COMPARISON OF ROBUST MOMENT METHODS FOR PARAMETER ESTIMATION IN AUTOREGRESSIVE PROCESS

SAMUEL FLIMMEL, JAN FOJTÍK, IVANA MALÁ, JIŘÍ PROCHÁZKA
University of Economics, Prague, Faculty of Informatics and Statistics,
Department of Statistics and Probability,
W. Churchill Sq. 4, Prague, Czech Republic
e-mails: samuel.flimmel@vse.cz, xfojj00@vse.cz, malai@vse.cz, xproj16@vse.cz

Abstract

Autoregressive process AR(p) is very popular and frequently used when working with time series, especially in financial mathematics. One of the requirements for working with AR(p) is the ability to estimate parameters of the model correctly. However, we currently often deal with big data, which can lead, among others, to a higher probability of outlier presence. As it is known, standard methods for parameter estimation are often not able to work correctly with outliers, and, consequently, standard estimates are usually biased. Therefore, working with sufficiently robust methods has increased in importance. In this paper, we present several robust moment methods for parameter estimation in AR(p) and we compare them using a simulation study. Outliers in the simulations are modelled using two most frequently used outlier models: additive outlier (AO) and innovative outlier (IO). For the simulation study, we use the R statistical software.

Keywords: *robust methods, parameter estimation, autoregressive process*

JEL Codes: *C02, C22, G10*

1. Introduction

Autoregressive processes AR(p) are well known and widely used (not only) in the financial world. They are one of the basic econometric tools used for time series modeling, or, more often, to explain the residue of randomness in a random process.

To estimate an AR(p) process from time series data, firstly, it is necessary to solve the stationarity and seasonality of the given time series, what is often done by decomposing the process. The next important step is to determine the order of the AR(p) process. The order estimation using robust methods was the main point of interest of the paper (Flimmel *et al.*, 2017) at the last conference. After the order determination, we can focus on the final step, which is the estimation of parameters. To describe this last step is the goal of this paper.

Due to extensive usage of big data, we face nowadays problems related to an increased probability of outlier presence. Outliers complicate AR process estimation because they can cause an estimator to be biased. There are several robust methods for parameter estimation that take into account outlier presence. These methods should be less sensitive to outlier presence and, generally, they should give better results (Chan and Wei, 1992).

Maronna *et al.* (2006), Rousseeuw and Croux (1992), and others suggest various approaches that should help with the problem. Naturally, every method has certain advantages as well as disadvantages. We choose four of these approaches, and, after their brief introduction, we compare them by performing a simulation study.

In Section 2, we establish the notation that we work with in this paper. In Section 3, we introduce two basic outlier models: the additive outlier model and the innovative outlier model. In Section 4, we give a brief introduction to the robust methods that we work with. In Section 5, we show results of the simulation study used to compare the methods.

2. Definitions and notation

Let us define white noise, which is a zero-mean mutually uncorrelated time series $\{\varepsilon_n, n \in N_0\}$ with an unknown constant variance $\sigma_\varepsilon^2 > 0$.

We define an autoregressive process AR(p) by the equation:

$$X_n = \varphi_1 X_{n-1} + \varphi_2 X_{n-2} + \cdots + \varphi_p X_{n-p} + \varepsilon_n, \tag{1}$$

where $\boldsymbol{\varphi} = \varphi_1, \varphi_2, \ldots, \varphi_p \in R^p$ are parameters, $\{\varepsilon_n, n \in N_0\}$ is the white noise and $\varphi_p \neq 0$.

We define an autocovariance function of the lag k $R(k)$ of the stationary process $\{X_n, n \in N_0\}$ as:

$$R(k) = E(X_k - \mu)(X_0 - \mu), \tag{2}$$

where μ is the expected value of the process.

We define an autocorrelation function (ACF) of the lag k of the stationary process $\{X_n, n \in N_0\}$ as:

$$\rho(k) = \frac{R(k)}{\sigma_X^2}, \tag{3}$$

where σ_X^2 is the variance of the process.

3. Outlier models

There are several models for simulating outliers in a time series (e.g. Maronna *et al.*, 2006). Let us introduce two of them: the additive outlier model and the innovative outlier model.

