Quasi-likelihood estimation of GARCH models with missing values

Marcos H. Cascone
State University of Campinas, Campinas, Brazil - marcos.cascone@gmail.com

Luiz K. Hotta
State University of Campinas, Campinas, Brazil - hotta@ime.unicamp.br

Abstract

This work presents a new method to deal with missing values in financial time series. Previous works are generally based in state space models and Kalman filter and few consider ARCH family models. The traditional approach is to bound the data together and perform the estimation without considering the presence of missing values. The existing methods generally consider missing values in the returns. The proposed method considers the presence of missing values in the price of the assets instead of in the returns. The performance of the method in estimating the parameters and the volatilities is evaluated through a Monte Carlo simulation. Value at risk is also considered in the simulation. An empirical application to NASDAQ 100 Index series is presented.

Keywords: Financial time series; incomplete time series; conditional expectation and variance; volatility of aggregated returns.

1. Introduction

Autoregressive conditionally heteroscedastic (ARCH) models, introduced by Engle (1982), and the Bollerslev's (1986) generalized ARCH (GARCH) model, have been two of the most popular models applied to estimate the volatility in finance series. One reason for their popularity is the fact that depending on some past observations and disturbance, the likelihood can be written in an analytical way when there are no missing observations.

However, a stylized fact of financial series is the presence of missing observations, which can be attributed to different reasons. There are mainly three types of approach in this case: (i) ignore the missing data and bound the observed data together; (ii) fill in the missing data with an imputation method; or (iii) take into account the missing observations without any type of imputation.

A review of an important method of imputation can be found in Dempster et al. (1977), where the authors introduce the EM algorithm (expectation-maximization). Furthermore, the Kalman filter can be used to input data in a set with incomplete observations, which can be found in Brockwell and Davis (2009) and Ossandón and Bahamonde (2012), in the last case applied to the ARCH models.

The goal of this paper is to deal with missing values in GARCH models taking them into account, but in the sense of (iii) approach, i.e., no imputation will be done. To circumvent the problem, we will extend the quasi maximum likelihood, rewriting the conditional expectations and conditional variances in the neighborhood of the missing observation. To construct the conditional expectation and variance, the dependence structure is take into account and, in a certain way, the information contained before the missing value is carried to the next observations. With this methodology, we expect to obtain better results when compared to the method where data is bound together, with little computational resources.

The outline of the paper is the following: Section 2 introduces the quasi log-likelihood; Section 3 presents part of the results of the Monte Carlo experiment considering Gaussian and Student-t innovations; Section 4 presents an application to real data with missing values and in Section 5 presents our final remarks.

2. Quasi-likelihood in the presence of missing values

Let $\{r_t\}$, the returns of an asset be a time process with zero mean and let $\mathcal{F}_{t-1}$ be the information available from the infinite past to the present, that is, $\{r_{t-i},\ i = 1, 2, \ldots\}$. The GARCH $(1, 1)$ model (Bollerslev, 1986) is given by

$$r_t = \sigma_t \epsilon_t, \tag{1}$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{2}$$

where $\sigma_t^2 = Var(r_t|\mathcal{F}_{t-1}) = E(r_t^2|\mathcal{F}_{t-1})$ is the conditional variance function and $\{\epsilon_t\}$ is a sequence of independent and identically distributed (i.i.d.) random variables, with $E(\epsilon_t) = 0$ and $Var(\epsilon_t) = 1$. Furthermore, ϵ_t is independent of $\mathcal{F}_{t-1}$, α_0, α_1 and β_1 are non-negative, $\alpha_0 > 0$ and $\alpha_1 + \beta_1 < 1$. In finance, GARCH models are usually applied to returns (log returns) of an asset price.

The log-likelihood function for the GARCH(1,1) model, conditional on σ_1^2 and r_1, is given by

$$\ell(\boldsymbol{\theta}; r_n, r_{n-1}, ..., r_1 | r_1, \sigma_1^2) = \sum_{t=2}^{n} [\log f_\epsilon(r_t|\boldsymbol{\theta}, \mathcal{F}_{t-1}) - \log(\sigma_t)] = \sum_{t=1}^{n} \ell_t(\boldsymbol{\theta}), \tag{3}$$

where $f_\epsilon(\cdot)$ is related to the distribution of ϵ_t, $\boldsymbol{\theta}$ is the parameter vector $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta})$, and $\mathcal{F}_{t-1}$ is the information up to time $(t-1)$.

