FORECASTING FINANCIAL RISKS BY MODIFIED GRID-BASED DECOMPOSITION ALGORITHM FOR NORMAL VARIANCE-MEAN MIXTURES

Alexander Korchagin
Moscow State University,
Leninskie Gory, Moscow, Russia
Email: sasha.korchagin@gmail.com

Victor Korolev
Moscow State University,
Leninskie Gory, Moscow, Russia
Institute of Informatics Problems,
FRC CSC RAS
Email: vkorolev@cs.msu.su

KEYWORDS

Risk prediction; Volatility decomposition; Variance-mean mixtures; Generalized Hyperbolic Distribution; Grid-based algorithms for separation of mixtures.

ABSTRACT

We describe an algorithm to forecast financial risks using parametrized models of normal variance-mean mixtures.

The proposed method takes a set of vectors as the input, containing a fixed number of the distribution parameters – the final result of the modified two-step grid-based decomposition algorithm applied to a moving time window. In this article we use the class of generalized hyperbolic (GH) distributions as an example for method demonstration. Practical applications of the method proposed and processing speed are discussed in detail. We also describe the process of calibrating the method as well as provide detailed instructions on how to find the best fitting model. Using real market data we illustrate the accuracy of the resulting forecasts depending on the method settings, including long-term forecasts.

INTRODUCTION

One of the most important hands-on problems of the financial market analysis is the problem of forecasting the risks associated with the excess of certain critical thresholds by specific indicators. Statistical analysis of available financial data often concludes that many distributions are subject to have so-called heavy tails. Precise estimation of those tails will enable us to adequately estimate the associated risks.

In addition to study of the distributions themselves, any financial organization is very interested in getting more or less accurate forecasts based on the data observed. Forecasting carries a large speculative factor, but some requirements for any meaningful forecasting method can be set in advance: a method should work fast enough so that the decision makers have sufficient time to take action based on the forecast; and the method should demonstrate good results on a randomly selected historical data sample.

To simplify the problem of estimating and forecasting distributions an approach is often used to narrow the acceptable distribution classes, significantly reducing the dimension of the task. In particular, models based on the normal variance-mean mixtures demonstrated extremely high highest adequacy when solving practical problems associated with the analysis of various financial indexes evolution. Moreover, these models have a substantial theoretical justification due to the fact they are limit distributions for certain random walks with random intensities [6].

In this paper we propose a forecasting algorithm for the parameters (in particular, for risk estimation) of special subordinated Wiener processes whose distributions are normal variance-mean mixtures. This class of mixtures is very large and includes generalized hyperbolic distributions which were introduced by O.-E. Barndorff-Nielsen in 1977–78 [2], [3].

Let $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$. If the distribution function of the generalized hyperbolic law with parameters α, β, ν, μ, λ is denoted as $P_{GH}(x;\alpha,\beta,\nu,\mu,\lambda)$, then, by definition

$$P_{GH}(x;\alpha,\beta,\nu,\mu,\lambda) = \int_0^\infty \Phi\Big(\frac{x-\beta-\alpha z}{\sqrt{z}}\Big) \cdot p_{GIG}(z;\nu,\mu,\lambda)dz, \quad x \in \mathbb{R}, \quad (1)$$

where $\Phi(x)$ is the standard normal distribution function:

$$\Phi(x) = \int_{-\infty}^{x} \varphi(z)dz, \quad \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R},$$

$p_{GIG}(x;\nu,\mu,\lambda)$ is the density of the generalized inverse Gaussian distribution:

$$p_{GIG}(x;\nu,\mu,\lambda) = \frac{\lambda^{\nu/2}}{2\mu^{\nu/2}K_\nu\big(\sqrt{\mu\lambda}\big)} \cdot x^{\nu-1} \cdot \exp\Big\{-\frac{1}{2}\Big(\frac{\mu}{x}+\lambda x\Big)\Big\}, \quad x > 0.$$

Here $\nu \in \mathbb{R}$,

$$\begin{aligned} &\mu > 0, \quad \lambda \geqslant 0, \quad \text{if } \nu < 0,\\ &\mu > 0, \quad \lambda > 0, \quad \text{if } \nu = 0,\\ &\mu \geqslant 0, \quad \lambda > 0, \quad \text{if } \nu > 0, \end{aligned}$$

