An Application of a Fuzzy Smoothing Filter for FX Rates

Michal Holčapek and Tomáš Tichý

Abstract— **Data smoothing is an important step within a data processing procedure that allows one to stress the most important pattern of a function relation between a studied object and given variables. Recently, Holčapek and Tichý (2011) suggested a smoothing filter based on fuzzy transform approach of Perfilieva (2004) and compared it to Nadaraya-Watson estimator. In this contribution, we provide some results on smoothing of FX rates.**

Index Terms—**Fuzzy smoothingfilter, FX rates, volatility.**

I. Introduction

Many issues of financial modeling require to present the price evolution in time as well as its probability distribution. Since it is often difficult to asses the problem quickly according to raw data, which can potentially consists of several thousand observations, and calculation of basic descriptive statistics might lead to too simplifying conclusions, it can be useful to provide a chart as a first insight into the problem.

Obviously, if we know the model, which is followed by a given random variable, ie. the market price, we can estimate its parameters and draw a function. Such approach is called parametrical. Within the real world problems, however, we cannot be sure about the properness of a given model. Since the evolution at financial markets is strongly related to the psychology of market participants, it can happen that a model, which has been identified as valid at one time instant, will not prove to be reliable at next time instant.

A natural way how to provide a first picture about the data therefore is a non-parametric smoothing. The most standard way of smoothing financial data is based on kernel regression. As an alternative, one can consider a smoothing filter based on fuzzy-transform approach of Perfilieva (2006) [1]. Both these approaches will be applied in the following text on a chosen part of FX rate time series aiming on their visual as well as computational comparison.

II. Data Smoothing

Let us assume observations of random variable, such as market price of foreign currency or equity index, with fixed length among particular observations, $\{S_t\}_{t=0}^{T}$. When dealing with such kind of data in finance, we might be interested in approximation of:

1) the evolution of the price in time, ie. a functional relationship of S_t and t;
2) the evolution of price returns x_t, $x_t = \ln \frac{S_t}{S_{t-1}}$;
3) the probability density function of returns$\{x_t\}_{t=1}^{T}$.

Specifically, the reason why we might be interested in (b) is that the evolution of price returns in time is more comparative than the price itself; next, the returns can be further standardized, eg. to get zero mean and unit variance. It follows, that also in (c) we are interested in the distribution of returns. Moreover, since many financial applications crucially depends on tails, it may be useful to transform the probability function f into ln$f$.

Generally speaking, we need to study variable Y_t in dependency on X_t. Obviously, in the most simple case, X_t becomes t so that Y_t is a price dependent solely on the time. To explore the functional relationship of both variables, we usually take the advantage of non-parametric methods, since there is no presumption about the kind of the dependency. The main disadvantages, by contrast, are strong data dependency and potential over fitting.

In the following sections we first review basic terminology so that we can later define probably the most popular approach to data smoothing – the kernel regression. After that an alternative approach of fuzzy-transform is presented. Both these two approaches are applied in order to provide illustrative examples on financial data. Finally, some further alternatives, including references, are provided.

III. Basic Methodology

Assuming two variables, Y_t and X_t, their relationship in time can be described as follows:

$$Y_t = g\{X\} + \epsilon_t. \quad (1)$$

where g is arbitrary, continuous, smooth, but unknown function and ϵ is a white noise.

The core idea of smoothing is that by the law of large numbers repeating the observation of Y, $\{Y = y_i\}_{i=1}^{n}$ 1, n-times for a fixed $X = x$ with sufficiently large n we get:

$$\frac{\sum_n y_i}{n} = g\{x\} + \frac{\sum_n \epsilon_i}{n} \quad (2)$$

Obviously, any sufficiently large white noise sequence must have zero mean. It means that the left-hand side of (2) is a consistent estimate of $g(x)$.

