Finance Stochast. 6, 49–61 (2002)

Finance and
Stochastics

Fourier series method for measurement of multivariate volatilities

Paul Malliavin[1], **Maria Elvira Mancino**[2]

[1] 10 rue Saint Louis en l'Isle, 75004 Paris, France (e-mail: sli@ccr.jussieu.fr)
[2] DIMAD, Università di Firenze, Via C.Lombroso 6/17, 50134 Firenze, Italy
(e-mail: mancino@mail.dm.unipi.it)

Abstract. We present a methodology based on Fourier series analysis to compute time series volatility when the data are observations of a semimartingale. The procedure is not based on the Wiener theorem for the quadratic variation, but on the computation of the Fourier coefficients of the process and therefore it relies on the integration of the time series rather than on its differentiation. The method is fully model free and nonparametric. These features make the method well suited for financial market applications, and in particular for the analysis of high frequency time series and for the computation of cross volatilities.

Key words: Volatility, Fourier series, financial time series

JEL Classification: C14, C32, C63

Mathematics Subject Classification (1991): 42B05, 62H12, 62M15, 62P20

Introduction

This work belongs to the realm of the econometric study of market evolution. It presents a new method based on Fourier analysis to measure volatility both in a univariate and in a multivariate setting. The method is completely model free, it only assumes that the data from the market (returns, prices, etc.) are observations of a semimartingale. We will refer to this assumption as the *Bachelier hypothesis*. No further assumption is needed. In particular, we will not parametrize the coefficients of the process and we will not suppose any long range stationarity, as it is usually done in the statistical study of time series in order to compute, by using the ergodic theorem, from a *single realization of the process* its spectral measure or some other invariant. At the first glance, it seems impossible to

Manuscript received: October 2000; final version received: January 2001

build a reasonable statistical estimator on very short time sample of a time series. Nevertheless, *volatilities* are statistical quantities that can be measured with full mathematical rigour. This point can be stated *theoretically*: in the context of the *Bachelier paradigm* the quadratic variation is perfectly defined; *empirically*: there exists a large literature on the estimate of the volatility through the quadratic variation formula which provides us with an unbiased estimator for it. In some cases, the volatility of the process is parametrized and then the parameters are estimated via the maximum likelihood method or the general method of moments, e.g. see [D],[GCJ],[LO],[FT], there are also some nonparametric estimators of the volatility, see [GLP],[H],[FZ],[NE],[JK],[AS]. For a survey of the literature on financial market applications we refer the reader to [CLM]. Almost all these methods are based on the quadratic variation formula which is implemented by omitting the limit and therefore considering the observations with a given frequency (daily, weekly, even monthly). Depending on the frequency of the sample, the estimate can be quite different, as a matter of fact the algorithm being based on a "differentiation procedure" is quite unstable. In particular, problems arise in the computation of cross-volatilities and using high frequency data which are not equally and regularly spaced and therefore are not well suited for the computation of the quadratic variation. We will discuss these drawbacks in Sect. 5.

To address these problems, we propose here a robust algorithm based on Fourier analysis to determine the multivariate volatility evolution from the observations of the market. This algorithm will reconstruct the *volatility as a function of time*. The method, being based on the integration of the time series, is more robust than the classical one and is well suited to compute volatility for a high frequency time series.

We heartly thanks the referee whose constructive report leads us to a complete reorganisation of the work.

1 Wiener theorem on quadratic variation

The material exposed in this section is quite classical. We recall the definition of *quadratic variation*. Given two $\mathbb{R}$-valued functions ϕ, ψ defined on $[0,1]$, we define the function

$$< \phi, \psi >_t := \lim_{n\to\infty} \left(\sum_{0 \le k \le 2^n t} (\phi(2^{-n}(k+1)) - \phi(2^{-n}k))(\psi(2^{-n}(k+1)) - \psi(2^{-n}k)) \right).$$

To shorthand the notation, we make the convention that each time we write the symbol $< \phi, \psi >_t$, we do assume that the written limit exists. We have the Schwarz inequality

$$| < \phi, \psi >_t |^2 \le < \phi, \phi >_t < \psi, \psi >_t \ .$$

Consider now two functions u, v satisfying an Hölder condition of exponent $\alpha > \frac{1}{2}$, then

$$< u, u >_t = 0, \quad < \phi + u, \psi + v >_t = < \phi, \psi >_t .$$

The following theorem holds.

