Journal of Siberian Federal University. Engineering & Technologies 1 (2012 5) 111-121
~ ~ ~

УДК 336.763

Extreme Value Theory and Peaks Over Threshold Model in the Russian Stock Market

Vladimir O. Andreev[a]*, Sergey E. Tinykov[b], Oksana P. Ovchinnikova[a] and Gennady P. Parahin[c]
[a] Oryol Regional Academy of State Service
5a Pobedy st., Oryol, 302028 Russia
[b] Branch of Siberian Federal University, Zheleznogorsk
12a Kirov st., Zheleznogorsk, Russia
[c] TelekomStroyService, Ltd.
61 Tsarev Brod st., Zarechensky, 302528 Russia [1]

Received 6.02.2012, received in revised form 13.02.2012, accepted 20.02.2012

Traditional research methods adopts normal distributions as a pattern of the stock market behavior. This paper utilized POT model of extreme value theory, and GPD distribution which can give more accurate description on tail distribution of financial returns/losses. EVT and POT techniques are applied to a series of daily losses of the RTS index (RTSI) over a 15-year period (1995-2009), RTSI is total index of 50 largest Russian stocks. The focus is on the use of proposed methods to asses tail related risk providing a modeling tool for modern risk management.

Keywords: Extreme Value Theory, General Pareto Distribution, Peaks over Threshold, Tail Distribution, Value at Risk.

Introduction

The study of extreme events has attracted the special attention in connection with the global crisis of 2008–2009. The Russian stock market has been dramatic volatile over 15-year period (from 38 points on 05.10.1998 to around 2487 points on 19.05.2008 and back to about 498 points on 23.01.2009). In irregular financial market, it is necessary to set up models and systems to evaluate and control risks. In this paper we focus on the extreme behavior of financial series, unraveling the volatilities in the financial markets has always been an decipherable mystery. One of the purposes of this chapter is to test the validity of a popular risk management instrument: Value-at-Risk estimator in Russian equity market, which is a widely adopted technique in the developed countries for quantifying market risk. We have to deal with extreme events when a risk takes values from the tails of its probability distribution. In the field of market risk management it is a great concern the day by day determination of the Value-at-Risk (VaR) [1]. VaR is a high

* Corresponding author E-mail address: oragsob@mail.ru

quantile of the distribution of losses (for example the 95th percentile): $VaR_p=F^{-1}(p)$, where F is the loss cumulative distribution function and p the selected probability level. Traditional procedure calculating VaR based on normal distribution has limitations. VaR model reflects that it is to asses the possible maximum loss under regular market environments. Risk managers have become more concerned with events occurring under extreme market conditions [2,3]. This paper argues that extreme value theory (EVT) and POT (Peaks Over Threshold) model provide tools for estimating measures of tail risk under irregular volatility in market. We consider a fully parametric model, based on the GPD (Generalized Pareto Distribution), which can be easily estimated by maximum likelihood method [4,5].

I. Theoretical framework of the extreme value approach

Extreme value theory is a powerful and fairly robust framework to study the tail behavior of a distribution. There have been a number of extreme value studies in the finance literature in recent years: quantile estimation using the extreme value theory [6]; the estimation of the tails of loss severity distributions and the estimation of the quantile risk measures for financial time series using extreme value theory [7,8]; overview the extreme value theory as a risk management tool [9]; potentials and limitations of the extreme value theory [10,11]; an extensive overview of the extreme value theory for risk managers [12]; the estimation of tail-related risk measures for heteroskedastic financial time series [13]; comprehensive source of the extreme value theory to the finance and insurance literature [14,15].

