

VOLATILITY COMPUTED BY TIME SERIES OPERATORS AT HIGH FREQUENCY

ULRICH A. MÜLLER

Olsen & Associates,

Seefeldstr. 233, CH - 8008 Zürich, Switzerland

E-mail: ulrichm@olsen.ch

Most financial markets produce inhomogeneous (i. e. unequally spaced) tick-by-tick data at high frequency. Recently developed time series operators can be used to directly compute statistical variables such as volatility from inhomogeneous data. This is not possible with traditional time series methods. Value-at-Risk computations require measurements of current volatility, but the conventional calculation from daily data, sampled at a certain daytime, is strongly sensitive to the choice of this daytime, revealing a high amount of stochastic noise. An alternative calculation from high-frequency, tick-by-tick data with time series operators is shown to have similar results, except for two advantages: distinctly reduced noise and up-to-date results at each tick. The time series operator method is flexible and computationally efficient. It can be used to express generating process equations and to compute the Value-at-Risk in real time.

1 Introduction

For banks and other financial institutions, risk measurement plays a central role. Risk levels must conform to the capital adequacy rule. An error in the computed risk level may thus affect a bank's investment strategy. The state of the art is measuring risk by analyzing daily data: using one market price per working day and per financial instrument. In this paper, the stochastic error of such a risk measure is demonstrated in a new way, concluding that using only daily data is insufficient.

The challenge for statisticians is to analyze the limitations of risk measures based on daily data and to develop better methods based on high-frequency data. This paper meets this challenge by introducing the time series operator method, applying it to risk measurement and showing its superiority when compared to a traditional method based on daily data.

Intra-day, high-frequency data is available from many financial markets nowadays. Many time series can be obtained at tick-by-tick frequency, including every quote or transaction price of the market. These time series are inhomogeneous because market ticks arrive at random times. Irregularly spaced series are called inhomogeneous, regularly spaced series are homogeneous. An example of a homogeneous time series is a series of daily data, where the data points are separated by one day (on a business time scale which omits the

weekends and holidays).

Inhomogeneous time series by themselves are conceptually simple; the difficulty lies in efficiently extracting and computing information from them. In most standard books ^{3,11,4} on time series analysis, the field of time series is restricted to homogeneous time series already in the introduction. This restriction induces numerous simplifications, both conceptually and computationally, and was almost inevitable before cheap computers and high-frequency time series were available.

In a recent paper, Zumbach and Müller ¹² have presented a new time series operator technique, together with a computationally efficient toolbox, to directly analyze and model inhomogeneous as well as homogeneous time series. This method has many applications, among them volatility or Value-at-Risk (VaR) computations tick by tick. Traditional methods such as RiskMetricsTM ⁶ can be shown to be special cases of the operator technique.

A comparison is made between VaR results based on daily data, sampled at a certain daytime, and results based on tick-by-tick data and the new time series operator technique. If using daily data, a surprising and (for practitioners) alarming sensitivity against the choice of the sampling daytime is observed. The stochastic noise seems higher than acceptable to risk managers. An alternative VaR computation based on tick-by-tick data and a new time series operator technique is shown to have similar properties, except for two advantages: distinctly reduced noise and availability of up-to-date results at each tick.

These results may encourage both researchers and practitioners to study and use high-frequency tick-by-tick data and, thereby, to apply methods based on time series operators. There exist some applications of the operator technology, such as the “Scale of Market Shocks” by Zumbach et al. ¹³ and some indicators of automated technical trading algorithms as described by Pictet et al. ¹⁰. An obvious step would be the implementation of real-time, tick-by-tick VaR computations in practice.

The time series operators can also be used in the formulation of old and new generating processes of time series. This opens new ways to develop process equations with new properties, also for inhomogeneous time series.

2 The time series operator technique

In this paper, only a minimum of a description is given, so the applications of the following sections can be understood. The theory of the time series operators has been derived and explained by Zumbach and Müller ¹².

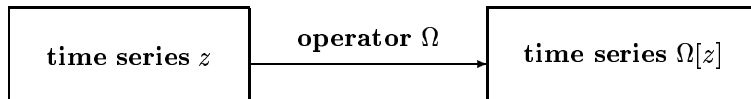


Figure 1: Time series operator Ω

2.1 Inhomogeneous time series

A time series z consists of elements or ticks z_i at times t_i . The sequence of these time points is required to be growing, $t_i > t_{i-1}$.

A general time series is inhomogeneous, meaning that the sampling times t_i are irregular. For a homogeneous time series, the sampling times are regularly spaced, $t_i - t_{i-1} = \delta t$.

For some discussions and derivations, a continuous-time version of z has to be assumed: $z(t)$. However, the operator methods that are eventually applied only need the discrete time series (t_i, z_i) .

The letter x is used to represent the time series of logarithmic middle prices, $x = (\ln p_{\text{bid}} + \ln p_{\text{ask}})/2$. This quantity is used in the applications.

2.2 Operators

An operator Ω , mapping from the space of time series into itself, is visualized in Figure 1. The resulting time series $\Omega[z]$ has a value of $\Omega[z](t)$ at time t . Important examples are moving average operators and more complex operators that construct a time series of volatility from a time series of prices.

Linear and translation-invariant operators are equivalent to a *convolution* with a kernel ω :

$$\Omega[z](t) = \int_{-\infty}^t \omega(t-t') z(t') dt' \quad (1)$$

A *causal* kernel has $\omega(t) = 0$ for all $t < 0$. No information from the “future” is used. If $\omega(t)$ is non-negative, $\Omega[z]$ is a weighted *moving average* of z whose kernel should be normalized:

$$\int_{-\infty}^t \omega(t-t') dt' = \int_0^{\infty} \omega(t) dt = 1 \quad (2)$$

The kernel $\omega(t)$ is the weighting function of the past. Figure 2 gives an example.