Theorem 1.1 (Wiener theorem) *Consider an $\mathbb{R}^d$ valued Brownian motion $(x^1, \ldots, x^d)$, then for almost all the trajectories we have*

$$< x^j, x^k >_t = \delta^j_k \; t, \quad \forall \; t \in [0, 1], \quad \forall \; j, k = 1, \ldots, d$$

where $\delta^j_k = 0$ if $j \neq k$ and equal to 1 otherwise.

Remark 1.2 The remarkable result expressed by the Wiener theorem is that it is true on each trajectory of the Brownian motion.

Proof It is sufficient to prove the theorem for the scalar valued Brownian motion x. We shall present a proof which underlines how it is possible to pass from an averaged result to an almost sure result. Without loss of generality, we can assume that $t = 1$. Since

$$E\left((x(2^{-n}(k+1) - x(2^{-n}k))^2\right) = 2^{-n},$$

we deduce that

$$E(S_n) = 1 \text{ where } S_n := \sum_{k=0}^{2^n - 1} (x(2^{-n}(k+1)) - x(2^{-n}k))^2. \tag{1}$$

By

$$E\left((x(2^{-n}(k+1)) - x(2^{-n}k))^2 (x(2^{-n}(s+1) - x(2^{-n}s))^2\right) = c2^{-2n}$$

where $c = 1$ if $s \neq k$ and $c = 3$ if $s = k$, we deduce

$$E(S_n^2) = 1 + 3 \times 2^{-n}.$$

Therefore we have

$$E((S_n - 1)^2) = 3 \times 2^{-n},$$

which implies that

$$P(|S_n - 1| > 2^{1-\frac{n}{4}}) < 2^{-\frac{n}{4}},$$

and, by the Borel-Cantelli theorem, that almost surely S_n converges to 1. □

2 Quadratic variation under the Bachelier paradigm

Our study will be based on the following *Bachelier hypothesis* [B] which can be formulated nowadays as :

All measurable economic data u^* are driven by semi martingales which have their Itô stochastic differential given by

$$du^j = \sum_{i=1}^{d} \alpha_i^j \ dx^i + \beta^j \ dt \tag{2}$$

where x^* are independent Brownian motions, and $\alpha_*^*,\ \beta^*$ are functions depending upon time.

Our hypothesis encompasses all the models supposing that the coefficients $\alpha_i^j,\ \beta^j$ are random processes satisfying some hypothesis; in fact in all these models the coefficients are *functions of time* which is precisely our hypothesis. No parametrization of the process is assumed, this renders our approach completely general and furthermore model free.

The following result holds true for almost all the trajectories.

Theorem 2.1 *Assume that the functions $\alpha_i^j,\ \beta^j$ are uniformly bounded, then almost surely we have the identity*

$$< u^j, u^k >_t = \int_0^t \sum_{i=1}^{d} \alpha_i^j(s)\alpha_i^k(s)\ ds\ . \tag{3}$$

Proof Without loss of generality, it is sufficient to prove the result for $t = 1$; by polarization it is sufficient to prove it when $j = k$; we shall shorthand the notation by omitting the upper indices j. The term β^j will not contribute to the quadratic variation and we shall forget it. We consider the partition of $[0, 1]$ given by the points $k2^{-n}$, $k = 0, \ldots, 2^n$, and denote

$$a_{n,k} := (u(2^{-n}(k+1)) - u(2^{-n}k))^2 - \int_{k2^{-n}}^{(k+1)2^{-n}} \sum_{i=1}^{d} \alpha_i^2(s)\ ds;$$

we denote by $\mathcal{N}_t$ the Itô filtration associated to the Brownian motion x. By the Itô energy identity on stochastic integrals, we have

$$E^{\mathcal{N}_{k2^{-n}}}(a_{n,k}) = 0;$$

therefore for $k' < k$,

$$E(a_{n,k'}a_{n,k}) = E(a_{n,k'}E^{\mathcal{N}_{k2^{-n}}}(a_{n,k})) = 0. \tag{4}$$

