ATLANTIS PRESS
Advances in Computer Science Research (ACSR), volume 72
IV International Research Conference "Information Technologies in Science, Management, Social Sphere and Medicine" (ITSMSSM 2017)

Combined Identification and Prediction Algorithms*

Yury G. Dmitriev, Gennady M. Koshkin, Vadim Yu. Lukov
Institute of Applied Mathematics and Computer Science
National Research Tomsk State University
Tomsk, Russia
dmit@mail.tsu.ru, kgm@mail.tsu.ru, v-lukov@rambler.ru

***Abstract*— In many applied problems it is required to construct a mathematical model of the dependence of output variables on input variables of the stochastic object. To solve this problem, both parametric and nonparametric approaches are used. Each of these approaches has advantages and disadvantages. In the paper, we consider combined algorithms for the identification and prediction of stochastic objects using jointly nonparametric and parametric estimates of regression.**

Keywords—Nadaraya–Watson nonparametric estimate; parametric estimate; a prior guess; regression; combined algorithm; identification; prediction; bootstrap.

I. INTRODUCTION

There are many papers on the estimation of the probability characteristics with using additional information (prior guess). Combined statistical adaptive estimators with a prior guess and their properties have been considered in [1-8]. In this paper, we consider case when there exists an assumption on the form of the estimated function. The expected form of this function we will assume as a prior guess.

Suppose that a stochastic object is described by a regression function

$$r(\vec{x}) = E\left(Y \mid \vec{X} = \vec{x}\right) = \int y \cdot p(y \mid \vec{x})dy = \frac{\int y \cdot p(x,y)dy}{p(x)}, \qquad (1)$$

where $\left(\vec{X}, Y\right) = \left(X^{\langle 1\rangle}, ..., X^{\langle p\rangle}, Y\right)$ is a (p+1)-dimensional vector of p object's inputs and output, $p(\vec{x}, y)$ is their joint distribution density, $p(\vec{x})$ is a distribution density of inputs, and $p(y \mid \vec{x})$ is the conditional distribution density.

Let there be independent observations $\left(\vec{X}_i, Y_i\right) = \left(X_i^{\langle 1\rangle}, ..., X_i^{\langle p\rangle}, Y_i\right)$, $i = 1, ..., n$, of the random vector $\left(\vec{X}, Y\right)$. Let us consider the nonparametric Nadaraya–Watson estimate [9-13] of the regression function (1)

$$\hat{r}(\vec{x}) = \hat{r}(\vec{x}; X_1, ..., X_n) = \frac{\sum_{i=1}^{n} Y_i \cdot K\left(\frac{\vec{x} - \vec{X}_i}{\vec{h}_n}\right)}{\sum_{i=1}^{n} K\left(\frac{\vec{x} - \vec{X}_i}{\vec{h}_n}\right)}, \qquad (2)$$

where $K\left(\frac{\vec{x} - \vec{X}_i}{\vec{h}_n}\right) = K\left(\frac{x_1 - X_i^{\langle 1\rangle}}{h_n^{\langle 1\rangle}}\right) ... K\left(\frac{x_p - X_i^{\langle p\rangle}}{h_n^{\langle p\rangle}}\right)$ is a p-dimensional kernel (the product of p one-dimensional kernels), $\vec{h}_n = \left(h_n^{\langle 1\rangle}, ..., h_n^{\langle p\rangle}\right)$ is a p-dimensional vector of bandwidth parameters.

Usually the researcher has some information about the nature of the dependence of the output of the object from the inputs. Suppose that he can express this knowledge in the form of a given function $\varphi\left(\vec{x}, \vec{\theta}\right)$, where $\vec{\theta} = \left(\theta^{(1)}, ..., \theta^{(s)}\right)$ is the vector of the known parameters. This type of information we call as a prior guess. Consider the task of sharing the nonparametric estimation of regression and a prior guess. The approach using combinations of different estimates was studied, for example, in [14-18].

