KYBERNETIKA — VOLUME 54 (2018), NUMBER 6, PAGES 1138–1155

ROBUST RECURSIVE ESTIMATION OF GARCH MODELS

Tomáš Cipra and Radek Hendrych

The robust recursive algorithm for the parameter estimation and the volatility prediction in GARCH models is suggested. It seems to be useful for various financial time series, in particular for (high-frequency) log returns contaminated by additive outliers. The proposed procedure can be effective in the risk control and regulation when the prediction of volatility is the main concern since it is capable to distinguish and correct outlaid bursts of volatility. This conclusion is demonstrated by simulations and real data examples presented in the paper.

Keywords: GARCH model, Kalman filter, outlier, robust recursive estimation, volatility

Classification: 62M10, 62F35, 91G70

1. INTRODUCTION

Financial time series (in particular returns of financial assets) typically exhibit significant kurtosis and volatility clustering. The assets are usually stocks or stock indices (see e. g. [52]) or currencies (see e. g. [45]). The GARCH models introduced by [4] and [19] are commonly applied in order to model these typical properties with the aim to describe dynamics of conditional variances and forecast financial volatility. However, when fitted to real time series the residuals of the estimated models have frequently excess kurtosis explainable by the presence of outliers which are not captured by the GARCH models, see e. g. [8, 9, 10, 23] (on the other hand, some authors argue that extreme observations are not outliers and they should be incorporated into the model, see e. g. [2] or [20]).

The parameters of the GARCH models are routinely estimated by the (conditional) maximum likelihood but they are rarely calibrated recursively. Nevertheless, recursive estimates performed using recursive algorithms are undoubtedly advantageous. To evaluate the parameter estimates at a time step, recursive estimation methods operate only with the current measurements and parameters estimated in previous steps (see e. g. [1, 18, 26, 32, 37]). It is in sharp contrast to the non-recursive estimation where all data are collected at first and then the model is fitted. Therefore, recursive estimation techniques are effective in terms of memory storage and computational complexity. This efficiency can be employed just in the framework of (high-frequency) financial time series data. Alternatively, it is possible to adopt these methods to monitor or forecast volatility on-line, to evaluate risk measures (e. g. Value at Risk or Expected Shortfall), to detect

DOI: 10.14736/kyb-2018-6-1138

faults, to check model stability including detection of structural changes, etc. Moreover due to the previous arguments, the recursive GARCH estimation should be resistant (robust) to outliers. The primary goal of this paper is to suggest a robust recursive algorithm which is effective enough in the context of GARCH models to estimate and forecast volatility of contaminated (high-frequency) financial data.

Various methods of non recursive estimation of GARCH parameters and volatility in presence of outliers consist either in (*i*) identifying and correcting additive outliers (AO) or innovative outliers (IO) in (residual) time series (see e. g. [9, 10, 23, 24, 27, 28, 34]), or in (*ii*) robustifying classical statistical estimators of the type LS or ML to the form of M estimators and similar robust versions (see e. g. [7, 33, 44, 49, 55]), or in (*iii*) applying estimators with robust properties of the type LAD or median MAD (see e. g. [3, 36, 39, 45, 46, 58]).

As robust recursive estimation of GARCH model is concerned, initially one should remind a close connection to robustification of Kalman filter which is desirable including various engineering applications in the context of state space modelling with outliers (see e. g. [6, 12, 14, 21, 29, 30, 35, 38, 47, 48, 50, 53, 54]). Moreover, a special case of Kalman filter robustification is the robust exponential smoothing including Holt-Winters method (see e. g. [11, 13, 15, 16, 17, 25, 31, 43]).

The following sections of the paper deal in sequence with (1) the presentation of self weighted recursive estimation algorithm for GARCH models, (2) its robustification, (3) the simulation study for various types of outliers, and finally (4) real data applications.

2. GARCH MODELS: CONSTRUCTION OF RECURSIVE ESTIMATOR

The GARCH(p, q) process $\{y_t\}_{t\in\mathbb{Z}}$ in financial applications is commonly defined as (refer to e. g. [22])

$$y_t = \sigma_t \varepsilon_t,\ \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \tag{1}$$

where $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is a sequence of *i.i.d.* random variables with zero mean and unit variance, and ω, $\alpha_1, \dots, \alpha_p$, $\beta_1, \dots, \beta_q$ are the parameters of the process. The first two conditional and unconditional moments can be simply calculated as

$$\mathbb{E}(y_t|\mathcal{F}_{t-1}) = 0,\ \mathbb{E}(y_t) = 0,\ \mathsf{var}(y_t|\mathcal{F}_{t-1}) = \sigma_t^2,\ \mathsf{var}(y_t) = \mathbb{E}(\sigma_t^2), \tag{2}$$

where $\mathcal{F}_t$ denotes the smallest σ-algebra with respect to which y_s is measurable for all $s \le t$. Sufficient conditions for σ_t^2 being positive are $\omega > 0$, $\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q \ge 0$. If $\beta_1 = \cdots = \beta_q = 0$, the model is reduced to the ARCH(p) process. Additionally, sufficient conditions for y_t being (weakly) stationary are $\omega > 0$, $\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q \ge 0$, and $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$. The stationary GARCH$(p, q)$ model has the finite variance:

$$\mathsf{var}(y_t) = \frac{\omega}{1 - \sum_{i=1}^{p} \alpha_i - \sum_{j=1}^{q} \beta_j}. \tag{3}$$

The one-step ahead prediction of σ_t^2 is expressed as

$$\hat{\sigma}^2_{t+1|t} = \omega + \sum_{i=1}^{p} \alpha_i y^2_{t+1-i} + \sum_{j=1}^{q} \beta_j \sigma^2_{t+1-j}. \tag{4}$$

The GARCH models are routinely estimated by the non-recursive conditional maximum likelihood method with normal distribution being usually preferred since the corresponding estimates stay consistent. Aknouche and Guerbyenne [1] proposed a couple of two-stage recursive estimation schemes appropriate for the standard GARCH(p, q) models extending the ideas presented in [5]. However, they focused mainly on the derivation and convergence analysis of the algorithm and not on its numerical evaluation, which might be regarded as a (crucial) objection (consult [42]). In particular, the whole computational implementation is based on recursive pseudo-linear regression estimation scheme applied to the following representation of the GARCH(p, q) process y_t (see e. g. [22]):

$$y_t^2 = \sigma_t^2(\boldsymbol{\theta}_0) + \nu_t, \tag{5}$$

where $\sigma_t^2(\boldsymbol{\theta}_0)$ is defined in (1) for the true values of model parameters collected in the vector $\boldsymbol{\theta}_0$ and $\nu_t = \sigma_t^2(\boldsymbol{\theta}_0)(\varepsilon_t^2 - 1)$ is white noise with $\mathbb{E}(\nu_t) = 0$ and $\mathsf{var}(\nu_t) = \mathbb{E}(\nu_t^2)$. Independently, such a recursive pseudo-linear regression algorithm has been also proposed in [37] following analogical derivation schemes (see [26]).

Hendrych and Cipra [32] derived an alternative recursive formulas for estimating parameters of the standard GARCH(p, q) model which is based on the principle of self-weighted estimation (see also [40] and [57], but Hendrych and Cipra [32] made use of the opportunity to formulate this principle conveniently in a recursive way). Even though the Gaussian QMLE approach was applied for this purpose, theoretically it could be generalized to other types of QMLE (see e. g. [56] for an overview), but with much higher computational complexity.