The range of an operator is the first moment of its kernel:

$$r = \int_{-\infty}^{\infty} \omega(t) t dt \quad (3)$$

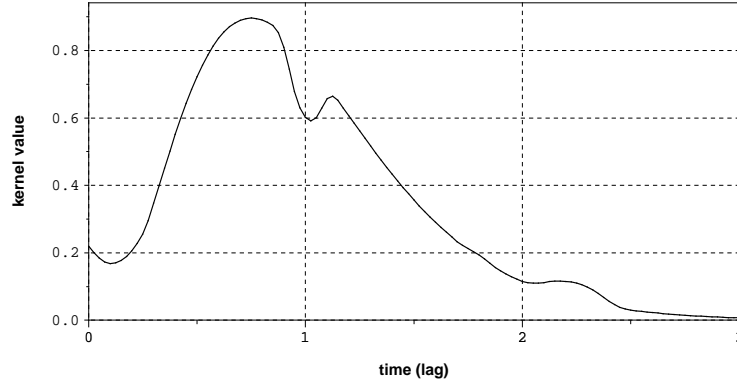


Figure 2: Example of the causal kernel $\omega(t)$ of a moving average

This indicates the characteristic depth of the past covered by the kernel.

Operators are useful for several reasons, as to be shown. One important aspect is to replace individual ticks from the market by local short-term averages of ticks. This mirrors the view of traders who consider not only the most recent tick but also the prices offered by other market makers within a short time interval.

2.3 The simple EMA operator

The Exponential Moving Average (EMA) operator is a simple example of an operator. It is written $\text{EMA}[\tau; z]$ and has an exponentially decaying kernel, as shown in Figure 3:

$$\omega(t) = \text{ema}(t) = \frac{e^{-t/\tau}}{\tau} \quad (4)$$

According to eqs. (3) and (4), the range of the operator $\text{EMA}[\tau; z]$ and its kernel is

$$r = \tau \quad (5)$$

The variable τ thus characterizes the depth of the past of the EMA.

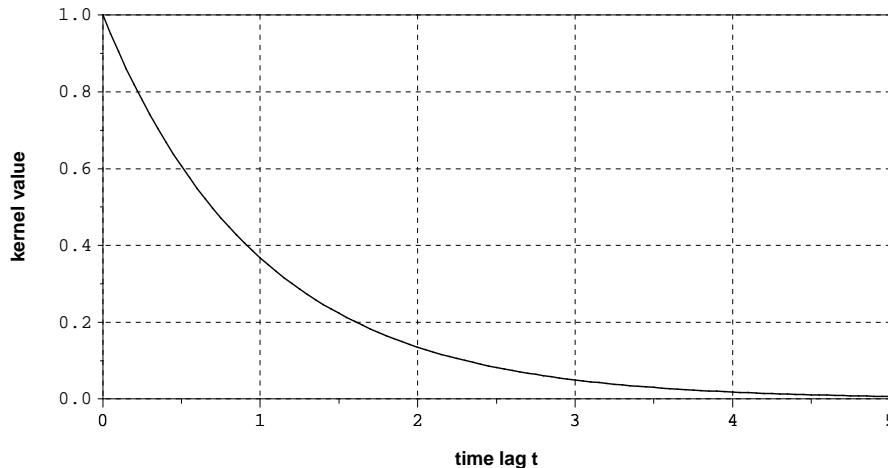


Figure 3: Kernel of the simple EMA operator $\text{EMA}[\tau; z]$. Here: $\tau = 1$.

The values of $\text{EMA}[\tau; z](t)$ can be computed by the convolution of eq. (1), if $z(t)$ is known in continuous time. This implies an integration whose numerical computation for many time points t is costly. Fortunately, there is an *iteration formula* that makes this computation much more efficient and, at the same time, solves the problem of discrete data. This means that we do not need to know the whole function $z(t)$; we just need the discrete time series values $z_i = z(t_i)$ at irregularly spaced time points t_i . The EMAs are calculated by the following iteration formula:

$$\text{EMA}[\tau; z](t_n) = \mu \text{EMA}[\tau; z](t_{n-1}) + (1-\mu) z(t_n) + (\mu-\nu) [z(t_n) - z(t_{n-1})] \quad (6)$$

with

$$\mu = e^{-\alpha} \quad , \quad \alpha = \frac{t_n - t_{n-1}}{\tau} \quad (7)$$

and

$$\nu = \frac{1-\mu}{\alpha} \quad (8)$$

This variable ν is related to the problem of using discrete data in a convolution defined in continuous time. We need an assumption on the behavior of $z(t)$ between the discrete time points t_i . Eq. (8) is based on the assumption of linear interpolation between points; other formulas for ν are implied by other

interpolation assumptions, as explained by Zumbach and Müller¹². In the case of assuming the value of the old tick for the whole interval before the new tick, the correct formula is

$$\nu = 1 \tag{9}$$

For a homogeneous time series, μ and ν are constants. A homogeneous time series can alternatively be regarded as a truly discrete time series to which interpolation does not apply. This is mentioned here because it is a popular approach used by traders (in their “Technical Analysis”) and by RiskMetrics. For such a discrete time series, $t_n - t_{n-1}$ is defined to be 1, and the following definition is appropriate:

$$\mu = \nu = \frac{1}{1 + \alpha} = \frac{\tau}{\tau + t_n - t_{n-1}} = \frac{\tau}{\tau + 1} \tag{10}$$

as shown by Müller (1991)⁷. The range of an operator for a genuine discrete time series has a new definition: $r = \sum_{i=0}^{\infty} \omega_i i$. For EMA, this means $r = \mu/(1 - \mu) = \tau$ with $\omega_i = (1 - \mu)\mu^i$. The μ and ν values resulting from eq. (10) are very similar to those of eqs. (7) and (8) as long as α is small.