For the sake of simplicity, suppose that the price of the underlying asset is observed from time one to time n, but it is not observed at time $m+1$. In terms of returns, this means that we observe only $\boldsymbol{r} = (r_1, ..., r_m, r_{m+1} + r_{m+2}, r_{m+3}, ..., r_n)$. The log-likelihood is given by

$$\begin{aligned} \ell(\boldsymbol{\theta}; \boldsymbol{r} | r_1, \sigma_1^2) &= \sum_{t=2}^{m} \log f(r_t | r_{t-1}, ..., r_1, \sigma_1^2) + \log f(r_{m+1} + r_{m+2} | r_m, ..., r_1, \sigma_1^2) \\ &+ \sum_{t=m+3}^{n} \log f(r_t | r_{t-1}, , ..., r_{m+3}, r_{m+1} + r_{m+2}, r_m, ..., r_1, \sigma_1^2). \end{aligned} \tag{4}$$

For the QGLL we consider that all the distributions are Gaussian and the volatilities are given by Equation 2, except for the second term of 4 and for σ_{m+3}^2, which is presented later. From $m+3$ until n the Gaussian distribution is only true if $V_{m+2}(r_{m+3})$ (given later by Equation 9) is the true value of σ_{m+3}^2. Because this is not true, we have an approximation. Since the conditional mean is zero, we only have to evaluate the conditional variances.

Denoting expectation and variance conditional to $\mathcal{F}_t$ as $E_t(X) = E(X|\mathcal{F}_t)$ and $V_t(X) = V(X|\mathcal{F}_t)$, respectively, we have:

$$V_m(r_{m+1}) = E_m(r_{m+1}^2) = \alpha_0 + \alpha_1 r_m^2 + \beta_1 \sigma_m^2 \tag{5}$$
$$V_m(r_{m+k+1}) = E_m(r_{m+k+1}^2) = \sum_{j=0}^{k-1} \left[\alpha_0 (\alpha_1 + \beta_1)^j\right] + (\alpha_1 + \beta_1)^k V_m(r_{m+1}),\ k = 1, \ldots. \tag{6}$$

Because conditional to $\mathcal{F}_m$, the returns r_{m+1} and r_{m+2} are uncorrelated, the conditional variance of $r_{m+1} + r_{m+2}$ is equal to the sum of the conditional variances, i.e.,

$$V_m(r_{m+1} + r_{m+2}) = \alpha_0 + (\alpha_1 + \beta_1 + 1) V_m(r_{m+1}), \tag{7}$$

where $V_m(r_{m+1})$ is observed when the parameters are known.

To find the distribution of r_{m+3} conditional on $\mathcal{F}_{m+2} = \{\mathcal{F}_m, r_{m+1} + r_{m+2}\}$, we need to evaluate its variance, which is given by

$$V_{m+2}(r_{m+3}) = E_{m+2}(r_{m+3}^2) = E_m(r_{m+3}^2 | r_{m+1} + r_{m+2}). \tag{8}$$

Since $r_{m+3}^2 = \sigma_{m+3}^2 \epsilon_{m+3}^2$ and the terms σ_{m+3}^2 and ϵ_{m+3}^2 are independent of each other, the last expectation can be written as $E(\sigma_{m+3}^2 | \mathcal{F}_m, r_{m+1} + r_{m+2})$. Then

$$\begin{aligned} V(r_{m+3}|\mathcal{F}_{m+2}) &= E_m[(\alpha_0 + \alpha_1 r_{m+2}^2 + \beta_1 \sigma_{m+2}^2) | r_{m+1} + r_{m+2}] \\ &= \alpha_0 + \alpha_1 \left[V_m(r_{m+2}|r_{m+1} + r_{m+2}) + E_m^2(r_{m+2}|r_{m+1} + r_{m+2})\right] + \\ &\quad \beta_1 E_m(\alpha_0 + \alpha_1 r_{m+1}^2 + \beta_1 \sigma_{m+1}^2 | r_{m+1} + r_{m+2}) \\ &= \alpha_0 + \alpha_1 \left[V_m(r_{m+2}|r_{m+1} + r_{m+2}) + E_m^2(r_{m+2}|r_{m+1} + r_{m+2})\right] + \\ &\quad \beta_1 \left\{\alpha_0 + \alpha_1 \left[V_m(r_{m+1}|r_{m+1} + r_{m+2}) + E_m^2(r_{m+1}|r_{m+1} + r_{m+2}\right] + \beta_1 \sigma_{m+1}^2\right\}. \end{aligned} \tag{9}$$