The recursive identification instruments introduced by [41, 42], and [51] were applied to deliver one-stage recursive estimation procedures (in contrast to two-stage procedure suggested in [1]):

$$\widehat{\boldsymbol{\theta}}_t = \widehat{\boldsymbol{\theta}}_{t-1} + \frac{\widehat{\boldsymbol{P}}_{t-1}\widehat{\boldsymbol{\psi}}_t(y_t^2 - \widehat{\boldsymbol{\varphi}}_t^\top \widehat{\boldsymbol{\theta}}_{t-1})}{\lambda_t(\widehat{\boldsymbol{\varphi}}_t^\top \widehat{\boldsymbol{\theta}}_{t-1})^2 + \widehat{\boldsymbol{\psi}}_t^\top \widehat{\boldsymbol{P}}_{t-1}\widehat{\boldsymbol{\psi}}_t}, \tag{6a}$$

$$\widehat{\boldsymbol{P}}_t = \frac{1}{\lambda_t}\left\{\widehat{\boldsymbol{P}}_{t-1} - \frac{\widehat{\boldsymbol{P}}_{t-1}\widehat{\boldsymbol{\psi}}_t\widehat{\boldsymbol{\psi}}_t^\top \widehat{\boldsymbol{P}}_{t-1}}{\lambda_t(\widehat{\boldsymbol{\varphi}}_t^\top \widehat{\boldsymbol{\theta}}_{t-1})^2 + \widehat{\boldsymbol{\psi}}_t^\top \widehat{\boldsymbol{P}}_{t-1}\widehat{\boldsymbol{\psi}}_t}\right\}, \tag{6b}$$

$$\widehat{\boldsymbol{\varphi}}_{t+1} = (1, y_t^2, \ldots, y_{t+1-p}^2, \widehat{\boldsymbol{\varphi}}_t^\top \widehat{\boldsymbol{\theta}}_t, \ldots, \widehat{\boldsymbol{\varphi}}_{t+1-q}^\top \widehat{\boldsymbol{\theta}}_{t+1-q})^\top, \tag{6c}$$

$$\widehat{\boldsymbol{\psi}}_{t+1} = \widehat{\boldsymbol{\varphi}}_{t+1} + \sum_{j=1}^{q} \widehat{\beta}_{j,t}\widehat{\boldsymbol{\psi}}_{t+1-j}, \tag{6d}$$

$$\lambda_t = \tilde{\lambda}\cdot\lambda_{t-1} + (1-\tilde{\lambda}), \quad \lambda_0,\ \tilde{\lambda} \in (0,1),\ t \in \mathbb{N}, \tag{6e}$$

where the recursive estimates are collected in $\widehat{\boldsymbol{\theta}}_t$ (the parameters in (1) are ordered to a single vector $\boldsymbol{\theta} = (\omega, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q)^\top$). The *forgetting factor* $\{\lambda_t\}_{t\in\mathbb{N}}$ is a deterministic sequence of positive real numbers less or equal to one. It represents the observation weight over time. One commonly puts $\lambda_0 = 0.95$ and $\tilde{\lambda} = 0.99$. The initialization of the algorithm is thoroughly discussed in [32] (see p. 320). Finally, one should introduce the simple projection, which completes the algorithm (6) and ensures that it will not degenerate:

$$\left[\widehat{\boldsymbol{\theta}}_t\right]_{\boldsymbol{D}_{\mathcal{M}}} = \begin{cases} \widehat{\boldsymbol{\theta}}_t & \text{if } \widehat{\boldsymbol{\theta}}_t \in \boldsymbol{D}_{\mathcal{M}}, \\ \widehat{\boldsymbol{\theta}}_{t-1} & \text{if } \widehat{\boldsymbol{\theta}}_t \notin \boldsymbol{D}_{\mathcal{M}}, \end{cases} \tag{7}$$

where $\boldsymbol{D}_{\mathcal{M}} := \{\boldsymbol{\theta} \in \mathbb{R}^{p+q+1} | \tilde{\delta}_1 \leq \theta_1 \leq \tilde{\Delta}_1; \theta_i \geq 0, i = 2, \ldots, p+q+1; \sum_{j=2}^{p+q+1} \theta_j \leq 1 - \tilde{\delta}_2\}$ and one usually puts $0 < \tilde{\delta}_1 \leq \tilde{\Delta}_1 < \infty$, $0 < \tilde{\delta}_2 < 1$, e. g. $\tilde{\delta}_1 = 10^{-9}$ and $\tilde{\Delta}_1 = 10^2$.

The performance of this algorithm was (fairly) compared by means of the various Monte Carlo experiments with other methods mentioned above. The one-stage recursive estimate has proven to be at least competitive amongst the others (and usually better, see [32]). The simulations have also shown that the initialization of the algorithms must be handled carefully, since it can significantly influence the speed of its convergence. For more details, consult [32] and the references therein.

Remark that the introduced recursive estimation technique (6) might be an appealing alternative to the moving-window estimation approach which can be also applied to detect outliers and structural changes. The latter approach is based on repeating minimization of the negative conditional log-likelihood function (after excluding the constant term), which corresponds to the model (1) when assuming normally distributed innovations ε_t. It can be expressed as follows:

$$\widehat{\boldsymbol{\theta}}_t^M = \arg \min_{\boldsymbol{\theta} \in \boldsymbol{D}_{\mathcal{M}}} \sum_{\tau=t-M+1}^{t} \left[\frac{y_\tau^2}{\sigma_\tau^2} + \log(\sigma_\tau^2) \right], \ t \geq M, \tag{8}$$

where $M \in \mathbb{N}$ denotes the moving-window width. At each time the minimum (8) is repeatedly calculated and $\widehat{\sigma}_{t+1}^2$ is evaluated using the most recent estimate $\widehat{\boldsymbol{\theta}}_t^M$ and M consecutive observations. The estimation can be initialized similarly as above. It is obvious that the estimation can start only after observing at least M financial returns. This scheme or its alternatives are frequently applied in practice. However, such an estimator is computationally very complex since the optimization task is obviously solved by an iterative procedure based on M consecutive measurements at each time. On the contrary, recursive estimation algorithms of the type (6) are computationally much more effective (as we stated in Section 1). Therefore, they are primarily preferred in real time applications (especially in the context of high-frequency data).

3. ROBUST RECURSIVE ESTIMATION OF GARCH MODELS

Using GARCH models, it is necessary to be concerned about outliers that may occur in data (see also the first introduction section and the last section on real data applications). Outliers can be caused by many reasons, e. g. by additive innovations, measurement failures, operational risk problems, management decisions, etc. They can influence the estimation and prediction in the applied model considerably if no specific action is taken.

Therefore, if such defects are expected in the data set, one should modify the estimation algorithms to make them more robust. The outliers tend to appear as spikes in the sequence of $\{y_t/\sqrt{\sigma_t^2}\}$, which obviously result in large contributions to the loss function. There exist various ways how to robustify recursive estimation algorithms (refer to the first introduction section above). In this contribution, a simple way of handling outliers is applied based on testing a measurement at each time t. If it is large compared with a given limit, it is indicated as erroneous and substituted immediately by another value (see e. g. [13, 42, 47], and others). According to simulations, this strategy seems to be efficient for additive outliers (AD) mainly.

Under the previous arguments, the algorithm (6) can be robustified to the following form:

$$\widehat{\boldsymbol{\theta}}_t^{rob} = \widehat{\boldsymbol{\theta}}_{t-1}^{rob} + \frac{\widehat{\boldsymbol{P}}_{t-1}^{rob}\widehat{\boldsymbol{\psi}}_t^{rob}\left[\left(\hat{y}_t^{rob}\right)^2 - (\widehat{\boldsymbol{\varphi}}_t^{rob})^\top\widehat{\boldsymbol{\theta}}_{t-1}^{rob}\right]}{\lambda_t\left[(\widehat{\boldsymbol{\varphi}}_t^{rob})^\top\widehat{\boldsymbol{\theta}}_{t-1}^{rob}\right]^2 + (\widehat{\boldsymbol{\psi}}_t^{rob})^\top\widehat{\boldsymbol{P}}_{t-1}^{rob}\widehat{\boldsymbol{\psi}}_t^{rob}}, \tag{9a}$$