Proceedings 29th European Conference on Modelling and Simulation ©ECMS Valeri M. Mladenov, Petia Georgieva, Grisha Spasov, Galidiya Petrova (Editors)
ISBN: 978-0-9932440-0-1 / ISBN: 978-0-9932440-1-8 (CD)

$K_\nu(z)$ is the modified Bessel function of the third kind with index ν,

$$K_\nu(z) = \frac{1}{2}\int_0^\infty y^{\nu-1}\exp\Big\{-\frac{z}{2}\Big(y+\frac{1}{y}\Big)\Big\}dy,$$
$$z\in\mathbb{C},\ \mathrm{Re}\, z>0.$$

DESCRIPTION OF THE METHOD FOR FORECASTING FINANCIAL RISKS AND ITS CHARACTERISTICS

The proposed approach aims at forecasting the *distribution* of the random process under consideration rather than at forecasting its *values*. To obtain point-wise forecasts within this approach, one just can take, say, the expectation of the predicted distribution (the best mean-square predictor) or its median (the best mean-absolute-deviation predictor). Moreover, this approach makes it possible to forecast the risks related to large deviations (say, VaR) by forecasting the behavior of the tails of the distributions or its quantiles.

We will describe the method using the example of the five-parameter class of generalized hyperbolic processes. Note that the method uses only the provided values of distributions parameters and does not take into account any other information about the class so that this method can be directly applied to any other parameterized class of distributions.

The first step is to take the analyzed time series and apply the standard approach of mixture analysis with a moving sample (or "window"). In order to do this, we need to fix the window size w and the shift of that window $s\ lew$ (i. e., how many new values from the time series will the next window have as compared to the previous one). After that, we apply the modified two-step grid-based method of normal variance-mean mixtures decomposition proposed in [1].

At the first step we select the *main* part of the mixing distribution support on $\mathbb{R}_+$, i. e. we specify an interval, the probability of which (calculated in accordance with the mixing distribution) is approximately equal to 1. Once the interval is selected, we cover it with a finite grid containing a *known* number K of *known* nodes $u_1,\dots,u_K$. Assume that the shift parameter $beta$ is zero. Then the approximation of the unknown generalized hyperbolic distribution by a finite mixture of normal laws is

$$P_{GH}(x;\,\alpha,0,\nu,\mu,\lambda)\approx\sum\nolimits_{i=1}^{K}p_i\Phi\Big(\frac{x-\alpha u_i}{\sqrt{u_i}}\Big),\quad x\in\mathbb{R}. \tag{2}$$

In the mixture on the right-hand side of (2), only the parameters $p_1,\dots,p_{K-1},\alpha$ are unknown. Let $x_1,\dots,x_n$ be the analyzed sample of random variable values which has the generalized hyperbolic distribution being estimated. The iterative process that defines the grid EM (Expectation Maximization) algorithm is defined as follows.

Let $p_1^{(m)},\dots,p_{K-1}^{(m)},\alpha^{(m)}$ be the estimates of the parameters $p_1,\dots,p_{K-1}$ and α respectively obtained on the mth iteration, $p_K^{(m)}=1-p_1^{(m)}-\ldots-p_{K-1}^{(m)}$. Denote

$$\varphi_{ij}^{(m)}=\frac{1}{\sqrt{u_i}}\varphi\Big(\frac{x_j-\alpha^{(m)}u_i}{\sqrt{u_i}}\Big),\quad g_{ij}^{(m)}=\frac{p_i^{(m)}\varphi_{ij}^{(m)}}{\sum_{r=1}^{K}p_r^{(m)}\varphi_{rj}^{(m)}},$$
$$i=1,\dots,K;\ j=1,\dots,n.$$

Then, using the standard procedures that define the EM-algorithm formulas for the parameters of a finite normal mixture (see, e.g., [4], sections 5.3.7–5.3.8), we obtain

$$p_i^{(m+1)}=\frac{1}{n}\sum\nolimits_{j=1}^{n}g_{ij}^{(m)},\quad i=1,\dots,K. \tag{3}$$

Denote $\overline{x}=\frac{1}{n}\sum_{j=1}^{n}x_j$. Using formula (5.3.24) from [4] and taking into account the obvious equation $\sum_{i=1}^{K}g_{ij}^{(m)}=1$ we notice that the updated α parameter estimation is

$$\alpha^{(m+1)}=\frac{\overline{x}}{\sum_{i=1}^{K}u_ip_i^{(m+1)}}, \tag{4}$$

i.e., it is equal to the ratio of the general sample mean and the current empirical mean of the mixing distribution.

