CONSTRUCTION OF PREDICTION OF DYNAMIC RISK MEASURES VAR AND CVAR FOR FINANCIAL TIME SERIES WITH DIFFERENT VOLATILITY

Nataly Zrazhevska & Alex Zrazhevsky
UKRAINE

ABSTRACT

The article considers the method of constructing of predictive values of dynamic risk measures *VaR* and *CVaR* on the basis of an optimizational smoothing of autocorrelation function, proposed by the authors in the previous works. The method is based on the heteroscedastic time series model and is designed to predict the time series with long range dependence. For dispersion modeling the FIGARCH model is used, which is reduced to the AR model of the infinite order. In this paper, we study the feasibility and effectiveness of the proposed method for obtaining predicted values of dynamic risk measures VaR and CVaR for time series with different volatility. We consider tree time series of logarithmic return on a daily basis of the Nasdaq-100 index over the period 2005-2015: the original time series data, and two modified time series obtained from the original by deleting time periods with high volatility. So we have the opportunity to compare the forecasts built for data with similar statistical characteristics, but different volatility. The forecast values of risk measures are built in accordance with the procedure of direct multi-step prediction. Analysis of the predicted values is carried out using the Kupiec test, the Kristoffersen, the V-test and analysis of Probability of Exceedance values. The obtained predicted values and test results are shown in the figures and are displayed in tables. The analysis of the test results shows the effectiveness of this approach for obtaining risk measures VaR and CVaR prediction values for time series with long range dependence in a wide range of volatility. The proposed algorithm allowes to obtain the forecast that qualitatively repeats both the regular behavior and emissions of time series.

Keywords: Dynamic risk measures VaR and CVaR, long range dependence, heteroscedastic model, prediction.

INTRODUCTION

VaR and *CVaR* are the risk measures that are most commonly used by various financial institutions to analyze and predict the stock risks. This may explain the great interest of experts to their evaluating and predicting. In the previous studies [3, 4] a classification scheme for the choice of a method for dynamic risk measures *VaR* и *CVaR* estimating is proposed. On the basis of this scheme an algorithm for constructing predictive values of risk measures based on the optimization procedure of the autocorrelation function (ACF) smoothing is developed [5]. Verification of the proposed algorithm has been carried out by comparing the results of applying of the method to the real time series of index of the Tokyo Stock Exchange and the artificially generated time series with known characteristics.

The aim of this work is to demonstrate the capability and efficiency of the application of the developed technique to real time series with a wide range of volatility. As the data the time series of daily log return (2575 values) of the Nasdaq-100 Index (the NDX Index) for the period from 08.02.2005 to 11.02.2015 is considered. The index represents the dynamics of shares of the largest and most actively traded companies outside the financial sector, that are in the list of Exchange Nasdaq Stock Market. For comparison, besides the time series NDX_1 with high enough volatility, two another time series are used. The series NDX_2 (2275 values) is obtained from NDX_1 by removing the values for the time period of high volatility from 13.05.2008 to 22.07.2009, and the series NDX_3 (1975 values) obtained from NDX_2 by removing the values for the volatility time period from 22.02.2011 to 01.05.2012.

It should be emphasized that the algorithm is developed for time series with long range dependence. All considered time series, as it is shown below, satisfy this condition.

Key Definitions

Consider a non-stationary time series $\{X_t, t \in Z\}$ with finite mean defined on the probability space $(\Omega, \Psi_t, \mathrm{P}_t)$, where Ψ_t is the information set containing all available at the time t information about the time series. The series $\{X_t^2, t \in Z\}$ (TSV) is considered to be stationary.
The time series has the property of long range dependence [6] if there are $0 < \alpha < 1$ and $c_r > 0$ such that:

$$\lim_{k \to \infty} \rho(k)/(c_r k^{-\alpha}) = 1, \rho(k) = Corr(X_t, X_{t+k}), \ k \in N \cup \{0\}. \tag{1}$$

