Journal of Environmental Statistics
February 2013, Volume 4, Issue 2. http://www.jenvstat.org

Kernel Regression Model for Total Ozone Data

Horová I., Koláček J., Lajdová D.
Department of Mathematics and Statistics
Masaryk University Brno

Abstract

The present paper is focused on a fully nonparametric regression model for autocorrelation structure of errors in time series over total ozone data. We propose kernel methods which represent one of the most effective nonparametric methods.

But there is a serious difficulty connected with them – the choice of a smoothing parameter called a bandwidth. In the case of independent observations the literature on bandwidth selection methods is quite extensive. Nevertheless, if the observations are dependent, then classical bandwidth selectors have not always provided applicable results. There exist several possibilities for overcoming the effect of dependence on the bandwidth selection. In the present paper we use the results of Chu and Marron (1991) and Koláček (2008) and develop two methods for the bandwidth choice. We apply the above mentioned methods to the time series of ozone data obtained from the Vernadsky station in Antarctica. All discussed methods are implemented in Matlab.

Keywords: total ozone, kernel, bandwidth selection.

1. Introduction

Antarctica is significantly related to many environmental aspects and processes of the Earth. And thus its impact on the global climate system and water circulation in the world ocean is essential.

The stratosphere ozone depletion over Antarctica was discovered at the beginning of the 1990s. The lowest total ozone contents (TOC) in Antarctica are usually observed in the first week of October. The formation of ozone depletion begins approximately in the second half of August, culminates in the first half of October, and dissolves in November. During the ozone depletion, the average ozone concentration varied at the time of its culmination in October from the original value over 300 Dobson Units (DU) in 1950s and 1960s to a level between 100 and 150 DU in 1990-2000 (see Láska *et al.* (2009)). One DU is set as a 0.001 mm strong layer of ozone under the pressure 1013 hPa and temperature 273 K.

One of the issues resolved within the Czech–Ukrainian scientific cooperation implemented on the Vernadsky Station in Antarctica is the measurement of total ozone content (TOC) in the stratosphere. The Vernadsky station is located on the west coast of Antarctic peninsula (65°S, 64°W). These data were obtained from ground measurements predominantly taken with the Dobson No 031 spectrophotometer. Data can be found at UAC (2012).

The data sets were processed as time points measuring the average daily amount of ozone. In order to analyze these data we have to take into account the autocorrelation structure of errors on such time series. We focus on kernel regression estimators of series of ozone data. These estimators depend on a smoothing parameter and it is well-known that selecting the correct smoothing parameter is difficult in the presence of correlated errors. There exist methods which are modifications of a classical cross-validation method for independent errors (the modified cross-validation method or the partitioned cross-validation method - see Chu and Marron (1991), Härdle and Vieu (1992)).

In the present paper we develop a new flexible plug-in approach for estimating the optimal smoothing parameter. The utility of this method is illustrated through a simulation study and application to TOC data measured in periods August to April 2004-2005, 2005-2006, 2006-2007.

2. Procedure Development

2.1. Kernel regression model

In nonparametric regression problems we are interested in estimating the mean function $\mathrm{E}(Y|x) = m(x)$ from a set of observations (x_i, Y_i), $i = 1, \ldots, n$. Many methods such as kernel methods, regression splines and wavelet methods are currently available. The papers in this filed have been mostly focused on case where an unknown function m is hidden by a certain amount of a white noise. The aim of a regression analysis is to remove the white noise and produce a reasonable approximation to the unknown function m.

Consider now the case when the noise is no longer white and instead contains a certain amount of a structure in the form of correlation. In particular, if data sets have been recorded over time from one object under a study, it is very likely that another response of the object will depend on its previous response. In this context we will be dealing with a time series case, where design points are fixed and equally spaced and thus our model takes the form

