UDC 519.24
DOI: 10.20535/SRIT.2308-8893.2025.2.07

TIME SERIES FORECASTING USING THE NORMALIZATION MODEL

VIKTOR BONDARENKO, VALERIIA BONDARENKO

Abstract. Empirical constructions of time series models based on the reduction of initial data to normally distributed values have been proposed. The goal of a normalization method is to construct an optimal forecast that is linear for the updated data, and the forecasted original data is recovered through the inverse transformation. The different variants of such transformations have been considered, including the reduction of initial data to Gaussian fractional Brownian motion and a one-dimensional transformation using a strictly monotonic function. The computational experiment based on real data, which allows for a stationary model, confirms the higher quality of the forecast by the normalization method compared to traditional models.

Keywords: optimal forecast, stochastic model, parameter estimation, fractional Brownian motion.

PRELIMINARY INFORMATION AND STATEMENT OF THE PROBLEM

Time forecasting is defined as estimation the future values of some function of a time variable based on known observations up to the current moment. Other-words, if we observe the trajectory $x(t)$, $0 \le t \le T$, then it is necessary to evaluate the value $x(s)$, $T \le s \le T + \tau$. The estimated values are called forecasts and are denoted by $\hat{x}(s)$, . The stochasticity of the trajectory is an essential circumstance, so $x(\cdot)$ does not represent a deterministic function.

As a rule, the trajectory values are observed at discrete moments

$$0 \le t_1 < t_2 < \ldots < t_{n-1} < t_n \le T,$$

that is, the terms of the sequence are known $\{x_1, x_2, \ldots, x_n\}$, where $x_k = x(t_k)$, and "future" values are to be assessed $x(t_{n+1}), \ldots, (t_{n+r})$,

$$T \le t_{n+1} < t_{n+2} < \ldots < t_{n+r} \le T + \tau .$$

If it is known a priori that $x(t)$ is an implementation of a random process $\xi(t)$, which corresponds to a finite-dimensional distribution

$$P\{\xi(t_1) < x_1, \ldots, \xi(t_m) < x_m\} = \int_{-\infty}^{x_1} dy_1 \int_{-\infty}^{x_2} dy_2 \ldots \int_{-\infty}^{x_m} f_m(t_1,..,t_m, y_1, y_2,..y_m)\, dy_m$$

with a density $f_m(t_1, \ldots, t_m, x_1, \ldots, x_m) \equiv f_m(x_1, \ldots, x_m)$, then the optimal forecast for r steps is the conditional average:

$$\begin{pmatrix} \hat{\xi}_{n+1} \\ \vdots \\ \hat{\xi}_{n+r} \end{pmatrix} = \mathrm{E}(\xi_{n+1}, \xi_{n+2}, \ldots, \xi_{n+r}\ \xi_1 = x_1, \ldots, \xi_n = x_n) = \boldsymbol{g}(x_1, x_2, \ldots, x_n), \qquad (1)$$

$$\xi_j = \xi(t_j).$$

The coordinates of the conditional mean are calculated by the formula:

$$\hat{\xi}_{n+j} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} y_j p(y_1, y_2, \dots, y_r\, x_1, \dots, x_n) dy_1 \dots dy_r ,$$

where the conditional density

$$p(y_1, y_2, \dots, y_r\, x_1, \dots, x_n) =$$

$$= f_{n+r}(x_1, x_2, \dots, x_n, y_{n+1}, \dots, y_{n+r}) \times (f_n(x_1, x_2, \dots, x_n))^{-1} ,$$

$$f_n(x_1, x_2, \dots, x_n) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{n+r}(x_1, x_2, \dots, x_n, z_1, \dots, z_r) dz_1 \dots dz_r .$$

Optimality means that for random vectors

$$\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}, \quad \eta = \begin{pmatrix} \xi_{n+1} \\ \vdots \\ \xi_{n+r} \end{pmatrix}$$

and the fair ratio

$$\mathrm{E}\eta - g(\xi)^2 = \min_{\boldsymbol{h}} \mathrm{E}\eta - h(\xi)^2 ,$$

where the function g is defined by the equality (1), $\|\cdot\|$ is a Euclidian distance in R^n.