3.1 Additive outliers

The additive outlier (AO) model was originally introduced by Fox (1972). In the AO model, we assume that we do not observe the process of interest $\{X_n, n \in N_0\}$ but, actually, we observe a process$\{Y_n, n \in N_0\}$ defined as:

$$Y_n = X_n + Z_n, \tag{4}$$

where processes $\{X_n, n \in N_0\}$ and $\{Z_n, n \in N_0\}$ are assumed to be independent of one another.

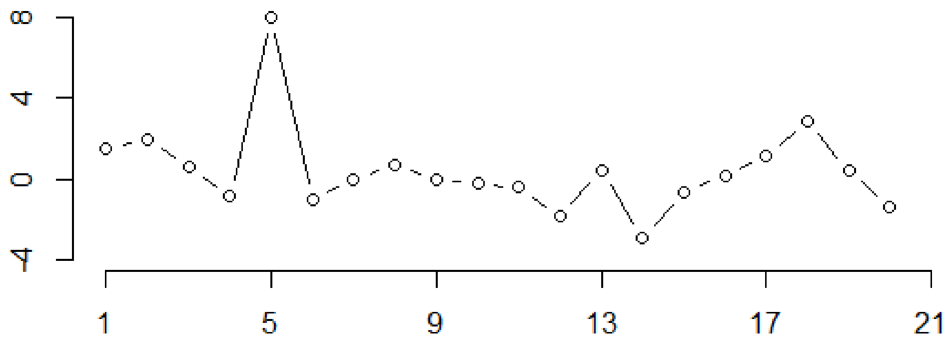
Let $\{Z_n, n \in N_0\}$ be a process with independent and identically distributed (i.i.d) random variables that have a normal mixture distribution with a degenerate central component:

$$Z_n \sim (1-\beta)\delta_0 + \beta N(\mu_Z, \sigma_Z^2), \tag{5}$$

where δ_0 is the point mass distribution located at zero, and we assume that the normal component $N(\mu_Z, \sigma_Z^2)$ has a variance significantly higher than the process $\{X_n, n \in N_0\}$, $\sigma_Z^2 \gg \sigma_X^2$.

The probability of outlier occurrence is represented by β, which is usually small. Consequently, the probability of occurrence of 2 outliers in a row is a much smaller β^2, which means that the AO model generates mostly isolated outliers.

Figure 1: Example of an additive outlier.



Source: The authors' work

3.2 Innovative outliers

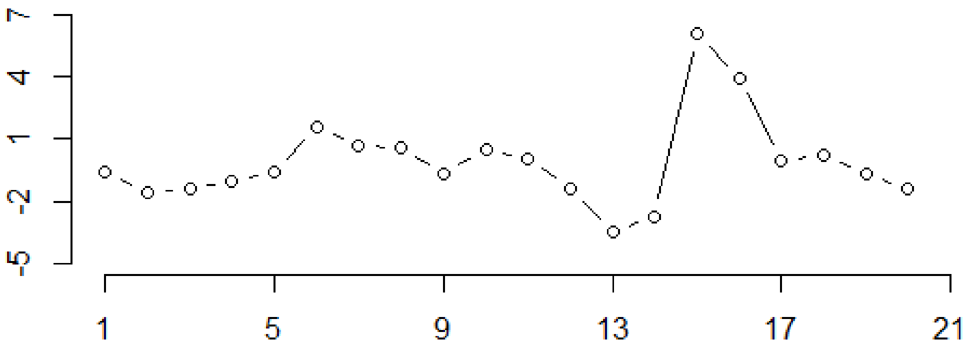
The innovative outlier (IO) model was originally introduced as well by Fox (1972), who used the term "type II outliers". The IO model works with a highly specialized form of outliers that can occur in a linear processes such as AR(*p*), ARMA(*p*,*q*) or ARIMA(*p*,*d*,*q*).

For simplicity, we introduce only a special case that we use for this paper. The white noise process $\{\varepsilon_n, n \in N_0\}$ from the definitions of the AR(*p*) is sometimes also called the innovation process. The IO model generates outliers directly in the innovation process. For the IO model, we assume independent identically distributed (i.i.d.) random variables in the process $\{\tau_n, n \in N_0\}$ with a normal mixture distribution

$$\tau_n \sim (1-\beta)N(0,\sigma_\tau^2) + \beta N(0,\sigma_I^2), \tag{6}$$

where $\sigma_I^2 \gg \sigma_\tau^2$. The IO outlier affects not only the current observation but also subsequent observations, so the IO gives greater problems at correct estimation in AR processes.

