Localising temperature risk

Wolfgang Karl Härdle*, Brenda López Cabrera†, Ostap Okhrin ‡, Weining Wang§

February 26, 2013

Abstract

On the temperature derivative market, modelling temperature volatility is an important issue for pricing and hedging. In order to apply the pricing tools of financial mathematics, one needs to isolate a Gaussian risk factor. A conventional model for temperature dynamics is a stochastic model with seasonality and intertemporal autocorrelation. Empirical work based on seasonality and autocorrelation correction reveals that the obtained residuals are heteroscedastic with a periodic pattern. The object of this research is to estimate this heteroscedastic function so that, after scale normalisation, a pure standardised Gaussian variable appears. Earlier works investigated temperature risk in different locations and showed that neither parametric component functions nor a local linear smoother with constant smoothing parameter are flexible enough to generally describe the variance process well. Therefore, we consider a local adaptive modelling approach to find, at each time point, an optimal smoothing parameter to locally estimate the seasonality and volatility. Our approach provides a more flexible and accurate fitting procedure for localised temperature risk by achieving nearly normal risk factors. We also employ our model to forecast the temperature in different cities and compare it to a model developed in Campbell and Diebold (2005).

Keywords: Weather derivatives, localising temperature residuals, seasonality, local model selection
JEL classification: G19, G29, G22, N23, N53, Q59

1 Introduction

The pricing of contingent claims based on stochastic dynamics, for example, stocks or FX rates, is well known in financial engineering. An elegant approach to such a pricing task is based on

*Professor at Humboldt-Universität zu Berlin, Ladislaus von Bortkiewicz chair of statistics and Director of C.A.S.E. - Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany. Email:haerdle@wiwi.hu-berlin.de

†Assistant professor at the Ladislaus von Bortkiewicz chair of statistics of Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany. Email:lopezcab@wiwi.hu-berlin.de

‡Assistant professor at the Ladislaus von Bortkiewicz chair of statistics of Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany. Email:ostap.okhrin@wiwi.hu-berlin.de

§Research associate at the Ladislaus von Bortkiewicz chair of statistics, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany. Email:wangwein@cms.hu-berlin.de

self-financing replication arguments. An essential element of this approach is the tradeability of the underlying. This however does not apply to weather derivatives, contingent on temperature or rain, since the underlying is not tradeable. In this context, the proposed pricing techniques are based on either equilibrium ideas (Horst and Mueller (2007)) or econometric modelling of the underlying dynamics Campbell and Diebold (2005) and Benth, Benth and Koekebakker (2007) followed by risk neutral pricing.

The equilibrium approach relies on assumptions about preferences (with explicitly known functional forms) though. In this study we prefer a phenomenological approach since the underlying (temperature) we consider is of varying local nature and our analysis aims at understanding the pricing at different locations and different time points around the world. A time series approach has been taken by Benth et al. (2007), who corrects for seasonality (in mean), then for intertemporal correlation and finally as in Campbell and Diebold (2005), for seasonal variations. After these manipulations, a Gaussian risk factor needs to be isolated in order to apply continuous time pricing techniques, Karatzas and Shreve (2001).

Empirical studies following this econometrical route show evidence that the resulting temperature risk factor deviates severely from Gaussianity, which in turn challenges the pricing tools, Benth, Härdle and López Cabrera (2011). In particular, for Asian cities, like for example Kaohsiung (Taiwan), one observes very distinctive non-normality in the form of clearly visible heavy tails caused by extended volatility in peak seasons. This is visible from Figure 1 where a log density plot reveals a nonnormal shoulder structure (kurtosis= 3.22, skewness= -0.08, JB= 128.74).

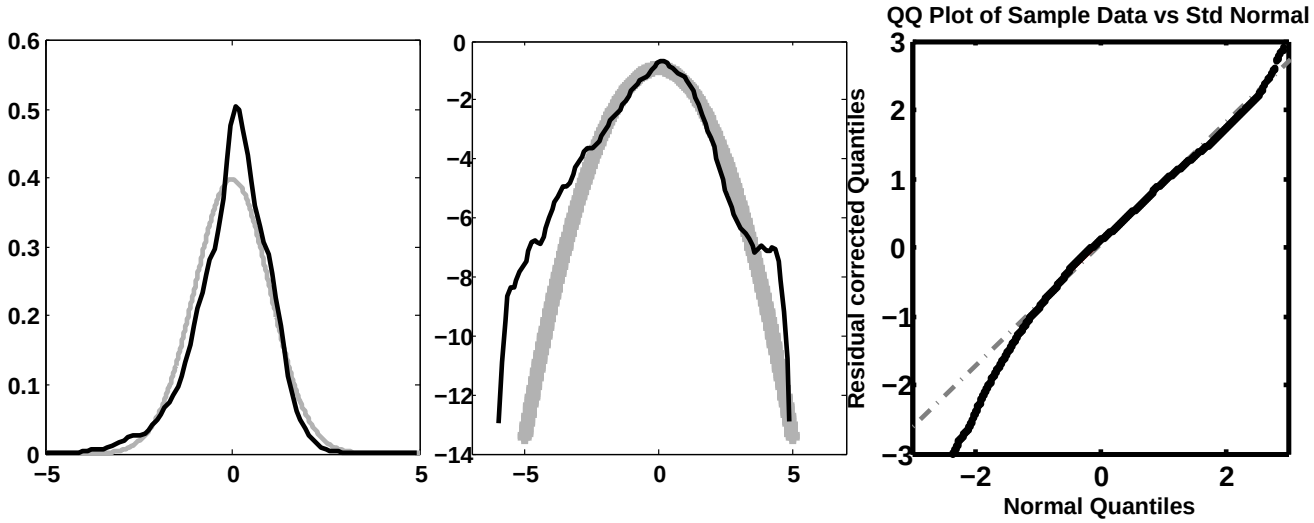


Figure 1: Kernel density estimates (left panel), log kernel density estimates (middle panel) and QQ-plots (right panel) of normal densities (grey lines) and Kaohsiung standardised residuals (black line)

The econometric analysis we apply follows Benth et al. (2007) where temperature T_t is decomposed into a seasonality term Λ_t and a stochastic part with seasonal variance σ_t^2. The fitted seasonality trend Λ_t and seasonal variance σ_t^2 are approximated with truncated Fourier series (and an additional

GARCH term):

$$\Lambda_t = a + bt + \sum_{l=1}^{L_1} c_l \cos\left\{\frac{2\pi(t-d_l)}{l\cdot 365}\right\}, \tag{1}$$

$$\sigma^2_{t,FTSG} = c_{10} + \sum_{l=1}^{L_2}\left\{c_{2l}\cos\left(\frac{2l\pi t}{365}\right) + c_{2l+1}\sin\left(\frac{2l\pi t}{365}\right)\right\} + \alpha_1(\sigma_{t-1}\eta_{t-1})^2 + \beta_1\sigma^2_{t-1}, \tag{2}$$

$$\eta_t \sim iid(0,1).$$

The upper panel of Figure 2 displays the seasonality and deseasonalised residuals over two years in Kaohsiung. The lower panel RHS displays the empirical and seasonal variance function, while the lower panel LHS shows the smoothed seasonal variance function over years. The series expansion (1) and (2) failed though in the volatility peak seasons. Even incorporating an asymmetry term for the dip of temperature in winter does not improve the closeness to normality.

One may of course pursue a fine tuning of (1) and (2) with more and more periodic terms but this will increase the number of parameters. We therefore propose a local parametric approach. The seasonality Λ_s and σ_s are approximated with a Local Linear Regression (LLR) estimator:

$$\arg\min_{e,f} \sum_{t=1}^{365} \left\{\bar{T}_t - e_s - f_s(t-s)\right\}^2 K\left(\frac{t-s}{h}\right), \tag{3}$$

$$\arg\min_{g,v} \sum_{t=1}^{365} \left\{\hat{\varepsilon}^2_t - g_s - v_s(t-s)\right\}^2 K\left(\frac{t-s}{h}\right), \tag{4}$$

where $\bar{T}_t$ is the mean (over years) of daily averages temperatures, $\hat{\varepsilon}^2_t$ the squared residual process (after seasonal and intertemporal fitting), h the bandwidth and $K(\cdot)$ is a kernel. Note, that due to the spherical character of the data, the kernel weights in (3) and (4) may be calculated from "wrapped around observations" thereby avoiding boundary bias. The estimates $\hat{\Lambda}_s$, $\hat{\sigma}^2_s$ are given by the minimisers $\hat{e}_s$, $\hat{g}_s$ of (3) and (4). The upper panel of Figure 2 shows the seasonality in mean and the bottom panel on the RHS the variance estimated with truncated Fourier series and local linear regression using the quartic kernel. We observe high variance in winter and early summer and low variance in spring and late summer.

The scale correction of the obtained residuals (after seasonal and intertemporal fitting) is apparently not identical over the year. A very structured volatility pattern up to April is followed by a moderately constant period until an increasing peak starting in September. This motivates our research to localise temperature risk. The local smoothness of σ^2_t is of course not only a matter of one location (here Kaohsiung) but varies also over the different cities around the world that we are analysing in this study. Our study is local in a double sense: local in time and space. We use adaptive methods to localise the underlying dynamics and with that being able to achieve Gaussian risk factors. This will justify the pricing via standard tools that are based on Gaussian risk drivers. The localisation in time is based on adjusting the smoothing parameter h. For a general framework on local parametric approximation we refer to Spokoiny (2009). As a result we obtain better approximations to normality and therefore less biased prices.

This paper is structured as follows. Section 2 describes the localising approach. In section 3, we present the data and conduct the analysis to different cities. Section 4 presents a forecasting exercise and the following section is devoted to an application where the pricing of weather derivative

contract types is presented. Section 6 concludes the paper. All quotations of currency in this paper will be in USD unless otherwise stated and therefore we will omit the explicit notion of the currency. All the computations were carried out in Matlab version 7.6 and R. The temperature data for different cities in US, Europe and Asia were obtained from the National Climatic Data Center (NCDC), the Deutscher Wetterdienst (DWD), Bloomberg Professional Service and the Japanese Meteorological Agency (JMA). All data is converted to Celsius degrees. Weather derivative data from CME was extracted from Bloomberg. To simplify notation, dates are denoted with yyyymmdd format.

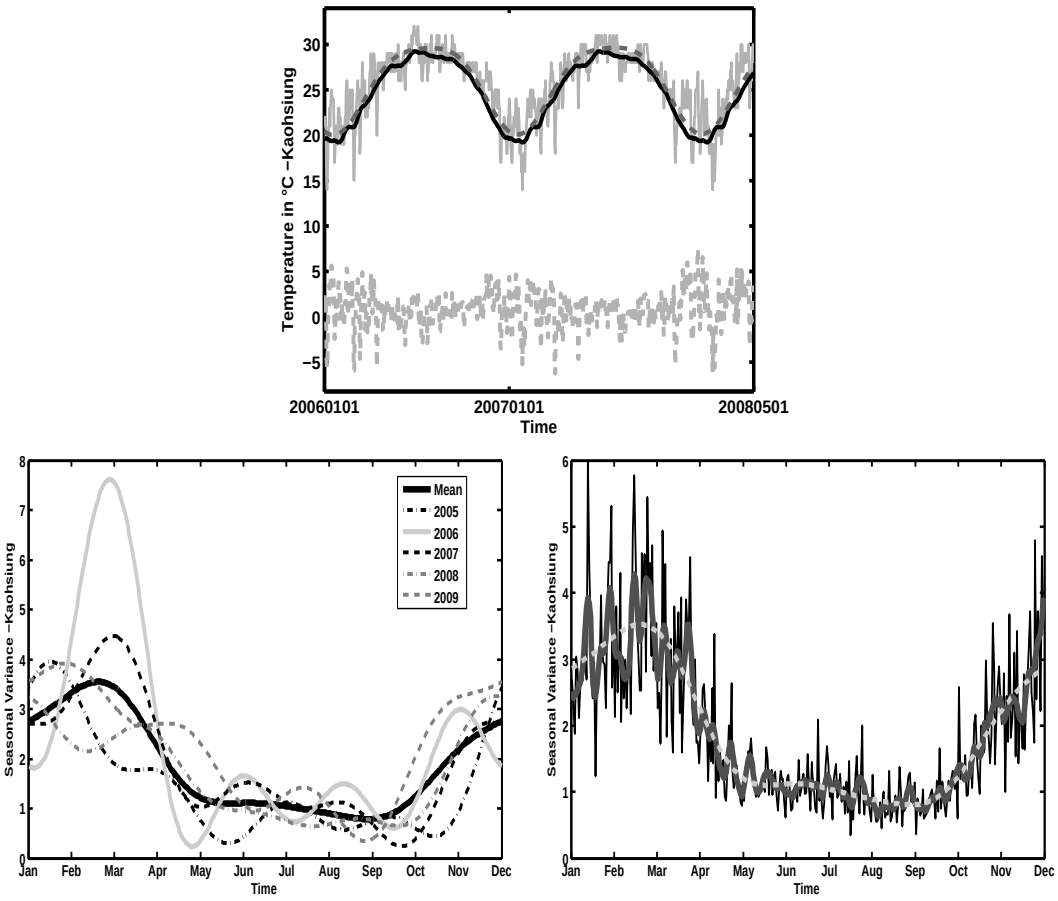


Figure 2: Upper panel: Kaohsiung daily average temperature (black line), Fourier truncated (dotted grey line) and local linear seasonality function (grey line), Residuals in lower part. Lower left panel: truncated Fourier seasonal variation ($\hat{\Lambda}_t$) over time. Lower right panel: Kaohsiung empirical (black line), truncated Fourier (dotted grey line) and local linear (grey line) seasonal variance ($\hat{\sigma}_t^2$) function.

2 Model

Let us first refine our notation from t to (t, j), with $t = 1, \ldots, \tau = 365$ days, $j = 0, \ldots, J$ years. The time series decomposition we consider is given as:

$$\begin{aligned} X_{365j+t} &= T_{t,j} - \Lambda_t, \\ X_{365j+t} &= \sum_{l=1}^{L} \beta_{lj} X_{365j+t-l} + \varepsilon_{t,j}, \\ \varepsilon_{t,j} &= \sigma_t e_{t,j}, \\ e_{t,j} &\sim \mathsf{N}(0,1), \\ \hat{\varepsilon}_{t,j} &= X_{365j+t} - \sum_{l=1}^{L} \hat{\beta}_{lj} X_{365j+t-l}, \end{aligned} \tag{5}$$

where $T_{t,j}$ is the temperature at day t in year j, Λ_t denotes the seasonality effect and σ_t the seasonal variance. Motivation of this approach can be found in Campbell and Diebold (2005) (CD), who proposes the model

$$\begin{aligned} T_t &= \text{Trend}_t + \text{Seasonality}_t + \sum_{l=1}^{L} \rho_{t-l} T_{t-l} + \sigma_t \varepsilon_t, \\ \text{Trend}_t &= \sum_{m=0}^{M} \beta_m t^m, \\ \text{Seasonality}_t &= \sum_{p=1}^{P} \left[\delta_{c,p} \cos\left\{2\pi p \frac{d(t)}{365}\right\} + \delta_{s,p} \sin\left\{2\pi p \frac{d(t)}{365}\right\} \right], \\ \sigma_t^2 &= \sum_{q=1}^{Q} \left[\gamma_{c,q} \cos\left\{2\pi q \frac{d(t)}{365}\right\} + \gamma_{s,q} \sin\left\{2\pi q \frac{d(t)}{365}\right\} \right] + \sum_{r=1}^{R} \{\alpha_r (\sigma_{t-r} \varepsilon_{t-r})^2 + \sum_{s=1}^{S} \beta_s \sigma_{t-s}^2\}. \end{aligned}$$

We will use the CD model as a benchmark model for further analysis. Later studies, e.g., Benth et al. (2007) and Härdle and López Cabrera (2011), have provided evidence that the parameters β_{lj} are likely to be j independent and hence estimated consistently from a global autoregressive process model $AR(L_j)$ with $L_j = L$. Since the stylised facts of temperature re-occur every year, our focus is on the flexible estimation of Λ_t and σ_t^2, see Figure 2.

The seasonal trend function Λ_t and the seasonal variance function σ_t^2 affect, of course, the Gaussianity of the resulting normalised residuals. The commonly used approaches 1. truncated Fourier series, and 2. local polynomial regression (with fixed bandwidth) are rather restrictive and do not fit the data well since they do not necessarily yield normal risk factors. These observations motivated us to consider a more flexible approach. The main idea is to fit a local parametric model for the trend and variance with adaptively chosen window sizes. Specifically, we use kernel smoothing and employ an adaptive technique to choose the bandwidth over days. Other examples of this technique can be found in Cízek, Härdle and Spokoiny (2009) and Chen, Härdle and Pigorsch (2010).

2.1 How does the adaptation work?

The time series $T_{t,j}$ are approximated at a fixed time point $s \in [1, 365]$. Our goal is to find a local window that possesses certain optimality properties, to be defined below. Specifically, for a specified weight sequence, we conduct a sequential LRT to choose an appropriate bandwidth. Different procedures of estimating seasonality and volatility are studied. Suppose that the object to be approximated is the seasonal variance $\theta(t) = \{\sigma_t^2\}$. A weighted maximum likelihood approach is given by:

$$\begin{aligned}\tilde{\theta}_k(s) &\stackrel{\text{def}}{=} \arg\max_{\theta} L\{W^k(s), \theta\} \\ &= \arg\min_{\theta} \sum_{t=1}^{365} \sum_{j=0}^{J} \{\log(2\pi\theta)/2 + \hat{\varepsilon}_{t,j}^2/2\theta\} w(s, t, h_k), \end{aligned} \tag{6}$$

with the "localising scheme" $W^k(s) = \{w(s, 1, h_k), w(s, 2, h_k), \ldots, w(s, 365, h_k)\}^\top$, where $w(s, t, h_k) = h_k^{-1} K\{(s-t)/h_k\}$, $k = 1, \ldots, K$, $h_1 < h_2 < h_3 < \ldots < h_K$ a prescribed sequence of bandwidths, and $K(u) = 15/16(1 - u^2)^2 \mathbf{I}(|u| \leq 1)$ (quartic kernel).