$$\widehat{\boldsymbol{P}}_t^{rob} = \frac{1}{\lambda_t}\left\{\widehat{\boldsymbol{P}}_{t-1}^{rob} - \frac{\widehat{\boldsymbol{P}}_{t-1}^{rob}\widehat{\boldsymbol{\psi}}_t^{rob}(\widehat{\boldsymbol{\psi}}_t^{rob})^\top\widehat{\boldsymbol{P}}_{t-1}^{rob}}{\lambda_t\left[(\widehat{\boldsymbol{\varphi}}_t^{rob})^\top\widehat{\boldsymbol{\theta}}_{t-1}^{rob}\right]^2 + (\widehat{\boldsymbol{\psi}}_t^{rob})^\top\widehat{\boldsymbol{P}}_{t-1}^{rob}\widehat{\boldsymbol{\psi}}_t^{rob}}\right\}, \tag{9b}$$

$$\widehat{\boldsymbol{\varphi}}_{t+1}^{rob} = \left(1, \left(\hat{y}_t^{rob}\right)^2, \ldots, \left(\hat{y}_{t+1-p}^{rob}\right)^2, (\widehat{\boldsymbol{\varphi}}_t^{rob})^\top\widehat{\boldsymbol{\theta}}_t^{rob}, \ldots, (\widehat{\boldsymbol{\varphi}}_{t+1-q}^{rob})^\top\widehat{\boldsymbol{\theta}}_{t+1-q}^{rob}\right)^\top, \tag{9c}$$

$$\widehat{\boldsymbol{\psi}}_{t+1}^{rob} = \widehat{\boldsymbol{\varphi}}_{t+1}^{rob} + \sum_{j=1}^{q}\widehat{\beta}_{j,t}^{rob}\widehat{\boldsymbol{\psi}}_{t+1-j}^{rob}, \tag{9d}$$

$$\lambda_t = \tilde{\lambda}\cdot\lambda_{t-1} + (1-\tilde{\lambda}), \quad \lambda_0,\ \tilde{\lambda}\in(0,1),\ t\in\mathbb{N}, \tag{9e}$$

where the recursive estimates are collected in $\widehat{\boldsymbol{\theta}}_t^{rob}$ (the $\{\lambda_t\}_{t\in\mathbb{N}}$ is the same as in (6e)). To complete (9), one defines the outlier-corrected series $\{\hat{y}_t^{rob}\}$ as follows:

$$\left(\hat{y}_t^{rob}\right)^2 = \begin{cases} (\widehat{\boldsymbol{\varphi}}_t^{rob})^\top\widehat{\boldsymbol{\theta}}_{t-1}^{rob} \\ \quad + \operatorname{sign}\left(y_t^2 - (\widehat{\boldsymbol{\varphi}}_t^{rob})^\top\widehat{\boldsymbol{\theta}}_{t-1}^{rob}\right)(u_{1-\alpha/2})^2\sqrt{\left[(\widehat{\boldsymbol{\varphi}}_t^{rob})^\top\widehat{\boldsymbol{\theta}}_{t-1}^{rob}\right]^2 + (\widehat{\boldsymbol{\psi}}_t^{rob})^\top\widehat{\boldsymbol{P}}_{t-1}^{rob}\widehat{\boldsymbol{\psi}}_t^{rob}/\lambda_t} \\ \qquad\quad \text{for } \left|y_t^2 - (\widehat{\boldsymbol{\varphi}}_t^{rob})^\top\widehat{\boldsymbol{\theta}}_{t-1}^{rob}\right| > (u_{1-\alpha/2})^2\sqrt{\left[(\widehat{\boldsymbol{\varphi}}_t^{rob})^\top\widehat{\boldsymbol{\theta}}_{t-1}^{rob}\right]^2 + (\widehat{\boldsymbol{\psi}}_t^{rob})^\top\widehat{\boldsymbol{P}}_{t-1}^{rob}\widehat{\boldsymbol{\psi}}_t^{rob}/\lambda_t}, \\ y_t^2 \qquad \text{otherwise.} \end{cases} \tag{9f}$$

Note that $\hat{y}_t^{rob} = \operatorname{sign}(y_t)\sqrt{(\hat{y}_t^{rob})^2}$ and that $u_{1-\alpha/2}$ denotes the corresponding quantile of the standard normal distribution, where one usually puts $\alpha = 0.05$. The initialization settings and projection rule (7) remain similar as in the previous section. The algorithm is based on the robustified version of Kalman filter derived in [13] (see Appendix) which is applied to the GARCH models written in the form (5). The assumption of normality can be replaced by other distributions.

Parallelly one can construct the robust recursive prediction of volatility, e. g. the one-step ahead prediction has the form (compare with (4)):

$$(\hat{\sigma}_{t+1|t}^{rob})^2 = \hat{\omega}_t^{rob} + \sum_{i=1}^{p}\hat{\alpha}_{it}^{rob}(\hat{y}_{t+1-i}^{rob})^2 + \sum_{j=1}^{q}\hat{\beta}_{jt}(\hat{\sigma}_{t+1-j}^{rob})^2. \tag{10}$$

A theoretical analysis of the introduced algorithm can follow the general schemes considered by [42]; it is rather technical and uses instruments known mainly in the ordinary differential equation theory. Under the corresponding (mostly technical) general assumptions, it can be shown that the estimated parameters converge to their true counterparts. Additionally, they are asymptotically normally distributed (consult [42]).

4. SIMULATIONS

The suggested procedure (9) has been studied by means of simulations $\{y_t^*\}$ using seven simulation scenarios described in Table 4:

$$y_t^* = y_t + \delta_t,\ t = 1, \ldots, T,\ T = 20000, \tag{11a}$$

$$y_t = \sigma_t \varepsilon_t,\ \sigma_t^2 = 0.0001 + 0.05 y_{t-1}^2 + 0.94 \sigma_{t-1}^2,\ \varepsilon_t \sim i.i.d.\ \mathsf{N}(0,1), \tag{11b}$$

where the term δ_t represents the additive outlier. According to Table 4, this simulation study covers various ways of contamination of the generic processes in the first scenario (Model 0) without outliers (e. g., the t-distribution with the degree of freedom one denoted as $t(1)$ represents the contamination by outliers with heavy tails). For each scenarios one has simulated 1000 realizations of length 20000 and applied the robust recursive estimation procedure (9) with the same initialization settings and projection rule (7) (in particular, $\alpha = 0.05$, $\tilde{\delta}_1 = \tilde{\delta}_2 = 10^{-9}$, and $\tilde{\Delta}_1 = 10^2$). The simulation results for innovative outliers (IO) are not presented in this paper.

Figures 1 − 7 present the corresponding probability densities of estimated parameters $\omega = 0.0001$, $\alpha_1 = 0.05$, $\beta_1 = 0.94$ in time $t = 20000$ calculated over 1000 realizations (see also the median absolute deviations MAD in Table 2). Moreover, Figure 8 shows the time records of medians of estimated parameters ω, α_1, β_1 in Model 1 (i. e., the single additive outlier of size 10 in time $t = 10000$) before and after the robustification. Figure 9 presents the corresponding boxplots in times $t = 5000$, 10000, 20000 for this model.

Similar results have been obtained for other configurations of true parameters (the presented case corresponds to the extreme situation close to the border of model stability). In general, the simulations show that the robustification improves the behaviour of suggested recursive procedure in a substantial way.

Model 0	$\delta_t = 0$ for all t
Model 1	$\delta_t = 10 \times I_t,\ I_t = \begin{cases} 1, & t = 10000 \\ 0, & \text{otherwise} \end{cases}$
Model 2	$\delta_t = 10 \times I_t,\ I_t \sim i.i.d.\ Alt\left(\frac{1}{20000}\right)$ for all t
Model 3	$\delta_t = 10 \times I_t,\ I_t \sim i.i.d.\ Alt\left(\frac{4}{20000}\right)$ for all t
Model 4	$\delta_t = s_t \times I_t,\ I_t \sim i.i.d.\ Alt\left(\frac{4}{20000}\right),\ s_t \sim i.i.d.\ t(1)$ for all t
Model 5	$\delta_t = s_t \times I_t,\ I_t \sim i.i.d.\ Alt\left(\frac{20}{20000}\right),\ s_t \sim i.i.d.\ t(1)$ for all t
Model 6	$\delta_t = s_t \times I_t,\ I_t \sim i.i.d.\ Alt\left(\frac{200}{20000}\right),\ s_t \sim i.i.d.\ t(1)$ for all t

Tab. 1. Various simulation scenarios for additive outliers in (11)
[$Alt \sim$ the alternative distribution, $t(1) \sim$ the Student t_1-distribution].