Since it is not feasible to repeat the observations at a given time when dealing with financial time series, we get the estimate by using values of any X_t located nearby to x. It is also natural that observations located further from x are less

Manuscript received February 28, 2013; revised June 28, 2013. This work was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070) and the European Social Fund (CZ.1.07/2.3.00/20.0296). The research of the second author was supported by SGS project of VSB-TU Ostrava under No. SP2013/3 and Czech Science Foundation through project No. 13-13142S. The support is greatly acknowledged.

The authors are with the Department of Finance, Technical University Ostrava, 701 21 Czech Republic (e-mail: michal.holcapek@vsb.cz, tomas.tichy@vsb.cz).

DOI: 10.7763/IJIET.2013.V3.339

important. A standard approach, how to take this natural consequence into account, is to replace the simple average (2) by a weighted one:

$$\hat{g}\{x\} = \sum_t w_t(x) y_t \quad (3)$$

where w_t already includes $1/T$ and therefore sum to unity. Particular smoothing approaches differ due to the ways to calculate the distance between x_t and x and its subsequent transformation into a weight measure.

IV. Smoothing via Kernel Regression

Kernel regression, probably the most common approach to smoothing of financial data, uses a kernel to assign the weights to particular observations. Typically, it is a probability density function, $K(x)$, satisfying:

$$K(\mathrm{x}) \geqslant 0, \int K(z)\mathrm{d}z = 1 \quad (4)$$

It is often useful to rescale (4) by a constant h called bandwidth as follows:

$$K_h(x) = 1/hK(x/h) \quad (5)$$

Finally, the weights can be obtained as follows:

$$w_t(x) = \frac{K_h(x-x_t)}{\sum_t K_h(x-x_t)} \quad (6)$$

This leads to Nadaraya-Watson kernel estimator [2], [3]:

$$\hat{g}(x) = \sum_t w_t(x) y_t = \frac{K_h(x-x_t)y_t}{\hat{g}\{x\}} \quad (7)$$

Useful choices of kernels are either Gaussian kernel,

$$K_h(x) = \frac{1}{h\sqrt{2\pi}} \exp\left(-\frac{x^2}{2h^2}\right) \quad (8)$$

or Epanechnikov kernel [4],

$$K_h(x) = \frac{0.75}{h}\left(1 - \frac{x^2}{h^2}\right) I_{|x/h|\leq 1} \quad (9)$$

where I is the indicator function.

Selection of the proper bandwidth h is even more importantthan the choice of the kernel. Observe what happens in theextreme cases:

1) If h is very large, potentially going to infinity, all observationshas the same weight so that we receive a singleline – a sample mean. This is called *oversmoothing*.
2) By contrast, as h approach to zero, the level of smoothingwill be decreasing and finally, we get the original curve.Moreover, if we increase h, the Bias of the estimation willincrease, too. Similarly, with lowering h, we loose the opportunityto decrease the variance of the estimator.

One approach, how to select optimal bandwidth h isbased on mean integrated square error criterion (MISE),which, according to [5], leads to $h_{Gauss}^{opt} = 1.06sT^{-1/5}$ or $h_{Epn}^{opt} = 2.34sT^{-1/5}$ for Gaussian and Epanechnikov kernel, respectively, with T being the length of the data as usually and the sample standard deviation.

V. Smoothing Based on Fuzzy Transformation

Sometimes dealing with the data in the original spaceis complicated. It might be therefore fruitful to transformthe data into a different space, in which the data handlingis simpler. After constructing the approximation, themodel is transformed back to the original space via inversetransformation.

In this section we will focus on the technique of fuzzytransform originally proposed by [6], [1] as a tool for imageprocessing. Recently, in [7] it was suggested to use this method forfinancial time series smoothing. Moreover, in [8] and [9] theconditions for corresponding behavior of fuzzy transform andkernel estimator were derived and subsequently applied onsmoothing of FX rate series.

Following [1], when we apply fuzzy transform approach, aswe can guess from the term fuzzy, the independent variables are fuzzy fied according to the proximity to a given point, which can be compared to the concept of weights obtained viaprobability distribution function in kernel regression approach.