Introducing

$$J_n := \sum_{0 \le k < 2^n} a_{n,k},$$

we have by (4)

$$E(J_n^2) = \sum_{0 \le k < 2^n} E(a_{n,k}^2).$$

Consider

$$Y(t) := \int_{2^{-n}k}^{t} \sum_{i=1}^{d} \alpha_i(s)\, dx^i(s),$$

then by Itô calculus we have

$$Y^2(t) - \int_{2^{-n}k}^{t} \sum_{i=1}^{d} \alpha_i^2(s)\, ds = \int_{2^{-n}k<r<s<t} \sum_{i,j=1}^{d} \alpha_j(r)\alpha_i(s)\, dx^j(r)\, dx^i(s)$$

which implies that

$$E((a_{n,k})^2) < c2^{-2n}.$$

We conclude that $E(J_n^2) < c2^{-n}$ which implies that almost surely J_n converges to 0. □

Going back to the representation (2) we denote

$$\Sigma^{j,k}(t) = \sum_i \alpha_i^j(t)\alpha_i^k(t). \tag{5}$$

We shall call $\Sigma^{*,*}$ the volatility matrix. We observe that, while there are several ways to represent u_* by a system of the form (2), the formula (3) gives to the volatility matrix the status of an intrinsic object which can be exactly measured at each time by empiric observations. To summarize, we can say that the volatility matrix is independent of the model which has been constructed in order to implement the Bachelier paradigm. We emphasize that under our hypothesis the volatility matrix depends upon time.

3 Fourier series computation of multivariate volatilities

The formula (3) has a nice conceptual meaning, but it has the disadvantage of being based on "a differentiation procedure". Therefore it is not well suited to provide a good estimate of the volatility. As a matter of fact, it is well known by the numerical analysis literature that differentiation of empirical functions have the tendency to lead to unstable numerical algorithms. In what follows we propose a new approach based on energy identity for stochastic integrals which avoids all these problems.

To facilitate our exposition, we consider firstly a single semi-martingale u on a fixed time window. We want to determine the evolution of the volatility $\Sigma(t)$ on this window. By change of the origin of time and by rescaling the unit of time, we can always reduce ourselves to the case where the time window is $[0, 2\pi]$. Then

$$du = \sum_{i=1}^{d} \alpha_i(t)\, dx^i + \beta(t)\, dt, \quad \Sigma(t) = \sum_{i=1}^{d} (\alpha_i)^2(t)$$

Our method will be the following: first compute the Fourier coefficients of du, then obtain a mathematical expression of the Fourier coefficients of Σ using the Fourier coefficients of du. Finally, classical results in Fourier theory allows to reconstruct Σ from its Fourier coefficients.

We compute the Fourier coefficients of du defined by

$$\begin{aligned} a_0(du) &= \frac{1}{2\pi}\int_0^{2\pi} du(t), \quad a_k(du) = \frac{1}{\pi}\int_0^{2\pi} \cos(kt)\; du(t), \\ b_k(du) &= \frac{1}{\pi}\int_0^{2\pi} \sin(kt)\; du(t). \end{aligned}$$

We consider the Fourier coefficients of the volatility

$$\begin{aligned} a_0(\Sigma) &= \frac{1}{2\pi}\int_0^{2\pi} \Sigma(t)\; dt, \quad a_k(\Sigma) = \frac{1}{\pi}\int_0^{2\pi} \cos(kt)\Sigma(t)dt, \\ b_k(\Sigma) &= \frac{1}{\pi}\int_0^{2\pi} \sin(kt)\; \Sigma(t)\; dt. \end{aligned}$$