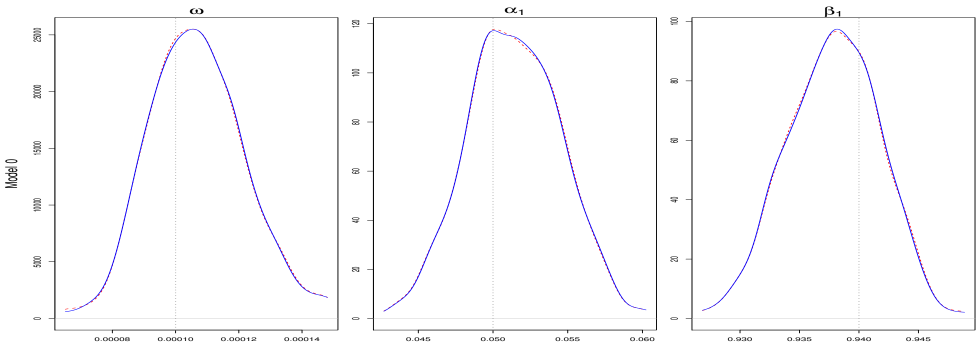

Fig. 1. Probability densities of estimated parameters: **Model 0**
[non-robustified estimates ∼ dashed lines, robustified estimates ∼ solid lines].

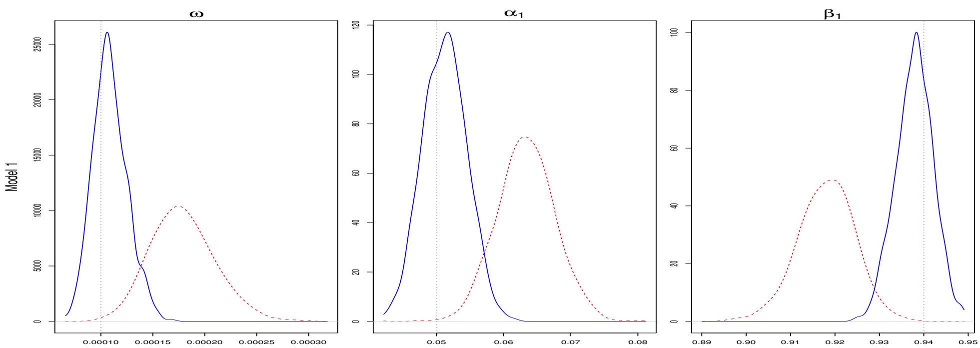

Fig. 2. Probability densities of estimated parameters: **Model 1**
[non-robustified estimates ∼ dashed lines, robustified estimates ∼ solid lines].

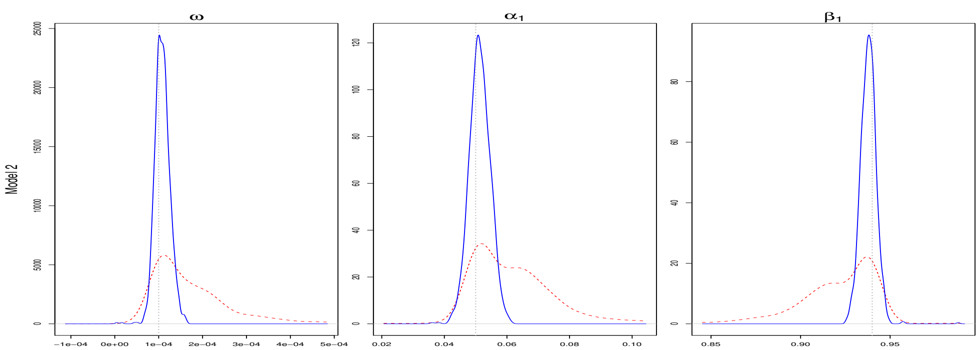

Fig. 3. Probability densities of estimated parameters: **Model 2**
[non-robustified estimates ∼ dashed lines, robustified estimates ∼ solid lines].

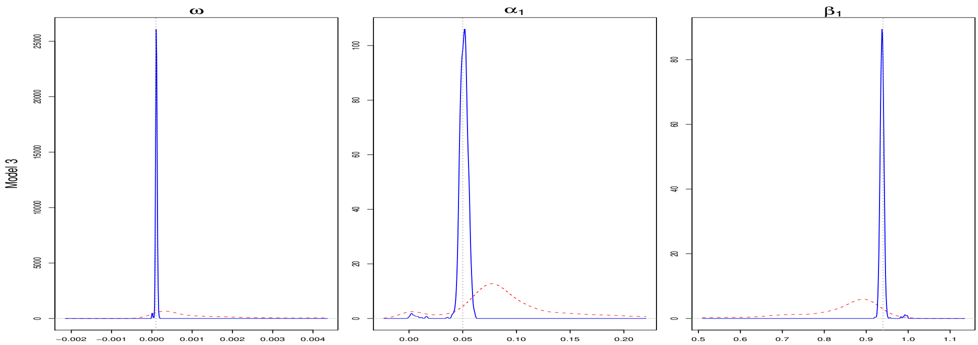

Fig. 4. Probability densities of estimated parameters: **Model 3**
[non-robustified estimates ∼ dashed lines, robustified estimates ∼ solid lines].

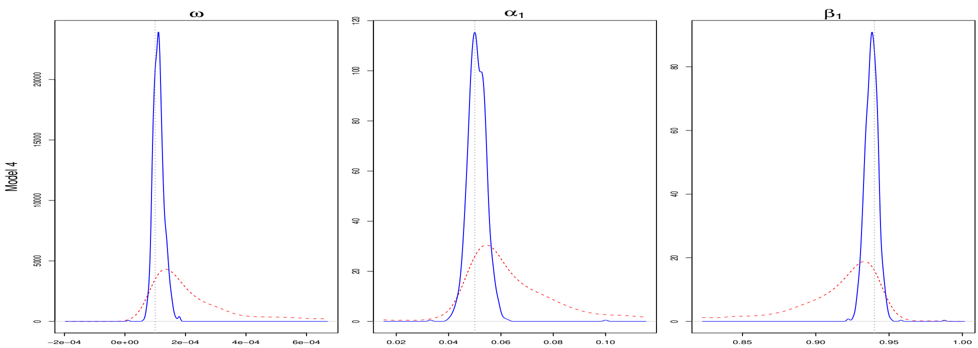

Fig. 5. Probability densities of estimated parameters: **Model 4**
[non-robustified estimates ∼ dashed lines, robustified estimates ∼ solid lines].

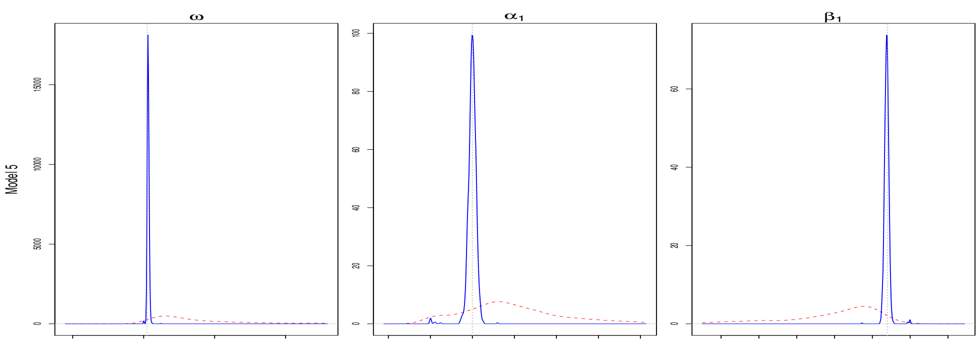

Fig. 6. Probability densities of estimated parameters: **Model 5**
[non-robustified estimates ∼ dashed lines, robustified estimates ∼ solid lines].