$$Y_i = m(i/n) + \varepsilon_i, \quad i = 1, \ldots, n, \tag{1}$$

and ε_i is an unknown ARMA process, i.e.,

$$\begin{aligned} \mathrm{E}(\varepsilon_i) &= 0, \quad \mathrm{var}(\varepsilon_i) = \sigma^2, \quad i = 1, \ldots, n, \\ \mathrm{cov}(\varepsilon_i, \varepsilon_j) &= \gamma_{|i-j|} = \sigma^2 \rho_{|i-j|}, \quad \mathrm{corr}(\varepsilon_i, \varepsilon_j) = \rho_{|i-j|} \end{aligned} \tag{2}$$

and the stationary process

$$\gamma_0 = \sigma^2, \quad \rho_t = \frac{\gamma_t}{\gamma_0},$$

where ρ_t is an autocorrelation function and γ_t is an autocovariance function. We consider the simplest situation (Opsomer *et al.* (2001), Chu and Marron (1991))

$$\rho_{t/n} = \rho_t.$$

Simple and the most widely used regression smoothers are based on kernel methods (see e.g. monographs Müller (1987), Härdle (1990), Wand and Jones (1995)). These methods are local weighted averages of the response Y. They depend on a kernel which plays the role of a weighted function, and a smoothing parameter called a bandwidth which controls the smoothness of the estimate.

Appropriate kernel regression estimators were proposed by Priestley and Chao (1972), Nadaraya (1964) and Watson (1964), Stone (1977), Cleveland (1979) and Gasser and Müller (1979).

These estimators were shown to be asymptotically equivalent (Lejeune (1985), Müller (1987), Wand and Jones (1995)) and without the lost of generality we consider the Nadaraya–Watson (NW) estimators $\widehat{m}$ of m. The NW estimator of m at the point $x \in (0,1)$ is defined as

$$\widehat{m}(x,h) = \frac{\sum_{i=1}^{n} K_h(x_i - x) Y_i}{\sum_{i=1}^{n} K_h(x_i - x)}, \tag{3}$$

for a kernel function K, where $K_h(.) = \frac{1}{h}K(\frac{.}{h})$, and h is a nonrandom positive number $h = h(n)$ called the bandwidth.

Before studying the statistical properties of $\widehat{m}$ several additional assumptions on the statistical model and the parameters of the estimator are needed:

I. Let $m \in \mathrm{C}^2[0,1]$.

II. Let K be a real valued function continuous on $\mathbb{R}$ and satisfying the conditions:

- **(i)** $|K(x) - K(y)| \leq L|x - y|$ for a constant $L > 0$, $\forall x, y \in [-1,1]$,
- **(ii)** $\mathrm{support}(K) = [-1,1]$, $\quad K(-1) = K(1) = 0$,
- **(iii)** $\int_{-1}^{1} x^j K(x)\mathrm{d}x = \begin{cases} 1 & j = 0, \\ 0 & j = 1, \\ \beta_2 \neq 0 & j = 2. \end{cases}$

  Such a function is called a kernel of order 2 and a class of these kernels is denoted as S_{02}.

III. Let $h = h(n)$ be a sequence of nonrandom positive numbers, such that $h \to 0$ and $nh \to \infty$ as $n \to \infty$.

IV. $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{\infty} |\rho_k| < \infty$, $\quad$ i.e., $R = \sum\limits_{k=1}^{\infty} \rho_k$ exists,

V. $\frac{1}{n} \sum\limits_{k=1}^{\infty} k|\rho_k| = 0$.

Remark. The well-known kernels are, e.g.,

Epanechnikov kernel $K(x) = \frac{3}{4}(1 - x^2)I_{[-1,1]}$,

quartic kernel $K(x) = \frac{3}{4}(1 - x^2)^2 I_{[-1,1]}$,

triweight kernel $K(x) = \frac{35}{32}(1 - x^2)^2 I_{[-1,1]}$,

Gaussian kernel $K(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{\frac{-x^2}{2}}$,

where $I_{[-1,1]}$ is an indicator function.

Though the Gaussian kernel does not satisfy the assumption II.(ii), it is very popular in many applications.

There is no problem with a choice of a suitable kernel. Symmetric probability density functions are commonly used (see Remark above). But choosing the smoothing parameter is a crucial problem in all kernel estimates. The literature on bandwidth selections is quite extensive in case of independent errors.

