The Use of The Kolmogorov–Wiener Filter for Prediction of Heavy-Tail Stationary Processes

Vyacheslav Gorev, Alexander Gusev and Valerii Korniienko

Dnipro University of Technology, 19 Dmytra Yavornytskoho Ave., 49005 Dnipro , Ukraine

Abstract

We investigate the possibility of the practical use of the Kolmogorov–Wiener filter for the prediction of a heavy-tail stationary random process. A discrete process and a discrete filter are considered. Nowadays telecommunication traffic in telecommunication systems with data packet transfer is considered to be a heavy-tail random process, so the problem under consideration may be applied to the prediction of telecommunication traffic, which may be important, for example, for the prevention of network congestion, for the maximization of the network utilization rate and for cyber security, because a comparison of the actual traffic with the predicted one may help to detect cyber-attacks. There are a lot of different and rather sophisticated approaches to traffic prediction, for example, the ARIMA approach, neural network approaches and so on, which may be applicable to the prediction of a non-stationary traffic in various cases. However, in the rather simple case of a stationary telecommunication traffic, more simple approaches may be applied. For example, such a simple prediction approach as the Kolmogorov–Wiener filter is not sufficiently developed in the literature. In this paper it is shown that if a stationary heavy-tail random process is smooth enough, then the Kolmogorov–Wiener filter may be used for its practical prediction. The obtained results may be taken into account for practical telecommunication traffic prediction in telecommunication systems with data packet transfer.

Keywords

Kolmogorov–Wiener filter, prediction, heavy-tail stationary random process, power-law correlation function, telecommunication traffic

1. Introduction and related works

The problem of telecommunication traffic prediction is important for telecommunications. For example, it is important for the prevention of network congestion and for the maximization of the network utilization rate [1]; it is significant for understanding future market dynamics and reducing the decision risks [2]. The telecommunication traffic prediction is also important for cyber security [3] because the comparison of the actual traffic with the predicted one may help to detect cyber-attacks.

There are a lot of different approaches to traffic prediction. For example, the following ones can be indicated: Auto Regressive Integrated Moving Average (ARIMA), Markov Modulated Poisson Process models (MMPP), Kalman filtering, Seasonal ARIMA (SA), a neural network approach (including deep neural networks [4]), wavelet transforms [1], the least-squares support vector machine (LSSVM), gray models [2], Holt-Winters models [3]. Of course, rather complicated approaches should be used for non-stationary randomly fluctuating traffic prediction. But if the traffic is stationary and rather smooth, sophisticated approaches may not be needed. For example, in [2] some methods are presented for a description of rather simple cases. In [2] it is stressed that in stationary

cases the ARMA approach may be used too, and in the case of a smooth monotone process the gray model may be applied.

As is known [5], such a simple filter as the Kolmogorov–Wiener one may be used for the prediction of stationary random processes. However, as far as we know, such an approach is not sufficiently developed in the literature for traffic prediction even for rather simple cases. The Kolmogorov–Wiener filter is widely used for signal extraction in different fields of knowledge [6]. It is widely used in econometric analyses [7, 8] and in image restoration [9]. The theoretical fundamentals of the Kolmogorov–Wiener filter for continuous telecommunication traffic prediction are developed in our recent paper [10]. The paper [10] is dedicated to the solution of the Wiener–Hopf integral equation in the unknown filter weight function for two telecommunication traffic models: the power-law structure function model and the model of fractional Gaussian noise; the solutions based on the truncated polynomial expansion method and the truncated trigonometric Fourier series method are obtained.

However, the possibility of using the Kolmogorov–Wiener filter for practical traffic prediction is still under question. The aim of this work is to show that the Kolmogorov–Wiener filter may be applicable to traffic prediction if the traffic is stationary and smooth enough. As is known [11, 12], the telecommunication traffic in systems with data packet transfer is considered to be a self-similar heavy-tail random process. So, if we show that the Kolmogorov–Wiener filter is applicable to the prediction of simulated data of a stationary random self-similar heavy-tail process, then we will be able to conclude that it may be applied to practical telecommunication traffic prediction. In this paper we restrict ourselves to the investigation of a discrete process and a discrete filter. The corresponding simulated data may be generated via the symmetric moving average approach [13], the generated process is in fact similar to the fractional Gaussian noise process, which may describe telecommunication traffic, see [14].

The paper is organized as follows. In Sec. 1 the introduction and the literature review are given. In Sec. 2 the discrete Kolmogorov–Wiener filter and the symmetric moving average approach for obtaining simulated stationary heavy-tail data are described. In Sec. 3 heavy-tail simulated data are obtained. In Sec. 4 the prediction results are described, and in Sec. 5 conclusions are made.