From Equation 6 and 7, we have, respectively, the variance of r_{m+2} and $r_{m+1}+r_{m+2}$ conditional on $\mathcal{F}_m$. Assuming that the pair $(r_{m+j}, r_{m+1}+r_{m+2})$ has bivariate normal distribution, with $j=1,2$, the terms of Equation 9 are given by

$$V_m(r_{m+j}|r_{m+1}+r_{m+2}) \;=\; V_m(r_{m+j})(1-\rho_j^2) \;\text{ and }\; E_m(r_{m+j}|r_{m+1}+r_{m+2}) = \rho_j^2(r_{m+1}+r_{m_j+2})$$

where ρ_j is equal to $V_m(r_{m+j})/\{V_m(r_{m+j}) * [V_m(r_{m+1})+V_m(r_{m+2})]\}^{1/2}$, $j=1,2$. The assumption of bivariate normal distribution is slightly stronger than normality of the errors, so we have a quasi-likelihood even when the distribution of ϵ_t is Gaussian due to this approximation.
In general, suppose that after the m-th observation there are k missing values, i.e., we observe $r_1, ..., r_m, r_{m+1}+ ...+r_{m+k+1}, r_{m+k+2}, ..., r_n$. In order to evaluate the quasi-log-likelihood, we need to calculate $V(\sum_{z=1}^{k+1} r_{m+z}|\mathcal{F}_m)$ and $V(r_{m+k+2}|\mathcal{F}_{m+k+1})$. Because conditional to $\mathcal{F}_m$, r_{m+j}, $j=1,\ldots,k+1$, are uncorrelated, the conditional variance of the sum is equal to the sum of the conditional variances given by Equations 5 and 6. The second conditional variance is given by

$$\begin{aligned} V(r_{m+k+2}|\mathcal{F}_{m+k+1}) \;=\; & \sum_{j=0}^{k}\beta_1^j\Big\{\alpha_0+\alpha_1\Big[V_m(r_{m+k-j+1}|\sum_{z=1}^{k+1} r_{m+z}) \\ & + E_m^2(r_{m+k-j+1}|\sum_{z=1}^{k+1} r_{m+z})\Big]\Big\} + \beta_1^k V_m(r_{m+1}) \end{aligned} \tag{10}$$

where $V_m(r_{m+1})$ is given by Equation 5, $V_m(r_{m+k-j+1}|\sum_{z=1}^{k+1} r_{m+z}) = V_m(r_{m+k-j+1})(1-\rho_{k-j+1}^2)$ and $V_m(r_{m+k-j+1})$ is given by Equation 6, $j=0,1,\ldots,k$. Furthermore,

$$E_m(r_{m+k-j+1}|\sum_{z=1}^{k+1} r_{m+z}) = \rho_{k-j+1}^2\sum_{z=1}^{k+1} r_{m+z} \;\text{ and }\; \rho_{k-j+1} = \frac{V_m(r_{m+k-j+1})}{[V_m(r_{m+k-j+1})\sum_{z=1}^{k+1} V_m(r_{m+z})]^{1/2}}, \tag{11}$$

with $j=0,1,\ldots,k$.
The approximate values of $V(\sum_{z=1}^{k+1} r_{m+z}|\mathcal{F}_m)$ and $V(r_{m+k+2}|\mathcal{F}_{m+k+1})$ are important on their own. The first one gives the conditional volatility of the sum of k steps ahead returns. In fact, although the k-step ahead of the conditional volatility of the returns is generally given, in practice one could be interested in the sum of returns.
In this paper, the GARCH model with standardized Student-t distribution is also used. Now, the process $\{\epsilon_t\}$ assumes the referred distribution, and the log-likelihood is given by

$$\ell_t(\boldsymbol{\theta}) \;=\; (n-r)\log\frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{(\nu-2)\pi}\;\Gamma(\frac{\nu}{2})} - \frac{\nu+1}{2}\sum_{t=r}^{n}\log\left[1+\frac{y_t^2}{\sigma_t^2(\nu-2)}\right] - \sum_{t=r}^{n}\log\sigma_t, \tag{12}$$

where ν is the degree of freedom of a standardized Student-t distribution, with domain given by $\nu>2$, in order to have finite second moment, and $\boldsymbol{\theta}$ is the parameter vector now defined by $\boldsymbol{\theta}=(\boldsymbol{\alpha},\boldsymbol{\beta},\nu)$.
However, there are some issues related to the use of this distribution. First, because it is not easy to evaluate the distribution of the sum of two i.i.d. Student-t random variables, we consider in the quasi-log-likelihood that it is also a Student-t with the same degree of freedom. Second, we assume normality to evaluate the conditional variances $V(\sum_{z=1}^{k+1} r_{m+z}|\mathcal{F}_m)$ and $V(r_{m+k+2}|\mathcal{F}_{m+k+1})$ as in the Gaussian case. Thus, even if the innovations have Student-t distribution, we have a quasi-Student-t log-likelihood (QSLL). Fiorentini et al. (2003) propose to estimate the reciprocal of the degrees of freedom, that is, $1/\nu$ instead ν in order to avoid substantial distortions in the estimation.

