Summary

This paper contributes to the existing literature on the topic of trend estimation in temperature series by applying a non-standard estimation procedure to data from cities all over Europe. It seems that statements about the important topic of global warming are oftentimes based on linear trend estimation, where a straight line is fitted to the temperature data. It is argued in this paper that this approach of linear trend estimation is too restrictive in the case of temperature data and that a different approach, allowing for a more flexible form of the trend, yields more accurate results. For instance, regressing the European temperature data on a linear trend provides evidence of no global warming for all series under consideration, whereas the non-parametric approach advocated in this paper shows a recent upward trend for all series underlining the importance of taking the complexity of temperature data into account.

1.1 Introduction

A change in the Earth's climate can be caused by various different forces, which can be divided into two broad classes according to their origin. Natural and so-called anthropogenic factors are considered by the World Meterological Organization (WMO) to lead to climate change. In particular, the latter class receives attention from scientists as well as politicians all over the world, since this is the category consisting of human-induced forces. Among these anthropogenic factors of climate change are greenhouse gas emissions, atmospheric aerosols and a change in land use - like deforestation [WMO (2012)].

A rapid climate change that cannot be explained by natural causes only is often referred to as global warming [WMO (2012)]. The term climate change is officially defined by the Intergovernmental Panel on Climate Change (IPCC) as a change in the mean or variability of climate properties that last over decades or longer and are caused by natural processes or by external forces [IPCC (2012)]. Climate change can lead to climate extremes, which occur when a climate variable lies above or below a certain threshold, determined by observed values for this variable. This, in turn, can lead to a climate disaster, which is a physical event that causes severe changes in the functioning of a community or society and requires emergency help [IPCC (2012)]. This is a formal definition of what can be seen in real life as earthquakes, heat waves, hurricanes, tsunamis and other extreme weather events.

The IPCC is one of the leading organizations in the analysis and assessment of climate change founded by the United Nations Environment Programme (UNEP) and the WMO. Its main concern is to provide important information to policy makers from many countries in order to help them base their decisions concerning the environment on scientific grounds. In 2007, the IPCC published their most recent assessment report, which contained estimates on trends in air temperature. If a significant, positive trend is found in temperature data, empirical evidence of the presence of global warming is created. The probably most frequently cited estimate taken from this report is that "global mean surface temperatures have risen by $0.74°C \pm 0.18°C$ when estimated by a linear trend over the last 100 years (19062005)" ([IPCC (2007)], p.237). The report further states that global warming is now unequivocal. The method, on which these statements are based and which is mentioned in the quote, is linear trend estimation. Linear trend estimation assumes that a rise in temperature, if any, takes place at a constant pace over the whole period under consideration. A straight line is fit to temperature data and the slope of this line can be seen as the speed of global warming. However, as temperature data usually are more complex, the trend might be more complex and can take on a non-linear shape.

In times of more frequently occuring climate disasters, the topic of trend estimation in temperature data is highly sensitive and the form of the trend should be carefully estimated. Linear trend estimation - as applied by the IPCC - restricts the trend to have a specific, pre-defined form. Other research has adapted to the complexity of temperature data by allowing the trend to have other forms, for example, a quadratic trend. A way of estimating trends with even more flexibility is non-parametric estimation, where no specific form has to be agreed upon in advance. Although it is more advanced to obtain confidence bands for this method, [Bühlmann (1998)] has shown that a sieve bootstrap method can be applied in this context and that it provides valid confidence intervals.

It is the purpose of this paper to show that the flexibility of non-parametric estimation is needed in the context of temperature data. It adds to the existing literature by applying a non-parametric approach in combination with the method proposed by [Bühlmann (1998)] to several temperature series from Europe. Additionally, these temperature series provide a spread over different geographical and climatic areas existing in Europe.

A summary of selected academic literature on the topic of trend estimation is provided in Section 2. In Section 3, the method of non-parametric estimation is introduced and in Section 4, the sieve bootstrap proposed by [Bühlmann (1998)] is presented. The subsequent Section 5 describes the data to which the two methods are applied. Results are given in Section 6.

1.2 Related Literature

In the academic literature, estimation of trends in temperature data has become more and more popular over the last decades. This increase in popularity is in line with the rising importance of the topic of global warming all over the world. Various methods have been proposed, how a trend can be extracted from time series data and how forecasting can be done on the basis of this trend. Particularly, on the one hand, linear trend estimation or more generally, parametric trend estimation, where the form of the trend is pre-specified, has received a lot of attention from scientists. On the other hand, some authors found that parametric estimation has a lot of limitations in climatology and that a non-parametric approach might be more appropriate for trend estimation in this field. First, a few results from research using parametric estimation are presented and after this, the focus will switch from parametric to non-parametric trend estimation for the remainder of this paper.

An important issue that researchers are confronted with when dealing with temperature data is autocorrelation of the error terms, of which a formal definition is given in Section 4. [Fomby and Vogelsang (2002)] address this issue in depth and stress the significance of serial correlation in temperature data. Wrong conclusions would be drawn, if serial correlation was not taken into account. For example, without any measure to account for autocorrelated errors, inference can be misled in a way that global warming is expected to continue in the future, even when it actually does not. The presence of serial correlation makes the use of standard t-tests invalid and nothing can be said about the signifiance of the trend coefficients from a regression. Using a Monte Carlo simulation, [Fomby and Vogelsang (2002)] show that when serial correlation is strong, the t-test, as well as the autocorrelation robust Newey-West estimator, tend to overreject the null hypothesis of no global warming in favor of global warming, even when the data are simulated from a process without a trend. In short, this means such tests would provide spurious evidence of the presence of global warming in case of strongly integrated errors.

To overcome this overrejection problem, [Fomby and Vogelsang (2002)] review a recent test proposed by [Vogelsang (1998)], which is a trend test that is robust to any kind of serial correlation. They apply this test to various temperature series starting in the late 19th century, ending in the late 20th century and they find a rejection of the null hypothesis of no significant trend in all but one series. The conclusion they draw is that there has been a rise in temperature of about 0.5°C over the period mentioned. In this paper, the non-parametric approach to be used deals with serial correlation in a different way, using a bootstrap method to find confidence intervals. This issue is discussed in Section 4.

A different result confirming an increase in temperature, in particular over the 20th century, is provided by [Franses and Vogelsang (2005)]. They use a similar parametric approach of linear trend estimation that avoids the use of t-tests, but uses new critical values accounting for autocorrelation to judge significance. This study analyzes a global temperature series that is divided into 12 annual series - one for each month - showing that there is an overall increase in global temperature, but that winters are getting warmer faster. Similar results also hold for the United Kingdom and the Netherlands [Franses and Vogelsang (2005)].

A final paper to be presented in this section on the topic of parametric trend estimation is a recent paper by [McKitrick and Vogelsang (2011)], which investigates mean shifts in temperature data. When weather stations move to a different location or when there is a change in the thermometers used, breaks can occur and it is important to model such shifts to avoid biased results [McKitrick and Vogelsang (2011)]. The two authors model a structural break at a pre-specified point in time and compare this model to the same model without a break. They specify the break to be the Pacific Climate Shift in 1977, which was a major change in temperature patterns on the surface of the Pacific Ocean. [McKitrick and Vogelsang (2011)] estimate their comparable models using lower- and mid-troposphere data ranging from 1958 to 2010. The results of this study indicate that significant trends become insignificant, once a mean shift is included in the model. To be more specific, in the model *without* a mean shift, the null hypothesis of no global warming is rejected, whereas in the model *with* a mean shift, the same null hypothesis is not rejected [McKitrick and Vogelsang (2011)]. This analysis shows how sensitive the trend estimation is to the shape of the trend being estimated.

To take this sensitivity into account, [Harvey and Mills (2003)] use non-parametric trend estimation for Central England Temperature (CET) data. After first using a parametric approach by fitting a quadratic trend, they find that it is too restrictive to pre-specifiy a particular form of the trend in advance, since usually the behavior of temperature data is more complex and there is a need for a more flexible trend estimation. Non-parametric trend estimation has this feature and the next section provides an introduction to the theory behind this method. Applying a non-parametric method to the CET data series, [Harvey and Mills (2003)] come to the conclusion that there is no general upward trend in the CET data. However, they only investigate data from England, so their analysis is much more restricted than other studies using global data.