The reconstruction of Σ from its Fourier coefficients is given by the classical Fourier-Féjer inversion formula:

$$\Sigma(t) = \lim_{N\to\infty} \Sigma_N(t) \tag{6}$$

where

$$\Sigma_N(t) := \sum_{k=0}^{N}(1 - \frac{k}{N})(a_k(\Sigma)\cos(kt) + b_k(\Sigma)\sin(kt)), \;\; \forall t \in (0, 2\pi).$$

We notice here that there are many inversion formulae which allow us to reconstruct a function from the data of its first Fourier coefficients. The advantage of Féjer inversion formula is that if Σ is a positive function, then the approximation given by (6) will be again positive.

In what follows, our main result is stated.

Theorem 3.1 *Fix an integer $n_0 > 0$, then the Fourier coefficients of the volatility are given by the following formulae:*

$$a_0(\Sigma) = \lim_{N\to\infty} \frac{\pi}{N+1-n_0} \sum_{s=n_0}^{N}(a_s^2(du) + b_s^2(du)), \tag{7}$$

$$a_q(\Sigma) = \lim_{N\to\infty} \frac{2\pi}{N+1-n_0} \sum_{s=n_0}^{N} a_s(du)a_{s+q}(du), \forall q > 0, \tag{8}$$

$$b_q(\Sigma) = \lim_{N\to\infty} \frac{2\pi}{N+1-n_0} \sum_{s=n_0}^{N} a_s(du)b_{s+q}(du), \;\forall\; q \geq 0. \tag{9}$$

Proof Firstly, we precise the meaning of these identities. Under the Bachelier paradigm the observed historical evolution of the market is one of the possible evolutions described by an underlying probability space Ω, on which the driving Brownian motions x^i and the adapted processes $\alpha_i(\omega, *)$, $\beta(\omega, *)$ are realized. As said before, we do not make any hypothesis on the structure of this probability space Ω; the process u is defined on this probability space, that is $u = u_\omega$. We denote by ω_0 the sample of the probability space Ω which corresponds to the observed evolution of the market.

We construct an auxiliary probability space X which is the Wiener space generated by an $\mathbb{R}^d$ valued Brownian motion considered on the time interval $[0, 1]$. We have a natural probability preserving map $\Phi : \Omega \mapsto X$. Define *the deterministic functions* $\tilde{\alpha}_i(t) := \alpha_i(\omega_0, t)$, $\tilde{\beta}(t) := \beta(\omega_0, t)$ and consider the process $\tilde{u}_x$ defined on X by

$$\tilde{u}_x(t) = \int_0^t \sum_i \tilde{\alpha}_i(t)\, dx^i(t) + \int_0^t \tilde{\beta}(t)\, dt,$$

where the stochastic integrals are Wiener type stochastic integrals of the Brownian motion integrating deterministic functions. Then we have

$$u_{\omega_0} = \tilde{u}_{x_0} \text{ where } x_0 := \Phi(\omega_0),\ a_k(du_{\omega_0}) = a_k(d\tilde{u}_{x_0}),\ b_k(du_{\omega_0}) = b_k(d\tilde{u}_{x_0}). \quad (10)$$

Our proof will be based on three steps: first we shall prove that the formulae (7),(8) and (9) hold true for the process $\tilde{u}_x$ if $x \notin S \subset X$ where $P(S) = 0$; then granted (10), the same formulae will hold true if $\omega_0 \notin \Phi^{-1}(S)$; finally, as $P(\Phi^{-1}(S)) = P(S) = 0$, we shall get that the formulae (7), (8) and (9) holds true on Ω almost surely.

We denote by E the expectation on the probability space X.