For a fixed confidence level α risk measures VaR and $CVaR$ are defined as:

$$VaR_\alpha^t(t+h) = \inf\left\{x \in R \middle| \mathrm{P}_t\left[X_{t+h} \le x\right] \ge \alpha\right\}, \ CVaR_\alpha^t(t+h) = E_{\Psi_t}[X_{t+h} \,|\, X_{t+h} \ge VaR_\alpha^t(t+h)],$$

where $E_{\Psi_t}[\cdot]$ denotes the conditional expectation with respect to Ψ_t [7].
As a model for the time series $\{X_t, t \in Z\}$ a stochastic process of the form:

$$X_t = \mu_t + \varepsilon_t = \mu_t + \sigma_t Z_t \tag{2}$$

is considered, where μ_t and σ_t^2 are the conditional expectation and variance, respectively, defined on the information space Ψ_t, $\{Z_t\} \overset{iid}{\sim} F_t(0,1)$.
Then the dynamic risk measures can be found by the formulas [8]:

$$VaR_\alpha^t = \mu_t + F^{-1}(\alpha)\sigma_t = \mu_t + VaR_\alpha(Z)\sigma_t, \ CVaR_\alpha^t = \mu_t + CVaR_\alpha(Z)\sigma_t, \tag{3}$$

and, accordingly, the predicted values for the P-steps ahead are defined as:

$$VaR_\alpha^{t+P} = \mu_{t+P} + VaR_\alpha(Z)\sigma_{t+P}, \ CVaR_\alpha^{t+P} = \mu_{t+P} + CVaR_\alpha(Z)\sigma_{t+P}, \tag{4}$$

where Z is a random variable with the same distribution as any random variable from $\{Z_t\}$. Hereinafter it is assumed that the trend, that defines μ_t, is missing (or it is removed from the data) [8].

METHODOLOGY

To build a predictive model a standard procedure is used. The sample is divided into two parts - In Sample and Out of Sample [6, 8]. In this study, the size of In Sample and Out of Sample is equal and the time interval In Sample is preceded by Out of Sample. In Sample is used for the series research and for the model constructing. After that the P-steps ahead prediction is built and the results are compared with the real values obtained on Out of Sample. The considered procedure assumes the construction of only short-term predictions, as it is based on an autoregressive time series model. In this paper we construct a 5-steps ahead prediction. To confirm the stability of the result the procedure (the window with the accumulation) is repeated many times. To build a prediction of risk measures for NDX_1 256, for NDX_2 226 and for NDX_3 196 windows are considered. To obtain the final prediction all 5-steps ahead predictions are combined into one. The confidence level for all risk measures is $\alpha = 0.9$.

In accordance with the proposed methodology in the first stage the standard statistical analysis of the time series and their squares is performed. The analysis confirms that the considered series have different volatility: $\sigma_{NDX_1} = 0.43$, $\sigma_{NDX_2} = 0.32$, $\sigma_{NDX_3} = 0.25$. The values of the asymmetry coefficient (Skewness) and kurtosis (Kurtosis) for all-time series indicate that PDF have fat tails.

In the second stage TSV are analyzed on the property of long range dependence. To estimate the Hurst parameter 5 standard methods are used - the method of absolute values of the aggregated series, the aggregated variance method, the method of residuals of regression, the periodogram method and the R/S method [1]. Table 1 shows minimum $H_{\min}$, maximum $H_{\min}$ and average H_{mean} values for the windows obtained by each method. The results indicate that different methods determine the value of the parameter with a significant scatter. However, all methods confirm that all series have the property of long range dependence. The scatter in the values of the Hurst parameter makes inefficient the standard modeling methods [9].

In the third stage, in view of the long range dependence of the TSV, for modeling and predicting the model $FIGARCH$ is used, which is reduced to the model AR (∞) [10, 2]. Autoregression coefficients $(a_1, ..., a_N, ...)'$ are determined by the least squares method that leads to the necessity of solving the infinite system of Yule-Walker equations:

$$\sum_{j=0}^{\infty} \rho_{|i-j|} a_j = \rho_{i+1},\ i = 0, ..., \infty . \tag{5}$$

To estimate ρ_i the regression equation for the ACF is used. The equation is based on the definition of the long range dependence (1): $\rho(k) = \alpha_1 H(2H-1)k^{2H-2} + \alpha_2 + \varepsilon_k$, $\varepsilon_k - iid$, $k_0 \le k \le N$. The optimization procedure [1] allows to specify the estimate $\hat{H}_{optim}$ (Table 1) and to correct the estimates $\widehat{\rho}(k)$.