Next, the observations of the dependent variable are averaged, which forms a functional relation for a given point. This stepof fuzzy transform is called direct fuzzy transform. Obviously, the second step is inverse fuzzy transform, within which wereturn back to the original crisp space and obtain a smoothedfunction describing the relation of both variables.Bellow, we will focus on the relation between fuzzytransformation and kernel smoothing in line with[1] and [8].

Assuming that R is a real interval, g is a finite real functiongiven at the nodes $x_1 < \cdot\cdot\cdot < x_n$ with Dom(g) $\subseteq R$ and$A = \{A_i | i \in I\}$ is a fuzzy r-partition of R determined by(T, S) such that Dom(g) is sufficiently dense with respect to$A$, one can say that a collection of real numbers $\{F_i | i \in I\}$is discrete fuzzy (F-)transform of g with respect to A, if

$$F_i = \frac{\sum_{j=1}^{n} g(x_j) A_i(x_j)}{\sum_{j=1}^{n} A_i(x_j) h} \quad (10)$$

The numbers F_i are called components of the discrete F-transform.

Moreover, if g is a real function as above, then F_i minimizesthe weighted least square criterion

$$\Phi_i(y) = \sum_{j=1}^{n} \left(g(x_j) - y\right)^2 A_i(x_j) \quad (11)$$

Thus, FT-smoothing filter determined by A is a mappingF_A : F(I_n,A) $\rightarrow CF(A)$ defined by

$$\mathrm{F}_A(g)(x) = \frac{1}{r} \sum_{i \in I} F_i A_i(x) \quad (12)$$

for any $x \in R$, where F_k are the components of the discrete Ftransformand the set of all continuous real functions g definedon R.

VI. Comparison Examples

Applying the asymptotic mean square error (AMSE) of the form

$$AMSE\big(\hat{g}_{FT}(x)\big) = \frac{k^2\sigma^2}{r^2nh(b-a)}R(K) + \frac{h^4\mu_2(K)^2}{r^2u^2}g''(x)^2 \quad (13)$$

the optimal value of bandwidth h can be derived by settingits derivative with respect to h equal to zero. Thus, withelementary calculus one can obtain

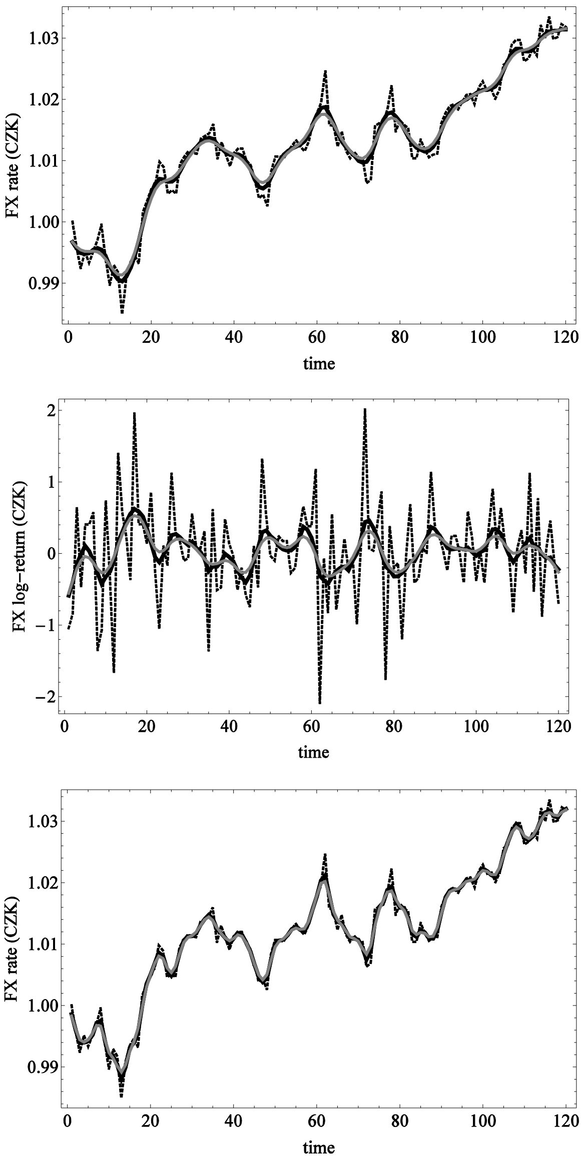
$$h_{AMSE} = \frac{k^2u^2\sigma^2}{rng''(x)^2(b-a)}C(K)^{1/5} \quad (14)$$