3. Monte Carlo simulation

The simulation process is performed in the following way: first, we generate a series of size n which follows a GARCH(1,1) process. Afterwards, we have to include the missing observations to simulate an incomplete data set. In this paper, we consider missing values in the prices, not in the returns. For the case of isolated missing observation, the i-th price missing value occurs at the $25i$-th observation, where i is the index of the missing values, with $i = 1, \ldots, n/25 - 1$. Then, the returns observed around the i-th missing value are $\{\ldots, r_{m_i}, r_{m_i+1} + r_{m_i+2}, r_{m_i+3}, \ldots\}$, where $m_i = 25i - 1$. In the case of two consecutive missing values, the prices in positions $25i$ and $25i + 1$ are removed and we observe the returns $\{\ldots, r_{m_i}, r_{m_i+1} + r_{m_i+2} + r_{m_i+3}, r_{m_i+4}, \ldots\}$.

We compare the method proposed in this paper with the estimation with the complete dataset using the method of maximum likelihood (ML), and the method where the missing observations are removed and the data is bound together (BT), which means that the length of the series is smaller than the complete data set. The true parameter values are $\alpha_0 = 0.04$, $\alpha_1 = 0.1$ and $\beta_1 = 0.86$, and the data is generated with Gaussian distribution. Table 1 reports the bias as well as the MSE values of the estimates, with sample size $n = 1050, 2050$. The benchmark model is the ML.

Table 1: Bias and MSE of the estimates of series with 1 and 2 missing values with Gaussian distributions for the innovations, by QGLL and BT methods and with the complete data via maximum likelihood (ML). The values in parentheses are the increase or decrease (-) in percentage in relation to the estimation with complete data. The sample size is n. Bias and MSE are pre-multiplied by 10^{-1} and 10^{-2}, respectively.

		Bias	MSE	Bias	MSE
		1 missing value		2 missing values	
		$n = 1050$			
	ML	0.0726	0.0450	0.0726	0.0450
α_0	QGLL	0.0754 (3.8)	0.0502 (11)	0.0740 (1.9)	0.0552 (23)
	BT	0.1224 (68)	0.0691 (53)	0.2037 (180)	0.1538 (241)
	ML	-0.0194	0.0541	-0.0194	0.0541
α_1	QGLL	-0.0072 (-)	0.0609 (12)	-0.0127 (-)	0.0648 (20)
	BT	-0.0297 (53)	0.0619 (14)	-0.0568 (192)	0.0817 (51)
	ML	-0.0693	0.1328	-0.0693	0.1328
β_1	QGLL	-0.0844 (22)	0.1502 (13)	-0.1236 (78)	0.1814 (36)
	BT	-0.0883 (27)	0.1633 (23)	-0.1148 (65)	0.2669 (101)
		$n = 2050$			
	ML	0.0359	0.0171	0.0359	0.0171
α_0	QGLL	0.0381 (6.1)	0.0195 (14)	0.0321 (-)	0.0189 (10)
	BT	0.0819 (128)	0.0283 (65)	0.1467 (308)	0.0593 (246)
	ML	-0.0054	0.0265	-0.0054	0.0265
α_1	QGLL	0.0057 (5.5)	0.0299 (13)	0.0018 (-)	0.0309 (17)
	BT	-0.0172 (218)	0.0305 (15)	-0.0465 (761)	0.0420 (58)
	ML	-0.0361	0.0597	-0.0361	0.0597
β_1	QGLL	-0.0499 (38)	0.0696 (16)	-0.0855 (136)	0.0776 (30)
	BT	-0.0503 (39)	0.0733 (23)	-0.0590 (63)	0.1051 (76)

The simulation shows that for the isolated missing value for all sample sizes, the bias of the proposed method is always smaller than of the BT method. In fact, in the estimation of α_1, the bias is even smaller than with the ML method, for $n = 1050$. The relative increase of the bias in the estimation of β_1 is similar between the competing methods, but QGLL has a slight advantage.