Figure 2: Example of an innovative outlier.



Source: The authors' work

4. Robust methods

The moment method is based on an autocorrelation function. When we have an ACF estimator, then we are able to estimate the parameters of the process using the relation:

$$\widehat{\boldsymbol{\varphi}} = \begin{pmatrix} 1 & \hat{\rho}(1) & \cdots & \hat{\rho}(p-1) \\ \hat{\rho}(1) & 1 & \cdots & \hat{\rho}(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\rho}(p-1) & \hat{\rho}(p-2) & \cdots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \hat{\rho}(1) \\ \hat{\rho}(2) \\ \vdots \\ \hat{\rho}(p) \end{pmatrix}. \tag{7}$$

It means that, if we have a robust ACF estimator, practically, we have as well a robust parameter estimator.

In the following subsections, we introduce 4 robust methods for ACF estimation. We use them in the equation (7) and compare them by means of a simulation study. For an overview of robust methods for ACF estimators, you can see Dürre *et al.* (2015).

4.1 Method based on median correlation

The method based on median correlation was introduced by Chakhchoukh (2010). This method is quite intuitive because, instead of using the mean, we work only with the median from the equation (4). The median is well known as a robust estimator of location. Firstly, we centre our observations $X_0, X_1, \dots, X_m$ by:

$$\tilde{X}_i = X_i - med(X_0, X_1, \dots, X_m), \tag{8}$$

where $med(\cdot)$ gives the median of the observations.

Then we estimate the ACF using:

$$\hat{\rho}_{med}(k) = \frac{med\left(\tilde{X}_0\tilde{X}_k, \tilde{X}_1\tilde{X}_{k+1}, \dots, \tilde{X}_{m-k}\tilde{X}_m\right)}{med\left(\tilde{X}_0^{\ 2}, \tilde{X}_1^{\ 2}, \dots, \tilde{X}_m^{\ 2}\right)}. \tag{9}$$

For a consistent estimation of $\rho(k)$, a nonlinear transformation of $\hat{\rho}_{med}(k)$, which has to be determined numerically, is necessary (Dürre *et al.*, 2015). This nonlinear transformation can be based on a Monte Carlo simulation.

4.2 Method based on trimming

The method based on trimming is described in several papers (e.g. Dürre *et al.*, 2015). This method is based on omitting certain terms in the calculation of the standard ACF. Firstly, we estimate an autocovariance function:

$$\hat{R}_{trim}^{(\tau)}(k) = \frac{1}{\sum_{i=0}^{m-k} L_i^{(\tau)} L_{i+k}^{(\tau)}} \left(\sum_{i=0}^{m-k} \left(X_i - \bar{X}^{(\tau)}\right)\left(X_{i+k} - \bar{X}^{(\tau)}\right) L_i^{(\tau)} L_{i+k}^{(\tau)} \right), \tag{10}$$

where

$$\bar{X}^{(\tau)} = \frac{1}{\sum_{i=0}^{m} L_i^{(\tau)}} \sum_{i=0}^{m} X_i\, L_i^{(\tau)} \text{ and } L_i^{(\tau)} = \begin{cases} 1, & X_t < X_i < X_{m-t} \\ 0, & else \end{cases}, \tag{11}$$

with $t = \lfloor \tau(m+1) \rfloor$-1 for some $0 \le \tau < 0.5$. Chan and Wei (1992) proposed $0.01 \le \tau \le 0.1$, depending on the suspected percentage of outliers.

The ACF estimator $\hat{\rho}_{trim}^{(\tau)}(k)$ is calculated as the ratio of the trimmed autocovariance and trimmed variance $\hat{R}_{trim}^{(\tau)}(0)$. Similarly, as for $\hat{\rho}_{med}(k)$, a nonlinear transformation is necessary to obtain a consistent estimation of $\rho(k)$. In addition, $\hat{\rho}_{med}(k)$ is a limiting case of $\hat{\rho}_{trim}^{(\tau)}(k) \xrightarrow[\tau \to 0.5]{} \hat{\rho}_{med}(k)$.

