta_7 cos\{4\pi(t-\zeta_2)/365\} \\
&+ \beta_9 cos\{10\pi(t-\zeta_5)/365\} \\
\Lambda_{t,R} &= \beta_1 + \beta_3(3t^2-1) + \beta_5 cos\{2\pi(t-\zeta_1)/365\} + \beta_7 cos\{4\pi(t-\zeta_2)/365\} \\
&+ \beta_9 cos\{10\pi(t-\zeta_5)/365\} \\
\Lambda_{t,P} &= \beta_1 + \beta_2 t + \beta_5 cos\{2\pi(t-\zeta_1)/365\} + \beta_7 cos\{4\pi(t-\zeta_2)/365\} \\
&+ \beta_9 cos\{10\pi(t-\zeta_5)/365\} \\
\Lambda_{t,A} &= \beta_1 + \beta_3(3t^2-1) + \beta_5 cos\{2\pi(t-\zeta_1)/365\} + \beta_7 cos\{4\pi(t-\zeta_2)/365\} \\
&+ \beta_{11} cos\{10\pi(t-\zeta_7)/(365\cdot 11)\}
\end{aligned}$$

The Lasso algorithm in (2) takes automatically, for every location, the basis functions that minimize the sum of squared errors. The parameters $\{\beta_k\}_{k=1}^K$ of the proposed models are estimated with the non-linear least squares algorithm and are displayed in Table 1. The interpretation to each parameter is defined as follows. The estimated parameter β_1 stands for the average of the temperature, which is higher for Rome as one would expect. The coefficients β_2 and β_3 of linear and quadratic terms represent the global warming effect. Parameters β_5, β_7, β_9 and β_{11} are the maximum displacements of the periodic terms

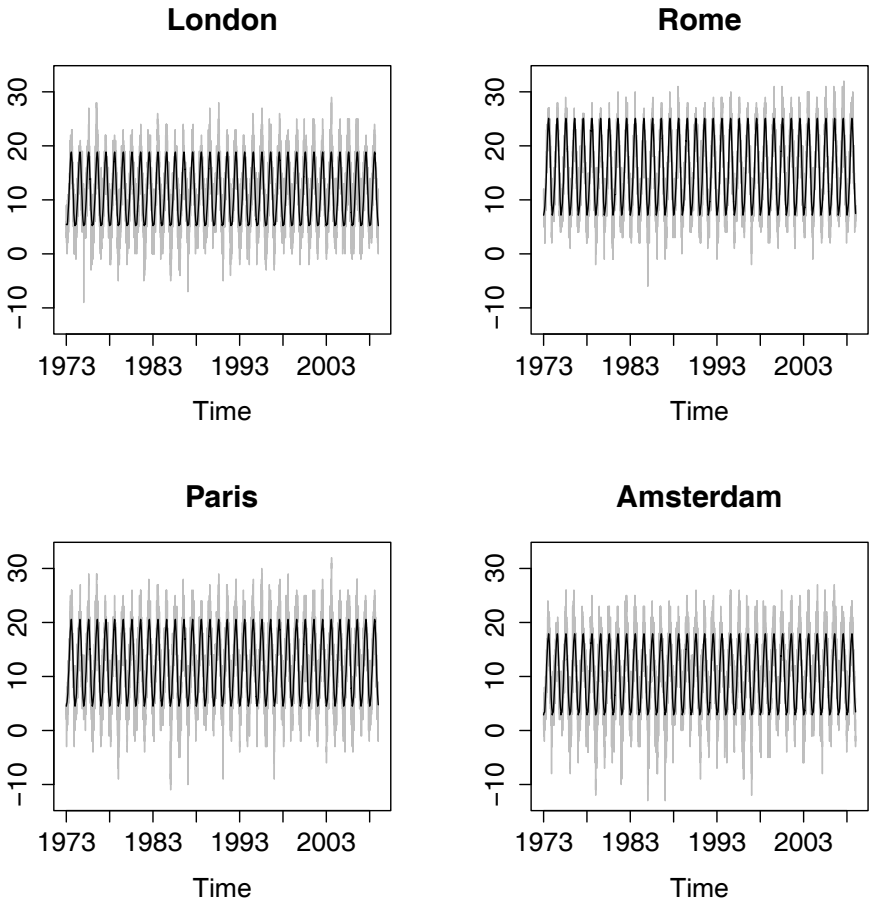

Figure 2: Daily average temperatures (gray line) and seasonality effect (black line), estimated with LLR, for the four european cities.

(cosine functions) while ζ_1, ζ_2, ζ_5 and ζ_7 are their shifts. Since most of the coefficients are positive, one can claim that there is an evidence for global warming. Figure 4 also shows an upward trend of the temperature for all cities, particularly for Amsterdam.

Before we fit the AR process to the deseasonalized residuals ($X_t = T_t - \Lambda_t$), we check whether the residuals are stationary with the ADF test and KPSS test. According to Table 1, both tests indicate that X_t is stationary. Since significant inter-temporal correlations and partial autocorrelations were identified in residuals, AR(p) models with order p were fitted according to the BIC, see Figure 5. The estimated coefficients are also reported in Table 1 and are consistent with other studies in temperature data.

The ACF of squared residuals of the AR(p) process (Figure 7) present a clear seasonal pattern which motivates us to calibrate σ_t^2 from the models in (7), (8), (9) and (10). Figure 6 displays that the seasonal variation is higher in winter and autumn and lower during the summer for all cities. The resulted ACF of the squared residuals after removing volatility (Figure 8) reveal that the multiplicative model removes seasonal pattern as good as the other estimators. Table 2 reports the normality tests (Anderson-Darling (AD) and the Jarque-Bera (JB)) for the standardized residuals. We find that, the multiplicative model with GARCH effects isolates gaussian factors, in particular for cities like London and Rome.

Parameters	London (p=4)	Rome (p=3)	Paris (p=3)	Amsterdam (p=3)
β_1	10.752	15.019	11.797	9.389
β_2	$8 \cdot 10^{-5}$	-	$5 \cdot 10^{-5}$	-
β_3	-	$3 \cdot 10^{-9}$	-	$4 \cdot 10^{-9}$
β_5	6.806	8.725	7.826	7.444
β_7	-8.208	-0.933	-7.222	-0.460
β_9	243.601	0.244	0.296	-
β_{11}	-	-	-	-0.366
ζ_1	204.307	207.161	201.241	203.146
ζ_2	126.571	131.614	136.806	136.171
ζ_5	1.498	146.65	-1095.272	-
ζ_7	-	-	-	305.919
α_1	0.759	0.818	0.909	0.888
α_2	-0.070	-0.085	-0.194	-0.187
α_3	0.016	0.033	0.065	0.084
α_4	0.036	-	-	-
ADF: $\hat{\tau}$(p-value)	-20.29(< 0.010)	-18.67(< 0.010)	-20.66(< 0.010)	-20.05(< 0.010)
KPSS: $\hat{k}$(p-value)	0.167(< 0.100)	0.094(< 0.100)	0.221(< 0.100)	0.070(< 0.100)

Table 1: Estimated parameters with nonlinear least squared of the seasonality models w.r.t. the basis functions selected by Lasso. ADF and KPSS stationarity tests. $\alpha's$ coefficients of the AR(p) process (the order p is displayed at the top of the table).

City		$\frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,FTS}}$	$\frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,LLR}}$	$\frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,FTSG}}$	$\frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,MFTSG}}$
London	AD	13.724	13.268	11.527	10.649
	JB	593.779	360.609	61.531	78.338
	Kurtosis	4.034	3.797	3.216	3.272
	Skewness	0.062	0.076	0.128	0.131
Rome	AD	18.382	16.482	14.449	10.770
	JB	615.057	509.693	405.052	197.498
	Kurtosis	4.036	3.943	3.853	3.571
	Skewness	-0.113	-0.102	-0.054	-0.094
Paris	AD	0.952	1.010	1.975	1.703
	JB	13.121	12.237	26.768	22.339
	Kurtosis	2.960	2.933	2.797	2.832
	Skewness	-0.074	-0.067	-0.044	-0.056
Amsterdam	AD	9.354	9.015	7.270	7.032
	JB	57.701	50.169	25.782	23.804
	Kurtosis	3.253	3.221	3.051	3.088
	Skewness	0.102	0.103	0.105	0.095

Table 2: Anderson-Darling (AD) and Jarque Bera (JB) normality tests as well as skewness and kurtosis for the standardized residuals under different models.