In the formulas above, K is a symmetric Kernel,$\mu_2(K) = \int_{-1}^{1} z^2K(z)dz$, $R(K) = \int_{-1}^{1} K^2(z)dz$, $C(K) = \frac{R(K)}{\mu_2(K)^2}$, $u = t_{i+1} - t_i$ and k denotes the number ofbasic functions obtained by a kernel. Since $t_i \notin R$ for some$i = 1,\ldots, k$, we can deduce $ku \geqslant b-a$. One can notice that$ku \approx b-a$ for small h and

$$h_{AMSE} = \frac{(b-a)^2\sigma^2}{rng''(x)^2(b-a)}C(K)^{1/5} = \frac{(b-a)\sigma^2}{rng''(x)^2}C(K)^{1/5} \quad (15)$$

Comparing this result of [8] with the well known result for theNadaraya-Watson estimator, we get the following approximaterelation:

$$h^{NW}_{AMSE} \approx \frac{h^{FT}_{AMSE}}{0.76} \quad (16)$$



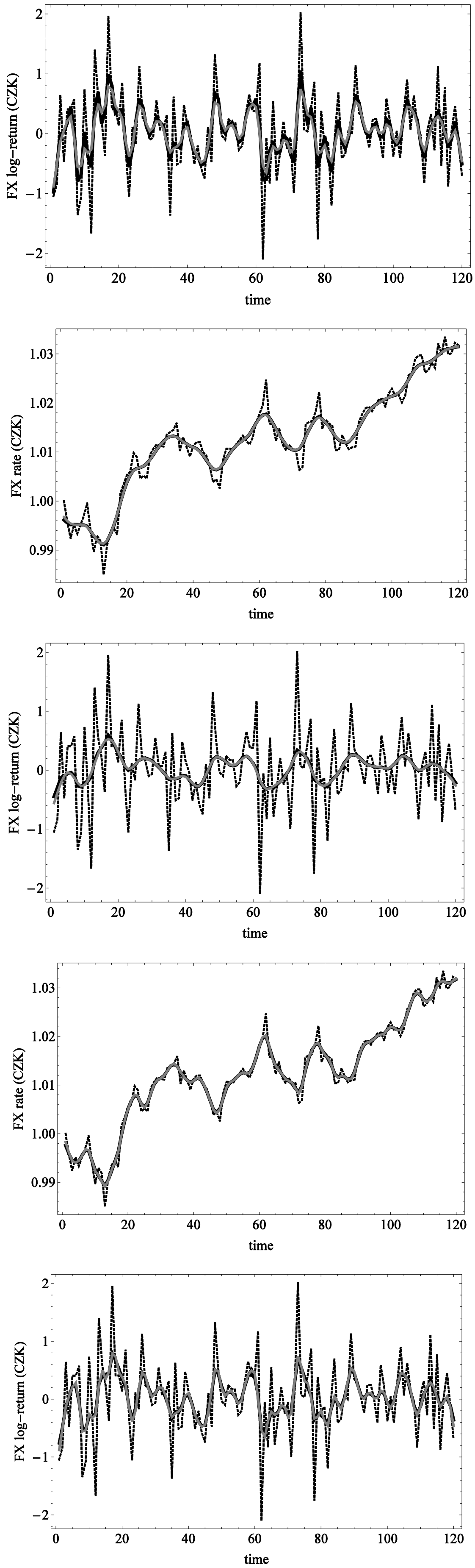


Fig. 1. Fuzzy-transform (black) and Nadaraya-Watson (gray) filterappliedwiththesamebandwidth (left) and adjustedbandwidth (right) to FX ratelogreturnsof CZK (hFT= 2 atthe top and hFT= 4 atthebottom); original data depicted by dottedcurve.

where $0,76 \approx \sqrt[5]{1/4}$. Hence, we need theoretically a higher bandwidth to obtain an optimal model for unknown function.