For consecutive missing values, the QGLL estimates of α_0 and α_1 keep the same advantage of isolated missing value. On the other hand, in the estimation of β_1, the QGLL method has larger bias with sample sizes 1050 and 2050. It is important to see that although there is a larger bias in the estimation of β_1, the MSE is smaller for all sample sizes. This means that although the QGLL is underestimating the true parameter, in general the estimates are closer to the true value of β_1 than with the BT method.

When the sample size is increased, the MSE has a slight variation although they are very close to each other. Furthermore, the increase of the MSE for the QGLL estimates with respect to ML varies from 11 to 16%, while in BT it is about 12 to 65%. In addition, for consecutive missing values the MSE values in the estimation of α_0 and β_1 with the QGLL method have significant changes downward, while for α_1 they are almost constant. Except for β_1, the increase in the MSE of our method is at least two times smaller than in the BT method.

We repeated the process for a different setting of parameters and also for the Student-t distribution. We also used a sample size $n = 550$ for all the cases. In general, the proposed method produced smaller increase in bias and MSE with respect to ML than the BT method. The volatilities $V_{t_0}(r_{t_0+1} + \ldots + r_{t_0+k+1})$ and $V_{t_{0+k+1}}(r_{t_0+k+2})$, $k = 1, 2$, were also estimated, but they are exhibited only in the empirical application.

4. Application to NASDAQ 100 composite stock prices

In order to illustrate the performance of our method, here we apply it to a real dataset, namely the NASDAQ 100 Index. The data consist of approximately 8 years, for a total of 2050 observations, from 10/17/2006 to 12/09/2014.

The experiment with real data is carried out just as with the simulated data, in which one and two observations are removed in a space of 25 observations. We assume that the pseudo true parameter values are the estimates obtained with the complete data. The performance is measured by means of estimating parameters and volatilities. To be able to measure the accuracy of the volatility estimation, we must establish a proxy value as reference for the true volatility. In our case, we consider the volatility filtered by the complete data. The local correlation is removed by an AR(1) filter, and the residuals are modeled by a GARCH(1,1) model. To avoid numerical problems, the residuals are divided by their standard errors. Table 2 presents the estimated volatilities $V_{m_i}(r_{m_i+1} + \ldots + r_{m_i+k+1})$ and $V_{m_i+k+1}(r_{m_i+k+2})$, $k = 1, 2$ missing values, as well as the parameter estimates for the QGLL and BT methods. As can be seen, our method produces smaller bias

Table 2: Filtered volatility and parameter estimation of Nasdaq 100. The reference measure of Bias and MSE are the filtered volatilities with the complete data.

		Bias		MSE	Bias		MSE
		$k=1$ missing value			$k=2$ missing values		
$V_{m_i}(r_{m_i+1} + \ldots + r_{m_i+k+1})$	QGLL	-0.001		0.006	0.087		0.114
	BT	-0.987		3.126	-1.763		9.830
$V_{m_i+k+1}(r_{m_i+k+2})$	QGLL	-0.005		0.113	0.023		0.056
	BT	-0.036		0.209	0.162		0.124
		$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$
Parameter estimation	No MV	0.013	0.096	0.890	0.013	0.094	0.891
	QGLL	0.013	0.100	0.886	0.013	0.103	0.878
	BT	0.014	0.101	0.883	0.019	0.109	0.871

and MSE in the estimation of the volatilities after omitting the missing values. The better estimation can be seen in Figure 1, where we plot the comparison between the estimated volatility with the complete data and the estimated volatility with the proposed method and BT method with one and two missing values. We only plot the estimated volatilities $\hat{V}_{m_i}(r_{m_i+1} + \ldots + r_{m_i+k+1})$ and $\hat{V}_{m_i+k+1}(r_{m_i+k+2})$, $k = 1, 2$, where $m_i + 1$ is the time of the first missing value of the i-th sequence of missing values in the series. We have isolated missing values when $k = 1$ and two consecutive missing values when $k = 2$. Figure 1 indicates that our method provides much better estimates of the volatilities next to the missing values, and the estimated volatilities $\hat{V}_{m_i}(r_{m_i+1} + \ldots + r_{m_i+k+1})$, $k = 1, 2$, of the QGLL method almost coincide with those of the complete data. The competing method underestimates this volatility. At time $\{m_i + k + 2\}$, both methods perform similarly, with an advantage of QGLL.

