Time series technical analysis via new fast estimation methods: a preliminary study in mathematical finance

Michel FLIESS *,** **Cédric JOIN** *,***

* *INRIA-ALIEN*
** *LIX (CNRS, UMR 7161), École polytechnique*
91128 Palaiseau, France
Michel.Fliess@polytechnique.edu
*** *CRAN (CNRS, UMR 7039), Nancy-Université*
BP 239, 54506 Vandœuvre-lès-Nancy, France
Cedric.Join@cran.uhp-nancy.fr

Abstract: New fast estimation methods stemming from control theory lead to a fresh look at time series, which bears some resemblance to "technical analysis". The results are applied to a typical object of financial engineering, namely the forecast of foreign exchange rates, via a "model-free" setting, *i.e.*, via repeated identifications of low order linear difference equations on sliding short time windows. Several convincing computer simulations, including the prediction of the position and of the volatility with respect to the forecasted trendline, are provided. $\mathcal{Z}$-transform and differential algebra are the main mathematical tools.

Keywords: Time series, identification, estimation, trends, noises, model-free forecasting, mathematical finance, technical analysis, heteroscedasticity, volatility, foreign exchange rates, linear difference equations, $\mathcal{Z}$-transform, algebra.

1. INTRODUCTION

1.1 Motivations

Recent advances in estimation and identification (see, *e.g.*, (Fliess & Sira-Ramírez, 2003; Fliess & Sira-Ramírez, 2004; Fliess, Join & Sira-Ramírez, 2004; Fliess & Sira-Ramírez, 2008; Fliess, Join & Sira-Ramírez, 2008) and the references therein) stemming from mathematical control theory may be summarized by the two following facts:

- Their algebraic nature permits to derive exact non-asymptotic formulae for obtaining the unknown quantities in real time.
- There is no need to know the statistical properties of the corrupting noises.

Those techniques have already been applied in many concrete situations, including signal processing (see the references in (Fliess, 2008)). Their recent and successful extension to discrete-time linear control systems (Fliess, Fuchshumer, Schöberl, Schlacher & Sira-Ramírez, 2008) has prompted us to study their relevance to financial time series.

Remark 1.1. The relationship between time series analysis and control theory is well documented (see, *e.g.*, (Box, Jenkins & Reinsel, 1994; Gouriéroux & Monfort, 1995; Hamilton, 1994) and the references therein). Our viewpoint seems nevertheless to be quite new when compared to the existing literature.

Remark 1.2. The title of this communication is due to its obvious connection with some aspects of *technical analysis*, or *charting* (see, *e.g.*, (Aronson, 2007; Béchu, Bertrand & Nebenzahl, 2008; Kaufman, 2005; Kirkpatrick & Dahlquist, 2006; Murphy, 1999) and the references therein), which is widely used among traders and financial professionals.[1]

1.2 Linear difference equations

Consider the univariate time series $\{x(t) \mid t \in \mathbb{N}\}$: $x(t)$ is not regarded here as a stochastic process like in the familiar ARMA and ARIMA models but is supposed to satisfy "approximatively" a linear difference equation

$$x(t+n) - a_1 x(t+n-1) - \cdots - a_n x(t) = 0 \qquad (1)$$

where $a_1, \ldots, a_n \in \mathbb{R}$. Introduce as in digital signal processing the additive decomposition

$$x(t) = x_{\text{trendline}}(t) + \nu(t) \qquad (2)$$

where

- $x_{\text{trendline}}(t)$ is the *trendline*[2] which satisfies Eq. (1) exactly;
- the additive "noise" $\nu(t)$ is the mismatch between the real data and the trendline.

Thus

$$x(t+n) - a_1 x(t+n-1) - \cdots - a_n x(t) = \epsilon(t) \qquad (3)$$

where

$$\epsilon(t) = \nu(t+n) - a_1 \nu(t+n-1) - \cdots - a_n \nu(t) \qquad (4)$$

[1] Technical analysis is often severely criticized in the academic world and among the practitioners of mathematical finance (see, *e.g.*, (Paulos, 2003)).

[2] Compare, *e.g.*, with (Durlauf & Philips, 1988).

