

# Application of methods of neural networks in the analysis of the market of mortgage housing loan

Aleksey V. Burkov<sup>1</sup>, Anastasiya Yu. Cheremunina<sup>2</sup>, Tatiana V. Yalyalieva<sup>3</sup>, Elena A. Murzina<sup>4</sup>

<sup>1</sup>Department of Applied Statistics and Informatics, Mari State University, Yoshkar-Ola, Russia, Moscow  
alexey.burkov@gmail.com

<sup>2</sup>Department of Applied Statistics and Informatics, Mari State University, Yoshkar-Ola, Russia  
psheni4nov@yandex.ru

<sup>3</sup>Department of Management and Law, Volga State University of Technology, Yoshkar-Ola Russia  
yal05@mail.ru

<sup>4</sup>Department of Management and Law, Volga State University of Technology, Yoshkar-Ola Russia

## ABSTRACT

The article analyzes the state of the mortgage lending market in the Russian Federation. The classification of regions by indicators characterizing the mortgage lending market is given. The factors affecting the state of the mortgage lending market are studied. A forecast for 12 months is proposed. The methods of neural networks were used as the main analysis methods. In particular, the paper presents examples of the use of neural networks for classifying, evaluating factors, and constructing regression models. The choice of this technology is due to the fact that over the past few years, interest in the use of neural networks in the study of statistical data, analysis and forecasting has increased significantly. Such popularity is due to the fact that neural networks significantly complement the usual traditional statistical methods of data analysis, which can be used to build complex nonlinear models and dependencies. The article is intended for researchers in the mortgage market who want to get acquainted with the use of neural network technology for analyzing the mortgage market.

**Key words** :multidimensional regression models, artificial neural network technologies, radially basic neural network, linear neural network, multilayer perceptron

## 1. INTRODUCTION

Russia's transition to a market economy has exacerbated the housing problem. The priority in public policy was the development of mortgage lending. Thanks to which she received much greater social support. In addition, mortgage lending is inextricably linked with important social functions, primarily with providing citizens of the country with quality and affordable housing and improving housing conditions. The real estate market is developing by increasing the volume of housing construction and improving the conditions for the

growth of solvent population. But despite the fact that there are a lot of state programs for the development of housing, without taking additional measures, these housing and social problems will be impossible to solve.

## 2. RESEARCH METHODOLOGY

A study of scientific articles revealed that quite a lot of attention is paid to the housing mortgage lending market. Among the authors touching upon this topic, one can name Patlatenko A.A. [6], Kochetkova A.A [7]., Abdullaeva M.A [8]., Orlova A.A. [8], Rimsky L.A.[8], Khovalyg N.N., Aizinova I.M.[12], Makeykinina C.M [11]. and others [13, 14, 15].

The theoretical basis of the study was the work of leading domestic and foreign scientists on the study of housing mortgage lending, Bank of Russia regulations and collections of the Federal State Statistics Service.

Over the past few years, interest in the use of neural networks in the study of statistical data, analysis and forecasting has increased significantly. Such popularity is justified by the fact that neural networks significantly complement the usual traditional statistical methods of data analysis, which can be used to build complex nonlinear models and dependencies [1, 2, 3]. For a deeper study, we used the classification of subjects into typological groups using neural networks. The advantage of neural networks relative to previously used statistical methods of analysis is the fact that they are based on parallel processing of information and, moreover, allow you to obtain a reasonable result, in the form of data not found in the learning process, as they differ in their ability to self-learn. In this connection, neural networks have the ability to solve intractable, large-scale tasks.

Neural networks are a new and very promising computing technology that provides new approaches to the study of dynamic problems in the financial field. The ability to model non-linear processes, work with noisy data and adaptability make it possible to use neural networks to solve a wide class of financial problems [2, 4].

An artificial neural network is a set of neurons interconnected.

A neuron is a special nerve cell capable of perceiving, transforming and distributing signals. As a rule, the transfer, activation functions of all neurons in the network are fixed, and the weights are network parameters and can vary. Some neuron inputs are labeled as external network inputs, and some outputs are labeled external network outputs. Neurons are divided into three types according to the functions performed by them on the network. Input neurons (neurons of the input layer) receive data from the external environment and in a certain way distribute them further on the network. The intermediate neurons (neurons of the hidden layer) are assigned the role of the main participants in the process of solving the problem. Output neurons (neurons of the output layer) transmit the results of the network to the external environment (to the consumer) [5].