The explicit solution of (6) is in fact a Nadaraya-Watson estimator:

$$\begin{aligned}\tilde{\theta}_k(s) &= \sum_{t,j} \hat{\varepsilon}_{t,j}^2 w(s, t, h_k) / \sum_{t,j} w(s, t, h_k) \\ &= \sum_{t} \hat{\varepsilon}_{t}^2 w(s, t, h_k) / \sum_{t} w(s, t, h_k), \end{aligned}$$

with

$$\hat{\varepsilon}_t^2 \stackrel{\text{def}}{=} (J+1)^{-1} \sum_{j=0}^{J} \hat{\varepsilon}_{t,j}^2.$$

From a smoothing perspective we are in a comfortable situation here since the boundary bias is not an issue, as we are dealing with a periodic function $\theta(t) = \theta(t + 365)$. We use mirrored observations: assume $h_K < 365/2$, then the observation set, for example for the seasonal variance, is extended to $\hat{\varepsilon}_{-364}^2, \hat{\varepsilon}_{-363}^2, \ldots, \hat{\varepsilon}_0^2, \hat{\varepsilon}_1^2, \ldots, \hat{\varepsilon}_{730}^2$, where

$$\begin{aligned}\hat{\varepsilon}_t^2 &\stackrel{\text{def}}{=} \hat{\varepsilon}_{365+t}^2, -364 \leq t \leq 0, \\ \hat{\varepsilon}_t^2 &\stackrel{\text{def}}{=} \hat{\varepsilon}_{t-365}^2, 366 \leq t \leq 730. \end{aligned}$$

Since the location s is fixed, we drop s for simplicity of notation.

For $\ell < k$, the accuracy of the estimation is measured by the fitted likelihood ratio (LR):

$$L(W^\ell, \tilde{\theta}_\ell, \tilde{\theta}_k) \stackrel{\text{def}}{=} L(W^\ell, \tilde{\theta}_\ell) - L(W^\ell, \tilde{\theta}_k). \tag{7}$$

For the Gaussian risk factor situation the variance σ_t^2 (or trend Λ_t) estimation is carried out within an exponential family framework, so the LR can be written in closed form:

$$\begin{aligned}L(W^k, \tilde{\theta}_k, \theta^*) &\stackrel{\text{def}}{=} N_k \mathcal{K}(\tilde{\theta}_k, \theta^*) \\ &= -\{\log(\tilde{\theta}_k/\theta^*) + 1 - \theta^*/\tilde{\theta}_k\}/2, \end{aligned} \tag{8}$$

where $N_k = J\sum_{t=1}^{365} w(s,t,h_k)$ and $\mathcal{K}(\tilde{\theta}_k, \theta^*)$ is the Kullback-Leibler divergence (9) between two normal distributions with variances $\tilde{\theta}_k$ and θ^*. Note that (8) is the divergence for exactly this case. For trend Λ_t estimation, it has to be replaced by $(\tilde{\theta}_k - \theta^*)^2/(2\sigma^2)$.

Recall that the Kullback-Leibler divergence of two distributions with densities $p(x)$ and $q(x)$ is

$$\mathcal{K}\{p(x), q(x)\} \stackrel{\text{def}}{=} \mathsf{E}_{p(.)} \log \frac{p(x)}{q(x)} \quad . \tag{9}$$

To guarantee the feasibility of the tests, we need moment bounds and confidence sets for the LR that will guarantee that the MLE is concentrated in the level set of the likelihood ratio process (indexed by the number of observations) around the true parameter, see Polzehl and Spokoiny (2006) and Mercurio and Spokoiny (2004). Below we state a result along this line for the variance (a similar bound can be derived for the mean).

Theorem 2.1 *[Spokoiny (2009)] Assuming that $\theta(t) = \theta^*$ for any $t \in [1, 365]$, then for $\mathfrak{z} > 0$ and $k \in 1, \ldots, K, r > 0$, denote by $\mathrm{P}_{\theta^*}(.)$ the measure corresponding to (6). We obtain*

$$\mathrm{P}_{\theta^*}\left\{L(W^k, \tilde{\theta}_k, \theta^*) > \mathfrak{z}\right\} \leq 2\exp\left(-\mathfrak{z}\right) \tag{10}$$

and a risk bound for a power loss function:

$$\mathsf{E}_{\theta^*}|L(W^k, \tilde{\theta}_k, \theta^*)|^r \leq \mathfrak{r}_r, \tag{11}$$

where $\mathfrak{r}_r = 2r\int_{\mathfrak{z}\geq 0} \mathfrak{z}^{r-1}\exp(-\mathfrak{z})d\mathfrak{z}$. This polynomial bound applies to all localising schemes W^k simultaneously.

The risk bound (11) allows us to define likelihood based confidence sets since together with (10) it tells us that the likelihood process is stochastically bounded. Define therefore confidence sets with critical values (CVs) $\mathfrak{z}_k$ to level α:

$$\mathfrak{E}_{\alpha,k} = \{\theta : L(W^k, \tilde{\theta}_k, \theta) \leq \mathfrak{z}_k\}. \tag{12}$$

Equipped with confidence sets (12), we launch the Local Model Selection (LMS) algorithm:

- Fix a point $s \in \{1, 2, \ldots, 365\}$.
- Start with the smallest interval h_1: $\hat{\theta}_1 = \tilde{\theta}_1$
- For $k \geq 2$, $\tilde{\theta}_k$ is accepted and $\hat{\theta}_k = \tilde{\theta}_k$ if $\tilde{\theta}_{k-1}$ was accepted and $\tilde{\theta}_k \in \mathfrak{E}_{\alpha,l}, \forall \ell = 1, \ldots, k-1$, i.e.

  $$L(W^k, \tilde{\theta}_\ell, \tilde{\theta}_k) \leq \mathfrak{z}_\ell, \forall \ell = 1, \ldots, k-1.$$

  Otherwise, set $\hat{\theta}_k = \hat{\theta}_{k-1}$, where $\hat{\theta}_k$ is the latest accepted after first k steps.
- Define $\hat{k}$ as the kth step we stopped, and $\hat{\theta}_\ell = \tilde{\theta}_{\hat{k}}, \ell \geq k$.

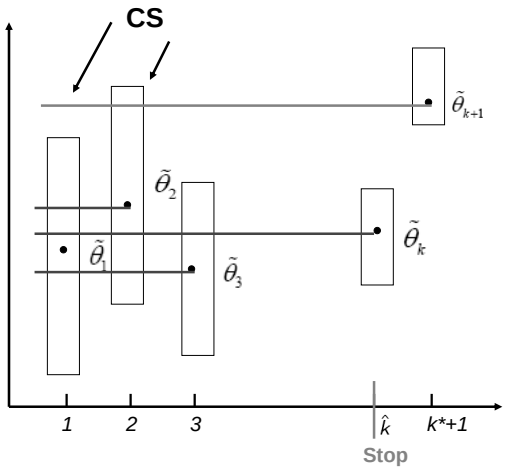


Figure 3: Illustration of the LMS

The LMS algorithm is illustrated in Figure 3. For every estimate $\tilde{\theta}_k$ the corresponding confidence set is shown. If the horizontal line originating in $\tilde{\theta}_k$ does not cross all the preceding intervals then the selection algorithm terminates.

A further integrated approach is to consider an iterative algorithm, which iterates between estimating the seasonal component and the variance $\theta(t) = \{\Lambda_t, \sigma_t^2\}$. This algorithm can further cope with heteroscedasticity in the corrected residuals after seasonality in mean and variance component. The procedure is:

Step 1. Estimate $\hat{\beta}$ in an initial Λ_t^0 using a truncated Fourier series or any other deterministic function;

Step 2. For fixed $\hat{\Lambda}_{s,\nu} = \{\hat{\Lambda}'_{s,\nu}, \hat{\Lambda}''_{s,\nu}\}^\top, s = \{1, \ldots, 365\}$ from last step ν, and fixed $\hat{\beta}$, get $\hat{\sigma}^2_{s,\nu+1}$ by

$$\begin{aligned}\hat{\sigma}^2_{s,\nu+1} &= \arg\min_{\sigma^2} \sum_{t=1}^{365} \sum_{j=0}^{J} [\{T_{365j+t} - \hat{\Lambda}'_{s,\nu} - \hat{\Lambda}''_{s,\nu}(t-s) \\ &- \sum_{l=1}^{L} \hat{\beta}_l X_{365j+t-l}\}^2/2\sigma^2 + \log(2\pi\sigma^2)/2] w(s,t,h'_k);\end{aligned}$$

Step 3. For fixed $\hat{\sigma}^2_{s,\nu+1}$ and $\hat{\beta}$, we estimate $\hat{\Lambda}_{s,\nu+1}, s = \{1, \ldots, 365\}$ via another local adaptive procedure:

$$\hat{\Lambda}_{s,\nu+1} = \arg\min_{\{\Lambda', \Lambda''\}^\top} \sum_{t=1}^{365} \sum_{j=0}^{J} \left\{ T_{365j+t} - \Lambda' - \Lambda''(t-s) - \sum_{l=1}^{L} \hat{\beta}_l X_{365j+t-l} \right\}^2 w(s,t,h'_k)/2\hat{\sigma}^2_{s,\nu+1},$$

where $\{h'_1, h'_2, h'_3, \ldots, h'_{K'}\}$ is a sequence of bandwidths;

Step 4. Repeat steps 2 and 3 until both $|\hat{\Lambda}_{t,\nu+1} - \hat{\Lambda}_{t,\nu}| < \pi_1$ and $|\hat{\sigma}^2_{t,\nu+1} - \hat{\sigma}^2_{t,\nu}| < \pi_2$ for some constants π_1 and π_2.

Our empirical implementation suggests that one iteration is enough. The LMS methods require CVs $\mathfrak{z}_k$, which define the significance for the LR statistics $L(W^\ell, \tilde{\theta}_\ell, \tilde{\theta}_k)$ or alternatively the length of the confidence interval (see (10)) at each step. The CVs are calibrated from the "propagation condition" below which ensures a desired level of type one error. To be more specific, for every step k, define $\hat{\theta}_k$ as the "survived estimator" after the kth step (if the estimator is not rejected up to step k, then $\hat{\theta}_k = \tilde{\theta}_k$, else if the estimator has been rejected at step $l < k$, then $\hat{\theta}_k = \tilde{\theta}_l$). Measure the closeness of $\tilde{\theta}_k$ and $\hat{\theta}_k$ by

$$\mathsf{E}_{\theta^*} |L(W^k, \tilde{\theta}_k, \hat{\theta}_k)|^r \leq \alpha \mathfrak{r}_r \tag{13}$$

for $k = 1, \ldots, K$ with $\mathfrak{r}_r$ the parametric risk bound in (11) and α a control parameter corresponding to the type one error. In fact

$$\mathsf{E}_{\theta^*} |L(W^k, \tilde{\theta}_k, \hat{\theta}_k)|^r \to \mathrm{P}_{\theta^*}(\tilde{\theta}_k \neq \hat{\theta}_k)$$

for $r \to 0$, therefore α can be interpreted as a false alarm probability.

More precisely, if step k is accepted as described in Figure 3, then $\tilde{\theta}_k = \hat{\theta}_k$ and a nonzero loss $\mathsf{E}_{\theta^*} L(W^k, \tilde{\theta}_k, \hat{\theta}_k)$ can only occur if the estimator has been rejected before or at step k, which under the homogeneous parametric model case, is denoted as a "false alarm".

With the "propagation condition" (13) CVs are constructed as follows:

- Consider first $\mathfrak{z}_1$ and let $\mathfrak{z}_2 = \mathfrak{z}_3 = \ldots = \mathfrak{z}_{K-1} = \infty$. This leads to the estimates $\hat{\theta}_k(\mathfrak{z}_1)$ and the value $\mathfrak{z}_1$ is selected as the minimal one for which

$$\sup_{\theta^*} \mathsf{E}_{\theta^*} |L\{W^k, \tilde{\theta}_k, \hat{\theta}_k(\mathfrak{z}_1)\}|^r \leq \frac{\alpha \mathfrak{r}_r}{K-1}, k = 2, \ldots, K. \tag{14}$$

- Suppose $\mathfrak{z}_1, \ldots, \mathfrak{z}_{k-1}$ have been fixed, and set $\mathfrak{z}_k = \ldots = \mathfrak{z}_{K-1} = \infty$. With estimate $\hat{\theta}_m(\mathfrak{z}_1, \ldots, \mathfrak{z}_k)$ for $m = k+1, \ldots, K$. select $\mathfrak{z}_k$ as the minimal value which fulfills

$$\sup_{\theta^*} \mathsf{E}_{\theta^*} |L\{W^m, \tilde{\theta}_m, \hat{\theta}_m(\mathfrak{z}_1, \ldots, \mathfrak{z}_k)\}|^r \leq \frac{k\alpha \mathfrak{r}_r}{K-1} \tag{15}$$

  for $m = k+1, \ldots, K$.

Inequality (14) describes the impact of the k CV to the risk, while the factor $\frac{k\alpha}{K-1}$ in (15) ensures that every $\mathfrak{z}_k$ has the same impact. The values of $(\alpha, r, h_1, \ldots, h_K)$ are prespecified hyperparameters for which robustness and sensitivity issues will be discussed in Section 3. The following theorem provides insight into the form of $\mathfrak{z}_k$.

A risk bound for a constant model ($\theta(t) = \theta^*$) has been given in (13). In order to expand this to a nonparametric $\theta(t)$, the "Small Modeling Bias (SMB)" condition is employed:

$$\Delta(\theta) \stackrel{\text{def}}{=} \sum_{t=1}^{365} \mathcal{K}\{\theta(t), \theta\} \mathbf{I}\{w(s,t,h_k) > 0\} \leq \Delta, \forall k < k^*, \tag{16}$$

where k^* is the maximum k satisfying (16), also called "oracle". Consequently the estimation risk for $\theta(t)$ is described for $k \leq k^*$ by the "propagation" property:

$$\mathsf{E}_{\theta(\cdot)} \log\{1 + |L(W^k, \tilde{\theta}_k, \hat{\theta}_k)|^r / \mathfrak{r}_r\} \leq \Delta + \alpha. \tag{17}$$

An estimate for the oracle k^* is given via the adaptive estimate $\hat{\theta}_{\hat{k}}$. The estimate $\hat{\theta}_{\hat{k}}$ behaves similarly to the oracle estimate $\tilde{\theta}_{k^*}$ since it is "stable" in the sense that even if the described selection scheme (14), (15) overshoots k^*, the resulting estimate $\hat{\theta}_{\hat{k}}$ is still close to the oracle $\tilde{\theta}_{k^*}$. In fact the attained quality of estimation during "propagation" is not lost at further steps:

$$L(W^{k^*}, \tilde{\theta}_{k^*}, \hat{\theta}_{\hat{k}})\, \mathbf{I}\{\hat{k} > k^*\} \leq \mathfrak{z}_{k^*}$$

In other words, $\hat{\theta}_{\hat{k}}$ lies in the confidence set of $\tilde{\theta}_{k^*}$. A combination of the propagation and stability property leads to the "oracle" property:

$$\mathsf{E}_{\theta(\cdot)} \log\left\{1 + \frac{|L(W^{k^*}, \tilde{\theta}_{k^*}, \theta)|^r}{\mathfrak{r}_r}\right\} \leq \Delta + 1,$$
$$\mathsf{E}_{\theta(\cdot)} \log\left\{1 + \frac{|L(W^{k^*}, \tilde{\theta}_{k^*}, \hat{\theta}_{\hat{k}})|^r}{\mathfrak{r}_r}\right\} \leq \Delta + \alpha + \log\left\{1 + \frac{\mathfrak{z}_{k^*}}{\mathfrak{r}_r}\right\},$$

for θ with $\Delta(W^k, \theta) \leq \Delta$ and $k \leq k^*$. These bounds show that the risk of estimating adaptively is composed into three parts: the SMB, the false alarm rate, and a small term corresponding to the risk of overshooting.

3 Empirical analysis

We conduct an empirical analysis of temperature patterns for different cities (Figure 4). The data set contains the daily average temperatures for different cities in Europe (1900-2011), Asia (1900-2011), and the US (1900-2011): Atlanta, Beijing, Berlin, Essen, Houston, Kaohsiung, New York, Osaka, Portland, Taipei, and Tokyo.

Figure 4: Map of locations where temperature are collected

We first check seasonality, intertemporal correlation, and seasonal variation. Table 1 provides the coefficients of the Fourier truncated seasonal function (1) for some cities for different time periods. The coefficient a can be seen as the average temperature, the coefficient b as an indicator for global warming. The latter coefficients are stable even when the estimation is done in a window length of 10 years. In the sense of capturing volatility peak seasons, the left panel of Figure 5 visualises the power of capturing volatility peak seasons by the seasonal local smoother (3) using the quartic kernel over the estimates modeled under Fourier truncated series (1).