It is a known fact that the classic EM algorithm is monotone. If the grid nodes $u_1,\dots,u_K$ are not equal to each other, are non-negative and known, then the iterative process (3) - (4) is monotone, i.e. every iteration of the algorithm does not reduce the target grid likelihood function

$$L(p_1,\dots,p_K,\alpha;x_1,\dots,x_n)=$$
$$=\prod\nolimits_{j=1}^{n}\Big[\sum\nolimits_{i=1}^{K}\frac{p_i}{\sqrt{u_i}}\varphi\Big(\frac{x_j-\alpha^{(m)}u_i}{\sqrt{u_i}}\Big)\Big].$$

In section 5.7.4 of [4] it is demonstrated that for each fixed α the grid likelihood function $L(p_1,\dots,p_{K-1},\alpha;\,x_1,\dots,x_n)$ is concave in $p_1,\dots,p_{K-1}$. Therefore, at each step of the iterative process instead of (3) we can use any other faster algorithm to maximize the function $L(p_1,\dots,p_{K-1},\alpha^{(m)};\,x_1,\dots,x_n)$ by variables $p_1,\dots,p_{K-1}$. For example, the estimates of the weights $p_1,\dots,p_K$ can be found using conditional gradient method [4].

As the result, on the first step we are able to produce an estimate for α and weights for each point u_i of the grid, covering the majority of the mixing Generalized Inverse Gaussian (GIG) distribution $P_{GIG}(z;\,\nu,\mu,\lambda)$.

On the second step we apply any standard method of fitting a parametrized distribution $P_{GIG}(z;\,\nu,\mu,\lambda)$ to the empirical data represented by the histogram $(u_1,p_1),\dots,(u_K,p_K)$. For example, we can find ν, μ and λ parameters solving the least squares problem

$$(\nu^*,\mu^*,\lambda^*)=\arg\min_{\nu,\mu,\lambda}\sum\nolimits_{i=1}^{K}\Bigg[p_i-$$
$$-\int_{\frac{1}{2}(u_{i-1}+u_i)}^{\frac{1}{2}(u_i+u_{i+1})}p_{GIG}(u;\,\nu,\mu,\lambda)du\Bigg]^2,$$

given $u_0=0$, $u_{K+1}=\infty$. As an alternative, these parameters can be found from the condition

$$(\nu^*,\mu^*,\lambda^*)=\arg\min_{\nu,\mu,\lambda}D_{KL}\big[p_{GIG}(u;\,\nu,\mu,\lambda),\,h(u)\big],$$

where

$$h(u)=\begin{cases}0, & u\leqslant\frac{1}{2}u_1,\\ \dfrac{u_{i+1}-u_{i-1}}{2p_i}, & \frac{1}{2}(u_{i-1}+u_i)<u\leqslant\frac{1}{2}(u_i+u_{i+1}),\\ 0, & u>\frac{1}{2}(3u_K-u_{K-1}),\end{cases}$$

is the histogram constructed from the values $(u_1, p_1), \ldots, (u_K, p_K)$, and $D_{KL}\big[p_{GIG}(u;\, \nu, \mu, \lambda),\, h(u)\big]$ is the Kullback–Leibler divergence.

A specific method of grid selection is described in [1].