Using the fact that the trigonometric system is an orthonormal basis of $L^2([0, 2\pi])$ we get

$$\int_0^{2\pi} \tilde{\beta}^2(t)\, dt = \sum_{k=0}^{\infty} (a_k^2(\tilde{\beta}) + b_k^2(\tilde{\beta}))$$

therefore the contribution of $\tilde{\beta}$ to the formulae (7),(8),(9) is zero. We can replace $a_*(d\tilde{u})$ by $a_*(dv)$ where $dv = d\tilde{u} - \tilde{\beta}\, dt$. We introduce the gaussian variables

$$G_k := a_k(dv),\ G'_k := b_k(dv);$$

the Itô energy identities for stochastic integrals leads to the following expression for the covariance :

$$E(G_k G_l) = \frac{1}{\pi^2} \int_0^{2\pi} \Sigma(t) \cos(kt) \cos(lt)\, dt.$$

Using the identity

$$\cos(kt)\cos(lt) = \frac{1}{2}(\cos(k - l)t + \cos(k + l)t) \quad (11)$$

we get

$$E(G_k G_l) = \frac{1}{2\pi}(a_{|k-l|}(\Sigma) + a_{k+l}(\Sigma)). \tag{12}$$

The energy identity is

$$\|\Sigma\|_{L^2}^2 = \sum_k (a_k(\Sigma))^2 + (b_k(\Sigma))^2.$$

For $q > 0$ consider the random variable

$$U_N^q := \frac{1}{N}\sum_{k=1}^{N} G_k G_{k+q}.$$

Then using (12) we get

$$E(U_N^q) = a_q(\Sigma) + R_N$$

where $|R_N| = \frac{1}{N}|\sum_{k=1}^N a_{2k+q}(\Sigma)| \leq \frac{1}{\sqrt{N}}\|\Sigma\|_{L^2}$, the last inequality being obtained through Schwarz inequality. Compute now

$$E((U_N^q)^2) = \frac{1}{N^2}\sum_{0\leq k,k'\leq N} E(G_k^2 G_{k'+q}^2).$$

Consider an $\mathbb{R}^2$-valued gaussian variable (G_1, G_2), denote

$$\lambda_i := E(G_i^2), \quad \mu := E(G_1 G_2),$$

then $Z := G_2 - \frac{\mu}{\lambda_1}G_1$ is independent of G_1 and

$$E(G_1^2 G_2^2) = E(Z^2)E(G_1^2) + \frac{\mu^2}{\lambda_1^2}E(G_1^4) = E(G_1^2)E(G_2^2) + 2\mu^2.$$

Therefore

$$E\left((U_N^q - E(U_N^q))^2\right) = \frac{1}{2\pi^2 N^2}\sum_{0\leq k,k'\leq N}(a_{|k-k'+q|}(\Sigma) + a_{k+k'+q}(\Sigma))^2 \leq \frac{1}{N}\|\Sigma\|_{L^2}^2,$$

an inequality that proves (8). Granted the previous methodology, relations (7) and (9) are consequence of the identities

$$E(b_s^2(dv)) = \frac{1}{\pi}(a_0(\Sigma) - \frac{1}{2}a_{2s}(\Sigma)), \quad E(a_s(dv)b_{s+q}(dv)) = \frac{1}{2\pi}(b_q(\Sigma) + b_{2s+q}(\Sigma)).$$

□

By polarization of the one dimensional result, we get the following expression of the multivariate volatility matrix:

Theorem 3.2 *The coefficient* $\Sigma^{i,j}$ *of the volatility matrix has the following Fourier coefficients*

$$a_k(\Sigma^{i,j}) = \lim_{N\to\infty}\frac{\pi}{N+1-n_0}\sum_{s=n_0}^{N}\frac{1}{2}(a_s(du^i)a_{s+k}(du^j) + a_s(du^j)a_{s+k}(du^i)), \tag{13}$$

with the corresponding analogous formulae of (7) and (9).