City	Period	$\hat{a}$	$\hat{b}$	$\hat{c}_1$	$\hat{d}_1$	$\hat{c}_2$	$\hat{d}_2$	$\hat{c}_3$	$\hat{d}_3$
Berlin	(19480101–20080527)	9.2173	0.0000	9.8932	-157.9123	0.2247	261.2850	0.1591	-127.7303
	(19730101–20080527)	9.3050	0.0001	10.0070	-161.2493	0.4601	-66.0530	-0.3723	-416.4776
	(19730101–20080527)	9.3050	0.0001	10.0070	-161.2493	0.4601	-66.0530	-0.3723	-416.4776
	(19830101–20080527)	9.4581	0.0001	10.0969	-161.7129	0.5205	-51.9929	0.3734	42.0874
	(19930101–20080527)	9.5923	0.0002	10.1995	-162.9774	0.6564	-37.1548	0.4241	41.9970
	(20030101–20080527)	9.6948	0.0007	10.1954	-162.3343	0.5554	-43.2293	0.3269	1.5998
Kaohsiung	(19730101–20081231)	24.2289	0.0001	0.9157	-145.6337	-4.0603	-78.1426	-1.0505	10.6041
	(19730101–19821231)	24.4413	0.0001	2.1112	-129.1218	-3.3887	-91.1782	-0.8733	20.0342
	(19830101–19921231)	25.0616	0.0003	2.0181	-135.0527	-2.8400	-89.3952	-1.0128	20.4010
	(19930101–20021231)	25.3227	0.0003	3.9154	-165.7407	-0.7405	-51.4230	-1.1056	19.7340
New York	(19490101–20081204)	53.1473	0.0001	18.6810	-143.4051	-3.3872	271.5072	-0.4203	-16.3125
	(19730101–20081204)	53.6992	0.0001	18.0092	-148.4124	-3.5236	279.6876	-0.4756	-21.8090
	(19730101–19821204)	53.6037	-0.0000	17.7446	-155.2453	-3.7769	289.7932	-0.8326	-4.2257
	(19830101–19921204)	54.8740	-0.0003	17.6924	-152.7461	-3.4245	284.6412	-0.4933	-218.9204
	(19930101–20021204)	53.8050	0.0003	17.6942	-153.3997	-3.4246	285.7958	0.5753	-315.2792
	(20030101–20081204)	52.9177	0.0012	17.8425	-151.2977	-3.8837	287.2022	-0.1290	-216.7298
Tokyo	(19730101–20081231)	15.7415	0.0001	8.9171	-162.3055	-2.5521	-7.8982	-0.7155	-15.0956
	(19730101–19821231)	15.8109	0.0001	9.2855	-162.6268	-1.9157	-16.4305	-0.5907	-13.4789
	(19830101–19921231)	15.4391	0.0004	9.4022	-162.5191	-2.0254	-4.8526	-0.8139	-19.4540
	(19930101–20021231)	16.4284	0.0001	8.8176	-162.2136	-2.1893	-17.7745	-0.7846	-22.2583
	(20030101–20081231)	16.4567	0.0001	8.5504	-162.0298	-2.3157	-18.3324	-0.6843	-16.5381

Table 1: Seasonality estimates $\hat{\Lambda}_t$ of daily average temperature. All coefficients are nonzero at 1% significance level.

City	Period	ADF	KPSS
Atlanta	19480101–20081204	-55.55+	0.21***
Beijing	19730101–20090831	-30.75+	0.16***
Berlin	19480101–20080527	-40.94+	0.13**
Essen	19700101–20090731	-23.87+	0.11*
Houston	19700101–20081204	-38.17+	0.05*
Kaohsiung	19730101–20091210	-37.96+	0.05*
New York	19490101–20081204	-56.88+	0.08*
Osaka	19730101–20090604	-18.65+	0.09*
Portland	19480101–20081204	-45.13+	0.05*
Taipei	19920101–20090806	-32.82+	0.09*
Tokyo	19730101–20090831	-25.93+	0.06*

Table 2: ADF and KPSS-Statistics for the detrended daily average temperature time series for different cities. '+', '*', '**' and '***' corresponds to significance levels of 0.01, 0.1, 0.05 and 0.01 respectively.

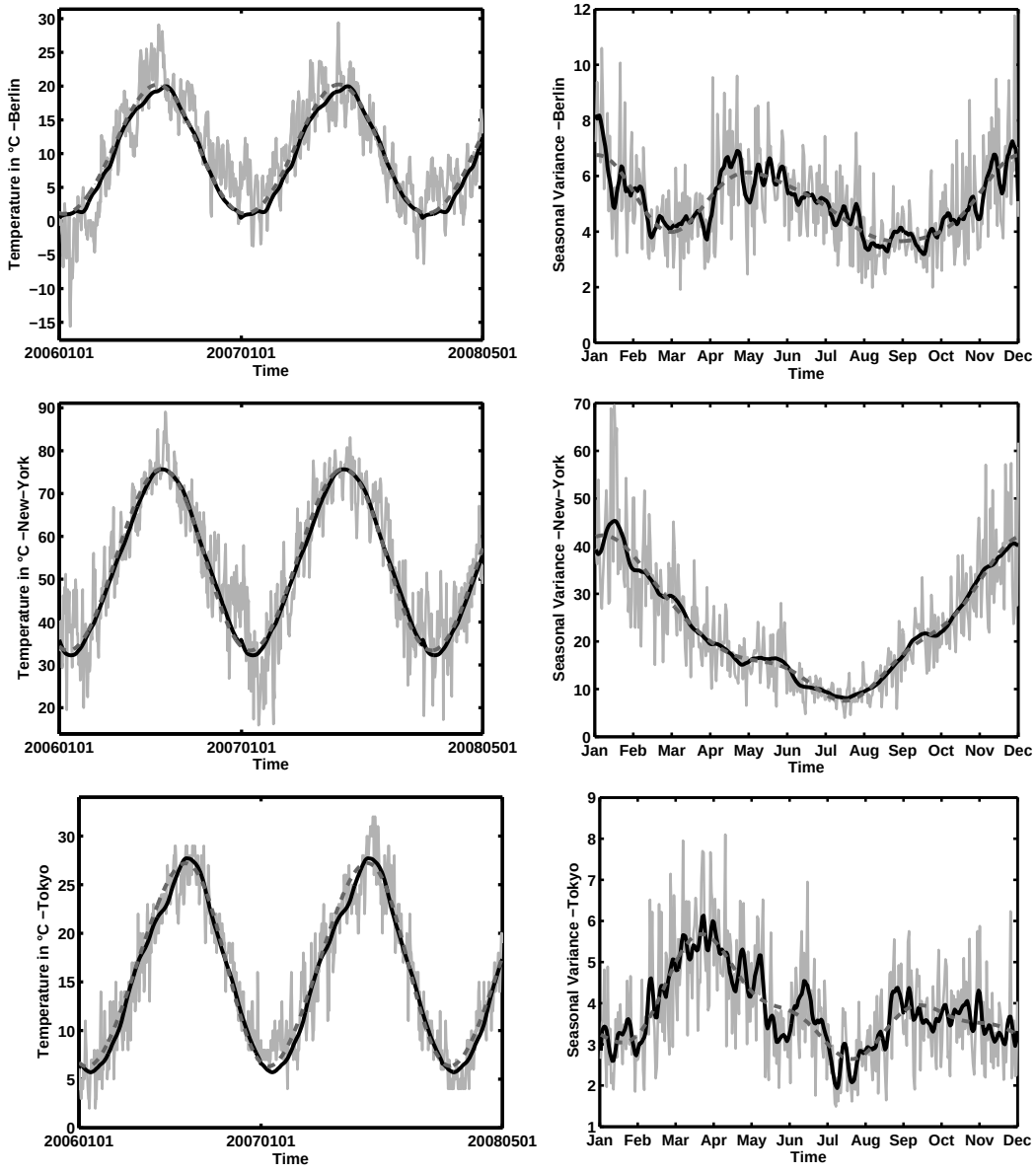


Figure 5: The empirical (black line), the Fourier truncated (dotted grey line), and the local linear (grey line) seasonal mean (left panel) and variance component (right panel) using quartic kernel and bandwidth $h = 4.49$.

After removing the local linear seasonal mean (3) from the daily average temperatures ($X_t = T_t - \Lambda_{t,LLR}$), we check that X_t is a stationary process with the Augmented Dickey-Fuller (ADF) and the KPSS tests. The analysis of the partial autocorrelations and the Akaike Information criterion (AIC) suggest that an $AR(3)$ model fits the temperature evolution well. Table 2 presents the results of the stationarity tests. The empirical seasonal variation (square residuals after seasonal and intertemporal fitting), the seasonal variation curves (2) and (4) are displayed on the right panel in Figure 5, while normality tests for the residuals are displayed in Table 5 (see there truncated Fourier method). All seasonal variance estimators lead to residuals that are far from being normally distributed. These facts are of course not an ideal platform for risk neutral pricing (based on standard stochastic financial models). The heavytailedness, as seen in Figure 1, may be attributed to an unsatisfactory extraction of the heteroscedasticity (or mean) function. As a solution we employ a localisation scheme.

The adjustment in the smoothing parameter h will provide the localisation in time. The bandwidth sequences are selected from six candidates: $(1, 2, 3, 4, 5, 6, 7)$, $(1, 2, 3, 5, 7, 10, 13)$, $(3, 5, 7, 9, 11, 13, 15)$, $(3, 5, 8, 12, 17, 23, 30)$, $(5, 7, 10, 14, 19, 25, 32)$, and $(7, 9, 11, 14, 17, 10, 24)$. These candidates are chosen according to the lowest Anderson–Darling (AD) statistic. The best candidate for the bandwidth sequence is the one which yields a residual distribution closest to the normal one. Smoothing the selected bandwidths gives another adaptive estimator, implemented, but not discussed here, due to space limitations.

The CVs as calibrated from (14) and (15) are given in Figure 6. The left side provides CVs simulated from a sample of 10000 observations for a quartic kernel for both mean with $\theta^* = 0$ and variance with $\theta^* = 1$, $r = 0.5$ and different values of significance level α. The CVs for different bandwidth sequences are displayed in the right side of Figure 6. The CVs, as one observes, are relatively robust to the choice of r and α.

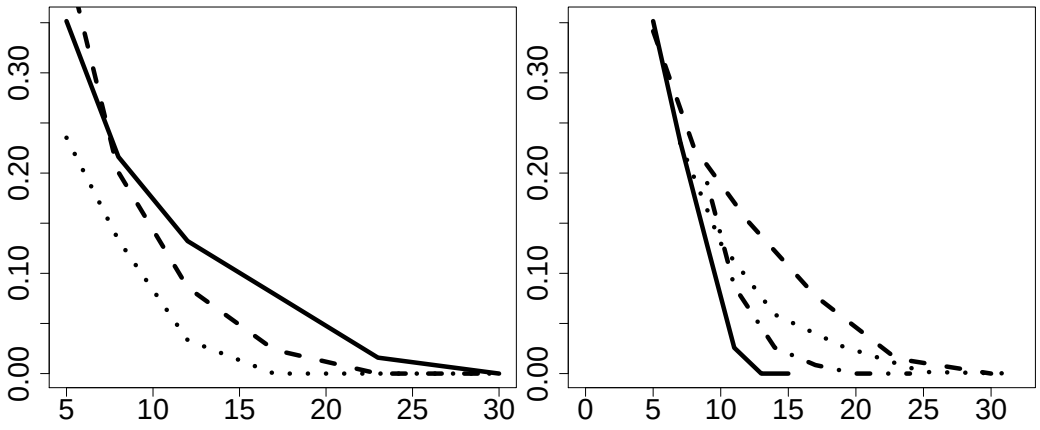
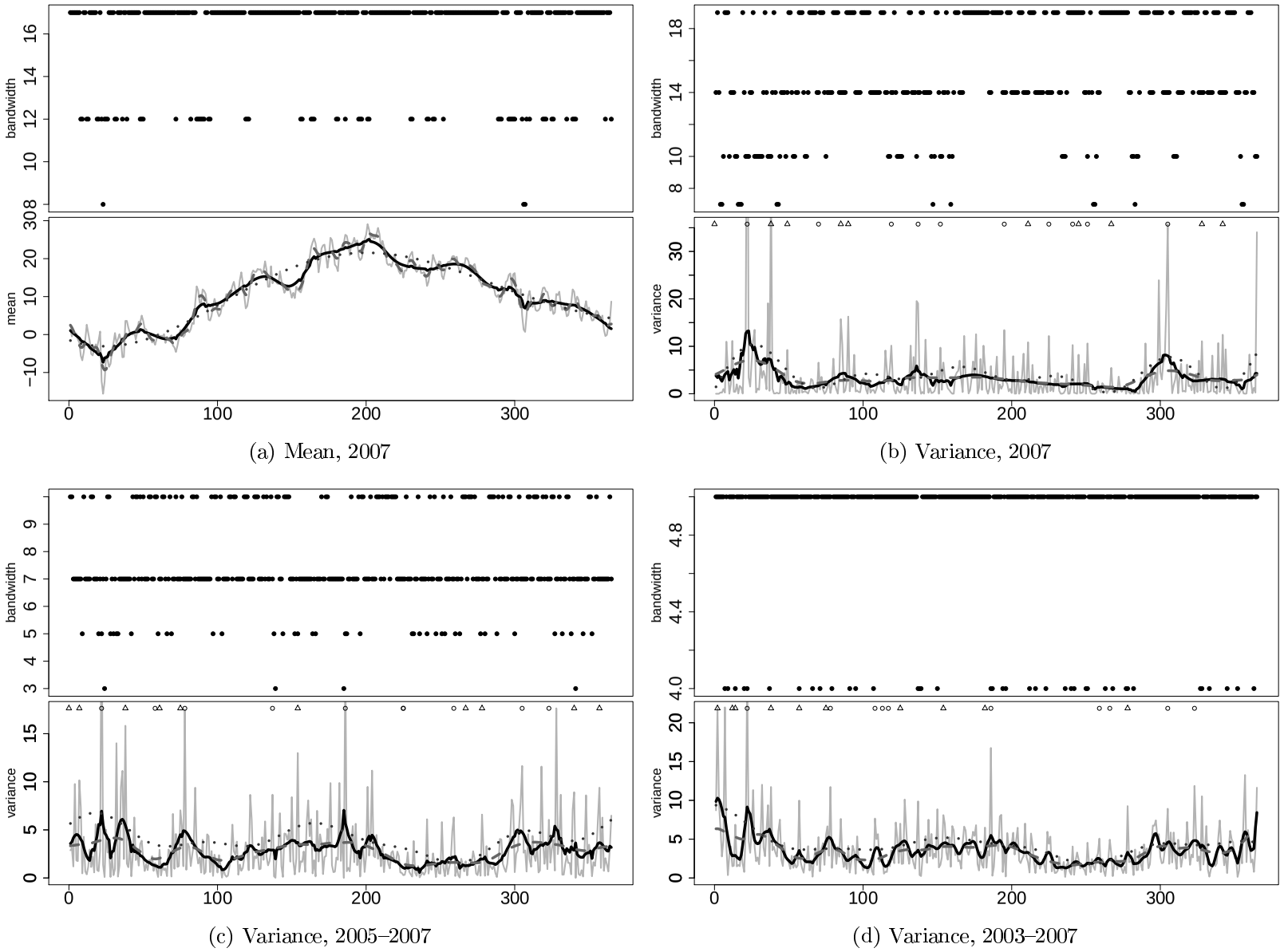


Figure 6: Simulated CVs for likelihood of seasonal variance (6) with $\theta^* = 1$, $r = 0.5$, $MC = 10000$ with $\alpha = 0.3$ (dotted), 0.5 (dashed), 0.7 (solid) for the bandwidth sequence $(3, 5, 8, 12, 17, 23, 30)$ on the left plot and with $\alpha = 0.3$ and for sequences $(3, 5, 7, 9, 11, 13, 15)$ (solid), $(3, 5, 8, 12, 17, 23, 30)$ (dashed), $(5, 7, 10, 14, 19, 25, 32)$ (dotted), and $(7, 9, 11, 14, 17, 10, 24)$ (dot-dashed) on the right plot.