1.3 Non-parametric trend estimation

In the previous section about related literature, two major estimation techniques have been distinguished: parametric and non-parametric regression. This section focuses on the latter method and provides a reason, why this method works well in the context of trend analysis in temperature data. To see the main difference between the two approaches, it is important to note how the regressors enter the regression equation in both methods. First, consider a standard linear regression on a trend,

$$y_t = \alpha + \beta t + \epsilon_t. \tag{1.1}$$

Clearly, (1.1) is a parametric regression, as it specifies the series y_t as a linear function of a trend t. More specifically, this equation is a linear regression, where the dependent variable has to be a linear function of the parameters, here α and β. Linear regressions are a specific form of parametric regressions, which are very restrictive, since in the case of trend estimation, it is assumed that the series linearly depends on some kind of trend. The regression provides an estimate of the coefficient β and a standard t-test could be used to test significance, if the assumptions of ordinary least squares regression apply. In short, the conditions on ϵ_t in ordinary least squares regression are homoscedasticitiy and nonautocorrelation. They concern the variance and covariance of the disturbances. Homoscedasticity assumes a constant variance of the disturbance terms ϵ_t over time and nonautocorrelation, in general terms, means that observations are not related over time [Greene (2008)]. As already indicated in the previous section, the assumptions do not hold for temperature data, as the random component of temperature data at one point in time is dependent on this component for other points in time. Thus, the error terms are serially correlated. This is a serious problem for the determination of confidence intervals, which is further addressed in the next section. For now, the focus lies on the form of the equation.

In more general forms of parametric regression, the variable of interest does not have to be a linear function of the parameters, but some prespecified function of the parameters and the independent variables. As a second model, consider the following regression of the temperature on a trend and a trend squared.

$$y_t = \ln(\beta_1 t + \beta_2 t^2) + \epsilon_t \tag{1.2}$$

This equation is non-linear, because of the presence of the natural logarithm. It still belongs to the class of parametric regressions, since the form of the trend is specified before estimating the equation. To see the difference to non-parametric regression, an equation that belongs to this class is

$$y_t = g(t) + u_t. \tag{1.3}$$

Here, the error terms u_t again contain serial correlation, which is taken into account when obtaining confidence bands. They are called u_t to be able to distuingish them later on from the error terms of the other models. The function $g(t)$ is the trend component, which can take on almost any form and is not specified in advance. The only assumptions made on $g(t)$ are assumptions concerning the smoothness of the function. Specifically, these assumptions state that $g(t)$ has to be twice continously differentiable as well as that the first and second derivatives of $g(t)$ are finite. This last assumption can formally be stated as $\sup\left|g^{(j)}(x)\right| < \infty$ for $j = 0, 1, 2$

[Bühlmann (1998)]. Smoothness assumptions on $g(t)$ are crucial, since otherwise, the method of approximating this function by local weighted averages would not work. An explanation of why this is the case can only be given after an introduction of the principle underlying non-parametric estimation and is postponed to the next section. It can be seen from (1.3) that estimating this equation allows for a much more flexible inference of the trend component. However, something has to be given up to get this additional benefit and in this case, reliability and the simplicity of obtaining confidence bands are sacrificed for the benefit of flexibility.

1.3.1 Local Estimators

Estimating the function $g(t)$, or in a more genreal setting $g(x_t)$, where x_t for $t = 1, ..., n$ is a regressor, is more complex than using ordinary least squares estimation. Kernel estimation is one way to do this. In short, non-parametric kernel estimation is a weighted least squares estimation with a weighting function called the *kernel function*. Different kernel functions are available for non-parametric estimation, but only one will be introduced in this paper.

Consider a fixed point x. The function $g(x_t)$ needs to be estimated for every point $x_t = x$ and the correspoding estimate is denoted by $\hat{g}(x)$. In the non-parametric approach, this in done by considering points in a certain interval around x and taking a weighted average of points falling in this interval. This weighted average is the estimator $\hat{g}(x)$ and can generally be written as

$$\hat{g}(x) = \frac{\sum_{t=1}^{n} k(\frac{x_t - x}{h}) y_t}{\sum_{t=1}^{n} k(\frac{x_t - x}{h})}, \tag{1.4}$$

where $\hat{g}(x)$ is called **local constant** estimator and h is the **bandwidth**, whose value has to be carefully chosen. The value of h determines the length of the interval for the estimation, because for every x, only observations are considered that fall in the interval $[x - h, x + h]$. At the beginning and at the end of the sample, one sided intervals are used. The kernel function $k(u)$ is formally defined in the following definition taken from [Hansen (2009)].

Definition 1. *A second-order* **kernel function** $k(u)$ *satisfies* $0 \leq k(u) \leq \infty$, $k(u) = k(-u)$, $\int_{-\infty}^{\infty} k(u)du = 1$ *and* $\sigma_k^2 = \int_{-\infty}^{\infty} u^2 k(u)du < \infty$.

Basically, the kernel function is a symmetric and bounded probability densitiy function. A frequently used kernel is the **Gaussian kernel**

$$k_\phi(u) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{u^2}{2}), \tag{1.5}$$

which represents the Gaussian pdf. Roughly speaking, this kernel function places a lot of weight on the point x, around which the estimation is done, and gradually less weight is given to points further away from x. A similar and also widely used kernel is the **Epanechnikov kernel**

$$k_1(u) = \frac{3}{4}(1-u^2)\mathbf{1}_{\{|u|\leq 1\}}. \tag{1.6}$$

This kernel function has similar properties as the Gaussian kernel, since it also puts the highest weight on x and decreasing weight on other points in the interval. In comparison to the Gaussian kernel, this decrease in weight is less pronounced with the Epanechnikov kernel. In the remainder of this paper, the Epanechnikov kernel is used for the bandwidth selection and kernel estimation.

At this stage, it is convenient to return to the question of why a smoothness assumption on the function $g(t)$ is important. Now, the basic principle underlying non-parametric estimation has been introduced as locally approximating the function $g(t)$ by taking local weighted averages. This principle only works, if $g(t)$ is smooth. Imagine, for example, there was a jump in $g(t)$ at some point. Non-parametric estimation applied to this particular point would not provide a reasonable approximation, because constructing a weighted average of points within the interval $[x-h, x+h]$ would never recover the jump, which actually is present in $g(t)$. However, if the smoothness assumption applies, non-parametric estimation works.

From a different perspective, the local constant estimator presented in equation (1.4) is the result of the constant - $\hat{\alpha}$ - from a weighted regression of y_t on an intercept. Formally,

$$\hat{g}(x) = \arg\min_{\alpha} \sum_{t=1}^{n} k(\frac{x_t - x}{h})(y_t - \alpha)^2 \tag{1.7}$$

is the local constant estimator.

An alternative to the estimator presented in (1.7) is the **local linear** estimator, which is similar to the previous estimator, but instead of being a constant approximation of $g(x)$, it approximates it by a linear function. Comparable to (1.7), the next equation shows a weighted least squares problem of y_t on a constant and a slope,

$$\left\{\hat{\alpha}(x), \hat{\beta}(x)\right\} = \arg\min_{\alpha,\beta} \sum_{t=1}^{n} k(\frac{x_t - x}{h})(y_t - \alpha - \beta(x_t - x))^2, \tag{1.8}$$

where $\hat{\alpha}$ is the local linear estimator of $g(x)$ [Hansen (2009)].

1.3.2 Bandwidth selection

As already indicated, the value of the bandwidth has to be carefully selected and plays an important role. For larger values of h, both estimators become more smooth and for smaller values of h, they become less smooth. Therefore, simply choosing a value for the bandwidth at random is not an option. A value should be taken that results in a good fit of the regression. However, in this setting, a problem of overfitting can easily arise, as for $h \to 0$, a perfect fit is obtained. Just taking the scatter plot of y_t - as it is done when $h \to 0$ - is not the goal of non-parametric estimation. This shows that minimizing the mean-squared error is not an option, either. A way around this problem is provided by the *leave-one-out residuals* [Hansen (2009)]. But before explaning this solution in detail, one formal assumption on h has to be introduced. This assumption deals with increasing sample sizes. Intuitively, as n grows large, a larger value for h should be selected to account for the increase in the number of observations. Formally, the order of h should be $h(n) \sim n^{-1/5}$ [Bühlmann (1998)].