The main goal of creating a neural network is to classify entities according to the volume of mortgage lending. The training series is an initial series of observations indicating input and output variables. The categorical output variable in this case will be the cluster number obtained by cluster analysis by the usual statistical method. Continuous output values selected a resultant indicator characterizing the volume of mortgage lending of subjects, and 3 factor signs identified as a result of the principal component method, namely, factors characterizing the demographic status of the subject, describing the financial situation of the population in the subject and combining bank interest rates on mortgage lending.

Classification building problems can be solved using various types of networks, shown in Figure 1.

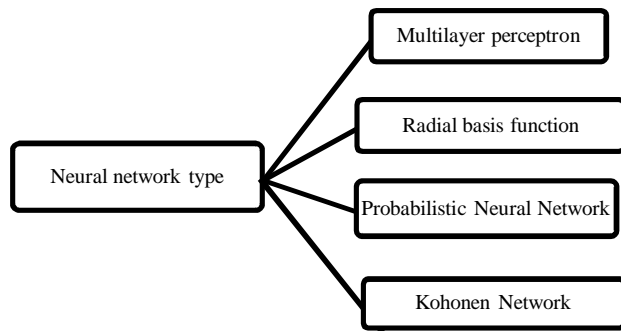


Figure 1: Topology of neural networks

### 3. RESULTS

To solve the problem of classifying entities according to indicators characterizing the volume of mortgage lending, we used networks based on the radial basis function and a multilayer perceptron, since they are an example of the solution to the problem of a nonlinear multilayer direct distribution network. The number of hidden layers in these networks corresponds to 1. It is important to select the correct number of elements in the intermediate layer. The required number of synaptic scales for calculating the range of elements in the intermediate layer is calculated by the formula 1.

$$\frac{N_y N_x}{1 + \log_2(N_p)} \leq N_w \leq N_y \left( \frac{N_p}{N_x} + 1 \right) (N_x + N_y + 1) + N_y, \tag{1}$$

where  $N_y$  – is the dimension of the output signal;

$N_x$  – is the dimension of the input signal;

$N_p$  – the number of elements of the training sample;

$N_w$  – the required number of synaptic weights.

Thus,  $111 \leq N_w \leq 769$ .

Using the identified number of synaptic weights, the range of the number of elements in the intermediate layer is calculated by the formula 2.

$$N = \frac{N_w}{N_x + N_y} \tag{2}$$

As a result, a range of the number of elements in the intermediate layer was obtained  $14 \leq N \leq 96$ .

Next, a network was constructed using minimization of the squared error. The duality method was used, in particular the error back propagation method, which is one of the gradient descent algorithms. During experiments with the type of neural network and the number of elements in the intermediate layers, the best models were identified.

Out of 9 evaluated models of the multilayer perceptron type and radial basis function, was chosen the best model, since among the identified models it has the best performance and the smallest construction errors. The architecture of the identified network has the parameters MLP 4:4-35-4:1, thus, the network is a multilayer perceptron type, has 4 input variables, 35 elements in the intermediate layer and 4 output variables.

The architecture of the constructed MLP 4:4-35-4:1 network is shown in Figure 2.

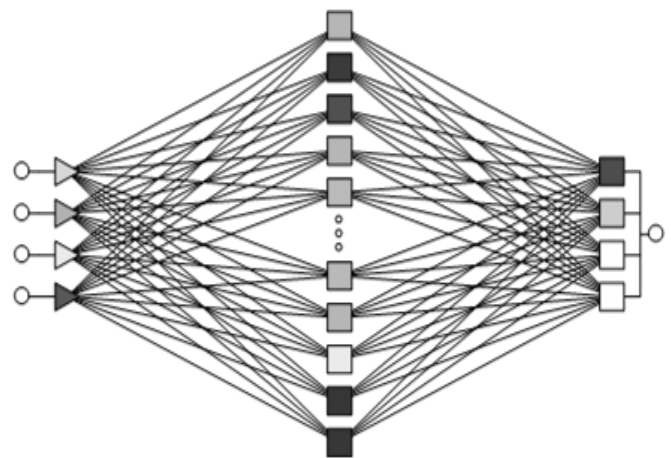


Figure 2: Architecture of the constructed neural network MLP 4:4-35-4:1

This neural network determined the correct and incorrect distribution of subjects across clusters. The classification results are presented in table 1.