As the input for the forecasting method we use the result of the two-step method described above, i.e. we use the set of vectors $\theta_1, \theta_2, \ldots, \theta_N$ containing the distribution parameters calculated on $\hat{N}$ known frames; $\theta_i = (\alpha_i, \beta_i, \nu_i, \mu_i, \lambda_i)^T$

The goal of the forecast is to estimate $\theta_{n+1}, \theta_{n+2}, \ldots$ for the frames which will partially or fully consist of the unknown data from the future. To achieve this, we use

$$\tilde{\theta}_{i+1} = F_1\theta_i + F_2\theta_{i-1} + \ldots + F_r\theta_{i-r+1},$$

where $r \in \mathbb{N}$ is a pre-defined constant meaning the *order* of the regression, $F_j \in \mathbb{R}^{5\times 5}$ – regressor-matrices.

The procedure to find the parameters is:

$$(F_1, \ldots, F_r) = \arg\min \sum\nolimits_{i=r+1}^{N-1} \left(\theta_{i+1} - \tilde{\theta}_{i+1}\right)^2. \qquad (5)$$

To put it simply, we have a regression model, where the search of the matrices F_j (learning process) is realized via the minimization of the residual sum of squares (RSS) using $\hat{N} - r$ known data points. To make the notation more readable, let $N = \hat{N} - r$.

Formula (5) is a variation of a linear regression, so implementing this calculation is not a complex task. Moreover, many statistical programs and packages already have built-in functions for the calculation of F_j when $r = 1, 2$.

ACCURACY OF THE FORECAST

The residual sum of squares, calculated in the process of finding F_j can be treated as a criteria of model fitness (if we divide it by N), but this value does not take into account the assumption regarding the form of the distribution itself. It is important that some parameters might have significantly bigger impact on the shape of the resulting distribution than the others.

In order to determine the accuracy of forecasts more precisely let us compare the actual forecasted *GH-distribution* with parameters $\tilde{\theta}_{i+1}$ and the historically calculated GH-distribution with parameters θ_{i+1} (as opposed to direct comparison of those parameters to each other). In order to do the comparison, we propose to use a number of metrics.

Let $\tilde{F}(x)$ and $F(x)$ – cumulative distribution functions (CDF) of generalized hyperbolic distributions with parameters $\tilde{\theta}_{i+1}$ and θ_{i+1} respectively, let $\tilde{f}(x)$ and $f(x)$ – corresponding probability density functions.

1. C, L_1, L_2 metrics
These metrics are often used to determine how close distributions are. They are respectively defined as

$$C(\tilde{f}, f) = \sup_{x\in\mathbb{R}} |\tilde{f}(x) - f(x)|,$$

$$L_1(\tilde{f}, f) = \int_{-\infty}^{+\infty} |\tilde{f}(x) - f(x)|dx,$$

$$L_2(\tilde{f}, f) = \left[\int_{-\infty}^{+\infty} (\tilde{f}(x) - f(x))^2 dx\right]^{1/2}.$$

The smaller the values C, L_1 or L_2, the closer the compared distributions are to each other.

2. aIntersect metric
This metric has the meaning of the graphical discrepancy (anti-intersect) of the density functions:

$$I(\tilde{f}, f) = 1 - \int_{-\infty}^{+\infty} \min(\tilde{f}(x), f(x))dx$$

The closer aIntersect to zero, the closer the compared distributions are.

3. Tail estimations
Specifically for the estimation of the tails of the distribution related to the corresponding risks, it is reasonable to introduce the following measures of tail prediction quality:

$$W_{0.05} = \tilde{F}(x_{0.05}), \quad W_{0.95} = \tilde{F}(x_{0.95}),$$

x_α stands for the α-quantile of the distribution $F(x)$.

Since the GH-distributions are absolutely continuous, the above can be presented in a more usable format:

$$W_{0.05} = \tilde{F}(F^{-1}(0.05)), \quad W_{0.95} = \tilde{F}(F^{-1}(0.95)).$$

The closer $T_{0.05}$ is to 0.05, the more accurate is the estimated left tail of the distribution. The same applies to the right tail for $T_{0.95}$ and 0.95

In addition to the estimation of the left and right tails, it is also reasonable to analyze the accuracy of the quantile estimation itself:

$$S_{0.05} = |\tilde{x}_{0.05} - x_{0.05}|, \quad S_{0.95} = |\tilde{x}_{0.95} - x_{0.95}|,$$

or using the density function

$$S_{0.05} = |\tilde{F}^{-1}(0.05) - F^{-1}(0.05)|,$$
$$S_{0.95} = |\tilde{F}^{-1}(0.95) - F^{-1}(0.95)|.$$

The lower S_α, the more accurate the corresponding quantile estimation is.