4.3 Method based on the Gnanadesikan-Kettenring approach

The method based on the Gnanadesikan-Kettenring approach, named after the researchers who introduced it (Gnanadesikan and Kettenring, 1972), exploits an idea that can be formulated as:

$$R(k) = \frac{1}{4}\big(var(X_0 + X_k) - var(X_0 - X_k)\big). \tag{12}$$

This method is also called a scale approach. Equation (15) is written here already in a simplified form. For a more general formula, you can see e.g. Huber (1981).

In the context of scale estimation, Rousseeuw and Croux (1992) proposed a robust estimator $\mathrm{Q_m}$:

$$Q_m = c\big[|X_i - X_j|, i < j\big]_l, \tag{13}$$

where *c* is a factor included for consistency (at the Gaussian distribution $c = 2.2191$), and $[\cdot]_l$ is the *l*-th order statistic, where *l* is defined as:

$$l = \left\lfloor \frac{\binom{m}{2} + 2}{4} \right\rfloor + 1, \tag{14}$$

where the $\lfloor\cdot\rfloor$ function denotes the integer part, and *m* is the number of observations.

Using the formulas above, we obtain an estimator:

$$\hat{\rho}_{GK}(k) = \frac{Q_{m-k}^2(u + v) - Q_{m-k}^2(u - v)}{Q_{m-k}^2(u + v) + Q_{m-k}^2(u - v)}, \tag{15}$$

where *u* is the vector $(X_{m-k}, X_{m-k+1}, \dots, X_m)$ and v is the vector $(X_0, X_1, \dots, X_k)$. The method of this robust ACF estimator was presented by Ma and Genton (2000).

4.4 Method based on robust filtering

The method based on robust filtering was described by Maronna *et al.* (2006). This approach takes the time series structure into account. The idea is to have robustly filtered values instead of the original observation and to calculate ACF from these filtered values. Principally, we replace outliers by some reasonable values.

Firstly, we estimate the order of the AR process which we use for robust filtering. It can be done by a robust AIC criterion that was proposed also by Maronna *et al.* (2006). Alternatively, we can use a “long” AR process instead.

Secondly, we obtain fitted values using a robustly filtered τ-scale estimate. Finally, we calculate the autocorrelation function.

5. Simulation study

The simulation study was designed in the R software (R Core Team, 2013). We use the R package *robts* (see Dürre *et al.*, 2016), which, however, had not been yet approved by CRAN at the time of the study. Therefore, certain functions were coded by the authors of this paper to validate the correctness of the package. After the validation was succesfull, we used functions from the package to obtain estimations in the simulation study.

To evaluate the estimation accuracy, we use two criteria: *mean absolute error* (MAE) and *mean absolute percentage error* (MAPE). The mean absolute error is given by:

$$MAE = \frac{\sum_{i=1}^{s} \left| \hat{\varphi}_j^i - \varphi_j \right|}{s}, \tag{16}$$

where $\hat{\varphi}_j^i$ is an estimation of the *i*-th simulation and the *j*-th component of the vector of parameters, and *s* is the number of simulations.

The mean absolute percentage error is defined by

$$MAPE = \frac{100}{s} \sum_{i=1}^{s} \left| \frac{\hat{\varphi}_j^i - \varphi_j}{\varphi_j} \right|, \tag{17}$$

where $\hat{\varphi}_j^i$ is an estimation of the *i*-th simulation and the *j*-th component of the vector of parameters, and *s* is the number of simulations.

Firstly, we have the 3 simplest models: AR(1), AR(2) and AR(3). The percentage of outliers present in a single simulation is chosen randomly with a uniform distribution, i.e. $\beta \sim U([0.00, 0.05])$. The value of μ_Z was set to zero and σ_Z, σ_I were set to 10. For every examined case, we run 10 000 simulations and have 1000 observations. In each case, we estimate all parameters of the model using all 4 described methods. The accuracy of the method is evaluated by the 2 mentioned criteria: MAE and MAPE.

Absolute values of the parameters of the AR(*p*) process are generated randomly with a uniform distribution, i.e. $\varphi_i \sim U((0.2, 1.0))$. Values of φ_i being close to zero are not taken into account because they are difficult to observe. Parameter signs are generated randomly using a Bernoulli's distribution with the probability of success $\pi = 0.5$. Subsequently, we check whether these parameters lead to a stationary process, and, if necessary, we repeat the procedure.