As a rule, the density of the distribution f_m is unknown, and to calculate the forecast, it is necessary to create a stochastic time series model $\{x_1, x_2, \dots, x_n\}$, that is, to construct a description of the relationship between the values of the series based on previous observations.

This model can be described

$$\xi_k = \Psi(\xi_1, \dots, \xi_{k-1}, \varepsilon) ,$$

where $\{\xi_1, \xi_2, \dots, \xi_n\}$ are the sequence of random variables for which the observation x_k is assigned by some value ξ_k, and ε is a random vector. Note that standard models do not always take into account the specifics of observations and the corresponding forecast turns out to be far from reality (as shown in the example below).

The forecast constructions based on traditional time series models are discussed in detail in the reference [1].

The preferred option of forecasting is a situation where the time series admits a Gaussian model, i.e. $\begin{pmatrix} \xi \\ \eta \end{pmatrix} \sim \aleph(0, Q)$, where the matrix Q is depicted in the form of

$$Q = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

and block elements A, B, C, D are determined by the ratios:

$$a_{jk} = E\xi_j \xi_k , \quad 1 \le j,\ k \le n \Leftrightarrow \xi \sim \aleph(0; A);$$

$$d_{jk} = E\eta_j \eta_k \equiv E\xi_{n+j}\xi_{n+k} , \quad 1 \le j,\ k \le r \Leftrightarrow \eta \sim \aleph(0; D)$$

and the elements of matrix $B, C,\ B = C^*$ are represent the mutual covariance of ξ and η, so

$$b_{jk} = E\xi_j\xi_{n+k}\ .1 \le j \le n, 1 \le k \le r;$$

$$c_{jk} = E\xi_{n+j}\xi_k\ ,\ 1 \le j \le r, 1 \le k \le n.$$

In this case, the optimal forecast is linear and is determined by the formula:

$$\hat{\eta} = \mathrm{E}(\eta \,|\, \xi) = CA^{-1}\xi\ . \tag{2}$$

In particular, the sample $\{x_1, \ldots, x_n\}$ can be considered the values of the fractional Brownian motion $B_H(t)$, which is defined as a Gaussian random process with characteristics

$$\mathrm{E}B_H(t) = 0, \quad \mathrm{E}B_H(t)B_H(s) = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}),$$

where $0 < H < 1$ is called a Hurst exponent.

The properties of fractional Brownian motion are discussed in [2]. The experience with temporal data shows that identifying a series as fBm values is a rather rare phenomenon.

TIME SERIES NORMALIZATION METHOD

The idea of the method consists of transforming the original data $\{x_1, x_2, \ldots, x_n\}$. into Gaussian $\{u_1, u_2, \ldots, u_n\}$..

In some cases, the original data can be converted to the fractional Brownian motion values.

Let there be a continuous one-to-one mapping

$$\varphi : R^n \to R^n\ ,\ \varphi(x_1, x_2, \ldots, x_n) = (u_1, u_2, \ldots, u_n)$$

such that the increments of the transformed data

$$z_1 = u_1,\ z_2 = u_2 - u_1, \ldots, z_n = u_n - u_{n-1}$$

form a stationary sequence (i.e., they admit a stationary model).

The statistics

$$\sigma^2 = \frac{1}{n}\sum_{k=1}^{n} z_k^2 \quad \text{and} \quad \theta = \frac{1}{n-1}\sum_{j=1}^{n-1} z_j z_{j+1}$$

are consistent estimates of the variance of the increments and their one-step covariance ([3, 4]).