It is well known that when the kernel method is used to recover m, that correlated errors trouble bandwidth selection severely (see Altman (1990), Opsomer *et al.* (2001)). De Brabanter *et al.* (2010) developed a bandwidth selection procedure based on bimodal kernels which successfully removes the error correlation without requiring any prior knowledge about its structure.

The global quality of the estimate $\widehat{m}$ can be expressed by means of the Mean Integrated Squared Error (Altman (1990), Opsomer *et al.* (2001)). However more mathematically tractable is the Asymptotic Mean Integrated Squared Error (AMISE):

$$\mathrm{AMISE}(\widehat{m}, h) = \underbrace{\frac{V(K)}{nh} S}_{\mathrm{AIV}(\widehat{m},h)} + \underbrace{\frac{\beta_2^2}{4} h^4 A_2}_{\mathrm{AISB}(\widehat{m},h)},$$

where

$$V(K) = \int K^2(x)\mathrm{d}x, \quad S = \sigma^2(1 + 2\sum_{k=1}^{\infty} \rho_k) = \sigma^2(1 + 2R), \quad A_2 = \int_0^1 m''(x)^2 \mathrm{d}x.$$

The first term is called the asymptotic integrated variance (AIV) and the second one the asymptotic integrated squared bias (AISB). This decomposition provides an easier analysis and interpretation of the performance of the kernel regression estimator.

Using a standard procedure of mathematical analysis one can easily find that the bandwidth h_{opt} minimizing the AMISE is given by the formula

$$h_{opt} = \left(\frac{V(K)S}{n\beta_2^2 A_2}\right)^{1/5} = O(n^{-1/5}). \tag{4}$$

This formula provides a good insight into an optimal bandwidth, but unfortunately it depends on the unknown S and A_2.

Let us explain the impact of assuming an uncorrelated model.

If $R > 0$ (error correlation is positive), then $\text{AIV}(\widehat{m}, h)$ is larger than in the corresponding uncorrelated case and $\text{AMISE}(\widehat{m}, h)$ is minimized by a value h that is larger than in the uncorrelated case. It means that assuming wrongly uncorrelated errors causes that the bandwidth becomes too small.

If $R < 0$ (error correlation is negative), then $\text{AIV}(\widehat{m}, h)$ is smaller and $\text{AMISE}(\widehat{m}, h)$ optimal bandwidth is smaller than in the uncorrelated case.

In the next section the choosing of parameters S and A_2 will be treated.

2.2. Choosing the parameters

There are a number of data-driven bandwidth selection methods, but it can be shown that they fail in the case of correlated errors.

Among the earliest fully automatic and consistent bandwidth selectors are those based on cross-validation ideas. The cross-validation method employs an objective function

$$CV(h) = \frac{1}{n} \sum_{j=1}^{n} \Big(\widehat{m}_{-j}(x_j, h) - Y_j \Big)^2, \tag{5}$$

where $\widehat{m}_{-j}(x_j, h)$ is the estimate of $\widehat{m}(x_j, h)$ with x_j deleted, i.e., the leave-one-out estimator. The estimate of h_{opt} is then

$$\widehat{h}_{opt} = \arg \min_{h \in H_n} CV(h),$$

where $H_n = [an^{-1/5}, bn^{-1/5}], \quad 0 < a < b < \infty.$

Remark. If the design points are equally spaced then a recommended interval is $[\frac{1}{n}, 1)$.

However, this ordinary method is not suitable in the case of correlated observations. As it was shown in the papers Altman (1990) and Opsomer *et al.* (2001), if the observations are positively correlated, then the CV method produces too small a bandwidth, and if the observations are negatively correlated, then the CV method produces a large bandwidth.

We demonstrate this fact by the following example.

Consider the regression model (1), where

$$m(x) = \cos(3.15\pi x), \quad \varepsilon_i = \phi\varepsilon_{i-1} + e_i,$$

e_i – i.i.d. normal random variables $\text{N}(0, \sigma^2)$,

ε_1 – $\text{N}(0, \sigma^2/(1 - \phi^2))$,

$\phi = 0.6, \quad \sigma = 0.5,$

i.e, the regression errors are AR(1) process.