Secondly, we work with models AR(*p*) for *p* = 1,...,5. We use the same distribution for the percentage of outliers present in a single simulation and, similarly, the same distribution of the parameters in the model AR(*p*). Subsequently, all estimated parameters are evaluated together using a box plot, in order to have a graphical representation of results.

We use the additive outlier model and the innovative outlier model (e.g. Maronna *et al.*, 2006) in the simulation study, therefore, we split the following section into 2 subsections.

5.1 Additive outlier model

We work with the additive outlier model in this subsection. We work only with stationary processes in our simulations, however, the estimated processes are not necessarily stationary. In Table 1 we can see the percentage of estimated processes that are stationary.

The simulations evaluated in Table 1 are used in described simulations. It means that there are 10 000 simulations for the models AR(1), AR(2) and AR(3), and, moreover, another 10 000 simulations of randomly chosen AR(*p*), where *p* = 1,2,3,4,5. It implies a higher representation of the 3 simplest models AR(1), AR(2) and AR(3) (approximately 12 000 simulations for every model), and a lower representation of the models AR(4) and AR(5) (approximately 2 000 simulations).

As the first observation deduced from Table 1, there is the fact that the percentage of being stationary, in general, decreases with the growing order *p* of the autoregressive model. The lowest ratio of processes estimated as stationary is obtained by the trimming approach, for example, for the AR(5), it is only two thirds. In total it is 92%.

The second worst result is reached using the median approach, where, in total, the rate of being stationary for the estimated processes is 95.4%.

Almost every process, when using the GK approach, is estimated as stationary. The robust filtering approach has the full percentage.

Table 1: Stationarity of the estimated processes in AO model.

AR(p)	median approach	trimming approach	GK approach	robust filtering approach
AR(1)	99.2%	99.4%	100.0%	100.0%
AR(2)	97.6%	96.4%	99.9%	100.0%
AR(3)	93.0%	86.5%	99.8%	100.0%
AR(4)	87.8%	80.0%	99.2%	100.0%
AR(5)	80.7%	65.3%	98.7%	100.0%
Total	**95.4%**	**92.0%**	**99.8%**	**100.0%**

Source: The authors' work.

In Table 2, we can see a comparison of the 4 above described methods. The worst results are given by the method based on trimming. Only in the case of φ_2 in the AR(2) model, the median approach is worse than the trimming approach, but the difference is negligible.

Very similar results to the trimming approach are given by the median approach. A slightly higher difference can be seen in the case of the AR(3), but it is still in the same scale.

There is a notable difference between the two simple methods and the method based on the GK approach. Values of the MAE criterion are 2 times higher for the AR(1), when compared to the GK approach, and even 5 times higher in the case of the AR(3).

The robust filtering approach gives very similar, maybe even slightly better results, if compared with the GK approach. It is natural because the robust filtering approach represents a more sophisticated method. However, surprisingly, for the AR(2) model, the results are different and significantly worse. Additional simulations for this special case were made. These additional simulations confirmed the worse result for the AR(2) in comparison with the GK approach. The reason why, for the AR(1) and the AR(3), the robust filtering approach is better, and for the AR(2) it is not, is unknown and will be explored.

Table 2: Comparison of 4 methods using AO model.

AR(p)	φ_k	median approach		trimming approach		GK approach		robust filtering approach	
		MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE
AR(1)	φ_1	.0462	9.3%	.0421	9.8%	.0229	4.9%	.0205	4.6%
AR(2)	φ_1	.1178	21.6%	.1256	24.4%	.0336	6.6%	.0634	14.7%
	φ_2	.1160	29.1%	.1158	28.9%	.0325	7.0%	.0754	14.3%
AR(3)	φ_1	.3288	55.9%	.4012	74.8%	.0573	10.9%	.0300	6.3%
	φ_2	.3209	74.7%	.3904	91.9%	.0579	12.2%	.0368	8.8%
	φ_3	.3423	83.3%	.4238	97.9%	.0625	14.7%	.0362	8.2%