We only assume that the "ergodic mean" of $\nu(t)$ is 0, *i.e.*,

$$\lim_{N \to +\infty} \frac{\nu(0) + \nu(1) + \cdots + \nu(N)}{N+1} = 0 \quad (5)$$

It means that, $\forall\ t \in \mathbb{N}$, the moving average

$$\mathrm{MA}_{\nu,N}(t) = \frac{\nu(t) + \nu(t+1) + \cdots + \nu(t+N)}{N+1} \quad (6)$$

is close to 0 if N is large enough. It follows from Eq. (4) that $\epsilon(t)$ also satisfies the properties (5) and (6). Most of the stochastic processes, like finite linear combinations of i.i.d. zero-mean processes, which are associated to time series modeling, do satisfy almost surely such a weak assumption. Our analysis

- does not make any difference between non-stationary and stationary time series,
- does not need the often tedious and cumbersome trend and seasonality decomposition (our trendlines include the seasonalities, if they exist).

1.3 A model-free setting

It should be clear that

- a concrete time series cannot be "well" approximated in general by a solution of a "parsimonious" Eq. (1), *i.e.*, a linear difference equation of low order;
- the use of large order linear difference equations, or of nonlinear ones, might lead to a formidable computational burden for their identifications without any clear-cut forecasting benefit.

We adopt therefore the quite promising viewpoint of (Fliess & Join, 2008) where the control of "complex" systems is achieved without trying a global identification but thanks to elementary models which are only valid during a short time interval and are continuously updated.[3] We utilize here low order difference equations.[4] Then the window size for the moving average (6) does not need to be "large".

1.4 Content

Sect. 2, which considers the identifiability of unknown parameters, extends to the discrete-time case a result in (Fliess, 2008). The convincing computer simulations in Sect. 3 are based on the exchange rates between US Dollars and €uros. Besides forecasting the trendline, we predict

- the position of the future rate w.r.t. the forecasted trendline,
- the standard deviation w.r.t. the forecasted trendline.

Those results might lead to a new understanding of volatility and risk management.[5] Sect. 4 concludes with a short discussion on the notion of *trend*.

[3] See the numerous examples and the references in (Fliess & Join, 2008) for concrete illustrations.

[4] Compare with (Markovsky, Willems, van Huffel, de Moor & Pintelon, 2005).

[5] See (Taleb, 1997) for a critical appraisal of the existing literature on this subject, which is of utmost importance in financial engineering. (Extreme) risks are discussed in (Bouchaud & Potters, 1997; Dacorogna, Gençay, Müller, Olsen & Pictet, 2001; Mandelbrot & Hudson, 2004; Sornette, 2003) from quite different perspectives. It is the trendline which would exhibit abrupt changes in our setting (compare with the probabilistic standpoint; see, *e.g.*, (Wilmott, 2006) and the references therein). Our estimation techniques permit an efficient change-point detection (Mboup, Join & Fliess, 2008), which needs to be extended, if possible, to some kind of forecasting.

2. PARAMETER IDENTIFICATION

2.1 Rational generating functions

Consider again Eq. (1). The $\mathcal{Z}$-transform X of x satisfies (see, *e.g.*, (Doetsch, 1967; Jury, 1964))

$$\begin{array}{l} z^n[X - x(0) - x(1)z^{-1} - \cdots - x(n-1)z^{-(n-1)}] \\ - \cdots - a_{n-1}z[X - x(0)] - a_n X = 0 \end{array} \quad (7)$$

It shows that X, which is called the *generating function* of x, is a rational function of z, *i.e.*, $X \in \mathbb{R}(z)$:

$$X = \frac{P(z)}{Q(z)} \quad (8)$$

where

$$\begin{array}{l} P(z) = b_0 z^{n-1} + b_1 z^{n-2} + \cdots + b_{n-1} \in \mathbb{R}[z] \\ Q(z) = z^n - \cdots - a_{n-1}z - a_n \in \mathbb{R}[z] \end{array}$$

Hence

Proposition 2.1. $x(t)$, $t \geq 0$, satisfies a linear difference equation (1) if, and only if, its generating function X is a rational function.

It is obvious that the knowledge of P and Q permits to determine the initial conditions $x(0), \ldots, x(n-1)$.

