Prediction and Estimation of Highway Construction Cost using Machine Learning

Abbas M. Abd
Department of Highway and Airport Engineering, College of Engineering, University of Diyala, Iraq
abbas_mahde@uodiyala.edu.iq

Yassi A. Kareem
Department of Highway and Airport Engineering, College of Engineering, University of Diyala, Iraq
yassir_nashat@uodiyala.edu.iq

Raquim N. Zehawi
Department of Highway and Airport Engineering, College of Engineering, University of Diyala, Iraq
raquim_zehawi@uodiyala.edu.iq (corresponding author)

Received: 2 July 2024 | Revised: 22 July 2024 | Accepted: 24 July 2024

 DOI: https://doi.org/10.48084/etasr.8285

ABSTRACT

Cost estimation and prediction are crucial processes for the success of construction projects, especially for infrastructure development. This study analyzes historical data collected between 2011 and 2023 and investigates the relationship between construction elements and the final cost of highway construction projects in Iraq. Different cost analysis approaches, including statistical assessment and machine learning techniques, were applied to a dataset of 291 highway projects. Cost estimation is a time-consuming and risky process that requires many qualitative and quantitative parameters to be well analyzed. However, machine learning provides a comprehensive assessment tool to predict future costs. Four ANN-based models were investigated and precision was improved by combining RMSE and the correlation coefficient (R) as a controller. The results showed improvements in performance metrics, such as error reduction rate and correlation coefficient, for the models developed. The best performance was achieved at an R of 0.989. The proposed model can be effectively adapted to predict road construction costs. Despite the need for more data, the implication of the proposed model can ensure a sustainable application, saving the time and resources required by construction professionals to predict road project costs during the planning phase.

Keywords-cost prediction; machine learning; ANN; estimation; statistical analysis

I. INTRODUCTION

Cost estimation and analysis of road construction are essential. Given the need to assess price trends and time data, cost estimation has become an increasingly time-consuming and complex process. This study uses Artificial Neural Networks (ANNs) in a locally collected dataset, taking advantage of their dynamic and nonlinear nature to simplify and efficiently resolve prediction issues. This study aimed to simplify and streamline the prediction process of road costs and help to make precise predictions, saving both time and effort. Traditional cost estimation in construction works includes the assessment of direct cost, indirect expenses, and contractor profits. Contractors often use the unit cost method to calculate initial project costs, which requires high experience and accuracy in materials, labor, and equipment costs. In [1], an approach was presented to estimate project costs, and the challenges of using the traditional approach and the impact on reliability were discussed [1]. The biggest challenge in the cost estimation process is the wide variety of construction items. Another source of challenges lies in the multiplicity of construction processes, as a project includes dozens or hundreds of diverse items. This requires extensive experience in estimating quantities and prices, taking into account procurement time, the effect of inflation, and other influencing factors [2]. Figure 1, illustrates the method followed in this study. A comprehensive literature review on road construction and cost estimation was conducted to investigate the most important factors. Then, Machine Learning (ML) methods and applications in construction and cost estimation were investigated. Subsequently, data collection and analysis were performed to explore the trends of different cost and price indices. The main focus was on the modeling process, model features, development, evaluation, and validation.

This study was carried out with the presence of the following limitations:

- Study area and scope limitations: The dataset includes projects carried out mainly in the Diyala region of Iraq, especially for road construction projects. However, for generalization, a public or global database must be used for diverse project types and locations.

- Time limitations: This study investigated the cost trends for the period 2011-2023, and the data collected covers projects implemented during this period.
- Technical limitations: The models developed were constructed based on ANN algorithms, exploring all possible combinations of features and activation functions to achieve improved precision in the model output.

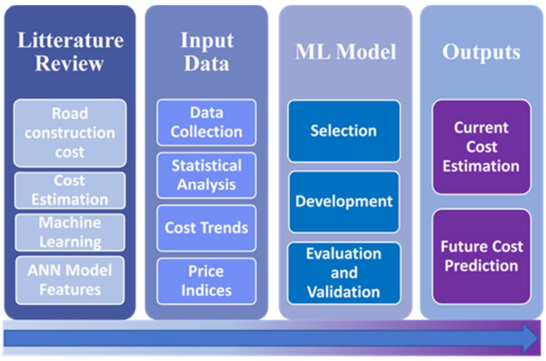

Fig. 1. Research method.

II. PRELIMINARY ESTIMATION

Cost estimates are made in the early stages of project development without design and engineering knowledge and with limited information at the project level. Preliminary estimates are conceptually required for many reasons, such as feasibility studies, budgeting and financing initiation, and investment profitability investigation. Therefore, experience and judgment are required to obtain a reliable cost estimate. Conceptual or preliminary estimation is a type of cost estimation in the early stages of any construction project [3, 4]. It is a process that depends on personal experience and involves evaluating the different relationships between all cost factors. The first step in the initial cost estimation of a project is to choose the right input factors, and it is important to obtain a good performance benchmark or improve the predictive power of the model [5]. Many studies have used different approaches, including statistical and non-statistical methods, to identify and select the main factors that are necessary to estimate the conceptual costs of highway construction projects [4, 5]. A comprehensive cost comparison should investigate totally or partially the type of project, customer type, and land usage, besides the material cost, work size, design changes, duration, type of contract, and any impacts on contract prices [6]. In [7], future costs were estimated for construction projects using input data for 12 years, considering the prices of construction materials and using a linear regression model to show the relationship between year and price. In [8], the development of cost predictions using highway agency statistics was discussed, showing that the main input variables had a significant impact on total cost and that the proposed cost estimation method required little effort. The crucial factor that affects the precision of cost estimation is the experience and skill of the estimator in many detailed construction cost details [8, 9]. Important factors include project type, construction services, location, and costs introduced in the modeling algorithm to predict construction costs. Cost-affecting factors have been identified during the implementation of construction projects. The expected cost is the first stage in the estimation process, aiming to provide conceptual information for decision-making actions and further cost modification [10. 11].