II. COMBINED ESTIMATORS. STATIC MODEL

In the case of static models [19-21] as a combined regression estimate, we take [16-18]

$$\hat{R}_\lambda(\vec{x}) = (1 - \lambda) \cdot \hat{r}(\vec{x}) + \lambda \cdot \varphi\left(\vec{x}, \vec{\theta}\right), \qquad (3)$$

where λ is the weight coefficient determined from minimum of the criterion

$$M\left\{\hat{R}_\lambda(\vec{x}) - r(\vec{x})\right\}^2. \qquad (4)$$

So, from (4) we obtain the optimal λ

$$\lambda(\vec{x}) = \frac{M\left\{\left(\hat{r}(\vec{x}) - r(\vec{x})\right)\left(\hat{r}(\vec{x}) - \varphi\left(\vec{x}, \vec{\theta}\right)\right)\right\}}{M\left\{\hat{r}(\vec{x}) - \varphi\left(\vec{x}, \vec{\theta}\right)\right\}^2}. \qquad (5)$$

Substituting (5) into (4) and making the transformations, we get the mean square error (MSE) of the combined estimate $\hat{R}_\lambda(\vec{x})$:

$$M\left\{\hat{R}_\lambda(\vec{x}) - r(\vec{x})\right\}^2 = M\left\{\hat{r}(\vec{x}) - r(\vec{x})\right\}^2 -$$

*The reported study was funded by the TSU Competitiveness Improvement Program for 2013-2020.

$$-\frac{\left[M\left\{\left(\hat{r}(\vec{x})-r(\vec{x})\right)\cdot\left(\hat{r}(\vec{x})-\varphi\left(\vec{x},\vec{\theta}\right)\right)\right\}\right]^2}{M\left\{\hat{r}(\vec{x})-\varphi\left(\vec{x},\vec{\theta}\right)\right\}^2}. \qquad (6)$$

The second term in (6) shows how much the MSE of the combined estimate $\hat{R}_\lambda(\vec{x})$, taking into account the prior guess $\varphi\left(\vec{x},\vec{\theta}\right)$, decreases compared to MSE of $\hat{r}(\vec{x})$ for each $\vec{x}\in R^p$. Since the optimal $\lambda(\vec{x})$ (5) is usually unknown, it becomes necessary to construct an estimate $\hat{\lambda}(\vec{x})$ of this coefficient, which leads to an adaptive combined estimate

$$\hat{R}_{\hat{\lambda}}(\vec{x})=\left(1-\hat{\lambda}(\vec{x})\right)\hat{r}(\vec{x})+\hat{\lambda}(\vec{x})\varphi\left(\vec{x},\vec{\theta}\right). \qquad (7)$$

Let us consider an estimate of a weight coefficient by a bootstrap method. We write (5) in the form:

$$\lambda(\vec{x})=\frac{M\psi_1(\vec{x})}{M\psi_2(\vec{x})}, \qquad (8)$$

where

$$M\psi_1(\vec{x})=M\left\{\left(\hat{r}(\vec{x})-r(\vec{x})\right)\left(\hat{r}(\vec{x})-\varphi\left(\vec{x},\vec{\theta}\right)\right)\right\},$$

$$M\psi_2(\vec{x})=M\left\{\hat{r}(\vec{x})-\varphi\left(\vec{x},\vec{\theta}\right)\right\}^2.$$

Generate a bootstrap sample $\left(\vec{X}_j^*,Y_j^*\right)$, $\vec{X}_j^*=\left(\vec{X}_{1,j}^*,...,\vec{X}_{n,j}^*\right)$, $j=1,...,B$, for the numerator and denominator in (8). Then we have:

$$M\psi_1(\vec{x})\approx\frac{1}{B}\cdot\sum_{j=1}^{B}\left\{\left(r\left(\vec{x};\vec{X}_j^*\right)-\hat{r}(\vec{x})\right)\cdot\left(r(\vec{x};\vec{X}_j^*)-\varphi\left(\vec{x};\vec{\theta}\right)\right)\right\},$$

$$M\psi_2(\vec{x})\approx\frac{1}{B}\cdot\sum_{j=1}^{B}\left\{r(\vec{x};\vec{X}_j^*)-\varphi\left(\vec{x};\vec{\theta}\right)\right\}^2.$$