All the above can be applied to 0.025- and 0.975-quantiles as well, covering 95% of the distribution.

By calculating the metrics mentioned above for different forecast parameters (regression order, historical data size used for learning) we can find the optimal model for each individual forecasting challenge.

RESULTS FOR REAL MARKET DATA

The paper [7] contains a detailed description of things to consider when using the modified two-step grid-based decomposition method as well as practical advice on how to get the values $\theta_1, \theta_2, \ldots, \theta_n$ in the fastest and most accurate manner.

As the source data for our experiment we use the KOSPI market index (Korea Composite Stock Price Index) is the main weighted indicator of Korean Stock Exchange which includes all publicly traded companies on that particular Exchange. This

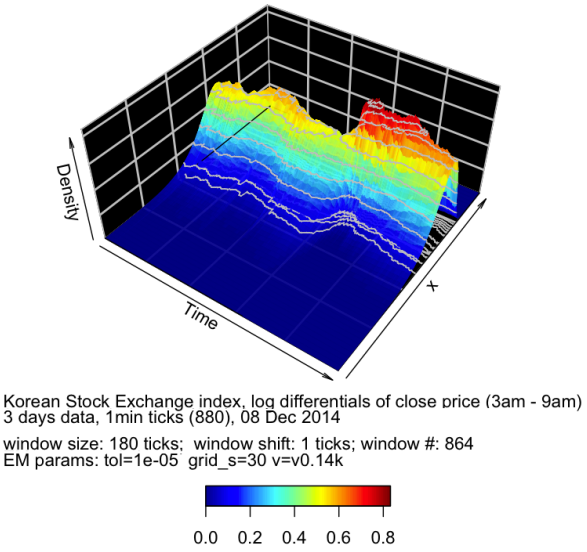

Fig. 1. Evolution of GH-distribution approximation of KOSPI index over time.

index was introduced in 1983 and is principal for Korean markets ever since, so it is often used for the analysis and forecast exercises. More details can be found on Bloomberg - www.bloomberg.com/quote/KOSPI:IND.

We will analyze the dynamic of the logarithms of the index values with one minute frequency (tick) during three working days, starting from market opening on 8 December, 2014.

The window size is set equal to 3 hours: $w = 180$, the window shift is minimal and equals one tick, $s = 1$. Once we apply the two-step grid algorithm to the above, it produces the parameter estimates $\theta_i = (\alpha_i, \beta_i, \nu_i, \mu_i, \lambda_i)^T$. In this experiment without significantly limiting the model we used all $\beta_i = 0$, leaving only 4 parameters in business.

The algorithm is robust to minor changes in input parameters, and the window shift is minimal, so the vector θ_i is "smooth" enough, i. e., for each next step the parameter values are close to the previous stem. Fig. 1 presents the dynamic of the GH-distribution in time (reconstructed using θ_i).

FORECASTING PROCEDURE AND SELECTED MODELS

Let us select a specific point in time (window T=300), that we will treat as *the current point in time* among available 864 frames. This will be the borderline between the data which is known to the researcher and which is yet unknown: the data on the left will be used for the calculation of the parameters of different models ('learning'), the data on the right will be used to compare the forecasted values against the real recorded data.

For the assessment of the forecast accuracy we will calculate the metrics for forecasts built for certain frames in the future:

- T+1 - frame shifted by 1 tick, minimal possible forecast, 99% values known as of moment T,

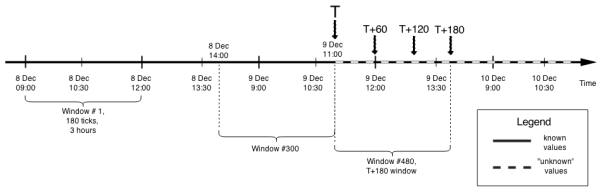

Fig. 2. Experiment timeline, local Korea Stock Exchange time.

TABLE I. MODELS IN SCOPE, PARAMETERS r AND N.