While our method's estimator of $V_{m_i}(r_{m_i+1} + \ldots + r_{m_i+k+1})$ can be considered (almost) unbiased, the same does not happen with the BT method, which presents a high negative bias for isolated missing values and even larger for consecutive missing values. Also, the MSE shows that our estimated volatilities are very close to the reference ones in the two cases, and the BT method produces a huge MSE. However, for the estimated volatility at time $\{m_i + k + 2\}$, $k = 1, 2$, both methods perform well, although QGLL produces MSE values around two times smaller than the BT method, and also produces lower bias, which is around seven times smaller.

In the sense of parameter estimation, our method is more accurate in the presence of missing values, and $\hat{\alpha}_0$ has the same value compared with the complete data. $\hat{\alpha}_1$ and $\hat{\beta}_1$ are not exactly the same, but they are closer than in the BT method. For one and two missing values, the behavior of the estimates is similar.

5. Conclusions

In this paper we develop a method to deal with missing values without compromising the correlation structure of financial time series. Through some approximations, we were able to evaluate conditional expectations and variances, in order to transfer the information contained in the series before the missing values to the observations after them. We also showed through an extensive Monte Carlo experiment and an application

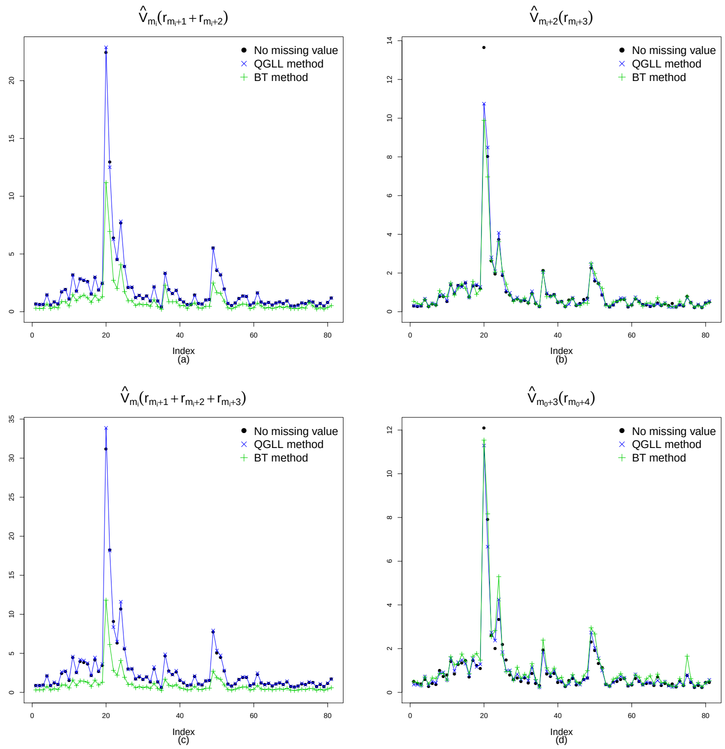

Figure 1: Plots **(a)** and **(c)** present the estimated volatilities $\hat{V}_{m_i}(r_{m_i+1} + \ldots + r_{m_i+k+1})$, $k = 1, 2$, respectively. Plots **(b)** and **(d)**, on the other hand, present the volatilities $\hat{V}_{m_i+k+1}(r_{m_i+k+2})$, $k = 1, 2$, respectively. $m_i + 1$ is the time of the first missing value in the i-th sequence of missing values in the series.

to real data that our method is better for parameter and volatility estimation, with results either very close to the complete data estimation or with an increase in bias or MSE, but still better than the usual method of estimation, such as binding the data together. Due to the flexibility of the quasi-likelihood method, it is possible to extend the method to a GARCH(p,q) model, but since the simple GARCH is sufficiently accurate in the most common situations, it is left for future works. Finally, as a byproduct, we estimate the variance of the prediction of the sum of the next k returns after the missing values.

References

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307–327.

Brockwell, P. J. & Davis, R. A. (2009). Time Series: Theory and Methods. Springer Series in Statistics. Springer, New York.

Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B, 39, 1–38.

Ossandón, S. & Bahamonde, N. (2012). A new nonlinear formulation for GARCH models. Comptes Rendus Mathematique, 351(5–6), 235–239.

Fiorentini, G., Sentana, E. & Calzolari, G. (2003). Maximum likelihood estimation and inference in multivariate conditionally heteroscedastic dynamic regression models with Student-t innovations. Journal of Business & Economic Statistics, 21(4), 532–546.