POT model and Generalized Pareto distribution

We use of Extreme Value Theory to model the tail returns and then show how our EVT estimates are incorporated into the risk measures. Two main approaches are proposed in the literature [16]: the Block Maxima (BM) and the Peaks-over-Threshold models (POT). The group of models for threshold exceedances are more modern and powerful than the BM models [16], we focus on this approach and its application to the losses on the RTSI stock index. We apply the parametric POT method based on the Generalized Pareto distribution (GPD) to describe tail behaviour. Begin by assuming that market losses represent the realizations x of a random variable X over an enough high threshold u. More particularly, if X has the cumulative distribution function $F(x)$, we are interested in the cumulative distribution function $F_u(x)$ of exceedances of X over a high threshold u, i.e. the conditional excess distribution function is defined as:

$$F_u(x) = P(X - u \le x \mid X > u) = \tfrac{F(x+u) - F(u)}{1 - F(u)}. \qquad (1)$$

As to the sufficient large u, EVT provides us with a powerful key result, which states for a large class of underlying distributions $F(x)$ [2]:

$$F_u(x) \approx G_{\xi,\beta}(x),\ u \to \infty,$$

where Generalized Pareto Distribution is defined by:

$$G_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \xi x / \beta)^{-1/\xi}, if\ \xi \ne 0 \\ 1 - \exp(-x/\beta), if\ \xi = 0 \end{cases} \qquad (2)$$

$$where\ x \in \begin{cases} [0,\infty),\ \text{если}\ \xi \geq 0 \\ [0,\text{-}\beta/\xi),\ \text{если}\ \xi < 0 \end{cases} .$$

GPD subsumes three other distributions under its parameterization [2]. So, when tail index $\xi=0$, we obtain a Type 1 (exponentially declining) distribution. If $\xi<0$, we have a Type 2 (power declining). For $\xi>0$, we obtain a Type 3 (constant declining) distribution. Given these three types of distribution, one of our tasks in this paper will be to uncover which type best describes the extremes of stock returns on the emerging Russian market.

Identification of GPD parameters

Let $(X_1,X_2,...,X_{k(u)})$ be a sequence of iid random variables from an unknown distribution F, $X_i>u$. Shape ξ and scale $\beta>0$ parameters are then defined on the threshold u [16]. These GPD parameters can be determined by maximum likelihood (ML) methods. The log likelihood function of the GPD for $\xi\neq0$ is:

$$L(\xi,\beta) = -k(u)(\log\beta) - (1+1/\xi)\sum_{i=1}^{n}\log(1+\xi x_i/\beta), \tag{3}$$

where x_i satisfies the constraints specified for x_i. If $\xi=0$, the log likelihood function is:

$$L(\beta) = -k(u)\ln(\beta) - \beta^{-1}\sum_{i=1}^{n} x_i\,. \tag{4}$$

ML estimates are then found by maximizing the log-likelihood function using numeral optimization methods. We can get these ξ and β estimates through solving simultaneous equations (ξ≠0):

$$\frac{\partial L(\xi,\beta)}{\partial\xi} = \frac{1}{\xi^2}\sum_{i=1}^{k(u)}\log(1+\xi\frac{x_i}{\beta}) - (\frac{1}{\xi}+1)\sum_{i=1}^{k(u)}\frac{x_i}{\beta+\xi x_i} = 0$$
$$\frac{\partial L(\xi,\beta)}{\partial\beta} = -\frac{k(u)}{\beta} + (\frac{1}{\xi}+1)\sum_{i=1}^{k(u)}\frac{\xi x_i}{\beta^2+\beta\xi x_i} = 0$$

Tail evaluation formula

Assuming that u is sufficiently high, by combining expressions (2) and (3) the distribution function $F(x)$ for exceedances can be written as:

$$\begin{aligned} F(x) &= (1-F(u))F_u(x) + F(u) \\ &= 1 - \tfrac{k(u)}{n}(1+\tfrac{\xi}{\beta}(x-u)^{-1/\xi}) \quad . \end{aligned} \tag{5}$$

$F_u(x)$ is a GPD and $F(u)$ is given by $[n\text{-}k(u)]/n$; n is the total number of observations, $k(u)$ the number of observations above the threshold u, ξ and β are the parameters of the GPD.