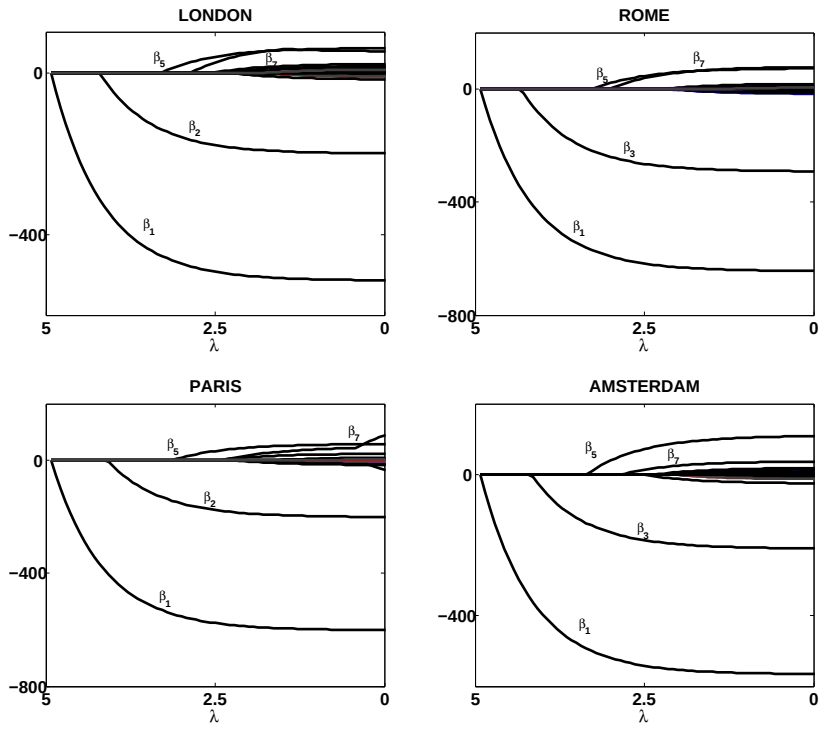

Figure 3: Shrinkage of coefficients via the Lasso penalty for the four European cities.

4.3 Model Validation

We generate one step ahead predictions from the 283 out of sample observations (January 1, 2009 to October 10, 2009) for the multiplicative model of Fourier and GARCH terms. The observed and predicted values are shown in Figure 10. The deviations between the circles and the discs correspond to the prediction errors (PE). The lines correspond to the 95% pointwise confidence intervals. Table 3 shows clearly that the normality of PEs can not be rejected at 5% significance level for all analyzed cities. The kurtosis and skewness of PEs are reported in Table 3. Since the PEs' skewness of London and Amsterdam is greater than 0 we conclude that the prediction values, derived by the fitted model, are more often below the observed temperature. PEs' of Rome is negative therefore the forecasted temperatures are more often above the real observations. For Paris the skewness is close to 0. The kurtosis of the PEs' distributions is leptokurtic for all cities. Moreover, QQ-plots of Figure 11 suggest that the PEs are close to the normal distribution with Paris to satisfy the best approximation.

To test the out-of-sample forecast, we apply the the root mean squared prediction error (RMSE) given by:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2}$$

and the mean absolute error (MAE) defined by Hyndman and Koehler (2006) as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |e_i|$$

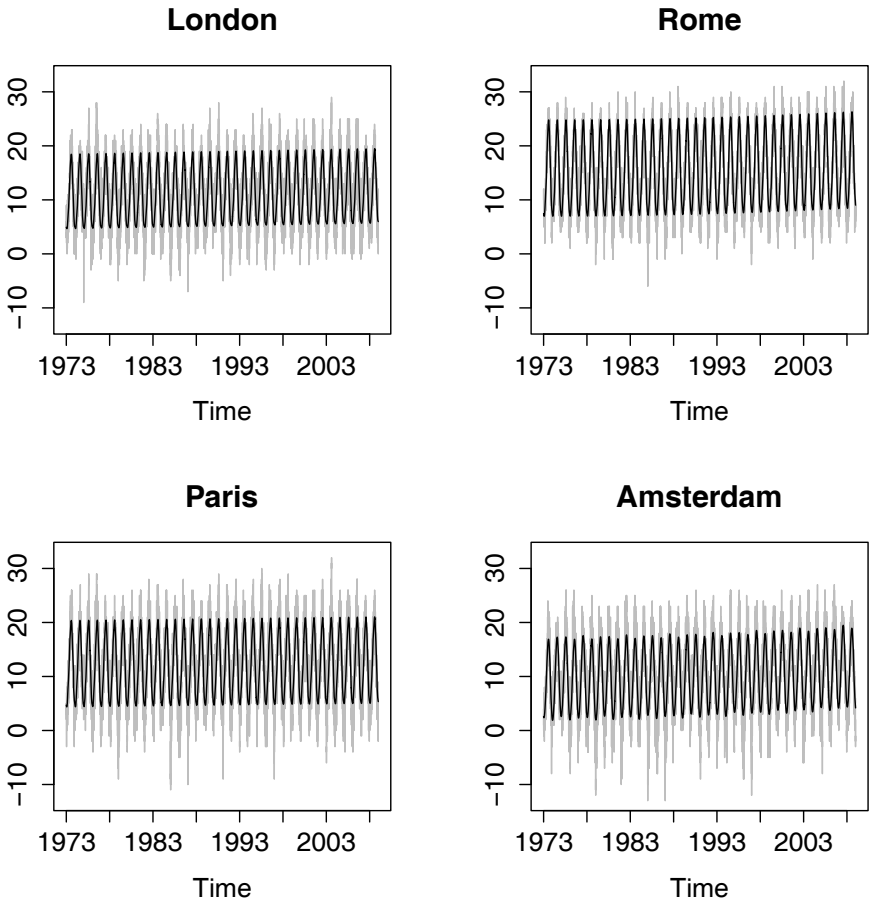

Figure 4: Daily average temperatures (gray line) and seasonality effect (blackline), estimated with Lasso, for the four European cities.

We describe the accuracy of the one-day-ahead forecasts in Table 3. The results for Rome outperform the ones of London, Paris and Amsterdam. The prediction power of the fitted model for the other cities is relatively similar. Additionally from Table 3 it is clear the RMSE and MAE have very small values. Therefore, we conclude that the multiplicative model gives us quite precise one day ahead predictions and it is a good model for forecasting.

Moreover, 95% and 80% predictions intervals (PI) were calculated from the model. Concerning the calculation of PI 283 random innovations were generated. Secondly, a series of values was built from the model. This iteration was repeated 1000 times, 1000 realisations of the model were simulated. The PI were then computed as a corresponding pointwise (for all 283 data points) empirical quantile. In the next step we once more simulated 1000 trajectories and investigated the robustness of constructed PI. Table 3 shows additionally the percentages of the simulated observations which lie outside the constructed PI.