The leave-one-out estimator, $\tilde{g}_{-t}(x)$, leaves the observation out that receives the most weight in the regression so that overfitting is avoided. Without loss of generality, it is assumed that the observation with the highest weight is observation t and the leave-one-out estimator is

$$\tilde{y}_t = \tilde{g}_{-t}(x_t) = \frac{\sum_{i \neq t} k(\frac{x_i - x_t}{h}) y_i}{\sum_{i \neq t} k(\frac{x_i - x_t}{h})} \tag{1.9}$$

for y_t at the fixed point $x = x_t$. The leave-one-out residuals are defined to be $\tilde{e}_t = y_t - \tilde{y}_t$ and present an adequate alternative measure of the goodness of fit. In contrast to the normal residuals, $\hat{e}_t$, it is possible to minimize the leave-one-out residuals over different values of h, without automatically overfitting the regression. The optimal value for h is selected by minimizing the mean squared leave-one-out error as a function of h, which is called the **cross-validation criterion** [Hansen (2009)],

$$CV(h) = \frac{1}{n} \sum_{t=1}^{n} \tilde{e}_t(h)^2. \tag{1.10}$$

After the optimal bandwidth has been found, the local constant or the local linear estimator can be used to non-parametrically estimate a trend in a given series of temperature data.

1.3.3 Application to temperature data

In the introduction of non-parametric estimation, it has been mentioned that a major advantage of using this approach is flexibility. Since no particular form of a trend has to be determined in advance, fewer assumptions are made than in a parametric approach. Obviously, a linear regression like equation (1.1) makes an extreme assumption concerning the form of the trend. It assumes that a linear relationship between the temperature series and a deterministic trend accurately describes the properties of this series. Otherwise, this model would not be appropriate. This is a restrictive assumption and it has been argued that temperature data have a much more complex structure [Harvey and Mills (2003)]. A few simple examples can be thought of to underline this argument. In particular, there is no reason to assume that a linear trend, which by definition has a constant slope, is adequate in this context. A constant slope means that an increase in temperature would be equally fast over the whole period. However, the development of warming can differ for summer and winter or from one subperiod to another. In addtion - as studied by [McKitrick and Vogelsang (2011)] - there can be breaks in global warming. Climate shifts are not considered in a linear trend model, unless the model is expicitly changed to take them into account.

These examples show that flexibility of non-parametric kernel estimation is beneficial in the context of temperature data, because the complex structure can be taken into account. Nevertheless, the benefit of flexibility comes at a cost. The results from non-parametric estimation are less reliable than using a simple parametric approach and it is difficult to judge the significance of these estimates. This topic is addressed in the next section.

1.4 Constructing Confidence Intervals

In the context of temperature data, autocorrelated error terms are an important issue. There clearly exists a dependence of the error terms over different observations. Considering again the simple model (1.1), fit on a series of monthly temperature data, it is hard to believe that a huge deviation from the trend in one month is completely independent of a deviation in the previous months or in the subsequent months. This is because the temperature on one day depends on the temperature on the preceding day and this will, in turn, influence the temperature on the next day. This dependence is carried over to monthly data. A formal definition of the concept of autocorrelation is given by Definition 2.

Definition 2. *A time series is called autocorrelated, if* $Cov[\epsilon_t, \epsilon_{t-j}] \neq 0$ *for some* $j \neq 0$.

Nonautocorrelation is one of the assumptions underlying ordinary least squares regression and as soon as this assumption does not hold, a measure has to be taken to account for this missing assumption. Without accounting for this problem, spurious evidence for a trend in the data is likely to be the outcome. In the academic literature, various measures against autocorrelation have been suggested. Among these measures are using an autocorrealtion robust estimator like the Newey-West estimator, a new estimator proposed [Vogelsang (1998)] or fitting an appropriate ARMA(p,q) model. As mentioned in Section 2, some of these measures do not work well with temperature data. In particular, applying a Newey-West estimator might lead to overrejection of a true null hypothesis and similar to using a standard t-test, false evidence of global warming could be created [Fomby and Vogelsang (2002)]. This is a result of judging the significance of an estimator, using inaccurate confidence intervals.

A different way of dealing with serial correlation is applying a *bootstrap* method to obtain confidence intervals. In general, bootstrap is the name of a method that takes random samples of the data with replacement to be able to make inference about the distribution of a test statistic. In other words, this method resamples from the sample. Usually, confidence intervals are determined using the asymptotic distribution of an estimator. In the presence of serial correlation, this is difficult and an additional complication exists when dealing with non-parametric estimators. The asympotic distribution of a non-parametric estimator is non-trivially biased [Hansen (2009)]. This makes asympotitic distributions infeasible to construct confidence bands, since it is impossible to account for this bias without a significant loss of precision. Applying a bootstrap method to get information about the distribution of the estimator solves this problem. A quote taken from a chapter by [Horowitz (2001)] describes, how a general bootstap method works.

> *"It amounts to treating the data as if they were the population for the purpose of evaluating the distribution of interest."(Horrowitz, 2009, p.1)*

This quote gives an intution of the principle underlying the method. Instead of using mathematical tools to calculate the asympotitc distribution, the data is used to construct bootstrap samples, which, in turn, are used to make inference about a particular test statistic or distribution.

Many different methods of bootstrapping exist. Although it is usually referred to "the bootstrap", it is important to distinguish between alternative methods, as some might be more or less appropriate in certain cases [McKinnon (2006)]. In general, two main classes of bootstrap methods are *parametric* and *non-parametric* bootstrap. A parametric bootstrap method is used when the distribution of the data is known, but the parameters specifying this distrubtion are unknown. If no information about the distribution exists, a non-parametric bootstrap should be used [Smeekes (2009)]. In the context of this paper, the second method applies so that in the remainder, the focus is on non-parametric bootstrap methods. Among these methods are the i.i.d. bootstrap, the block bootstrap and the sieve bootstrap. The abbreviation i.i.d. stands for *indepentent and identically distributed* and summarizes two characteristics that data can have. First, the observations are mutually independent and second, they come from the same distribution. The i.i.d. bootstrap is a standard method, which is only valid when it is applied to i.i.d. data. An algorithm of this method is presented below. The basic idea underlying the block bootstrap is shortly presented in the next section, whereas the main topic of this section is the sieve bootstrap. The i.i.d. bootstrap can be summarized in three steps.

- *Step 1*
  From the data $(x_1, x_2, ..., x_n)$, draw a bootstrap sample $(x_1^*, x_2^*, ..., x_n^*)$ randomly and with replacement.

- *Step 2*
  Apply the statistic of interest, T_n, to the bootstrap sample and obtain

$$T_n^* = T_n(x_1^*, x_2^*, ..., x_n^*).$$

- *Step 3*
  Produce B replications of Step 1 and Step 2 with $T_{n,b}^*$ denoting the statistic applied to the bth bootstrap replication. Compute

$$\frac{1}{B}\sum_{i=1}^{B} I(T_{n,b}^* \leq x)$$

  to estimate the distribution of the statistic T_n, which is dependent on the distribution of the data.

1.4.1 Bootstrapping Time Series Data

Not every bootstrap method can be used with time series data. In the case of temperature data, a method has to be applied that can keep the dependence structure that exists in the data over the sample period. Simply resampling with replacement would lead to a loss of the dependence structure [Smeekes (2009)]. In particular, a *block bootstrap* or a *sieve bootstrap* can be applied in this context. Both methods deal with serially correlated, stationary time series data. The idea behind the block bootstrap is that blocks are resampled instead of single observations so that the structure is maintained within these blocks. The focus in this paper, however, lies on the sieve bootstrap proposed by [Bühlmann (1998)]. The basic idea underlying this method substantially differs from the idea of the block bootstrap. In the sieve bootstrap, the dependece structure of the data is estimated with an appropriate model and by using the residuals of this estimation, the dependence structure can be neglected and an i.i.d. bootstrap approach can be used [Smeekes (2009)].

The Sieve Bootstrap

The bootstrap algorithm underlying the sieve bootstrap by [Bühlmann (1998)] consists of four steps. For a better understanding of the notation used, recall equation (1.3), $y_t = g(t) + u_t$. This equation can be estimated using a local constant or local linear estimator to obtain $\hat{g}(t)$. Bootstrap quantities are denoted by an asterisk as superscript.