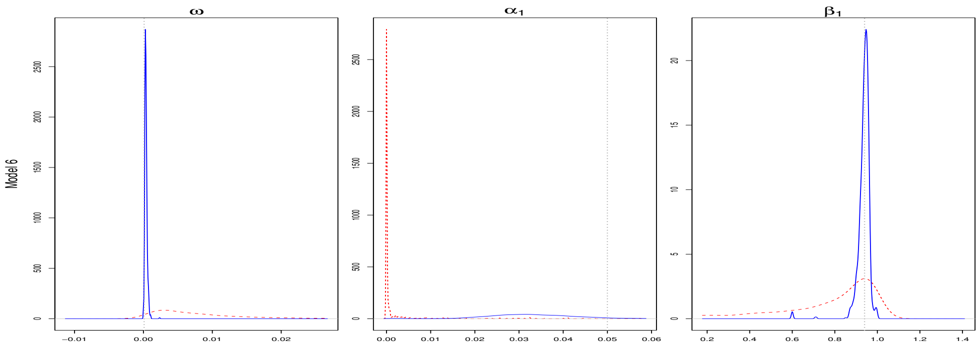

Fig. 7. Probability densities of estimated parameters: **Model 6**
[non-robustified estimates ∼ dashed lines, robustified estimates ∼ solid lines].

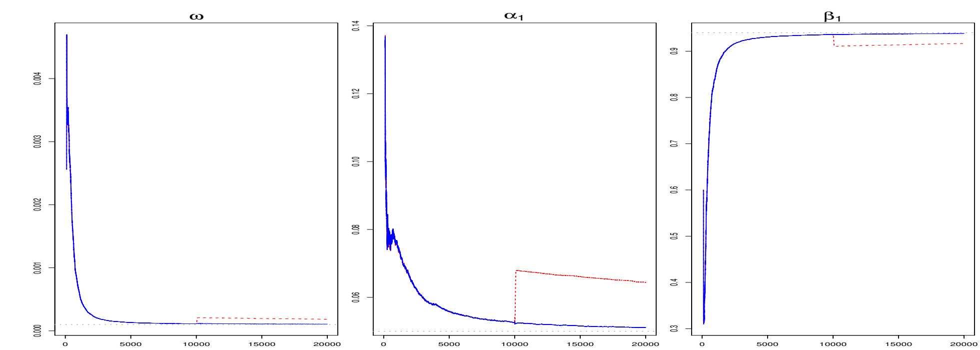

Fig. 8. Time records of medians of estimated parameters: **Model 1**
[non-robustified estimates ∼ dashed lines, robustified estimates ∼ solid lines].

Model 0	$\hat{\omega}_t$	$\hat{\omega}_t^{rob}$	$\hat{\alpha}_{1t}$	$\hat{\alpha}_{1t}^{rob}$	$\hat{\beta}_{1t}$	$\hat{\beta}_{1t}^{rob}$
5,000	0.00003	0.00004	0.00635	0.00636	0.00940	0.00939
10,000	0.00002	0.00002	0.00343	0.00341	0.00473	0.00480
20,000	0.00001	0.00001	0.00240	0.00238	0.00292	0.00292
Model 1	$\hat{\omega}_t$	$\hat{\omega}_t^{rob}$	$\hat{\alpha}_{1t}$	$\hat{\alpha}_{1t}^{rob}$	$\hat{\beta}_{1t}$	$\hat{\beta}_{1t}^{rob}$
5,000	0.00004	0.00004	0.00681	0.00673	0.01032	0.01022
10,000	0.00002	0.00002	0.00499	0.00397	0.00654	0.00497
20,000	0.00008	0.00001	0.01435	0.00227	0.02334	0.00298
Model 2	$\hat{\omega}_t$	$\hat{\omega}_t^{rob}$	$\hat{\alpha}_{1t}$	$\hat{\alpha}_{1t}^{rob}$	$\hat{\beta}_{1t}$	$\hat{\beta}_{1t}^{rob}$
5,000	0.00005	0.00004	0.00885	0.00688	0.01281	0.00989
10,000	0.00003	0.00002	0.00732	0.00371	0.00989	0.00478
20,000	0.00007	0.00001	0.01267	0.00229	0.01972	0.00303
Model 3	$\hat{\omega}_t$	$\hat{\omega}_t^{rob}$	$\hat{\alpha}_{1t}$	$\hat{\alpha}_{1t}^{rob}$	$\hat{\beta}_{1t}$	$\hat{\beta}_{1t}^{rob}$
5,000	0.00022	0.00004	0.02682	0.00694	0.05244	0.01073
10,000	0.00027	0.00002	0.03274	0.00363	0.06364	0.00527
20,000	0.00065	0.00001	0.04147	0.00235	0.08291	0.00321
Model 4	$\hat{\omega}_t$	$\hat{\omega}_t^{rob}$	$\hat{\alpha}_{1t}$	$\hat{\alpha}_{1t}^{rob}$	$\hat{\beta}_{1t}$	$\hat{\beta}_{1t}^{rob}$
5,000	0.00007	0.00004	0.01030	0.00703	0.01660	0.01101
10,000	0.00006	0.00002	0.00948	0.00370	0.01441	0.00523
20,000	0.00009	0.00001	0.01221	0.00242	0.02092	0.00318
Model 5	$\hat{\omega}_t$	$\hat{\omega}_t^{rob}$	$\hat{\alpha}_{1t}$	$\hat{\alpha}_{1t}^{rob}$	$\hat{\beta}_{1t}$	$\hat{\beta}_{1t}^{rob}$
5,000	0.00058	0.00007	0.03504	0.00765	0.09018	0.01327
10,000	0.00077	0.00004	0.04377	0.00413	0.09973	0.00619
20,000	0.00098	0.00003	0.04786	0.00280	0.10609	0.00378
Model 6	$\hat{\omega}_t$	$\hat{\omega}_t^{rob}$	$\hat{\alpha}_{1t}$	$\hat{\alpha}_{1t}^{rob}$	$\hat{\beta}_{1t}$	$\hat{\beta}_{1t}^{rob}$
5,000	0.00630	0.00050	0.05000	0.01550	0.25070	0.04070
10,000	0.00660	0.00040	0.05000	0.01440	0.17110	0.02000
20,000	0.00630	0.00020	0.05000	0.01710	0.08070	0.01230

Tab. 2. Median absolute deviations MAD of estimated parameters
[calculated in times $t = 5000$, 10000, 20000 over 1000 realizations].