Remark 2.2. Consider the inhomogeneous linear difference equation

$$\begin{array}{l} x(t+n) - a_1 x(t+n-1) - \cdots - a_n x(t) \\ = \sum_{\text{finite}} \varpi(t)\alpha^t + \sum_{\text{finite}} \varpi'(t)\sin(\omega t + \varphi) \end{array}$$

where $\varpi(t), \varpi'(t) \in \mathbb{R}[t]$, $\alpha, \omega, \varphi \in \mathbb{R}$. Then the $\mathcal{Z}$-transform $X \in \mathbb{R}(z)$ of $x(t)$ is again rational. It is equivalent saying that $x(t)$, $t \geq 0$, still satisfies a homogeneous difference equation.

2.2 Parameter identifiability

Generalities Let

$$\mathfrak{K} = \mathbb{Q}\left(a_1, \ldots, a_n, b_0, \ldots, b_{n-1}\right)$$

be the field generated over the field $\mathbb{Q}$ of rational numbers by $a_1, \ldots, a_n, b_0, \ldots, b_{n-1}$, which are considered as unknown parameters and therefore in our algebraic setting as independent indeterminates (Fliess & Sira-Ramírez, 2003; Fliess, Join & Sira-Ramírez, 2004; Fliess & Sira-Ramírez, 2008). Write $\bar{\mathfrak{K}}$ the algebraic closure of $\mathfrak{K}$ (see, *e.g.*, (Lang, 2002; Chambert-Loir, 2005)). Then $X \in \bar{\mathfrak{K}}(z)$, *i.e.*, X is a rational function over $\bar{\mathfrak{K}}$. Moreover $\bar{\mathfrak{K}}(z)$ is a *differential field* (see, *e.g.*, (Chambert-Loir, 2005)) with respect to the derivation $\frac{d}{dz}$. Its subfield of *constants* is the algebraically closed field $\bar{\mathfrak{K}}$.

Introduce the square Wronskian matrix $\mathcal{M}$ of order $2n+1$ (Chambert-Loir, 2005) where its χ^{th}-row, $0 \leq \chi \leq 2n$, is

$$\frac{d^\chi}{dz^\chi} z^n X, \ldots, \frac{d^\chi}{dz^\chi} X, \frac{d^\chi}{dz^\chi} z^{n-1}, \ldots, \frac{d^\chi}{dz^\chi} 1 \quad (9)$$

It follows from Eq. (7) that the rank of $\mathcal{M}$ is $2n$ if, and only if, x does not satisfy a linear difference relation of order strictly less than n. Hence

Theorem 2.3. If x does not satisfy a linear difference equation of order strictly less than n, then the parameters

$$a_1, \ldots, a_n, b_0, \ldots, b_{n-1}$$

are *linearly identifiable.* [6]

Identifiability of the dynamics For identifying the dynamics, *i.e.*, $a_1, \ldots, a_n$, without having to determine the initial conditions consider the $(n+1)\times(n+1)$ Wronskian matrix $\mathcal{N}$, where its μ^{th}-row, $0 \leq \mu \leq n+1$, is

$$\frac{d^{n+\mu}}{dz^{n+\mu}} z^n X, \ldots, \frac{d^{n+\mu}}{dz^{n+\mu}} X$$

It is obtained by taking the X-dependent entries in the $n+1$ last rows of type (9), *i.e.*, in disregarding the entries depending on $b_0, \ldots, b_{n-1}$. The rank of $\mathcal{N}$ is again n. Hence

Corollary 2.4. $a_1, \ldots, a_n$ are linearly identifiable.

Identifiability of the numerator Assume now that the dynamics is known but not the numerator P in Eq. (8). We obtain $b_0, \ldots, b_{n-1}$ from the first n rows (9). Hence

Corollary 2.5. $b_0, \ldots, b_{n-1}$ are linearly identifiable.

2.3 Some hints on the computer implementation

We proceed as in (Fliess & Sira-Ramírez, 2003; Fliess & Sira-Ramírez, 2008) and in (Fliess, Fuchshumer, Schöberl, Schlacher & Sira-Ramírez, 2008). The unknown linearly identifiable parameters are solutions of a matrix linear equation, the coefficients of which depend on x. Let us emphasize that we substitute to x its filtered value thanks to a discrete-time version of (Mboup, Join & Fliess, 2009). [7]

3. EXAMPLE: FORECASTING 5 DAYS AHEAD THE \$ - € EXCHANGE RATES

We are utilizing data from the European Central Bank, depicted by the blue lines in the Figures 1 and 2, which summarize the 2400 last daily exchange rates between the US Dollars and the €uros. [8]

3.1 Forecasting the trendline

In order to forecast the exchange rate 5 days ahead we apply the rules sketched in Sect. 2.3 and we utilize a linear difference equation (1) of order 3 (the filtered values of the exchange rates are given by the black lines in the Figures 1, 2). Fig. 3 provides the estimated values of the coefficients of the difference equation. The results on the forecasted values of the exchange rates are depicted by the red lines in the Figures 1 and 2, which should be viewed as a predicted trendline.