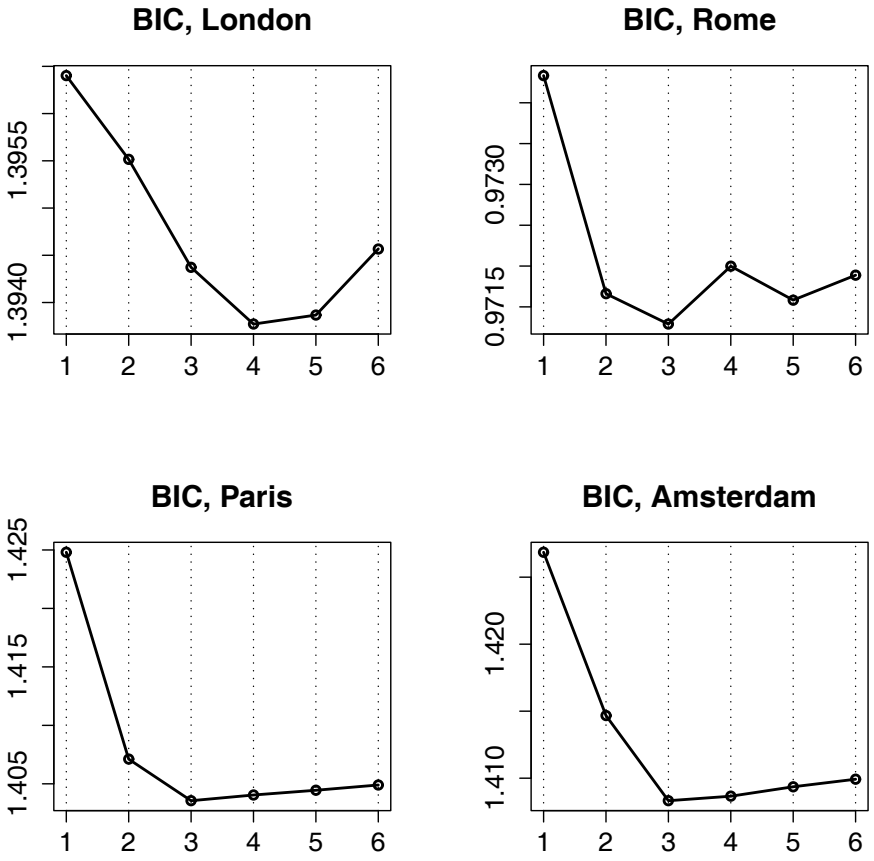

Figure 5: Bayesian Information Criterion (BIC) for the four european cities after the seasonal correction.

4.4 Testing shifts of variance over time

In this section we attempt to find the trends of the variances of residuals of daily average temperature over time. This is achieved with the help of ecpectile and quantile curves.

Since the empirical seasonal variance relies on the nature of the grouping resulted residuals of equation (3), we follow the methodology of Guo and Härdle (2012) and apply the expectiles (ECs) to the residuals for each 12 year period. In this case, we have $X = 1, \ldots, 365$ denotes the day of the year and Y are the model residuals within each 12-year subsample.

City	JB (p-value)	Kurtosis	Skewness	$RMSE$	MAE	95%	80%
London	44.733 (< 0.001)	4.889	3.235	1.999	1.489	5.069	20.076
Rome	27.326 (< 0.001)	4.075	-0.539	1.486	1.178	5.063	20.008
Paris	7.176 (0.028)	3.769	-0.067	2.025	1.555	5.052	20.055
Amsterdam	12.576 (0.002)	3.945	0.208	1.934	1.475	5.095	20.053

Table 3: Jarque-Bera Tests, kurtosis and skewness of prediction errors, forecast accuracy measures and prediction intervals (PI).

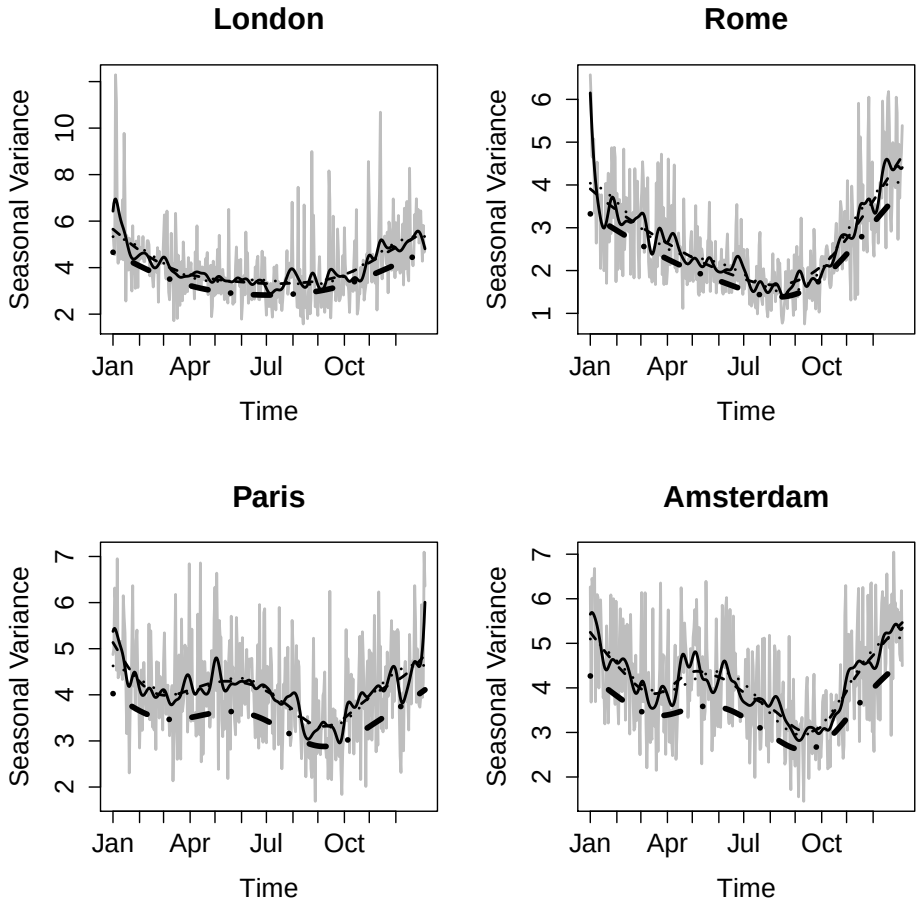

Figure 6: Estimated empirical seasonal variance (gray line), $\hat{\sigma}_{t,LLR}$ (solid line), $\hat{\sigma}_{t,FTS}$ (dotted line), $\hat{\sigma}_{t,FTSG}$ (dashed line) and $\hat{\sigma}_{t,MFTSG}$ (dashed with points line) for each city.

The upper panels of Figures 12, 13, 14 and 15 depict the estimated 0.9-EC for the seasonal variance in London, Rome, Paris and Amsterdam for different periods: $1973-1984$ (solid lines), $1985-1996$ (dashed-dotted lines), $1997-2008$ (dashed lines). For the sake of brevity the fitted $5\%-95\%$ confidence corridor is displayed only for one expectile curve. We attribute 0.9-EC to extreme temperatures, squared model residuals, observed within the sample. Analogously, the lower panels of 12, 13, 14 and 15 display the 0.1-EC and denote the smallest squared residuals, the observations well explained by the model of equation 3. It is worth to notice here, that expectiles by definition (12) are robust for very high and very low τ.

For each of the cities the ECs have similar spatial structure to seasonal variance curves in Figure 6: the variance is significantly higher for the winter-fall period. The maximum variance occurs mostly in January, and the lowest variation is reported in July. The structure of fitted expectiles show differences across the cities.

The most interesting findings coming from the fitted expectiles are the differences within each 12-year period. The extreme temperatures revealed by the 0.9-EC differ significantly over each subsample. We report that for any of the ECs fitted for the different periods, does not lay within the $5\%-95\%$ confidence corridor of the other EC. In general, except for London, the values of the ECs grow over time: the ECs for the period $1973-1985$ (solid

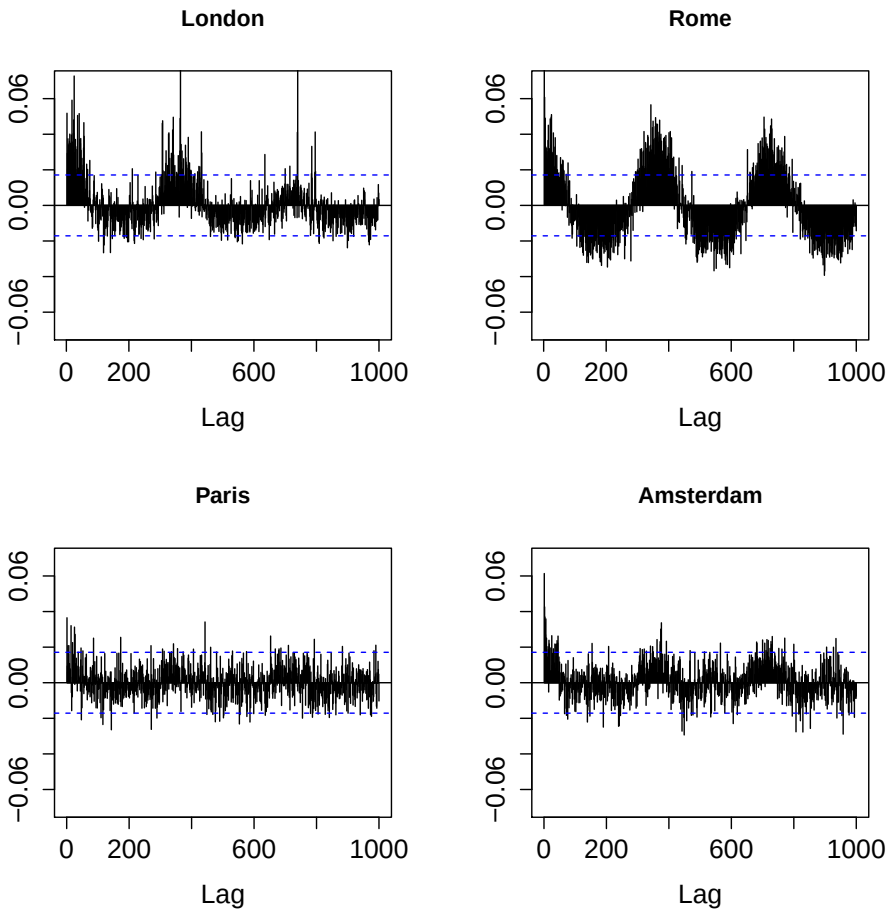

Figure 7: ACF of squared residuals after removing the seasonality and trend.