Estimating VaR

For a given probability $p>F(u)$ and threshold u, the value-at-risk (VaR) is calculated by inverting the tail estimation formula (5):

$$VaR_p = u + \tfrac{\beta}{\xi}\{[\tfrac{n}{k(u)}(1-p)]^{-\xi} - 1\}\cdot \tag{6}$$

Choosing threshold value

Choice of the threshold u is the important issue to deal with: u too high results in too few exceedances and consequently high variance estimators. On the other hand, u too small provides biased estimators and the approximation to a GPD could not be feasible. It is possible to choose an asymptotically optimal threshold by a quantification of a bias versus variance trade-off.

Mean excess function

One suggestion which is of immediate use in practice is based on the linearity of the mean excess function $e(u)$ for the GPD. From [2] we know that for a random value X with a GPD distribution function $G_{\xi,\beta}$ the mean excess function is:

$$\begin{aligned} & e(u) = E(X-u \mid X>u) = \tfrac{\beta+\xi u}{1-\xi}, \\ & \beta + u\xi > 0, \xi < 1. \end{aligned} \tag{7}$$

It suggests a graphical approach for choosing u: choose u >0 such that $e(x)$ is approximately linear for $x¸u$. Using plots to compare resulting estimates across a variety of u-values, due to the usual presence of multiple choice of the threshold, is recommended.

Hill plot

Let $X_1>X_2>\ldots>X_n$ be the order statistics of positive random variables iid. The Hill estimator of the tail index ξ using k+1 order statistics is defined by [17]:

$$\hat{\xi} = \tfrac{1}{k}\sum\nolimits_{j=1}^{k}(\ln X_{j,n} - X_{k,n}). \tag{8}$$

Hill plot is a good instrument to find the optimal threshold [18]. Over a specific range of exceedances, the Hill plot may be under the stationary series, and the turning point is a good choice of optimal threshold. We use the following intuitive ideas:

(1) The sequence of the turning point is less than ~n/10 [19].

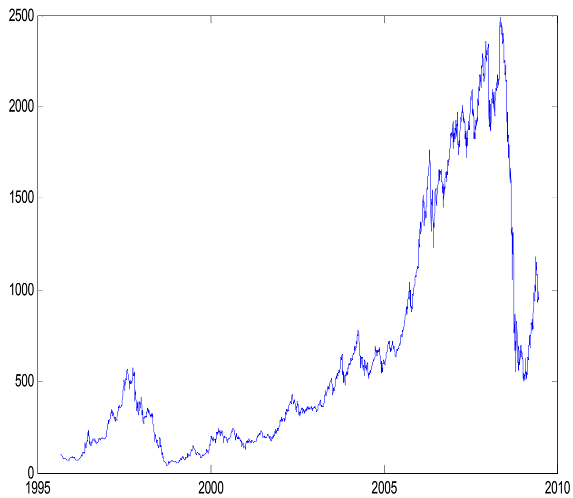
(2) The Hill estimator in the turning point has a relative large deviation from the fitted stationary straight line.

(3) The turning point is the last sequence of point that satisfies the two conditions stated above.

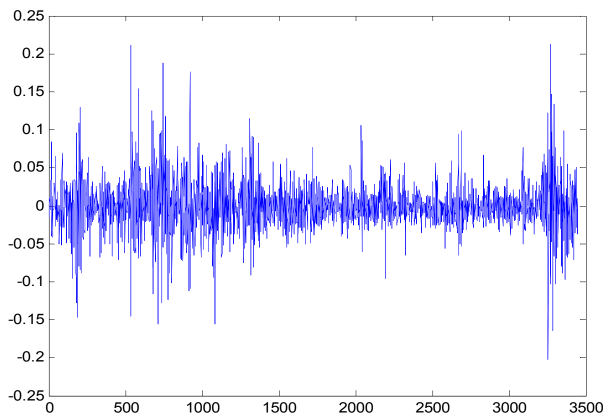
Empirical results

We consider a extreme value approach, working on the series of daily log losses (negative returns) of the Russian RTSI Index over a period of fifteen years (1995-2009). The Russian Trading System Index (RTSI) comprises of 50 of the largest stocks capturing 85% of the total market capitalization of the Russian Trading System exchange. The data used in this paper are obtained from RTS web site [21]. The empiric study uses the series of log daily losses of the RTSI Index, containing 3 447 trading days (closing prices). Fig.1 shows the plot of daily dynamics of RTSI index values, and log daily losses.