III. MACHINE LEARNING AND NEURAL NETWORKS

Artificial intelligence presents powerful modeling tools that can be used efficiently throughout the project life cycle before it is even constructed. Many ML algorithms, such as ANNs, model complicated data relationships, despite the absence of a straightforward equation to describe data variable relationships. In the construction sector, ANNs have proven to achieve human-like performance in complex situations, such as productivity estimation, corporate bankruptcy prediction, and evaluation of financial projects. Previous studies have attempted to apply ANNs in construction estimation with mixed results [11-13]. The main features of these algorithms can be summarized as follows:

- Neural networks can perform complex functions on data. Neural network modeling involves learning to map input to output values through a network of connected neurons.
- Kernel functions map the input data into a higher-dimensional space, where linear algorithms can solve non-linear problems.
- Learning representations: A neural network is an intermediate-level network that can learn classes of representations (features) from low- to high-level physical objects from training input, just as people do.

Neural networks discover patterns in the data during training and create new knowledge related to them [14-16]. In ANNs, the training functions continuously adjust the weights and biases of the model to guide the learning process and reach the lowest difference between the observations and the model output. Meanwhile, nonlinearity is introduced by activation functions to determine the output of a neuron, applied to the nonlinearity of the input's weighted sum of the previous layer, enabling the model to learn complex patterns [17-19]. Table I lists the training and activation functions that are used and combined in ANNs. This study will explore combinations of features and activation functions in the training phase to achieve improved precision of model output.

TABLE I. TRAINING AND ACTIVATION FUNCTIONS BUILT IN ANN

Training functions				
trainlm	trainrp	trainscg	trainbr	traincgb
Hidden activation functions				
logsig	tansig	softmax		
Output activation functions				
purelin	logsig	tansig	softmax	

The second approach focuses on the network's weights, which are continuously modified until the controller metric is reached. This study used two controllers: Root Mean Square

Error (RMSE) and correlation coefficient (R). The training process continues until the difference between the model and the actual outputs converges to an acceptable accuracy. Model training is stopped either when RMSE reaches as minimum as possible, R reaches as maximum as possible, or both.

IV. PERFORMANCE METRICS

The choice of metrics to evaluate the precision of an ANN model depends on the type of problem. The best metric depends on the features of the specific problem and the errors that are most critical to avoid [7, 20, 21]. It is often helpful to consider multiple metrics to get a well-rounded view of model performance. Table II provides a list of the main metrics used to assess the precision of an ANN model [22, 23].

TABLE II. STATISTICAL METRICS TO ASSESS AN ANN MODEL

Formula	Nomenclature
$R = \frac{n(\sum x.y) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2].[n\sum y^2 - (\sum y)^2]}}$	R = correlation coefficient σ = Standard Deviation x = Values in the first set of data y = Values in the second set of data n = Observation number Y_i = Actual values, $\widehat{Y}_i$ = Predicted values $\bar{Y}$ = Mean
Standard Error: $(SE) = \frac{\sigma}{\sqrt{n}}$	
$RMSE = \sqrt{\frac{\sum_1^n (Y_i - \widehat{Y}_i)^2}{n}}$	
Standard Deviation: $\sigma = \sqrt{\frac{\sum_1^n (Y_i - \bar{Y})^2}{n}}$	
Mean Absolute Error (MAE): MAE $= \frac{\sum_1^n \lvert x_o - x \rvert}{n}$	

V. COST ESTIMATION OF ROAD PROJECTS

A. Data Collection

The collection of historical data is the first step in estimating costs for any type of project. This step provides estimators with the required information on price trends and inflation effects on material and different work costs. For this study, data from approximately 350 projects implemented during the 2011-2023 period were collected. In the second step, the dataset was prepared.

TABLE III. VARIABLES CONSIDERED FOR COST ESTIMATION AND PREDICTION OF HIGHWAY PROJECTS

Input			Output
Years	Earthwork costs	Paving work costs	Total project cost
Quantity of land preparation	Quantity of sub-base works	Quantity of shoulder works	Cost per meter
Price of land preparation	Price of sub-base works	Price of shoulder works	
Cost of land preparation	Cost of sub-base works	Cost of shoulder works	
Quantity of earthworks	Quantity of paving works	Road length	
Price of earthworks	Price of paving works	Cost of other items	

Many projects were excluded due to missing data, abnormal pricing within the time interval, or not covering all the items targeted by this investigation. A set of 291 projects was collected, which satisfied the objectives of estimation and prediction. A set of 20 parameters was used: prices, quantities, and cost of the main items (land preparation, earth, sub-base layer, paving, shoulders, and complementary works). In addition, the total cost and the cost per meter were determined for each project. Table III presents details on the input and output variables considered, depending on the literature and the dataset collected.

B. Data Analysis

The frequency analysis of the construction items and their weights within the project gives a clear view of the effects of each item on the final project cost. Figure 2 shows the mean weights of the items considered for all the data collected. A normal distribution reflects the trend of weight for the six main items. The highest weight was for paving works (63%) and then sub-base works (17%). Earthworks and complementary works were 8 and 5%, respectively. The lowest rates belonged to road shoulders and land preparation works, with 4 and 3%, respectively. Figure 2 explains the cost breakdown analysis for each year and the share rate through a graphical representation of each item with the projects throughout the study period.

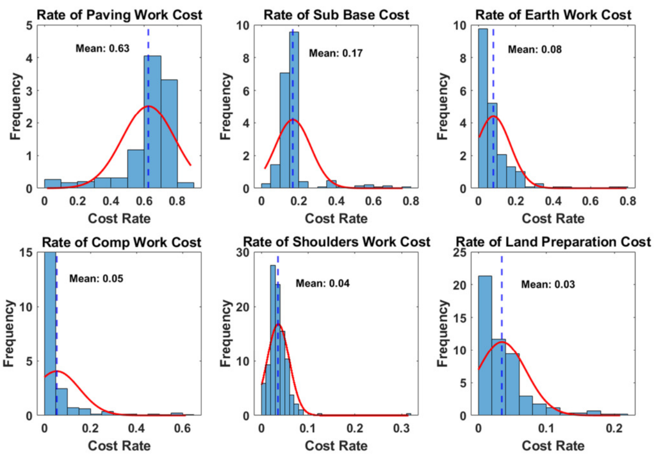

Fig. 2. Histograms of cost rates related to total project cost with normal distribution.