4 Numerical implementation

A basic question is the choice of the time window, where Fourier analysis will be applied. Let t_0 the center of a time window for the computation of the volatility $[t_0 - \pi L, t_0 + \pi L]$; for each L, Fourier analysis makes possible to compute a volatility matrix $\Sigma_L^{*,*}(t_0)$, the theory shows that $\Sigma_L^{*,*}(t_0)$ does not depend upon L. The implementation of the method confirms it, but as L is decreased, the noise of the estimate goes up. Precision of the estimates goes as $\frac{1}{\sqrt{N}}$ where N is the number of observations. Note that, differently from other classical estimation procedures, no preliminary analysis of the stationarity of the time series in the time window is needed.

The volatility matrix $\Sigma(t)$ is a symmetric positive definite matrix; the Féjer procedure of reconstruction of $\Sigma(t)$ choosen above is such that all partial sums $\Sigma_N(t)$ are also positive definite matrices. The implementation of the method has fully confirmed this point.

We now consider specific numerical procedures. Given a semi-martingale

$$du = \alpha\ dx + \beta\ dt,$$

it is convenient to compute numerically its Fourier coefficients by integration by parts as

$$a_k(du) = \frac{1}{\pi}\int_0^{2\pi} \cos kt\ du(t) = -\frac{k}{\pi}\int \sin kt\ u(t)\ dt + \frac{u(2\pi) - u(0)}{\pi}. \tag{14}$$

Remark that the expression of $a_k(du)$ in (14) is numerically stable, because it does not involve the differentiation of u. As observations are finite, to implement the method and in particular the integration we need an assumption on how data are connected. Our choice is $u(t)$ be equal to $u(t_i)$ in the interval $[t_i, t_{i+1}]$ (piecewise constant). With this choice, the integral in equation (14) in the interval $[t_i, t_{i+1}]$ becomes:

$$\frac{k}{\pi}\int_{t_i}^{t_{i+1}} \sin(kt)u(t)dt = u(t_i)\frac{k}{\pi}\int_{t_i}^{t_{i+1}} \sin(kt)dt = u(t_i)\frac{1}{\pi}\left(cos(kt_i) - cos(kt_{i+1})\right) \tag{15}$$

thus avoiding the multiplication by k which amplifies cancellation errors when k becomes large. An inconvenient of this approximation is that it has not finite quadratic variation. This theoretical drawback is not detrimental to the numerical implementation, however it is advisable to use other continuous approximating functions.

The larger significant Fourier frequency is at most

$$N_1 = \left[\frac{L}{\delta}\right] \tag{16}$$

where δ is the interval between two observations. Formulae (7),(8),(9) and (13) will be computed with $n_0 < N_1$. It is advisable to eliminate the first n_0 terms as too much sensitive to the effect of the coefficients β^i. Numerical experiments show that it is enough to eliminate the first coefficient. The Fourier coefficients of Σ will be computed for $n_0 \leq k \leq J$, such that $n_0 + J < N_1$.

The method proposed above allows us to reconstruct the volatility for each t in the time window, however Monte Carlo experiments with a time varying volatility function have shown that the precision of the estimate is higher in the central part of the window than nearby the extremities. To smooth the extremity effects, a C^∞ function φ equal to zero on a neighbourhood of 0 and 2π and equal to 1 on the neighbourhood of the center π of the interval can be employed in order to compute a_k as follows:

$$\hat{a}_k(du) = \int_0^{2\pi} \varphi(t) cos(kt) du(t).$$

There are other nonparametric methods to estimate the volatility through a series expansion with an orthonormal basis, an example is provided by a wavelet basis which has been used in [GLP] to estimate the volatility of a univariate process assuming it a deterministic function of time. The methodology developed above does not work with wavelets mainly because for that basis we cannot establish the equivalent of (11) which strongly exploits the properties of a trigonometric basis. The main novelties of the method proposed above with respect to the one proposed in [GLP] are the following: a) it works in a multivariate setting, b) it works assuming stochastic coefficients without any parametric restriction.