line) lay below the curves of periods $1985-1996$ (dashed-dotted lines) and $1997-2008$ (dotted lines). The ECs of the period $1997-2008$ are significantly higher than others. These seasonal results are consistent with the findings reported by the IPCC about global warming effect. The findings hold for most of the studied cities, indicating that all sources of trends other than mean and variance would rise trends over spatial scales. This means that high temperature is increasing far more often. The exception of London might be explained by the extensive human activities and industries localized in London area within $1973-1985$.

The study of the low, 0.01-ECs do not reveal significant differences within different periods. All of the fitted ECs differ significantly over each subsample. We report that the expectile lines fitted for different periods are not located within the $5\%-95\%$ confidence corridors. Moreover there is no seasonal pattern and curves do not fluctuate much within a year. The only exception is Rome, what might be attributed to the higher temperatures reported there, in comparison to Amsterdam, Paris and London.

The quantile regression, similarly to the expectiles, is applied to the daily squared residuals over each 12 year period after taking out seasonal and AR effects. X_i, with $i = 1, \ldots, n$ are the days of each year and Y_i are the daily squared residuals within each 12-year period.

Figures 16, 17, 18 and 19 display the estimate quantile curves (lines) and daily average

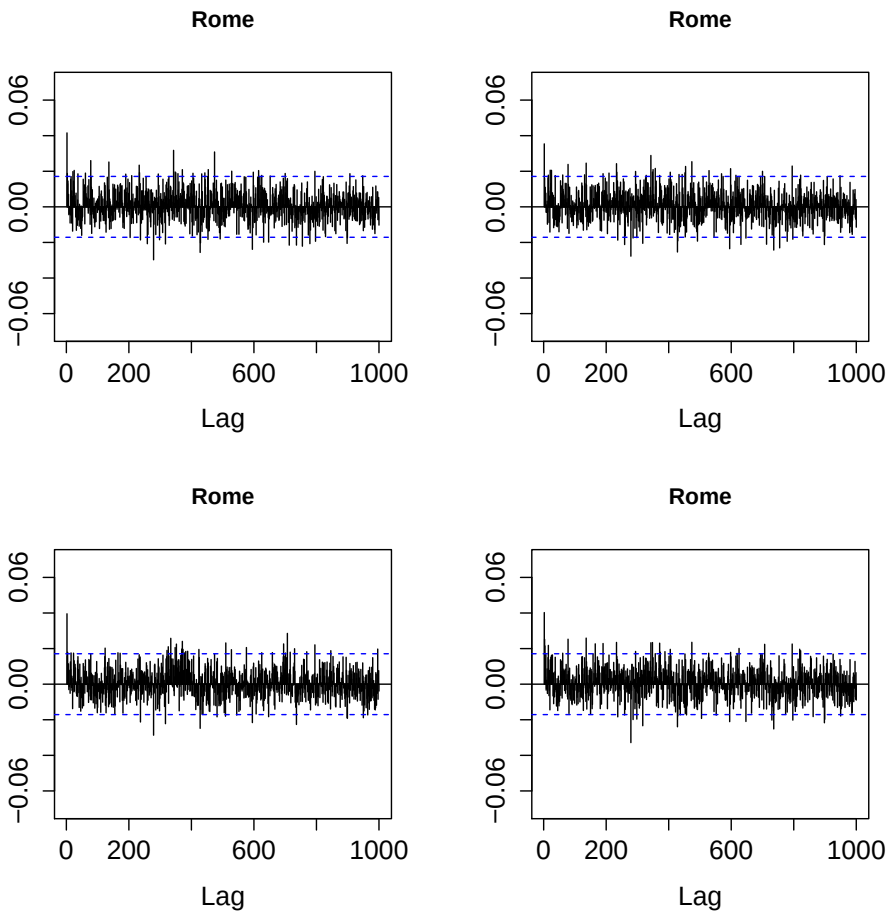

Figure 8: ACF of squared residuals after removing volatility with different models for Rome: $\frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,FTS}}$ (upper left), $\frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,LLR}}$ (upper right), $\frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,MFTSG}}$ (lower left), $\frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t,FTSG}}$ (lower right).

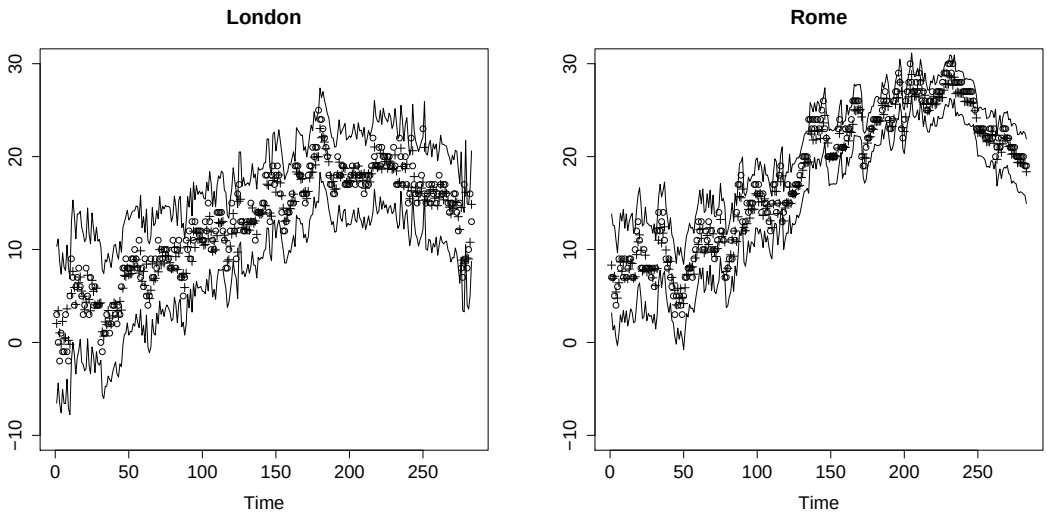

Figure 9: Observed (circles) and predicted (crosses) values with 95% prediction intervals (lines) for London and Rome.

squared residuals (discs) for each 12 year subsample as well as for all 36 years for the cities of London, Rome, Paris and Amsterdam. The lower line corresponds to the 0.05 quantile curve, the middle one to the 0.5 quantile curve and the upper one to the 0.95 quantile curve. The 0.05 and 0.5 quantile curves are not varying significantly in comparison to 0.95 quantile curves for all cities. Therefore, we focus our analysis on understanding the 0.95 quantiles which correspond to the greatest values of daily average squared residuals.

For London, Figure 16, we observe the very interesting phenomenon that although for the first 24 years the variances in summer months are not volatile, for latest 12 years there is an upward tendency which reaches its highest peak at the end of august. For the second 12 years we observe increase of the variances for the winter months. For Rome, Figure 17, we have the same yearly scheme of the quantile curves for each 12 year subperiod. The variance is higher in the beginning of the fall until the end of winter and lower from march until the end of summer. The same conclusion is remarked for the latest 12 year period of Paris, Figure 18, while for the previous years the variances are more volatile. For Amsterdam, Figure 19 depicts lower quantile levels during the spring and summer period and higher for fall and winter.