[6] It means following the terminology of (Fliess & Sira-Ramírez, 2003; Fliess & Sira-Ramírez, 2008) that $a_1, \ldots, a_n, b_0, \ldots, b_{n-1}$ are uniquely determined by a system of $2n$ linear equations, the coefficients of which depend on $\frac{d^\chi}{dz^\chi}X$ and $\frac{d^\chi}{dz^\chi}z^m$, $0 \leq m \leq n-1$.

[7] See, *e.g.*, (Gençay, Selçuk & Hafner, 2002) for an excellent presentation of various filtering techniques in economics and finance.

[8] The authors are perfectly aware that only computations dealing with high frequency data might be of practical value. This type of results will be presented elsewhere.

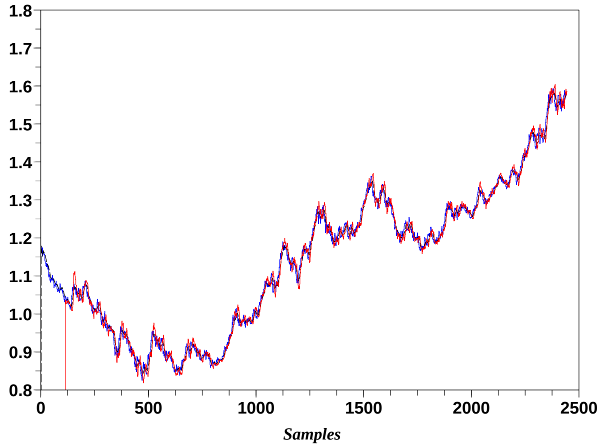

Figure 1. Exchange rates (blue –), filtered signal (black --), forecasted signal (5 days ahead) (red –)

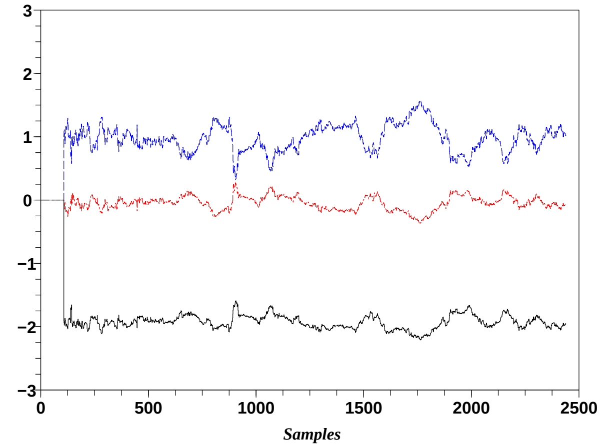

Figure 3. Parameter estimations a_1 (red –.–), a_2 (blue – –) and a_3 (black – –) (5 days ahead)

3.2 Above or under the predicted trendline?

Consider again the "error" $\nu(t)$ in Eq. (2) and its moving average $\text{MA}_{\nu,N}(t)$ in Eq. (6). Forecasting $\text{MA}_{\nu,N}(t)$ as in Sect. 3.1 tells us an expected position with respect to the forecasted trendline. The blue line of Fig. 4 displays the result for the window size $N = 100$. The meaning of the *indicators* $\triangle$ and ∇ is clear.

Table 1 compares for various window sizes the signs of the predicted values of $\text{MA}_{\nu,N}(t)$, which tells us if one should expect to be above or under the trendline, with the true positions of $x(t)$ with respect to the trendline. The results are expressed via percentages.