As a result, we obtain the following estimate of the weight coefficient (5):

$$\hat{\lambda}_B(\vec{x})=\frac{\sum_{j=1}^{B}\left(\hat{r}(\vec{x};\vec{X}_j^*)-\hat{r}(\vec{x})\right)\left(\hat{r}(\vec{x};\vec{X}_j^*)-\varphi\left(\vec{x},\vec{\theta}\right)\right)}{\sum_{j=1}^{B}\left[\hat{r}(\vec{x};\vec{X}_j^*)-\varphi\left(\vec{x},\vec{\theta}\right)\right]^2}, \qquad (9)$$

The usage of (9) in (7) leads to an adaptive combined estimate

$$\hat{R}_{\hat{\lambda}_B}(\vec{x})=\left(1-\hat{\lambda}_B(\vec{x})\right)\hat{r}(\vec{x})+\hat{\lambda}_B(\vec{x})\varphi\left(\vec{x},\vec{\theta}\right). \qquad (10)$$

If θ is evaluated by a sample, then the estimate (10) we will denote as $\tilde{R}_{\hat{\lambda}_B}(\vec{x})$.

III. Combined Estimators. Dynamic Model

Consider the dynamic model (cf. [22-27])

$$Y_t=f\left(Y_{t-1},...,Y_{t-s},X_{t-1},...,X_{t-q}\right)+\xi_t, \qquad (11)$$

where Y_t is the output of the object at the time moment *t*, X_t is the value of exogenous factor at the moment *t*, f is an unknown function, ξ_t is the sequence of the i.i.d. random variables with a nonnegative distribution density, and $E\xi_t=0,\ E\xi_t^{\,2}<\infty,\ E\xi_t^{\,3}=0,\ E\xi_t^{\,4}<\infty$ (see [22]).

Assume that f is bounded and its form does not change in the time interval under study. As an prior guess about the form of f, take the function $\varphi\left(x,\vec{\theta}\right)$ and consider the following combined adaptive estimate:

$$\hat{R}\left(\vec{y}_t,\vec{x}_t\right)=\left(1-\hat{\lambda}_B(\vec{y}_t,\vec{x}_t)\right)\hat{r}\left(\vec{y}_t,\vec{x}_t\right)+\hat{\lambda}_B(\vec{y}_t,\vec{x}_t)\,\varphi\left(\vec{y}_t,\vec{x}_t;\vec{\theta}\right), \qquad (12)$$

where

$$\hat{\lambda}_B(\vec{y}_t,\vec{x}_t)=$$

$$\frac{\sum_{j=1}^{B}\left(\hat{r}(\vec{y}_t,\vec{x}_t;\vec{Y}_j^*,\vec{X}_j^*)-\hat{r}(\vec{y}_t,\vec{x}_t)\right)\left(\hat{r}(\vec{y}_t,\vec{x}_t;\vec{Y}_j^*,\vec{X}_j^*)-\varphi\left(\vec{y}_t,\vec{x}_t;\vec{\theta}\right)\right)}{\sum_{j=1}^{B}\left[r(\vec{y}_t,\vec{x}_t;\vec{Y}_j^*,\vec{X}_j^*)-\varphi\left(\vec{y}_t,\vec{x}_t;\vec{\theta}\right)\right]^2}.$$

The estimate (12) is applied in section 4 for the analysis of stock prices on real data.

IV. Analysis of Real Data

The analysis of the prices of Gazprom's stocks for 2016 is carried out on the basis of the (11) with $s=1,\ q=1$:

$$Y_t=f\left(Y_{t-1},X_{t-1}\right)+\xi_t, \qquad (13)$$

where $t=1,...,n,$ Y_t is the stock price at the time moment *t*. We take the parametric function in the form $\varphi(X_{t-1};\theta^{(1)},\theta^{(2)})=\theta^{(1)}+\theta^{(2)}X_{t-1}$, where for simplicity we set $\theta^{(1)}=0,\ \theta^{(2)}=2,$ i.e. $\varphi(X_{t-1};0,2)=2X_{t-1}$.