On Fig. 1 we provide some illustration of both filters (Nadaraya-Watson and fuzzy-transform), and specifically the effect of various h, when applied to daily evolution of FX rate of CZK in terms of EUR over the period of about six months (first half of 2001). Charts in the first and third line provides the results obtained for the normalized FX rate, ie. it is set equal to 1.0 at the beginning, while the second and fourth line shows the results for daily log-returns.

In all cases, dotted black lines show the original data while solid gray and black lines are used for smoothed curve according to the Nadaraya-Watson estimator and fuzzy-transform, respectively. Comparing the left ($h_{NW} = h_{FT}$) and right panel ($h_{NW} = h_{FT}/0.76$), one can see the effect of the previous result on the relation between optimal h of both filters (as based on AMSE criterion). Note moreover, that the first two lines were obtained assuming basic $h = 2$, while the last two are for $h = 4$.

VII. Conclusion

Following the application results, we can clearly see that under certain circumstances both approaches can lead to equivalent results. However, Holčapek and Tichý [8] argued, that one should prefer fuzzy-transform estimator if a storage problem may arise. By contrast, standard approach due to Nadaraya-Watson is easier to understand and thus there is a low risk of inappropriate parameters selection.

Acknowledgement

This work was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070). The research of the second author was supported by SGS project of VSB-TU Ostrava under No. SP2013/3 and Czech Science Foundation through project No. 13-13142S.

References

[1] I. Perfilieva, "Fuzzy transforms: Theory and applications," *Fuzzy setssyst.*, vol. 157, no. 8, pp. 993–1023, 2006.

[2] E. Nadaraya, "On estimating regression," *TheoryProbab. Appl.*, vol. 9, pp. 141–142, 1964.

[3] G. Watson, "Smooth regression analysis," *Sankhya, Series A*, vol. 26, pp. 359–372, 1964.

[4] V. Epanechnikov, "Non parametric estimates of a multivariate probability density," *Theoryof Probability and itsapplications*, vol. 14, pp. 153–158, 1969.

[5] J. Fan and Q. Yao, "NonlinearTimeSeries: Nonparametric and Parametric Methods, ser," *Springer Series in Statistics*, Springer-Verlag, 2005.

[6] I. Perfilieva, "Fuzzy transforms," in J. F. Peters *et al.*, ed., *Transactions on RoughSets II. Roughsets and fuzzy sets*, *Lecture Notes in Computer Science 3135, Journal Subline*, Berlin: Springer, pp. 63-81, 2004.

[7] M. Holčapek and T. Tichý, "A probability density function estimation using F-transform," *Kybernetika*, vol. 46, no. 3, pp. 447– 458, 2010.

[8] M. Holčapek "A smoothing filter based on fuzzy transform," *Fuzzy Sets* Syst., vol. 180, no. 1, pp. 69–97, 2011.

[9] M. Holčapek, "Statistical analysis of a smoothing filter based on fuzzy transform," in *Proc. the 7th conferenceoftheEuropean Society for Fuzzy Logic and Technology (EUSFLAT-2011) and LFA-2011, ser. Advances in Intelligent Systems Research*, Atlantis Press, 2011, pp. 472–479.

Michal Holčapek obtained his Ph.D in mathematics at University of Ostrava, Czech Republic and currently focuses on fuzzy mathematics and logic, category theory, probability and statistics.
He commonly publishes research papers in top journals such as Fuzzy sets and systems or Information science and serve as a reviewer there as well.

Tomáš Tichý obtained his Ph.D in finance at Technical University Ostrava, Czech Republic and currently focuses on risk management and modeling of financial quantities via various kinds of statistical tools.
Currently, he serves as Editor or Editor-in-Chief of several journals that focuses especially on economic issues of Central Europe.