Expectiles have the advantage of capturing weather extremes, as we see in the shape of the seasonal variance (V shape). Figures 16, 17, 18 and 19 depict as well how the quantiles change over the whole period from 1973 to 2008. It is shown that extreme events are punished in the 95% quantile.

5 Discussion and Conclusion

The present study shows a purely data-driven approach to the important problem of detecting global warming, thus avoiding principles of structural atmospheric models. Since changes in variance might have greater impact than mean effects we attempt to find evidence of shifts in variance of residuals of daily average temperatures over the time.

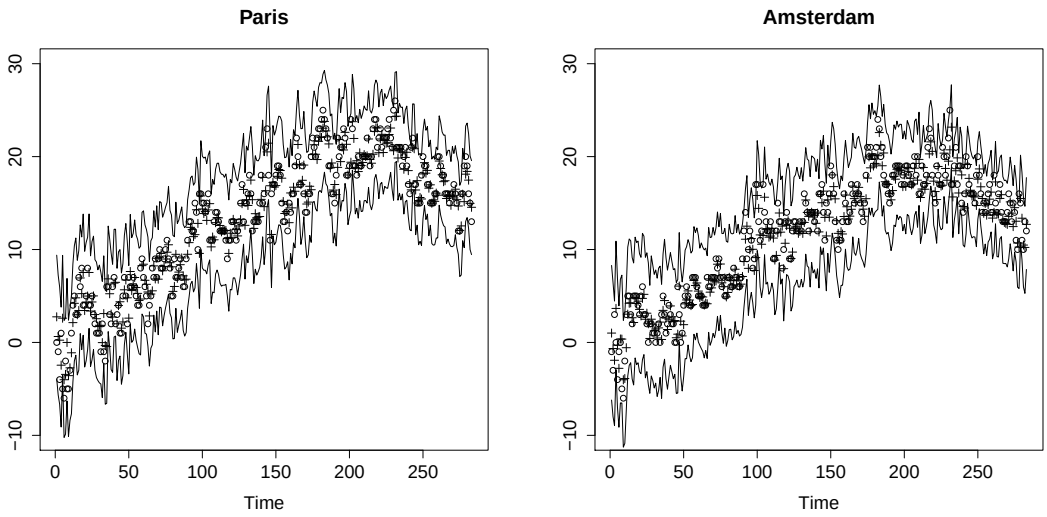

Figure 10: Observed (circles) and predicted (crosses) values with 95% prediction intervals (lines) for Paris and Amsterdam.

We present a time series approach for modelling temperature dynamics. The application is on temperature data of the industrial Blue Banana European area. We study the effect of seasonal variance change by implementing quantile and expectile curves to detrended and deseasonalized temperature residuals. We found, for most of the cities, that 0.9 expectile curves grow over time. This means a tendency for an increasing hot weather. These results are consistent with the findings reported by the IPCC about global warming effect. For the 0.01 expectile curve, the effects are not significant differences within different periods. The findings hold for most of the studied cities, indicating that all sources of trends other than mean and variance would rise trends over spatial scales. shifts in variance vary from location to location. Finally, it is important to remark the importance of statistical tools for assessing changes in weather extrem events over time.

The results provide evidence that global warming exists in the local scale.

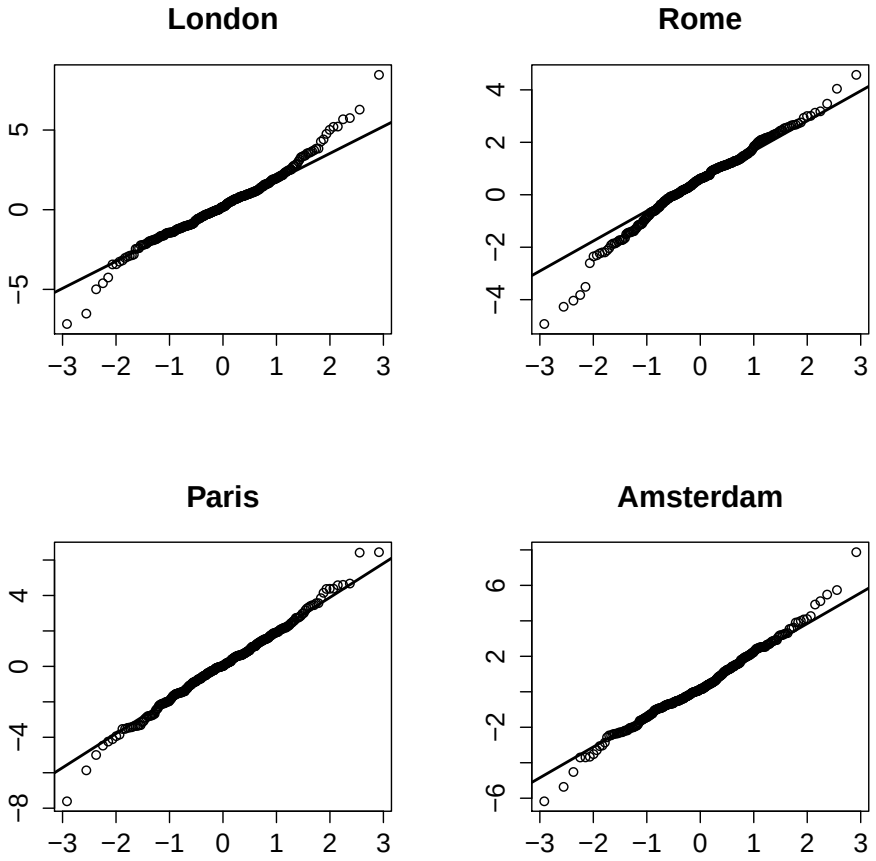

Figure 11: QQ-plots for prediction errors.

References

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein-Tank, A. M. G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Rupa-Kumar, K., Revadekar, J., Griffiths, J., Vincent, L., Stephenson, D., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M., and Vazquez-Aguirre, J. L. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res.*, 111 (D05109):1–22.

Beniston, M. and Goyette, S. (2007). Changes in variability and persistence of climate in switzerland: Exploring 20th century observations and 21st century simulations. *Global and Planetary Change*, 57:1–15.

Benth, F. E., Härdle, W. K., and López-Cabrera, B. (2011). *Pricing Asian temperature risk in Statistical Tools for Finance and Insurance 2nd. edition (Cizek, Härdle and Weron, eds.)*. Springer Verlag Heidelberg.

Benth, F. E. and Saltyte Benth, J. (2010). A critical view on temperature modelling for application in weather derivative markets. *Working paper, University of Oslo*.

Benth, F. E., Saltyte Benth, J., and Koekebakker, S. (2007). Putting a price on temperature. *Scandinavian Journal of Statistics*, 34:746–767.

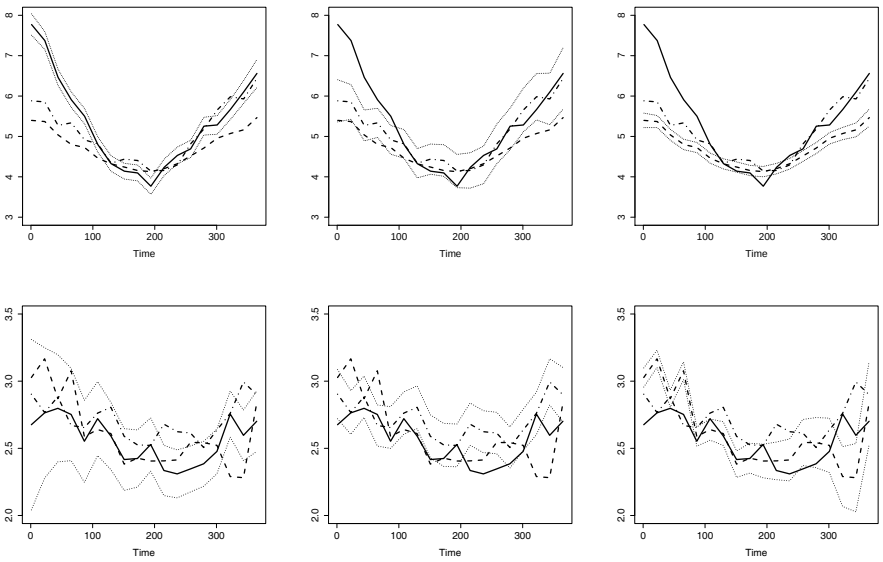

Figure 12: 0.9 (upper panel) & 0.01 (lower panel) expectile curves for London of seasonal temperature variation from 1973 to 2008, for different periods: $1973-1984$ (solid lines), $1985-1996$ (dashed-dotted lines), $1997-2008$ (dashed lines), with the 5% - 95% confidence corridors for the first 12 years expectile (left panel), the second 12 years expectile (middle panel) and the last 12 years expectile (right panel).