As the nonparametric estimate of the interpolation forecast for Y_t, we take the following modification of estimate (2):

$$\hat{Y}_t=\hat{r}\left(Y_{t-1},X_{t-1}\right)=$$

$$=\frac{\sum_{j\ge 2,\,j\ne t}Y_tK\left(\frac{Y_{t-1}-Y_{j-1}}{h_n^{\langle y\rangle}}\right)K\left(\frac{X_{t-1}-X_{j-1}}{h_n^{\langle x\rangle}}\right)}{\sum_{j\ge 2,\,j\ne t}K\left(\frac{Y_{t-1}-Y_{j-1}}{h_n^{\langle y\rangle}}\right)K\left(\frac{X_{t-1}-X_{j-1}}{h_n^{\langle x\rangle}}\right)}. \qquad (14)$$

The combined estimate, for which $\varphi(X_{t-1};0,2)=2X_{t-1}$, takes the form

$$\bar{Y}_t = \hat{R}_{\hat{\lambda}_B}\left(Y_{t-1},X_{t-1}\right)=$$

$$\left(1-\hat{\lambda}_B(Y_{t-1},X_{t-1})\right)\hat{r}\left(Y_{t-1},X_{t-1}\right)+2\hat{\lambda}_B(Y_{t-1},X_{t-1})X_{t-1}, \quad (15)$$

where

$$\hat{\lambda}_B(Y_{t-1},X_{t-1})=$$

$$=\frac{\sum_{j=1}^{B}\left(\hat{r}(Y_{t-1},X_{t-1};Y^*_{j,1},\dots,Y^*_{j,n-1})-\hat{r}(Y_{t-1},X_{t-1})\right)\left(\hat{r}(Y_{t-1};Y^*_{j,1},\dots,Y^*_{j,n-1})-2X_{t-1}\right)}{\sum_{j=1}^{B}\left[r(Y_{t-1},X_{t-1};Y^*_{j,1},\dots,Y^*_{j,n-1})-2X_{t-1}\right]^2}.$$

Based on the prices of stocks $Y_1,\dots,Y_n$, formulas (14) and (15), the estimates of forecasts by one step for price Y_{n+1} are defined as follows:

$$\hat{Y}_{n+1}=\hat{r}\left(Y_n,X_n;\right)=$$

$$=\frac{\sum_{i=2}^{n}Y_iK\left(\frac{Y_n-Y_{i-1}}{h_n^{\langle y\rangle}}\right)K\left(\frac{X_n-X_{i-1}}{h_n^{\langle x\rangle}}\right)}{\sum_{i=2}^{n}K\left(\frac{Y_n-Y_{i-1}}{h_n^{\langle y\rangle}}\right)K\left(\frac{X_n-X_{i-1}}{h_n^{\langle x\rangle}}\right)}, \quad (16)$$

$$\bar{Y}_{n+1}=\hat{R}_{\hat{\lambda}_B}\left(Y_n,X_n\right)=$$

$$\left(1-\hat{\lambda}_B(Y_n,X_n)\right)\hat{r}\left(Y_n,X_n\right)+2\hat{\lambda}_B(Y_n,X_n)X_n, \quad (17)$$

where

$$\hat{\lambda}_B(Y_n)=\frac{\sum_{j=1}^{B}\left(\hat{r}(Y_n;Y^*_{j,1},\dots,Y^*_{j,n-1})-\hat{r}(Y_n)\right)\left(\hat{r}(Y_n;Y^*_{j,1},\dots,Y^*_{j,n-1})-2X_n\right)}{\sum_{j=1}^{B}\left[r(Y_n;Y^*_{j,1},\dots,Y^*_{j,n-1})-2X_n\right]^2}.$$

Let there be $n+L$ stock prices. Estimates of forecasts $\hat{Y}_{n+2}$ and $\bar{Y}_{n+2}$ will be constructed at n prices $Y_3,\dots,Y_{n+1}$ by formulas (16) and (17). Similarly, at n prices, shifting by the required number of steps, make forecasts $\hat{Y}_{n+3},\dots,\hat{Y}_{n+L}$ and $\bar{Y}_{n+3},\dots,\bar{Y}_{n+L}$.