Figure 1 shows the result obtained by the CV method. It is evident, that the estimate is undersmoothed.

In order to overcome this problem, modified and partitioned CV methods were proposed by Härdle and Vieu (1992) and Chu and Marron (1991), respectively.

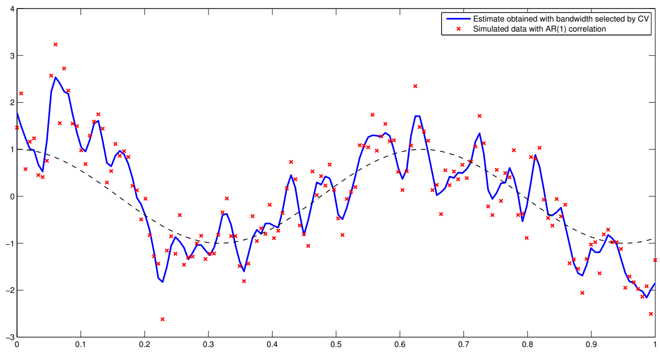

Figure 1: The estimate of simulated data with AR(1) errors

The modified cross-validation (MCV) method is a "leave-$(2l+1)$-out" version of CV $(l \geq 0)$. The idea consists in minimizing of the modified cross-validation score:

$$CV_l(h) = \frac{1}{n}\sum_{j=1}^{n}\Big(\widehat{m}_{-j}(x_j,h) - Y_j\Big)^2, \tag{6}$$

where $\widehat{m}_{-j}(x_j,h)$ is the "leave-$(2l+1)$-out" estimate of $\widehat{m}(x_j,h)$, i.e., the observations (x_{j+i}, Y_{j+i}), $-l \leq i \leq l$ are left out in constructing $\widehat{m}(x_j,h)$.
Then

$$\widehat{h}_{MCV} = \arg\min_{h\in H_n} CV_l(h).$$

The principle of the partitioned cross-validation method (PCV) can be described as follows. For any natural number $g \geq 1$, the PCV involves splitting the observations into g groups by taking every g-th observation, calculating the ordinary cross-validation score $CV_{0,k}(h)$ of the k-th group of observations separately, for $k = 1, 2, \ldots, g$, and minimizing the average of these ordinary cross-validation scores

$$CV^*(h) = \frac{1}{g}\sum_{k=1}^{g} CV_{0,k}(h). \tag{7}$$

Let $\widehat{h}^*_{CV}$ stand for the minimizer of $CV^*(h)$:

$$\widehat{h}^*_{CV} = \arg\min_{h\in H_n} CV^*(h).$$

Since $\widehat{h}^*_{CV}$ is appropriate for the sample size n/g, the partitioned cross-validated bandwidth $\widehat{h}_{PCV(g)}$ is defined to be rescaled $\widehat{h}^*_{CV}$:

$$\widehat{h}_{PCV(g)} = g^{-1/5}\widehat{h}^*_{CV}.$$

When $g = 1$, the PCV is an ordinary cross-validation.

Remark. The number of subgroups is g and the number of observations in each group is $\eta = n/g$. If n is not a multiplier of g, then the values Y_j, $1 \leq j \leq g[n/g]$ are applied and the rest of the observations are dropped out ($[n/g]$ is the highest integer less or equal to n/g).

The asymptotic behavior of $\widehat{h}_{MCV(l)}$ and $\widehat{h}_{PCV(g)}$ was studied in the paper by Chu and Marron (1991). Furthemore we focus on the PCV method.

The PCV method needs to determine the factor g. A possible approach for the practical choice of g is based on an analogue of the mean squared error. Using the asymptotic variance and the asymptotic mean of $\widehat{h}_{PCV(g)}/h_{opt}$, the asymptotic mean squared error (AMSE) of this ratio is defined by

$$AMSE\big(\widehat{h}_{PCV(g)}/h_{opt}\big) = n^{-1/5}\mathrm{VAR}_{PCV(g)} + \big[C_{PCV(g)}/C - 1\big]^2, \tag{8}$$

where $\mathrm{VAR}_{PCV(g)}$, $C_{PCV(g)}$, C depend on γ_k, K, A_2 (see Chu and Marron (1991)).