The correlation coefficient (R) is essential to assess the contribution of each construction item to the overall cost of the project, reflecting the statistical relationship between each cost and the output factors. This means that when developing a mathematical model to estimate the project cost, it will be very sensitive to the variation in the factors of high correlation (in this case; sub-base, paving, and earthworks), as shown in Figure 4. Figure 5 shows the fluctuation in the cost per meter of road work throughout the period. It is clear that for 2019 and 2022, there was the highest fluctuation in cost per meter of length for road construction. The reason behind that was the security problems and the deficiency in the country's budget that caused low exchange rates of the local currency relative to international ones, especially the US$. In the same context, minimal fluctuation occurred during the period 2011-2013. The reason behind this is that this period witnessed high economic activity, security, stability, and broad reconstruction projects.

Investigating the interaction and interoperability between factors affecting the final cost or cost per meter plays an important role for the estimators to assess the mutual relations among these factors. Figure 6 shows the correlation between all these factors. This plot shows the degree (if any) of influence of each factor on others and their relation.

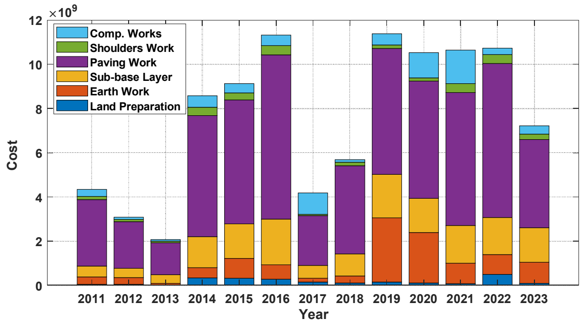

Fig. 3. Cost breakdown analysis for each year.

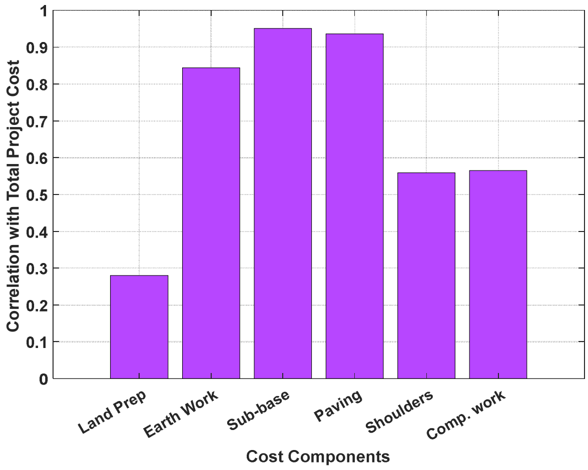

Fig. 4. Correlation analysis of cost components.

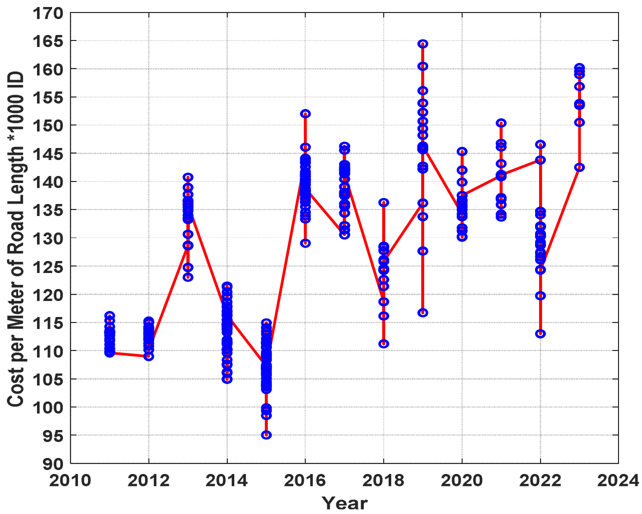

Fig. 5. Cost per meter of road length for each year.

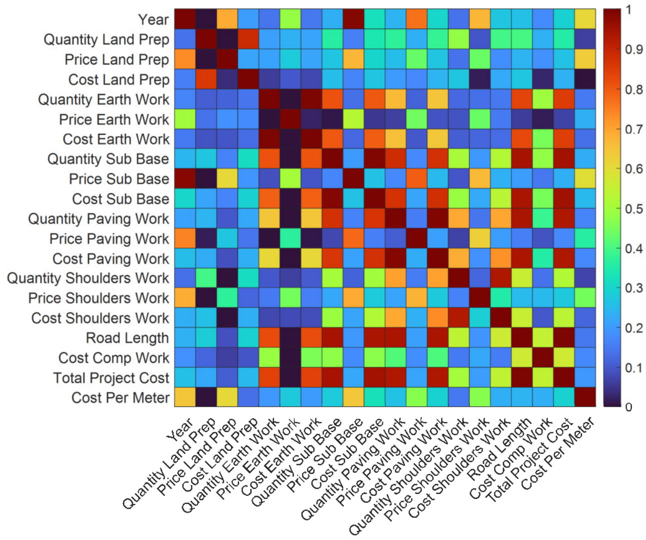

Fig. 6. Correlation heatmap matrix between variables,

VI. PRINCIPAL VISUALIZATION

The principal components analysis also gives a clear idea of the relations between the variables considered. This method groups the variables into multiple groups of interrelated effects and highlights the participation of each group in the final output. Rather than explaining the most critical variables, this method enables estimators to analyze the mutual impact of variables within each group (Table IV) to consider in future decisions when estimating the final project cost. Understanding variables' relations can guide specialists to state the reason behind the fluctuation in the pricing of some items compared to others, depending on the behavior of the interaction effects of the considered variables. As key concepts of principal visualization, this study considered the accumulative explained variance rate. The rate of explained variance represents the portion of each principal component on the total variance, and how these variables in specific components spread across the collected data.

TABLE IV. EXPLAINED VARIANCES FOR THE MOST IMPORTANT PRINCIPAL COMPONENTS

Principal component	Explained variance	Variable (rate)	Variable (rate)	Variable (rate)
1	82.45%	13 (0.8461)	7 (0.4487)	10 (0.2642)
2	13.17%	7 (0.8561)	13 (-0.4903)	18 (0.1253)
3	3.38%	18 (0.9842)	7 (-0.1707)	13 (-0.042)
4	0.81%	10 (0.9535)	13 (-0.197)	7 (-0.181)
5	0.15%	4 (0.9923)	10 (-0.1186)	7 (0.0298)
6	0.04%	16 (0.9962)	13 (-0.0549)	7 (0.0539)
All other components (7 to 19) have explained variance less than 0.01				

In this study, the first component has a rate of explained variance at 82.45%, and the rates for the second and third components are 13.17% and 3.38%, respectively. These three components capture 99% of the total variance for the dataset. The variables belonging to these three components were 7, 10, 13, and 18, which represent the cost of earthworks, cost of sub-base, cost of paving works, and cost of complementary works. These trend leader variables reflect the most important patterns of the collected data, as shown in Figure 7.