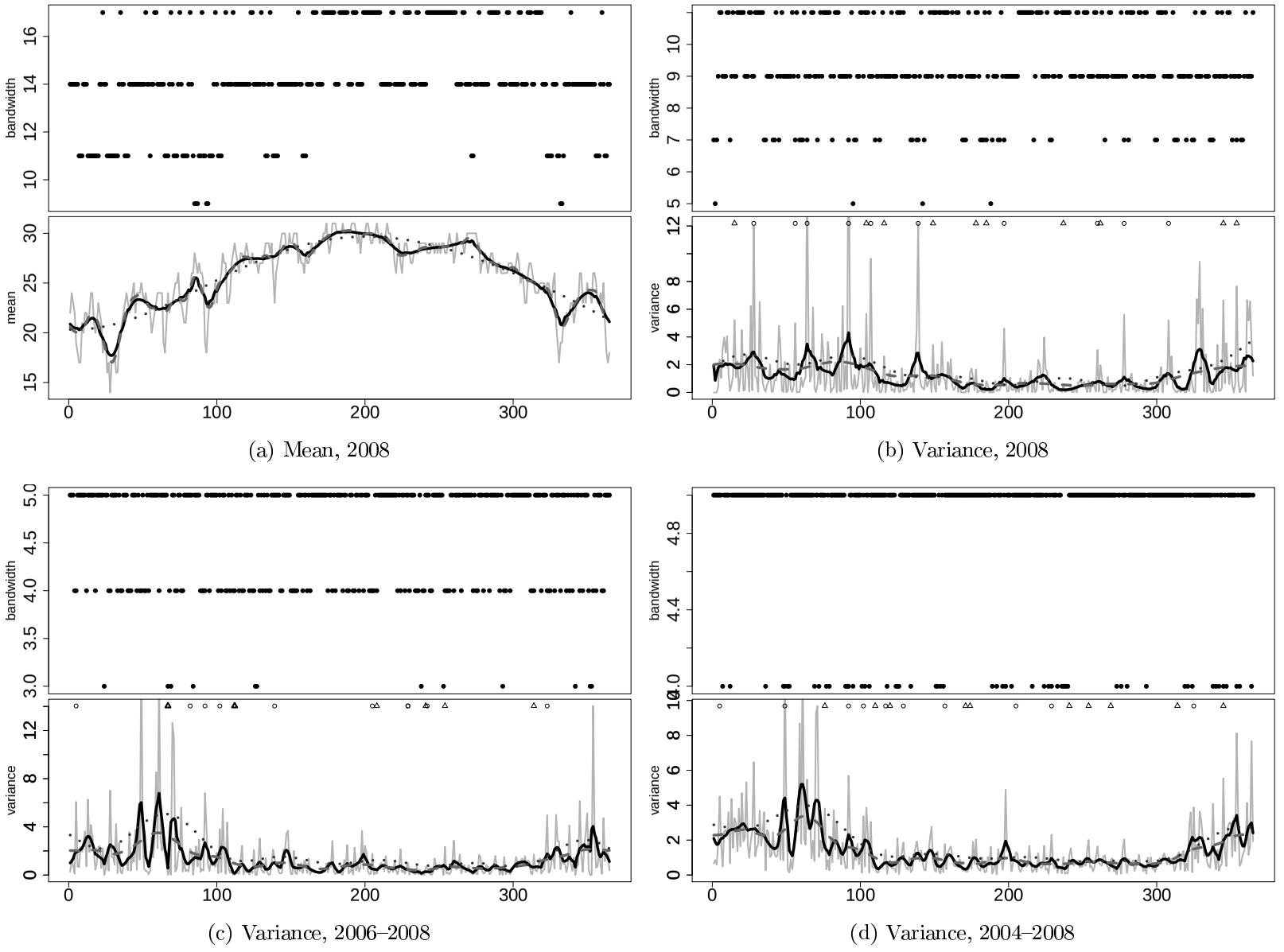
(a) Mean, 2007

(b) Variance, 2007

(c) Variance, 2005–2007

(d) Variance, 2003–2007

Figure 7: Estimation of mean and variance for Berlin. In each figure sequence of bandwidths (upper panel), averaged observations (solid grey line), nonparametric function estimation with fixed bandwidth (dashed grey line), adaptive bandwidth (solid black line) and truncated Fourier (dotted line) (bottom panel of each figure). Circles and triangles in each bottom panel for variance represents 10 smallest and 10 largest outliers respectively.



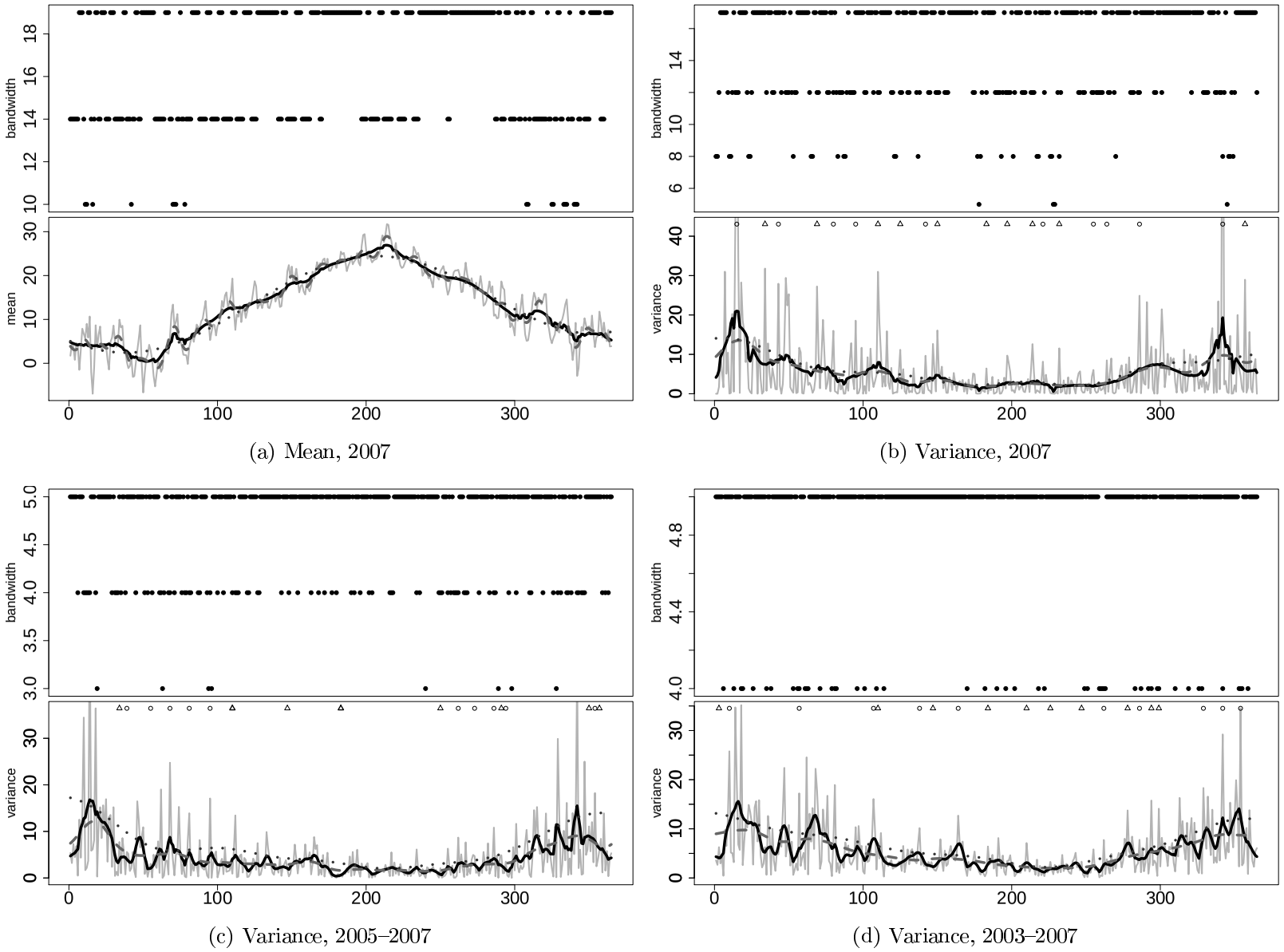
(a) Mean, 2008

(b) Variance, 2008

(c) Variance, 2006–2008

(d) Variance, 2004–2008

Figure 8: Estimation of mean and variance for Kaohsiung. In each figure sequence of bandwidths (upper panel), averaged observations (solid grey line), nonparametric function estimation with fixed bandwidth (dashed grey line), adaptive bandwidth (solid black line) and truncated Fourier (dotted line) (bottom panel of each figure). Circles and triangles in each bottom panel for variance represents 10 smallest and 10 largest outliers respectively.



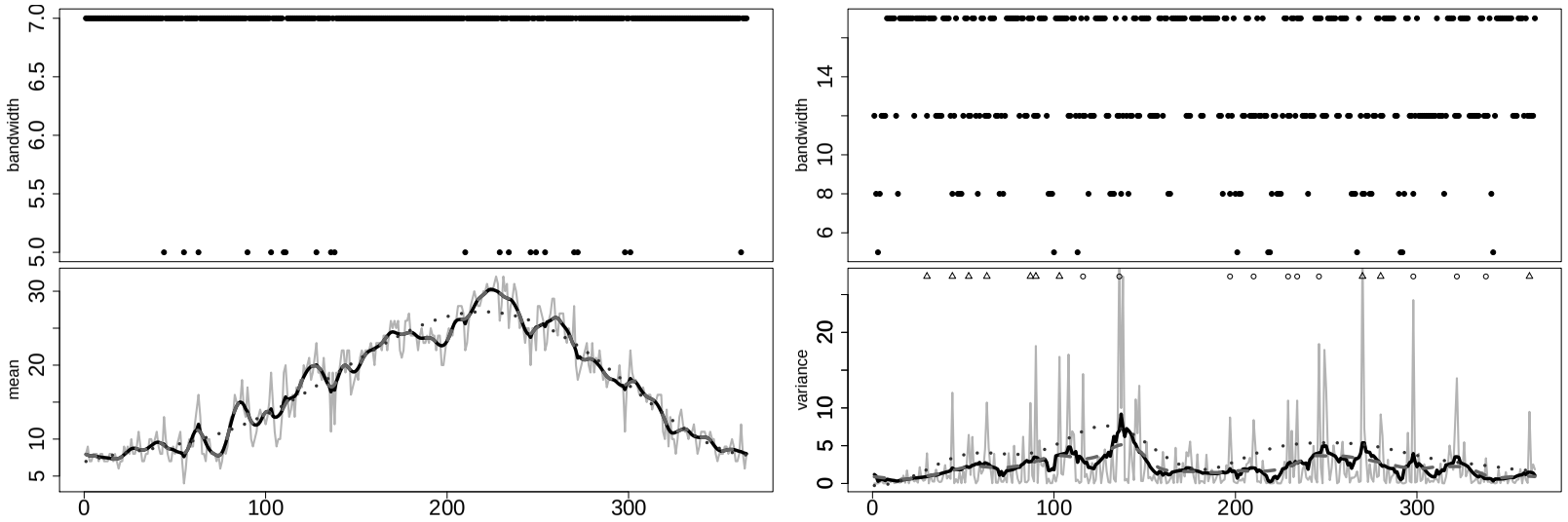
(a) Mean, 2007

(b) Variance, 2007

(c) Variance, 2005–2007

(d) Variance, 2003–2007

Figure 9: Estimation of mean and variance for New York. In each figure sequence of bandwidths (upper panel), averaged observations (solid grey line), nonparametric function estimation with fixed bandwidth (dashed grey line), adaptive bandwidth (solid black line) and truncated Fourier (dotted line) (bottom panel of each figure). Circles and triangles in each bottom panel for variance represents 10 smallest and 10 largest outliers respectively.



(a) Mean, 2008

(b) Variance, 2008

(c) Variance, 2006–2008

(d) Variance, 2004–2008

Figure 10: Estimation of mean and variance for Tokyo. In each figure sequence of bandwidths (upper panel), averaged observations (solid grey line), nonparametric function estimation with fixed bandwidth (dashed grey line), adaptive bandwidth (solid black line) and truncated Fourier (dotted line) (bottom panel of each figure). Circles and triangles in each bottom panel for variance represents 10 smallest and 10 largest outliers respectively.

A one year period is considered in the first place for demonstration purposes, while later we show how the results change with different time length periods. Figures 7, 8, 9, and 10 present the general results for different cities under different adaptive localising schemes for seasonal mean (Me) and seasonal variance (Va): with fixed bandwidth curve (fi), adaptive bandwidth curve (ad), and truncated Fourier (Fourier) for different time intervals. The seasonal mean is estimated jointly over the years, using $\alpha = 0.7$ and power level $r = 0.5$.

The upper panel of each variance plot in Figures 7–10 shows the sequence of bandwidths; the bottom panel displays the variance estimation with fixed bandwidth (dashed line), Fourier truncated method (dotted line), and adaptive bandwidth (solid black line). In all countries, one observes significant differences between the estimates. In particular, in cities like Kaohsiung and New York, one observes more variation of the seasonal variance curves during peak seasons (winter and summer times). The triangles and circles in the bottom panel of each variance plot help us trace the source of the non-normality over time, since they correspond to ten dots of the upper and lower tails of the QQ-plots of square residuals respectively (see Figure 11 for the Berlin results). The left top plots of Figures 7– 10 show the mean case. Unlike the seasonal variance function, we do not observe a big variation of smoothness in the mean function. One can see that in all cities, the bandwidths vary over the yearly cycle with a slight degree of non homogeneity for Kaoshiung.

An approach to cope with the non normality brought in by more observations is to estimate mean functions year by year (SeMe), and then aggregate the residuals for variance estimation. We therefore estimate the joint/separate seasonal mean (JoMe/SeMe) and seasonal variance (Va) curves with fixed bandwidth curve (fi), adaptive bandwidth curve (ad). The variance plots in Figures 7–10 display the behaviour of the variance function estimation when the period length changes. The average over years acts as a smoother when we consider more years. The estimated $AR(L)$ parameters for different cities using joint/separate mean (JoMe/SeMe) with different bandwidth curves are illustrated in Table 3. The results again show that an $AR(3)$ fits well the stylised facts of temperature.

We tackle the problem of loosing information when considering estimates at the individual level or averaging mean functions over time, with a refined approach that considers the minimum variance between the aggregation of yearly local mean function estimates and an optimal local estimate θ^o. Once the sets of local mean functions have been identified, the aggregated local function can be defined as the weighted average of all the observations in a given time set. Formally, if $\hat{\theta}^j(t)$ is the localised observation at time t of year j, the aggregated local function is given by:

$$\hat{\theta}_\omega(t) = \sum_{j=1}^{J} \omega_j \hat{\theta}^j(t). \tag{18}$$

With this aggregation step across J, we give the same weight to all observations, even to observations that were unimportant at the yearly level. Then a reasonable optimised estimate will be:

$$\arg\min_{\omega} \sum_{j=1}^{J} \sum_{t=1}^{365} \{\hat{\theta}_\omega(t) - \hat{\theta}_j^o(t)\}^2 \quad \text{subject to} \quad \Sigma_{j=1}^{J} \omega_j = 1; \omega_j > 0, j = 1, \ldots, J, \tag{19}$$

where the weights are assumed to be exogenous and nonstochastic, and $\hat{\theta}_j^o$ is defined as one of the following: 1 (Locave), $\hat{\theta}_j^o(t) = J^{-1} \sum_{j=1}^{J} \hat{\sigma}_j^2(t)$, the average of seasonal empirical variances over years, 2, (Locsep) $\hat{\theta}_j^o(t) = \hat{\sigma}_j^2(t)$, the yearly empirical variances, 3, one of above two approaches with

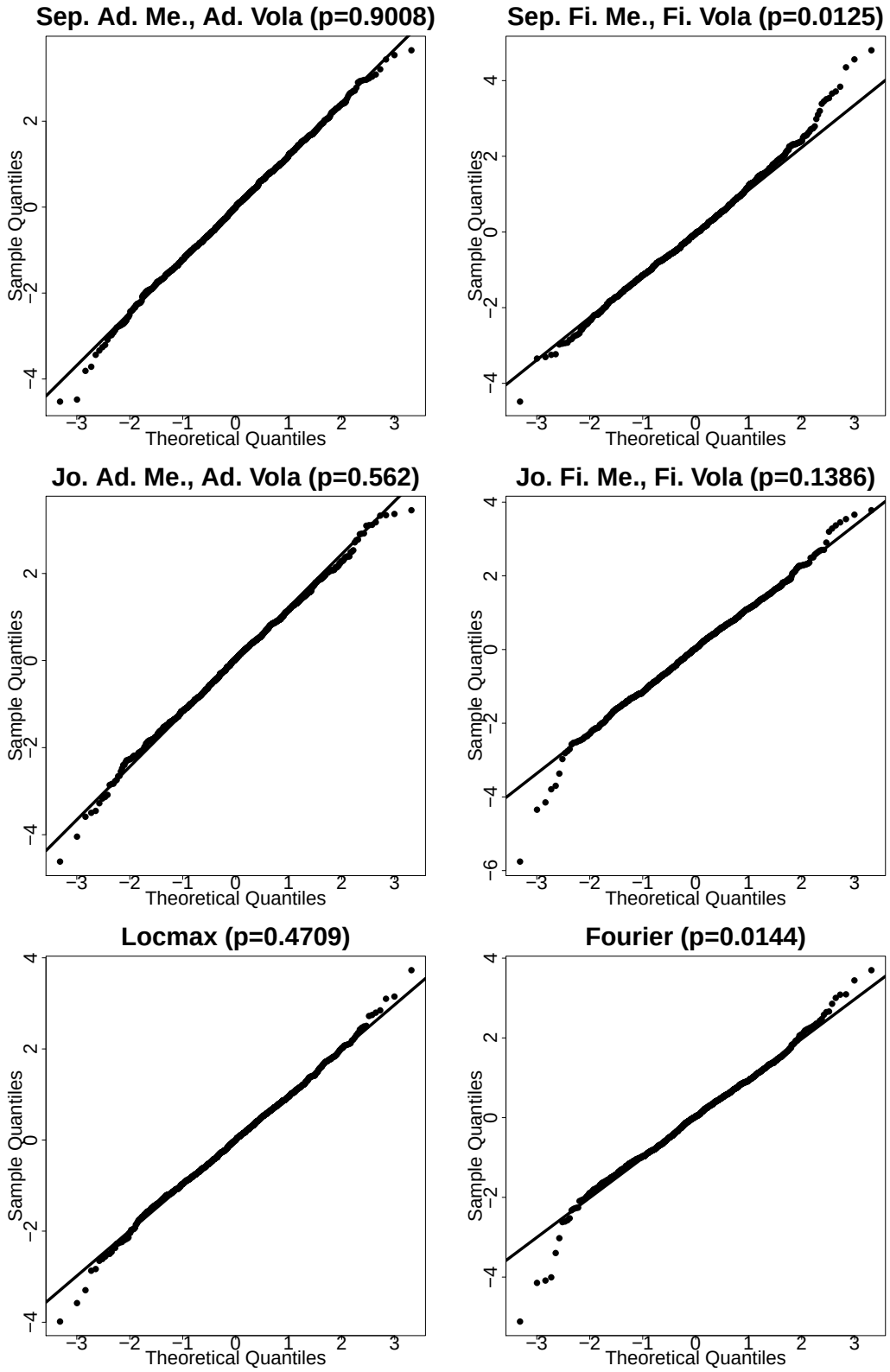


Figure 11: QQ-plot for standardised residuals from Berlin using different methods for the data from 2005-2007 (3 years).