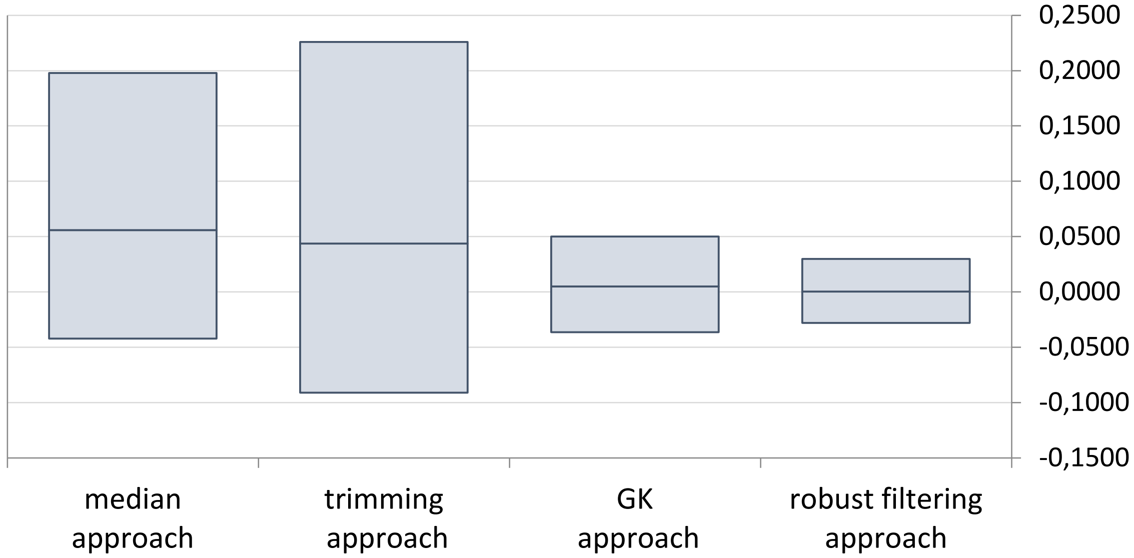
Source: The authors' work.

Finally, we show Figure 3 with boxplots of errors (simple difference between the estimates and the real parameters). The biggest boxplot size is reached with the trimming approach, however, we can see the median line is closer to zero than with the median approach. Nevertheless, the boxplot size for the median approach is somewhat smaller than in the case of the trimming approach.

The GK boxplot is much smaller in comparison, and its median is very close to zero. This suggests a smaller volatility of the estimations and a smaller bias in the sense of median, i.e. median-biased estimator (Brown, 1947).

The best results are given by the robust filtering approach. It gives even less volatile results than the GK approach. Furthermore, the median is almost zero, which would imply an unbiased estimator in the sense of median, i.e. median-unbiased.

Figure 3: Boxplots of 4 methods using AO model.



Source: The authors' work.

5.2 Innovative outlier model

We work with the innovative outlier model in this subsection. Again, we start with Table 3, which contains the percentages of the estimated processes being stationary. The same principles about simulations are applied here as in the case of the AO model.

Table 3: Stationarity of the estimated processes in IO model.

AR(p)	median approach	trimming approach	GK approach	robust filtering approach
AR(1)	97.8%	97.9%	100.0%	100.0%
AR(2)	96.3%	86.0%	99.6%	100.0%
AR(3)	89.4%	71.9%	97.7%	100.0%
AR(4)	81.8%	59.2%	94.0%	100.0%
AR(5)	76.4%	43.3%	89.7%	100.0%
Total	**93.0%**	**81.9%**	**98.4%**	**100.0%**

Source: The authors' work.

The results are in general a little bit worse, when we compare them with the AO model. This was expected because the IO model is designed to be more difficult to work with in the case of robust methods. Usually, the standard method is not affected by the IO model and robust methods are affected much more.

Again, the percentage of being stationary for an estimated process is decreasing with the growing order *p*, and the worst results are given by the trimming approach. Actually, for the case of the AR(5), the percentage is below 50%. In total it gives 81.9% success.

For the median approach, the percentage decreased to the level of 93%, which is much more than for the trimming approach. However, it is still relatively low in comparison with the two more sophisticated methods.

The GK approach, as well, is not as successful as it was previously. Nevertheless, the percentage is still quite high, except for the model AR(5), where every fifth simulation was evaluated as a non-stationary process. The robust filtering approach holds the full percentage.

Next, we show Table 4, where we see a comparison of the 4 above described methods.

Table 4: Comparison of 4 methods using IO model.