- *Step 1*
  Form the residuals from the estimation of the series on the non-parametric trend. This means, calculate

  $$\hat{u}_t = y_t - \tilde{g}(t), \quad t = 1, ..., n,$$

  where the estimate $\tilde{g}(t)$ can be obtained by either using the same estimation as in the previous section, $\hat{g}(t)$ - or a new estimation using a different bandwidth $\tilde{h}$. When a different bandwidth is used, the new bandwidth $\tilde{h}$ can be obtained from the old bandwidth, h, using the relation

  $$\tilde{h} = Ch^{5/9}, \quad C = 1/2, 1, 2.$$

- *Step 2*
  To the residuals $\hat{u}_t$ for $t = 1, ..., n$, fit an autoregressive model of order p and form the new series of residuals

  $$\hat{\epsilon}_t = \hat{u}_t - \sum_{j=1}^{p} \hat{\phi}_j \hat{u}_{t-j}, \quad t = p + 1, ..., n.$$

  Subtract the mean $\bar{\epsilon} = \frac{1}{n-p} \sum_{t=p+t}^{n} \hat{\epsilon}_t$ to form $\tilde{\epsilon}_t = \hat{\epsilon}_t - \bar{\epsilon}$.

- *Step 3*
  Draw randomly with replacement from $\tilde{\epsilon}_t$ to obtain ϵ_t^*.

- *Step 4*
  Calculate the boostrap errors u_t^* as

$$u_t^* = \sum_{t=1}^{p} \hat{\phi}_j u_{t-j}^* + \epsilon_t^*.$$

Now, generate the bootstrap observations by

$$y_t^* = \tilde{g}(t) + u_t^*, \qquad t = 1, ..., n,$$

where $\tilde{g}(t)$ is the same estimated value as in the first step.

Steps 3 and 4 are repeated B times, where B stands for the number of replications. The more replications are generated, the higher the precision of the estimation, but the more time is needed to do the computations [Smeekes (2009)]. Furthermore, in Step 1, a different bandwidth than during the non-parametric estimation can be used. [Bühlmann (1998)] suggests that the new bandwidth, $\tilde{h}$, should be related to the original bandwidth, h, by $\tilde{h} = Ch^{5/9}$, where $C = 1/2, 1$ or 2. In his simulation results [Bühlmann (1998)] shows that setting $C = 1/2$ is appropriate.
There are three assumptions about the noise process underlying the method by [Bühlmann (1998)]. First, the moments of ϵ_t are finite up to the fourth moment. Formally, $\mathbb{E}\left|\epsilon_t\right|^4 < \infty$ has to hold. Second, an assumption is made about the summability of the lag coefficients of the AR model. In particular, $\sum_{j=0}^{\infty} j\left|\phi_j\right| < \infty$ is the formal assumption and states that there has to be an appropriate decay of the lag coefficients. Intuitively, this means a lower contribution to the AR model of values that are far in the past. Third, [Bühlmann (1998)] assumes that, as the AR(p) model is an approximation of the infinite order AR(∞) model, the lag specification p has to increase with n at a certain speed. This assumption is presented in [Bühlmann (1998)] as

$$p(n) = o\left(\min\left\{(n/log(n))^{1/4}, (n\tilde{h})^{1/4}, \tilde{h}^{-1}\right\}\right), \qquad for\, n \to \infty.$$

In the application of the steps presented above, this is accounted for by taking a maximum value of p, which is dependent on the sample size and equals $10\log_{10}(n)$.

Pointwise Confidence Intervals

The non-parametric estimator, which is used to obtain $\hat{g}(t)$, is applied to all B series of bootstrap observations y_t^* to obtain $\hat{g}^*(t)$ and for every point t, the deviations of these B esimations to the value $\tilde{g}(t)$ are determined. The value $\tilde{g}(t)$ is set to be the *true* value in this case. The deviations are ordered from the largest negative to the highest positive deviation from the true value. From these ordered series, it is straightforward to determine the values for the pointwise two-sided confidence interval for a confidence level of $1-\alpha$. These are exactly the values for every t, between which $1-\alpha$ of the deviations fall. Formally, this can be stated as

$$I^*(t, 1-\alpha) = \left[\hat{g}(t) - \hat{q}_{1-\alpha/2}, \hat{g}(t) - \hat{q}_{\alpha/2}\right], \tag{1.11}$$

where $1-\alpha$ is the confidence level and $\hat{q}_\alpha = \inf\left\{u; P^*\left[\hat{g}^*(t) - \tilde{g}(t) \leq u\right] \geq \alpha\right\}$. From equation (1.11), it can be seen that the confidence bands only apply to a certain t, showing that they are only valid for this particular point in time. In trend estimation, this is not particularly useful, as the focus is not on the accuracy of the estimation of single observations, but on the estimation of the behavior over time. To achieve this, simultaneous confidence intervals are needed.

Simultaneous Confidence Intervals

Simultaneous confidence intervals are calculated for every point in time, but they are valid for either a certain neighborhood G or for the whole time covered by the sample. In this case G is simply set to contain the whole sample [Bühlmann (1998)]. This is the main difference to pointwise confidence intervals, which are only valid for one single observation.

The first step to obtain these intervals for a confidence level of $1-\alpha$ is similar to the calculation of pointwise bootstrap quantiles presented in the previous section. The additional step is that this is done for varying values of α_p according to an analogous formula than stated above, $\hat{q}_{\alpha_p} = \inf\left\{u; P^*\left[\hat{g}^*(t) - \tilde{g}(t) \leq u\right] \geq \alpha_p\right\}$. In addition, these quantiles are obtained for every point in the neighborhood G. Note that G may contain the whole sample, which is assumed throughout this paper. This step provides a pair of quantiles - $\hat{q}_{\alpha_p/2}(t), \hat{q}_{1-\alpha_p/2}(t)$ - for every $t \in G$ and for every α_p under consideration. The values of α_p range from $1/B$ to a final value of α in increments of $1/B$. In a second step, the value of α_p is selected, which guarantees that a fraction of around $1-\alpha$ of all bootstrap series are completely within the corresponding intervals $\left[\hat{q}_{\alpha_p/2}(t), \hat{q}_{1-\alpha_p/2}(t)\right]$ for all $t \in G$. Formally, this can be stated as to select the value of α_p such that

$$P^*\left[\hat{q}_{\alpha_p/2}(t) \leq \hat{g}^*(t) - \tilde{g}(t) \leq \hat{q}_{1-\alpha_p/2}(t)\ \forall x \in G\right] \approx 1-\alpha \tag{1.12}$$

is satisfied. This optimal value of α_p is then denoted by α_s and used to construct the simultanoues confidence bands as in [Bühlmann (1998)]

$$I_n(t) = \hat{m}(t) - \hat{q}_{1-\alpha_s/2}(t), \hat{m}(t) - \hat{q}_{\alpha_s/2}(t), \quad x \in G. \tag{1.13}$$

1.5 The Data

Non-parametric trend estimation and the sieve bootstrap method are applied to several series of temperature data from Europe. The data are taken from the European Climate Assessment & Dataset (ECA&D), which was founded in 1998 by the European Climate Support Network (ECSN). This Network supports the 25 members in their climate analysis and helps to provide the users of this network with climate services and products. The dataset was updated in April 2012 and contains data on changes in weather extremes from 62 countries. Series on temperature data are available at a daily frequency for various time periods and for stations all over Europe. In order to provide time series data that is as complete as possible, all data is available as blended and non-blended series. The blended series are more complete than non-blended series, as missing values are taken from neighboring stations that are at most 12.5 km away from the original station. Furthermore, the height difference between the stations should not exceed 25 m. All series used in this paper are taken from the blended data set, since having a (nearly) complete dataset is important for trend estimation.