Then consistent estimates of the correlation coefficient and Hurst parameter are calculated by the formulas:

$$\rho = \frac{\theta}{\sigma^2}\ , \quad H = \frac{1}{2} + \frac{ln(\rho + 1)}{2\,ln\,2}\ .$$

The necessary conditions for the hypothesis of "data $\{u_1, u_2, \ldots, u_n\}$ form the values of fractional Brownian motion" is the fulfillment of the limiting relations for the statistics A_n, B_n, D_n, F_n, which are defined by the following relations:

$$A_n = \frac{1}{n}\frac{1}{\sigma^4}\sum_{k=1}^{n} u_{k-1}z_k^3 \to -\frac{3}{2}, \quad H \in \left(0;\frac{1}{2}\right);$$

$$B_n = \frac{1}{n^{1+H}}\frac{1}{\sigma^5}\sum_{k=1}^{n} u_{k-1}^2 z_k^3 \to 3\eta, \;\; \eta \sim \aleph\left(0;\frac{1}{2H+2}\right), \;\; H \in \left(0;\frac{1}{2}\right);$$

$$D_n = \frac{1}{n^{2H}}\frac{1}{\sigma^4}\sum_{k=1}^{n} u_{k-1}z_k^3 \to \frac{3}{2}\varsigma^2 \;, \;\; \varsigma \sim \aleph(0;1), \quad H \in \left(\frac{1}{2};1\right);$$

$$F_n = \frac{1}{n^H}\frac{1}{\sigma^3}\sum_{k=1}^{n} z_k^3 \to 3\varsigma \;, \quad H \in \left(\frac{1}{2};1\right).$$

The proof of the limit relations is contained in [5].

The standard algorithm for testing the specified hypothesis using known H and σ is a following: let us assume that the hypothesis is fulfilled and we set the significance level α with comparing the value of the statistic with the tabular value β where $F(\beta) = 1 - \alpha$.

In particular, for the marginal distribution function of statistics D_n $(H > 0{,}5)$:

$F(x) = 2\Phi\left(\sqrt{x\frac{2}{3}}\right) - 1$, where Φ is a Laplace function.

Corresponding to the level of significance $\alpha = 0{,}05$ the critical value $\beta = 6$, and the hypothesis is accepted if $0 < D_n < 6$.

If the hypothesis $u_k = B_H(t_k)$ is true, then the optimal (linear) forecast for $\{u_{n+1}, u_{n+2}, \ldots, u_{2n}\}$ in *n* steps is calculated by the formula (2):

$$\begin{pmatrix} \hat{u}_{n+1} \\ \vdots \\ \hat{u}_{2n} \end{pmatrix} = CA^{-1}\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, \quad Q = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

$$q_{jk} == (j^{2H} + k^{2H} - |j-k|^{2H}), \;\; 1 \le j, \; k \le 2n,$$

and the forecast of the primary variables

$$(\hat{x}_{n+1}, \ldots, \hat{x}_{2n}) = \varphi^{-1}(\hat{u}_{n+1}, \ldots, \hat{u}_{2n}).$$

Note that the choice of transformation φ is a rather cumbersome procedure.

Let us consider another normalization method — the one-dimensional transformation of the real data $u_k = \varphi(x_k)$.

$$\varphi(x) = \Phi^{-1}(F_\xi(x)), \qquad \textbf{(3)}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \exp\left\{-\frac{z^2}{2}\right\}dz$, F_ξ is a distribution function corresponding to the sample $\{x_1, x_2, \ldots, x_n\}$.

Thereby, $\{u_1,\dots,u_n\}$ is a sample from the general population $\aleph(0,1)$ and the proposed procedure requires stationarity in the narrow sense of the real data and requires an assessment of their distribution. The sample size should not be large to prevent the detection of non-stationarity.

Actually, $\{x_1,x_2,\dots,x_n\}$ represents a sample of different random variables $\{\xi(t_1),\xi(t_2),\dots,\xi(t_n)\}$ with the same distribution F. This disadvantage is partially compensated by the dependence of the obtained data, which is determined by their correlation matrix.