The values $\hat{H}_{optim}$ show that the use of the optimization procedure gives much more stable value of the parameter. Thus, one of the main objectives of the developed technique - model is more stable - is achieved. So it is able to get more stable prediction.

Table 1. Hurst parameter values

Method/ $\widehat{H}$	NDX_1			NDX_2			NDX_3		
	$H_{\min}$	$H_{\max}$	H_{mean}	$H_{\min}$	$H_{\max}$	H_{mean}	$H_{\min}$	$H_{\max}$	H_{mean}
Abs.values of the aggregated series	0.924	0.948	0.938	0,844	0.774	0.875	0.799	0.766	0.833
Aggregated variance	0.905	0.880	0.931	0.759	0.601	0.784	0.771	0.726	0.792
Residuals of regresion	0.937	0.827	1.008	0.741	0.538	0.802	0.701	0.648	0.758
periodogram	1.038	1.012	1.068	0.800	0.723	0.885	0.854	0.801	0.915
R/S	0.831	0.782	0.863	0.781	0.729	0.825	0.742	0.719	0.769
optimization	0,775	0,772	0,779	0,763	0,726	0,783	0,732	0,726	0,755

The reduced system of normal equations (5) with the coefficients $\widehat{\rho}(k)$ is solved . The Akaike information criterion [10] is used to determine the lag M of the reduced autoregression model. As it is shown in [11], the solution of the reduced system converges to the exact solution. The variance ratio test [12] confirms the correctness of the constructed model for getting estimates σ_t. The built model is used to obtain the predictive values of variations:

$$\hat{\sigma}_{l+p}^2 = \sum_{i=p}^{M+p} \hat{a}_i \hat{\sigma}_{l-i+1}^2 ,\ l = N+1,\ldots,\ p = 1,\ldots,P \qquad (6)$$

In the fourth stage the obtained model estimates σ_t are used to find realizations of a random variable Z_t: $Z_t = X_t / \sigma_t$. The variance ratio test shows that they are iid. The Jarque-Bera test [8] indicates that the PDF of the model residuals have relatively thick tails. In accordance with the classification scheme presented in [3], the following methods for finding the estimates $VaR_{0.9}(Z)$, $CVaR_{0.9}(Z)$ are chosen: the historical simulation method (estimate $hist1$), the use of explicit formulas under the assumption of a normal distribution with the maximum likelihood estimates of the parameters (estimate $pd1$), the use of explicit formulas using the GEV function with the maximum likelihood estimates of the parameters (estimate $GEVq$), the Monte Carlo method for the obtained GPD function (estimate $GPDmc$), the empirical POT method (estimate $POTem$). Description of the methods can be found, for example, in [3].

Minimum, maximum and average values for the windows for $VaR_{0.9}(Z)$ are presented in Table 2 and for $CVaR_{0.9}(Z)$ in Table 3.

Table 2. Values of $VaR_{0.9}(Z)$

Method/ $VaR_{0.9}(Z)$	NDX_1			NDX_2			NDX_3		
	min	max	mean	min	max	mean	min	max	mean
*hist*1	1.657	1.627	1.743	1.603	1.541	1.754	1.546	1.522	1.613
*pd*1	1.658	1.634	1.709	1.623	1.582	1.740	1.580	1.564	1.615
GEVq	1.515	1.484	1.563	1.479	1.410	1.570	1.445	1.398	1.492
GPDmc	1.570	1.543	1.616	1.545	1.410	1.647	1.497	1.475	1.552
POTem	2.077	1.807	2.370	1.916	1.524	2.187	1.752	1.443	2.076

Table 3. Values of $CVaR_{0.9}(Z)$

Method/ $CVaR_{0.9}(Z)$	NDX_1			NDX_2			NDX_3		
	min	max	mean	min	max	mean	min	max	mean
*hist*1	2.512	2.465	2.603	2.462	2.408	2.666	2.414	2.366	2.508
*pd*1	2.267	2.238	2.233	2.217	2.162	2.373	2.381	2.138	2.205
GEVq	2.503	2.447	2.611	1.479	1.410	1.570	2.381	2.318	2.470
GPDmc	2.523	2.486	2.613	1.545	1.410	1.647	2.419	2.386	2.483
POTem	2.942	2.634	3.246	1.916	1.524	2.187	2.621	2.280	3.000

Using the obtained values of $VaR_{0.9}(Z)$ and $CVaR_{0.9}(Z)$ and (4) predictive estimates of $VaR_{0.9}^{t+p}$ and $CVaR_{0.9}^{t+p}$, $p = 1,\dots,5$ are built.