In addition, the daily data are averaged monthly in order to reduce volatility and the impact of missing daily data. For non-parametric estimation to be as accurate as possible, it is crucial to have time series data with many observations, which cover a sufficient period of time. The temperature series considered in this paper might have different ranges, since having exactly the same range for all series does not add any benefit to the results. However, most of the series start after World War II, because for many cities, temperature series were not available or complete for some years during the war. There is one exception to this range, which is the Czech city of Prague. This series starts in the late 18th century and provides seamless daily data until the end of 2004, which makes it particularly interesting to study. With this range, this city provides the series with the largest number of observations availabe in this dataset and therefore, it is the only series used to study a trend in temperature during the full range of the 19th and 20th century. The other series analyzed are obtained from weather stations in Perpignan, Zurich, Lerwick, St.Petersburg, Hannover, Karlstad and Prague.

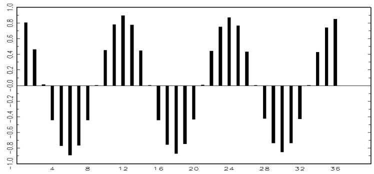
To seasonally adjust the data, averages are constructed for every month using the data for a particular month over the whole sample. These averages are subtracted from the corresponding observations in the sample and these differences constitute the adjusted series, which is used for the estimation and the application of the bootstrap. Important information regarding the series is summarized in Table 1.1. The sample periods, given in the third column of Table 1.1, start in January of the first year and end in December of the final year. The number of observations in the next column are stated in unit of months.

City	**Country**	**Range**	**Obs.**
Hannover	Germany	1946-2011	792
Karlstad	Sweden	1950-2011	744
Lerwick	Scotland	1945-2011	804
Perpignan	France	1946-2011	792
Prague	Czech Republic	1775-2004	2760
St.Petersburg	Russia	1943-2011	828
Zurich	Switzerland	1901-2009	1308

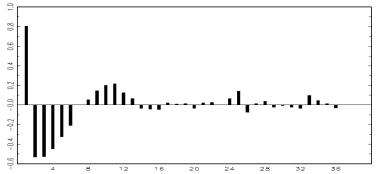
Table 1.1: Stations to be analyzed

Table 1.2 presents some descriptive statistics on the series. It shows that the selected series are diverse and reflect different geographic and climatic areas, since some of them differ substantially in mean temperature, as well as in minimal and maximal values. All temperature values in Table 1.2 are given in $^\circ$C and are calculated from monthly averages, which are not yet seasonally adjusted. The reason to use monthly averages to give descriptive statistics is to filter out some extreme values and to reduce volatility, as with daily data the standard deviation would be extremely high and the minimal and maximal values would be given by some extreme values. Furthermore, daily data series are usually incomplete, since some daily values are missing or are classified as unreliable. However, the number of missing values in the series analyzed was kept within a limit of 2 values per month and a maximum of 17 values for a complete series. Some series - like the one from Prague, Hannover, Perpignan and Zurich - had no missing values. The values in the final column are not given in $^\circ$C, since this column presents the first order estimated autocorrelation coefficient for each series. For all series, these coefficients lie between 0.8 and 0.83.

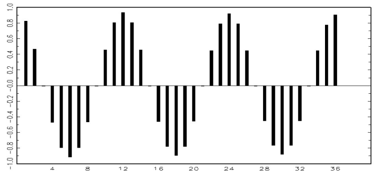
To give an impression about how a typical correlogram looks like, Figure 1.1a plots the estimated autocorrelation of the Hannover series and Figure 1.1c does the same for the Perpignan data. Figure 1.1b and Figure 1.1d display the corresponding partial autocorrelations. The two autoccorelation



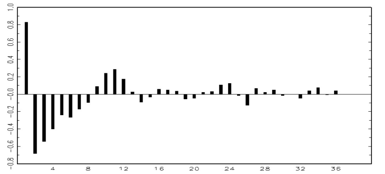
(a) Autocorrelation Hannover



(b) Partian Autocorrelation Hannover



(c) Autocorrelation Perpignan



(d) Partial Autocorrelation Perpignan

Figure 1.1: Correlograms of Hannover and Perpignan

City	Mean	Max	Min	Std.Dev.	AR(1) coeff.
Hannover	9.12	22.31	-9.32	6.36	0.806
Karlstad	5.93	20.51	-13.45	7.69	0.825
Lerwick	7.20	14.37	-0.34	3.34	0.809
Perpignan	15.51	27.61	1.56	5.74	0.827
Prague	9.60	25.05	-10.92	7.63	0.825
St.Petersburg	5.17	24.46	-17.94	9.17	0.819
Zurich	8.94	22.73	-8.73	6.69	0.812

Table 1.2: Descriptive Statistics (values in °C) - monthly averages

figures show an almost identical pattern of wave-like autocorrelation for lags up to a number of 36. The two figures presenting the partial autocorrelation show a decay after a few lags and some significant spikes for longer lags. They are also comparable for the two series. An inspection of correlograms for the other five series shows that the pattern is almost the same for all series.

Coming back to Table 1.2, with a standard deviation of 9.17°C, the series of St.Petersburg is the most volatile one in this set. The highest monthly average of 24.26°C for St.Petersburg is a July observation from 2010, whereas its month on average is January 1987 with an average of -17.94°C. At the same time, this observation is the lowest monthly average of all considered series. The highest monthly average of 27.61°C was reached in July 2006 in Perpignan, which is not surprising, as it is the most southern city studied in this paper.

1.6 Results

This section presents results of the application of the non-parametric local constant and local linear estimators to the data described in the previous section. On the basis of these estimates, the sieve bootstrap method is used to construct simultaneous confidence intervals for the trend functions. The local constant estimator is applied to the complete set of temperature series as presented in Section 5, whereas the local linear estimator is only applied to the series of Hannover and Perpignan. For these two series, the two estimators provide very similar trend estimations, suggesting that it is not a major limitation to rely on the local constant estimator for the five remaining series. Furthermore, according to [Hansen (2009)], none of the two estimators is strictly outperformed by the other.

The whole analysis presented in this paper is done under the maintained assumption of stationary data. This is underlined by a statement concerning the stationarity of temperature data taken from [Fomby and Vogelsang (2002)]. In this paper, it is stated that "nonstationarity can be ruled out on scientific grounds" ([Fomby and Vogelsang (2002)], p.120). A visual inspection of all the series and a unit root test further support this maintained assumption. An Augmented Dickey-Fuller unit root test has been conducted on the series - including an intercept and a trend - showing that the null hypothesis of having a unit root in the data is rejected at a 1% significance level for all series.

In the remainder of this section, a plot of the monthly averaged data is presented, combined with a fitted linear trend. In addition, for every series, two graphs present the results of the non-parametric local constant estimation. One graph includes the estimator, the simultaneous confidence intervals and the actual data, while the other graph does not plot the actual data to enable a better examination of the form of the trend. For the two series of Hannover and Perpignan, the local linear estimation is presented in a similar fashion.

Figures 1.2 to 1.3 plot the seasonally adjusted monthly data including a linear trend. The linear trend has been estimated according to model (1.1), presented in Section 3, where the residuals of this estimation are serially correlated for all series. The estimate of the linear trend coefficient β is presented underneath the corresponding figure using ordinary least squares regression. To take serial correlation into account, the standard deviation in brackets is obtained using the autocorrelation robust Newey-West estimator.

For all seven series, the trend coefficient is extremely small and not significantly different from zero according to the Newey-West standard deviation - using a 5% significance level. This would provide evidence of no global warming for the cities studied in this paper. However, allowing for a more flexible trend function changes these results, as can be seen in the next section.

1.6.1 Local Constant Estimation Results

The local constant estimator is applied to all seven series and this section, as described above, presents two plots for every city. It also provides the bandwidth used to obtain the estimation and the order of the AR(p) model selected during the process of the sieve bootstrap method. The first plot always includes the actual, seasonally adjusted data to present a first impression of the trend function next to the series. The second graph leaves out the actual data and only plots the non-parametric estimation together with the simultaneous confidence bands to allow a clear observation of the specific form of each trend function. All results are obtained using the Epanechnikov kernel function. As already mentioned, at the beginning and at the end of the sample, only one-sided or truncated intervals are used and this is why there is a more or less visible widening of the confidence bands for all series.