2. Description of the discrete Kolmogorov–Wiener filter and of the method of generation of heavy-tail simulated data

Let the filter input x_t be a stationary random process which is the sum of the signal s_t and the noise n_t:

$$x'_t = s_t + n_t. \tag{1}$$

The Kolmogorov–Wiener filter output y_t should be «the closest» to the value s_{t+z} where z is the number of points for which the prediction is made, so we have the following requirement:

$$\langle (y_t - s_{t+z})^2 \rangle \to \min. \tag{2}$$

The correlation function $R_{x'}(t)$ of the filter input x'_t and the cross-correlation function of the processes s_t and x'_t $R_{sx'}(t)$ are considered to be given. The Kolmogorov–Wiener filter is considered to be a linear one, so the filter output is expressed in terms of the filter input as follows:

$$y_t = \sum_{i=0}^{T} h_i x'_{t-i} \tag{3}$$

where h_i are the unknown filter weight coefficients and the input data are given for $t = 0,1,2,..,T$. The coefficients h_i should minimize expression (2). The term $\langle s_{t+z}^2 \rangle$ is a constant and does not depend on the weight coefficients h_i, so (2) can be rewritten as

$$\langle y_t^2 \rangle - 2\langle y_t s_{t+z} \rangle \to \min, \tag{4}$$

which in view of (3) gives

$$\sum_{i,j=0}^{T} h_i h_j \langle x'_{t-i} x'_{t-j} \rangle - 2 \sum_{i=0}^{T} h_i \langle x'_{t-i} s_{t+z} \rangle = f(h_0, h_1, \ldots, h_T) \to \min. \tag{5}$$

With account for the facts that

$$\langle x'_{t-i} x'_{t-j} \rangle = R_{x'}(i - j) \tag{6}$$

and

$$\langle x'_{t-i} s_{t+z} \rangle = R_{sx'}(i + z) \tag{7}$$

one can finally write

$$\sum_{i,j=0}^{T} h_i h_j R_{x'}(i - j) - 2 \sum_{i=0}^{T} h_i R_{sx'}(i + z) = f(h_0, h_1, \dots, h_T) \to \min. \tag{8}$$

The function $f(h_0, h_1, \dots, h_T)$ is a quadratic one, and thus it has one minimum, which is described by the conditions

$$\frac{\partial f(h_0, h_1, \dots, h_T)}{\partial h_k} = 0; \quad k = 0,1,2,\dots,T. \tag{9}$$

These conditions with account for the evenness of the correlation function and the fact that

$$\frac{\partial h_i}{\partial h_j} = \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases} \tag{10}$$

lead to

$$\sum_{i=0}^{T} h_i R_{x'}(i - k) = R_{sx'}(k + z); \quad k = 0,1,2,\dots,T, \tag{11}$$

which is a set of linear equations in the unknown coefficients h_i. In matrix form, this set may be presented as

$$R_{x'} \cdot h = R_{sx'} \tag{12}$$

where

$$R_{x'} = \begin{pmatrix} R_{x'}(0) & R_{x'}(1) & R_{x'}(2) & \dots & R_{x'}(T) \\ R_{x'}(1) & R_{x'}(0) & R_{x'}(1) & \dots & R_{x'}(T-1) \\ R_{x'}(2) & R_{x'}(1) & R_{x'}(0) & \dots & R_{x'}(T-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_{x'}(T) & R_{x'}(T-1) & R_{x'}(T-2) & \dots & R_{x'}(0) \end{pmatrix} \tag{13}$$

is the correlation matrix [5], h is the vector column of the unknown weight coefficients, and R_{sx} is the vector column of the free terms:

$$h = \begin{pmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_T \end{pmatrix}, \quad R_{sx'} = \begin{pmatrix} R_{sx'}(z) \\ R_{sx'}(z+1) \\ R_{sx'}(z+2) \\ \vdots \\ R_{sx'}(z+T) \end{pmatrix}. \tag{14}$$

So, the vector column h may be found as

$$h = R_{x'}^{-1} \cdot R_{sx'}. \tag{15}$$

Then the filter output may be obtained by formula (3).

It should be noticed that all the above-mentioned calculations are described in [6]. The Kolmogorov–Wiener filter may be used both for the extraction of a signal form the sum of a signal and a noise and for the signal prediction. In the case where the input signal is non-noisy, the Kolmogorov–Wiener filter may be used for the prediction of the stationary process given at the filter input. In the non-noisy case, the filter weight coefficients are given by formula (15) with account for the fact that

$$R_{sx'} = (R_{x'}(z) \quad R_{x'}(z+1) \quad R_{x'}(z+2) \quad \dots \quad R_{x'}(z+T))^T. \tag{16}$$

Now let us describe the method of the generation of heavy-tail simulated data, which is used in the paper. In this paper we use the symmetric moving average approach, which is described in detail in [13]. Such an approach was chosen because of its simplicity.