To show that the method performs well in computing volatility for a financial time series, the method has been applied to compute the daily volatility of the Dow Jones Industrial and of the Dow Jones Transportation from 1896 to 1998 (28000 data), the sample is divided into 28 periods of 1000 data. The results are shown in Fig. 1, where the volatilities of the two series and the cross volatility computed according to the new methodology and to the classical method are reported. The estimate according to the new methodology is in line with the one obtained with the quadratic variation method. Monte Carlo simulations of “theoretical” time series confirm the goodness of the methodology.

5 Financial market applications

The method proposed above is well suited for financial market applications and in particular to compute the volatility with high-frequency data (tick by tick observations). As it is very well known, tick by tick data are inequally and irregularly spaced. These features create problems in computing the volatility and in particular cross-volatilities. To use the classical methods based on the quadratic variation, equally and regularly spaced data are needed. While daily, weekly, or monthly time series have these features, problems arise when using tick by tick observations. To avoid these problems, data are usually handled to construct regularly and equally spaced times series. This is done through interpolation or imputation methods. In the first case a time length is fixed, the time axis is split according to that length and inside each interval the last observation is

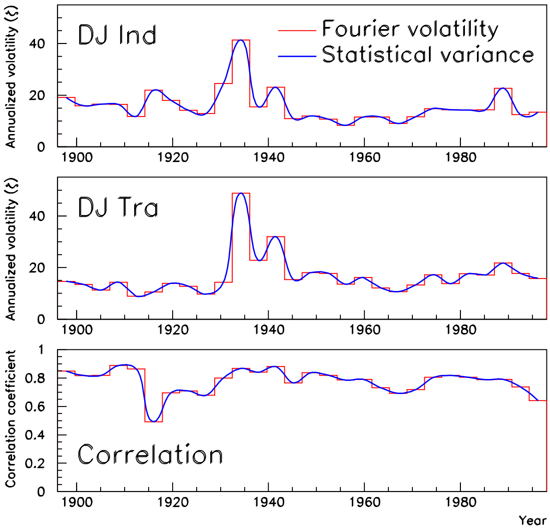

Fig. 1. Volatility matrix for two indexes: Dow Jones Industrial and Dow Jones Transportation

considered. This procedure gives us homogeneous and equally spaced time series but it entails two main drawbacks: it may happen that no observation is available in some intervals of time, furthermore some observations are thrown away and therefore only a subset of the data are used. Instead, interpolation requires to aggregate observations, e.g. a time series with observations every five minutes is built by taking at time t an average of the market prices in the time interval $[t-2.5, t+2.5]$. These procedures may cause some bias in the empirical analysis. First of all, only some data are employed, this fact reduces strongly the power of any statistical test. A second problem comes from the so called *non synchronous trading* effect: data at time t do not necessarily refer to market observations at that time. These two problems are particularly relevant in measuring cross volatility between two stocks with different trading frequency. In [LM] it is shown that non synchronous trading may be responsible for rejecting the random walk hypothesis for portfolio returns, as a matter of fact averaging induces a spurious mean reversion in returns; in a multivariate setting, non synchronous trading may generate some bias in evaluating lead-lag effects between the markets.

The problems pointed out above come from the fact that traditional methods based on the quadratic variation of the time series requires a “differentiation” of the process; our methodology, being based on the integration of the time series, is immune from these drawbacks. Volatility can be computed by using high frequence data using all of them without any manipulation. This feature allows us to use the method in order to correctly evaluate the performance of a model with high frequency data and the distortions due to interpolation or imputation methods, on them see also [ABDE, ABL]. Recently, new estimators of the volatility of a discrete time process have been developed which are not based on interpolation and imputation techniques, see [LM, DN]. They avoid manipulation of the data by assuming that returns are serially uncorrelated (in the first paper) and that the process generating the transaction times and the prices are independent (in both papers). As stressed by the authors, these assumptions are quite strong. Our method does not require them.