The iteration equation (6) is computationally efficient, extremely so when compared to a numerical convolution based on eq. (1). No other operator can be computed as efficiently as the simple EMA operator. However, there are means to use the iteration equation (6) as a tool to efficiently compute operators with other kernels, as shown below.

An iteration formula is not enough. We have to initialize $\text{EMA}[\tau; z](t_0)$ at the start of the time series. For this, we can take $z_0 = z(t_0)$ or another typical value of z . This choice introduces an initial error of $\text{EMA}[\tau; z](t_0)$ which decreases exponentially with time. Therefore, we also need a build-up period for $\text{EMA}[\tau; z]$: a time period over which the values of $\text{EMA}[\tau; z]$ should not yet be applied because of their initial error. Build-up periods should be multiples of τ , e. g. 5τ . The choice of a large enough build-up period is discussed by Zumbach and Müller¹².

2.4 The operator $\text{EMA}[\tau, n; z]$

Time series operators can be convoluted: a time series resulting from an operator can be mapped by another operator. This is a powerful method to generate new operators with different kernels.

The $\text{EMA}[\tau, n; z]$ operator results from the repeated application of the same simple EMA operator, as defined by Müller (1991)⁷ for the first time. The following recursive definition applies:

$$\text{EMA}[\tau, k; z] = \text{EMA}[\tau; \text{EMA}[\tau, k - 1; z]] \tag{11}$$

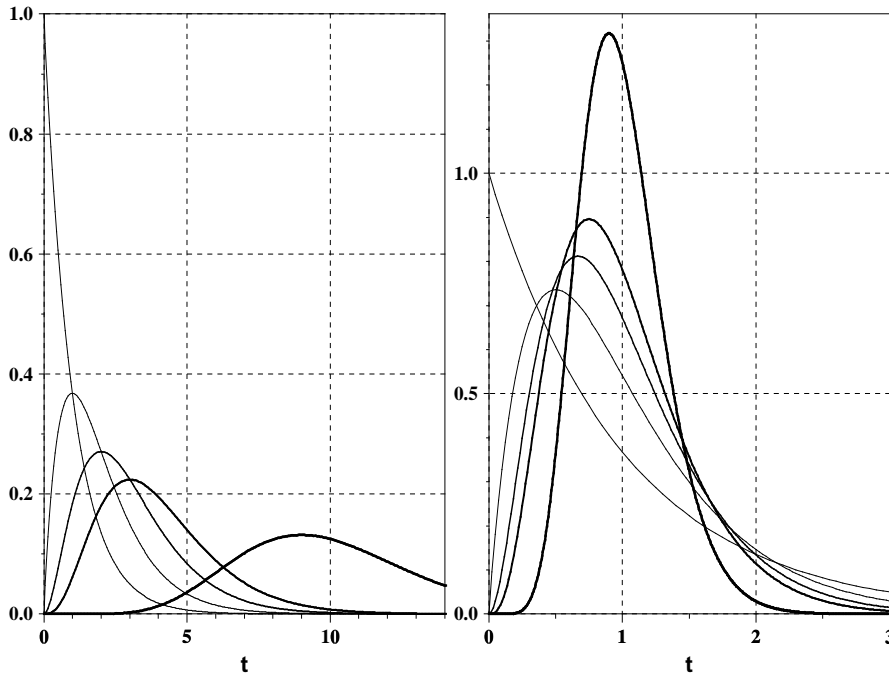


Figure 4: Kernels of $\text{EMA}[\tau, n; z]$. Curves for $n = 1$ (thin), 2, 3, 4, 10 (bold). Left graph: for $\tau = 1$; right graph: $r = n\tau = 1$.

with $\text{EMA}[\tau, 1; z] = \text{EMA}[\tau; z]$. The computationally efficient iteration formula of the simple EMA, eq. (6), can again be used; we have to apply it recursively (n times) for each new tick (t_i, z_i) . For ν , we insert eq. (8) which is based on a linear interpolation assumption between ticks. (This assumption is just a good approximation in some cases, as discussed by Zumbach and Müller¹²).

The operator $\text{EMA}[\tau, n; z]$ has the following kernel:

$$\text{ema}[\tau, n](t) = \frac{1}{(n-1)!} \left(\frac{t}{\tau}\right)^{n-1} \frac{e^{-t/\tau}}{\tau} \quad (12)$$

This kernel is plotted in Figure 4, for several n . For large n (e. g. $n = 10$ in Figure 4), the mass of the kernel is concentrated in a relatively narrow region around a time lag of $n\tau$. The corresponding operator can thus be seen as a smoothed backshift operator.

The family of functions of eq. 12 is related to Laguerre polynomials which

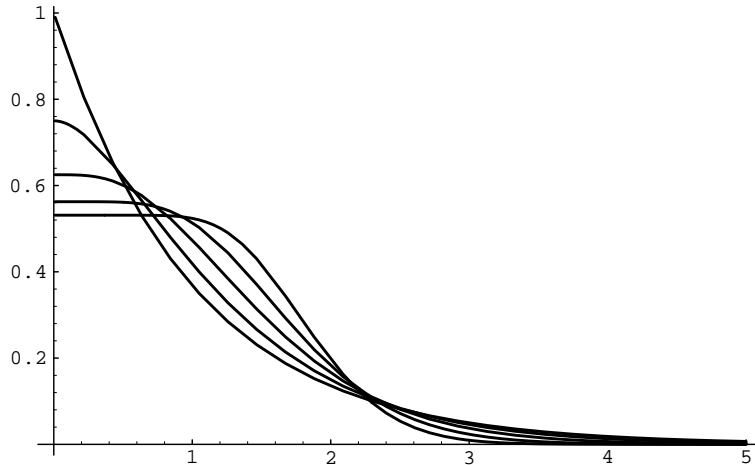


Figure 5: Kernels of $\text{MA}[\tau, n; z]$; $n = 1, 2, 4, 8, 16$, for $\tau = 1$.

are orthogonal with respect to the measure e^{-t} (for $\tau = 1$).