Predicted values are compared with the estimates of risk measures obtained by the respective methods using real data from Out of Sample on the forecast horizon. Error characteristics are given in Table 4 (for $VaR_{0.9}^{t+p}$) and in Table 5 (for $CVaR_{0.9}^{t+p}$).

Table 4. Results of analysis of prediction errors for $VaR_{0.9}^{t+p}$

Method/ quantity	NDX_1			NDX_2			NDX_3		
	ME $\times 10^{-3}$	MAE $\times 10^{-3}$	MSE $\times 10^{-5}$	ME $\times 10^{-3}$	MAE $\times 10^{-3}$	MSE $\times 10^{-5}$	ME $\times 10^{-3}$	MAE $\times 10^{-3}$	MSE $\times 10^{-5}$
*hist*1	-1.6	6.0	6	0.4	5.1	5	-0.3	3.6	2
*pd*1	-3.0	6.7	7	0.1	5.2	5	-0.5	4.0	3
GEVq	-3.1	6.3	6	-0.2	4.8	4	-0.8	3.8	2
GPDmc	-1.7	5.7	5	-0.3	4.9	5	-0.5	3.8	3
POTem	0.2	6.9	9	2.4	5.8	8	1.9	4.5	4

Table 5. Results of analysis of prediction errors for $CVaR_{0.9}^{t+p}$

Метод/ Оценка	NDX_1			NDX_2			NDX_3		
	ME $\times10^{-3}$	MAE $\times10^{-3}$	MSE $\times10^{-5}$	ME $\times10^{-3}$	MAE $\times10^{-3}$	MSE $\times10^{-5}$	ME $\times10^{-3}$	MAE $\times10^{-3}$	MSE $\times10^{-5}$
$hist1$	-4.5	10.1	16	0.3	7.9	13	-0.8	6.3	6.7
$pd1$	-4.3	9.2	13	0.1	7.1	10	-0.8	5.6	5.2
$GEVq$	-3.6	9.7	15	0.9	7.8	13	-0.06	5.9	6.3
$GPDmc$	-4.3	10.0	16	0.6	7.9	13	-0.4	6.1	6.6
$POTem$	-2.8	10.9	20	2.5	8.4	15	1.4	6.6	9.0

As can be seen, the constructed forecasts are stable and reliable. Forecast quality tests show the correctness of assessments built by all methods, with some advantage of $GPDmc$ for NDX_1, $GEVq$ for NDX_2 and $hist1$ for NDX_3. This can be explained by the different volatility as well as the features of the distribution functions.

Figure 1 shows the prediction of risk measures for NDX_1, where the estimates $GPDmc$ are used. Figure 2 shows the prediction of risk measures for NDX_2 with estimates $GEVq$ and Figure 3 demonstrates the prediction of risk measures for NDX_3 with estimates $hist1$.

