Int. J. Nonlinear Anal. Appl. In Press, 1–9
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2023.7374

Using support vector machine (SVM) technology to predict the duration of irrigation canal projects in western Iraq

Jumaa A. AL-Somaydaii[a], Aseel H. Abdaljader[a,*], Saadi Shartooh Sharqi[b], Nadhir Al Ansari[c]

[a]*Dams and Water Resources Department, University of Anbar, Ramadi, Iraq*
[b]*Civil Engineering Department, University of Anbar, Ramadi, Iraq*
[c]*Civil, Environmental and Natural Resources Engineering, Lulea University of Technology, Sweden*

(Communicated by Javad Vahidi)

Abstract

In this study, a support vector machine (SVM) based technique for timing irrigation projects is presented, and one of the most accurate predictive models in calculating the final project duration within the contract documents, where the research problem is projects are not completed within the contract period because most of the total project duration is determined In an unthoughtful manner by the employer. Linear regression models were applied to data and information for several projects, and a significant improvement in forecast accuracy was obtained.

Keywords: linear regression, support vector machine, construction time
2020 MSC: 62J05

1 Introduction

Within the framework of the literature, it is noted that there is a dearth of studies that deal with the issue of determining the factors affecting the duration estimation of irrigation projects using SVM, and from here a question arises concerning the most critical factors that affect the forecasting of project durations to reach a comprehensive view of the reasons for duration estimate of irrigation projects. Conducting a comprehensive review of the literature with the goal of the most significant factors influencing this estimation, which deals with this subject for a period of 25 years for the period from 1996 to 2022. It is expected that the study will be an addition to the scientific literature related to improving the performance of construction project management in terms of estimating project durations using SVM, especially since the literature still calls for more studies will be conducted to determine the most important factors affecting the estimation of project durations. Factors affecting the duration of infrastructure projects (irrigation and drainage projects) are: (channel length, channel width, channel depth, dimensions of the concrete lining, weather conditions, and ground conditions)

A support vector machine (SVM) is a type of machine learning that uses a recursive training algorithm to reduce the magnitude of a known error function:

- SVM classification type 1 (also known as C-SVM classification)

*Corresponding author
Email addresses: jah_eng@uoanbar.edu.iq (Jumaa A. AL-Somaydaii), aseel.abdulla67@uoanbar.edu.iq (Aseel H. Abdaljader), eng.saadish@uoanbar.edu.iq (Saadi Shartooh Sharqi), nadhir.alansari@ltu.se (Nadhir Al Ansari)

Received: November 2022 *Accepted:* January 2023

- SVM classification type 2 (also known as nu-SVM classification)
- SVM regression type 1 (also known as epsilon-SVM regression)
- SVM regression type 2 (also known as nu-SVM regression)

The study area is an irrigation project in western Iraq from Al-Qaim district to Al-Saqlawiya sub-district, Al-Tharthar arm - Euphrates. One of the most important projects in it is the Al-Ramadi irrigation project, which has an area of more than 150,000 acres on both sides of the Euphrates River. The areas of Ramadi Island, Khalidiya, Fallujah, and Garma, where the directorate secures the delivery of water quotas to all agricultural lands and beneficiaries in an equitable manner while continuing maintenance work for irrigation stations on a regular basis and supervising the implementation of the summer and winter agricultural plan.

1.1 Research problem

The problems of this research are as follows:

1. Helping the Project manager in predicting the duration of the project Add knowledge to the specialization field (Construction Management).
2. The construction sector in Iraq lacks accuracy in scheduling projects, Guessing mostly depends on personal experience and causes an extended estimation of project activities, thus estimating the duration of the project.

1.2 Research aims

Developing a mathematical model, the focus of this study that can forecast how long irrigation channel projects in the Iraqi sector would take. To achieve the aim of the study, the following objectives have been identified:

1. Determining the affected variable on the duration and driving mathematical equation.
2. Utilization the modern technologies such as SVM for prediction of the duration of Irrigation Channels Projects and Validity of a model.
3. Clarify the role of modern technologies and their applicability to resource management and planning Water in general, and in the study area, Anbar Governorate in particular.

2 Planning and scheduling management in irrigation project

To achieve the research objectives, two scientific methods with successive and overlapping stages were used, as follows:

2.1 Theoretical research method

1. A review of the literature on the implementation of irrigation projects in Iraq.
2. Investigating planning methodologies and managing irrigation project scheduling and their requirements, as well as the factors influencing them.