The bandwidth was carefully selected using the cross-validation criterion. As a solution to the minimization problem can only be found numerically, this is done by minimizing this criterion over a suitable range for h, which should not include unreasonably small values [Hansen (2009)]. Formally, [Hansen (2009)] states this as

$$\hat{h} = \underset{h \geq h_l}{\text{argmin}} CV(h),$$

where a value for $h_l > 0$ has to be chosen manually for every series. The selection of h_l has been done using plots of $CV(h)$ against different ranges of h. An example of such a plot is given for the Hannover data in Figure 1.4a. In this case, the range of h goes from 40 to 340 and clearly, the starting value of 40 is too small. A value of around $h_l = 80$ should be chosen to avoid an unreasonably small value for h to be selected. Taking this into account, the optimal bandwith for Hannover equals 238. However, this value will not be selected using the range of Figure 1.4a, because the minimum is attained at 40. This would lead to an estimation which is less smooth than the one using $h = 238$. Figure 1.4b plots the values of $CV(h)$

for h ranging from 80 to 380 and shows that this is a much more appropriate range. Using this range, the optimal value of 238 can finally be chosen with the help of the CV criterion.

After estimating the trend using an appropriate bandwidth, the sieve bootstrap method is applied according to the four step algorithm presented in Section 4. The residuals in Step 1 are calculated using the same bandwidth as determined by the cross-validation criterion for the non-parametric estimation. Applying $\tilde{h} = Ch^{5/9}$ with $C = \frac{1}{2}$ as suggested by [Bühlmann (1998)] does not provide significantly different results than using the same value of h as for the determination of $\hat{g}(t)$.

Figure 1.5 presents the results of Hannover. The value of the bandwidth is, as already mentioned, $h = 238$ and the lag length is $p = 26$ months. In general, these values are given in brackets below the graphs. The two graphs show that since the middle of the 20th century, there has been an upward trend in temperature for the city of Hannover. This result could not be extracted using linear trend estimation. Before around 1965, a slight downward trend is visible and for the most recent years, there is indication of a deceleration of the trend. The pattern that is exhibited by the Hannover series can be recognized in the graphs of some other cities. In particular, Figure 1.6 and Figure 1.7 show a similar trend for Lerwick and Perpignan. There, the upward trend starts a few years later as in Hannover - around 1970 in Lerwick and 1975 in Perpignan - and is preceded by a more pronounced downward trend since the start of the two series in 1945. A downward trend would actually be associated with a period of global cooling. The series of Lerwick does not show a recent slowdown of the trend, whereas for Perpignan, the increase in temperature seems to slightly decelerate. Whether this represents an actual slowdown can only be seen in the future.

The trend function of Zurich is estimated over a longer period than the ones of Hannover, Lerwick and Perpignan and is presented in Figure 1.8. However, over the period that these four series have in common - from 1945 to 2009 - the form of the trend in Zurich temperature data is comparable to the other three series. A period of global cooling until around 1970 is followed by an upward trend, indicating a period of global warming. Unfortunately, more recent data on this series is not available in the dataset used. The period between 1901 and 1945 is characterized by an upward trend, which is not as strong as the one occuring after 1970.

A different pattern can be seen in Figure 1.9 for St.Petersburg. This series displays a trend after a 20 year period of stagnation. For this series, global warming seems to take place at an almost constant pace after 1960 and no deceleration can be observed. This is in contrast to the trend estimation of Karlstad displayed in Figure 1.10. This most nothern city among the seven series shows the highest variation, although bandwidth and range are in the same order as for Perpignan. In addition, the simultaneous confidence bands are wider as for all the other series. This may be an indication of lower precision of the estimation. The range of values that the estimation is likely to be located in is much larger. For this series, it might be better to use a larger sample, which is not available in the dataset used. Looking at the form of the trend for this series, there does not seem to be a period of continued global warming or cooling. An upward trend started in the early 1980s, but it flattens out rapidly after less than 20 years. Overall, this estimation shows a completely different form compared to all other trend functions and has to be interpreted with care, as the confidence intervals are relatively wide.

The series of Prague by far covers the longest period of the seven series. It starts in 1775 and ends in 2004. The beginning of the period studied is characterized by an upward trend until 1800, followed by a downward trend and a period of stagnation until around 1900. Since 1900, the estimated trend function is increasing with a rise in the slope of this function in 1980. This suggests that in Prague, a period of global warming started in 1900 and is still continuing in 2011.

Overall, the trend functions estimated non-parametrically show a complex structure and most temperature series exhibit an upward trend over at least part of the period studied. For most of the cities, this upward trend started in the near past and seemed to persist until the end of 2011. None of the series showed a significant upward trend, when estimated using a linear trend. This clearly underlines the limitations of this approach. Some of the trends could probably have been recovered fitting a quadratic trend. In particular, the series of Perpignan and Lerwick display a near parabolic shape. However, this form has to be known or at least assumed before estimating a quadratic trend, whereas with non-parametric estimation, no knowledge or assumptions about the form of the trend have to be present. In the case of the Prague series, which covers a long period and contains a lot of observations, the complex shape of the trend function could most likely not have been recovered using some form of parametric trend estimation.

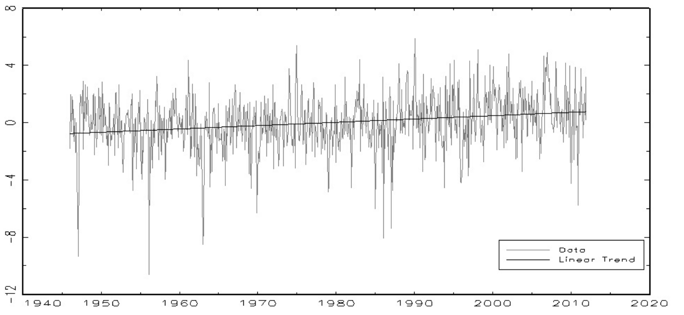
The analysis of this section together with the results of the linear trend estimation emphasize the advantage of non-parametric estimation. The flexibility of this estimation technique makes it possible to extract any form of a trend and no knowledge is needed in advance about a possible shape of the trend. After obtaining the results of non-parametric estimation, it is, of course, possible to conclude that this shape could have been estimated using a different, parametric approach. But this is only possible *after* having observed the results of the approach advocated in this paper.

1.6.2 Local Linear Estimation Results

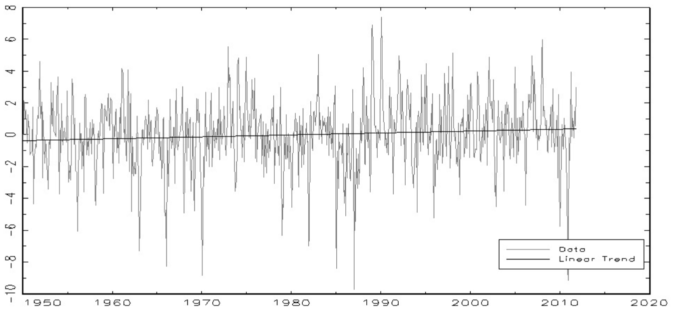
In the previous section, the results of the non-parametric approach have been presented using the local constant estimator as presented in Section 3.1. This section shows results of the local linear estimator, which was introduced in the same section. The main difference between the two estimators is that a local constant estimator approximates $g(x)$ for every x by a constant function, whereas the local linear estimator does the same by using a linear function. It has been argued by [Hansen (2009)] that it is not obvious which one of the two estimators should be used, because none of them performs strictly better than the other. There is indication of a better performance by the local linear estimator, when the function to be estimated is significantly non-constant. Similarly, the local constant estimator seems to outperform the local linear estimator, when the shape of the function is almost constant [Hansen (2009)].

This section presents the results of applying the local linear estimator to two out of the seven data series. In particular, results are given for the two cities of Hannover and Perpignan. The way this is done is similar to how the local constant estimation results have been presented. Each figure contains two graphs, one including the data, the estimation and the confidence bands and a second graph leaving out the data. From Figure 1.12 it can be inferred that the estimation results of Hannover are comparable to the results of the local constant estimation for this series. The trend has a very similar shape to the one in Figure 1.5, only the confidence bands widen even more at the beginning and at the end of the series with the local linear estimation. Similarly, Figure 1.13 displays a trend function almost identical to Figure 1.7 but with an extreme widening of the confidence bands.