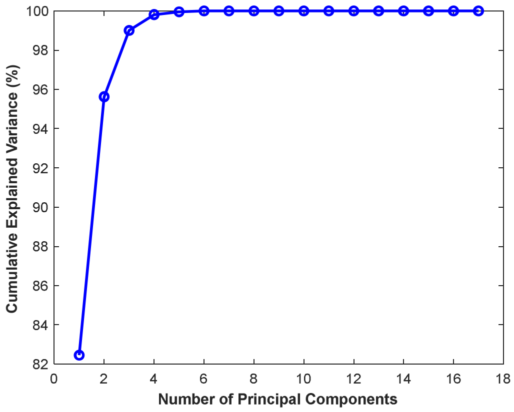

Fig. 7. Cumulative explained variance for all principal components.

VII. COST ESTIMATION

Figure 8 represents the trend line for historical data collected for highway construction projects.

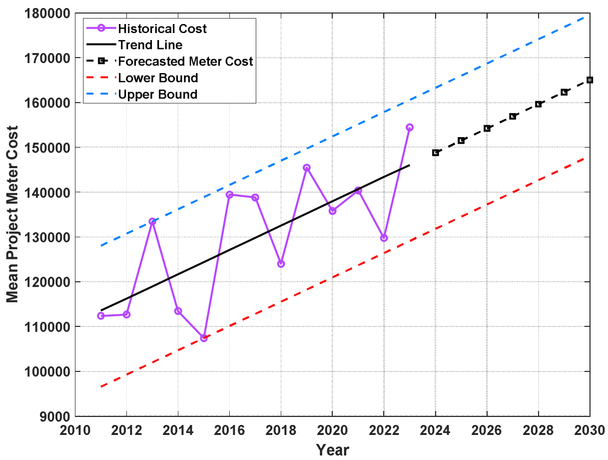

Fig. 8. Project cost per meter and future cost prediction.

TABLE V. PREDICTED COST PER METER FOR 2024-2030

Year	Lower bound	Trend line	Upper bound	PERT
2024	131,806	148,785	163,277	148,370
2025	134,516	151,495	165,988	151,081
2026	137,227	154,205	168,698	153,791
2027	139,937	156,916	171,408	156,501
2028	142,648	159,626	174,119	159,212
2029	145,358	162,337	176,829	161,922
2030	148,068	165,047	179,540	164,633

The upper and lower bounds reveal the fluctuation ranges of upward and downward values of the cost per meter in the period studied (2011-2023) and extended linearly up to 2030. There were specific periods of downward cost (2014, 2015, and 2022), due to security issues and the COVID-19 pandemic, as most projects stopped and there was a reduction in government expenditure and investments for many sectors, especially infrastructure projects. Competition between contractors resulted in price drops for these projects. On the other hand, there are upward trends in costs during 2016, 2019, and 2023. These periods witnessed economic improvement throughout the country, with the reconstruction of cities affected by war or security problems. Infrastructure projects, especially road rehabilitation, gained high funding and investment, which justifies the increase in the prices of highway projects. Using linear regression, three values are shown in Table V, representing the most likely, optimistic, and pessimistic predicted costs. Statistically, a three-point estimation technique or the Program Evaluation and Review Technique (PERT) is a preferred tool to find the mean cost using the extracted data shown in Figure 8. At this stage, the estimator can draw the estimated cost of the project, informed of the best and worst scenarios of the expected circumstances and conditions.

$$X = \frac{O+4M+P}{6}$$

where X is the mean value, O is the optimistic, M is the most likely, and P is the pessimistic.

VIII. COST PREDICTION USING ANN

This study aimed to evaluate and validate the newly developed framework for predicting cost models for highway projects (typical two-lane roads). This framework uses 18 input features to predict total cost and cost per meter as output. Three approaches were followed to develop and improve the neural network using MATLAB (2023b) software:

- Basic backpropagation method with default options as a reference model (M1.0).
- Optimizing the ANN model using three statistical controllers:
  - Correlation coefficient (R) (M2.1),
  - RMSE (M2.2)
  - R and RMSE (M2.3).
- Optimizing the ANN model by applying the combination of different activation and training functions (M3.0).

The first model (M1.0) used a basic ANN with backpropagation and default options. Figure 9 presents the actual mean cost per meter and the predicted cost, indicating a good prediction in the training stage. Figure 10 details the performance of the model for the training, validation, and testing processes (0.92, 0.85, and 0.80, respectively), with an overall correlation of 0.89. Figure 11 plots the predicted cost for future highway projects up to 2030 (161000 to 174000 ID per meter).

Model M2.1 used the ANN with the correlation coefficient (R) as a controller, enforcing the modeling process to continuously iterate until reaching the best R for the predicted cost. A notable improvement was observed compared to the M1.0 model. The correlation coefficient increased by 5% and RMSE decreased by 23%. Taking the M1.0 as a reference, Figure 12 presents the best convergence for the actual mean cost per meter with predicted cost, indicating better prediction capabilities. Figure 13 indicates its higher performance, with high correlations for the training (0.95), validation (0.90), and testing (0.80) processes, respectively, and the overall correlation was 0.94. Figure 14 shows the predicted cost for highway projects up to 2030 (207500 to 218500 ID per meter).

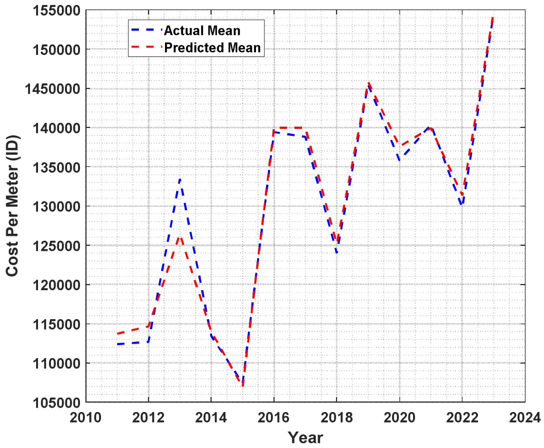

Fig. 9. M1.0: Mean actual vs predicted cost per meter.