Operators, i. e. their kernels, can be linearly combined. This is a powerful method to generate more operators. Linear combinations of $\text{EMA}[\tau, n; z]$ operators with different n but identical τ values have kernels that correspond to expansions in Laguerre polynomials. This means that any kernel can be expressed as such a linear combination. The convergence of the Laguerre expansion may however be slow.

In practice, a small set of useful operators can be prepared with all the kernels needed. Aside from the discussed expansion, it is also possible to linearly combine kernels with different τ values. Some useful types of combined operators are presented by Zumbach and Müller¹².

2.5 The operator $\text{MA}[\tau, n; z]$

The moving average (MA) operator has kernels with useful properties as shown in Figure 5. It is constructed as a sum of $\text{EMA}[\tau, n; z]$ operators:

$$\text{MA}[\tau, n] = \frac{1}{n} \sum_{k=1}^n \text{EMA}[\tau', k] \quad \text{with } \tau' = \frac{2\tau}{n+1} \quad (13)$$

The variable τ' is chosen such that the range of $\text{MA}[\tau, n]$ is $r = \tau$, independent of n . The kernel form depends on the choice of n . For $n = 1$, we obtain a simple

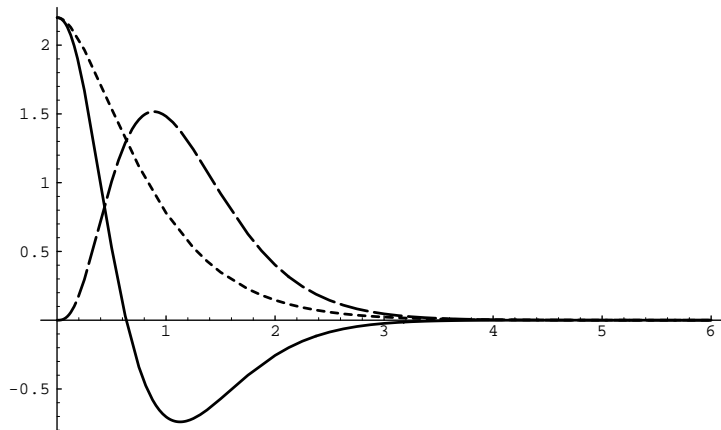


Figure 6: Kernel of the differential operator $\Delta[\tau]$: full curve. Here: $\tau = 1$; dotted curve: first two terms of the operator; dashed curve: the last term $2 \text{EMA}[\alpha\beta\tau, 4]$.

EMA operator, for $n = \infty$ the rectangularly shaped kernel of a simple moving average with constant weight up to a limit of 2τ . This simple rectangular moving average has a serious disadvantage in its dynamic behavior: additional noise when old observations are abruptly dismissed from the rectangular kernel area. Kernels with finite n are better because of their smoothness; the memory of old observations fades gradually rather than abruptly.

The formula for the $\text{MA}[\tau, n]$ kernel is

$$\text{ma}[\tau, n](t) = \frac{n+1}{n} \frac{e^{-t/\tau'}}{2\tau} \sum_{k=0}^{n-1} \frac{1}{k!} \left(\frac{t}{\tau'}\right)^k \quad (14)$$

Many other kernel forms can be constructed through different linear combinations of $\text{EMA}[\tau, n; z]$ and other operators.

2.6 From returns to differentials

Most statistics in finance is based on returns: price changes rather than prices. Simple returns have a rather noisy behavior over time; we often want differences between local averages of x : *smoothed* returns.

Smoothed returns are computed by differential operators. Examples:

- $x - \text{EMA}[\tau, n; x]$, where the EMA replaces $x(t - \tau)$. This is used by the application of section 3.2.

- $\text{EMA}[\tau_1, n_1] - \text{EMA}[\tau_2, n_2]$, with $\tau_1 < \tau_2$ or $n_1 < n_2$.
- $\Delta[\tau] = \gamma \{ \text{EMA}[\alpha\tau, 1] + \text{EMA}[\alpha\tau, 2] - 2 \text{EMA}[\alpha\beta\tau, 4] \}$, with $\gamma = 1.22208$, $\beta = 0.65$ and $\alpha^{-1} = \gamma(8\beta - 3)$. This is normalized, so $\Delta[\tau, 1] = 0$, $\Delta[\tau, t] = 1$. The kernel of this differential operator proposed by Zumbach and Müller¹² is plotted in Figure 6.

The expectation value of squared smoothed returns may differ from that of the corresponding simple returns. This has to be accounted for when comparing the two concepts, for example in terms of a factor c in eq. (20).

2.7 Volatility measured by operators

Volatility is a central term in risk measurement and finance in general, but there is no unique, universally accepted definition. There are volatilities derived from option market prices and volatilities computed from diverse model assumptions. In this paper, the focus is on historical volatility: a volatility computed from recent data of the underlying instrument with a minimum of parameters and model assumptions.

For computing the time series of such a volatility, a time series operator is again the suitable tool. We first define the nonlinear moving norm operator:

$$\text{MNorm}[\tau, p; z] = \{ \text{MA}[\tau; |z|^p] \}^{1/p} \quad (15)$$

This operator is based on a linear MA operator (where we are free to choose any positive, causal kernel), it is nonlinear only because a nonlinear function of the basic time series variable z is used. $\text{MNorm}[\tau, p; z]$ is homogeneous of degree 1.

The volatility of a time series x can now be computed with the help of the moving norm operator:

$$\text{Volatility}[\tau_1, \tau_2, p; x] = \text{MNorm} \left[\frac{\tau_1}{2}, p; \Delta[\tau_2; x] \right] \quad (16)$$

$$= \left\{ \text{MA} \left[\frac{\tau_1}{2}; |\Delta[\tau_2; x]|^p \right] \right\}^{1/p} \quad (17)$$

This is the moving norm of (smoothed) returns. With $p = 2$, it is a particular version of the frequently used RMS value. However, some researchers had and have good reasons to choose a lower value such as $p = 1$ in their special studies.