		AR(p) parameters															CAR(p) parameters		
		1 year			2 years			3 years			4 years			5 years			2 years		
		β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3	α_1	α_2	α_3
Berlin	JoMe adMe	0.744	-0.244	-0.008	0.972	-0.303	0.060	0.971	-0.286	0.072	0.960	-0.252	0.073	0.985	-0.288	0.101	2.015	1.318	0.202
	JoMe fiMe	0.301	-0.330	-0.267	0.908	-0.283	0.113	0.937	-0.277	0.113	0.939	-0.252	0.097	0.976	-0.289	0.114	2.024	1.337	0.199
	SeMe adMe adVo				0.795	-0.255	-0.057	0.807	-0.293	0.000	0.727	-0.248	-0.038	0.768	-0.290	0.000	2.204	1.664	0.517
	SeMe fiMe fiVo				0.344	-0.333	-0.196	0.346	-0.334	-0.176	0.294	-0.332	-0.175	0.326	-0.342	-0.165	2.655	2.644	1.185
	Locave				0.795	-0.255	-0.057	0.971	-0.286	-0.072	0.960	-0.252	-0.073	0.768	-0.290	0.000	2.204	1.664	0.517
	Locsep				0.795	-0.255	-0.057	0.971	-0.286	-0.072	0.960	-0.252	-0.073	0.768	-0.290	0.000	2.204	1.664	0.517
	Locmax				0.795	-0.255	-0.057	0.971	-0.286	-0.072	0.960	-0.252	-0.073	0.768	-0.290	0.000	2.204	1.664	0.517
	Fourier	0.968	-0.292	0.065	1.017	-0.309	0.075	1.017	-0.297	0.074	0.981	-0.264	0.079	1.000	-0.297	0.101	1.982	1.273	0.215
Kaohsiung	JoMe adMe	0.601	-0.132	-0.152	0.735	-0.105	-0.071	0.826	-0.204	0.000	0.807	-0.155	0.000	0.785	-0.139	0.000	2.215	1.569	0.354
	JoMe fiMe	0.507	-0.164	-0.200	0.725	-0.112	-0.080	0.806	-0.195	0.000	0.788	-0.146	0.000	0.773	-0.139	0.000	2.227	1.593	0.366
	SeMe adMe adVo				0.442	-0.138	-0,152	0.401	-0.200	-0.154	0.412	-0.159	-0.132	0.316	-0.194	-0.146	2.557	2.253	0.084
	SeMe fiMe adVo				0.504	-0.152	-0.210	0.502	-0.219	-0.206	0.479	-0.184	-0.196	0.457	-0.189	-0.190	2.495	2.143	0.857
	Locave				0.442	-0.138	-0.152	0.401	-0.200	-0.154	0.412	-0.159	-0.132	0.316	-0.194	-0.146	2.557	2.253	0.848
	Locsep				0.442	-0.138	-0.152	0.401	-0.200	-0.154	0.412	-0.159	-0.132	0.316	-0.194	-0.146	2.557	2.253	0.848
	Locmax				0.442	-0.138	-0.152	0.401	-0.200	-0.154	0.412	-0.159	-0.132	0.316	-0.194	-0.146	2.557	2.253	0.848
	Fourier	0.820	-0.116	-0.020	0.791	-0.102	-0.051	0.853	0.185	-0.000	0.820	-0.128	0.011	0.802	-0.126	0.001	2.208	1.520	0.363
New-York	JoMe adMe	0.601	-0.207	0.000	0.636	-0.197	0.161	0.728	-0.212	0.127	0.688	-0.212	0.165	0.708	-0.190	0.140	2.292	1.774	0.342
	JoMe fiMe	0.342	-0.277	-0.118	0.644	-0.205	0.138	0.692	-0.213	0.150	0.695	-0.222	0.158	0.711	-0.195	0.139	2.289	1.773	0.345
	SeMe adMe adVo				0.554	-0.247	0.000	0.495	-0.282	0.000	0.481	-0.287	0.000	0.475	-0.247	-0.050	2.289	1.773	0.345
	SeMe fiMe fiVo				0.278	-0.312	-0.118	0.283	-0.322	-0.133	0.299	-0.326	-0.120	0.288	-0.308	-0.152	2.712	2.732	1.172
	Locave				0.000	0.554	-0.247	0.495	-0.282	0.000	0.481	-0.287	0.000	0.475	0.247	-0.050	2.525	1.803	0.328
	Locsep				0.000	0.554	-0.247	0.495	-0.282	0.000	0.481	-0.287	0.000	0.475	0.247	-0.050	2.525	1.803	0.328
	Locmax				0.000	0.554	-0.247	0.495	-0.282	0.000	0.481	-0.287	0.000	0.475	0.247	-0.050	2.525	1.803	0.328
	Fourier	0.720	-0.184	0.066	0.763	-0.193	0.117	0.751	-0.205	0.134	0.750	-0.212	0.148	0.756	-0.190	0.128	2.244	1.678	0.306
Tokyo	JoMe adMe	0.152	-0.157	-0.294	0.431	-0.074	0.000	0.492	-0.087	0.000	0.541	-0.092	0.072	0.567	-0.102	0.083	2.433	1.968	0.452
	JoMe fiMe	0.158	-0.150	-0.296	0.452	-0.074	-0.054	0.512	-0.103	0.000	0.564	-0.097	0.045	0.575	0.105	0.072	2.425	1.745	0.248
	SeMe adMe adVo				0.360	-0.106	-0.162	0.330	-0.177	-0.136	0.307	-0.190	-0.138	0.317	-0.198	-0.124	2.639	2.384	0.907
	SeMe fiMe fiVo				0.225	-0.177	-0.245	0.265	-0.211	-0.184	0.269	-0.212	-0.171	0.252	-0.232	-0.171	2.774	2.726	1.197
	Locave				0.360	-0.106	-0.162	0.333	-0.177	-0.136	0.307	-0.190	-0.138	0.317	-0.198	-0.124	2.639	2.384	0.907
	Locsep				0.360	-0.106	-0.162	0.333	-0.177	-0.136	0.307	-0.190	-0.138	0.317	-0.198	-0.124	2.639	2.384	0.907
	Locmax				0.360	-0.106	-0.162	0.333	-0.177	-0.136	0.307	-0.190	-0.138	0.317	-0.198	-0.124	2.639	2.384	0.907
	Fourier	0.534	0.038	-0.039	0.561	-0.039	-0.017	0.581	-0.091	0.044	0.597	-0.088	0.060	0.615	-0.096	0.079	2.438	1.916	0.495

Table 3: $AR(L)$ parameters for Berlin, Kaohsiung, New York, Tokyo using joint/separate mean (JoMe/SeMe) with fixed bandwidth curve (fi), adaptive bandwidth curve (ad) seasonal mean (Me)curve. $CAR(L)$ parameters estimated with the last 2 years.

	Method	KS	JB	AD		KS	JB	AD
1 Year	JoMe adMe adVa	*0.033*	*0.233*	*0.200*				
	JoMe fiMe fiVa	*0.033*	0.300	0.333				
	Fourier	0.066	0.633	0.600				
	CD	0.066	0.733	0.700				
2 Years	JoMe adMe adVa	1.000	0.466	0.366	3 Years	0.900	0.600	0.600
	JoMe fiMe fiVa	0.966	0.766	0.633		0.700	0.866	0.866
	SeMe adMe adVa	1.000	*0.233*	*0.266*		1.000	*0.300*	*0.400*
	SeMe fiMe fiVa	1.000	0.666	0.533		0.866	0.766	0.800
	Locave	*0.033*	0.366	0.333		0.200	0.533	0.633
	Locsep	*0.033*	0.366	0.333		0.200	0.533	0.633
	Locmax	0.066	0.266	0.300		*0.166*	0.566	0.566
	Fourier	0.266	0.800	0.766		0.333	0.866	0.900
	CD	1.000	0.833	0.866		1.000	0.900	0.900
4 Years	JoMe adMe adVa	0.666	0.633	0.633	5 Years	0.566	0.800	0.766
	JoMe fiMe fiVa	0.500	0.933	0.933		0.533	0.933	0.933
	SeMe adMe adVa	0.900	*0.400*	*0.500*		0.600	*0.566*	*0.600*
	SeMe fiMe fiVa	0.633	0.800	0.866		0.400	0.933	0.866
	Locave	0.233	0.633	0.600		0.400	0.700	0.700
	Locsep	0.233	0.633	0.600		0.400	0.700	0.700
	Locmax	*0.200*	0.633	0.600		*0.333*	0.733	0.700
	Fourier	0.500	0.933	0.933		0.600	1.000	0.933
	CD	1.000	0.933	0.933		1.000	1.000	0.900

Table 4: Rejection rates of the normality at 5% level for 30 cities with different history, methods of estimation and normality tests. Tests for normality are Kolmogorov–Smirnov (KS), Jarque–Bera (JB) and AD. Methods used: joint/separate mean (JoMe/SeMe) with fixed/adaptive (fi/ad) bandwidth for the mean/variance (Me/Va), Locave, Locsep, Locmax, truncated Fourier (Fourier) and CD model. Highlighted with italic are models with smallest rejection rate for each goodness-of-fit (GoF) test and each history.

maximised p-values over year. One may interpret this normalisation of weights as an optimisation with respect to different frequencies (yearly, daily).

Kolmogorov–Smirnov (KS), Jarques–Bera (JB) and AD normality tests are taken to test the normality of the corrected residuals (after seasonal mean and variance). For each city, a rejection at 0.05 level is counted as 1 (else 0). The rejection rates over all cities under different estimation techniques are displayed in Table 4. Also the results are compared for different periods ($1-5$ years). A higher rejection rate would indicate a poorer performance of the relevant method. To make our conclusion more general, we add 18 cities more selected all around the world, including, Cordoba (Argentina), Adelaide (Australia), Sydney (Australia), Brasilia (Brazil), Larnaca (Cyprus), Cairo (Egypt), Paris (France), London (UK), Manchester (UK), Wellington (New Zealand), Moscow (Russia), Perm (Russia), Barcelona (Spain), Borlange (Sweden), Zurich (Switzerland), Istanbul (Turkey), Kyiv (Ukraine) and Lusaka (Zambia). The additional data is taken from NNDC Climate Data Online from $1999-2012$. Results for only 12 original cities can be found in the Supplementary material. We observe a superior performance of SeMe adMe adVa (for the AD and JB tests) and Locmax (best for the KS test and second best for the AD and JB tests) methods over the other alternatives like truncated Fourier, CD method and methods with fixed bandwidth. More detailed information on the p-values is represented in the Table 5 for the cities Berlin, Kaohsiung, New-York and Tokyo. Results on other 8 cities is given in the Supplementary material.

		1 year			2 years			3 years			4 years			5 years		
		KS	JB	AD	KS	JB	AD	KS	JB	AD	KS	JB	AD	KS	JB	AD
Berlin	JoMe adMe adVa	0.99	7.15e-01	0.99	5.06e-06	1.91e-01	0.55	0.01	2.41e-01	0.56	0.03	2.45e-01	0.18	0.13	2.17e-01	2.33e-01
	JoMe fiMe fiVa	0.87	3.71e-01	0.53	3.49e-03	1.81e-10	0.06	0.09	6.55e-08	0.13	0.06	7.29e-13	0.00	0.37	9.59e-14	8.22e-03
	SeMe adMe adVa				8.15e-06	1.52e-01	0.85	0.00	3.42e-01	0.90	0.02	4.18e-01	0.81	0.22	3.72e-01	7.85e-01
	SeMe fiMe fiVa				5.16e-04	2.89e-03	0.04	0.01	1.59e-05	0.01	0.11	3.88e-05	0.01	0.11	7.08e-05	3.52e-02
	Locave				8.48e-01	3.09e-01	0.88	0.93	1.47e-02	0.44	0.89	6.05e-04	0.31	0.68	1.27e-05	9.56e-02
	Locsep				8.48e-01	3.09e-01	0.88	0.93	1.47e-02	0.44	0.89	6.06e-04	0.31	0.68	1.27e-05	9.55e-02
	Locmax				9.79e-01	3.30e-01	0.94	0.87	1.60e-02	0.47	0.77	4.85e-04	0.35	0.47	3.85e-06	1.52e-01
	Fourier	0.77	8.04e-11	0.00	3.14e-01	0.00e+00	0.01	0.60	2.22e-16	0.01	0.66	2.88e-15	0.00	0.72	2.14e-12	1.16e-02
	CD	0.54	7.94e-11	0.01	4.94e-07	0.00e+00	0.00	0.00	0.00e+00	0.01	0.00	0.00e+00	0.00	0.00	0.00e+00	1.72e-06
Kaohsiung	JoMe adMe adVa	0.32	1.11e-02	1.82e-03	1.55e-07	9.90e-03	1.78e-02	2.38e-05	1.04e-11	1.57e-07	0.00	8.88e-16	6.88e-12	8.85e-06	0.00e+00	1.69e-15
	JoMe fiMe fiVa	0.54	2.97e-04	5.26e-04	1.83e-05	0.00e+00	2.76e-09	2.25e-03	0.00e+00	1.13e-14	0.00	0.00e+00	5.74e-19	4.80e-05	0.00e+00	8.87e-22
	SeMe adMe adVa				1.59e-07	1.62e-01	1.55e-02	8.61e-04	3.39e-05	6.71e-05	0.01	5.98e-07	5.00e-07	1.01e-02	4.61e-06	1.33e-05
	SeMe fiMe fiVa				6.29e-06	3.10e-07	9.60e-05	1.12e-03	1.26e-08	4.93e-06	0.02	2.88e-15	1.42e-09	9.87e-03	8.32e-15	3.55e-10
	Locave				6.18e-02	1.36e-04	8.55e-05	1.15e-02	1.56e-06	1.16e-06	0.01	2.29e-10	4.23e-09	2.14e-02	8.35e-10	2.56e-07
	Locsep				6.18e-02	1.36e-04	8.55e-05	1.15e-02	1.56e-06	1.16e-06	0.01	2.29e-10	4.24e-09	2.14e-02	8.39e-10	2.56e-07
	Locmax				5.92e-02	1.11e-04	4.44e-04	9.05e-03	1.57e-05	4.46e-06	0.00	3.58e-09	1.24e-08	7.30e-03	1.00e-07	9.03e-07
	Fourier	0.08	6.75e-09	6.12e-05	6.29e-03	0.00e+00	3.03e-10	3.89e-04	0.00e+00	2.01e-14	0.00	0.00e+00	2.15e-19	1.09e-05	0.00e+00	9.86e-23
	CD	0.12	3.80e-08	3.18e-05	1.49e-05	0.00e+00	1.95e-10	0.00e+00	0.00e+00	6.72e-20	0.00	0.00e+00	1.19e-24	0.00e+00	0.00e+00	3.84e-36
New-York	JoMe adMe adVa	0.51	0.02	0.02	3.45e-05	0.46	0.40	0.00	1.77e-01	0.26	0.08	0.08	0.30	0.06	3.09e-02	0.05
	JoMe fiMe fiVa	0.56	0.20	0.02	4.28e-04	0.01	0.11	0.01	2.77e-03	0.01	0.10	0.00	0.02	0.06	6.42e-04	0.00
	SeMe adMe adVa				1.83e-07	0.08	0.90	0.00	9.72e-01	0.70	0.00	0.50	0.61	0.00	4.48e-01	0.35
	SeMe fiMe fiVa				1.27e-06	0.03	0.02	0.00	5.50e-01	0.05	0.00	0.13	0.01	0.02	2.25e-03	0.00
	Locave				9.92e-01	0.57	0.74	0.97	5.83e-01	0.61	0.94	0.84	0.51	0.54	6.55e-01	0.15
	Locsep				9.92e-01	0.57	0.74	0.97	5.83e-01	0.61	0.94	0.84	0.51	0.54	6.55e-01	0.15
	Locmax				7.56e-01	0.82	0.97	0.95	6.37e-01	0.81	0.94	0.82	0.53	0.76	4.57e-01	0.24
	Fourier	0.99	0.42	0.81	7.97e-01	0.01	0.12	0.52	3.94e-03	0.01	0.23	0.00	0.00	0.18	2.45e-04	0.00
	CD	0.98	0.26	0.87	1.68e-05	0.06	0.03	0.00	8.64e-06	0.00	0.00	0.00	0.00	0.00	2.56e-05	0.00
Tokyo	JoMe adMe adVa	0.90	0.66	0.74	7.50e-03	2.62e-01	0.58	3.17e-02	3.24e-01	0.29	0.21	3.30e-01	3.44e-01	0.35	1.21e-01	7.33e-02
	JoMe fiMe fiVa	0.33	0.00	0.06	1.25e-02	1.77e-04	0.05	1.23e-01	2.45e-06	0.01	0.21	5.42e-07	1.13e-03	0.33	1.15e-08	1.50e-04
	SeMe adMe adVa				2.19e-06	8.69e-01	0.93	5.03e-04	7.82e-01	0.42	0.06	3.99e-01	2.74e-01	0.11	1.13e-01	1.60e-01
	SeMe fiMe fiVa				8.85e-04	5.22e-08	0.00	9.99e-02	0.00e+00	0.00	0.18	7.20e-13	2.80e-05	0.34	1.36e-14	1.23e-05
	Locave				4.56e-01	2.86e-02	0.17	1.96e-01	3.65e-03	0.07	0.17	6.48e-04	1.68e-02	0.38	2.22e-05	1.84e-02
	Locsep				4.56e-01	2.86e-02	0.17	1.96e-01	3.65e-03	0.07	0.17	6.48e-04	1.68e-02	0.38	2.21e-05	1.84e-02
	Locmax				3.32e-01	6.40e-02	0.18	1.55e-01	3.00e-03	0.08	0.22	1.65e-04	2.09e-02	0.44	1.17e-05	2.02e-02
	Fourier	0.62	0.00	0.00	3.71e-01	8.27e-07	0.00	2.21e-01	2.58e-10	0.00	0.15	3.47e-10	3.89e-04	0.29	4.15e-12	2.30e-04
	CD	0.18	0.00	0.00	7.11e-05	3.93e-13	0.00	3.33e-16	3.99e-13	0.00	0.00	2.22e-15	2.43e-05	0.00	8.34e-13	5.94e-05

Table 5: p-values for different models and GoF tests for Berlin, Kaohsiung, New-York and Tokyo.