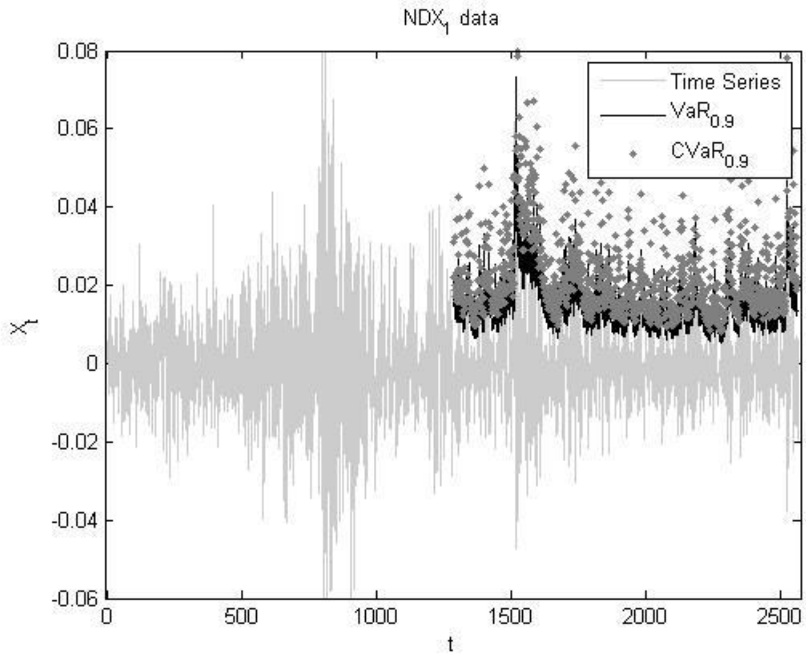

Fig 1. Predictive estimates of risk measures for NDX_1

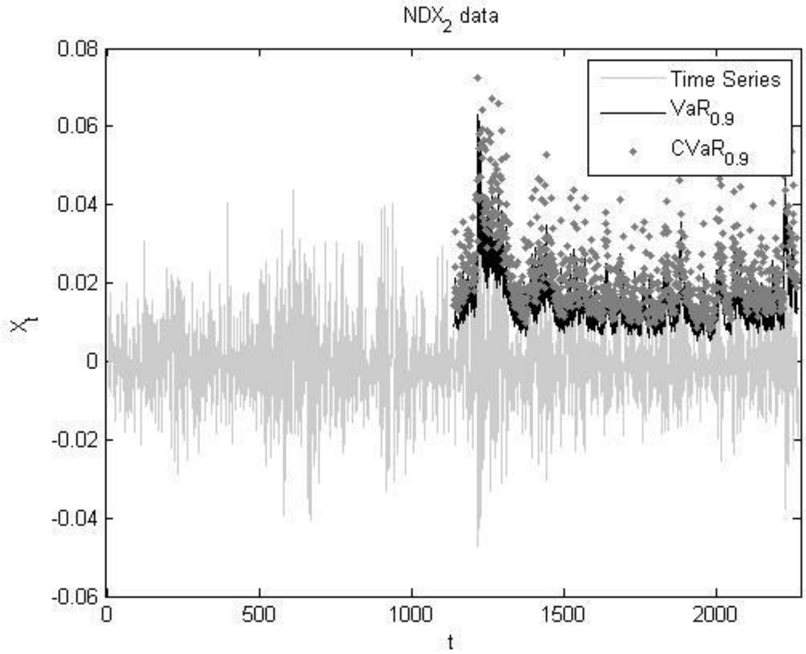

Fig 2. Predictive estimates of risk measures for NDX_2

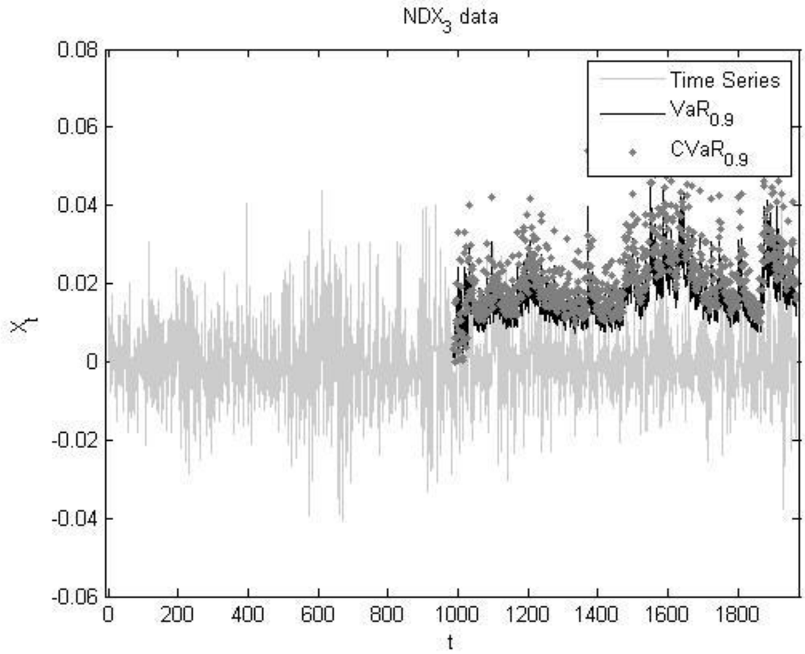

Fig 2. Predictive estimates of risk measures for NDX_3

Analysis of the results leads to the conclusion that the values of risk measures qualitatively monitor the dynamics of the time series, the risk jumps repeat measures emissions with minimal delay. $CVaR$, as an integral characteristic, smoothes VaR jumps. It should be noted that VaR is not a convex function [8], and, as it is seen from Table 2, the increase of data volatility does not automatically result its growth. At the same time, $CVaR$ is a convex function and its values grow when data volatility increases.