**Table 1:** Classification results

	cluster 1	cluster 2	cluster 3	cluster 4
Total	17	12	31	21
Right	15	12	26	20
Wrong	2	0	5	1
Unknown	0	0	0	0
Right (%)	88,24	100,00	83,87	95,24
Erroneous (%)	11,76	0,00	16,13	4,76
Unknown (%)	0	0	0	0

It can be noted that the percentage of correctly classified entities is rather high in each of the clusters: 88.24%, 100%, 83.87%, 95.24%, respectively, according to the cluster number. The percentage of erroneously classified was 9.88%, that is, 8 of the total number of subjects. There are no unclassified entities.

Table 2 shows the matrix of classification errors for correctly assigning subjects to their clusters. It reflects information on the classification results and the classification of an element to a particular cluster. The columns of the matrix represent the source classes, the rows predicted.

**Table 2:** Matrix of classification errors

	cluster 1	cluster 2	cluster 3	cluster 4
cluster 1	15	0	0	1
cluster 2	0	12	0	0
cluster 3	2	0	26	0
cluster 4	0	0	5	20

Let us consider the contents of clusters by subjects of the Russian Federation:

The first cluster contains 16 constituent entities of the Russian Federation: Voronezh Oblast, Leningrad Oblast, Volgograd Oblast, Republic of Dagestan, Stavropol Territory, Udmurt Republic, Orenburg Oblast, Saratov Oblast, Yamalo-Nenets Autonomous Okrug, Altai Krai, Irkutsk Oblast, Kemerovo Oblast, Omsk Region, Republic of Sakha (Yakutia), Primorsky Territory, Khabarovsk Territory.

The second cluster contains 12 entities; it remained unchanged: Krasnodar Territory, Rostov Region, Republic of Bashkortostan, Republic of Tatarstan, Perm Territory, Nizhny Novgorod Region, Samara Region, Sverdlovsk Region, Khanty-Mansi Autonomous Area, Chelyabinsk Region, Krasnodar Territory, Novosibirsk Region.

The third cluster has replaced most of the subjects and now contains 28: Belgorod region, Bryansk region, Vladimir region, Ivanovo region, Kaluga region, Kursk region, Lipetsk region, Ryazan region, Smolensk region, Tambov region, Tver region, Tula region, Yaroslavl region, Komi Republic, Arkhangelsk Region, Vologda Region, Kaliningrad Region, Murmansk Region, Astrakhan Region, Chuvash Republic, Kirov Region, Penza Region, Ulyanovsk Region, Republic of Buryatia, Trans-Baikal Territory, Tomsk Region, Amur Region, Sakhalin Region.

The fourth cluster includes 25 constituent entities of the Russian Federation: Kostroma Oblast, Oryol Oblast, Republic of Karelia, Nenets Autonomous Okrug, Novgorod Oblast, Pskov Oblast, Republic of Adygea, Republic of Kalmykia, Republic of Crimea, Sevastopol, Republic of Ingushetia, Kabardino-Balkaria, Karachay-Cherkess Republic, Republic of North Ossetia, Chechen Republic, Republic of Mari El, Republic of Mordovia, Kurgan Region, Republic of Altai, Republic of Tuva, Republic of Khakassia, Kamchatka Territory, Magadan Region, Jewish Autonomous Region, Chukotka Autonomous Region.

Visually, the distribution of subjects by clusters is seen in Figure 3



**Figure 3:** Visual distribution of the constituent entities of the Russian Federation by cluster

Next, we consider the contents of clusters by federal district: The first cluster united regions from all federal districts: the Siberian Federal District (4), the Far Eastern Federal District (3), the Volga Federal District (3), the Southern Federal District (3) and the Central, Northwest and Ural Federal Districts, one at a time subject.

The second cluster contains only four federal districts: the Volga Federal District (5), the Ural Federal District (3), the Siberian (2) and Southern (2) federal districts.

Most of the subjects of the third cluster are from the Central Federal District (15), the Northwest District (5) and the Volga Federal District (4). And there are also subjects from the Siberian Federal District (3), the Far Eastern District (2) and the Southern Federal District (1). Only the subjects of the Ural Federal District were not included in this cluster.