Theoretically, if there exists a value $\widehat{g}$ which minimizes AMSE over $g \geq 1$, then this value is taken as the optimal value of g in the sense of AMSE:

$$\widehat{g}_{opt} = \arg\min_{g\geq 1} \mathrm{AMSE}\big(\widehat{h}_{PCV(g)}/h_{opt}\big).$$

Unfortunately the minimization of AMSE also depends on the unknown γ_k and A_2.

As far as the estimation of the variance component S is concerned, a common approach is the following (see e.g. Herrmann *et al.* (1992), Hart (1991), Opsomer *et al.* (2001), Chu and Marron (1991)):

$$\begin{aligned}\widehat{S} =& \hat{\gamma_0}\Big(1 + 2\sum_{k=1}^{n-1}\hat{\rho_k}\Big), \quad \hat{\gamma_0} = \hat{\sigma}^2, \quad \hat{\rho_k} = \frac{\hat{\gamma_k}}{\hat{\gamma_0}},\\ \hat{\gamma_k} =& \frac{1}{n-k}\sum_{t=1}^{n-k}\Big(Y_t - \overline{Y}\Big)\Big(Y_{t+k} - \overline{Y}\Big), \quad k = 0, \dots, n-1.\end{aligned} \tag{9}$$

Nevertheless there is still a problem of how to estimate A_2. In paper Chu and Marron (1991) a simulation study was only conducted and no idea of estimating A_2 was given there.

We complete this method by adding a suitable estimate of A_2 and recommend to use an estimate of A_2 proposed by Koláček (2008). By means of the Fourier transformation he derived a suitable estimate $\widehat{A_2}$ of A_2. Therefore, A_2 in the AMSE formula is replaced by $\widehat{A_2}$. This approach is commonly known as a plug-in method.

Plug-in methods are also commonly used for selecting the bandwidth in the kernel regression. But these methods perform badly when the errors are correlated. In the paper Herrmann *et al.* (1992) a modified version of an existing plug-in bandwidth selectors is proposed. This method is based on the Gasser–Müller estimator of the second derivative and an iterative process is constructed. It is shown that under some additional assumptions this iterative process converges to a suitable estimate of the optimal bandwidth.

However we do not use this iterative method and propose to directly plug-in A_2 in the formula (4). This new version of a plug-in method is denoted as PI and the bandwidth estimate takes the form:

$$\widehat{h}_{PI} = \Big(\frac{V(K)\widehat{S}}{n\beta_2^2\widehat{A_2}}\Big)^{1/5}.$$

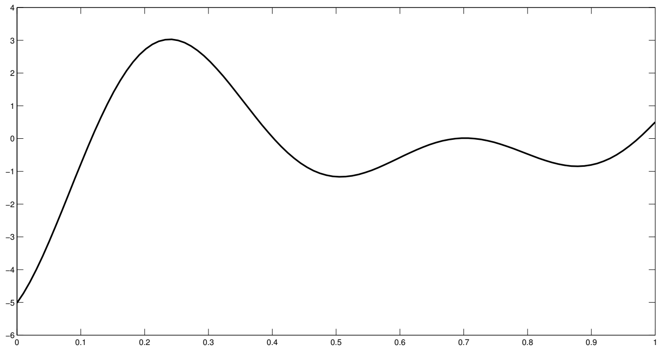

Figure 2: The regression function $m(x)$

$h_{opt} = 0.759$		
	$\mathrm{E}(\widehat{h})$	$\mathrm{std}(\widehat{h})$
PCV	0.1927	0.0649
PI	0.1513	0.0083

Table 1: The estimates $\widehat{h}$

We would like to point out the computational aspect of the plug-in method. It has preferable properties to classical methods, because it does not need any additional calculations such as the PCV method (see Koláček (2008) for details).

3. Case study

We conduct a simulation study to compare the PCV method and the PI method. The Epanechnikov kernel is used both in simulations and in applications.