N =	10	20	50	100	200
$r = 1$ (16 params)	✓	✓	✓	✓	✓
$r = 2$ (32 params)			✓	✓	✓
$r = 3$ (48 params)				✓	✓

- T+10 - forecast of 10 ticks (10 minutes), 94% values known as of moment T,
- T+60 - forecast of 60 ticks (1 hour), 66% values known as of moment T,
- T+120 -forecast of 120 ticks (2 hours), 33% values known as of moment T,
- T+180 - forecast of 180 ticks (3 hours), 0% values known as of moment T.

Fig. 2 illustrates the experiment timeline. The moment T is marked on the timeline, the events on the left are available for us and are historical values, the events on the right need to be forecasted (and later compared to the actual values).

As the fitted models we will use models with $r = 1, 2$ 3, i. e., matrix-based regression of the first, second and third orders. In addition to changing r we will also vary the number of historical events N used for calculations. For example, the models that use less data will reflect more recent trends but will probably be less accurate for the long-term forecasts.

In this case, the calculation of regressor-matrices reduces to the calculation of 16, 32 or 48 floating point parameters for $r = 1, 2$ and 3 respectively. Taking into account the number of available data, we will specify which models is included into consideration in Table I.

SELECTING THE BEST MODEL FOR FIXED FORECAST INTERVAL

Once all the required model parameters are calculated (this process is very rapid), we will fix the most relevant prediction interval ("the horizon"): for example, we are mostly interested in forecasting one hour ahead or using the T-terminology, T+60. The comparison of different models for this selected horizon can be found in Tables II, III and IV.

To better visualize the results, the best mertic values are highlighted in Tables II, III and IV for each of the groups $r = 1, 2$ and 3.

Based on the results above, we make the conclusion that for $r = 1$ and $r = 2$ the model with minimum RSS is in fact not the best for forecasting the distribution at T+60. This fact also confirms the necessity to use the metrics introduced above.

TABLE II. FORECAST QUALITY COMPARISON FOR DIFFERENT MODELS, +1 HOUR (T+60). METRICS C, L_1, L_2, I.

			Forecast T+60			
r	N	RSS	C	L_1	L_2	I,%
1	10	0.583	0.080	0.106	0.070	5.68
1	20	0.377	0.091	0.118	0.079	6.23
1	50	0.350	**0.062**	**0.082**	**0.054**	**4.40**
1	100	**0.315**	0.123	0.159	0.106	8.36
1	200	0.920	0.286	0.355	0.235	18.05
2	50	**0.275**	0.130	0.169	0.112	8.81
2	100	0.291	**0.078**	**0.116**	**0.073**	**6.22**
2	200	0.903	0.272	0.340	0.225	17.33
3	100	**0.271**	**0.064**	**0.102**	**0.063**	**5.55**
3	200	0.758	0.269	0.332	0.221	16.94

TABLE III. FORECAST QUALITY COMPARISON FOR DIFFERENT MODELS, +1 HOUR (T+60). TAIL METRICS FOR 90%.

			T+60 forecast, 90% interval			
R	N	RSS	$W_{0.05}$	$W_{0.95}$	$S_{0.05}$	$S_{0.95}$
1	10	0.583	0.059	0.935	0.112	0.051
1	20	0.377	0.055	0.931	0.065	0.063
1	50	0.350	**0.055**	**0.938**	**0.057**	**0.041**
1	100	0.315	0.058	0.920	0.102	0.092
1	200	0.920	0.041	0.866	0.121	0.193
2	50	0.275	0.052	0.916	0.026	0.105
2	100	0.291	0.067	0.934	0.201	0.054
2	200	0.903	0.041	0.870	0.117	0.188
3	100	0.271	0.068	**0.937**	0.206	**0.045**
3	200	0.758	**0.043**	0.873	**0.089**	0.181

Among all selected models our top picks are $r = 1, N = 50$ and $r = 2, N = 100$, the corresponding rows in the table are highlighted with gray color. Based on the metric values alone, the simple linear regression of the first order $r = 1, N = 50$ provides the best fit.

Fig. 3 demonstrates the forecasted distribution vs. the actual one for the best selected model. Based on the calculated metrics we may conclude that the model found by the proposed method is suitable for the predictions of at least one hour ahead. Also we have a strong evidence that these forecasts provide a very reliable estimation of the distribution tails, which is the key for our initial goal, future risk assessment.