Values of Probability of Exceedance (POE) [13] are obtained to analyze the predicted values of dynamic VaR (Table. 6).

Table. 6. Comparison of PoE for real and predicted values of $VaR_{0.9}^{t}$

Method/ Statistics	NDX_1		NDX_2		NDX_3	
	real	*pred*	*real*	*pred*	*real*	*pred*
*hist*1	0.061	0.091	0.087	0.096	0.079	0.094
*pd*1	0.051	0.090	0.080	0.093	0.071	0.090
GEVq	0.052	0.104	0.087	0.105	0.082	0.103
GPDmc	0.059	0.099	0.089	0.101	0.080	0.100
POTem	0.066	0.063	0.078	0.072	0.079	0.081

The values *real* are obtained by the model using real data. The proximity of the obtained values to the value $1-\alpha=0.1$ demonstrates the possibility to use the considered methods for the determination of the dynamic VaR .As expected, the worst result is obtained for NDX_1 with high volatility. The values *pred* are obtained using the predicted values and describe the dynamic behavior of VaR more adequately. The method *POTem* demonstrates the worst result that corresponds to the results shown in Table 4.

To analyze the obtained predicted values of VaR the quality tests are used. Table 7 shows the p-values of the statistics for the unconditional Kupiec test (*LRpof*), Kristoffersen test for independence (*LRind*) and the combined statistics *LRcc* [14].

Table. 7. Results of the quality tests for the predicted values of $VaR_{0.9}^{t}$

Method/p-value	NDX_1			NDX_2			NDX_3		
	LRpof	*LRind*	*LRcc*	*LRpof*	*LRind*	*LRcc*	*LRpof*	*LRind*	*LRcc*
*hist*1	0.257	0.013	0.054	0.618	0.047	0.046	0.519	0.057	0.046
*pd*1	0.119	0.035	0.045	0.423	0.024	0.056	0.279	0.032	0.055
GEVq	0.643	0.051	0.071	0.555	0.061	0.144	0.750	0.054	0.053
GPDmc	0.126	0.080	0.060	0.921	0.0159	0.062	1.000	0.066	0.062
POTem	0.001	0.020	0.000	0.001	0.395	0.003	0.037	0.031	0.011

Test result is positive if p-value is greater than the selected Value of Reliability equal 0.05. As can be seen, the method *POTem* demonstrates the lowest quality. The estimates obtained with the method *pd*1 also have not enough high quality that can be explained by the fat tails of the PDF.

V-test with statistics V_1, V_2, V [7,4] is used to analyze the quality of the predicted values of $CVaR_{0.9}^{t}$ (Table 8). Estimates have good quality, if the statistics are close to 0. The results in the table show the methods are chosen correctly and forecasts have high quality.

Table. 8. Results of the V-test for the predicted values of $CVaR_{0.9}^{t}$

Method/ statistics	NDX_1			NDX_2			NDX_3		
	V_1 $\times 10^{-3}$	V_2 $\times 10^{-3}$	V $\times 10^{-3}$	V_1 $\times 10^{-3}$	V_2 $\times 10^{-3}$	V $\times 10^{-3}$	V_1 $\times 10^{-3}$	V_2 $\times 10^{-3}$	V $\times 10^{-3}$
$hist1$	0.04	-18.18	9.11	-0.24	-17.92	9.08	0.15	-16.61	8.38
$pd1$	1.90	-15.19	8.54	1.85	-14.82	8.34	2.31	-13.72	8.02
$GEVq$	-0.95	-18.10	9.53	-0.77	-17.58	9.17	-0.28	-16.34	8.31
$GPDmc$	-8.00	-18.30	9.56	-0.75	-18.11	9.43	-0.31	-16.67	8.41
$POTem$	-0.51	-22.83	11.68	-0.52	-21.48	10.99	-0.20	-18.66	9.43