In the fourth cluster, subjects from all federal districts are also found, most of the subjects from the Southern Federal District (9), 4 from each of the North-West and Far Eastern Federal Districts, the least from the Siberian Federal District (3), Central, Ural and Volga region - 2 subjects each.

Sensitivity analysis makes it possible to identify the significance of the selected variables for compiling a specific neural network.

**Table 3:** Sensitivity analysis

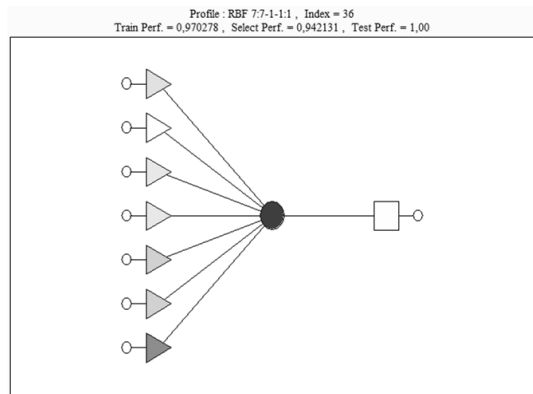
	Volume of mortgage housing lending	factor 1	factor 2	factor 3
Attitude	3,58	2,85	1,89	2,02
Rank	1	2	4	3

Of great importance for building a network, as expected, is the volume of mortgage lending, the second most important factor is the factor characterizing the demographic state of the subject. The least importance in constructing the model is a variable that characterizes the financial situation of the population of the subject.

Let's move on to the construction of regression models. To build a regression model of the mortgage lending market, the neural network method was used. Based on the previously considered factors, 20 models of each type of neural networks were constructed, where the volume of mortgage lending was selected as an effective attribute.

The highest quality models of each type are analyzed. Linear (Linear) and generalized regression neural networks (GRNN) showed rather weak results. The best results were obtained using models of radial basis functions (RBF).

The best neural network of the radial basis function RBF 7:7-1-1:1 is the highest quality (Figure 4). The proportion of correctly predicted values on the test sample was 100%, on the training sample 97%, and on the control 94%. But, despite the high performance, the errors of this neural network are slightly higher (0.02) than that of other analyzed networks.



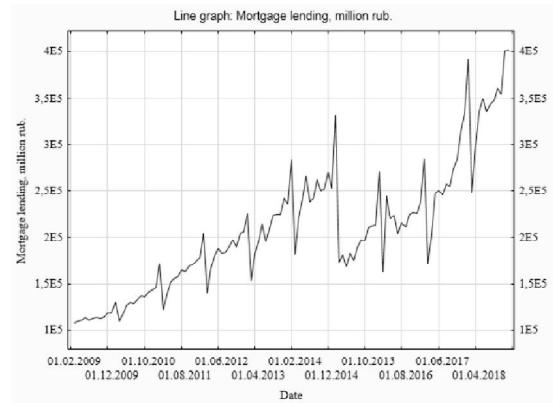
**Figure 4:** Neural network architecture type RBF 7:7-1-1:1

The sensitivity of the neural network showed that in this model,  $Y_2$  turned out to be the most important indicator - the number of mortgage housing loans and factor  $X_6$  the number of employed population is less significant.

A neural network of the RBF 7: 7-1-1: 1 type can be considered high-quality, since the histogram of residues gives values similar to the normal distribution. In addition, the correlation between the observed and predicted values is greatest (0.89), which indicates the adequacy of the model.

Dale was forecasting the time series of the volume of mortgage lending using neural networks. The main role of the neural network in this task is the preparation and prediction of the forecast for its previous behavior. For a more accurate forecast, a more extensive amount of data is taken.

There are monthly data on the volume of mortgage housing lending, measured in million rubles. from 2009 to 2018 inclusive. The dynamics of changes in this indicator using the line graph (Figure 5).