Eq. (17) is based on a moving average (MA) and a differential (Δ) operator. In principle, we may choose any MA and Δ operator according to our preference. In the applications of section 3, this choice is made explicit.

The volatility definition of eq. (17), as any definition of historical volatility, necessarily has *two* timing parameters:

1. the size of the return measurement intervals: τ_2 ;
2. the size of the total moving sample: τ_1 , often $\gg \tau_2$; defined as the double range of the used MA. The MA operator has a range (center of gravity of the kernel) of $\tau_1/2$.

3 Application: volatility computation in risk management

Computing recent volatility is a central ingredient of risk assessment in risk management. Here it serves as an example to demonstrate the usefulness and superiority of time series operators.

The RiskMetricsTM method⁶ is chosen as a well-known example. First it is shown to be a special application of the time series operator technique. Then a better volatility computation method, also based on the time series operator technique, is proposed. Thus two approaches are compared:

1. The RiskMetrics method, based on an IGARCH model with *working-daily data*.
2. A *tick-by-tick* alternative, following RiskMetrics as closely as possible, based on time series operators.

In both cases, squared volatility is defined as the expectation σ^2 of squared, working-daily changes of the logarithmic middle price x . This is for the sake of a meaningful comparison; it does not imply that using squared returns is necessarily the best choice for an optimal volatility definition.

3.1 Conventional computation based on daily data

The RiskMetrics method is based on an IGARCH model. Its volatility formula gives the conditional expectation of the squared return assuming IGARCH:

$$\sigma^2(t) = \mu \sigma^2(t - 1 \text{ wday}) + (1 - \mu) [x(t) - x(t - 1 \text{ wday})]^2 \quad (18)$$

with $\mu = 0.94$. This is just an EMA iteration which can also be written in our operator notation:

$$\sigma^2(t) = \text{EMA} [\tau = 15.67 \text{ wdays}; [x(t) - x(t - 1 \text{ wday})]^2] \quad (19)$$

evaluated at discrete time points separated by 1 working day (= 1 wday); with EMA range $\tau = \mu/(1 - \mu)$ in working days, following eq. (10).

Thanks to the regularity of the underlying homogeneous time series, $\mu = 0.94$ is a constant. In general, the constancy of μ makes the operator technique particularly efficient for homogeneous time series.

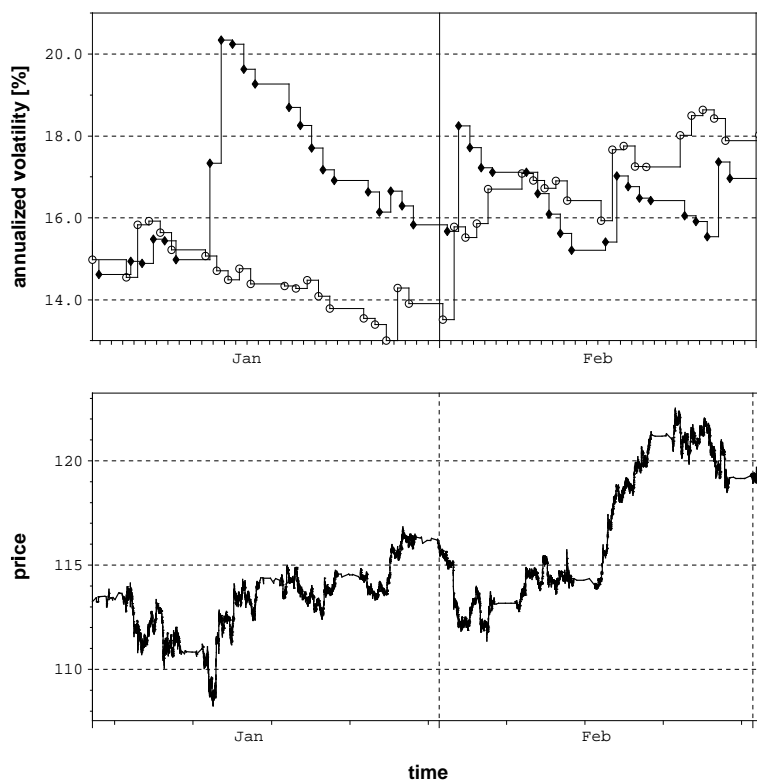


Figure 7: Standard RiskMetrics volatility. Circles: data sampled at 7 am GMT; diamonds: data sampled at 5 pm GMT. Independent computations. FX rate: USD/JPY; period: Jan 1999 - Feb 1999. The price is plotted against time on the lower graph.

Figure 7 shows the resulting volatility as a function of time, in an empirical example from the foreign exchange (FX) market: USD/JPY data in January and February 1999. The volatility is computed only once per working day, at a given daytime; the resulting volatility value is valid until it is replaced by a new one, one working day later.

In Figure 7, two such volatilities are plotted. The difference between the two curves solely originates from the choice of daytime when the raw data x is sampled and the volatility is computed by eq. (18) or (19). One curve is sampled at 7 am GMT which is a time in the late afternoon of East Asian time zones – a suitable daytime for the daily risk calculations of an East Asian risk

manager. The other curve is sampled at 5 pm GMT – a suitable daytime for a risk manager in London.

The differences between the two curves are surprisingly large: up to 25%, an alarming uncertainty for risk managers. Risk levels are linked to a bank's capital through the capital adequacy rule, so differences in risk measurements have a major impact on banking! In our case, two risk managers measure very different volatility and thus risk levels for the *same* financial instrument, just because they live in different time zones. A difference can persist over weeks, as shown in Figure 7. This figure is just an example. The same surprisingly strong effect can be found also for other financial instruments, sampling periods and choices of daytime.