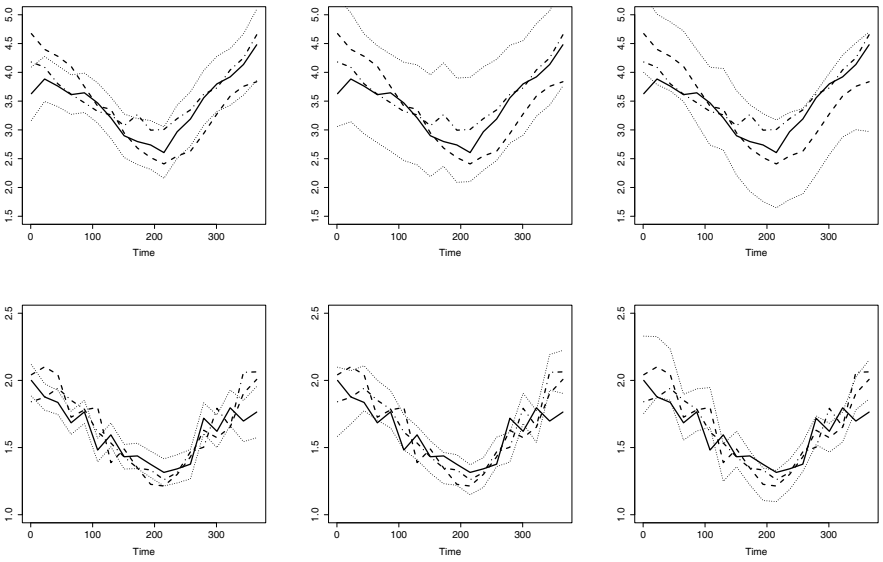

Figure 13: 0.9 (upper panel) & 0.01 (lower panel) expectile curves for Rome of seasonal temperature variation from 1973 to 2008, for different periods: $1973-1984$ (solid lines), $1985-1996$ (dashed-dotted lines), $1997-2008$ (dashed lines), with the 5% - 95% confidence corridors for the first 12 years expectile (left panel), the second 12 years expectile (middle panel) and the last 12 years expectile (right panel).

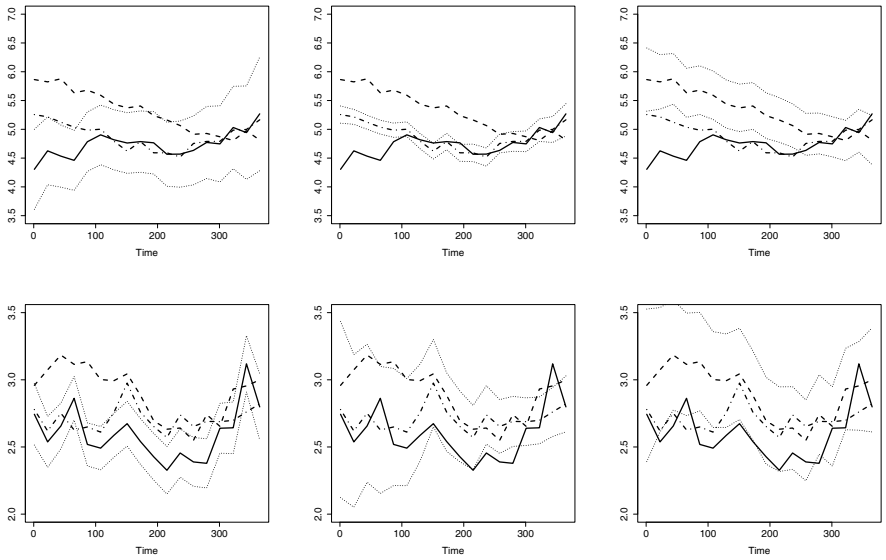

Figure 14: 0.9 (upper panel) & 0.01 (lower panel) expectile curves for Paris of seasonal temperature variation from 1973 to 2008, for different periods: $1973-1984$ (solid lines), $1985-1996$ (dashed-dotted lines), $1997-2008$ (dashed lines), with the 5% - 95% confidence corridors for the first 12 years expectile (left panel), the second 12 years expectile (middle panel) and the last 12 years expectile (right panel).

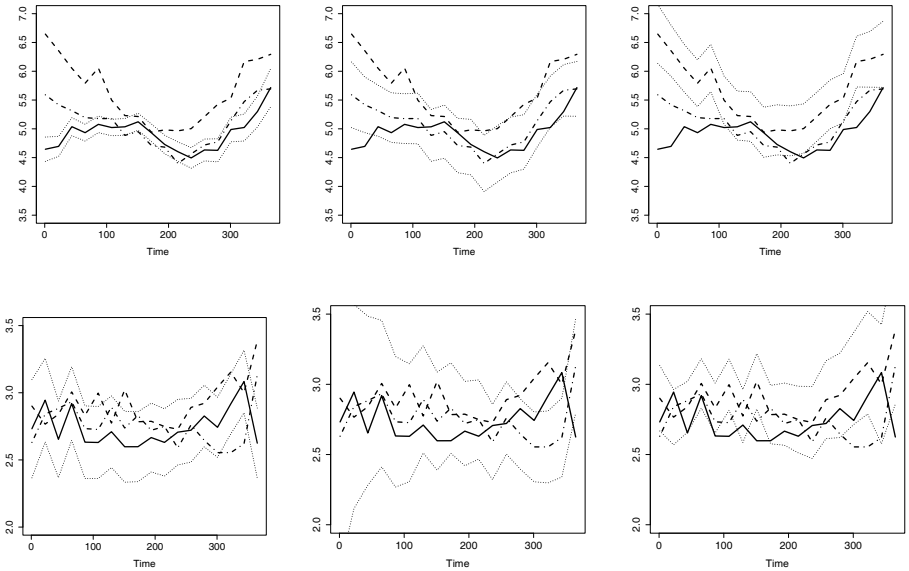

Figure 15: 0.9 (upper panel) & 0.01 (lower panel) expectile curves for Amsterdam of seasonal temperature variation from 1973 to 2008, for different periods: $1973-1984$ (solid lines), $1985-1996$ (dashed-dotted lines), $1997-2008$ (dashed lines), with the 5% - 95% confidence corridors for the first 12 years expectile (left panel), the second 12 years expectile (middle panel) and the last 12 years expectile (right panel).