Let us note the following property of the transformation.

Proposition. Let φ is a strictly increasing function $\xi(t)$ is a stationary process in the narrow sense, i.e.

$$P\{\xi(t_1)<x_1,\dots,\xi(t_n)<x_n\}=F(x_1,\dots,x_n,t_2-t_1,\dots,t_n-t_{n-1}).$$

Then the process $\eta(t)=\varphi(\xi(t))$ is also stationary in the narrow sense.

Proof. Under the condition of stationarity

$$P\{\xi(t_1)<x_1,\dots,\xi(t_n)<x_n\}=F(x_1,\dots,x_n,t_2-t_1,\dots,t_n-t_{n-1}).$$

Distribution of the process $\eta(t)$:

$$P\{\varphi(\xi(t_1))<x_1,\dots,\varphi(\xi(t_n))<x_n\}=P\{\xi(t_1)<y_1,\dots,\xi(t_n)<y_n\}=$$
$$=F(y_1,\dots,y_n,t_2-t_1,\dots,t_n-t_{n-1}),\quad y_j=\varphi^{-1}(x_j)$$

also satisfies the definition of stationarity.

Let us formulate a modeling and forecasting algorithm using transformation (3).

1. Checking the data for stationarity, for example, using the Dickey–Fuller criterion and determining the size of the training sample (in the case of checking the adequacy of the model, determine the size of the training plus the predicted sample).

2. Estimation of the distribution function F_ξ of random value ξ by the sample $\{x_1,\dots,x_n\}$.

3. Construction of the sample $\{u_1,\dots,u_n\}$. from a normal population $\aleph(0,1)$ by the formula

$$u_k=\Phi^{-1}(F_\xi(x_k)).$$

4. Calculation of the sample correlation coefficients

$$\rho_j=\frac{1}{n-j}\sum_{k=1}^{n-j}u_k u_{k+j},$$

with consctruction of correlation matrix Q and defining the forecasting horizon r.

5. Construction of the forecast $\{\hat{u}_{n+1,\dots,}\hat{u}_{n+r}\}$ of transformed data by the formula (2).

6. Calculating predicted values $\{\hat{x}_{n+1,\dots,}\hat{x}_{n+r}\}$ of primary data according to the formula

$$\hat{x}_k=\varphi^{-1}(\hat{u}_k)\equiv F_\xi^{-1}(\Phi(\hat{u}_k)).\quad \hat{x}_{n+1},\dots,\hat{x}_{n+r}.$$

EXAMPLE OF FORECAST CALCULATION

The following example illustrates the application of the proposed model:

The meteorological data files Precipitation–Florida Climate Center.

The sample size is determined by stationarity, which is tested using the Dickey–Fuller criterion.

In the given example, stationarity occurs in the interval $x_1 - x_{40}$; let us put the data into the training sample $x_1 - x_{30}$. The data values and their graph are shown in Table 1 and Fig. 1, respectively.

Table 1. The value of the series $\boldsymbol{x}_1 - \boldsymbol{x}_{40}$

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
3.03	**3.43**	**3.54**	**2.98**	**2.13**	**– 2.62**	**–0.61**	**–0.15**	**–2.15**	**–2.65**
x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}
–3.58	–2.13	–2.32	–2.43	–2.8	–2.42	–3.15	–2.62	–2.81	**–2.46**
x_{21}	x_{22}	x_{23}	x_{24}	x_{25}	x_{26}	x_{27}	x_{28}	x_{29}	x_{30}
–1.2	**–0.33**	**–1.81**	**–2.18**	**–0.2**	**0.6**	**3.07**	**5.48**	**6.34**	**8.81**
x_{31}	x_{32}	x_{33}	x_{34}	x_{35}	x_{36}	x_{37}	x_{38}	x_{39}	x_{40}
6.2	**4.04**	**2.86**	**1.53**	**0.702**	**1.7**	**3.72**	**5.33**	**6.27**	**3.32**