The quality of identification and forecasting will be characterized by means of the average relative errors $\delta_{real}\left(\hat{r}\right)$ and $\eta_{real}\left(\hat{r}\right)$:

$$\delta_{real}\left(\hat{r}\right)=\frac{1}{n-1}\sum_{i=2}^{n}\frac{\left|Y_i-\hat{r}(Y_{i-1},X_{i-1})\right|}{Y_i}\cdot 100\%,$$

$$\eta_{real}\left(\hat{r}\right)=\frac{1}{L}\sum_{i=n+1}^{n+L}\frac{\left|Y_i-\hat{r}(Y_{i-1},X_{i-1})\right|}{Y_i}\cdot 100\%.$$

Consider the case n = 100. In Fig. 1, there are presented the results of identification and prediction for the combined model using estimates (15) and (17), and the behavior of their weight coefficients are shown in Fig. 2.

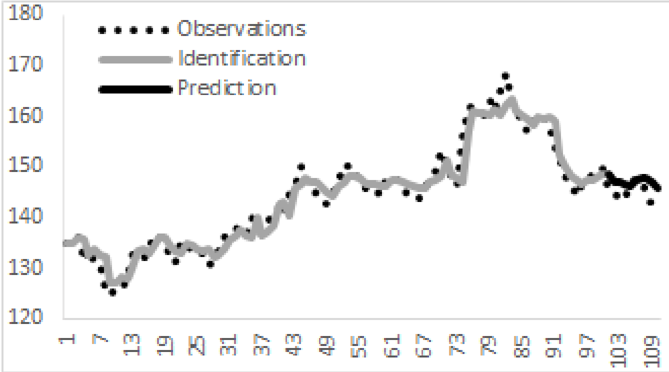

Fig. 1. Identification and forecasting using combined estimates (15) and (17)

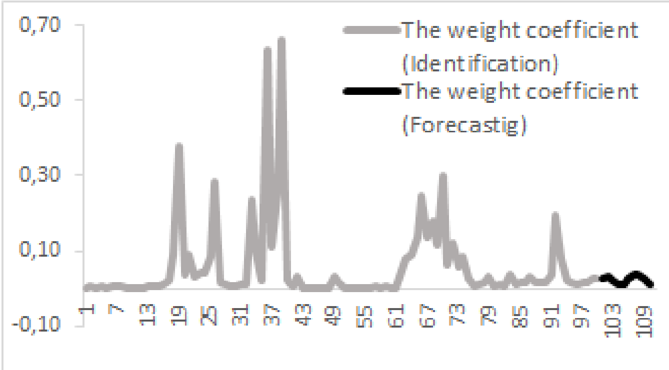

Fig. 2. Plot of the dependence of the estimates of the weight coefficients for identification and forecasting

Let $\varphi(X_{t-1};\hat{\theta}^{(1)},\hat{\theta}^{(2)})=\hat{\theta}^{(1)}+\hat{\theta}^{(2)}X_{t-1}$, where $\hat{\theta}^{(1)},\hat{\theta}^{(2)}$ are the least square method (LSM) estimates. For this case, we denote estimators (10) like $\tilde{R}$.

Table I gives average relative errors of identification and prediction for different volumes of observations.