AR(*p*)	$\boldsymbol{\varphi_k}$	**median approach**		**trimming approach**		**GK approach**		**robust filtering approach**	
		MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE
AR(1)	φ_1	.0442	9.5%	.0602	12.8%	.0318	6.2%	.0214	4.8%
AR(2)	φ_1	.1566	28.7%	.1787	32.9%	.0381	7.8%	.0820	18.2%
	φ_2	.1502	40.1%	.1712	37.7%	.0396	8.9%	.0801	15.9%
AR(3)	φ_1	.3956	72.5%	1.8777	355.3%	.0963	18.2%	.0621	13.3%
	φ_2	.3812	95.5%	1.3219	236.1%	.0866	18.6%	.0742	18.3%
	φ_3	.4130	94.5%	1.9221	496.7%	.0924	22.4%	.0658	14.9%

Source: The authors' work.

Firstly, we realize that all the results are worse than those obtained using the AO model. This was already discussed above.

Secondly, we can see a similar order of the estimation accuracy. The trimming approach seems to be the worst method. Especially the results for the AR(3) are really bad. This is caused by the fact that there are many simulations which are evaluated with the absolute error greater than 2 (approximately 2.8% cases), and the biggest absolute error is higher than 3000. In these cases, we realize the absurdity of these results and we consider the situation not to be estimable with the trimming approach.

Thirdly, the median approach gives more reasonable values, however, in comparison with the two sophisticated approaches, it is still not good enough.

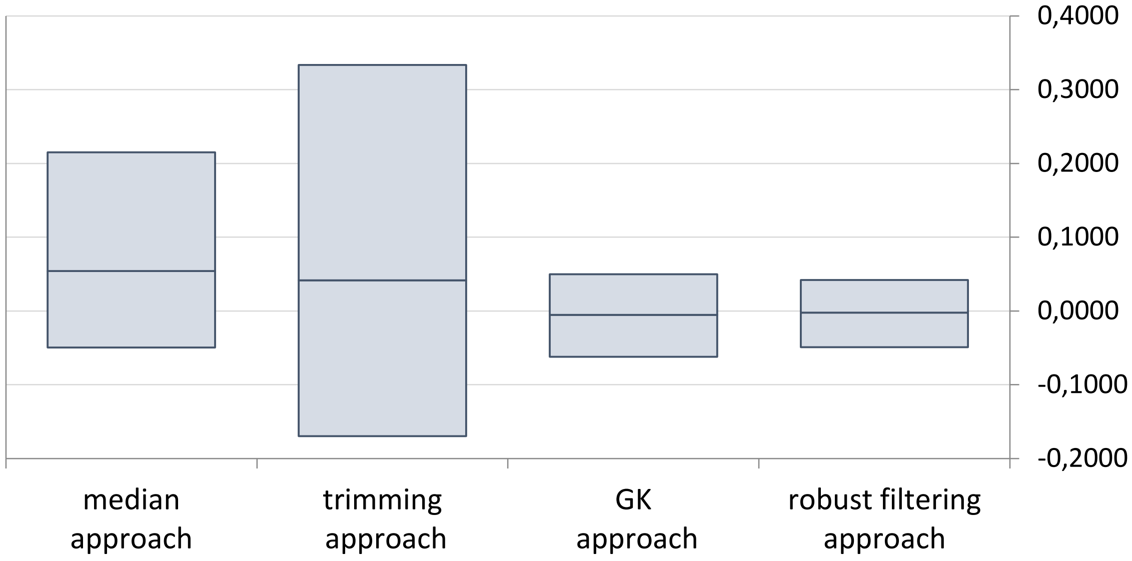
Finally, the robust filtering approach gives better results for the AR(1) and AR(3) models. Furthermore, the specific situation with the AR(2) model is present again. For the AR(1) and AR(2) models, the results are only slightly worse than with the AO model. Nevertheless, in the case of AR(3) model, the MAE is twice as big for the IO model as for the AO model.

At last, we show Figure 4 with boxplots of errors (simple difference between the estimates and the real parameters). We have to point out the different scale of the figure.

Naturally, we can see more volatile results in this figure. The trimming approach gives the biggest boxplot again. Similarly, the related median is closer to zero than for the median approach.

Significantly smaller boxplot sizes are given by the two more sophisticated robust methods, and, furthermore, their medians are almost equal to zero. The robust filtering approach looks a little less volatile than the GK approach, but we could see the same also in the case of the AO model. The comparison is very similar for both outliers model, nevertheless, the results are worse for the IO model.