CONCLUSIONS

Comprehensive approach consisting of the classification scheme for the choice of method for dynamic risk measures VaR and $CVaR$ estimating, the algorithm and the new method of forecasting of risk measures for time series with long range dependence is proposed by the authors in studies [1-5]. In this article the approach is used to obtain the predicted values of dynamic risk measures for real data describing the Nasdaq-100 Index (NDX index) for the period from 08.02.2005 to 11.02.2015. To demonstrate the efficacy of this approach for time series with a wide range of volatility forecasts are constructed for three time series: the original data NDX and two modified time series obtained from the original by deleting time periods with high volatility. This allowes to compare forecasts built for data with similar statistical characteristics, but different volatility.

Dynamic risk measures forecasts are constructed in accordance with the procedure of direct multi-step prediction, based on the model $FIGARCH$. Constructing the model the new method [1-5], based on the optimizational smoothing of the ACF and reduction of Yule-Walker equations, is used. To check the quality of the forecasts standard statistical tests and specific tests, designed to test the dynamic risk measures, are used.

Obtained predicted values and test results are shown in the figures and are displayed in tables. The analysis of the results confirms the effectiveness of the proposed approach for dynamic risk measures prediction. The classification scheme helps to determine the best method of forecasting for each time series.

Application of the method of the optimizational smoothing of the ACF leads to the construction of sustainable forecasting of dynamic variations, that qualitatively repeats both the regular behavior and emissions of time series. The quality of forecasts is practically the same for all-time series regardless of their volatility that is confirmed by quality tests. Thus, the effectiveness of the developed approach for the forecasting of dynamic risk measures for time series with long range dependence in a wide range of volatility is confirmed.

REFERENCES

1. Pankratova, N.D. and Zrazhevska, N.G. (2015) Model of the autocorrelation function of the time series with strong dependence, *Problemi upravleniya i informatiki*, 5, 102-112.
2. Zrazhevska, N.G. (2015) Smoothed autocorrelation function method for predicting of variation of heteroscedastic time. *Systemni doslidzhennya ta informacijni texnologiyi*, 2, 97–108. (In Russian)
3. Zrazhevskaja, N.G., Zrazhevskij, A.G. (2016) Classification of methods for risk measures VaR and CVaR calculation and estimation. *Systemni doslidzhennya ta informacijni texnologiyi*, 3.
4. Zrazhevska, N. (2016) Construction and application of the classification scheme of dynamic risk measures estimating. *Eureka: Physics and Engineering*, 5, 67–80.
5. Pankratova N.D., Zrazhevska N.G. (2017) Method of dynamic VaR and CVaR risk measures forecasting for long range dependent time series on the base of the heteroscedastic model. *Intelligent Control and Automation (ICA*), (in press).
6. Palma, W. (2007) *Long-Memory Time Series: Theory and Methods.* NewJersey: John Wiley& Sons, Inc.
7. Embrechts, P., Kaufmann, R. & Patie, P. (2005) Strategic long-term financial risks: single risk factors. *Computational Optimization and Applications*, 32, 61–90.
8. Tsay, R.S. (2010) *Analysis of Financial Time Series.* Hoboken: John Wiley & Sons, Inc.
9. Beran, J. (1994) *Statistics for Long-Memory Processes.* New York: Chapman and Hall.
10. Zivot, E., Wang, J. (2003) *Modeling Financial Time Series with S-PLUS.* New York: Springer-Verlag.
11. Zrazhevskij, A. G. (2011) System approach to restoring functional dependencies of non-stationary time series with different structures. *Abstract of dissertation for scientific degree of Ph.D*, 20. (In Ukrainian)
12. Cochrane, J.H. (1988) How big is the random walk in GNP? *Journal of Political Economy*, 96, 893–920.
13. Shang, D., Kuzmenko, V., & Uryasev. S. (2015) Cash Flow Matching With Risks Controlled by bPOE and CVaR., *ISE Dept., University of Florida*, *Research Report* 2015-3.
14. Karadzic, J., Cerovic, J. (2014) *Market risk of the Western Balkans countries during the global financial crisis*. Economic Annals-XXI.