Table 1 shows the summary statistics for the series of log daily changes. This table shows that kurtosis value is 9.7024 and skewness value is 0.3752. Relative value of Normal distribution is 3 and 0, respectively. So we can see empirical distribution of log daily losses and normal distribution is not compatible.



a) daily dynamics of index values



b) log daily losses

Fig. 1. RTSI Index – sample period 01.09.1995 – 30.06.2009 (closing values)

In addition to this, Jarqua-Bera statistic shows that law of log daily losses is obviously different from normal distribution. The JB test statistics is defined as [10]:

$$JB = \frac{n}{6}\left[\frac{STD^2}{6} + \frac{(Kurtosis-3)^2}{24}\right].$$

The JB statistic has approximately a chi-squared distribution, with two degrees of freedom. The Jarqua-Bera test depends on skewness and kurtosis statistics. If the JB test statistic equals zero, it means that the distribution has zero skewness and kurtosis is about equal 3, and so it can be concluded that the normality assumption holds. Skewness values far from zero and kurtosis values far from 3 lead to an increase in JB values. The test returns the logical value $h = 1$ if it rejects the null hypothesis at the $p<0.05$ significance level, and h =0 if it cannot. We have for data of Table 1: *JB value=6532.8*, $p\sim0$, h=1. It means that we can reject the hypothesis that the distribution of daily losses is normal.

Table 1. Summary statistics for daily losses in RTS

mean	*min*	*max*
-0.0007	-0.2020	0.2120
std	*skewness*	*kurtosis*
0.0289	0.3752	9.7024
variance	*JB test*	*n*
-0.0008	6532.8 *h=1,p*<0.001	3447

In Fig.2 we represent the Hill graph, which plots the Hill estimator of ξ, versus the k upper order statistics (and threshold u, respectively). We select the last area to k~0.1*3447~350, where the Hill estimator is more stable.

The mean excess function (7) allows to establish the behavior of the distribution tails [23]: we choose threshold u looking at the linear shape (with positive slope) of the graph (Fig.3). Considering Hill plot and the mean excess function, we choose u=0.0334 (the number of observation exceeding threshold u is equal k=294).

The results of ML estimation of the GPD parameters (on chosen threshold u=0.0334) are ξ=0.1492 and β = 0.0206:

```
Maximum Likelihood (ML) estimates of ξ,β:
out =
 par_ests: [0.1492 0.0206]
 funval: -803.6979
 par_ses: [0.0688 0.0018]
 threshold: 0.0334
 data: [1x294 double]
 p_less_thresh: 0.9675
```

QQ-plot graph makes us able to evaluate the goodness of fit of the empirical series to a parametric GPD model (Fig.4) [24]. Notice that a concave departure from the straight line in the QQ-plot (Fig.4a) is an indication of heavy tailed distribution, whereas a convex departure is an indication of a thin tail.