3.3 Predicted volatility w.r.t. the trendline

Introduce the *moving standard deviation*

$$\text{MSTD}_{\nu,N}(t) = \sqrt{\frac{\sum_{\tau=0}^{N}(\nu(t+\tau) - \text{MA}_{\nu,N}(t-N+\tau))^2}{N+1}}$$

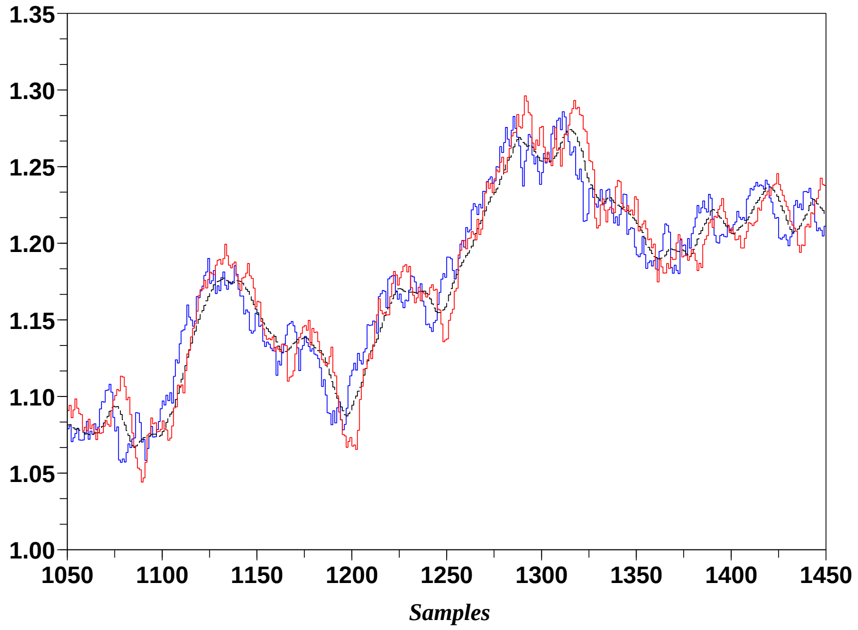

Figure 2. Zoom of Figure 1

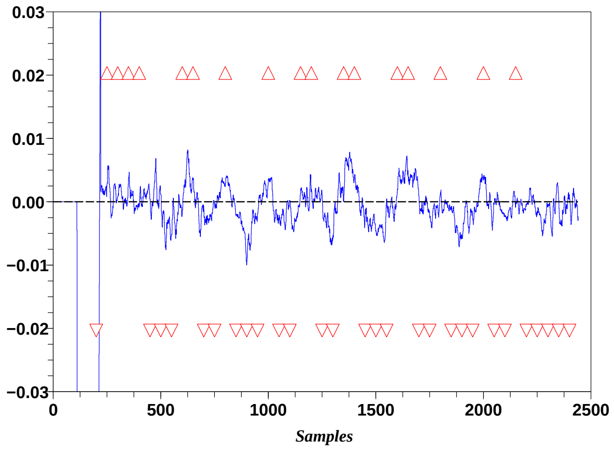

Figure 4. Predicted position w.r.t. the trendline (5 days ahead) – ∇: above, △: under

Window's size	Percentage
50	65.6%
100	88.3%
200	62.3%
300	67.1%

Table 1. Comparison between the sign of the predicted value of $\mathrm{MA}_{\nu,N}(t)$ and the true position of $x(t)$ w.r.t. the trendline (5 days ahead).

and forecast it as in Sect. 3.2. The results, which are displayed for a window size $N = 100$ in Table 2 and Fig. 5 via the familiar confidence intervals,[9] confirm the time-dependence of the variance, *i.e.*, the *heteroscedasticity*.

[9] There is of course no need for the underlying statistics to be Gaussian. Lack of space prevents us from exhibiting forecasts of quantities like *skewness* and *kurtosis*, which would be obtained by similar calculations. This will be done in some future publications.