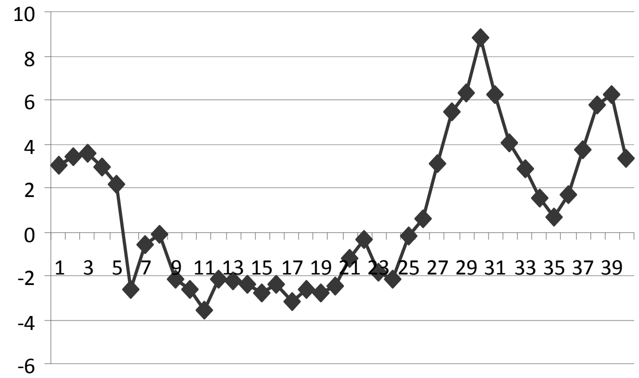

Fig. 1. Meteorological data $x_k = x(t_k)$

If we consider the given data as values of one random variable, then from the analysis of the sample $\{x_1, x_2, \ldots, x_n\}$ the hypothesis follows that this random variable ξ is distributed by Gumbel's law:

$$F_\xi(x) = \exp\left\{-\exp\left(\frac{\lambda - x}{\beta}\right)\right\},$$

with the moments:

$$E\xi = \lambda + \beta\gamma, \quad \gamma \approx 0{,}577 \text{ is a Euler's constant,}$$

$$\mathrm{E}\xi^2 = \int_0^\infty (\lambda - \beta \ln z)^2 e^{-z} dz = \frac{\pi^2}{6}\beta^2 + (\lambda + \beta\gamma)^2,$$

$$E\xi^3 = \int_0^\infty (\lambda - \beta \ln z)^3 e^{-z} dz = (\lambda + \beta\gamma)^3 + \frac{1}{2}\beta^2\pi^2(\lambda + \beta\gamma) + 2\beta^3\zeta(3),$$

$\zeta(s)$ is a Reamann zeta function, $\zeta(3) \approx 1{,}202$.

The values of sample moments:

$$\bar{x} = -0.037 \approx 0\,, \quad \overline{x^2} = \frac{1}{30}\sum_{k=1}^{30} x_k^2 = 10.19\,;$$

$$\overline{x^3} = \frac{1}{30}\sum_{k=1}^{30} x_k^3 = 33{,}76\,,$$

lead to estimates of the distribution parameters $\beta = 2{,}49$; $\lambda = -1{,}44$, so $F_\xi(x) = \exp\left\{-\exp\left(-\frac{x+1{,}44}{2{,}49}\right)\right\}$.

Density distribution graph

$$f_\xi(x) = \frac{1}{2{,}49}\ \exp\left(-\frac{x+1{,}44}{2{,}49}\right) F_\xi(x) \quad \text{shown in Fig. 2.}$$

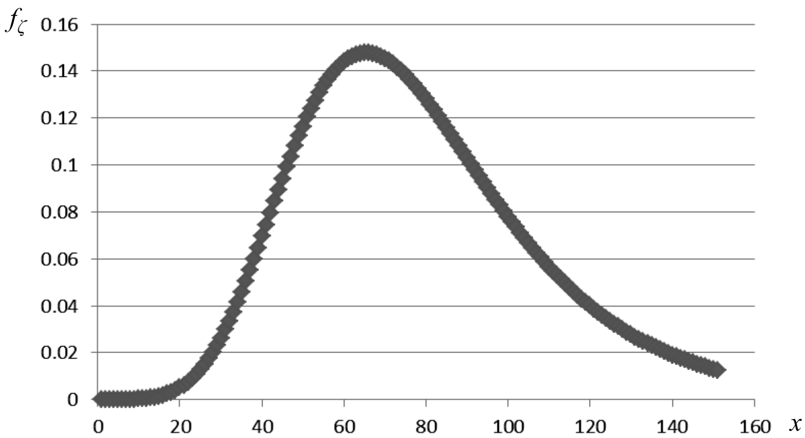

Fig. 2. Density distribution $f_\xi(x)$

Let us construct the transformed data using the formula

$$u_k = \Phi^{-1}\left(\exp\left\{-\exp\left(-\frac{x_k+1{,}44}{2{,}49}\right)\right\}\right) = \Phi^{-1}(y_k)\,,$$

where the values $\Phi^{-1}(y_k)$ are calculated using the Laplace function table. The result of the calculations is given in Table 2.