After we get estimates $\xi,\beta,$ use them in (5), get the formula for of tail evaluation:

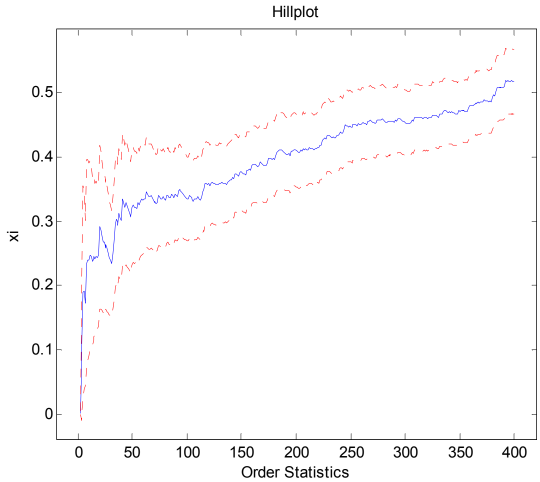
$$p = F(x) = 1 - \frac{k(u)}{n}\left(1 + 0.1492 * \frac{x - 0.0334}{0.0206}\right)^{-\frac{1}{0.1492}}$$

$k(u) = 294$, n=3447, k(u)/n=8.53%

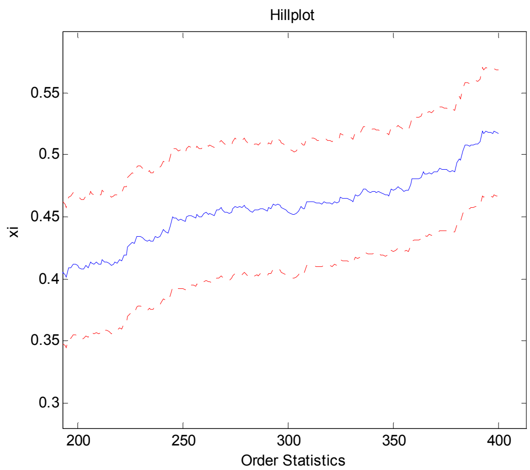
Employ the result in (6), get the VaR formula on GPD model:

$$VaR_p = 0.0334 + \frac{0.0206}{0.1492}\{[\frac{3447}{294}(1-p)]^{-0.1492} - 1\}.$$

In Table 2 we report 95%, 99%, 99.5%, 99,9% Value-at-Risk estimates of three different VaR estimation methods. The performance of the different VaR estimation methods can be evaluated by comparing the estimates with the actual losses observed, in particular by computing (and testing) the number of VaR violations. VaR approaches based on the assumption of normal distribution are



a) Hill estimator versus k upper order statistics (k≤350)



b) Hill estimator versus k upper order statistics (k≤350) (plot zooming)

Fig. 2. Hill estimator versus k upper order statistics (probability level p=0.95)

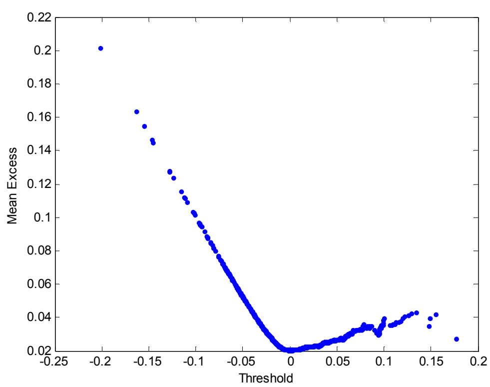
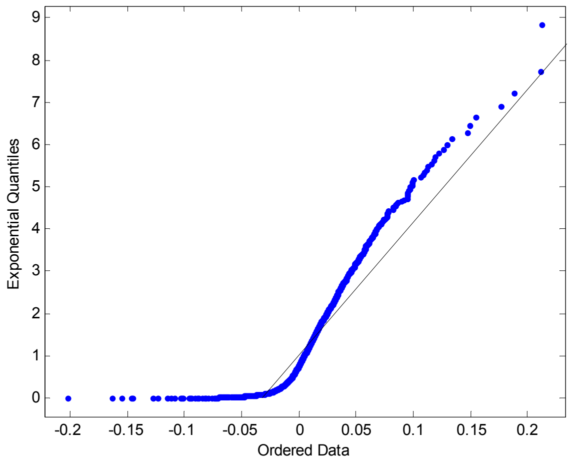
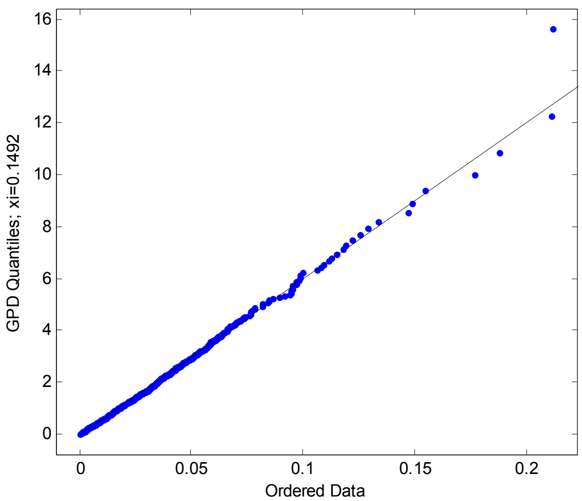