Consider the regression model (1), where

$$m(x) = \frac{-6\sin 11x + 5}{\mathrm{cotg}(x-7)}, \quad \varepsilon_i = \phi\varepsilon_{i-1} + e_i$$

$$e_i - \text{i.i.d. normal random variables } \mathrm{N}(0, \sigma^2)$$

$$\varepsilon_1 - \mathrm{N}(0, \sigma^2/(1-\phi^2))$$

$$\phi = 0.6, \quad \sigma = 0.5,$$

for $i = 1, \ldots, n = 100$.

The graph of the regression function m is presented in Figure 2.

One hundred series are generated. For each data set, the optimal bandwidth is estimated by the PCV and PI method. Table 1 shows the comparison of means and standard deviations for these two methods.

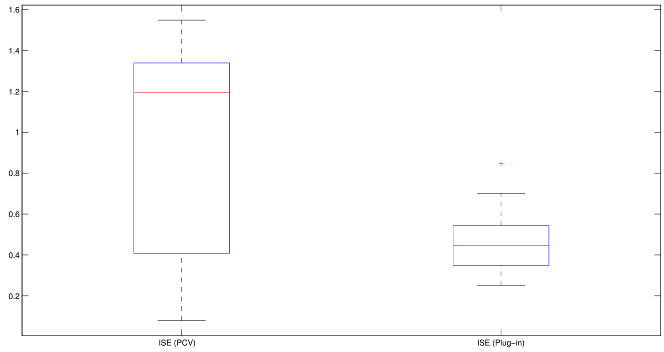

Figure 3: $ISE(\widehat{m}(.,h)) = \int_0^1 \Big(\widehat{m}(x,h) - m(x)\Big)^2 \mathrm{dx}$.

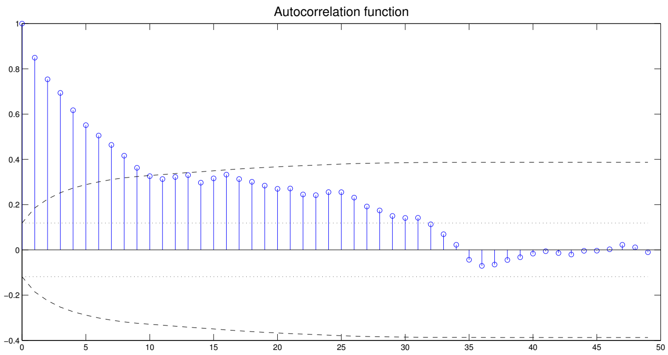

Figure 4: The autocorrelation function of the data set August 2004 – April 2005

The Integrated Square Error (ISE) is calculated for each estimate $\widehat{m}(.,h)$:

$$ISE(\widehat{m}(.,h)) = \int_0^1 \Big(\widehat{m}(x,h) - m(x)\Big)^2 \mathrm{dx}$$

for both PCV and PI methods and the results are displayed by means of the boxplots in Figure 3.

4. Results and discussion

In this section we apply the methods described above to ozone data. We analyze data which were measured in the period August to April in years 2004–2005, 2005–2006, 2006–2007. The sample size is $n = 273$ days. The observations are correlated as it can be seen in Figure 4. We transform data to the interval [0,1] and use the PCV method and the PI method to get the optimal bandwidth. Then we re-transform the bandwidth to the original sample and obtain the final kernel estimate.

Kernel estimates based on the PCV and PI methods are presented in Figure 6, Figure 7, or in Figure 8, respectively.

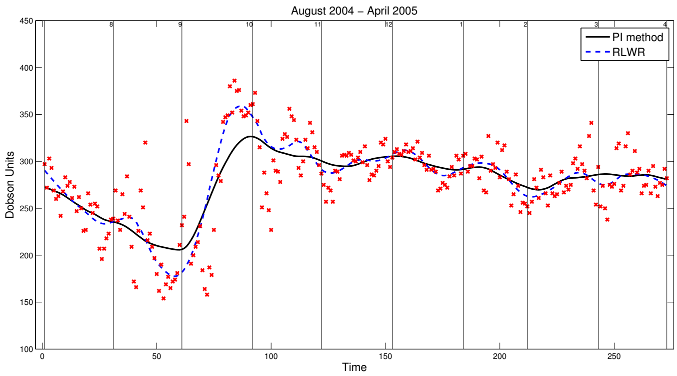

Figure 5: RLWR estimate with span = 40 (dashed line) and PI estimate with the bandwidth = 17.8 (solid line).