Table 2. Transformed data $u_1 - u_{40}$

u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}
1.02	**1.12**	**1.14**	**1.01**	**0.8**	**–0.84**	**–0.03**	**0.13**	**–0.63**	**–0.85**
u_{11}	u_{12}	u_{13}	u_{14}	u_{15}	u_{16}	u_{17}	u_{18}	u_{19}	u_{20}
–1.32	**–0.62**	**–0.65**	**–0.75**	**–0.92**	**–0.75**	**–1.09**	**–0.84**	**–0.92**	**–0.76**
u_{21}	u_{22}	u_{23}	u_{24}	u_{25}	u_{26}	u_{27}	u_{28}	u_{29}	u_{30}
–0.25	**0.07**	**–0.49**	**–0.64**	**0.11**	**0.37**	**1.03**	**1.56**	**1.72**	**2.13**
u_{31}	u_{32}	u_{33}	u_{34}	u_{35}	u_{36}	u_{37}	u_{38}	u_{39}	u_{40}
1.685	**–1.25**	**0.98**	**0.64**	**0.4**	**0.69**	**1.18**	**1.52**	**–1.71**	**1.09**

The evaluation of correlation coefficients

$$\rho_j = \frac{1}{30-j}\sum_{k=1}^{30-j} u_k u_{k+j}\,.$$

Leads to the results and for $k > 7$ the coefficients $\rho_k \approx 0$.

ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6	ρ_7
0.77	0.57	0.43	0.27	0.144	0.036	0.02

The elements of the matrix Q are determined by the ratios:

$$q_{jk} = \begin{cases} \rho_{|j-k|}, & |j-k| \le 7, \\ 0, & |j-k| > 7, \end{cases} \quad 1 \le j,\ k \le 37\,,\ \ \rho_0 = 1\,,\ \ a_{jk} = q_{jk}\,,\ \ 1 \le j,\ k \le 30\,.$$

The forecast of transformed data in 7 steps is calculated by the formula (2)

$$\begin{pmatrix} \hat{u}_{31} \\ \vdots \\ \hat{u}_{37} \end{pmatrix} = CA^{-1} \begin{pmatrix} u_1 \\ \vdots \\ u_{30} \end{pmatrix}.$$

The forecasting results are shown in Table 3.

Table 3. The values of forecast in 7 steps

u_{31}	u_{32}	u_{33}	u_{34}	u_{35}	u_{36}	u_{37}
1.685	**1.25**	**0.98**	**0.64**	**0.4**	**0.69**	**1.18**
Forecast						
$\hat{u}_{31}$	$\hat{u}_{32}$	$\hat{u}_{33}$	$\hat{u}_{34}$	$\hat{u}_{35}$	$\hat{u}_{36}$	$\hat{u}_{37}$
1.682	**1.155**	**0.838**	**0.443**	**0.205**	**0.6**	**0.32**
$\hat{x}_{31}$	$\hat{x}_{32}$	$\hat{x}_{33}$	$\hat{x}_{34}$	$\hat{x}_{35}$	$\hat{x}_{36}$	$\hat{x}_{37}$
6.169	**3.595**	**2.267**	**0.848**	**0.08**	**1.396**	**0.44**

where the forecast of the initial data

$\hat{x}_{31} - \hat{x}_{37}$ is calculated by the formula

$$\hat{x}_k = \varphi^{-1}(\hat{u}_k) \equiv F_\xi^{-1}(\Phi(\hat{u}_k))\,,\quad F_\xi^{-1}(y) = -2{,}49\ln(-\ln y) - 1{,}44\,.$$

Let us compare the quality of forecasting using a model, which have constructed using the normalization method and four classical discrete time series models.