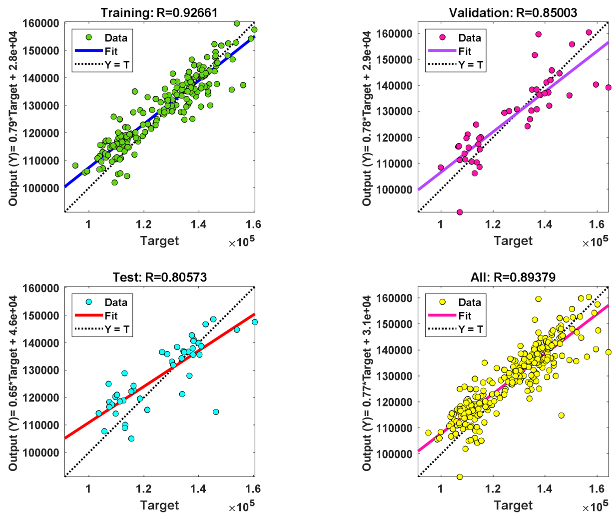

Fig. 10. M1.0: Correlation coefficients for training and test data.

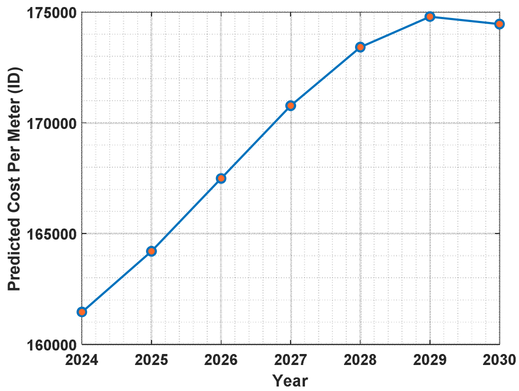

Fig. 11. M1.0: Predicted highway cost per meter.

The model using RMSE as the controller (M2.2) enforces the modeling process to iterate until reaching a minimum RMSE for the predicted cost. A significant improvement was recorded. The correlation coefficient (R) increased by 7.3% and RMSE decreased by 39% compared to M1.0. Figure 15 presents the continuous improvement of the prediction for the actual mean cost per meter. Figure 16 indicates a higher performance with high correlations for the training (0.97), validation (0.93), and testing (0.91) processes, with a correlation of 0.96 for the overall dataset. Figure 17 shows the expected cost range predicted for 2024-2030 (149500 to 154500 ID per meter).

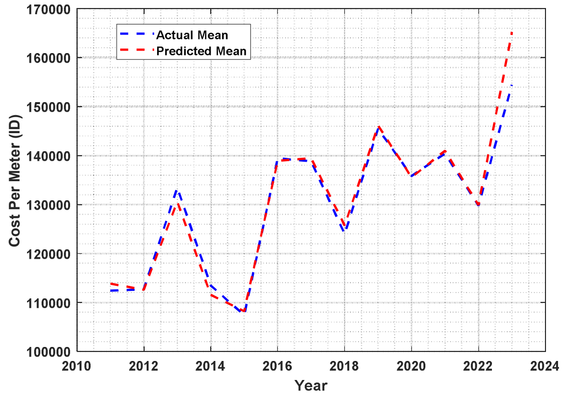

Fig. 12. M2.1: Mean actual vs predicted cost per meter.

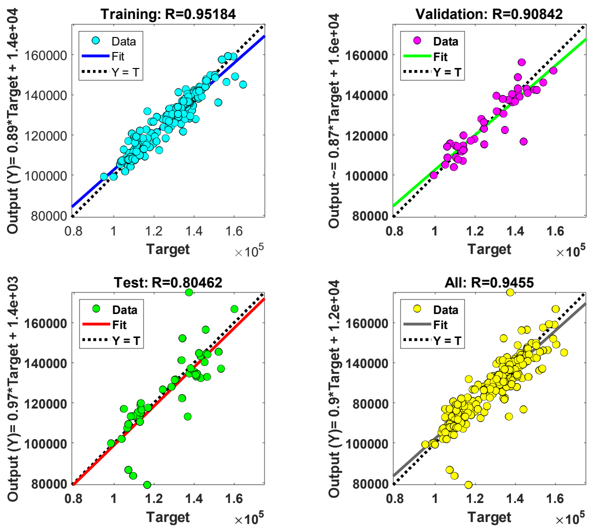

Fig. 13. M2.1: Correlation coefficients for training and test data.

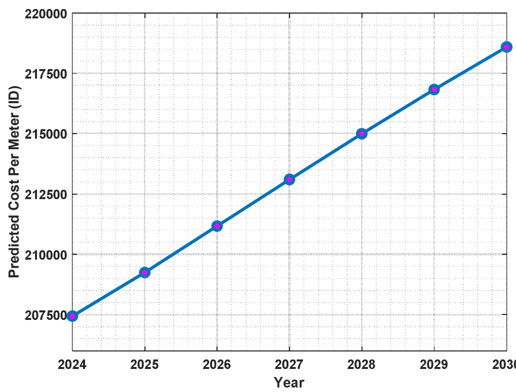

Fig. 14. M2.1: Predicted highway cost per meter.

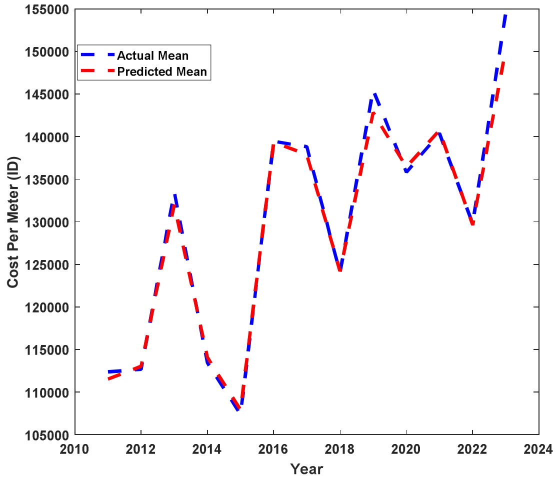

Fig. 15. M2.2: Mean actual vs predicted cost per meter.