A good planning strategy entails monitoring the new sources that emerge throughout the lifecycle of the project and identifying their delay causes, as a result, the proposal of robust and reliable research methods as an analysis tool capable of mimicking and comprehending dynamic input parameters is critical for this purpose [9].

Irrigation project planning is an important process for implementing irrigation projects. The construction plan, the work to be accomplished, and the sequence in which the work will be completed are all clearly outlined in a concise event management strategy created by executing the following five processes [5]:

1. Establishing a general strategy for irrigation projects (maintenance or implementation).
2. Second, there are both large-scale and small-scale irrigation projects.
3. Identifying the logical relationships between activities in an irrigation project.
4. Using modern planning methods to schedule and present the plan.
5. Approval of the scheduling plan by stakeholders.

The irrigation projects are different from the other construction projects in [6]:

1. Every irrigation project is unique because it does not include repeated activities, whereas Other activities include repeating (typical) operations.
2. The irrigation project has a time frame that has a start and end point, as to meet stakeholders' requirements.
3. There is a pre-established budget for the irrigation project.
4. The irrigation project requires particular resources, which are provided to it at the project's beginning in an agreed-upon quantity of labor, materials, and equipment.
5. There is a lack of certainty that makes some risks in irrigation projects.
6. Bringing a favorable change: any irrigation project's goal is to expand the area that can be farmed and provide water quotas to interested parties.
7. The irrigation project passes in many stages that are completed in order, beginning with the feasibility and technical studies and continuing with engineering designs, material purchases, and execution stages.

An irrigation canal is a man-made waterway that is built on the ground to convey water from its supply to the field, the canal segment has a trapezoid shape, and it can take different shapes based on the nature of the supply source. Canals are classified into several types [2], as in Table 1.

Table 1: Classification of irrigation channels in Iraq

<table>
<tr><td rowspan="21">Main Channels</td><td rowspan="2">Depending On The Nature Of The Processing Source</td><td colspan="2">Temporary Or Intermittent Channel</td></tr>
<tr><td colspan="2">Permanent Channels</td></tr>
<tr><td rowspan="4">By Size</td><td colspan="2">Feeding Channel</td></tr>
<tr><td colspan="2">Secondary Channel</td></tr>
<tr><td colspan="2">Sub Channel</td></tr>
<tr><td colspan="2">Main Channel</td></tr>
<tr><td rowspan="7">By Job</td><td colspan="2">Irrigation Water Transfer</td></tr>
<tr><td colspan="2">Transporting Water To Power Station</td></tr>
<tr><td colspan="2">Navigational Channel</td></tr>
<tr><td colspan="2">Canal For Drinking Water And Industrial Uses</td></tr>
<tr><td colspan="2">Channel Is Switched Over Another Channel</td></tr>
<tr><td colspan="2">Vector Channel</td></tr>
<tr><td colspan="2">Multipurpose Channel</td></tr>
<tr><td rowspan="3">According To The Channel Surface</td><td colspan="2">Lined Channel</td></tr>
<tr><td rowspan="2">Unlined Channel</td><td>Grena Channel</td></tr>
<tr><td>Grena Non-Channel</td></tr>
<tr><td rowspan="3">By Alignment</td><td colspan="2">Catchment Channel</td></tr>
<tr><td colspan="2">Contour Channel</td></tr>
<tr><td colspan="2">Side Slope Channel</td></tr>
<tr><td rowspan="2">According To The Financial Return</td><td colspan="2">Productivity Channel</td></tr>
<tr><td colspan="2">Production Channel</td></tr>
</table>

2.2 Fieldworks

There are several irrigation projects on the Euphrates, west of Iraq, some of which are upstream. In Fallujah city are almost small or medium projects irrigated by pumping. Then, in the rest of the Euphrates, there are the Great Abu Ghraib project, Great Mussayab, HillaKifil, some small projects, and Kifil-Shinafiyah projects. In addition, the Hilla branch, Iraq's largest division from the Euphrates, is irrigating several irrigation projects. Also, the Hilla branch is the largest in Iraq from the Euphrates, where this branch is irrigating several irrigation projects [2].