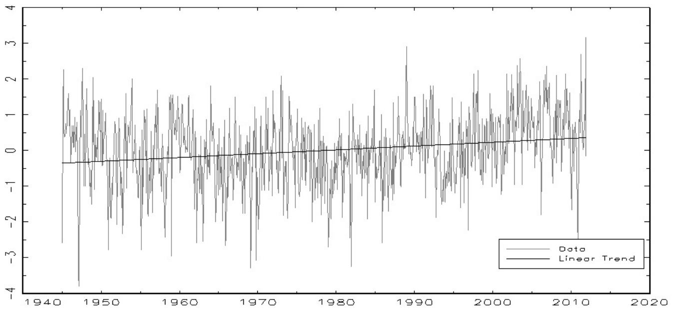
Overall, for these two series, the two different estimation techniques do not show significanlty different results. Due to computational limitations, the local linear estimator is not applied to the remaining five data series. However, more evidence has been collected in order to show the importance of flexibility in trend estimation. Both approaches - local linear and local constant estimation - show a complex form of the trend in these series of temperature data.



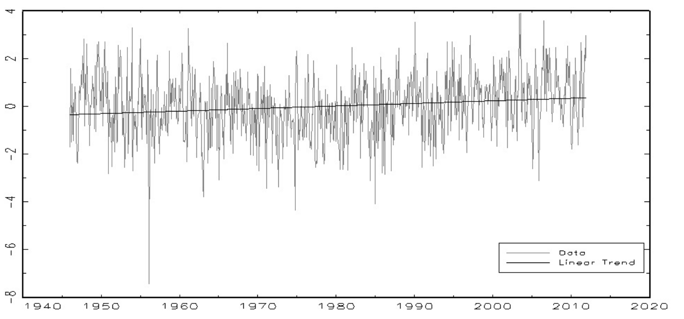
(a) Hannover, $\hat{\beta} = 0.001934(0.001404)$



(b) Karlstad, $\hat{\beta} = 0.0010041(0.001853)$

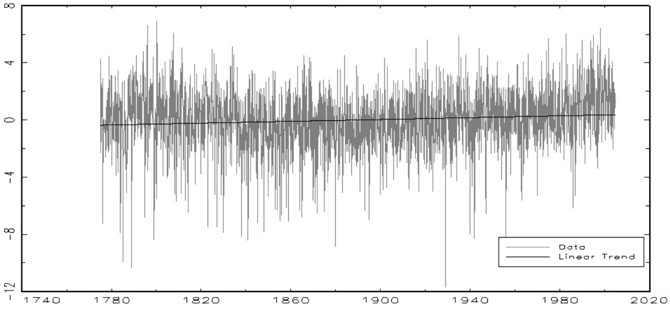


(c) Lerwick, $\hat{\beta} = 0.000897(0.000178)$

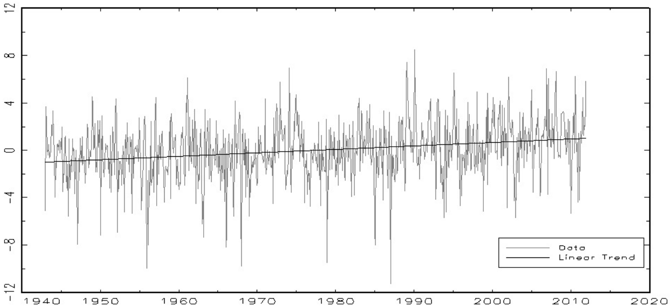


(d) Perpignan, $\hat{\beta} = 0.00897(0.001286)$

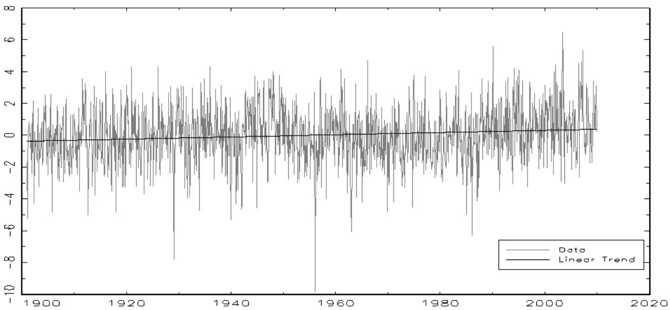
Figure 1.2: The data and a linear trend estimation



(a) Prague, $\hat{\beta} = 0.000267(0.000178)$

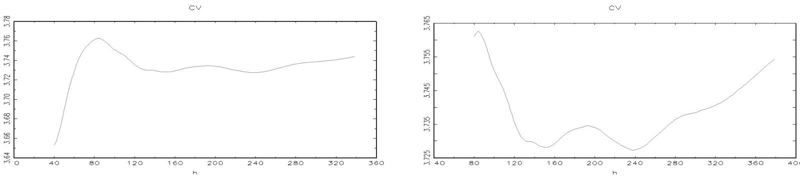


(b) St.Petersburg, $\hat{\beta} = 0.002438(0.001868)$



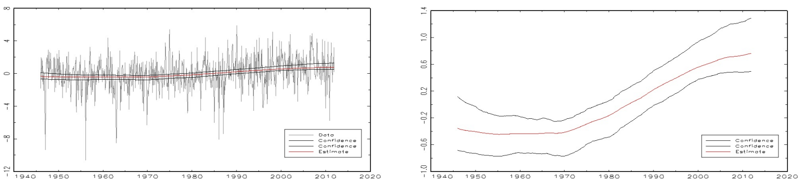
(c) Zurich, $\hat{\beta} = 0.000577(0.000578)$

Figure 1.3: The data and a linear trend estimation - continued



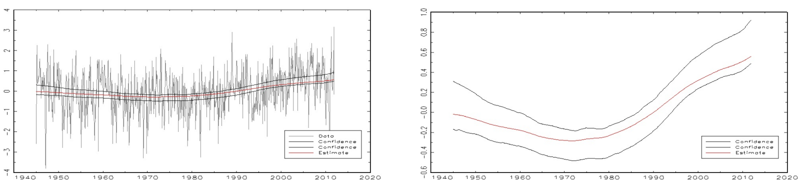
(a) h ranging from 40 to 340 (b) h ranging from 80 to 380

Figure 1.4: Example of Crossvalidation Minimization for the Hannover Series



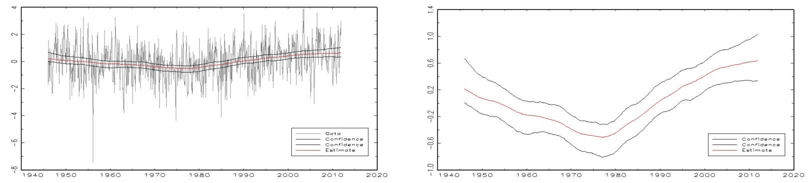
(a) Local constant estimator and actual data ($h = 238$, $p = 26$) (b) Local constant estimator ($h = 238$, $p = 26$)

Figure 1.5: Local constant estimation of the Hannover series



(a) Local constant estimator and actual data ($h = 222$, $p = 3$) (b) Local constant estimator ($h = 222$, $p = 3$)

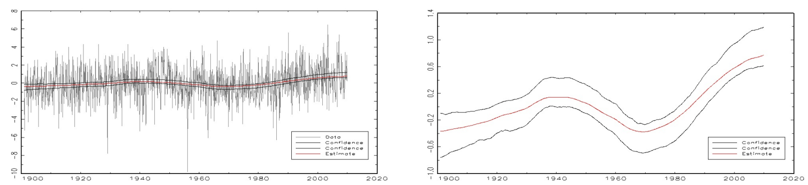
Figure 1.6: Local constant estimation of the Lerwick series



(a) Local constant estimator and actual data ($h = 120$, $p = 4$)

(b) Local constant estimator ($h = 120$, $p = 4$)

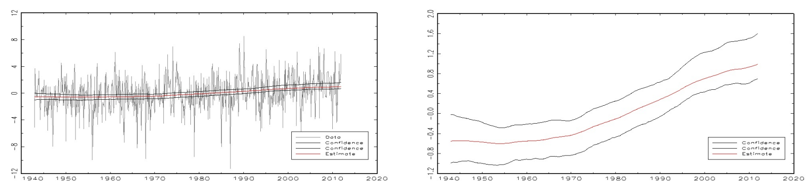
Figure 1.7: Local constant estimation of the Perpignan series



(a) Local constant estimator and actual data ($h = 223$,$p = 2$)

(b) Local constant estimator ($h = 223$,$p = 2$)