Let V_t be a stationary white noise process with an average value equal to zero and a variance equal to 1. Then a heavy-tail process X_i similar to the fractional Gaussian noise may be generated as follows [13]:

$$X_i = \sum_{j=-q}^{q} a_{|j|} V_{i+j} = a_q V_{i-q} + a_{q-1} V_{i-q+1} + \dots + a_q V_{i+q}, \tag{17}$$

theoretically, q should be infinite; in practical calculation it may be a rather large, but finite number; and the coefficients a_j are as follows:

$$a_0 = \frac{\sqrt{(2 - 2H)\gamma_0}}{1.5 - H} \qquad (18)$$

and

$$a_j = \frac{a_0}{2}((j + 1)^{H+0.5} + (j - 1)^{H+0.5} - 2j^{H+0.5}), \qquad (19)$$

here, γ_0 is the variance and H is the Hurst exponent of the process X_i. The number q may be very large, it is estimated as follows [13]:

$$q \geq \max\left(m, \left(\frac{H^2 - 0.25}{2\beta}\right)^{\frac{1}{1.5-H}}\right) \qquad (20)$$

where m is the number of correlation function points of the process X_i which should be obtained and a small number β is in fact the given accuracy of the coefficient a_j in (17); the values $a_{j>q}$ should be less than βa_0. The accuracy of this method depends on q, and the method is not exact even in the case where $q \to \infty$. However, for a rather large q the method may lead to good practical results [13].

3. The generation of non-smooth and smooth heavy-tail simulated data

10^6 points of the white noise process V_t with an average value equal to 0 and a variance equal to 1 are generated on the basis of the generator built in the Wolfram Mathematica package. The following parameters were chosen:

$$m = 10^5, \qquad \beta = 10^{-4}, \qquad H = 0.8, \qquad \gamma_0 = 1. \qquad (21)$$

The corresponding number $q = 3 \cdot 10^5$ is chosen. In fact, the inequality (21) holds even for $q = 10^5$, the value $q = 3 \cdot 10^5$ was chosen for a higher accuracy. On the basis of the idea (17)–(19), 10^5 points of the process X_i were generated as follows:

$$X_i = \sum_{j=-q}^{q} a_{|j|}V_{i+j+q}, \qquad (22)$$

in fact, the quantities V_{i+j+q} and V_{i+j} are independent because V_t is the white noise, no matter whether formula (17) or formula (22) is used; formula (22) is chosen in order to avoid indices beyond the array V_i bounds. The coefficients a_j are calculated on the basis of (19).

The average value of X_i is close to zero. We have to construct simulated data that may describe telecommunication traffic, which is obviously non-negative. So we build the array x_i as follows:

$$x_i = X_i + |\min(X)| + 10^{-3}, \qquad (23)$$

a small summand 10^{-3} is added in order to avoid obtaining an infinite value of the prediction mean average percentage error (MAPE). The process x_i is a non-negative random stationary heavy-tail process; its graph is given in Fig. 1.

Let us make sure that the generated process x_i is a heavy-tail one. Let us describe the corresponding centralized process xc_i:

$$xc_i = x_i - \langle x \rangle \qquad (24)$$

where the average value $\langle x \rangle$ is

$$\langle x \rangle = \frac{1}{10^5} \sum_{j=1}^{10^5} x_i, \qquad (25)$$

here we take into account the fact that the number of points of the generated array x_i is equal to 10^5. The correlation function of the process xc_i is built as follows:

$$R_x(\tau) = \langle xc_i \cdot xc_{i+\tau} \rangle = \frac{1}{10^5 - \tau} \sum_{i=1}^{10^5-\tau} (xc_i \cdot xc_{i+\tau}). \qquad (26)$$

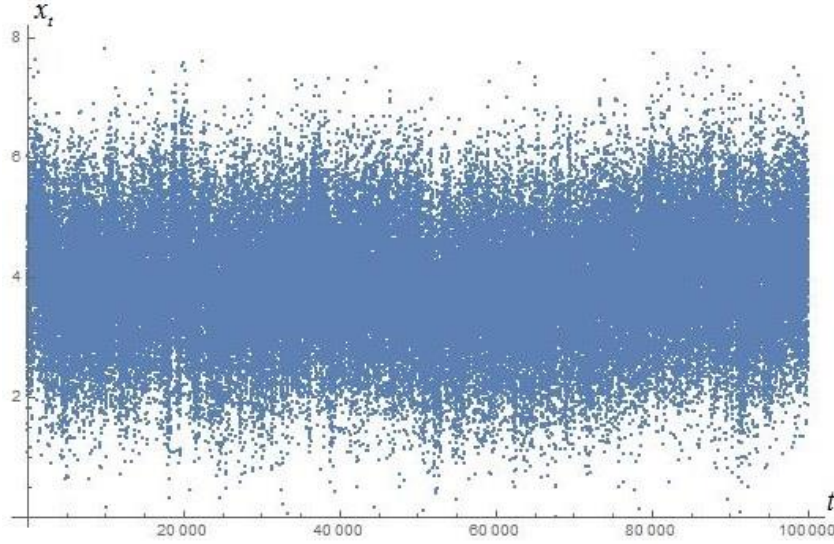The corresponding correlation function and its least-square fit are given in Fig.2.