Confidence intervals	Prediction	Real
mean-3×std,mean+3×std	99%	98.7%
mean-2×std,mean+2×std	95%	92.2%
mean-std,mean+std	68%	64.4%

Table 2. Confidence interval validations (5 days ahead)

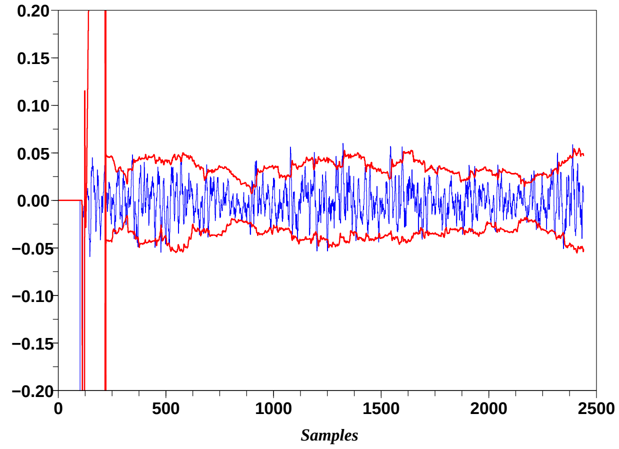

Figure 5. Confidence interval (95%) (5 days ahead)

3.4 Forcasting 10 days ahead

Figures 6, 7, 8, 9 display the same type of results as in Sections 3.1, 3.2, 3.3 via similar computations for a forecasting 10 days ahead. The quality of the computer simulations only slightly deteriorates.

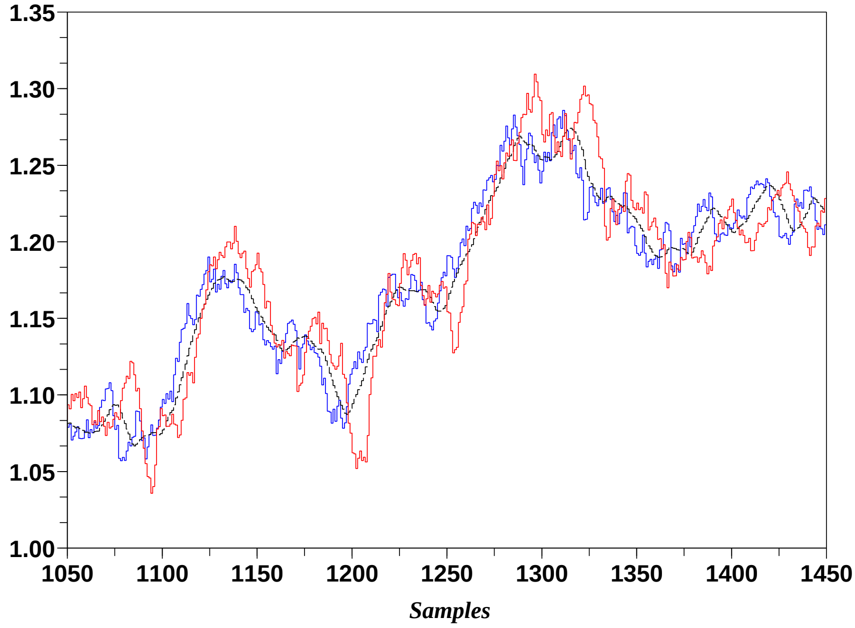

Figure 7. Zoom of Figure 6

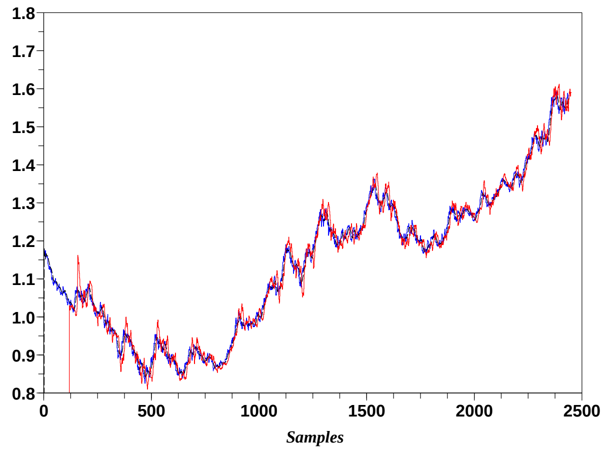

Figure 6. Exchange rates (blue –), filtered signal (black --), forecasted signal (red –) (10 days ahead)