Through personal interviews and field visits, data were collected, from some irrigation projects in the west of Iraq. At the stage of preparing and developing the questionnaire, many experts and specialists were consulted in the field of management scheduling and planning of irrigation projects. to collect the largest possible amount of information and data that lets a researcher have more understanding of the planning and management scheduling of irrigation projects in Iraq to enhance the study's value and its scientific credibility, this part included:

2.2.1 Data and information collection

It relied on two sources to collect data:

1. Brainstorming tools: This scientific approach is utilized when the mind has many thoughts that it wishes to filter down, modify, reorganize, or uncover interrelationships between these concepts so the individual may start preparing effectively for the job he is thinking about [4]. Following are some brainstorming techniques [1]: ”lists, figures, free writing, viewpoints, ideas map, cubism, parts press questions, and graphs”, use of knowledge sources, and a focus on the target and audience.
2. A questionnaire is a means of communication with experts, skilled engineers, and specialists, taking advantage of the information base at their disposal and building on it to make good choices [8, 3]. The questionnaire was created and designed in accordance with the study's research questions and hypotheses.
3. Books, periodicals, and previous studies related.

2.2.2 Selection of study tools

Several irrigation project sites in western Iraq's Anbar Governorate have been chosen; some of these projects are currently being implemented and maintained, while others have already been completed. Gathered information pertains to a wide range of problems encountered while developing and observing the scheduling management plan for these projects.

To create and monitor the irrigation projects' implementation timetable management plan, it is necessary to understand the nature of the work involved, we conducted one-on-one interviews with a wide range of experts in the field, including engineers, contractors, and consultants working at different levels of the relevant administrative structures. The goal of these conversations is to get their input on the technical accuracy, form, and validity of the chosen data axes so that we can find workarounds for issues and gain a more grounded understanding of the challenges that were faced during the scheduling preparation process.

3 Develop mathematical models for predicting the duration of irrigation projects using SVM

This study aims to create and assess a mathematical model that could be applied to determine, using the SVM technique, how long future irrigation projects will last, as a result of the research fieldwork, ended with a list of variables affecting the duration of irrigation projects, arranged in an order of importance, and to attempt in an account of developing and evaluate of a duration estimating model through the following steps:

- Diagnosing and identifying the variables in the SVM model.
- Development and assessment of the suggested SVM models for irrigation project implementation. This stage involves three steps: diagnosing and identifying neural network model variables, designing and modeling a model, and validating the development model.
- The creation and assessment of the suggested SVM models for irrigation maintenance.
- Assessing the significance of factors that have an impact on how long irrigation projects take to build and maintain.
- Examining the reliability and validity of the development model.

3.1 SVM variable identification

Six randomly selected and well-defined independent variables, as well as one dependent variable, were included in the historical data for each irrigation project. Furthermore, the SVM model has adopted two types of variables that determine the duration of the implementation project: dependent and independent variables:

1. In depended Variables (SVM Output): The six components of this type were as follows:

   LC refers to channel length in (m) 300

   AC refers to the contract's value (I.D)

   FI refers to the fill, in (m^3)

   CU refers to the cut in (m^3)

   CA refers to a concrete area in (m^2)

T.H refers to concrete thickness in (cm)

2. Dependent Variable (ANN Input): There was just one variable in this kind.

   D refers to contract duration in (day)

3.2 Model with Data Division SVM

Available data has been divided into three subsets: training, testing, and validation set, in this step of the SVM model building. Using the rescaling technique covariates, Data can be split in such a way that testing errors are minimized and coefficient correlations are maximized.

3.3 SVM Model Variables Identification

Historical data comprised six randomly selected and clearly specified independent factors and one dependent variable for each irrigation project. The SVM model employed two kinds of variables to predict project length. Variables dependence:

1. Independent Variables (SVM Output): There were six parameters in this type, and they were as follows: **LC**: refers to channel length in (m) 300, **AC**: refers to contract amount (I.D), **FI**: refers to the fill, in (m^3), **CU**: refers to the cut in (m^3), **CA**: refers to a concrete area in (m^2), **TH**: refers to concrete thickness in (cm)
2. Dependent Variable (SVM Input): As can be seen in the table below, only one variable existed for this category:

   D refers to contract duration in (day).

4 Predicting the periods of irrigation project using SVM

4.1 Support Vector Machines

It's supervised machine learning for classification or regression. It's used to classify. SVM displays each data point as a point in n-dimensional space, where n is the maximum number of features a feature may have and the value of a feature is a specific coordinate [7].