Figure 1: The values of the simulated non-smooth heavy-tail non-negative random process
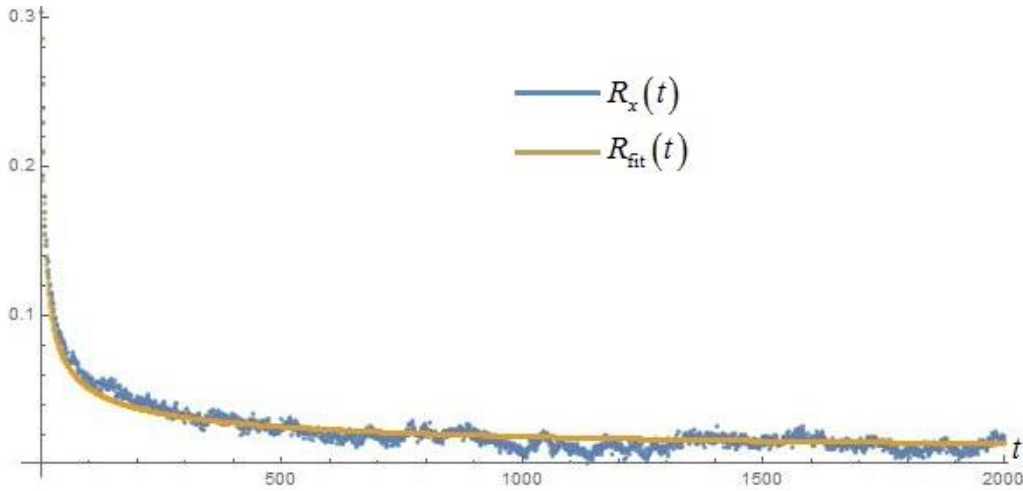

Figure 2: The correlation function of the simulated non-smooth heavy-tail random process and its least-square power-law fit; t≥1.

The corresponding least-square fit is sought as
$$R_{\text{fit}}(t) = a \cdot t^b , \tag{27}$$
The following numerical coefficients were obtained:
$$a = 0.39, \qquad b = -0.44, \tag{28}$$
here, the coefficients are rounded off to two significant digits. So,
$$R_{\text{fit}}(t) = 0.39 \cdot t^{-0.44} , \tag{29}$$
and on the basis of formula (29) and Fig.2 one can conclude that the correlation function exhibits a power law decay rather than an exponential one. So, indeed, the generated process is a heavy-tail one.

It should also be noticed that according to [13] the following property should be valid for ≥ 1 :
$$R_x(t) \sim t^{2H-2}, \tag{30}$$
so, according to the least-square fit
$$2H - 2 = -0.44, \tag{31}$$
which leads to
$$H = 0.78, \tag{32}$$
which is very close to the value 0.8, see (21). The variance of the process is equal to
$$R_x(0) = 0.93, \tag{33}$$
which is rather close to the value $\gamma_0 = 1$, see (21). So one can conclude that the generated process is close to the fractional Gaussian noise with given variance and Hurst exponent.

The generated process is non-smooth, i.e. it is really highly fluctuating, so it is rather difficult to predict it. So it is reasonable to investigate smooth heavy-tail processes. In order to obtain smoother processes, we use a very simple smoothing algorithm [15]:

$$\tilde{X}_i = \frac{1}{2l+1} \sum_{j=-l}^{l} X_{i+j} \qquad (34)$$

where $\tilde{X}_i$ are the values of a smooth process, expression (34) is valid for every point except for the first l and the last l ones. The first l and the last l points of the process $\tilde{X}_i$ may be obtained as the corresponding linear least-square fit of the first l and the last l points of the process X_i, respectively. The corresponding non-negative process may be expressed similarly to (23):

$$\tilde{x}_i = \tilde{X}_i + \left|\min(\tilde{X})\right| + 10^{-3}, \qquad (35)$$

and the corresponding centralized process

$$\widetilde{xc}_i = \tilde{x}_i - \langle \tilde{x} \rangle \qquad (36)$$

where the average value

$$\langle \tilde{x} \rangle = \frac{1}{10^5} \sum_{j=1}^{10^5} \tilde{x}_i . \qquad (37)$$

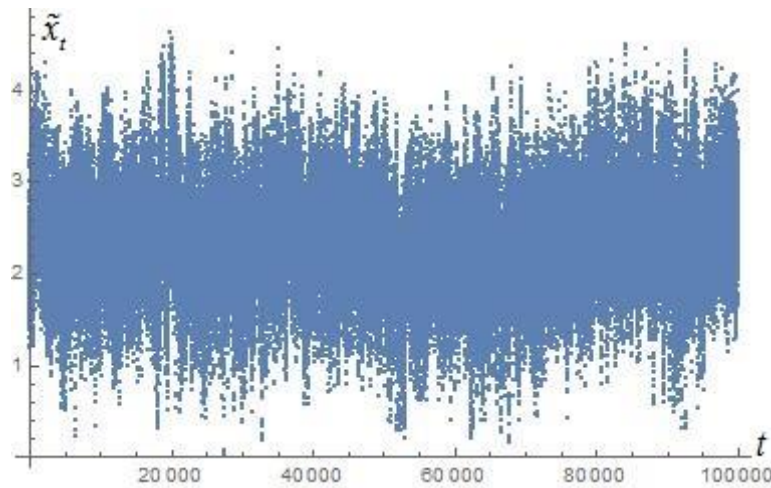The simulated data for the process $\tilde{x}_i$ for $l = 3$ are given in Fig.3.