FURTHER ENHANCEMENTS

The method described above can be enhanced in many ways. If we know up-front which metric is the most important

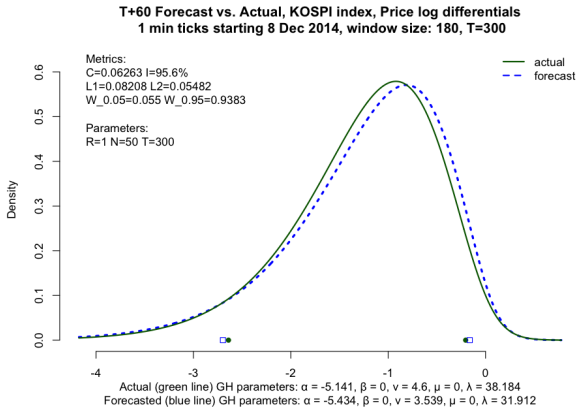

Fig. 3. Forecasted and actual dist., $r = 1$, $N = 50$.

TABLE IV. FORECAST QUALITY COMPARISON FOR DIFFERENT MODELS, +1 HOUR (T+60). TAIL METRICS FOR 95%.

		T+60 forecast, 95% interval			
R	N	$W_{0.025}$	$W_{0.975}$	$S_{0.025}$	$S_{0.975}$
1	10	0.033	0.968	0.173	0.033
1	20	0.030	0.966	0.119	0.045
1	50	0.029	**0.969**	0.096	**0.029**
1	100	0.033	0.960	0.177	0.072
1	200	**0.023**	0.927	**0.040**	0.161
2	50	0.028	0.956	0.085	0.087
2	100	0.038	0.968	0.283	0.037
2	200	0.023	0.930	0.040	0.157
3	100	0.039	**0.969**	0.283	**0.030**
3	200	**0.024**	0.932	**0.006**	0.149

for our task, we can directly minimize this metric on historical data instead of finding the least RSS (see eq. (5)). However, in most cases the calculation of regressor-matrices F_j is not a linear programming problem and requires significantly more computational resources. One of potential ways to overcome the increased computation time is to increase the window shift, for example, $s = 20$. By increasing the shift we decrease the frequency of recalculation procedures, allowing more time per each individual forecast.

This work was supported by Russian Science Foundation, grant no. 14-11-00364.

REFERENCES

[1] Korolev, V. Yu., A. Yu. Korchagin. *Modified grid-based method for decomposition of normal variance-mean mixtures (in Russian)*. Informatics and Its Applications 2014; 8(4), 11-19.

[2] *O. E. Barndorff-Nielsen.* Exponentially decreasing distributions for the logarithm of particle size // Proc. Roy. Soc. Lond., Ser. A, 1977. Vol. 353. P. 401–419.

[3] *O. E. Barndorff-Nielsen.* Hyperbolic distributions and distributions of hyperbolae // Scand. J. Statist., 1978. Vol. 5. P. 151–157.

[4] Korolev, V. Yu. 2011. *Probabilistic and statistical methods for the decomposition of volatility of chaotic processes (in Russian)*. Moscow: Moscow University Press. 510 p.

[5] Nazarov, A. L. 2013. *Approximate methods for the decomposition of volatility of chaotic processes (in Russian)*. PhD Thesis. Moscow State University.

[6] Korolev, V. Yu. 2013. *Generalized hyperbolic distributions as limiting for random sums (in Russian)*. Theory Probab. Appl., 58(1):117–132.

[7] Korchagin, A. Yu., I. I. Yaroshenko *On practical usage of modified grid-based method for decomposition of normal variance-mean mixtures (in Russian)*. Materials from young scientists conference "Zadachi sovremennoi informatiki" , 9 dec 2014 . M.: IPI RAN, 34–38.

AUTHOR BIOGRAPHIES

ALEXANDER KORCHAGIN is a PhD student of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University.

VICTOR KOROLEV is Doctor of Science in physics and mathematics, professor, Head of Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University; leading scientist, Institute of Informatics Problems, Russian Academy of Sciences.