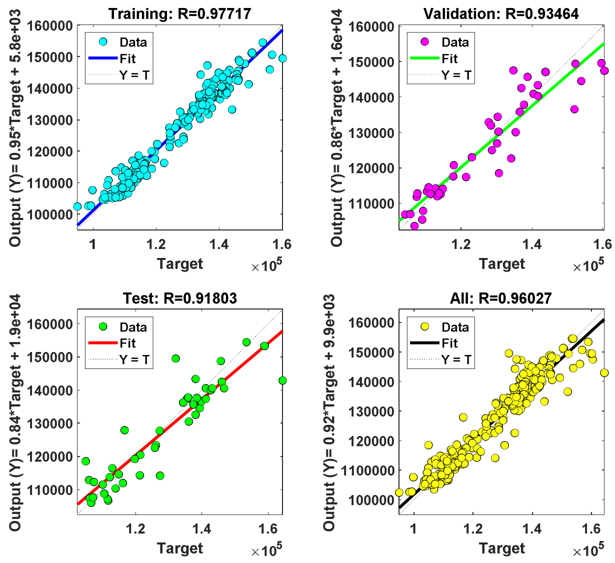

Fig. 16. M2.2: Correlation coefficients for training and test data.

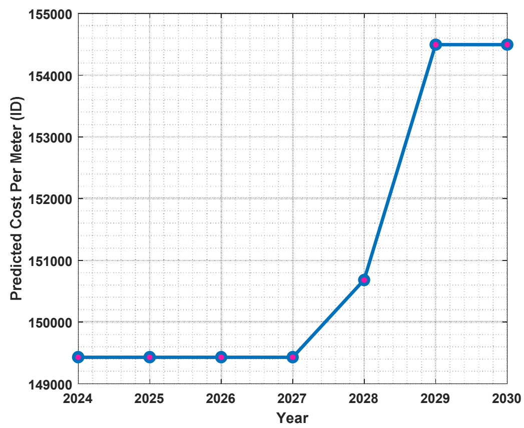

Fig. 17. M2.2: Predicted highway cost per meter.

M2.3 achieved the best performance in the second approach using both R and RMSE to control the training process. Figure 18 indicates the remarkable enhancement in cost prediction and model performance. The correlation coefficient increased by 9.5% and RMSE reduced by 48% compared to M1.0. Figure 19 shows a higher performance with high correlations for training (0.98), validation (0.94), and testing (0.93), and a correlation of 0.97 for the overall data set. Figure 20 shows the predicted cost for 2024-2030 (151700 to 153800 ID per meter).

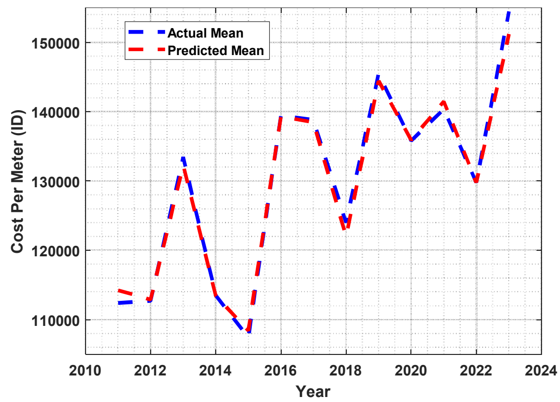

Fig. 18. Mean actual vs. predicted cost per meter.

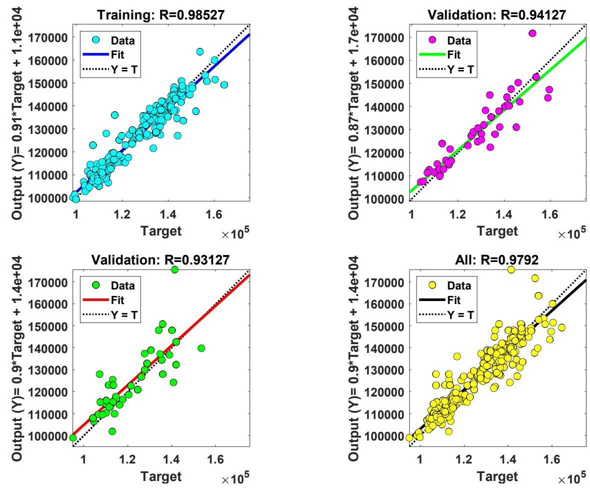

Fig. 19. Correlation coefficients for training and test data.

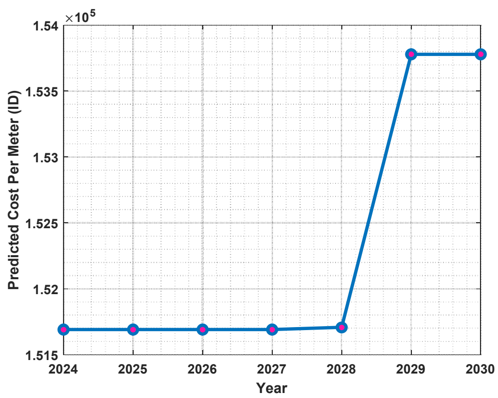

Fig. 20. Predicted highway cost per meter.

The third approach focused on neural network features, using different sets of functions for the input, hidden, and output layers for training and activation, resulting in 45 different combinations. Model M3.0 investigates the optimization of the use of these functions simultaneously. The results verify the remarkable improvement in all performance metrics. Figure 21 states an almost perfect agreement between the actual and predicted costs. Figure 23 shows the highest R = 98.9% and the highest RMSE reduction rate (58.4%) relative to the M1.0 model. Cost prediction for 2024-2030 reveals more reliable results compared to other models, as it converges closely to the linear prediction found earlier, as shown in Figure 22.

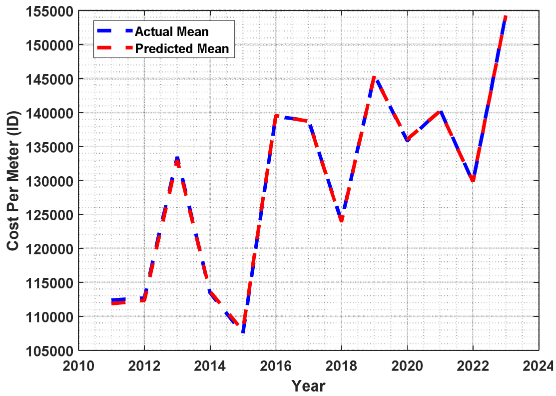

Fig. 21. Mean actual vs. predicted cost per meter.