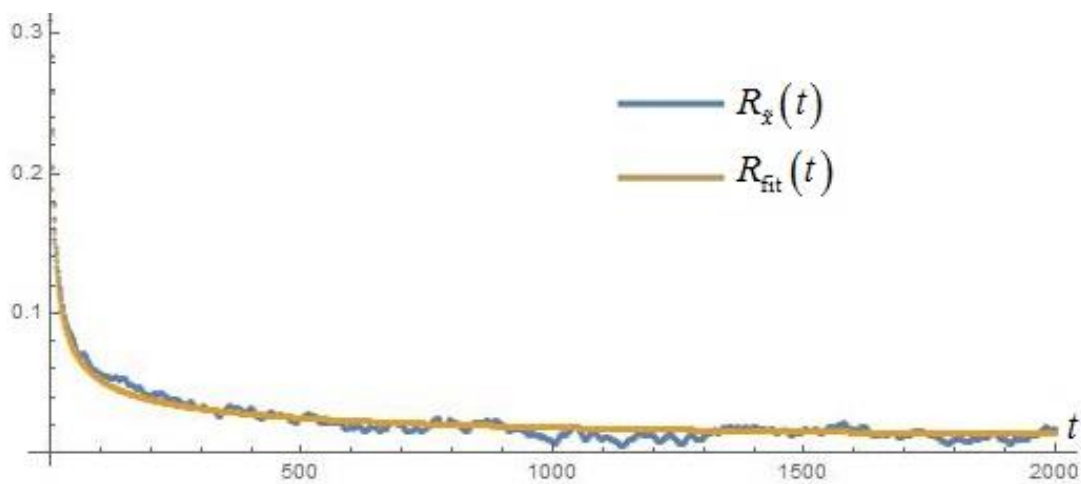**Figure 3:** The values of the simulated smooth heavy-tail non-negative random process for $l = 3$

Figure 4: The correlation function of the simulated smooth heavy-tail random process and its least-square power-law fit; t≥1.

It should be stressed that the obtained smooth process $\tilde{x}_i$ is also a heavy-tail one. Let us consider the corresponding correlation function:

$$R_{\tilde{x}}(\tau) = \langle \widetilde{xc}_i \cdot \widetilde{xc}_{i+\tau} \rangle = \frac{1}{10^5 - \tau} \sum_{i=1}^{10^5 - \tau} (\widetilde{xc}_i \cdot \widetilde{xc}_{i+\tau}). \tag{38}$$

For example, for $l = 3$ the following correlation function and its fit are obtained, see Fig. 4. The least-square fit is sought in the form (27), the following numerical coefficients were obtained:
$$a = 0.43, \qquad b = -0.46, \tag{39}$$
here, the coefficients are rounded off to two significant digits. So,
$$R_{\text{fit}}(t) = 0.43 \cdot t^{-0.46}. \tag{40}$$
As can be seen form Fig.4, the correlation function of a smooth process is also well described by a power-law function, the obtained smooth process $\tilde{x}_i$ is also a heavy-tail one, and, in fact, this process may also be roughly considered as fractional Gaussian noise.

4. Prediction on the basis of the Kolmogorov–Wiener filter

The prediction for non-smooth data is built as follows. In fact, the prediction for the centralized process is used. The filter weight coefficients are built on the basis of (13)–(16); the corresponding correlation function is taken in the form (26).

First of all, the points $xc_1, xc_2,\ldots, xc_{T+1}$ of the simulated process xc are taken as the filter input, and the points $xc_{T+2}, xc_{T+3},\ldots, xc_{T+z+1}$ are predicted. Then the points $xc_2, xc_3, \ldots, xc_{T+2}$ are taken from the simulated data, and the points $xc_{T+3}, xc_{T+4},\ldots, xc_{T+z+2}$ are predicted, and so on.