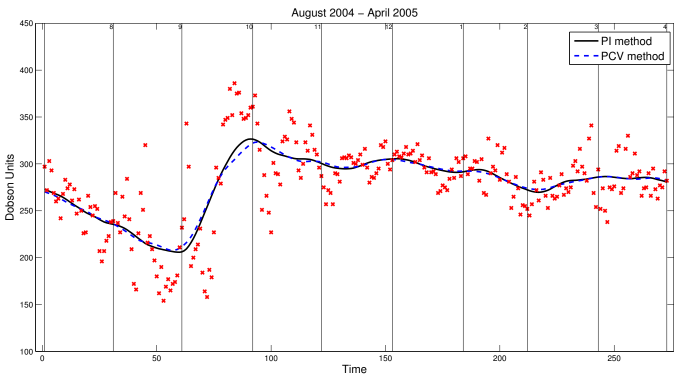

Figure 6: PCV estimate with the bandwidth = 20.9 (dashed line) and PI estimate with the bandwidth = 17.8 (solid line).

In paper Kalvová and Dubrovský (1995) the robust locally wighted regression (RLWR) is employed for data processing of TOC. They recommended to optimize h subjectively. This approach needs an experience and a special knowledge of the given data sets. The advantage of our methods consists in more complex approach. These methods are general and they allow to choose the value of h automatically. We used their methodology for data April 2004 - August 2005 and the comparison of the estimate obtained by the PI method and by the robust locally weighted regression can be seen in Figure 5. The PI method yields a rather oversmoothed estimate.

Our experience shows that both methods could be considered as a suitable tool for the choice of the bandwidth. But it seems that the PI method is sufficiently reliable and less time consuming than the PCV method.

Presented methods can be applied to other time series not only in environmetrics but also in economics or other fields.

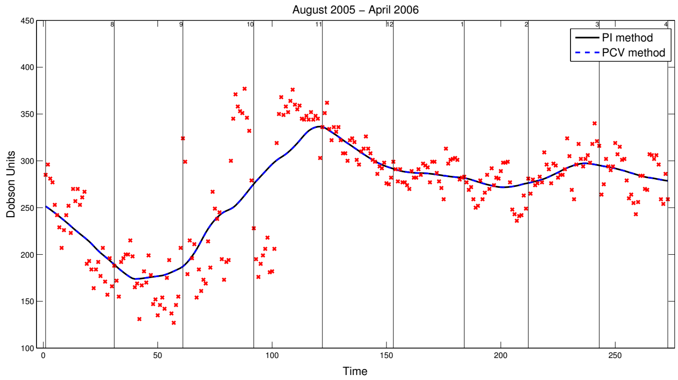

Figure 7: PCV estimate with the bandwidth = 20.4 (dashed line) and PI estimate with the bandwidth = 21.9 (solid line).

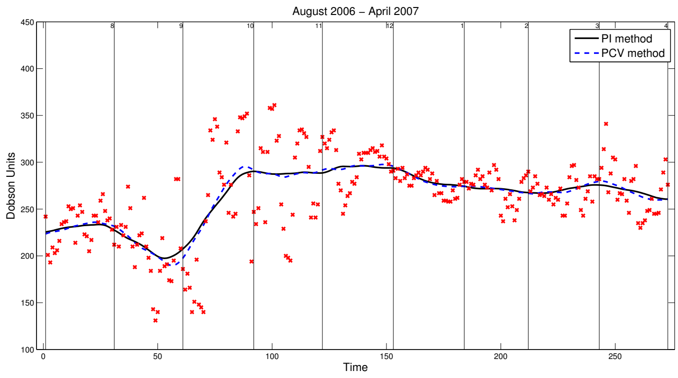

Figure 8: PCV estimate with the bandwidth = 17.2 (dashed line) and PI estimate with the bandwidth = 22.3 (solid line).