4 Forecast and comparison

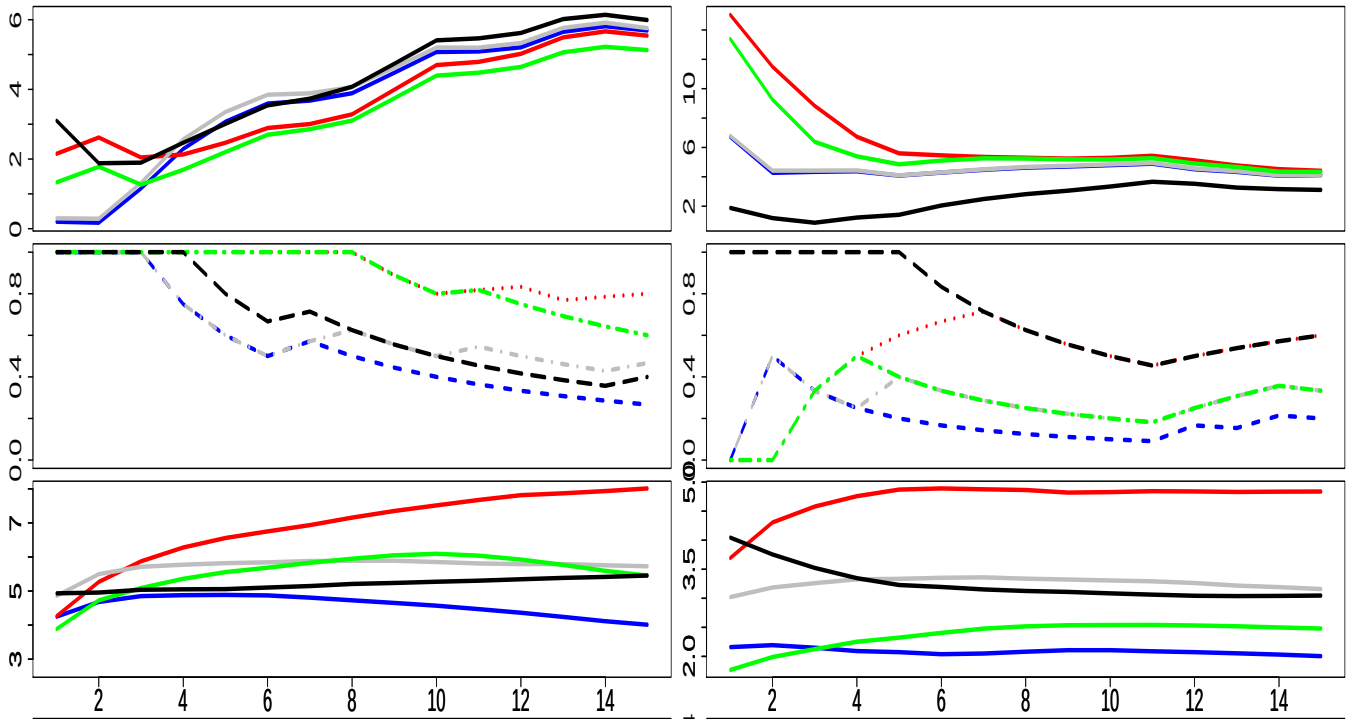
In this section we compare the forecasting accuracy of the proposed models to the CD model. CD mentioned that their point forecasts were always at least as good as the persistence and climatological forecasts, although not so good as the judgementally adjusted NWP forecast produced by EarthSat for a horizon of eight days. Therefore, good performance of the technique presented here could potentially suggest that our time series model is relevant for weather derivatives.

In Figures 12 and 13 we compare the out-of-sample forecast performance between five methods, namely SeMe adMe adVo, Locave, JoMe adMe adVo, truncated Fourier and CD. The comparison is provided at different time horizons ($1, \ldots, 15$ days) for Berlin, Kaohsiung, New York and Tokyo using 2 (Figure 12) and 3 (Figure 13) years of history. These figures contain information both on point forecast and interval forecast. The top panel of each plot shows the absolute deviation of the forecasted temperature from the true one, averaged over 1000 simulation pathes. This may be considered as the quality of the point forecast. In this terms, as we see in most cities and over all time horizons, we have at least one localising method better than or as good as the CD method. However, the CD model dominates others in the case of Kaohsiung. Bottom panel of each plot shows the averaged width of the point-wise confidence interval based on 1000 sample pathes. These curves represent the efficiency of the models. Although the truncated Fourier series method also look quite competitive in the point forecast, it usually has a very wide confidence interval, which is a sign of low efficiency. Other methods in this context are strictly better. The middle panel is showing the coverage of the true temperature by the confidence interval, where larger values represent higher quality. We do not see an outperforming behavior of the CD method over proposed adaptive techniques in almost all 12 cities. CD method performs well in Kaohsiung in terms of interval forecast, as one can see that it has a high coverage probability with a narrow confidence interval. As a conclusion, we do not claim strict superiority over the CD method in forecasting, but conclude, that both methods are quite competitive.

5 A temperature pricing example

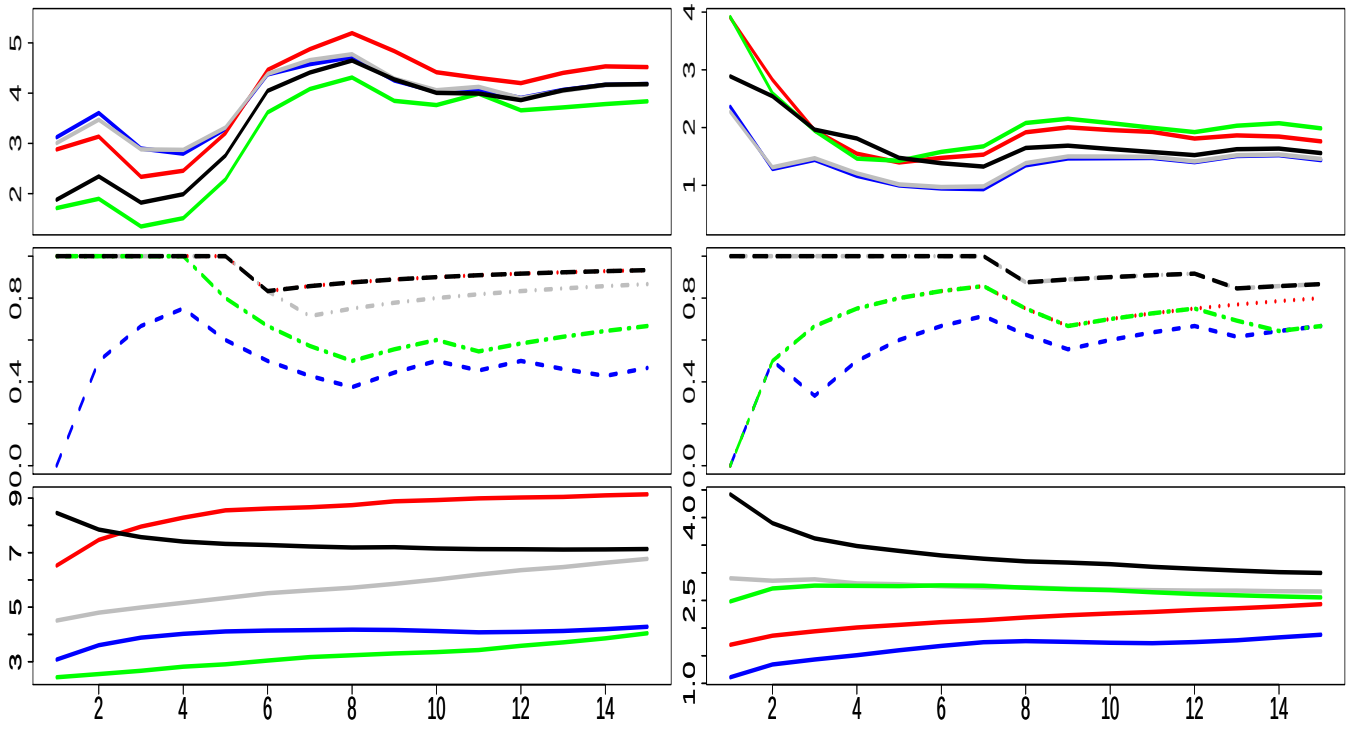
Based on a model for the daily temperature evolution, futures and European options written on temperature indices traded at the Chicago Mercantile Exchange (CME) can be calibrated. Temperature futures are contracts written on different temperature indices measured over specified periods $[\tau_1, \tau_2]$ such as weeks, months, or quarters of a year. Temperature futures allow one party to profit if the realized index value is greater than a predetermined strike level and the other party benefits if the index value is below. The owner of a call (put) option written on futures $F(t, \tau_1, \tau_2)$ with exercise time $t \leq \tau_1$ and measurement period $[\tau_1, \tau_2]$ will receive $\max\{F(t, \tau_1, \tau_2) - K, 0\}$ ($\max\{K - F(t, \tau_1, \tau_2), 0\}$), where K denotes the strike level. In other words, in exchange for the payment of the premium, the call (put) option gives the buyer a payoff based upon the difference between the realized index value and the strike level.

The most common temperature indices $I(\tau_1, \tau_2)$ are: Heating Degree Day (HDD), Cooling Degree Day (CDD), Cumulative Averages Temperatures (CAT) or Average Accumulative Temperatures



(a) Berlin (2006-2007)

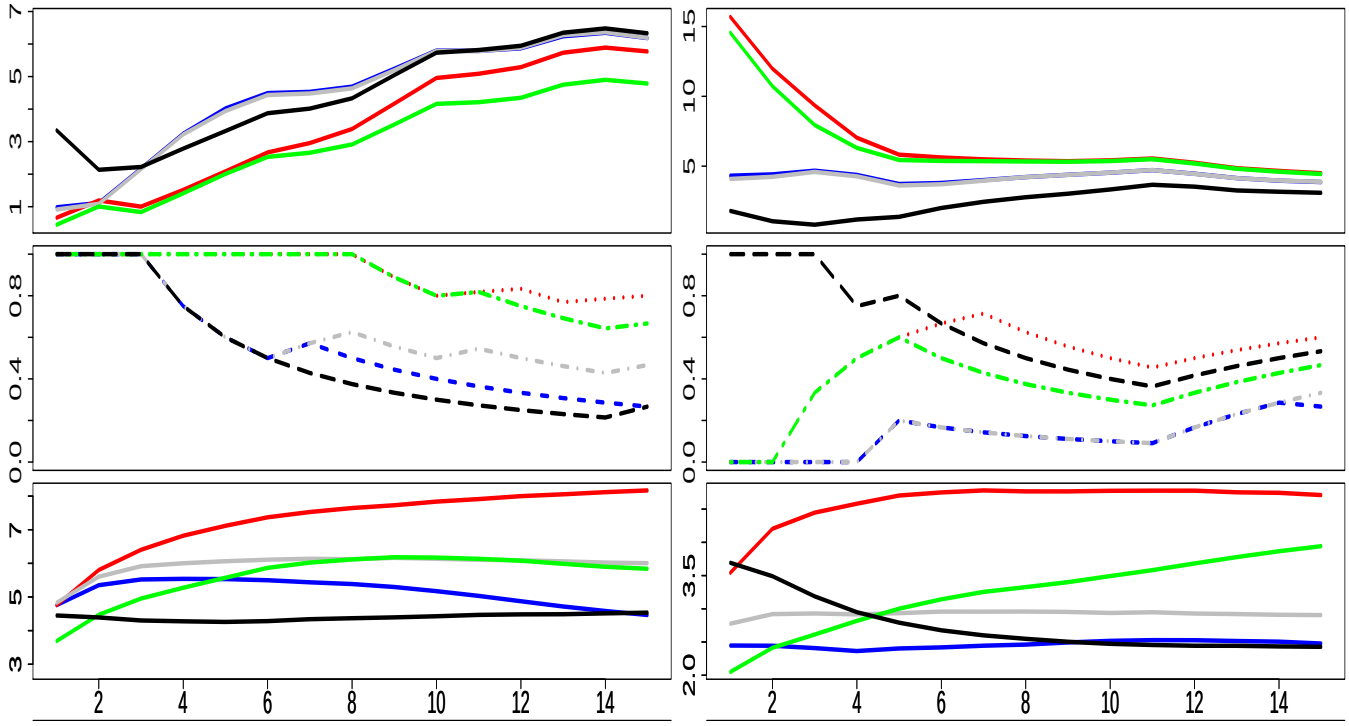
(b) Kaohsiung (2007-2008)



(c) New York (2006-2007)

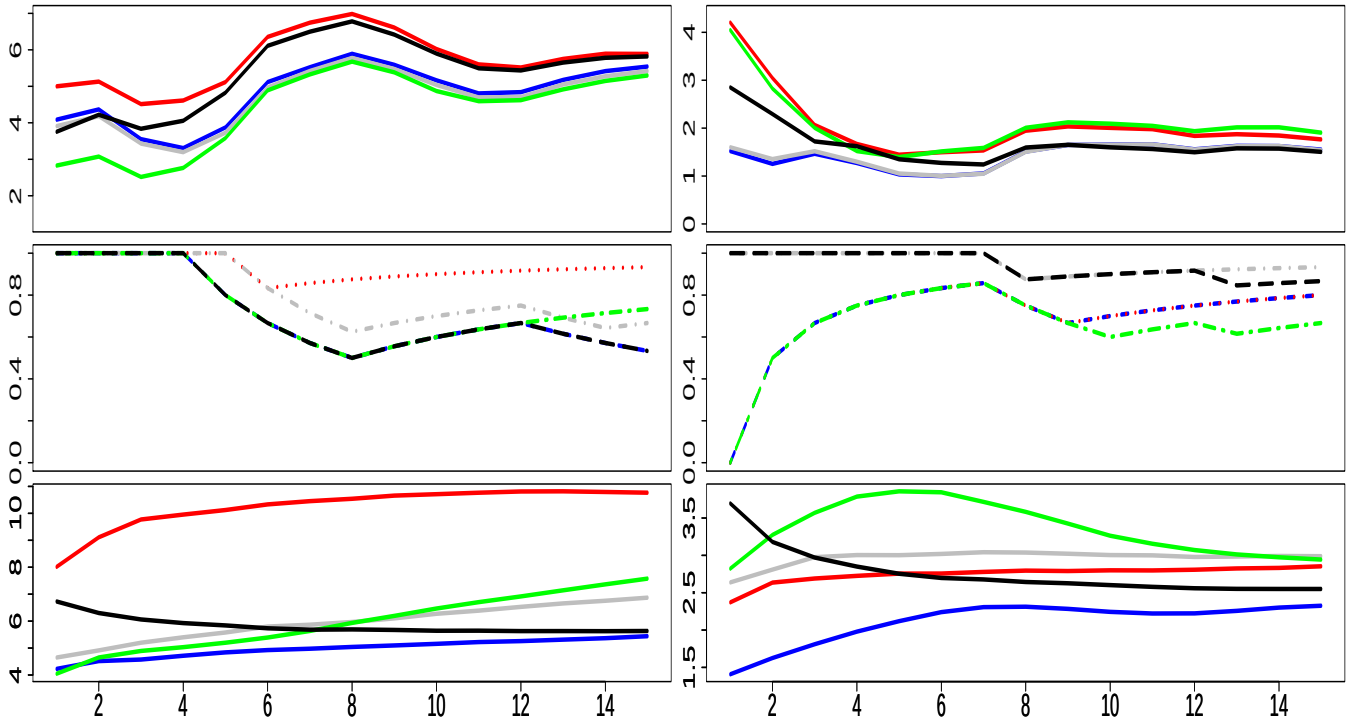
(d) Tokyo (2007-2008)

Figure 12: $h = 1, \ldots, 15$ days (X axis) ahead forecast for Berlin, Kaohsiung, New York and Tokyo (left to right, top to bottom); averaged absolute error (Y axis, upper panel), averaged coverage days (Y axis, middle panel), averaged width of the confidence 95% intervals (Y axis, lower panel), SeMe adMe adVo (blue), Locave (grey), JoMe adMe adVo (green), truncated Fourier (red), CD (black), fitted using 2 years of historical data and 1000 samples.



(a) Berlin (2005-2007)

(b) Kaohsiung (2006-2008)



(c) New York (2005-2007)

(d) Tokyo (2006-2008)

Figure 13: $h = 1, \ldots, 15$ days (X axis) ahead forecast for Berlin, Kaohsiung, New York and Tokyo (left to right, top to bottom); averaged absolute error (Y axis, upper panel), averaged coverage days (Y axis, middle panel), averaged width of the confidence 95% intervals (Y axis, lower panel), SeMe adMe adVo (blue), Locave (grey), JoMe adMe adVo (green), truncated Fourier (red), CD (black), fitted using 3 years of historical data and 1000 samples.