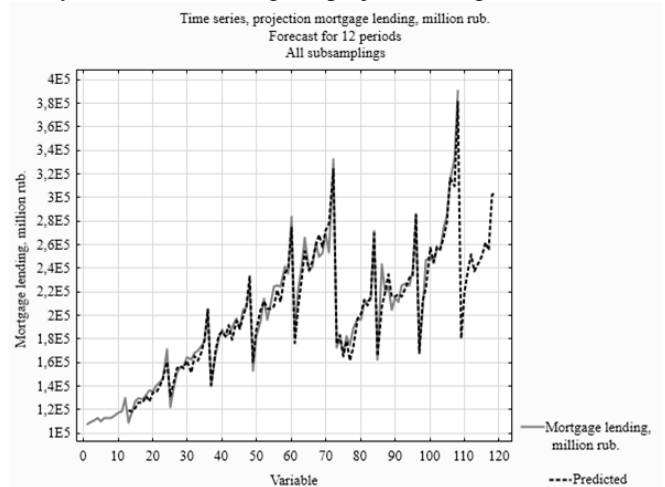


**Figure 5:** Dynamics of the volume of mortgage housing lending for 2009-2018

For forecasting, only one variable is used, which is both input and output. Primary training of the neural network was carried out without the last 12 observations in order to further use them for cross-checking the neural network. Therefore, the test sample is also excluded. The training and control samples are taken in a ratio of 70/30%. A multilayer perceptron is chosen as the network type, since a trend is observed in this network. As the default error function, the sum of the squares is always used.

After analyzing the table, you can see that the performance of trained networks is quite high. The first three models have training productivity and control performance at approximately the same level, therefore, these neural networks work quite evenly, both in training and in control. The first network with MLP 12-7-1 architecture was chosen as the best, with 12 input neurons and 7 layers at a hidden level, due to the fact that its performance is the highest. The functions of its activation on hidden and input neurons are logical and identical, respectively. For further quality control, a histogram of residuals was constructed for the control and training samples.

The possibility of forecasting the volume of mortgage lending for this network has been investigated. The projection for the next year is derived using the projection (Figure 6).

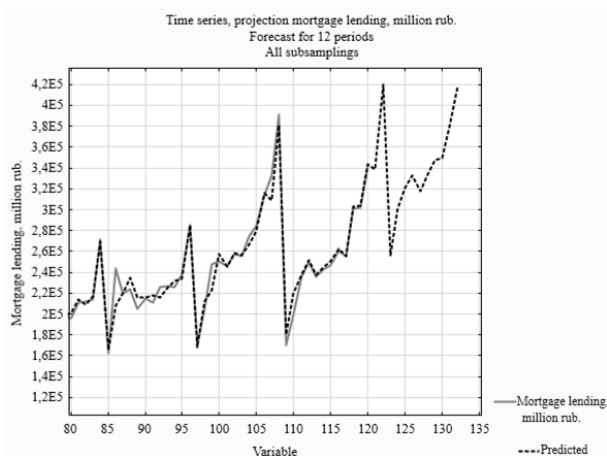


**Figure 6:** Projection for forecasting the volume of mortgage lending

The calculated projection visually resembles previous network behavior, but verification of this forecast is necessary. The predicted data were checked against a test sample that was skipped earlier, i.e. 12 observations for the last 2018 year.

The relative error in this case is 2.91, which is good enough. Thus, it is possible to build a further forecast using this neural network.

Visually, the forecast is presented in Figure 7.



**Figure 7:** Forecast of mortgage housing lending for 2019

#### 4. CONCLUSIONS

In the course of the study of the demographic situation in the Volga region in the context of the problems of economic clustering is the main objective factor affecting the balance of socio-economic development of this macro region, a significant differentiation of regions by population was revealed (the population of the Republic of Bashkortostan is 5.9 larger than the Republic of Mari El). This situation is exacerbated by the marked tendency to the decrease of population in all regions of the Volga Federal district with the exception of the Tatarstan Republic, ranked second in terms of population, and high proportion of pre-retirement and retirement age of the disabled population up to 15 years in the Republic of Mordovia, Kirov and Penza regions.

Another key factor that determines the prospects for the labor potential of the Volga region regions is insufficient population reproduction for balanced development, the values of the net population reproduction coefficient are in the range from 0.6 to 0.8, while only the Tatarstan Republic is characterized by a positive value of both population migration as a whole and interregional labor migration. The main factor for people leaving for work in other regions is the difference in the average wage between the regions of the Volga region and in comparison with other regions of Russia, as well as high tension in the labor market. It is proposed to use the following indicators of tension in the labor market of the region: the unemployment rate (%), the average time of searching for work by the unemployed (months), the share of unemployed looking for work for 12 months or more (%).