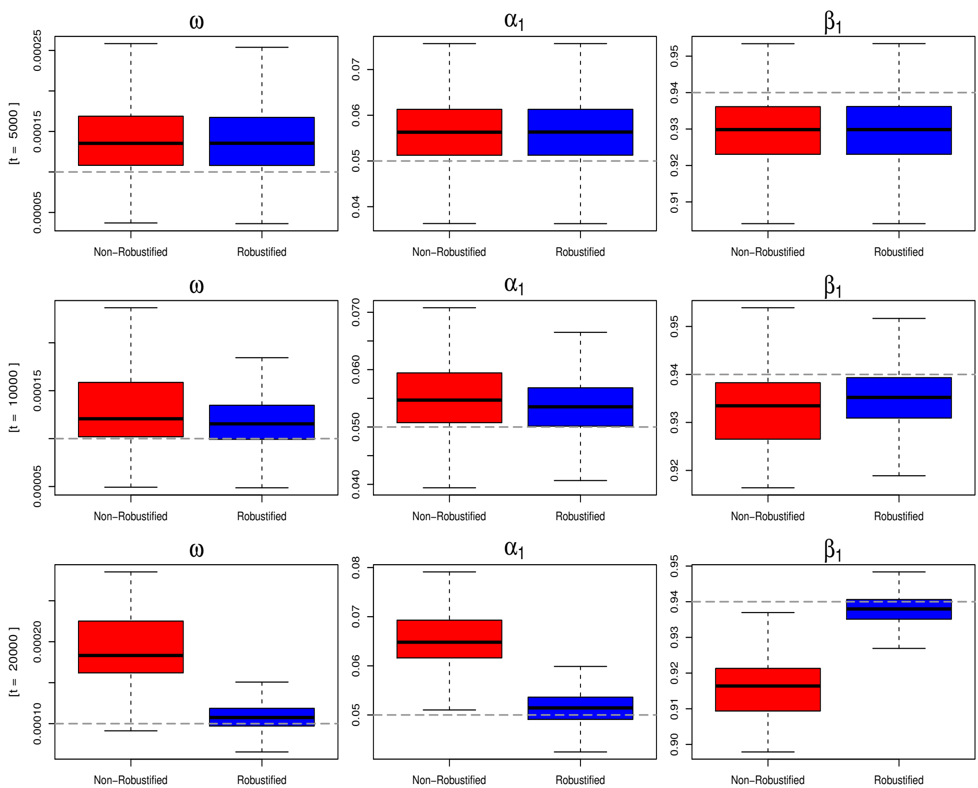

Fig. 9. Boxplots of estimated parameters in **Model 1** in times $t = 5000$, 10000, 20000 before and after robustification.

5. REAL DATA APPLICATIONS

Figure 10 plots the log returns of the daily currency rate CHF/EUR for the period January 2000 - May 2017 which shows an apparent burst of volatility in January 2015 (the initial segment of the data is not displayed due to initialization of the recursive algorithm: the recursive estimates generally tend to be volatile here).[1] It has a clear explanation, i. e. the end of currency regulation of CHF by the Swiss National Bank since 2015: "*CHF was pegged to the Euro for around two years, with the minimum rate (or the floor) at 1.2. As of today, this link has been removed (consequence of recent appreciation of USD against EUR and of CHF weakening against USD).*"[2]

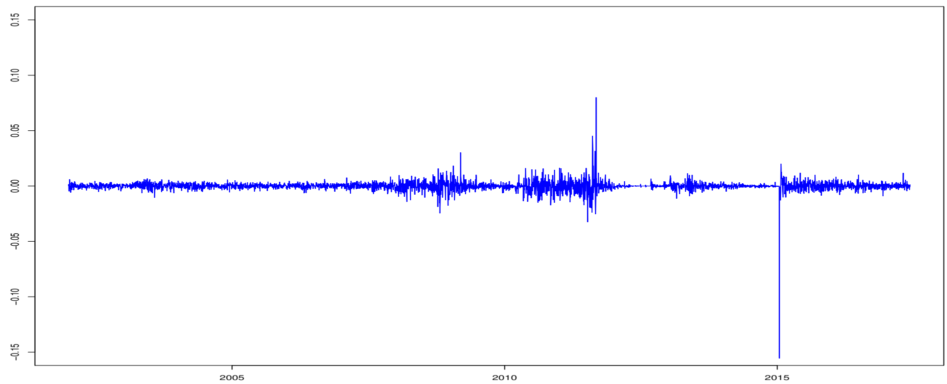

Fig. 10. Log returns of daily currency rate CHF/EUR for the period January 2002 – May 2017.

It was natural to apply the algorithm (9) to handle this time series. This estimation scheme has indeed identified the value of January 2015 as an outlier and corrected it in a proper way. Figure 11 with the recursively estimated parameters of the GARCH(1,1) model displays that one should not ignore the presence of outlier; otherwise there occurs a jump in the estimated parameters. Moreover, another legitimate reason for the application of the algorithm (9) is that it enables the adaptation in the case of parameter changes. Similarly, the one-day-ahead predictions of volatility (see (4) and (10)) would be out of reality without the robustification (see Figure 12). The estimation algorithm (9) has been applied to forty currency rates */EUR (daily log returns) and in some of them the declared robustification has been activated (see Table 3). The results for these daily log returns are not reported here since the figures are just of the type presented herein.

[1] `https://www.ecb.europa.eu/stats/policy_and_exchange_rates/euro_reference_exchange_rates/html/index.en.html`, last accessed 9^{th} June 2017

[2] `https://www.currencyfair.com/blog/what-happened-to-the-swiss-franc-chf-today/`, last accessed 31^{st} July 2017

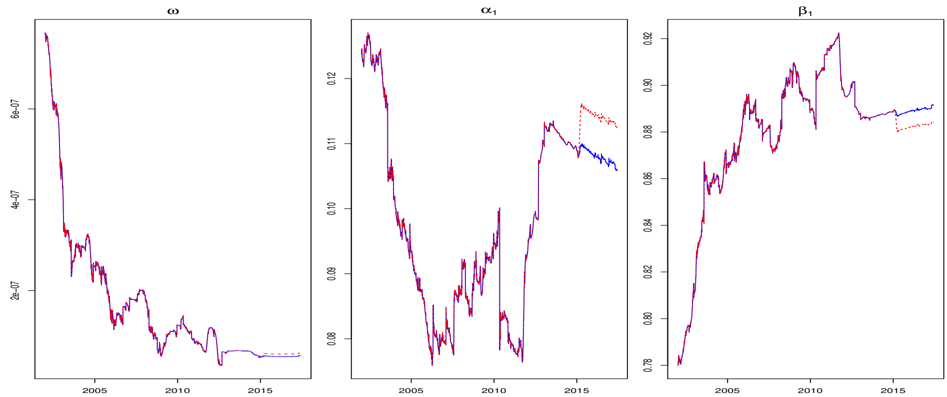

Fig. 11. Time records of estimated parameters for daily log returns of the currency rate CHF/EUR [non-robustified estimates ∼ dashed lines, robustified estimates ∼ solid lines].

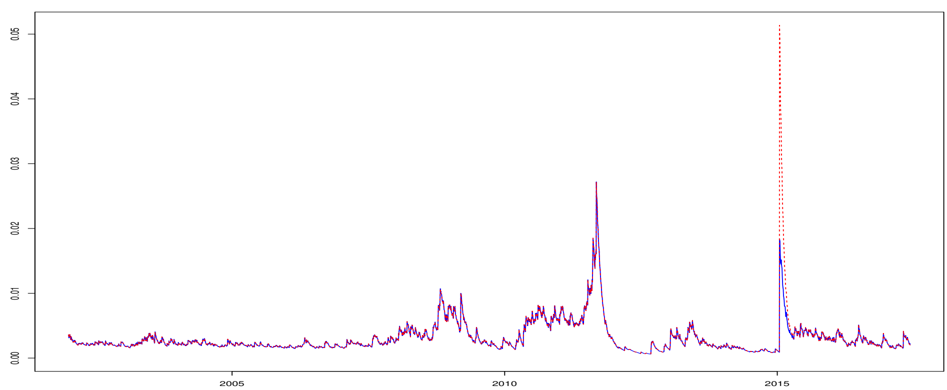

Fig. 12. One-day-ahead predictions of volatility for daily log returns of the currency rate CHF/EUR [non-robustified estimates ∼ dashed line, robustified estimates ∼ solid line].

Currency	Months of identified outliers
USD	1999-07-26
HUF	2003-01-17
ROL	2000-01-04
RON	2006-05-15
CHF	2015-01-15
ISK	2008-11-06
TRL	2001-02-22
TRY	2006-05-12
CAD	2000-01-04
CNY	2006-01-23
MYR	2006-04-18
MYR	2008-03-17
MYR	2008-03-20
NZD	1999-08-25

Tab. 3. Times of activation of robustification in the algorithm (9) [for some daily currency rates */EUR].

6. CONCLUSION

The robust recursive algorithm for the estimation parameters and the corresponding volatility prediction of the GARCH model suggested in this paper seems to be effective for financial data, especially for contaminated log returns in the risk control and regulation when the prediction of volatility is the main concern. The one-stage recursive estimation procedure introduced for the GARCH process in [32] is robustified in such a way that it can distinguish and correct outlaid bursts of volatility. The simulations and real data examples also demonstrate that the suggested procedure enables corresponding adaptations in the case parameter changes.