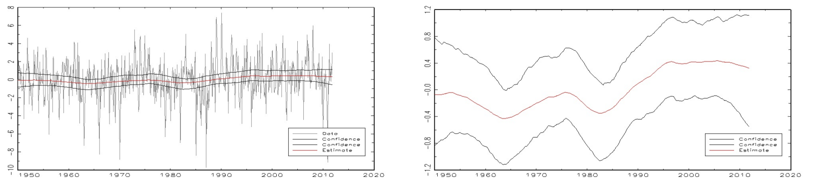
Figure 1.8: Local constant estimation of the Zurich series



(a) Local constant estimator and actual data ($h = 229.5$,$p = 2$)

(b) Local constant estimator ($h = 229.5$,$p = 2$)

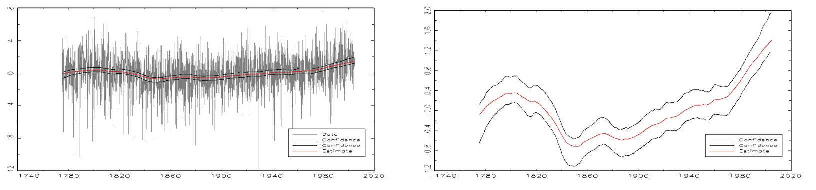
Figure 1.9: Local constant estimation of the St.Petersburg series



(a) Local constant estimator and actual data ($h = 114$, $p = 1$)

(b) Local constant estimator ($h = 114$, $p = 1$)

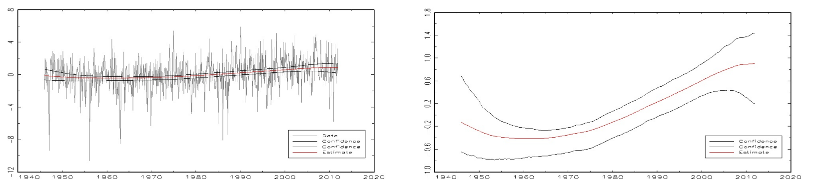
Figure 1.10: Local constant estimation of the Karlstad series



(a) Local constant estimator and actual data ($h = 224$, $p = 2$)

(b) Local constant estimator ($h = 224$, $p = 2$)

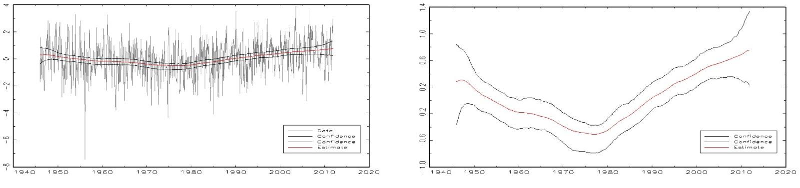
Figure 1.11: Local constant estimation of the Prague series



(a) Local linear estimator and actual data ($h = 299$, $p = 26$)

(b) Local linear estimator ($h = 299$, $p = 26$)

Figure 1.12: Local linear estimation of the Hannover series



(a) Local linear estimator and actual data ($h = 119$, $p = 4$) (b) Local linear estimator ($h = 119$, $p = 4$)

Figure 1.13: Local linear estimation of the Perpignan series

1.7 Conclusion

In this paper, non-parametric kernel estimation in combination with a sieve bootstrap method have been introduced. First, these two methods have been presented on a theoretical basis and second, they have been applied to seven series of temperature data. On theoretic grounds, the advantages of this approach have been argued to be a major gain in flexibility and the fact that valid confidence intervals for the non-parametric estimator can be found in the presence of serial correaltion. Particularly, the construction of confidence intervals around a non-parametric estimator is challenging, because the asymptotic distribution of such an estimator is non-trivially biased. This problem is avoided by using a bootstrap approach. Nevertheless, a bootstrap approach has to be carefully selected in this context. With dependent time series data, such a method has to maintain the dependence structure within the data. The sieve bootstrap method achieves this goal by estimating the dependence structure and using the residuals of this estimation for the resampling process. The validity of this approach has been shown in [Bühlmann (1998)].

The application to temperature data of different areas in Europe underlines the importance of flexibility. Fitting a linear trend to the seven series results in an insignificant trend coefficient for all series. This gives the impression that no global warming occured over the period. However, applying the non-parametric approach reveals a complex shape of the trend functions, which includes a period of global warming for most series. The series of Hannover, Perpignan, Lerwick and parts of the Zurich series display a similar trend pattern. After a period of global cooling, an upward trend begins, indicating the presence of global warming. The point in time, where the downward trend turns to an upward trend is different for the series mentioned - varying from 1965 to 1975. The series of St.Petersburg and Prague also show a period of global warming, but the pattern is slightly different. In St.Petersburg, an upward trend started after a period of stagnation in 1970 and in Prague, there has been a continuing upward trend after around 1900. The series of Lerwick constitutes an exception, since the confidence bands are wider than for all the other series, the reason of which remains an open question.

The whole analysis in this paper presents evidence on the complexity of trend estimation in temperature data. Trend estimation has been shown to be extremely sensitive to the form of the estimated trend. It has been argued that non-parametric estimation does not specifiy a form of a trend before estimating and is therefore able to reveal any form of a trend that is present in the data. Parametric trend estimation in general and linear trend estimation in particular do not take this complexity and sensitivity into account. The question that remains is why institutions like the IPCC still rely on these methods despite such major drawbacks. The topic of trend estimation in temperature data has to be carefully analyzed, because its interpretation can have widespread consequences on global warming. Environmental politics relies on such estimates when making important decisions concerning future activies against global warming. Therefore, it might be advisable to add non-parametric inference to the set of tools, when estimating temperature trends.

Bibliography

[Bühlmann (1998)] Bühlmann, P. (1998). *Sieve bootstrap for smoothing in nonstationary time series.* Annals of Statistics, 26, 48-83.

[Fomby and Vogelsang (2002)] Fomby, T. and T.J. Vogelsang (2002). *The application of size-robust trend statistics to global warming temperature series.* Journal of Climate, 15, 117-123.

[Franses and Vogelsang (2005)] Franses P.H. and T.J. Vogelsang (2005). *Are winters getting warmer?* Environmental Modelling & Software, 20, 1449-1455.

[Greene (2008)] Greene, W.H. (2008). *Econometric Analysis.* Pearson Prentice Hall, 6th Edition.

[Hansen (2009)] Hansen, B.E. (2009). *Econometrics.* Univeristy of Wisconsin.

[Harvey and Mills (2003)] Harvey, D.I. and T.C. Mills (2003). *Modelling trends in Central England temperatures.* Journal of Forecasting, 22, 35-47.

[Horowitz (2001)] Horowitz, J.L (2001). *The Bootstrap.* Handbook of Econometrics, Vol. 5, Ch. 52, 3159-3228.

[IPCC (2007)] Intergovernmental Panel on Climate Change (2007). *Observations: Surface and Atmospheric Climate Change. In: Climate Change 2007: The Physical Science Basis.* Trenberth, K.E., P.D. Jones, P. Ambenje, R. Bojariu, D. Easterling, A. Klein Tank, D. Parker, F. Rahimzadeh, J.A. Renwick, M. Rusticucci, B. Soden and P. Zha. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

[IPCC (2012)] Intergovernmental Panel on Climate Change (2012). *Summary for Policymakers. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation.* Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.). A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp. 1-19.

[McKinnon (2006)] McKinnon, J. (2006). *Bootstrap Methods in Econometrics.* Queen's University, Department of Economics, Working Paper 1028.

[McKitrick and Vogelsang (2011)] McKitrick, R. and T.J. Vogelsang (2011). *Multivariate trend comparisons between autocorrelated climate series with general trend regressors.* University of Guelph, Department of Economics, Working Paper 1109.

[Smeekes (2009)] Smeekes, S. (2009). *Bootstrapping Nonstationary Time Series.* Universitaire Pers Maastricht.

[Vogelsang (1998)] Vogelsang, T.J. (1998). *Trend function hypothesis testing in the presence of serial correlation.* Econometrica, 66, 123-148.

[WMO (2012)] World Meterological Organization (2012). *Causes of Climate Change.* Retrieved from http://www.wmo.int/pages/themes/climate/causes_of_climate_change.php on 23 June, 2012.