At the ith iteration of the algorithm the predcition is calculated as follows. The filter input data are
$$x'_0 = xc_i, x'_1 = xc_{i+1}, \ldots, x'_T = xc_{i+T}, \tag{41}$$
so
$$x'_j = xc_{i+j}. \tag{42}$$
The filter output y_t is the predicted value for x'_{t+z} (the non-noisy case is investigated). According to (3) we have
$$y_t = \sum_{k=0}^{t} h_k x'_{t-k}, \tag{43}$$
here, the upper bound of summation is changed in order to avoid obtaining indices beyond the array of the input data. Such a change of the bound does not lead to a significant error for the prediction under consideration. On the basis of (41)–(43) one can conclude that
$$\widehat{xc}_{i+t+z} = \sum_{k=0}^{t} h_k xc_{t+i-k} \tag{44}$$
where $\widehat{xc}_{i+t+z}$ is the predicted value of xc_{i+t+z}. Obviously, the prediction is made only for the values $i + t + z = T + 1 + i, \ T + 2 + i, \ldots, T + z + i$. We should also remember that we should make the prediction for the non-negative simulated data. So, the predicted non-negative data may be expressed as
$$\hat{x}_{i+t+z} = \widehat{xc}_{i+t+z} + \langle x \rangle = \langle x \rangle + \sum_{k=0}^{t} h_k xc_{t+i-k}, \ t = \overline{T + 1 - z, T}. \tag{45}$$
The MAPE and MAE errors for the corresponding prediction are calculated as
$$\text{MAPE} = \frac{1}{z} \sum_{t=T+z-1}^{T} \left| \frac{\hat{x}_{i+t+z} - x_{i+t+z}}{x_{i+t+z}} \right| \cdot 100\% \tag{46}$$
and
$$\text{MAE} = \frac{1}{z} \sum_{t=T+z-1}^{T} |\hat{x}_{i+t+z} - x_{i+t+z}|. \tag{47}$$
The corresponding prediction errors are calculated at each iteration.

Let us tell a few words why the above-mentioned change of the upper bound of summation has no significant effect on the result. In order to make the prediction for $\widehat{xc}_{T+1+i}$, one should calculate the sum of $T + 2 - z$ summands, in order to make the prediction for $\widehat{xc}_{T+2+i}$ one should calculate the

sum of $T + 3 - z$ summands, and so on. We obviously deal with the case where $T \gg z$, so the value $T + 1 - z$ is rather close to $T + 1$, so the above-mentioned change of the upper bound is not significant for the calculations.

Similarly, the prediction for the smooth heavy-tail process is made as follows. At the ith iteration of the algorithm the prediction is calculated as follows:

$$\hat{\tilde{x}}_{i+t+z} = \langle \tilde{x} \rangle + \sum_{k=0}^{t} h_k \widetilde{xc}_{t+i-k}, \ \ t = \overline{T + 1 - z, T} \tag{48}$$

and the corresponding MAPE and MAE errors are

$$\text{MAPE} = \frac{1}{z} \sum_{t=T+z-1}^{T} \left| \frac{\hat{\tilde{x}}_{i+t+z} - \tilde{x}_{i+t+z}}{\tilde{x}_{i+t+z}} \right| \cdot 100\% \tag{49}$$

and

$$\text{MAE} = \frac{1}{z} \sum_{t=T+z-1}^{T} \left| \hat{\tilde{x}}_{i+t+z} - \tilde{x}_{i+t+z} \right|. \tag{50}$$

The MAPE and MAE are calculated for each above-mentioned iteration. The $10^5 - T - z$ MAPE and MAE values both for smooth and for non-smooth processes are obtained. The following parameters are chosen:

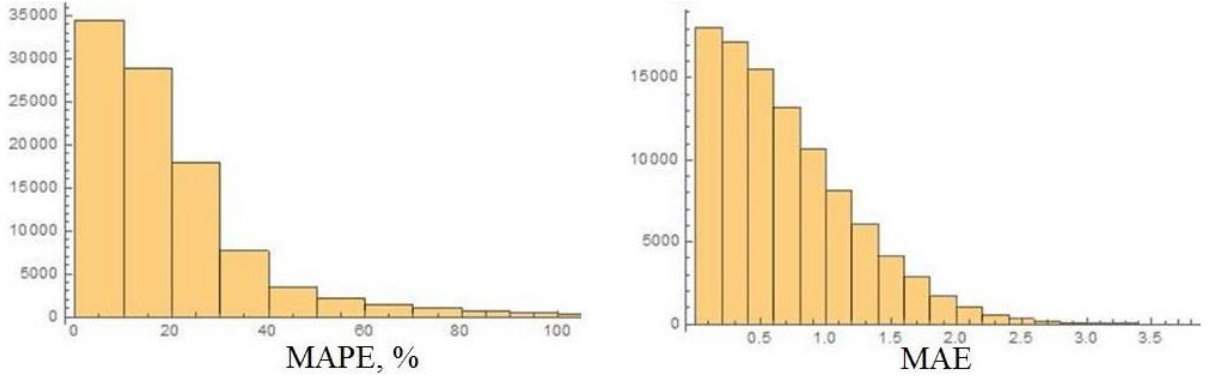$$\text{T} = 100, \quad \text{z} = 1. \tag{51}$$

Figure 5: The MAPE and MAE histograms for the prediction of a non-smooth heavy-tail process