The analysis of the processes of clusterization of the Volga region economy revealed a significant differentiation of

regions in terms of the effectiveness of cluster policy and the presence of successfully implemented cluster initiatives on the territory of the region. The leaders of the Volga region in the field of cluster development are the Tatarstan Republic, Perm, Samara, Penza and Ulyanovsk regions. The clustering processes do not cover the economic systems of the Mari El Republic, Kirov, Saratov and Orenburg regions.

The obtained values of the integral and private indexes of the labor potential of clustering the economy for the Volga region regions (IILPC) allowed us to draw the following conclusions. The Nizhny Novgorod region has the greatest potential for clustering the economy in terms of labor potential. At the same time, this potential remains largely unused, since the share of the employed population of the Volga regions working in the organizations participating in the cluster is 0.32%. The first place of the Nizhny Novgorod region is due to the high values of private indices of urbanization (the largest agglomeration of the Volga region), skills and scientific personnel, and the Nizhny Novgorod region is characterized by a fairly stable situation of labor migration in comparison with other Volga regions.

The group of leading regions also includes the Perm region, Tatarstan Republic, Samara and Ulyanovsk regions, which are marked by high values of labor clustering for the Volga Federal district. The above-mentioned regions have significant labor potential for further development of production clustering processes based on market mechanisms, which should be taken into account when choosing forms and methods of state cluster policy. The following regions have the lowest value of the integral index (IILPC): Orenburg region, Chuvash Republic, Mari El Republic, and Mordovia. Implementation of the state cluster policy in these regions should include mechanisms of Federal support for the integration of production, education and science, as well as use the potential of interregional cooperation with the leading regions in the framework of joint cluster projects.

The obtained results allowed to formulate the following tasks in the study of processes of a clustering of economy of regions of the Volga region, namely, impact assessment of implementation of cluster policy in the development of regional innovation systems, and the impact on clustering of economic space of a region of the development process of urban agglomerations.

#### ACKNOWLEDGMENTS

The reported study was funded by RFBR, project number №20-010-00472 A

#### REFERENCES

1. G.A.Gareeva, T.V.Gilyazeev, D.R. Grigoryev Applications of neural networks in the economy. Young scientist. 2018. 18. 306-309.
2. E.I. Citizen. Neural networks. Textbook.
3. J.Goodfellow, I.Bengio, A.Courville. Deep Learning. DMK Press. 2018. 652.
4. Information on the housing lending market in Russia. Statistical Digest (Internet version) / CBR. Moscow, 2019. 66.

5. L.R.Yunusova, A.R. Magsumova . Classification of artificial neural networks. Science, education and culture. 2019. 7 (41).
6. A.A. Patlatenko Bank mortgage housing lending in Russia. Young scientist. 2017. 3 (137). 389-392.
7. A.A.Kochetkova New opportunities for the development of mortgage lending in Russia. Problems of Economics and Management. 2015. 5 71-73.
8. M.A.Abdullaev, A.A.Orlov, L.A. Rimsky, Bu Hassan Hossam. Mortgage. Organization of mortgage lending. Textbook. 2015. 244.
9. N.N. Khovalyg Features of mortgage lending: concept, classification, risks. Young scientist. 2018. 32 (218). 47-51.
10. E.Yu.,Khrustalev O.G. Shramko Using the method of neural networks to predict the effectiveness of investment investments. Economic analysis: theory and practice. 2017. 8 (467). 1438-1451.
11. S.M.Makeykina, I.A. Khlynkova Monitoring mortgage housing lending in Russia at the present stage. Scientific almanac. 2016. 10.
12. Aizinova I.M. Mortgage lending as an instrument of housing policy. Scientific works: Institute of Economic Forecasting RAS. 2017. 15. 489-510.
13. I.S.Ulbieva, K.M.Uzhakhov, A.H.Tsechoeva, Dzangieva AR, Uzhakhov MK. The construction technology of preschool institutions: Eco-friendly and fire-resistant materials. Int J Emerg Trends Eng Res [Internet]. 2020;8(4):1227-31
14. O.O.Gorshkova. Construction of long-distance oil pipelines. Int J Emerg Trends Eng Res [Internet]. 2020;8(1):60-66Kulikov I,
15. K. Karpukhin Studying energy efficiency of thermal management systems designed for electric vehicles with in-wheel motors. Int J Emerg Trends Eng Res [Internet]. 2020;8(6):2654-6