(AAT). The CAT index takes the accumulated average temperature over $[\tau_1, \tau_2]$:

$$CAT(\tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} T_u du,$$

where $T_u = (T_{u,max} + T_{u,min})/2$ denotes the daily average temperature. The measurement period is usually defined in months or season. The HDD index measures the cumulative amount of average temperature below a threshold (typically 18°C or 65°F) over a period $[\tau_1, \tau_2]$: $\int_{\tau_1}^{\tau_2} \max(c - T_u, 0) du$. Similarly, the CDD index accumulates $\max(T_u - c, 0)$. At CME, CAT/CDD futures are traded for European cities, CDD/HDD for US, Canadian, and Australian cities, and AAT for Japanese cities. Note that these temperature indices are the underlying and not the temperature itself. The options at CME are cash settled, i.e., the owner of a future receives 20 times the Degree Day Index at the end of the measurement period, in return for a fixed price. At time t, CME trades different contracts $i = 1, \ldots, N$ with measurement period $0 \leq t \leq \tau_1^i < \tau_2^i$ (usually the length between τ_1^i and τ_2^i is one month). For example, a contract with $i = 7$ is six months ahead from the trading day t. For the US and Europe CAT/CDD/HDD futures, N is usually equal to 7 (April–November or November–April), while for Asia, $N = 12$ (January–December).

Although the temperature data is usually given in a discrete scale, temperature itself develops continuously over time. Thus, a continuous model for the futures price dynamics can be clearly formulated. We propose, as also suggested in Benth et al. (2007) and Härdle and López Cabrera (2011), a mean reverted Ornstein-Uhlenbeck process for the modeling of detrended temperature variations in continuous time $CAR(L)$:

$$d\mathbf{X}_t = \mathbf{A}\mathbf{X}_t dt + \mathbf{e}_L \sigma_t dB_t, \tag{20}$$

where $\sigma_t^2 > 0$ is a bounded deterministic seasonal variation, $\mathbf{X}_t \in \mathbb{R}^L$ for $L \geq 1$ denotes a vectorial Ornstein-Uhlenbeck process, $\mathbf{e}_k$ a kth unit vector in $\mathbb{R}^L$ for $k = 1, \ldots, L$, B_t a Brownian motion, and an $L \times L$-matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & & \ddots & & 0 & \vdots \\ 0 & \ldots & \ldots & 0 & 0 & 1 \\ -\alpha_L & -\alpha_{L-1} & \ldots & & -\alpha_2 & -\alpha_1 \end{pmatrix}.$$

In this framework, the autoregressive process $AR(L)$ in (5) can be seen as a discretely sampled continuous time process ($CAR(L)$) (20) driven by one dimensional Brownian motion. The continuous time process (20) is Markov and allows therefore standard applications of pricing tools. The last three columns of Table 3 display the $CAR(3)$ parameters for the temperature data. The fact that temperature's random factor is close to the normal distribution, as disclosed in the analysis of the residuals before, motivates the use of a Brownian motion as the noise in the Ornstein-Uhlenbeck process. This suggests us that the non-Gaussian shocks found in the literature are the result of the model mis-specification. The continuous analogue of the CD model is however difficult to estimate. Thus the model in (20) is simpler than CD's one and provides better fit to the data.

The temperature futures price is the risk adjusted index, given today's filtration $\mathcal{F}_t$

$$F_I(t, \tau_1, \tau_2) = \mathsf{E}^Q \left[I(\tau_1, \tau_2) | \mathcal{F}_t \right], \tag{21}$$

with $I(\tau_1, \tau_2)$ being one of the indices CAT, HDD or CDD. The expectation is computed under a risk neutral pricing probability Q and is equivalent to the physical measure P under which the discounted temperature index is a Q-martingale. To evaluate (21), we need to know the temperature index dynamics under Q. We restrict the class of pricing probabilities to those that can be parametrized via $Q = Q_\lambda$, where equivalent changes of measures are simply associated with changes of drift. Thus, in the modeling of the dynamics of futures prices written on temperature indices, it is natural to define a parameter measuring the market price of risk (MPR) λ_t, which can be calibrated from traded (CAT/CDD/HDD) derivative type contracts. The temperature dynamics in (20) under Q_λ become:

$$d\mathbf{X}_t = (\mathbf{A}\mathbf{X}_t + \mathbf{e}_L \sigma_t \lambda_t)dt + \mathbf{e}_L \sigma_t dB_t^\lambda, \tag{22}$$

where B_t^λ is a Brownian motion for any time before the end of the trading time and a martingale under Q_λ. Then, for $0 \leq t \leq \tau_1 < \tau_2$, the explicit form of an CAT futures price is given by:

$$\begin{aligned} F_{CAT}(t, \tau_1, \tau_2, \Lambda_t, \sigma_t, \lambda_t) &= \mathsf{E}^{Q_\lambda}\left[\int_{\tau_1}^{\tau_2} T_u du | \mathcal{F}_t\right] \\ &= \int_{\tau_1}^{\tau_2} \Lambda_u du + \mathbf{a}_{t,\tau_1,\tau_2}\mathbf{X}_t + \int_t^{\tau_1} \lambda_u \sigma_u \mathbf{a}_{t,\tau_1,\tau_2}\mathbf{e}_L du \\ &+ \int_{\tau_1}^{\tau_2} \lambda_u \sigma_u \mathbf{e}_1^\top \mathbf{A}^{-1}\left[\exp\left\{\mathbf{A}(\tau_2 - u)\right\} - I_L\right]\mathbf{e}_L du, \end{aligned} \tag{23}$$

with $\mathbf{a}_{t,\tau_1,\tau_2} = \mathbf{e}_1^\top \mathbf{A}^{-1}\left[\exp\left\{\mathbf{A}(\tau_2 - t)\right\} - \exp\left\{\mathbf{A}(\tau_1 - t)\right\}\right]$ and I_L the $L \times L$ identity matrix. Similarly one can compute the price dynamics of CDD and HDD, see Benth et al. (2007). The CAR model (20) provides the analytical formula (23). Note that all constituents except λ_t in the left and right side of (23) are known or estimable (Λ_t and σ_t are estimated as in the previous section), hence the calibration of the MPR from market data turns out to be an inverse problem in terms of λ_t.

Assuming that the parametrization of the MPR is of a constant form for each observed contract ($\lambda_u = \lambda_{t,\tau_1^i,\tau_2^i}$ in (23) for $u \in [\tau_1, \tau_2]$), one can calibrate the MPR for every combination of (t, τ_1^i, τ_2^i), $i = 1, \ldots, N$ contracts, by inverting the pricing formulae in (23) with the observed CME market prices at time t, $(F_{t,i,CME})$ with respect to λ as:

$$\hat{\lambda}_{t,\tau_1^i,\tau_2^i} = \arg\min_\lambda |F_{CAT}(t, \tau_1^i, \tau_2^i, \hat{\Lambda}_t, \hat{\sigma}_t, \lambda) - F_{t,i,CME}|. \tag{24}$$

We name $\hat{\lambda}_{t,\tau_1^i,\tau_2^i}$ as implied MPR. For fixed time t, assuming that λ_t remains the same for different contracts with different maturities, one can also obtain $\hat{\lambda}_{t,OLS}$ for all contracts $i = 1, \ldots, N$:

$$\hat{\lambda}_{t,\tau_1^i,\tau_2^i,OLS} = \hat{\lambda}_{t,OLS} = \arg\min_\lambda \sum_{j=1}^{N} \{F_{CAT}(t, \tau_1^j, \tau_2^j, \hat{\Lambda}_t, \hat{\sigma}_t, \lambda) - F_{t,j,CME}\}^2. \tag{25}$$

Moreover, to evaluate the estimation of $\hat{\lambda}_t$ for a particular contract i, the observed price $F_{t,i,CME}$ for this contract can be excluded for the estimation. We have then the cross validated estimation:

$$\hat{\lambda}_{t,\tau_1^i,\tau_2^i,CV} = \arg\min_\lambda \sum_{j=1;j\neq i}^{N} \{F_{CAT}(t, \tau_1^j, \tau_2^j, \hat{\Lambda}_t, \hat{\sigma}_t, \lambda) - F_{t,j,CME}\}^2. \tag{26}$$

Other specifications of the MPR for temperature derivatives have been explored in Härdle and López Cabrera (2011), where the authors argue that a constant MPR is sufficient for pricing purposes. This might be compared with complete markets, where the MPR is minus the Sharp ratio $(\mu_t - r)/\sigma_t^F$, where μ_t and σ_t^F denote the mean and standard deviation of traded futures, and r is the risk free interest rate. From now on, pricing follows (23) with an MPR from (24), (25) or (26) and with Λ_t and σ_t estimated via the localisation techniques.

Observe that calibration (24), (25) or (26) is only valid if there exists a weather derivative market, like e.g. for Berlin and Tokyo. In order to price temperature derivatives for regions with no weather derivative markets, like Kaohsiung, one can use the implied MPR of traded futures of a neighbour market, e.g. Tokyo AAT futures. Thus, by finding a relationship between the MPR and the seasonal variance one can use this as a proxy to price over the counter (OTC) AAT futures for Kaohsiung. This is acceptable since stylized facts of temperature in Tokyo reveal similarity to that in Kaohsiung. But generally, we are aware of arbitrage opportunities across the two different markets, therefore this approach cannot be generalized for every two weather derivative markets. Considering that the MPR is a risk premium per unit of volatility, one can project the implied MPR on the state variables related to volatility. An insight for Tokyo AAT futures, which can be employed for the Kaohsiung case, is by regressing the averaged implied MPR (24) against the variation:

$$\hat{\lambda}_{\tau_1^i,\tau_2^i} = 4.08 - 2.19\hat{\sigma}^2_{\tau_1^i,\tau_2^i} + 0.28\hat{\sigma}^4_{\tau_1^i,\tau_2^i},$$

where $\hat{\lambda}_{\tau_1^i,\tau_2^i} \stackrel{\text{def}}{=} (\tau_2^i - \tau_1^i)^{-1} \sum_{t=\tau_1^i}^{\tau_2^i} \hat{\lambda}_{t,\tau_1^i,\tau_2^i}$, $\hat{\sigma}^2_{\tau_1^i,\tau_2^i} \stackrel{\text{def}}{=} (\tau_2^i - \tau_1^i)^{-1} \sum_{t=\tau_1^i}^{\tau_2^i} \hat{\sigma}_t^2$, $\hat{\sigma}^4_{\tau_1^i,\tau_2^i} \stackrel{\text{def}}{=} (\tau_2^i - \tau_1^i)^{-1} \sum_{t=\tau_1^i}^{\tau_2^i} \hat{\sigma}_t^4$ and $R^2_{adj} = 0.71$. Plugging the corresponding $\hat{\sigma}_t^2$, $\hat{\sigma}_t^4$ values for Kaohsiung into this equation let us price such a non CME traded weather derivative via (23).

We compare the prices obtained with localisation procedures ('localised' prices) for Λ_t and σ_t (SeMe adMe adVo (AdaptBW), Locmax) with prices estimated under fixed bandwidth (SeMe fiMe fiVo (FixedBW)) and truncated Fourier series. As a benchmark, we compute the index values from realized temperatures $I(\tau_1, \tau_2)$ and compare them with our model prices and with CME data obtained from Bloomberg for the period 2009-2011. Note that prices from Benth et al. (2007), later denoted as Benth, assumed MPR equal to zero and truncated Fourier series for seasonal mean and variance. These prices are estimated out of sample.

Table 6 shows CME prices, realized indices for Berlin, Essen, London and Tokyo contracts, as well as their differences between CME and model prices. We offer both in sample fitting prices calculated with $\hat{\lambda}_{t,\tau_1^i,\tau_2^i,OLS}$ (25) and cross validated prices $\hat{\lambda}_{t,\tau_1^i,\tau_2^i,CV}$ for fixed time t and contract i (26). When MPR is different from zero ($\lambda_t = \lambda$), our estimated prices both with $\hat{\lambda}_{t,\tau_1^i,\tau_2^i,CV}$ and $\hat{\lambda}_{t,\tau_1^i,\tau_2^i,OLS}$ are close to the CME prices in most of the cases. 'Localised' prices with calibrated nonzero MPR are closer to CME market futures prices and offer an improvement over the fixed bandwidth and the truncated Fourier method. Our calibrated MPR embeds information on the risk and uncertainty in the market, which is helpful in analyzing market risk. Also, as mentioned before, this information may help to price OTC derivatives in the same market.

To judge the performance of the models, we compute the root mean squared errors (RMSE) between the indices of realized temperatures $RT = I(\tau_1, \tau_2)$ (benchmark) and the estimated model

prices $F_I(t, \tau_1^i, \tau_2^i, \hat{\Lambda}_t, \hat{\sigma}_t, \hat{\lambda}_{t,\tau_1^i,\tau_2^i})$ $(i = 1, \ldots, N)$:

$$RMSE(\tau_1^i, \tau_2^i) \quad = \quad \sqrt{|\mathfrak{T}|^{-1} \sum_{t \in \mathfrak{T}} \Big\{ F_I(t, \tau_1^i, \tau_2^i, \hat{\Lambda}_t, \hat{\sigma}_t, 0) - I(\tau_1, \tau_2) \Big\}^2},$$

in Table 7, where $\mathfrak{T}$ is the set of days when the contract i with the measurement period (τ_1^i, τ_2^i) was traded. In addition the RMSE between models' prices and F_{CME} is also presented in Table 7. The results show smaller RMSE when future prices are estimated via localisation techniques, which in general outperforms the prices based on the truncated Fourier series (Benth).

When considering the MPR equal to zero ($\lambda_t = 0$) our model's prices behave similar to the realized temperature indices and thus may be used as the weather forecast. The small difference between 'localised' prices with zero MPR to realized prices, suggests that our model's prices have learned the market condition of past weather surprises.

This brings, of course, investment opportunities: an investor who purchased a futures CAT contract for Berlin on 20070316 with $\tau_1 = 20070401$ and $\tau_2 = 20070430$ would have paid the value of the index 288 points (see CME column and first row in Table 6) or 5,760 EUR (1 index point = 20 EUR per contract). The realized temperature index, however resulted in 362.9 points (see RT column in Table 6), which is equivalent to 7,258 EUR. If the contract holder holds the contract until expiration, he would receive a payoff of 1,498 EUR (7,258-5,760 EUR). Taking one of the models from Table 6 as personal forecast, an investor could be tempted to act according to the difference between the CME price and his forecasted price. When the difference is positive, the strategy would be to go short on the future, and go long for a negative difference. Thus, if the personal forecast is based on the AdaptBW with $\lambda_t = 0$, then investor will go long, and had a profit of 1,498 EUR, otherwise, using Benth method as a personal forecast, will lead to a short position, thus loosing 1,498 EUR. More formally, the unit profit PR from the contract can be calculated as

$$PR = \text{sign}\{F_{t,i,CME} - F_{CAT}(t, \tau_1^i, \tau_2^i, \hat{\Lambda}_t, \hat{\sigma}_t, 0)\} \cdot |F_{t,i,CME} - I(\tau_1^i, \tau_2^i)|. \tag{27}$$

In a more general example, let us consider an investor, who invests in all the 39 contracts (one unit per contract) with $\lambda_t = 0$ from the Table 6. Without transaction costs and interest rate, the total profit will be a summation of the profits from each of the 39 contracts under different model specifications. Total profit for AdaptBW, FixedBW or Locmax models results in 11,836 EUR, while using the Benth model only into 10,328 EUR.

These results provide important insights on how weather is priced at CME and how the observed prices conform with the stylized facts (seasonal effect, inter-temporal correlation, etc.) of weather data. The meaning of the terms in the CAT futures price confirm this statement. To illustrate this point, consider, for example, the purchase of a May CAT contract for Berlin on 20070427, which starts measurement at time $\tau_1 = 20070501$ and finished at $\tau_2 = 20070531$. Setting a constant MPR (for example $\lambda = 0.20$), the first term of (23) is equal to 431.060, the second, third and fourth terms lead to 11.531, 0.8690 and 13.5390 respectively. The seasonal effect Λ_t plays an important role in the level of the futures price, as it explains almost 94% of the market price which is 457. However, as we get closer to the measurement period, temperature variations have more impact on the price, since the last two terms with the MPR contribute to the CAT futures price more.