APPENDIX

The robustified recursive algorithm for estimation of GARCH model (refer to (9)) is based on the robustified version of Kalman filter derived in [13] with the aim to robustify the classical exponential smoothing in time series analysis:

Let us consider the dynamic linear model (DLM) in the form:

$$\boldsymbol{x}_{t+1} = \boldsymbol{x}_t, \tag{A1a}$$

$$y_t = \boldsymbol{h}_t^\top \boldsymbol{x}_t + \nu_t,\ \nu_t \sim i.i.d.\ \mathsf{N}(0, w_t^2), \tag{A1b}$$

where $\boldsymbol{x}_t$ is the state vector in the signal equation (A1a) and $\boldsymbol{h}_t$ is the vector of coefficients in the observation equation (A1b). Then the classical Kalman filter recursion is:

$$\hat{\boldsymbol{x}}_t = \hat{\boldsymbol{x}}_{t-1} + \frac{\boldsymbol{P}_{t-1}\boldsymbol{h}_t}{\boldsymbol{h}_t^\top \boldsymbol{P}_{t-1}\boldsymbol{h}_t + w_t^2}\left(y_t - \boldsymbol{h}_t^\top \hat{\boldsymbol{x}}_{t-1}\right), \tag{A2a}$$

$$\boldsymbol{P}_t = \boldsymbol{P}_{t-1} - \frac{\boldsymbol{P}_{t-1}\boldsymbol{h}_t\boldsymbol{h}_t^\top \boldsymbol{P}_{t-1}}{\boldsymbol{h}_t^\top \boldsymbol{P}_{t-1}\boldsymbol{h}_t + w_t^2}, \tag{A2b}$$

which can be rewritten by means of the filter gain $\boldsymbol{k}_t$ as

$$\hat{\boldsymbol{x}}_t = \hat{\boldsymbol{x}}_{t-1} + \boldsymbol{k}_t\left(y_t - \boldsymbol{h}_t^\top \hat{\boldsymbol{x}}_{t-1}\right), \tag{A3a}$$

$$\boldsymbol{P}_t = \boldsymbol{P}_{t-1} - \boldsymbol{k}_t\boldsymbol{h}_t^\top \boldsymbol{P}_{t-1}, \tag{A3b}$$

$$\boldsymbol{k}_t = \frac{\boldsymbol{P}_{t-1}\boldsymbol{h}_t}{\boldsymbol{h}_t^\top \boldsymbol{P}_{t-1}\boldsymbol{h}_t + w_t^2}. \tag{A3c}$$

The robust version according to [13] replaces (A3a) by

$$\hat{\boldsymbol{x}}_t^{rob} = \hat{\boldsymbol{x}}_{t-1}^{rob} + \boldsymbol{k}_t\Psi(e_t), \tag{A4}$$

i. e. the prediction error

$$e_t = y_t - \hat{y}_{t+1|t} = y_t - \boldsymbol{h}_t^\top \hat{\boldsymbol{x}}_{t-1} \sim \mathsf{N}(0, \boldsymbol{h}_t^\top \boldsymbol{P}_{t-1}\boldsymbol{h}_t + w_t^2) \tag{A5}$$

is trimmed by means of the robustifying function $\Psi(\cdot)$ defined as

$$\Psi(e_t) = \begin{cases} \operatorname{sign}(e_t)u_{1-\alpha/2}\sqrt{\boldsymbol{h}_t^\top \boldsymbol{P}_{t-1}\boldsymbol{h}_t + w_t^2} & \text{for } |e_t| > u_{1-\alpha/2}\sqrt{\boldsymbol{h}_t^\top \boldsymbol{P}_{t-1}\boldsymbol{h}_t + w_t^2}, \\ e_t & \text{otherwise.} \end{cases} \tag{A6}$$

For instance, if applying this scheme to the classical model AR(1) then the strong consistency of the recursive formulas for the robust estimation of the autoregressive parameter can be shown (see [14]).

ACKNOWLEDGEMENT

This research was supported by the grant 17-00676S provided by the Grant Agency of the Czech Republic.

(Received November 26, 2017)

REFERENCES

[1] A. Aknouche and H. Guerbyenne: Recursive estimation of GARCH models. Comm. Statist. Simul. Comput. *35* (2006), 925–938. DOI:10.1080/03610910600880328

[2] N. S. Balke and T. B. Fomby: Large shocks, small shocks, and economic fluctuations: outliers in macroeconomics time series. J. Appl. Econometr. *31* (1994), 307–327. DOI:10.1002/jae.3950090205

[3] T. Bernholt, R. Fried, U. Gather, and I. Wegener: Modified repeated median filters. Statist. Comput. *16* (2006), 177–192. DOI:10.1007/s11222-006-8449-1

[4] T. Bollerslev: Generalized autoregressive conditional heteroskedasticity. J. Econometr. *31* (1986), 307–327. DOI:10.1016/0304-4076(86)90063-1

[5] A. Bose and K. Mukherjee: Estimating the ARCH parameters by solving linear equations. J. Time Series Anal. *24* (2003), 127–136. DOI:10.1111/1467-9892.00296

[6] L. E. Calvet, V. Czellar, and E. Ronchetti: Robust filtering. J. Amer. Statist. Assoc. *110* (2015), 1591–1606. DOI:10.1080/01621459.2014.983520

[7] M. A. Carnero, D. Peña, and E. Ruiz: Effects of outliers on the identification and estimation of GARCH models. J. Time Series Anal. *28* (2007), 471–497. DOI:10.1111/j.1467-9892.2006.00519.x

[8] M. A. Carnero, D. Peña, and E. Ruiz: Estimating GARCH volatility in the presence of outliers. Econom. Lett. *114* (2012), 86–90. DOI:10.1016/j.econlet.2011.09.023

[9] A. Charles: Forecasting volatility with outliers in GARCH models. J. Forecast. *27* (2008), 551–565. DOI:10.1002/for.1065

[10] A. Charles and O. Darné: Outliers and GARCH models in financial data. Econom. Lett. *86* (2005), 347–352. DOI:10.1016/j.econlet.2004.07.019

[11] T. Cipra: Robust exponential smoothing. J. Forecast. *11* (1992), 57–69. DOI:10.1002/for.3980110106

[12] T. Cipra: Robust recursive estimation in nonlinear time-series. Comm. Statist. Theory Methods *27* (1998), 1071–1082. DOI:10.1080/03610929808832146

[13] T. Cipra and T. Hanzák: Exponential smoothing for time series with outliers. Kybernetika *47* (2011), 165–178.

[14] T. Cipra and R. Romera: Robust Kalman filter and its applications in time series analysis. Kybernetika *27* (1991), 481–494.

[15] R. Crevits and C. Croux: Forecasting using robust exponential smoothing with damped trend and seasonal components. Working paper KBI_1714, KU Leuven, Leuven 2016 (DOI:10.13140RG.2.2.11791.18080).

[16] C. Croux and S. Gelper: Computational aspects of robust Holt-Winters smoothing based on M-estimation. Appl. Math. *53* (2008), 163–176. DOI:10.1007/s10492-008-0002-4

[17] C. Croux, S. E. C. Gelper, and K. Mahieu: Robust exponential smoothing of multivariate time series. Comput. Statist. Data Anal. *54* (2010), 2999–3006. DOI:10.1016/j.csda.2009.05.003

[18] R. Dalhaus and S. S. Rao: A recursive online algorithm for the estimation of time-varying ARCH parameters. Bernoulli *13* (2007), 389–422. DOI:10.3150/07-bej5009

[19] R. F. Engle: Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica *50* (1982), 987–1007. DOI:10.2307/1912773

[20] B. Eraker, M. Johannes, and N. Polson: The impact of jumps in volatility and returns. J. Finance *58* (2003), 1269–1300. DOI:10.1111/1540-6261.00566

[21] A. Fasso: Recursive least squares with ARCH errors and nonparametric modelling of environmental time series. Working Paper *6*, University of Bergamo 2009.