TABLE I. AVERAGE RELATIVE ERRORS OF IDENTIFICATION AND PREDICTION

Average errors	**Volumes of observations**		
	10	50	100
$\delta_{real}\left(\hat{r}\right)$	1,84	1,40	1,36
$\delta_{real}\left(\hat{R}\right)$	1,31	1,23	1,23
$\delta_{real}\left(\tilde{R}\right)$	1,40	1,25	1,24
$\eta_{real}\left(\hat{r}\right)$	1,94	1,27	1,69
$\eta_{real}\left(\hat{R}\right)$	1,70	1,17	1,65
$\eta_{real}\left(\tilde{R}\right)$	1,75	1,24	1,66

From the results obtained, in practice it is preferable using the combined evaluation in comparison with the nonparametric estimate, especially in the case of small sample sizes.

V. Conclusion

In this paper, the problem of identification of a stochastic object by means of a combined estimate is considered, which is a weighted sum of the nonparametric estimate of the regression and some function given by the researcher. Adaptive combined estimates are constructed on the basis of which algorithms for predicting static and dynamic objects are proposed.

Based on the results of numerical simulation, the advantage of adaptive combined estimates is shown in comparison with nonparametric regression estimates for small samples sizes and a large noise level. The expediency of applying the proposed approach in practice is illustrated in the analysis of the prices of Gazprom's stocks for 2016.

References

[1] Y.G. Dmitriev and P.F. Tarasenko, “The Use of a Priori Information in the Statistical Processing of Experimental Data.” Russian Physics Journal, 35 (9), 1992, pp. 888-893.

[2] S.S. Tarima and D. Pavlov, “Using auxiliary information in statistical function estimation.” ESAIM: Probability and Statistics, 10, 2006, pp. 11-23.

[3] S.S. Tarima and Y.G. Dmitriev, “Statistical estimation with possibly incorrect model assumptions.” The Bulletin of Tomsk State University : control, computing, informatics, 3(8), 2009, pp, 87-99.

[4] Sergey Tarima, Albert Vexler and Singh Siddhartha, “Robust mean estimation under a possibly incorrect log-normality assumption.” Communications in Statistics - Simulation and Computation, 42, 2013, pp. 316-326.

[5] Yuri Dmitriev, Peter Tarassenko and Yuri Ustinov, “On Estimation of Linear Functional by Utilizing a Prior Guess.” A. Dudin et al. (Eds.): ITMM 2014, CCIS 487, 2014, pp. 82–90,

[6] Yuri Dmitriev and Peter Tarassenko, “On Adaptive Estimation Using a Prior Guess.” Proceedings The International Workshop, Applied Methods of Statistical Analysis. Nonparametric Approach. Novosibirsk, Russia. Novosibirsk, 14-15 September, 2015: NSTU publisher, 2015, pp. 49-55,

[7] Yu. Dmitriev, F.Tarasenko and P.Tarasenko, “On improving statistical estimation by utilizing collateral (“guesses”): A case of the probability estimation.” Proceedings of the International Workshop, Applied Methods of Statistical Analysis. Nonparametric Methods in Cybernetics and System Analysis.- AMSA -2017. Krasnoyarsk, Russia, 18-22 September, 2017: NSTU publisher, 2017, pp. 262-269.

[8] Sergey Tarima, “Statistical estimation in the presence of possibly incorrect model assumptions.” Journal of Statistical Theory and Practic, 11(3), 2017, pp. 449-467.

[9] E, Nadaraya, “On Estimating of Regression.” Theory Probab. Appl. Vol. 9, No. 1, 1964, pp. 141–142.

[10] G.S. Watson, “Smooth Regression Analysis.” Sankhya. Indian J. Statist. Vol. A26, 1964, pp. 359–372.

[11] G.M. Koshkin, “Approach to the investigation of functionals of conditional distributions under statistical uncertainty.” Automat. and Remote Control, vol. 39, No. 8, 1978, pp. 1141-1151.

[12] G.M. Koshkin, ”Asymptotic properties of functions of statistics and their application to nonparametric estimation.” Automat. and Remote Control, vol. 51, No. 3, 1990, pp. 345--357.

[13] E.F. Camacho and C. Bordons, “Model Predictive Control”. London: Springer-Verlag, 2004.