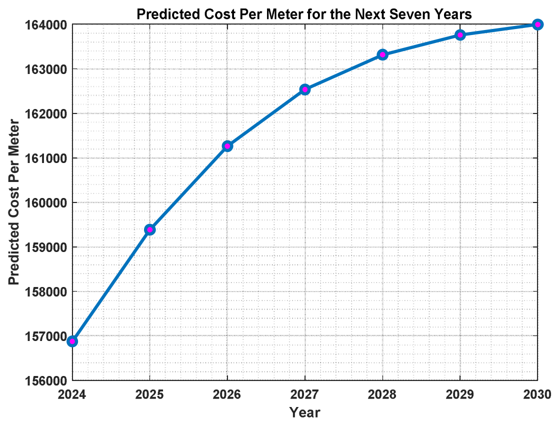

Fig. 22. Predicted highway cost per meter.

Figure 24 displays the distribution of errors between the predicted cost and the true cost. Each bin plotted along the x-axis represents a particular range of error values, and the height of the bin along the y-axis shows the number of data points falling in the respective error range.

The histogram is centered on zero error, with most data points clustered around bins near zero. This indicates that most points are well predicted and the model makes good predictions for a great proportion of the data points. Table VI presents the performance metrics for all the models tested.

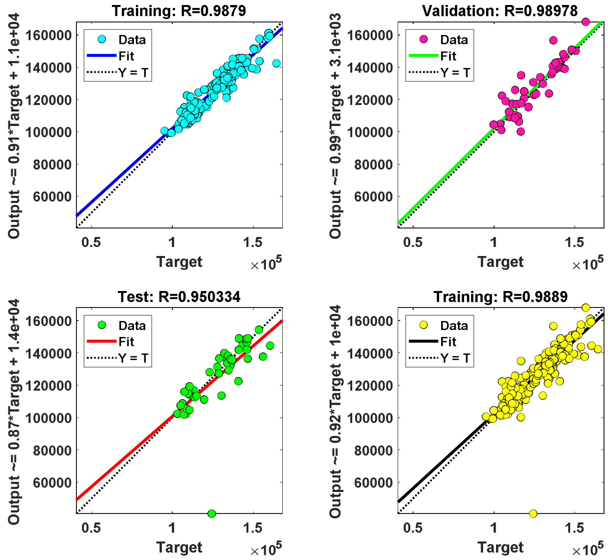

Fig. 23. Correlation coefficients for training and test data.

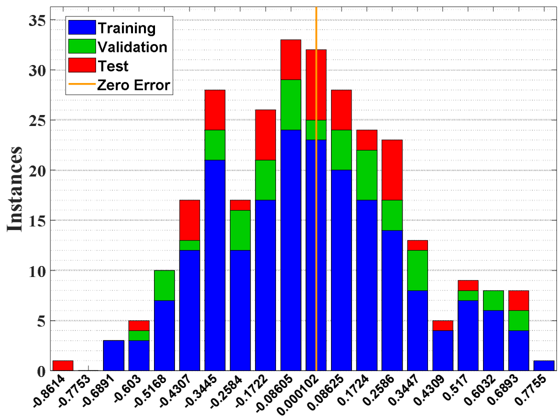

Fig. 24. Histogram of error's distribution.

TABLE VI. ANN MODEL METRICS FOR PREDICTED COST PER METER

	M1.0	M2.1	M2.2	M2.3	M3.0
R	0.894	0.946	0.960	0.979	0.989
SE	5644	4357	3051	2883	2371
RMSE	6885	5284	4187	3570	2866
SD	14998	16217	14356	12511	14962
MAE	3956	2999	2872	2662	1556

IX. CONCLUSIONS

This paper presented a new framework for estimating costs in the early stages of highway construction projects. This framework incorporates ANNs to simplify and enhance the estimation process. Paving and sub-base layer works were identified as crucial elements of the project cost, by 0.63 and 0.17, respectively. Cost analysis determined a reduction in government expenditure and investment for many sectors, especially infrastructure projects, for two periods (2013-2014 and 2020-2021) due to security issues and the COVID-19 pandemic, where most projects were stopped and the competition between contractors resulted in price drops.

Model M1.0, which was a basic backpropagation ANN, achieved acceptable performance for training, validation, and testing with R values of 0.92, 0.85, and 0.80, respectively, while its overall correlation was 0.89. Model M2.1, which uses the ANN with the correlation coefficient (R) as a controller, achieved better performance with high correlations of 0.95, 0.90, and 0.80 for training, validation, and testing, respectively, while its overall correlation was 0.94. Model M2.2 used RMSE as a controller, achieving higher performance with R values of 0.97, 0.93, and 0.91 for training, validation, and testing, respectively, and an overall R of 0.96. Enhanced performance was achieved by Model M 2.3, which used both R and RMSE as a double controller for the training process. This model achieved much better performance with R values of 0.98, 0.94, and 0.93 for training, validation, and testing, respectively, and an overall R of 0.97. Model M3.0 optimized the use of activation functions (trainbr, logsig, tansig) for the input, hidden, and output layers during training. This model achieved the best performance with R values of 0.95 for testing and 0.989 for training, validation, and for the overall dataset. This study incorporated multiple enhancements to the reference model to improve its performance. Compared to traditional estimation methods, cost prediction using ANN provides results with high precision and less effort. The results of this study are consistent with the findings of [18, 21], distinguished by its comprehensive cost analysis and future forecast of cost trends in the field of road projects.

As a recommendation, it will be worth studying the effect of early prediction of project cost on decision-making, quality assurance, and risk management of construction processes. In addition, it is recommended to investigate the generalization of the developed ML models, using reliable data to predict the cost of numerous types of construction projects in different sectors.

REFERENCES

[1] N. I. El-Sawalhi and O. Shehatto, "A Neural Network Model for Building Construction Projects Cost Estimating," *Journal of Construction Engineering and Project Management*, vol. 4, no. 4, pp. 9–16, Dec. 2014, https://doi.org/10.6106/JCEPM.2014.4.4.009.

[2] A. Alkhuadhan and S. Naimi, "A new model for cost estimation construction project using Hybrid importance regression ensemble method," *Acta logistica*, vol. 10, no. 2, pp. 191–198, Jun. 2023, https://doi.org/10.22306/al.v10i2.372.