Acknowledgments

The research was supported by The Jaroslav Hájek Center for Theoretical and Applied Statistics (MŠMT LC 06024). The work was supported by the Student Project Grant at Masaryk university, rector's programme no. MUNI/A/1001/2009.

References

Altman N (1990). "Kernel Smoothing of Data With Correlated Errors." *Journal of the American Statistical Association*, **85**, 749–759.

Chu CK, Marron JS (1991). "Choosing a Kernel Regression Estimator." *Statistical Science*, **6**(4), 404–419. ISSN 08834237.

Cleveland WS (1979). "Robust Locally Weighted Regression and Smoothing Scatterplots." *Journal of the American Statistical Association*, **74**(368), 829–836. ISSN 01621459.

De Brabanter K, De Brabanter J, Suykens J, De Moor B (2010). "Kernel Regression with Correlated Errors." *Computer Applications in Biotechnology*, pp. 13–18.

Gasser T, Müller HG (1979). "Kernel estimation of regression functions." In T Gasser, M Rosenblatt (eds.), *Smoothing Techniques for Curve Estimation*, volume 757 of *Lecture Notes in Mathematics*, pp. 23–68. Springer Berlin / Heidelberg.

Härdle W (1990). *Applied Nonparametric Regression.* 1st edition. Cambridge University Press, Cambridge.

Härdle W, Vieu P (1992). "Kernel Regression Smoothing of Time Series." *Journal of Time Series Analysis*, **13**(3), 209–232.

Hart JD (1991). "Kernel Regression Estimation with Time Series Errors." *Journal of the Royal Statistical Society*, **53**, 173–187.

Herrmann E, Gasser T, Kneip A (1992). "Choice of Bandwidth for Kernel Regression when Residuals are Correlated." *Biometrika*, **79**, 783–795.

Kalvová J, Dubrovský M (1995). "Assessment of the Limits Between Which Daily Average Values of Total Ozone Can Normally Vary." *Meteorol. Bulletin*, **48**, 9–17.

Koláček J (2008). "Plug-in Method for Nonparametric Regression." *Computational Statistics*, **23**(1), 63–78. ISSN 0943-4062.

Láska K, Prošek P, Budík L, Budíková M, Milinevsky G (2009). "Prediction of Erythemally Effective UVB Radiation by Means of Nonlinear Regression Model." *Environmetrics*, **20**(6), 633–646.

Lejeune M (1985). "Estimation Non-paramétrique par Noyaux: Régression Polynomiale Mobile." *Revue de Statistique Appliquée*, **33**(3), 43–67.

Müller HG (1987). "Weighted Local Regression and Kernel Methods for Nonparametric Curve Fitting." *Journal of the American Statistical Association*, **82**(397), 231–238. ISSN 01621459.

Nadaraya EA (1964). "On Estimating Regression." *Theory of Probability and its Applications*, **9**(1), 141–142.

Opsomer J, Wang Y, Yang Y (2001). "Nonparametric Regression with Correlated Errors." *Statistical Science*, **16**(2), 134–153.

Priestley MB, Chao MT (1972). "Non-Parametric Function Fitting." *Journal of the Royal Statistical Society. Series B (Methodological)*, **34**(3), 385–392. ISSN 00359246.

Stone CJ (1977). "Consistent Nonparametric Regression." *The Annals of Statistics*, **5**(4), 595–620. ISSN 00905364.

UAC (2012). "World Ozone and Ultraviolet Radiation Data Centre (WOUDC) [data]." URL http://www.woudc.org.

Wand M, Jones M (1995). *Kernel smoothing*. Chapman and Hall, London.

Watson GS (1964). "Smooth Regression Analysis." *Sankhya - The Indian Journal of Statistics, Series A*, **26**(4), 359–372. ISSN 0581572X.

Affiliation:

Ivana Horová
Masaryk University
Department of Mathematics and Statistics
Brno, Czech Republic
E-mail: horova@math.muni.cz
URL: https://www.math.muni.cz/~horova/

Journal of Environmental Statistics http://www.jenvstat.org
Volume 4, Issue 2 Submitted: 2012-03-31
February 2013 Accepted: 2012-10-09