[22] J. Franke, W. K. Härdle, and C. M. Hafner: Statistics of Financial Markets: An Introduction. Springer, Berlin 2011. DOI:10.1007/978-3-642-16521-4

[23] P. H. Franses and H. Ghijsels: Additive outliers, GARCH and forecasting volatility. Int. J. Forecast. *15* (1999), 1–9. DOI:10.1016/s0169-2070(98)00053-3

[24] P. Galeano and D. Peña: Finding outliers in linear and nonlinear time series. In: Robustness and Complex Data Structures (C. Becker, R. Fried, S. Kuhnt, eds.), Springer, Berlin 2013, pp. 243–260. DOI:10.1007/978-3-642-35494-6_15

[25] S. Gelper, R. Fried, and C. Croux: Robust forecasting with exponential and Holt–Winters smoothing. J. Forecast. *29* (2009), 285–300. DOI:10.1002/for.1125

[26] L. Gerencsér, Z. Orlovits, and B. Torma: Recursive estimation of GARCH processes. In: Proc. 19th International Symposium on Mathematical Theory and Systems – MTNS (A. Edelmayer, ed.), Eötvös Loránd University, Budapest 2010, pp. 2415–2422.

[27] A. Grané and H. Veiga: Wavelet based detection of outliers in financial time series. Comput. Statist. Data Anal. *54* (2010), 2580–2593. DOI:10.1016/j.csda.2009.12.010

[28] A. V. Gregory and J. J. Reeves: Estimation and inference in ARCH model in the presence of outliers. J. Financ. Econometr. *8* (2010), 547–569. DOI:10.1093/jjfinec/nbq028

[29] C. Grillenzoni: Optimized adaptive prediction. J. Ital. Statist. Soc. *6* (1997), 37–58. DOI:10.1007/bf03178900

[30] C. Grillenzoni: Recursive generalized M-estimators of system parameters. Technometrics *39* (1997), 211–224. DOI:10.2307/1270909

[31] R. Hendrych and T. Cipra: Robustified on-line estimation of the EWMA models: Simulations and applications. In: Proc. 33rd International Conference Mathematical Methods in Economics (D. Martinčák, J. Ircingová, and P. Janeček, eds.). University of West Bohemia, Pilsen 2014, pp. 237–242. DOI:10.3311/ppee.9684

[32] R. Hendrych and T. Cipra: Self-weighted recursive estimation of GARCH models. Comm. Statist. Simul. Comput. *47* (2018), 315–328. DOI:10.1080/03610918.2015.1053924

[33] J. B. Hill: Robust estimation and inference for heavy tailed GARCH. Bernoulli *21* (2015), 1629–1669. DOI:10.3150/14-bej616

[34] L. K. Hotta and R. S. Tsay: Outliers in GARCH processes. In: Economic time series: Modeling and seasonality (W. R. Bell, S. H. Holan, and T. S. McElroy, eds.). CRC Press, Boca Raton 2012, pp. 337–358. DOI:10.1201/b11823-20

[35] R. J. Hyndman, A. B. Koehler, J. K. Ord, and R. D. Snyder: Forecasting with Exponential Smoothing. The State Space Approach. Springer, Berlin 2008. DOI:10.1111/j.1751-5823.2009.00085_17.x

[36] J. Jiang, Q. Zhao, and Y. V. Hui: Robust modelling of ARCH models. J. Forecast. *20* (2001), 111–133. DOI:10.1002/1099-131x(200103)20:2¡111::aid-for786¿3.0.co;2-n

[37] J. Kierkegaard, J. Nielsen, L. Jensen, and H. Madsen: Estimating GARCH models using recursive methods. Working paper, Technical University of Denmark, Lyngby 2000 (`http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.40.703`).

[38] K. R. Koch and Y. Yang: Robust Kalman filter for rank deficient observation models. J. Geodesy *72* (1998), 436–441. DOI:10.1007/s001900050183

[39] V. Lanius and U. Gather: Robust online signal extraction from multivariate time series. Comput. Statist. Data Anal. *54* (2010), 966–975. DOI:10.1016/j.csda.2009.10.009

[40] S. Ling: Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/ IGARCH models. J. Econometr. *140* (2007), 849–873. DOI:10.1016/j.jeconom.2006.07.016

[41] L. Ljung: System Identification: Theory for the User. Prentice Hall PTR, Upper Saddle River 1999.

[42] L. Ljung and T. S. Söderström: Theory and Practice of Recursive Identification. MIT Press, Cambridge 1983.

[43] J. Michálek: Robust methods in exponential smoothing. Kybernetika *32* (1996), 289–306.

[44] N. Muler and V. Yohai: Robust estimates for GARCH models. J. Statist. Planning Inference *138* (2008), 2918–2940. DOI:10.1016/j.jspi.2007.11.003

[45] B.-J. Park: An outlier robust GARCH model and forecasting volatility of exchange rate returns. J. Forecast. *21* (2002), 381–393. DOI:10.1002/for.827

[46] L. Peng and Q. Yao: Least absolute deviations estimation for ARCH and GARCH models. Biometrika *90* (2003), 967–975. DOI:10.1093/biomet/90.4.967

[47] R. Romera and T. Cipra: On practical implementation of robust Kalman filtering. Comm. Statist. Simul. Comput. *24* (1995), 461–488. DOI:10.1080/03610919508813252

[48] P. Ruckdeschel, B. Spangl, and D. Pupashenko: Robust Kalman tracking and smoothing with propagating and non propagating outliers. Statist. Papers *55* (2014), 93–123. DOI:10.1007/s00362-012-0496-4

[49] S. Sakata and H. White: High breakdown point conditional dispersion estimation with application to S&P 500 daily returns volatility. Econometrica *66* (1998), 529–567. DOI:10.2307/2998574

[50] H. U. Shaolin, K. Meinke, H. Ouyang, and S. Guoji: Outlier-tolerant Kalman filter of state vectors in linear stochastic system. Int. J. Advanced Computer Sci. Appl. *2* (2011), 37–41. DOI:10.14569/ijacsa.2011.021206

[51] T. S. Söderström and P. Stoica: System Identification. Prentice Hall, New York 1989.

[52] R. S. Tsay: Analysis of Financial Time Series. Wiley, Hoboken 2013.

[53] Y. Yang: Adaptively robust Kalman filters with applications in navigation. In: Sciences of Geodesy (G. Xu, ed.), Springer, Berlin 2010, pp. 49–82. DOI:10.14569/ijacsa.2011.021206

[54] Y. Yang, W. Gao, and X. Zhang: Robust Kalman filtering with constraints: a case study for integrated navigation. J. Geodesy *84* (2010), 373–381. DOI:10.1007/s00190-010-0374-6

[55] R. Zhang and S. Ling: Asymptotic inference for AR models with heavy-tailed G-GARCH noises. Econometr. Theory *31* (2015), 880–890. DOI:10.1017/s0266466614000632

[56] K. Zhu and W. K. Li: A new Pearson-type QMLE for conditionally heteroskedastic models. J. Business Econom. Statist. *33* (2015), 552–565. DOI:10.1080/07350015.2014.977446

[57] K. Zhu and S. Ling: Global self-weighted and local quasi-maximum exponential likelihood estimators for ARMA-GARCH/IGARCH models. Ann. Statist. *39* (2011), 2131–2163. DOI:10.1214/11-aos895

[58] K. Zhu and S. Ling: LADE-based inference for ARMA models with unspecified and heavy-tailed heteroscedastic noises. J. Amer. Statist. Assoc. *110* (2015), 784–794. DOI:10.1080/01621459.2014.977386

Tomáš Cipra, Dept. of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Praha 8. Czech Republic.
e-mail: cipra@karlin.mff.cuni.cz

Radek Hendrych, Dept. of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Praha 8. Czech Republic.
e-mail: hendrych@karlin.mff.cuni.cz