Table 4. Comparison of forecast quality

Time series values $x_{31} - x_{37}$							
Actual data	x_{31}	x_{32}	x_{33}	x_{34}	x_{35}	x_{36}	x_{37}
	6.2	**4.04**	**2.86**	**1.53**	**0.702**	**1.7**	**3.72**
Forecast data	**The results of forecast**						
	$\hat{x}_{31}$	$\hat{x}_{32}$	$\hat{x}_{33}$	$\hat{x}_{34}$	$\hat{x}_{35}$	$\hat{x}_{36}$	$\hat{x}_{37}$
Normalization method	**6.169**	**3.595**	**2.267**	**0.848**	**0.08**	**1.396**	**0.44**
ARMA	7.16	9.53	9.902	7.286	9.672	10.06	9.9
Point Forecast	6.703	5.163	3.89	2,833	1.927	1.142	0.434
GARCH	8.379	8.908	7.437	7.565	7.9	8.02	9.21
ARIMA	8.652	8.063	7.44	6.85	6.328	5.808	4.35

CONCLUSIONS

Given the stationarity of the time series, modeling using the normalization method, which is defined by the relation (3), provides higher forecast quality compared to traditional forecasting methods.

REFERENCES

1. P.I. Bidyuk, O.S. Menyaylenko, O.V. Polovtsev, *Forecasting methods*. Luhansk: Alma Mater, 2008, 604 p
2. Y. Mishura, "Stochastic Calculus for Fractional Brownian Motion and Related Processes," *Lecture Notes in Mathematics 1929*, Springer 2008, 393 p. doi: 10.1007/978-3-540-75873-0
3. K. Kubilius, Y. Mishura, K. Ralchenko, *Parameter Estimation in Fractional Diffusion Models*, vol. 8, 2018, 390 p. doi: https://doi.org/10.1007/978-3-319-71030
4. V.G. Bondarenko, "On some statistics of fractional brownian motion," *System Research & Information Technologies*, no. 1, pp. 131–138, 2021. doi: https://doi.org/10.20535/SRIT.2308-8893.2021.1.11
5. I. Nourdin, *Selected Aspects of fractional Brownian motion*. Milano: Springer, 2012, 124 p. doi: https://doi.org/10.1007/978-88-470-2823-4

Received 12.04.2024

INFORMATION ON THE ARTICLE

Viktor G. Bondarenko, ORCID: 0000-0003-1663-4799, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e mail: bondarenvg@gmail.com

Valeriia V. Bondarenko, The University of the Littoral Opal Coast, France, e-mail: valeria_bondarenko@yahoo.com

ПРОГНОЗУВАННЯ ЧАСОВОГО РЯДУ ЗА МОДЕЛЛЮ НОРМАЛІЗАЦІЇ / В.Г. Бондаренко, В.В. Бондаренко

Анотація. Запропоновано емпіричні конструкції моделей часового ряду за схемою зведення первинних даних до нормально розподілених. Метою такого методу нормалізації є побудова оптимального прогнозу, який для оновлених даних є лінійним, а прогнозовані первинні дані відновлюються через обернене перетворення. Розглянуто варіанти таких перетворень — зведення первинних даних до гаусівського фрактального броунівського руху та одновимірне перетворення з використанням строго монотонної функції. Обчислювальний експеримент на базі реальних даних, що допускають стаціонарну модель, підтверджує вищу якість прогнозу методом нормалізації порівняно з традиційними моделями.

Ключові слова: оптимальний прогноз, стохастична модель, оцінювання параметрів, фрактальний броунівський рух.