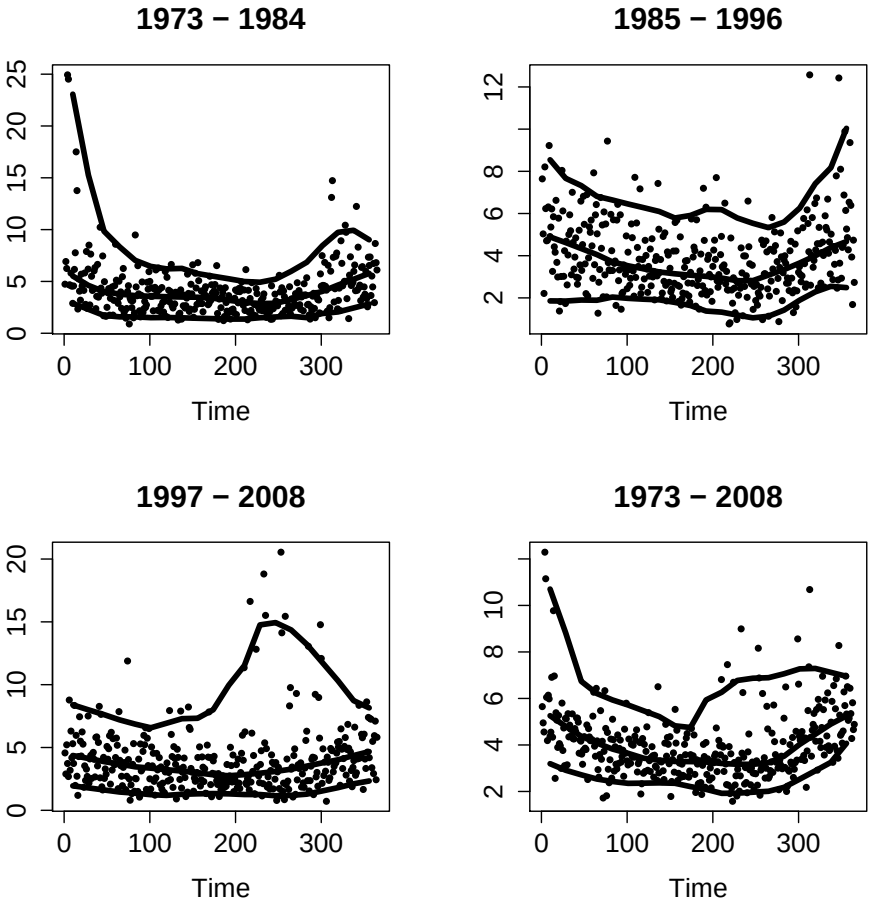

Figure 16: 0.05 (lower line), 0.5 (middle line) and 0.95 (upper line) - quantile curves for London.

Bowman, A. W. and Azzalini, A. (1997). *Applied Smoothing Techniques for Data Analysis*. Clarendon Press, Oxford.

Campbell, S. D. and Diebold, F. X. (2005). Weather forecasting for weather derivatives. *Journal of American Statistical Association*, 100(469):6–16.

Guo, M. and Härdle, W. K. (2012). Simultaneous confidence bands for expectile functions. *Advances in Statistical Analysis*. Available on `http://dx.doi.org/10.1007/s10182-011-0182-1` Version: December 3, 2011.

Härdle, W. K. and López-Cabrera, B. (2011). The implied market price of weather risk. *Applied Mathematical Finance*. Available on `http://dx.doi.org/10.1080/1350486X.2011.591170` Version: January 19, 2012.

Härdle, W. K., López-Cabrera, B., Okhrin, O., and Wang, W. (2011). Localising temperature risk. *Working paper, SFB 649 Humbold Universität zu Berlin*.

Härdle, W. K., Spokoiny, V., and Wang, W. (2011). Local quantile regression. *Working paper, SFB 649 Humboldt Universität zu Berlin*.

Horowitz, J. K. (2001). The income-temperature relationship in a cross-section of countries

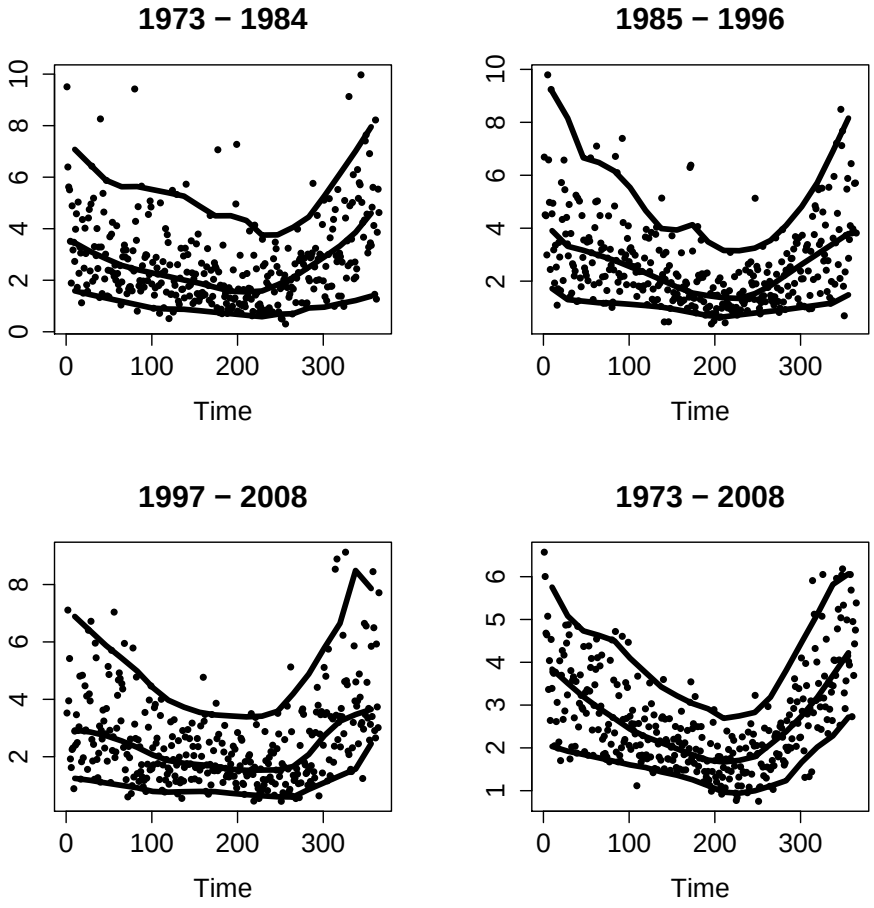

Figure 17: 0.05 (lower line), 0.5 (middle line) and 0.95 (upper line) - quantile curves for Rome.

and its implications for global warming. *University of Maryland Working Paper No. 01-02.*

Hurvich, C. M. and Tsai, C. L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76:299–307.

Hyndman, R. J. and Koehler, A. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22.

IPCC, I.-P.-o.-C.-C. (2007). *The IPCC fourth Assessment Report http://www.ipcc.ch/*, 1:1.

Mendelsohn, R. O., Morrison, W. N., Schlensiger, M. E., and Andronova, N. G. (2000). Country-specific market impacts of climate change. *Climatic Change*, 45:553–569.

Michaels, P. J., Balling-Jr., R. C., Vose, R. S., and P. C. Knappenberger, P. (1998). Analysis of trends in the variability of daily and monthly historical temperature measurements. *Climate Research*, 10:27–33.

Nordhaus, W. D. (2006). Country-specific market impacts of climate change. *Inaugural Article, Proceedings of the National Academy of Sciences*, 103:3510–3517.

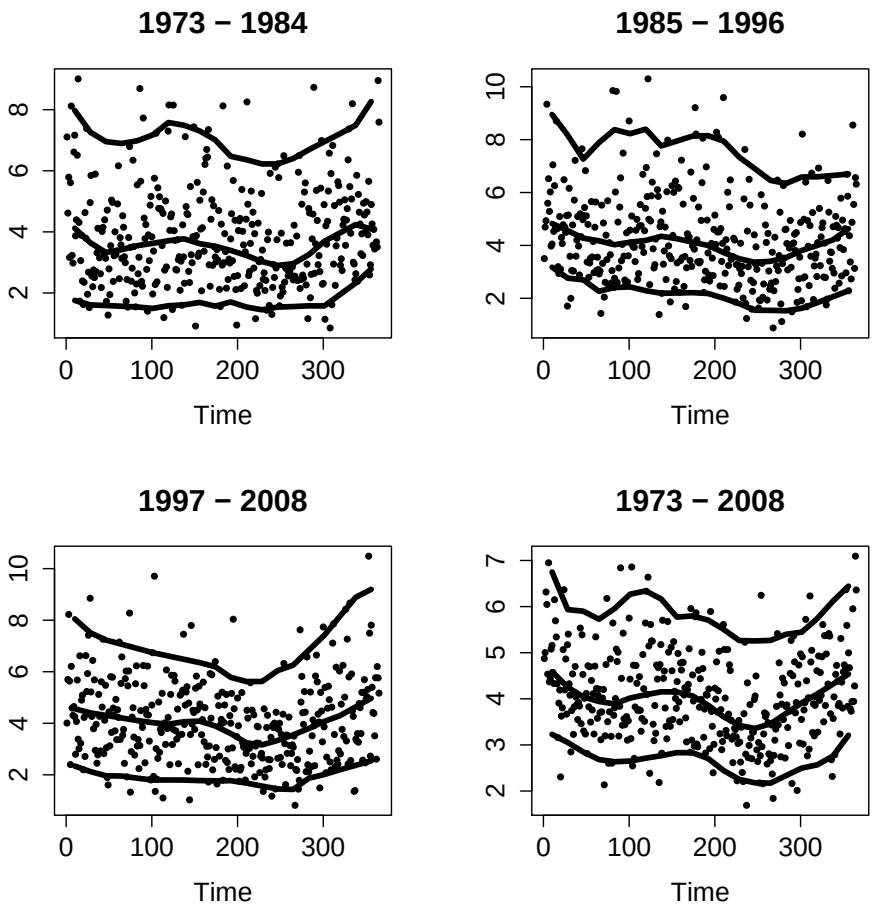

Figure 18: 0.05 (lower line), 0.5 (middle line) and 0.95 (upper line) - quantile curves for Paris.