Fig. 3. Mean excess function



a) QQ-plot: empirical vs exponential distribution



b) QQ-plot: empirical vs GPD distribution (ξ=0.1492, β=0.0206, u=0.0334)

Fig. 4. QQ-plot versus GPD distribution and exponential distribution

Table 2. VaR estimation for daily RTSI losses:one day horizon

VaR approach	p=0.950	p=0.975	p=0.990	p=0.995	p=0.999
Normal model	0.0394	0.0451	0.0520	0.0565	0.0663
Historical simulation	0.0452	0.0607	0.0849	0.1083	0.1771
GPD model	0.0499	0.0611	0.0856	0.1062	0.1620

definitely to underestimate high percentiles, while estimates based on historical simulation face with the problem of out of sample performance. The extreme value approach on GPD model seems appropriate and easy to implement.

Conclusions

Since last century, volatility of international financial system is getting severe. A stable financial system is so desirable. Therefore, risk management has aroused growing attention. As a measurement of market risk, VaR has been widely used in risk management. However, derivation between VaR estimation of normal hypothesis and abnormal distribution of practical benefit rate of financial always cause the bigger error in estimation. Aiming at this problem, through GPD model which fits tail distribution of financial products more accurately, this paper recalculates VaR by POT method. Compared with traditional method of risk study, this paper has made some progress in research approach and philosophy and more applicable in practice, which has been demonstrated by example of the Russian market analysis.

We use software systems: EVIM [25,26], MATLAB [27] and LOGOS-EVT, developed by authors of this paper.

References

[1] Jorison, P. (2001) Value at Risk: the New Benchmark for Managing Financial Risk, 2nd ed, McGraw-Hill

[2] Embrechts P., Kluppelberg C. and Mikosch T. (1997) Modelling extremal events: for insurance and finance, Springer, Berlin.

[3] Fernandez, V., 2005. Risk management under extreme events. International Review of Financial Analysis 14, 113-148.

[4] Kluppelberg, C. (2004) Risk management with extreme value theory. In: Finkenstaedt, B. and Rootzen, H. (Eds) Extreme Values in Finance, Telecommunication and the Environment, pp.101-168. Chapman and Hall/CRC, Boca Raton.

[5] Gilli, M., Kellezi, E., 2006. An application of extreme value theory for measuring financial risk. Computational Economics 27, 207-228.

[6] de Haan, L., Jansen, D. W., Koedijk, K. G., and de Vries, C. G. (1994). Safety first portfolio selection, extreme value theory and long run asset risks. In J. Galambos, J. Lechner and E. Simiu, eds., Extreme value Theory and Applications, Kluwer, Dordrecht, pages 471–488.

[7] McNeil, A. J. (1997). Estimating the tails of loss severity distributions using extreme value theory. ASTIN Bulletin, 27, 1117–137.

[8] McNeil, A. J. (1998). Calculating quantile risk measures for financial time series using extreme value theory. Manuscript, Department of Mathematics, ETH, Swiss Federal Technical University.