Table 1
The prediction results for a smooth heavy-tail process

l	$\langle \tilde{x} \rangle$	Average MAPE, %	Average MAE
1	2.98	9.11	0.235
2	2.52	6.26	0.142
3	2.34	4.85	0.103
4	2.31	3.92	0.081
5	2.22	3.37	0.067
6	2.11	2.98	0.057
7	2.04	2.68	0.050

The following results are obtained. The MAPE and MAE histograms in the case of the non-smooth process are shown in Fig.5. The y-axes of the histograms indicate the number of MAPE and MAE values that belong to the corresponding intervals. For the non-smooth process the average MAPE is 24.7%, and the average MAE is 0.70 (the average value of the process is $\langle x \rangle = 3.88$). It should also be stressed that for some points the MAPE are more than 100%. So one can conclude that the prediction accuracy is not high in the case of the non-smooth process. So, if the process is a highly fluctuating one, then the prediction based on the Kolmogorov–Wiener filter may not lead to good results.

But if the process is rather smooth, the prediction results are much better. The corresponding results are given in Table 1. In Table 1 l is the parameter used in (34), i.e. $2l + 1$ is the number of smoothing points. As can be seen, the smother the process is, the better the prediction results are, and the prediction accuracy increases with l. The corresponding histograms for $l = 3$ are given in Fig.6. The predictions for $l \geq 6$ have an average MAPE value less than 3%.
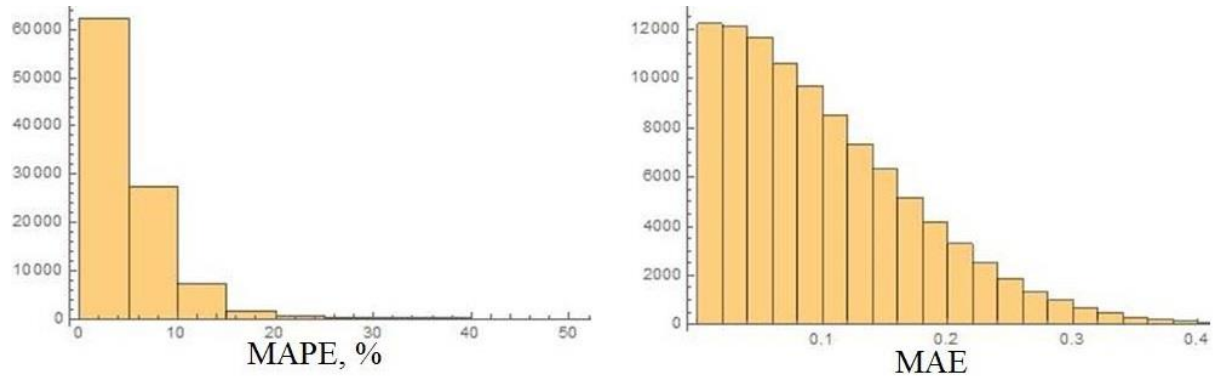

Figure 6: The MAPE and MAE histograms for the prediction of a non-smooth heavy-tail process

For example, for $l = 3$ the average MAPE is less than 5%. As can be seen from the corresponding histogram, the MAPE for the overwhelming majority of points is less than 10%. For some very rare points the MAPE may be rather high (up to 40%), but in our opinion this may be explained as follows. As can be seen from Fig. 3, the values for some points of the process $\tilde{x}$ are rather close to zero, and the MAPE may not be an adequate characteristic for the prediction of points close to zero. So, one can conclude that the Kolmogorov–Wiener filter may give good results for the prediction of a stationary heavy-tail random process if the process is smooth enough.

5. Conclusions and plans for the future

The use of the Kolmogorov–Wiener filter for the prediction of stationary random heavy-tail processes is considered. The attention is paid to the discrete case. The problem under consideration may be connected with the telecommunication traffic prediction, which is important, for example, for cyber security, see [3]. There are many rather sophisticated approaches to telecommunication traffic prediction [1]. For rather simple cases (stationary or smooth traffic) the ARMA or gray model approaches may be used [2]. The traffic in telecommunication systems with data packet transfer is considered to be a self-similar heavy-tail process, see [11]. Such a simple filter as the Kolmogorov–Wiener one may be used in the prediction of stationary random processes [6]. However, as far as we know, the corresponding approach for traffic prediction is not sufficiently developed in the literature.

In this paper we generate data for a stationary heavy-tail process on the basis of the symmetric moving average approach [13]. The corresponding non-smooth and smooth data are generated. The prediction for 1 point forward on the basis of the previous 101 points is investigated. It is shown that the Kolmogorov–Wiener filter is not good for non-smooth processes, but may give a good prediction for a stationary random heavy-tail process if the process is rather smooth. So, if the traffic is stationary and rather smooth, the Kolmogorov–Wiener filter may be used for its prediction. The advantage of the corresponding approach is the simplicity of the method in contrast with, for example, neural networks or ARIMA models.