[3] S. Yun, "Performance Analysis of Construction Cost Prediction Using Neural Network for Multioutput Regression," *Applied Sciences*, vol. 12, no. 19, Jan. 2022, Art. no. 9592, https://doi.org/10.3390/app12199592.

[4] Y. Zhang, R. E. Minchin, I. Flood, and R. J. Ries, "Preliminary Cost Estimation of Highway Projects Using Statistical Learning Methods," *Journal of Construction Engineering and Management*, vol. 149, no. 5, May 2023, Art. no. 04023026, https://doi.org/10.1061/JCEMD4.COENG-12773.

[5] M. G. Meharie, Z. C. A. Gariy, R. N. Ndisya Mutuku, and W. J. Mengesha, "An Effective Approach to Input Variable Selection for Preliminary Cost Estimation of Construction Projects," *Advances in Civil Engineering*, vol. 2019, no. 1, 2019, Art. no. 4092549, https://doi.org/10.1155/2019/4092549.

[6] C. R. Irnanda, I. Sadalia, and Nazaruddin, "Contract Analysis for Design and Build Lump Sum Price," presented at the 19th International Symposium on Management (INSYMA 2022), Dec. 2022, pp. 1162–1169, https://doi.org/10.2991/978-94-6463-008-4_143.

[7] A. O. Elfaki, S. Alatawi, and E. Abushandi, "Using Intelligent Techniques in Construction Project Cost Estimation: 10-Year Survey," *Advances in Civil Engineering*, vol. 2014, no. 1, 2014, Art. no. 107926, https://doi.org/10.1155/2014/107926.

[8] B. J. Gardner, D. D. Gransberg, and H. D. Jeong, "Reducing Data-Collection Efforts for Conceptual Cost Estimating at a Highway Agency," *Journal of Construction Engineering and Management*, vol. 142, no. 11, Nov. 2016, Art. no. 04016057, https://doi.org/10.1061/(ASCE)CO.1943-7862.0001174.

[9] J. Yuan and C. Yin, "Cost Estimation of Expressway Construction and Installation Engineering Based on Convolutional Neural Networks," *Journal of Engineering Studies*, vol. 15, no. 5, pp. 446–455, Oct. 2023, https://doi.org/10.3724/j.issn.1674-4969.23050401.

[10] Y. G. Abed, T. M. Hasan, and R. N. Zehawi, "Cost Prediction for Roads Construction using Machine Learning Models," *International Journal of Electrical and Computer Engineering Systems*, vol. 13, no. 10, pp. 927–936, Dec. 2022, https://doi.org/10.32985/ijeces.13.10.8.

[11] Y. Zhang, R. E. Minchin, I. Flood, and R. J. Ries, "Preliminary Cost Estimation of Highway Projects Using Statistical Learning Methods," *Journal of Construction Engineering and Management*, vol. 149, no. 5, May 2023, Art. no. 04023026, https://doi.org/10.1061/JCEMD4.COENG-12773.

[12] M. G. Meharie, W. J. Mengesha, Z. A. Gariy, and R. N. N. Mutuku, "Application of stacking ensemble machine learning algorithm in predicting the cost of highway construction projects," *Engineering, Construction and Architectural Management*, vol. 29, no. 7, pp. 2836–2853, Jan. 2022, https://doi.org/10.1108/ECAM-02-2020-0128.

[13] S. H. Ji, J. Ahn, H. S. Lee, and K. Han, "Cost Estimation Model Using Modified Parameters for Construction Projects," *Advances in Civil Engineering*, vol. 2019, no. 1, 2019, Art. no. 8290935, https://doi.org/10.1155/2019/8290935.

[14] S. Yun, "Performance Analysis of Construction Cost Prediction Using Neural Network for Multioutput Regression," *Applied Sciences*, vol. 12, no. 19, Jan. 2022, Art. no. 9592, https://doi.org/10.3390/app12199592.

[15] G. I. Parisi, R. Kemker, J. L. Part, C. Kanan, and S. Wermter, "Continual lifelong learning with neural networks: A review," *Neural Networks*, vol. 113, pp. 54–71, May 2019, https://doi.org/10.1016/j.neunet.2019.01.012.

[16] G. M. Tina, C. Ventura, S. Ferlito, and S. De Vito, "A State-of-Art-Review on Machine-Learning Based Methods for PV," *Applied Sciences*, vol. 11, no. 16, Jan. 2021, Art. no. 7550, https://doi.org/10.3390/app11167550.

[17] J. C. Ancheta *et al.*, "Road construction analysis using regression technique," *World Journal of Advanced Research and Reviews*, vol. 18, no. 3, pp. 658–664, 2023, https://doi.org/10.30574/wjarr.2023.18.3.1125.

[18] A. F. Khalil and S. Rostam, "Machine Learning-based Predictive Maintenance for Fault Detection in Rotating Machinery: A Case Study," *Engineering, Technology & Applied Science Research*, vol. 14, no. 2, pp. 13181–13189, Apr. 2024, https://doi.org/10.48084/etasr.6813.

[19] A. Satty, M. M. Y. Salih, A. A. Hassaballa, E. A. E. Gumma, A. Abdallah, and G. S. M. Khamis, "Comparative Analysis of Machine Learning Algorithms for Investigating Myocardial Infarction Complications," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12775–12779, Feb. 2024, https://doi.org/10.48084/etasr.6691.

[20] L. Gurung and M. Chhetri, "Cost estimation model for highway projects using ANN," *Zorig Melong- A Technical Journal of Science, Engineering and Technology*, vol. 8, no. 1, Jun. 2023, https://doi.org/10.17102/v8002.

[21] A. M. Abd and S. S. Sammen, "Artificial Neural Network Model for Managing and Forecasting Water Reservoir Discharge (Hemren Reservoir as A Case Study)," *Diyala Journal of Engineering Sciences*, pp. 132–143, Dec. 2014, https://doi.org/10.24237/djes.2014.07409.

[22] D. O. Mac-Barango and P. C. Nwogu, "Ranking of Variables Influencing Construction Project," *World Journal of Finance and*

Investment Research, vol. 6, no. 1, pp. 49–61, Aug. 2022, https://doi.org/10.56201/wjfir.v6.no1.2022.pg49.61.

[23] C. Magashi, J. Agbinya, A. Sam, and J. Mbelwa, "Prediction of SACCOS Failure in Tanzania using Machine Learning Models," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12887–12891, Feb. 2024, https://doi.org/10.48084/etasr.6696.