In [BR] the authors have applied the methodology outlined above to the estimate-forecast the volatility in the GARCH models framework. While it was widely recognized in the empirical literature that asset prices volatility is highly persistent and therefore is well described by an autoregressive process (ARCH-GARCH models), the forecasting performance of these methods was quite poor. In [AB], it is shown that a GARCH model performs well in forecasting volatility one day ahead when high frequency data are used and the volatility is computed through the cumulative squared intraday returns exploiting the classical quadratic variation result. To do it, the authors construct a time series with an observation every five minutes by means of an interpolation technique. In this way, the number of observations is smaller than the original time series and as a consequence the estimate is poorer than using all the data set. In [BR] it is shown that by employing the Fourier method proposed in this paper and all the data set the forecasting performance of the GARCH model is better than the one obtained in [AB] and also the daily volatility is estimated with a lower variance.

References

[AB] Andersen, T., Bollerslev, T.: Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. Inter. Econ. Rev. **39**, 885–905 (1998)

[ABL] Andersen, T., Bollerslev, T., Lange, S.: Forecasting financial market volatility: Sample frequency via-à-vis forecast horizon. J. Empirical Finance **6** (1999)

[ABDE] Andersen, T., Bollerslev, T., Diebold, F., Ebens, H.: The distribution of stock return volatility. NBER working paper, n. 7933, 2000

[AS] Ait-Sahalia, Y.: Nonparametric pricing of interest rate derivative securities. Econometrica **64**, 527–560 (1996)

[B] Bachelier, H.: Théorie de la spéculation. Gauthier-Villars 1900, pp. 200

[BR] Barucci, E. Renò, R.: A note on volatility estimate-forecast with GARCH models. Working paper, Dipartimento di Statistica e Matematica Applicata all'Economia, Università di Pisa 2000

[CLM] Campbell, J. Lo, A., MacKinlay, C.: The econometrics of financial markets. Princeton: Princeton University Press (1997)

[DN] de Jong, F., Nijman, T.: High frequency analysis of lead-lag relationships between financial markets. J. Empirical Finance **4**, 259–277 (1997)

[D] Donhal, G.: On Estimating the diffusion coefficient. J. Appl. Prob. **2**, 105–114 (1987)

[FZ] Florens-Zmirou, D.: Estimation de la variance à partir d'une observation discrètisée. C. R. Acad. Sci. **I**, 195–200 (1989)

[F] Föllmer, H.: Der Zufall in den Wirtschaftswissensschaften: Zur Rolle der Wahrscheinlichtkeitstheorie in der Theorie der Finanzmärte. Nova Acta Leopoldina **79**, 113–125 (1999)

[FT] Fournié, E., Talay,D.: Application de la statistique des diffusions à un modèle de taux d'intérêt. Finance **12**, 79–111 (1991)

[GCJ] Genon-Catalot, V., Jacod, J.: On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. Ann. Inst. Henri Poincaré **29**(1), 119–151 (1993)

[GLP] Genon-Catalot, V., Larédo, C., Picard, D.: Non parametric estimation of the diffusion coefficient by wavelets methods. Scand. J. Stat. **19**, 317–335 (1992)

[H] Hoffmann, M.: On estimating the diffusion coefficient: parametric versus nonparametric. Prépub. n. 521 Paris: Laboratoire de Probabilités 1999

[JK] Jiang, G., Knight, J.: A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model. Econ. Theory **13**, 615–645 (1997)

[LO] Lo, A.: Maximum likelihood estimation of generalized Itô processes with discretely sampled data. Econ. Theory **4**, 231–247 (1988)

[LM] Lo, A., MacKinlay, C.: An econometric analysis of nonsynchronous trading. J. Econ. **45**, 181–211 (1990)

[M76] Malliavin, P.: Stochastic calculus of variations and hypoelliptic operators. In: Itô, K. (ed.) International Symposium on SDE. Kyoto: 1976, pp. 195-263

[M97] Malliavin, P.: Stochastic analysis. (A Series of Comprehensive Studies in Mathematics, Vol. 313.) Berlin Heidelberg New York: Springer 1997

[NE] Nelson, D.: Filtering and forecasting with misspecified ARCH models I. J. Econ. **52**, 61–90 (1993)