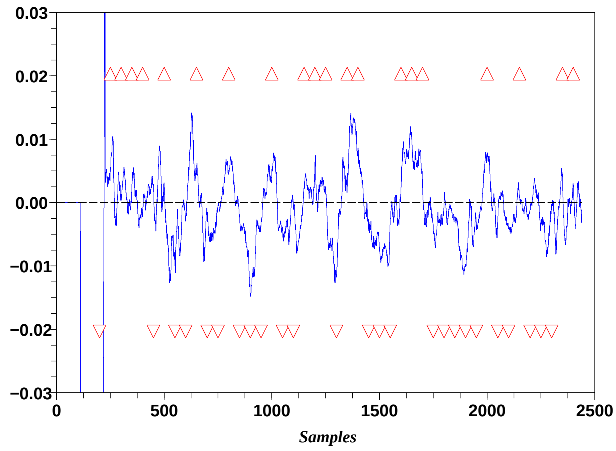

Figure 8. Predicted position w.r.t. the trendline (10 days ahead) – ∇: above, △: under

4. CONCLUSION

The existence of *trends*, which is

- the key assumption in technical analysis,[10]
- quite foreign, to the best of our knowledge, to the academic mathematical finance, where the paradigm of *random walks* is prevalent (see, *e.g.*, (Wilmott, 2006)),

is fundamental in our approach. A theoretical justification will appear soon (Fliess & Join, 2009).[11] We hope it will lead to a sound foundation of technical analysis,[12] which will bring as a byproduct easily implementable real-time computer programs.[13]

[10] Trends in technical analysis should not be confused with what are called *trends* in the time series literature (see, *e.g.*, (Gouriéroux & Monfort, 1995; Hamilton, 1994)).

[11] The existence of trends does not necessarily contradict a random character (see (Fliess & Join, 2009) for details).

[12] See also (Dacorogna, Gençay, Müller, Olsen & Pictet, 2001) for a most exciting study which employs high frequency data. There are also other types of attempts to put technical analysis on a firm basis (see, *e.g.*, (Lo, Mamaysky & Wang, 2000)). See (Blanchet-Scalliet, Diop, Gibson, Talay & Tanré, 2007) for a comparison between technical analysis and model-based approaches with parametric uncertainties.

[13] Our technics already lead to such computer programs in automatic control and in signal processing.

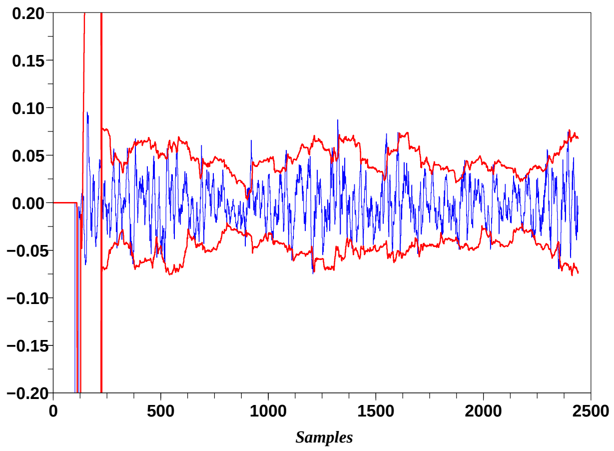

Figure 9. Confidence interval (95%) (10 days ahead)

Acknowledgement. The authors wish to thank G. Daval-Leclercq (Société Générale - Corporate & Investment Banking) for helpful discussions and comments.

REFERENCES

Aronson D. (2007). *Evidence-Based Technical Analysis*. Wiley.

Béchu T., Bertrand E., Nebenzahl J. (2008). *L'analyse technique* (6^e éd.). Economica.

Blanchet-Scalliet C., Diop A., Gibson R., Talay D., Tanré E. (2007). Technical analysis compared to mathematical models based methods under parameters mis-specification. *J. Banking Finance*, **31**, 1351–1373.

Bouchaud J.-P., Potters M. (1997). *Théorie des risques financiers*. Eyrolles. English translation (2000): *Theory of Financial Risks*. Cambridge University Press.

Box G.E.P., Jenkins G.M., Reinsel, G. (1994). *Time Series Analysis: Forecasting and Control* (3^{rd} ed.). Prentice Hall.

Chambert-Loir A. (2005). *Algèbre corporelle*. Édit. École Polytechnique. English translation (2005): *A Field Guide to Algebra*. Springer.

Dacorogna M.M., Gençay R., Müller U., Olsen R.B., Pictet O.V. (2001). *An Introduction to High Frequency Finance*. Academic Press.

Doetsch G. (1967). Funktionaltransformationen. In R. Sauer, L. Szabó (Eds): *Mathematische Hilfsmittel des Ingenieurs*, 1. Teil, pp. 232–484. Springer.