					Differences between F_{CME} and models' prices												
					AdaptBW			FixedBW			Locmax			Fourier			RT
Type	t	MP	CME	RT	$\lambda_t = 0$	$\hat{\lambda}_{t,OLS}$	$\hat{\lambda}_{t,CV}$	$\lambda_t = 0$	$\hat{\lambda}_{t,OLS}$	$\hat{\lambda}_{t,CV}$	$\lambda_t = 0$	$\hat{\lambda}_{t,OLS}$	$\hat{\lambda}_{t,CV}$	Benth	$\hat{\lambda}_{t,OLS}$	$\hat{\lambda}_{t,CV}$	$I(\tau_1, \tau_2)$
Berlin-CAT	20070316	200704	288.0	362.9	-15.328	-4.051	-0.011	-10.409	-0.265	0.001	-15.328	-7.284	-0.011	21.613	27.648	-3.117	-74.9
Berlin-CAT	20070316	200705	457.0	494.2	25.775	37.551	41.771	19.223	29.461	29.730	25.775	35.873	45.004	1.686	7.814	-23.422	-37.2
Berlin-CAT	20070316	200706	529.0	574.3	-17.359	-5.254	-0.917	-13.160	-1.776	-1.477	-17.359	-7.584	1.255	-37.066	-29.770	-66.958	-45.3
Berlin-CAT	20070316	200707	616.0	583.0	-51.305	-35.800	-30.244	-59.036	-45.561	-45.208	-51.305	-40.677	-31.068	-16.706	-10.136	-43.625	33.0
Berlin-CAT	20070316	200701	610.0	580.7	48.743	57.361	60.449	63.309	72.497	72.738	48.743	56.079	62.712	12.472	18.065	-10.446	29.3
Berlin-CAT	20070316	200709	472.0	414.8	-38.109	-28.589	-25.178	-47.924	-39.559	-39.340	-38.109	-31.484	-25.493	-12.911	-9.564	-26.626	57.2
Berlin-CAT	20070427	200705	457.0	494.2	14.253	27.861	31.275	18.924	29.203	29.553	14.253	25.875	34.487	-72.412	-41.125	-100.444	-37.2
Berlin-CAT	20070427	200706	529.0	574.3	-17.359	-4.215	-0.917	-13.160	-1.862	-1.477	-17.359	-6.668	1.255	-37.088	-3.717	-66.988	-45.3
Berlin-CAT	20070427	200707	616.0	583.0	-51.305	-34.469	-30.244	-59.036	-45.663	-45.208	-51.305	-39.681	-31.068	-16.706	13.338	-43.625	33.0
Berlin-CAT	20070427	200708	610.0	580.7	48.743	58.100	60.449	63.309	72.427	72.738	48.743	56.766	62.712	12.472	38.051	-10.446	29.3
Berlin-CAT	20070427	200709	472.0	414.8	-38.109	-27.772	-25.178	-47.924	-39.623	-39.340	-38.109	-30.863	-25.493	-12.911	2.396	-26.626	57.2
Berlin-CAT	20070524	200706	529.0	574.3	-16.741	3.854	-0.349	-13.103	5.546	-1.449	-16.741	2.495	1.814	-50.548	-24.838	-82.870	-45.3
Berlin-CAT	20070524	200707	616.0	583.0	-51.305	-24.844	-30.244	-59.036	-36.908	-45.208	-51.305	-30.326	-31.068	-16.711	4.704	-43.634	33.0
Berlin-CAT	20070524	200708	610.0	580.7	48.743	63.451	60.449	63.309	78.397	72.738	48.743	63.224	62.712	12.472	30.702	-10.446	29.3
Berlin-CAT	20070524	200709	472.0	414.8	-38.109	-21.862	-25.178	-47.924	-34.188	-39.340	-38.109	-25.030	-25.493	-12.911	-2.002	-26.626	57.2
Essen-CAT	20080430	200805	435.0	501.9	-22.790	9.350	12.521	-23.323	8.865	11.834	-22.790	9.350	12.521	-4.791	22.456	31.212	-66.9
Essen-CAT	20080430	200806	499.0	516.3	-43.409	-14.285	9.407	-43.326	-14.128	9.382	-43.409	-14.285	9.407	-43.315	-19.197	9.142	-17.3
Essen-CAT	20080430	200807	570.0	571.9	-73.972	-48.187	-31.891	-73.886	-48.041	-31.908	-73.972	-48.187	-31.891	-73.886	-51.744	-30.686	-1.9
Essen-CAT	20080430	200808	564.0	566.3	34.435	56.514	65.912	34.522	56.665	65.935	34.435	56.514	65.912	34.522	52.931	65.892	-2.3
Essen-CAT	20080430	200809	469.0	410.7	-16.797	5.050	28.439	-16.714	5.201	28.474	-16.797	5.050	28.439	-16.714	1.251	27.903	58.3
Essen-CAT	20080527	200806	510.0	516.3	-31.879	8.301	0.119	-32.004	8.135	-0.113	-31.879	8.301	0.119	-63.912	-8.855	-28.378	-6.3
Essen-CAT	20080527	200807	556.0	571.9	-87.972	-52.399	-41.211	-87.886	-52.357	-41.230	-87.972	-52.399	-41.211	-87.901	-42.550	-39.723	-15.9
Essen-CAT	20080527	200808	550.0	566.3	20.435	50.895	56.467	20.522	50.963	56.488	20.435	50.895	56.467	20.522	58.212	56.438	-16.3
Essen-CAT	20080527	200809	455.0	410.7	-30.797	-0.656	0.350	-30.714	-0.587	0.376	-30.797	-0.656	0.350	-30.714	6.068	-0.100	44.3
London-CAT	20071231	200804	228.0	285.0	-98.210	-91.116	-106.889	-98.473	-91.112	-106.913	-98.473	-91.112	-106.913	-98.473	-91.112	-106.913	-57.0
London-CAT	20071231	200805	430.0	485.0	15.849	22.936	26.897	15.577	22.929	26.895	15.577	22.929	26.895	15.577	22.929	26.895	-55.0
London-CAT	20071231	200806	507.0	497.0	-3.096	5.450	6.935	-3.359	5.497	6.993	-3.359	5.497	6.993	-3.359	5.497	6.993	10.0
London-CAT	20071231	200807	579.0	577.0	-9.456	-2.048	-2.440	-9.728	-2.049	-2.442	-9.728	-2.049	-2.442	-9.728	-2.049	-2.442	2.0
London-CAT	20071231	200808	579.0	552.0	18.967	26.508	31.812	18.695	26.515	31.824	18.695	26.515	31.824	18.695	26.515	31.824	27.0
London-CAT	20071231	200809	495.0	443.0	26.379	33.841	40.446	26.115	33.849	40.448	26.115	33.849	40.448	26.115	33.849	40.448	52.0
London-CAT	20080502	200806	515.0	495.0	9.495	-1.807	-1.205	9.232	-1.769	-1.160	9.232	-1.769	-1.160	9.232	-1.769	-1.160	20.0
London-CAT	20080502	200807	579.0	573.0	-10.742	-20.242	0.718	-11.014	-20.263	0.716	-11.014	-20.263	0.716	-11.014	-20.263	0.716	6.0
London-CAT	20080502	200801	579.0	557.0	17.119	7.556	25.804	16.847	7.532	25.817	16.847	7.532	25.817	16.847	7.532	25.817	22.0
London-CAT	20080502	200809	495.0	448.0	24.128	14.313	31.319	23.864	14.311	31.332	23.864	14.311	31.332	23.864	14.311	31.332	47.0
Tokyo-AAT	20081027	200903	450.0	305.0	112.508	4.109	4.922	112.439	4.014	4.810	112.508	4.109	4.922	112.439	4.300	5.149	145.0
Tokyo-AAT	20081027	200904	592.0	479.0	158.346	21.321	28.967	158.279	21.470	29.149	158.346	21.321	28.967	158.279	20.680	28.202	113.0
Tokyo-AAT	20081027	200905	682.0	623.0	111.742	-39.678	-58.553	111.673	-39.762	-58.714	111.742	-39.678	-58.553	111.673	-38.936	-57.221	59.0
Tokyo-AAT	20081027	200906	818.0	679.0	147.010	42.284	49.993	146.944	42.230	49.937	147.010	42.284	49.993	146.944	43.283	51.004	139.0
Tokyo-AAT	20081027	200907	855.0	812.0	60.073	-21.821	-24.093	60.004	-21.702	-23.953	60.073	-21.821	-24.093	60.004	-23.227	-25.738	43.0

Table 6: Differences between weather futures listed on date yyyymmdd at CME (Source: Bloomberg) and $F_I(t, \tau_1^i, \tau_2^i, \hat{\Lambda}_t, \hat{\sigma}_t, \lambda_t)$ with τ_1 and τ_2 being the first and the last day of the measurement period (MP, yyyymm) respectively. The prices are estimated with MPR equal to zero, $\hat{\lambda}_{t,OLS}$ and $\hat{\lambda}_{t,CV}$ under different estimations schemes ($\hat{\Lambda}_t, \hat{\sigma}_t$ under AdaptBW, FixedBW, Locmax and truncated Fourier). Last column provides comparison of F_{CME} against realised the temperature indices $I(\tau_1, \tau_2)$. Note, that Benth prices are estimated with truncated Fourier series and MPR equal to zero.

			RMSE	RMSE between models' prices and $I(\tau_1,\tau_2)$				RMSE between models' prices and F_{CME}							
				$\lambda_t = 0$				$\hat{\lambda}_{t,CV}$				$\hat{\lambda}_{t,OLS}$			
Type	MP	n	$F_{CME} - I(\tau_1,\tau_2)$	AdaptBW	FixedBW	Locmax	Benth	AdaptBW	FixedBW	Locmax	Fourier	AdaptBW	FixedBW	Locmax	Fourier
Berlin-CAT	200704	230	74.702	59.439	64.473	59.439	98.241	2.868	2.617	2.876	9.665	7.498	7.813	7.679	12.545
Berlin-CAT	200705	38	59.572	62.689	56.455	62.689	36.593	49.891	43.141	52.061	55.290	39.959	34.435	36.941	28.069
Berlin-CAT	200706	58	44.613	27.599	31.986	27.599	9.925	2.033	3.078	2.662	68.263	5.323	4.891	6.844	22.485
Berlin-CAT	200707	79	33.068	83.569	91.858	83.569	48.913	31.650	46.607	32.437	44.850	33.718	43.794	38.568	15.367
Berlin-CAT	200708	101	29.431	21.110	34.623	21.110	19.834	61.730	73.404	63.982	15.451	57.530	71.086	55.971	21.654
Berlin-CAT	200709	122	59.131	94.761	104.943	94.761	69.139	27.161	40.820	27.443	28.175	28.310	38.435	30.992	15.312
Berlin-CAT	200710	3	190.026	103.624	105.613	103.624	99.592	101.597	96.384	101.528	68.094	49.376	44.624	49.348	53.914
Essen-CAT	200704	230	119.610	57.976	58.413	57.976	58.125	3.739	3.738	3.739	4.197	9.546	9.503	9.546	10.277
Essen-CAT	200705	39	57.295	8.755	9.229	8.755	14.271	38.178	38.030	38.178	39.279	18.477	18.528	18.477	20.246
Essen-CAT	200706	59	53.700	3.845	4.044	3.845	6.643	9.752	9.104	9.752	13.284	21.311	21.575	21.311	21.125
Essen-CAT	200707	80	18.541	76.146	75.916	76.146	74.583	13.416	13.557	13.416	16.541	42.470	42.753	42.470	40.562
Essen-CAT	200708	102	40.444	6.603	6.162	6.603	13.987	90.912	90.528	90.912	96.121	56.878	56.903	56.878	57.785
Essen-CAT	200709	123	53.510	60.560	60.248	60.560	59.833	36.938	36.759	36.938	39.401	17.050	16.869	17.050	19.580
Essen-CAT	200710	3	150.938	62.750	62.324	62.750	62.308	151.181	151.345	151.181	151.117	39.922	39.433	39.922	42.142
Essen-CAT	200804	78	36.372	72.459	72.349	72.459	70.492	75.412	75.393	75.412	74.163	59.183	59.155	59.183	58.900
Essen-CAT	200805	100	72.379	43.211	43.285	43.211	44.080	21.871	21.845	21.871	22.628	17.084	17.061	17.084	17.200
Essen-CAT	200806	121	18.211	28.701	28.562	28.701	33.612	7.748	7.690	7.748	15.775	6.459	6.465	6.459	10.143
Essen-CAT	200807	142	8.230	71.045	70.963	71.045	70.967	40.302	40.319	40.302	38.949	39.423	39.455	39.423	37.940
Essen-CAT	200808	164	9.764	35.866	35.978	35.866	36.890	70.605	70.647	70.605	71.147	61.635	61.651	61.635	61.805
Essen-CAT	200809	185	55.805	76.590	76.503	76.590	76.913	19.490	19.513	19.490	19.307	15.855	15.889	15.855	16.109
Essen-CAT	200810	1	132.700	54.954	54.868	54.954	54.868	134.935	134.970	134.935	132.858	4.841	4.866	4.841	4.202
London-CAT	200804	78	57.000	40.572	40.842	40.842	40.842	106.252	106.281	106.281	106.281	90.929	90.933	90.933	90.933
London-CAT	200805	100	50.795	71.228	70.951	70.951	70.951	27.897	27.883	27.883	27.883	23.365	23.347	23.347	23.347
London-CAT	200806	121	15.123	11.981	12.238	12.238	12.238	12.646	12.688	12.688	12.688	6.650	6.668	6.668	6.668
London-CAT	200807	142	4.788	14.684	14.952	14.952	14.952	7.633	7.643	7.643	7.643	14.730	14.743	14.743	14.743
London-CAT	200808	164	24.536	6.825	7.051	7.051	7.051	27.422	27.429	27.429	27.429	18.123	18.122	18.122	18.122
London-CAT	200809	185	49.616	23.953	24.209	24.209	24.209	36.187	36.196	36.196	36.196	22.664	22.669	22.669	22.669
London-CAT	200904	18	223.919	31.783	31.815	31.815	31.815	337.930	348.371	348.371	348.371	60.637	66.326	66.326	66.326
London-CAT	200905	18	276.542	36.565	36.598	36.598	36.598	346.610	353.159	353.159	353.159	16.250	18.847	18.847	18.847
London-CAT	200906	18	317.106	21.817	21.850	21.850	21.850	434.103	424.054	424.054	424.054	30.357	24.764	24.764	24.764
London-CAT	200907	18	321.539	26.317	26.283	26.283	26.283	345.005	342.891	342.891	342.891	53.098	54.898	54.898	54.898
London-CAT	200908	18	345.055	30.484	30.518	30.518	30.518	348.115	351.479	351.479	351.479	53.659	50.813	50.813	50.813
London-CAT	200909	18	297.210	16.995	17.028	17.028	17.028	364.723	357.984	357.984	357.984	6.426	4.187	4.187	4.187
Tokyo-AAT	200903	18	145.000	32.492	32.561	32.492	32.561	4.922	4.810	4.922	5.149	4.109	4.014	4.109	4.300
Tokyo-AAT	200904	18	113.000	45.346	45.279	45.346	45.279	28.967	29.149	28.967	28.202	21.321	21.470	21.321	20.680
Tokyo-AAT	200905	18	59.000	52.742	52.673	52.742	52.673	58.553	58.714	58.553	57.221	39.678	39.762	39.678	38.936
Tokyo-AAT	200906	18	139.000	8.010	7.944	8.010	7.944	49.993	49.937	49.993	51.004	42.284	42.230	42.284	43.283
Tokyo-AAT	200907	18	43.000	17.073	17.004	17.073	17.004	24.093	23.953	24.093	25.738	21.821	21.702	21.821	23.227

Table 7: RMSE between the weather futures listed at CME and realised temperature; estimated weather futures $F_I(t,\tau_1^i,\tau_2^i,\hat{\Lambda}_t,\hat{\sigma}_t,0)$ and the realized temperature; estimated weather futures $F_I(t,\tau_1^i,\tau_2^i,\hat{\Lambda}_t,\hat{\sigma}_t,\hat{\lambda}_t)$ with $\hat{\lambda}_t = \hat{\lambda}_{t,OLS}$ and $\hat{\lambda}_t = \hat{\lambda}_{t,CV}$ and F_{CME}. τ_1^i and τ_2^i are the first and the last day of the measurement period (MP, yyyymm) respectively. Prices are estimated under different estimations schemes ($\hat{\Lambda}_t, \hat{\sigma}_t$ under AdaptBW, FixedBW, Locmax and truncated Fourier). n corresponds to the number of trading days for a given MP. Note that Benth prices are estimated with MPR equal to zero and truncated Fourier series.

6 Conclusions

We showed that temperature risk stochastics are closer to Gaussian when applying adaptive statistical methods for seasonal mean and seasonal variance. This suggest us that the non-Gaussian shocks found in the literature are truly a result of mis-specification. We found that the localisation method performs well, and it is robust to the specification given for Λ_t or σ_t. The proposed adaptive technique shows good performance over a calibration window of two years, for a longer calibration interval of three years. We also observed in most of the cases, that the proposed method outperforms the standard estimation methods. Finally, our results provide important insights on how weather is priced at CME and how the observed prices conform with the stylized facts of weather data.

References

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

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

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

Chen, Y., Härdle, W. K. and Pigorsch, U. (2010). Localized realized volatility modelling, *Journal of the American Statistical Association* **105(492)**: 1376–1393.

Cízek, P., Härdle, W. K. and Spokoiny, V. (2009). Adaptive pointwise estimation in time-inhomogeneous conditional heteroschedasticity models, *The Econometrics Journal* **12**: 248–271.

Davis, M. (2001). Pricing weather derivatives by marginal value, *Quantitative Finance* **1**: 305–308.

Diebold, F. X. and Inoue, A. (2001). Long memory and regime switching, *Journal of Econometrics* **105**: 131–159.

Girsanov, I. V. (1960). On transforming a certain class of stochastic processes by absolutely continuous substitution of measures, *Theory of Probability and its Applications* **5**(3): 285–301.

Granger, C. W. J. and Hyung, N. (2004). Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns, *Journal of Empirical Finance* **11**: 399–421.

Härdle, W. K. and López Cabrera, B. (2011). Infering the market price of weather risk, *Applied Mathematical Finance* **5**: 1–37.

Horst, U. and Mueller, M. (2007). On the spanning property of risk bonds priced by equilibrium, *Mathematics of Operation Research* **32(4)**: 784–807.

Karatzas, I. and Shreve, S. (2001). *Methods of Mathematical Finance*, Springer Verlag, New York.

Mercurio, D. and Spokoiny, V. (2004). Statistical inference for time-inhomogeneous volatility models, *The Annals of Statistics* **32(2)**: 577–602.

Polzehl, J. and Spokoiny, V. (2006). Propagation-separation approach for local likelihood estimation, *Probability Theory and Realated Fields* **135**: 335–362.

Spokoiny, V. (2009). Multiscale local change point detection with applications to value at risk, *The Annals of Statistics* **37**(3): 1405–1436.