Figure 4: Boxplots of 4 methods using IO model.



Source: The authors' work.

6. Conclusion

We introduced two models for outliers. The additive outlier model generates mostly isolated outliers. On the contrary, the innovative outlier model affects not only the current observation but also subsequent observations.

We introduced four robust methods for ACF estimation. Every robust method is based on a different idea, which we briefly described. When ACF robust estimators are available, we can, as well, robustly estimate parameters of the AR(p) model using the moment method.

We performed a simulation study in which we compared the four chosen methods. We noticed that the worst method is the trimming approach, which gave the worst results in almost all cases. For the IO model, there were several absurd estimations, what we consider as an inability to estimate certain simulations with the trimming approach.

Slightly better results were given by the median approach, however, the volatility of the estimations and the bias are still significantly higher than in the case of the two more sophisticated methods.

The GK approach gave quite good results, and for the case of the AR(2), it was the best approach. For other AR(p) models, it was a little worse than the robust filtering approach, but still very similar. For both outlier models, the estimator based on the GK approach was very close to the median-unbiased property.

Finally, the robust filtering approach showed the smallest volatility of the estimations and, as well, was the closest to zero in the sense of median, what could imply median-unbiased.

The MAE and MAPE criteria were the lowest for the robust filtering approach too. The only exception was the AR(2) model, for which the robust filtering approach gave suspiciously worse results. The reason for this exception is still unknown and it will be explored further. It is quite unusual that there exists such an exception because, for the AR(1) and AR(3) models, the robust filtering method gave significantly better results than the AR(2) model.

To conclude, based on the presented results of the performed simulation study, we consider the method based on robust filtering as the most appropriate to estimate parameters of the autoregressive processes, except for the special case of the AR(2) process. For the AR(2) case, we recommend to use the method based on the GK approach.

Acknowledgements

This study was supported by the grant F4/17/2017 (Robustnosť v úmyselnom useknutí časového radu), which has been provided by the Internal Grant Agency of the University of Economics in Prague.

References

[1] Brown, G. W. 1947. On Small-Sample Estimation. The Annals of Mathematical Statistics, vol. 18, iss. 4, pp. 582–585.

[2] Dürre, A., Fried, R., Liboschik, T. 2015. Robust estimation of (partial) autocorrelation. Wiley Interdisciplinary Reviews: Computational Statistics, vol. 7, iss. 3, pp. 205-222.

[3] Dürre, A et. al. 2016. robts: Robust Time Series Analysis. R package version 0.3.0/r220. http://R-Forge.R-project.org/projects/robts/

[4] Flimmel, S., Čamaj, M., Malá, I., Procházka, J. 2017. Comparison of Robust Methods for ARMA Order Estimation. In Applications of Mathematics and Statistics in Economics (AMSE 2017). Wroclaw: Publishing House of Wrocław University of Economics, 2017, s. 133--144. ISBN 978-83-7695-693-0.

[5] Fox, A.J. 1972. Outliers in time series. Journal of the Royal Society, vol. 34, iss. 3, pp. 350-363.

[6] Gnanadesikan, R., Kettenring, J. R. 1972. Robust estimates, residuals, and outlier detection with multiresponse data. Biometrics, vol. 28, pp. 81-124.

[7] Huber, P. J. 1981. Robust Statistics. New York: Wiley.

[8] Chakhchoukh, Y. 2010. A new robust estimation method for ARMA models. IEEE Transactions on Signal Processing, vol. 58, iss. 7, pp. 3512–3522.

[9] Chan, W.-S., Wei, W. W. 1992. A comparison of some estimators of time series autocorrelations. Computational statistics & data analysis, vol. 14, iss. 2, pp. 149–163.

[10] Ma, Y. and Genton M. 2000. Highly robust estimation of the autocovariance function, Journal of time series analysis, vol. 21, iss. 6, pp. 663–684.

[11] Maronna, R., Martin, D., Yohai, V. 2006. Robust Statistics: Theory and Methods. Chichester: John Wiley & Sons, Ltd. ISBN 978-0-470-01092-1.

[12] R Core Team 2013: R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, http://www.r-project.org/.

[13] Rousseeuw, P. J., Croux, C. 1992. Explicit scale estimators with high breakdown point. L1-Statistical Analyses and Related Methods, pp. 77-92.