Durlauf S.N., Philips P.C.B. (1988). Trends versus random walks in time series analysis. *Econometrica*, **56**, 1333–1354.

Fliess M. (2008). Critique du rapport signal à bruit en communications numériques. *ARIMA*, to appear. Online `http://hal.inria.fr/inria-00311719/en/`.

Fliess M., Fuchshumer S., Schöberl M., Schlacher K., Sira-Ramírez H. (2008). An introduction to algebraic discrete-time linear parametric identification with a concrete application. *J. europ. syst. automat.*, **42**, 211–232.

Fliess M., Join C. (2008). Commande sans modèle et commande à modèle restreint, *e-STA*, **5**. Online `http://hal.inria.fr/inria-00288107/en/`.

Fliess M., Join C. (2009). A mathematical proof of the existence of trends in financial time series. *Int. Conf. Systems Theory: Modelling, Analysis and Control*. Fes, Marocco. Soon online `http://hal.inria.fr/INRIA`.

Fliess M., Join C., Sira-Ramírez H. (2004). Robust residual generation for linear fault diagnosis: an algebraic setting with examples. *Int. J. Control*, **77**, 1223–1242.

Fliess M., Join C., Sira-Ramírez H. (2008). Non-linear estimation is easy. *Int. J. Modelling Identification Control*, **4**, 12–27.

Fliess M., Sira-Ramírez H. (2003). An algebraic framework for linear identification. *ESAIM Control Optimiz. Calculus Variat.*, **9**, 151–168.

Fliess, M., Sira-Ramírez, H. (2004). Reconstructeurs d'état. *C.R. Acad. Sci. Paris Ser. I*, **338**, 91–96.

Fliess M., Sira-Ramírez H. (2008). Closed-loop parametric identification for continuous-time linear systems via new algebraic techniques. In H. Garnier, L. Wang (Eds): *Identification of Continuous-time Models from Sampled Data*, pp. 363–391, Springer.

Gençay R., Selçuk F., Whitcher B. (2002). *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Academic Press.

Gouriéroux C., Monfort A. (1995). *Séries temporelles et modèles dynamiques* (2^e éd.). Economica. English translation (1996): *Time Series and Dynamic Models*. Cambridge University Press.

Hamilton J.D. (1994). *Time Series Analysis*. Princeton University Press.

Kaufman P.J. (2005). *New Trading Systems and Methods* (4^{th} ed.). Wiley.

Jury E.I. (1964). *Theory and Application of the z-Transform Method*. Wiley.

Kirkpatrick C.D., Dahlquist J.R. (2006). *Technical Analysis: The Complete Resource for Financial Market Technicians*. FT Press.

Lang S. (2002). *Algebra* (3^{rd} rev. ed.). Springer.

Lo A.W., Mamaysky H., Wang J. (2000). Foundations of technical analysis: computational algorithms, statistical inference, and empirical implementation. *J. Finance*, **55**, 1705-1765.

Mandelbrot B.B., Hudson R.L. (2004). *The (Mis) Behavior of Markets*. Basic Books.

Markovsky I., Willems J.C., van Huffel S., de Moor B., Pintelon R. (2005). Application of structured total least squares for system identification and model reduction. *IEEE Trans. Automat. Control*, **50**, 1490-1500.

Mboup M., Join C., Fliess M. (2008). A delay estimation approach to change-point detection. *Proc. 16^{th} Medit. Conf. Control Automation*, Ajaccio. Online `http://hal.inria.fr/inria-00179775/en/`.

Mboup M., Join C., Fliess M. (2009). Numerical differentiation with annihilators in noisy environment. *Numer. Algor.*, DOI: 10.1007/s11075-008-9236-1.

Murphy J.J. (1999). *Technical Analysis of the Financial Markets* (3^{rd} rev. ed.). New York Institute of Finance.

Paulos J.A. (2003). *A Mathematician Plays the Stock Market*. Basic Books.

Sornette D. (2003). *Why Stock Markets Crash: Critical Events in Complex Financial Systems*. Princeton University Press.

Taleb N. (1997). *Dynamic Hedging*. Wiley.

Wilmott P. (2006). *Paul Wilmott on Quantitative Finance* (2^{nd} ed., 3 volumes). Wiley.