Both deviating volatility values cannot be right at the same time; there must be an error in these values. This error is of stochastic nature; there is no systematic bias dependent on the daytime. In Figure 7, the difference between the two curves is neither always positive nor negative; it changes its sign.

Figure 7 demonstrates the large stochastic error of the RiskMetrics method. The large size of this error has two main reasons:

1. The rather small range of the kernel of some 16 working days. The number of independent observations is limited. We cannot essentially change this fact, because the choice of a short range is also motivated by the goal of fast adaptivity to new market events.
2. The results depend on only one observation per day, taken at a certain daytime. All the other information on prices of the day is thrown away. The value at that daytime may be little representative for the full day: it may be located on top of a short-lived local peak of the price curve. This is indeed the reason for the large deviations of the two curves in Figure 7. The effect is exacerbated by the known fact that returns have a heavy-tailed distribution function: extreme (intra-day) events dominate the statistics.

The focus of this study is not so much the behavior of RiskMetrics (IGARCH), but the problems of using homogeneous, daily data in general, no matter which GARCH-type or other model is investigated. The significance of most results can be improved by using all the available information, tick by tick, as shown in the next section.

3.2 Tick-by-tick volatility computation

For the sake of a fair comparison, a tick-by-tick volatility computation is introduced that follows RiskMetrics as closely as possible. There are only two

innovative modifications:

- The squared volatility $\sigma^2(t)$ is computed at every available tick, not only once per working day.
- Simple returns are replaced by operator-based, smoothed returns.

Nothing is changed otherwise; the sampling range of 15.67 working days and the working-daily nature of (smoothed) returns are preserved.

The new volatility measure is again defined in operator notation (where “wdays” stands for working days):

$$\sigma^2 = c \text{EMA} \left[\tau = 15.67 \text{ wdays}; (x - \text{EMA}[\tau = 1 \text{ wday}, 4; x])^2 \right] \quad (20)$$

This is just a special case of eq. (16). The computation is efficiently done at every new tick, repeatedly using the iteration formula (6). This works not only for the simple EMA but also for $\text{EMA}[\tau, 4; x]$ as explained in section 2.4.

The constant c compensates for the fact that we use smoothed returns $x - \text{EMA}[\tau, 4; x]$ as introduced in section 2.6 instead of the simple returns of section 3.1. In the case of x following a Gaussian random walk, the theoretically correct value is $c = 128/93$. Using this factor eliminates a systematic bias of the tick-by-tick volatility as compared to the RiskMetrics volatility.

Eq. (20) is computed on a special business time scale defined as follows. The 49 weekend hours from Friday 8 pm GMT to Sunday 9 pm GMT are compressed to the equivalent of only 1 hour outside the weekend. This fully corresponds to the time scale of RiskMetrics which omits the weekend days. A more sophisticated and appropriate choice of the business time scale would be the ϑ -time of Dacorogna et al. (1992)¹, but this is avoided here in order to keep the approach as close to RiskMetrics as possible.

Figure 8 shows the resulting volatility as a function of time. The same financial instrument and sampling period is studied as in Figure 7. High-frequency data is available here. Now, the large differences between values at 7 am GMT and 5 pm GMT have vanished. The observations at these daytimes appear as points on one continuous, consistent curve. In fact, we can obtain volatility values at any daytime now, not just once or twice a day. A risk manager in London essentially measures the same risk of the instrument as a risk manager in East Asia, as should be expected in normal situations. Only if a dramatic event between the two daytimes of measurement happens, the risk levels deviate. This is natural; the operator-based volatility quickly reacts to dramatic events, as can be seen in Figure 8.

The variations of the volatility level over time are moderate in Figure 8. The extreme volatility minima and maxima of Figure 7 which are mostly due

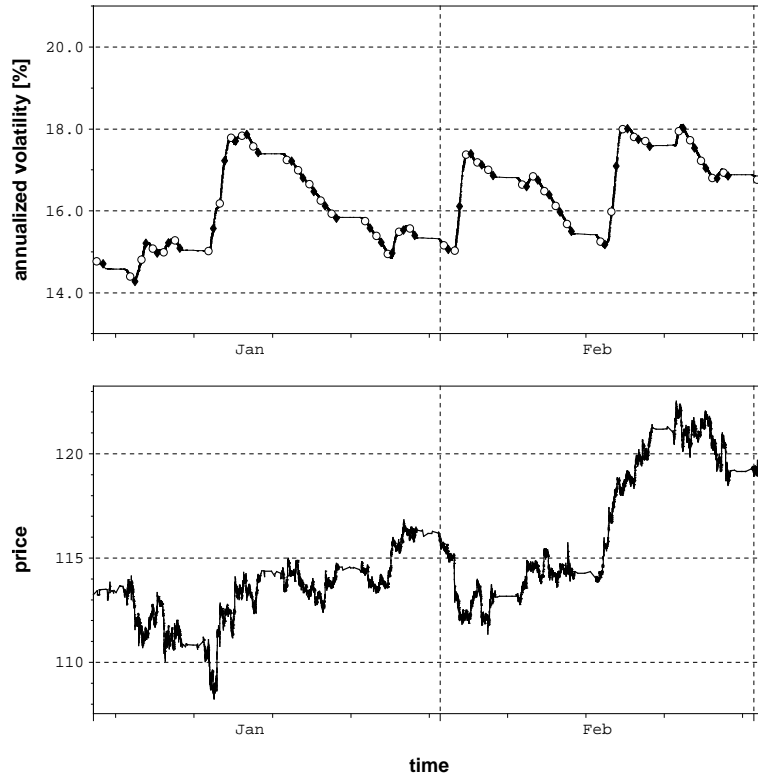


Figure 8: Operator-based tick-by-tick volatility. Circles: values at 7 am GMT (7:00); diamonds: values at 5 pm GMT (17:00). FX rate: USD/JPY; period: Jan 1999 - Feb 1999. The price is plotted against time on the lower graph.

to stochastic noise have vanished. The new tick-by-tick volatility has less stochastic noise than the RiskMetrics volatility, although the moving sample range of 15.67 working days is the same.