Nordhaus, W. D. and J. G. Boyer, J. (2000). *Geography and Macroeconomics: New Data and New Findings*. MIT Press, Cambridge, Massachusetts.

Parton, W. and Logan, J. (1981). A model for diurnal variation in soil and air temperature. *Agricultural Meteorology*, 23:205–216.

Racskoa, P., Szeidla, L., and Semenovb, M. (1991). A serial approach to local stochastic weather models. *Ecological Modelling*, 57(1):27–41.

Smith, C. A. and Sardeshmukh, P. D. (2000). The effect of enso on the intraseasonal variance of surface temperatures in winter. *International Journal of Climatology*, 20(1):1543–1557.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. Royal. Statist. Soc B.*, 58(1):267–288.

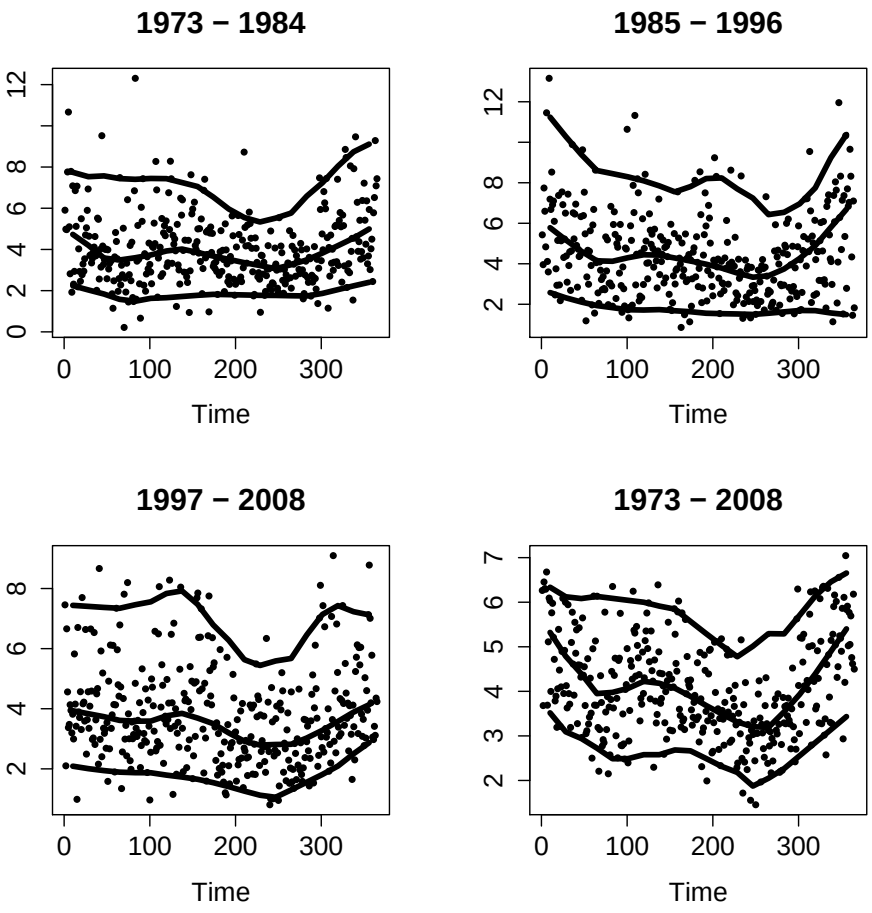

Figure 19: 0.05 (lower line), 0.5 (middle line) and 0.95 (upper line) - quantile curves for Amsterdam.

SFB 649 Discussion Paper Series 2012

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "HMM in dynamic HAC models" by Wolfgang Karl Härdle, Ostap Okhrin and Weining Wang, January 2012.
002 "Dynamic Activity Analysis Model Based Win-Win Development Forecasting Under the Environmental Regulation in China" by Shiyi Chen and Wolfgang Karl Härdle, January 2012.
003 "A Donsker Theorem for Lévy Measures" by Richard Nickl and Markus Reiß, January 2012.
004 "Computational Statistics (Journal)" by Wolfgang Karl Härdle, Yuichi Mori and Jürgen Symanzik, January 2012.
005 "Implementing quotas in university admissions: An experimental analysis" by Sebastian Braun, Nadja Dwenger, Dorothea Kübler and Alexander Westkamp, January 2012.
006 "Quantile Regression in Risk Calibration" by Shih-Kang Chao, Wolfgang Karl Härdle and Weining Wang, January 2012.
007 "Total Work and Gender: Facts and Possible Explanations" by Michael Burda, Daniel S. Hamermesh and Philippe Weil, February 2012.
008 "Does Basel II Pillar 3 Risk Exposure Data help to Identify Risky Banks?" by Ralf Sabiwalsky, February 2012.
009 "Comparability Effects of Mandatory IFRS Adoption" by Stefano Cascino and Joachim Gassen, February 2012.
010 "Fair Value Reclassifications of Financial Assets during the Financial Crisis" by Jannis Bischof, Ulf Brüggemann and Holger Daske, February 2012.
011 "Intended and unintended consequences of mandatory IFRS adoption: A review of extant evidence and suggestions for future research" by Ulf Brüggemann, Jörg-Markus Hitz and Thorsten Sellhorn, February 2012.
012 "Confidence sets in nonparametric calibration of exponential Lévy models" by Jakob Söhl, February 2012.
013 "The Polarization of Employment in German Local Labor Markets" by Charlotte Senftleben and Hanna Wielandt, February 2012.
014 "On the Dark Side of the Market: Identifying and Analyzing Hidden Order Placements" by Nikolaus Hautsch and Ruihong Huang, February 2012.
015 "Existence and Uniqueness of Perturbation Solutions to DSGE Models" by Hong Lan and Alexander Meyer-Gohde, February 2012.
016 "Nonparametric adaptive estimation of linear functionals for low frequency observed Lévy processes" by Johanna Kappus, February 2012.
017 "Option calibration of exponential Lévy models: Implementation and empirical results" by Jakob Söhl und Mathias Trabs, February 2012.
018 "Managerial Overconfidence and Corporate Risk Management" by Tim R. Adam, Chitru S. Fernando and Evgenia Golubeva, February 2012.
019 "Why Do Firms Engage in Selective Hedging?" by Tim R. Adam, Chitru S. Fernando and Jesus M. Salas, February 2012.
020 "A Slab in the Face: Building Quality and Neighborhood Effects" by Rainer Schulz and Martin Wersing, February 2012.
021 "A Strategy Perspective on the Performance Relevance of the CFO" by Andreas Venus and Andreas Engelen, February 2012.
022 "Assessing the Anchoring of Inflation Expectations" by Till Strohsal and Lars Winkelmann, February 2012.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

SFB 649 Discussion Paper Series 2012

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

023 "Hidden Liquidity: Determinants and Impact" by Gökhan Cebiroglu and Ulrich Horst, March 2012.
024 "Bye Bye, G.I. - The Impact of the U.S. Military Drawdown on Local German Labor Markets" by Jan Peter aus dem Moore and Alexandra Spitz-Oener, March 2012.
025 "Is socially responsible investing just screening? Evidence from mutual funds" by Markus Hirschberger, Ralph E. Steuer, Sebastian Utz and Maximilian Wimmer, March 2012.
026 "Explaining regional unemployment differences in Germany: a spatial panel data analysis" by Franziska Lottmann, March 2012.
027 "Forecast based Pricing of Weather Derivatives" by Wolfgang Karl Härdle, Brenda López-Cabrera and Matthias Ritter, March 2012.
028 “Does umbrella branding really work? Investigating cross-category brand loyalty” by Nadja Silberhorn and Lutz Hildebrandt, April 2012.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".