The plans for the future are as follows. In this paper only the values $T = 100$ and $z = 1$ are investigated. So the prediction investigation for a wider range of parameters may be a plan for the future. In our recent paper [10] the theoretical approach to the Kolmogorov–Wiener filter construction in the continuous case is considered. In this paper we generated a large number of data points, which may allow one to try to investigate the continuous case, so the investigation of the applicability of the method [10] may be another plan for the future. This paper is based on the generation of simulated data, so the investigation of real experimental traffic data may be another plan for the future. It should also be stressed that the use of the Kolmogorov–Wiener filter for the prediction of stationary

processes may be useful not only in telecommunications, but also in other fields of knowledge, for example, in electrical engineering, see [16].

6. References

[1] Q. H. Do, T. T. H. Doan, T. V. A. Nguyen, N. T. Duong, V. Van Linh, Prediction of Data Traffic in Telecom Networks based on Deep Neural Networks, Journal of Computer Science 16 (2020) 1268-1277. doi: 10.3844/jcssp.2020.1268.1277.

[2] J.-X. Liu, Z.-H. Jia, Telecommunication Traffic Prediction Based on Improved LSSVM, International Journal of Pattern Recognition and Artificial Intelligence, 32, No. 3 (2018) 1850007 (16 pages), doi: 10.1142/S0218001418500076.

[3] H. Brugner, Holt-Winters Traffic Prediction on Aggregated Flow Data, Proceedings of the Seminars Future Internet and Innovative Internet Technologies and Mobile Communication Focal Topic: Advanced Persistent Threats. Summer Semester 2017 (2017), 25-32. doi: 10.2313/NET-2017-09-1_04.d

[4] P. Kaushik, S. Singh, P. Yadav, Traffic Prediction in Telecom Systems Using Deep Learning, Proceedings of 7th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), August 29-31, 2018, Noida, India (2018), 302-207, doi: 10.1109/ICRITO.2018.8748386.

[5] P. S. R. Diniz, Adaptive Filtering Algorithms and Practical Implementation, 5th ed., Springer Nature Switzerland AG, Cham, 2020, doi: 10.1007/978-3-030-29057-3.

[6] T. Bao, J. Duffy, Signal extraction: experimental evidence, Theory and Decision 90 (2021), 219–232. doi: 10.1007/s11238-020-09785-x

[7] S. G. Pollock, Filters, Waves and Spectra, Econometrics 6 (2018), 35 (33 pages). doi: 10.3390/econometrics6030035

[8] S. G. Pollock, E. Mise, A Wiener–Kolmogorov Filter for Seasonal Adjustment and the Cholesky Decomposition of a Toeplitz Matrix, Computational Economics 59 (2022), 913–933. doi: 10.1007/s10614-020-10087-1

[9] V. Pronina, F. Kokkinos, D.V. Dylov, S. Lefkimmiatis, Microscopy Image Restoration with Deep Wiener-Kolmogorov Filters, in: A. Vedaldi, H. Bischof, T. Brox, JM. Frahm (Eds.), Lecture Notes in Computer Science, vol 12365, Springer, Cham, 2020, pp. 185–201. doi:10.1007/978-3-030-58565-5_12

[10] V. Gorev, A. Gusev, V. Korniienko, M. Aleksieiev, Kolmogorov–Wiener Filter Weight Function for Stationary Traffic Forecasting: Polynomial and Trigonometric Solutions, in: P. Vorobiyenko, M. Ilchenko, I. Strelkovska (Eds.), Lecture Notes in Networks and Systems, vol 212, Springer, 2021, pp. 111–129. doi:10.1007/978-3-030-76343-5_7

[11] D. Zhuang, C. Li, Loss Analysis for Networks based on Heavy-Tailed and Self-Similar Traffic, Journal of Physics: Conference Series 1584 (2020), 012054 (8 pages). doi: 10.1088/1742-6596/1584/1/012054.

[12] D. Radev, I. Lokshina, Advanced models and algorithms for self-similar IP network traffic simulations and pefformance analysis, Journal of Electrical Engineering 61, No. 6 (2010), 341-349. doi: 10.2478/v10187-010-0053-0.

[13] D. Koutsoyiannis, The Hurst phenomenon and fractional Gaussian noise made easy, Hydrological Sciences Journal, 47 (2002), 573-595. doi: 10.1080/02626660209492961.

[14] M. Li, Generalized fractional Gaussian noise and its application to traffic modeling, Physica A 579 (2021), 126138 (22 pages). doi: 10.1016/j.physa.2021.126138.

[15] K. Molugaram, G. S. Rao, Statistical Techniques for Transportation Engineering, Butterworth-Heinemann (Elsevier), Oxford, 2017, doi: 10.1016/B978-0-12-811555-8.00012-X.

[16] Yu. A. Papaika, O. H. Lysenko, Ye. V. Koshelenko, I. H. Olishevskyi, Mathematical modeling of power supply reliability at low voltage quality, Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, No. 2 (2021), 97-103. doi: 10.33271/nvngu/2021-2/097.