The curves of Figures 7 and 8 are combined in Figure 9. Here, we can see that the tick-by-tick volatility (bold curve) has a similar dynamic behavior as the RiskMetrics volatilities while avoiding extreme oscillations due to stochastic noise.

The lower noise level of the tick-by-tick volatility is now plausible, but we need scientific evidence for this. In the general case, such evidence can be gained through Monte-Carlo studies based on a certain process assumption,

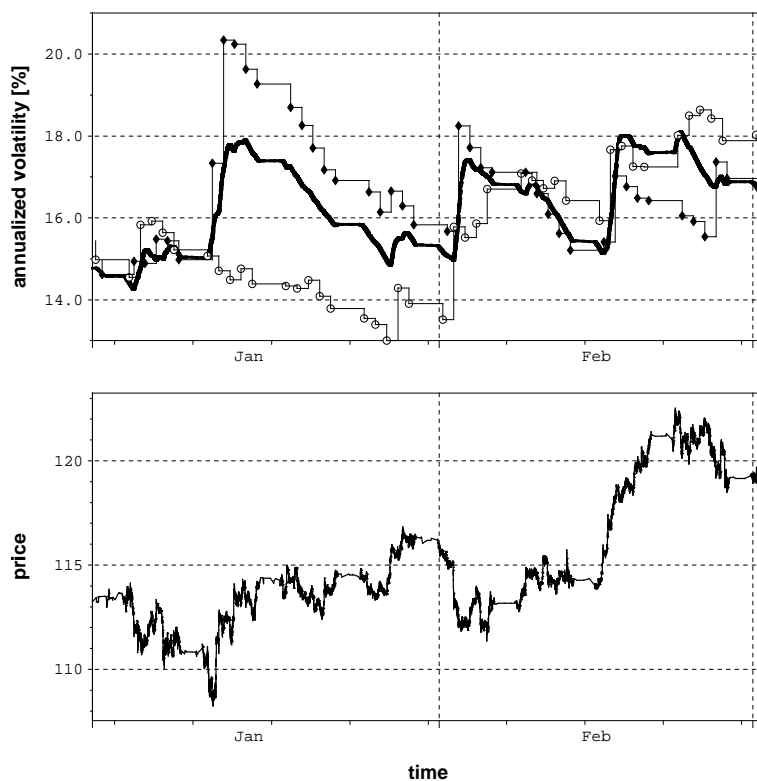


Figure 9: Comparing RiskMetrics to operator-based volatility. Bold curve: operator-based; circles: RiskMetrics 7 am GMT; diamonds: RiskMetrics 5 pm GMT. FX rate: USD/JPY; period: Jan 1999 - Feb 1999. The price is plotted below.

comparing the error variances of the RiskMetrics volatility and the tick-by-tick volatility. In the case of Gaussian random walk, we have even stronger evidence: an exact result by Müller (1993)⁸. By using continuously overlapping returns instead of non-overlapping returns, the error variance of the empirically determined σ^2 is reduced to $2/3$ of the original value.

The tick-by-tick operator is indeed using (almost) continuously overlapping returns. In addition to this, it is based on smoothed rather than simple returns, which also leads to a reduction of stochastic noise.

Other advantages of tick-by-tick, operator-based methods are the efficient computation based on iterations and the updating at every new tick. Thanks

to fast updates, the volatility measure can quickly react to new market events such as shocks, at any daytime.

3.3 VaR computation in real-time

Conventional Value-at-Risk (VaR) computations are done once a day, usually in the evening. Traders and portfolio managers typically do not know their current risk; they just know yesterday evening's risk. What they would really need is a *real-time* VaR computation, updated as quickly as they can change their positions.

The tick-by-tick operator proposed in section 3.2 and eq. (20) is a tool to make a real-time VaR possible. In addition to this, a corresponding tool to compute moving covariances or correlations is needed. This is possible with similar operators; some operators are proposed by Zumbach and Müller¹².

A real-time VaR computed according to these guidelines would still be rather similar to RiskMetrics, except for lower noise and a higher updating frequency. There are many criticisms of RiskMetrics that would apply to it, too. Some researchers, for example, replace the IGARCH model by another GARCH-type model. Other researchers such as McNeil et al.⁵ focus on the behavior of extreme price changes which may follow other laws than average-size changes. Moreover, return observations over intervals other than daily, for example hourly or weekly returns, contain valuable information that should also be used in a VaR computation.

All these possible refinements of risk measurement are not discussed in this paper. They have to be addressed in future research studies that will eventually lead to a more mature solution of real-time VaR computation. But it seems clear that this optimal solution will also be based on tick-by-tick operators.

4 Process equations based on operators

Processes of the ARMA and GARCH families can be expressed in terms of time series operators, as we have seen for IGARCH in eq. (19). The squared conditional volatility of a GARCH(1,1) process, for example, can be written as follows:

$$\sigma^2(t) = c + a\sigma^2(t') + b[x(t') - x(t' - \Delta t)]^2 \quad (21)$$

where $t' = t - \Delta t$. The following alternative notation is based on a simple EMA operator:

$$\sigma^2(t) = \frac{c}{1-a} + \frac{b}{1-a} \text{EMA} \left[\frac{a}{1-a}; [x(t') - x(t' - \Delta t)]^2 \right] \quad (22)$$

This rephrased form of GARCH(1,1) is a starting point of interesting new developments. Initially, it applies to a discrete, homogeneous time series in the sense of eq. (10), but it allows for a direct and efficient computation of the GARCH(1,1) volatility from inhomogeneous data, since the operator technique is also suited to inhomogeneous time series. This is not possible with conventional methods.

Moreover, eq. (22) can be modified to obtain other processes. The kernel of the EMA operator can, for example, be replaced by other kernels. The return $x(t') - x(t' - \Delta t)$ can be replaced by a smoothed return computed by a differential operator which reflects the perception of market participants better than the simple return.