[9] Embrechts, P., Resnick, S., and Samorodnitsky, G. (1998). Extreme value theory as a risk management tool. Manuscript, Department of Mathematics, ETH, Swiss Federal Technical University.

[10] Embrechts, P. (1999). Extreme value theory in finance and insurance. Manuscript, Department of Mathematics, ETH, Swiss Federal Technical University.

[11] Embrechts, P. (2000a). Extreme value theory: Potentials and limitations as an integrated risk management tool. Manuscript, Department of Mathematics, ETH, Swiss Federal Technical University.

[12] McNeil, A. J. (1999). Extreme value theory for risk managers. Internal Modeling and CAD II, Risk Books, pages 93–113.

[13] McNeil, A. J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. Journal of Empirical Finance, 7, 271–300.

[14] Embrechts, P., Kluppelberg, C., and Mikosch, C. (1997). Modeling Extremal Events for Insurance and Finance. Springer, Berlin.

[15] Embrechts, P. 2000. Extremes and Integrated Risk Management. Risk Books and UBS Warburg, London.

[16] McNeil A.J., Frey R. and Embrechts P. Quantitative risk management: Concepts, techniques and tools. 2005. Princeton University Press.

[17] B. M. Hill, A simple general approach to inference about the tail of a distribution, Annals Statistics 3 (1975) 1163–1173.

[18] Reiss, R.-D.and Thomas, M.(1997). Statistical Analysis of Extreme Values. Birkhauser, Basel (2nd ed. 2001).

[19] Zhou, Chunyang, Wu, Chongfeng, Liu, Hailong and Liu, Fubing,A New Method to Choose the Threshold in the Pot Model. Available at SSRN.

[20] Loretan, M., Phillips, P., 1994, Testing the Covariance Stationarity of Heavy tailed Time Series, Journal of Empirical Finance 1, 211-248.

[21] www.rts.ru

[22] Brys G, Hubert M, Struyf A. A Robustification of the Jarqua-Bera Test of Normality. COMPSTAT 2004 Symposium, Section: Robustness, 2004.

[23] Erik Brodin, Claudia Klueppelberg Extreme Value Theory in Finance. In: Everitt, B. and Meiinick, E. (Eds.) Encyclopedia of Quantitive Risk Assessment.

[24] Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. Springer-Verlag, London.

[25] Gencay, R., Selcuk, F., and Ulugaulyagci, A. (2003). EVIM: a software package for extreme value analysis in Matlab. *Studies in Nonlinear Dynamics and Econometrics*, 5:213-239.

[26] www.sfu.ca/~rgencay/

[27] www.mathworks.com

Применение метода превышений порога и теории экстремальных значений для моделирования рисков на российском фондовом рынке

В.О. Андреев[а], С.Е. Тиняков[б], О.П. Овчинникова[а], Г.П. Парахин[в]

[а] Орловская региональная академия государственной службы Россия 302028, Орел, ул. Победы, 5а

[б] Железногорский филиал СФУ Россия, Железногорск, ул. Кирова, 12а

[в] ТелекомСтройСервис Россия 302528, Орловская обл., п. Зареченский, ул. Царев Брод, 61

При моделировании поведения фондового рынка обычно используется нормальное распределение Гаусса. Однако, метод превышений порога (POT) теории экстремальных значений (EVT), обобщенное распределение Парето (GPD) позволяют более точно описывать финансовые доходности (потери) от операций с ценными бумагами, особенно на хвостах вероятностных распределений. В данной работе описывается применение предложенных методов к моделированию и анализу потерь на основе индекса РТС, представляющего собой усредненную цену акций 50 крупнейших российских компаний, за 15-летний период (1995 - 2009 г.). Особое внимание уделено использованию предложенных методов для оценки экстремального рыночного риска на хвостах распределений, что позволяет получить современный инструмент моделирования для системы управления рисками.

Ключевые слова: теория экстремальных значений, обобщенное распределение Парето, метод превышений порогового значения, вероятностное распределение на хвосте, стоимость под риском.