[14] Yu.G. Dmitriev and G.M. Koshkin, "On the Use of a Priori Information in Nonparametric Regression Estimation". IFAC Proceedings Series., 1987, vol. 2, pp. 223– 228.

[15] Yu.G. Dmitriev and G.M. Koshkin, “Using Additional Information in Nonparametric Estimation of Density Functionals”. Automat. and Remote Control, vol. 48, No. 10, 1987, pp. 1307–1316.

[16] S.V. Skripin, “Properties of a Combined Regression Estimate for Finite Sample Sizes.” Bulletin of the Tomsk Polytechnic University, vol. 313, No. 5, 2008, pp. 10–14. (in Russian).

[17] Yury G. Dmitriev, Gennady M. Koshkin and Vadim Yu. Lukov, “Combined Identification Algorithms”. Applied Methods of Statistical Analysis. Nonparametric Methods in Cybernetics and System Analysis - AMSA’2017, Krasnoyarsk, Russia, 18-22 September, 2017: Proceedings of the International Workshop. – Novosibirsk: NSTU publisher, 2017, pp. 19-27.

[18] Yu. Dmitriev and G. Koshkin, “On Distribution Functionals Estimation with Auxiliary Information.” Proceedings of the International Workshop, Applied Methods of Statistical Analysis. Nonparametric Methods in Cybernetics and System Analysis - AMSA -2017. Krasnoyarsk, Russia, 18-22 September, 2017: NSTU publisher, 2017, pp. 9-18.

[19] Gennady Koshkin and Anna Kitayeva, “Nonparametric Identification of the Production Functions”. Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering 2011, WCE 2011, (6-8 July, 2011, Imperial College London, London, U.K.). Hong Kong: Newswood Limited. 2011, vol. 1, pp. 276-280.

[20] A.V. Kitayeva and G.M. Koshkin, “Nonparametric semirecursive identification in a wide sense of strong mixing processes.” Problems Inform. Trans., 2010, vol. 46, No. 1, pp. 22-37.

[21] A.V. Kitaeva and G.M. Koshkin. “Recurrent Nonparametric Estimation of Functions from Multidimensional Functional Density and their Derivatives”. Automat. and Remote Control, vol. 70, No. 3, 2009, pp. 389–407.

[22] A.V. Kitayeva and G.M. Koshkin, “Semi-recursive nonparametric identification in the general sense of a nonlinear heteroscedastic autoregression.” Automat. and Remote Control, 2010, vol. 71, No. 2, pp. 257-274.

[23] Gennady Koshkin and Anna Kitayeva, “Nonparametric Identification of Static and Dynamic Production Functions”. IAENG International Journal of Applied Mathematics, 2011, vol. 41. Issue 3, pp. 228-234.

[24] Gennady Koshkin and Anna Kitayeva, “Nonparametric Identification of the Production Functions”. Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering 2011, WCE 2011, (6-8 July, 2011, Imperial College London, London, U.K.). Hong Kong: Newswood Limited. Vol. 1, 2011, pp. 276-280.

[25] G.M.Koshkin, V.Yu. Lukov and I.G. Piven, “Nonparametric Algorithms of Identification and Prediction in the ARX-Models.” Proceedings. The Second International Symposium on Stochastic Models, in Reliability Engineering, Life Science, and Operations Management. Beer Sheva, Israel. Conference Publishing Services The Institute of Electrical and Electronics Engineers, 2016, pp. 620–623.

[26] G.M. Koshkin and F.P. Tarasenko, “Nonparametric Algorithms for Identifying and Control of Continuous-discrete Stochastic Objects”. 8-th IFAC-IFORS Symposium on Identification and System Parameter Estimation. Beijing. Pargamon Press, No. 2, 1988, pp. 882–887.

[27] V.A. Vasiliev and G.M. Koshkin, “Nonparametric Identification of Autoregressions.” Theory Probab. Appl. vol. 43, No.3, 1998, pp. 507–517.