Dacorogna et al. (1998)² have introduced the EMA-HARCH process to model some empirical facts of high-frequency data in finance: the long memory of volatility, the fat tails of the return distribution and the asymmetric causality between fine-grained (high-resolution) and coarse-grained (low-resolution) volatilities as found by Müller et al.⁹. This is one of the first processes whose equation is written with the help of a time series operator:

$$\sigma^2(t) = c + \sum_{j=1}^n C_j \sigma_j^2(t) \quad (23)$$

The “partial volatilities” σ_j^2 correspond to market segments and are written in terms of the EMA operator:

$$\sigma_j^2(t) = \text{EMA} \left[\frac{k_{j+1} - k_j}{2} \Delta t; [x(t') - x(t' - k_j \Delta t)]^2 \right] \quad (24)$$

with $k_1 = 1$ and $k_j = 4^{j-2} + 1$ for $j > 1$.

There are many more possibilities and good reasons to introduce time series operators in process equations to a much larger extent in the future.

5 Conclusions

Most financial markets produce inhomogeneous data, irregularly spaced in time. The time series operators presented in this paper are able to directly use inhomogeneous data to estimate statistical variables such as volatilities. This computation is efficient thanks to simple iteration formulas. The operator technique efficiently works also for homogeneous, equally spaced data.

Starting from the simple exponential moving average (EMA) operator, large families of operators with different kernels and different purposes can be constructed. A wider overview of these more complex operators, which are still

computationally efficient, is given by Zumbach and Müller¹². One example is a tick-by-tick Fourier analysis on a moving sample.

Thanks to averaging, the operator technique often produces results with less noise (lower stochastic errors) than conventional methods based on homogeneous time series. This is also the case for the main application of this paper: volatility of daily returns as needed for Value-at-Risk (VaR) computations. The conventional RiskMetrics method has a rather high stochastic error which is demonstrated in a new way: volatility is computed twice with daily data. In one case, the data is always sampled at 7 am GMT (late afternoon in East Asia), in the other case at 5 pm (late afternoon in London). The results of the two computations can differ by some 25% for many days in a row - an embarrassing fact for risk managers. The proposed tick-by-tick alternative based on time series operators does not have this sensitivity against the choice of the daytime and has less noise, while keeping the essential characteristics of the RiskMetrics method: it is still based on daily returns and still has a range (center of gravity of the kernel) of around 16 working days.

The logical next step will be to design a complete market risk analysis based on the time series operator technique. Its result will be a real-time VaR, updated with every new incoming tick from a market, with less noise than the corresponding results from conventional methods.

Finally, the operator technique can be used to formulate time series generation process equations. This is possible for well-known processes such as GARCH and new, more complex processes such as EMA-HARCH. The formulation in terms of operators has many advantages: flexibility and applicability to irregularly spaced time series.

Acknowledgments

This paper resulted from my presentation at the Hong Kong International Workshop on Statistics in Finance, 5 - 9 July 1999. I would like to thank the organizing committee of this successful workshop for the excellent organization and for having given me the opportunity to present the time series operator method. The presented work is not mine alone; I am grateful for the help and the contributions of Gilles Zumbach and my other colleagues at Olsen & Associates.

References

1. Dacorogna M. M., Müller U. A., Nagler R. J., Olsen R. B., and Pictet O. V., 1993, *A geographical model for the daily and weekly seasonal*

- volatility in the FX market*, Journal of International Money and Finance, **12**(4), 413–438.
2. Dacorogna M. M., Müller U. A., Olsen R. B., and Pictet O. V., 1998, *Modelling short-term volatility with GARCH and HARARCH models*, published in “Nonlinear Modelling of High Frequency Financial Time Series” edited by Christian Dunis and Bin Zhou, John Wiley, Chichester, 161–176.
 3. Granger C. W. J. and Newbold P., 1977, *Forecasting economic time series*, Academic Press, London.
 4. Hamilton J. D., 1994, *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.
 5. McNeil A. J. and Frey R., 1998, *Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach*, Preprint from the ETH Zürich, August 27, 1–28.
 6. J. P. Morgan, 1996, *RiskMetrics – technical document*, Technical report, J. P. Morgan and International marketing – Reuters Ltd.
 7. Müller U. A., 1991, *Specially weighted moving averages with repeated application of the EMA operator*, Internal document UAM.1991-10-14, Olsen & Associates, Seefeldstrasse 233, 8008 Zürich, Switzerland.
 8. Müller U. A., 1993, *Statistics of variables observed over overlapping intervals*, Internal document UAM.1993-06-18, Olsen & Associates, Seefeldstrasse 233, 8008 Zürich, Switzerland.
 9. Müller U. A., Dacorogna M. M., Davé R. D., Olsen R. B., Pictet O. V., and von Weizsäcker J. E., 1997, *Volatilities of different time resolutions – analyzing the dynamics of market components*, Journal of Empirical Finance, **4**(2-3), 213–239.
 10. Pictet O. V., Dacorogna M. M., Müller U. A., Olsen R. B., and Ward J. R., 1992, *Real-time trading models for foreign exchange rates*, Neural Network World, **2**(6), 713–744.
 11. Priestley M. B., 1989, *Non-linear and non-stationary time series analysis*, Academic Press, London.
 12. Zumbach G. O. and Müller U. A., 1998, *Operators on inhomogeneous time series*, submitted for publication; internal document GOZ.1998-11-01, Olsen & Associates, Seefeldstrasse 233, 8008 Zürich, Switzerland.
 13. Zumbach G. O., Dacorogna M. M., Olsen J. L., and Olsen R. B., 1998, *Introducing a scale of market shocks*, submitted for publication; internal document GOZ.1998-10-01, Olsen & Associates, Seefeldstrasse 233, 8008 Zürich, Switzerland.