4.2 Factors affecting the periods of the irrigation project

Through the field study and review of the literature, which is related to the study of the factors influencing the estimated duration of irrigation channel projects, the researchers identified the variables affecting the periods of irrigation projects as in Table 2.

Table 2: Affecting variable's the periods of the irrigation project

Variables	**Description**	**Max**	**Min**	**S.D**
V1	Security Condition	1	1	0.14
V2	Channel length	55	22	9.53
V3	Work location	32	4	8.32
V4	The number of machines used	4	1	0.59
V5	Efficiency of machines used	2	1	0.32
V6	Irrigation channel type	2	1	0.27
V7	The type of soil	2	1	0.47
V8	Topography of the area	8	4	1.14
V9	Investigation accuracy	2	1	0.42
V10	Weathers condition	3	1	0.59
V11	Bill of Quantity Accuracy	3	1	0.57
V12	Fewer errors in channel designs	2	1	0.48

The data were split into training and validation divisions, full training is used in the testing set, as in Table 3.

Table 4, shows the factors affecting the prediction implementation periods of channel irrigation projects, which are (15) influencing (Bill of quantity accuracy) because this factor has a relative importance of 84%, then comes the factor (Availability of cash flow) with a relative importance of 82%, then the factor (Efficiency of machine used) with an important rate of 81%, until we reach the least relative importance Weathers condition.

V13	Work Quality	2	1	0.44
V14	Availability of cash flow	2	1	0.19
V15	Skill of machine operators	2	1	0.39

Table 3: Effect of data division on the periods of irrigation project SVM model

Training%	Validation%
77	23

Table 4: Affecting variable's the periods of the irrigation project

Variables	Description	RII
V1	Security Condition	75%
V2	Channel length	77%
V3	Work location	67%
V4	The number of machines used	80%
V5	Efficiency of machines used	81%
V6	Irrigation channel type	72%
V7	The type of soil	76%
V8	Topography of the area	65%
V9	Investigation accuracy	63%
V10	Weathers condition	59%
V11	Bill of Quantity Accuracy	84%
V12	Fewer errors in channel designs	81%
V13	Work Quality	78%
V14	Availability of cash flow	82%
V15	Skill of machine operators	79%

4.3 Kernel-selection of SVM-CRP model

The kernel functions, like those used in kernel tables, return the inner product between any two points in a suitable feature space, defining a notion of similarity in the process, Table 5, is determined. The best option is considered to be a poly kernel of the model with a Root square mean error (0.406) As a result, it was chosen for this model.

Table 5: Performance of the kernel on the (SVM) model

Kernel Type	RMSR	MAE	Coefficient Correlation(r)%
poly kernel	0.456	0.26	82%
normalized poly kernel	0.692	0.477	72%
RBF kernel	0.491	0.32	78%

4.4 Parameter SVM (C and Epsilon) SVM model

Parameter C affects the SVM Model as in Table 6; the parameter with the greatest value C=2 produced the highest correlation coefficient (82%), consequently, and applied to the model.

The SVM Model performance is mostly unaffected by the different values of parameter C as in Figure 1, shown below, especially in the range of (1 to 10). C= 2 when it with a mean absolute error of 0.252, a root means a square error of 0.466, and a maximum correlation coefficient (r) of 80.68%; for these reasons, it was utilized in the mode in Figure 1.

4.5 Parameter Epsilon of SVM duration model

The parameter epsilon affects displays of the SVM model as in Table 7. The parameter with the best value, epsilon=0.02, had the highest correlation coefficient (81.71%) and was therefore chosen in the model in Figure 2.

Figure 2, indicates that the SVM Model's effect on various parameter epsilon is largely unaffected by them, especially in the region of (0.001-0.02) skilling toward rising. Epsilon=0.02 has a mean absolute error value is 0.254,

Table 6: The effort of changing the parameter C on the SVM model performance

parameter C	RMSE	MAE	Correlation Coefficient
1	0.434	0.2507	80.23%
2	0.446	0.252	80.68%
3	0.435	0.2504	79.26%
4	0.437	0.25	79.29%
5	0.433	0.25	79.21%
6	0.436	0.25	79.22%
7	0.4347	0.25	79.12%
8	0.4146	0.25	79.19%
9	0.4147	0.25	79.24%
10	0.4146	0.25	79.23%

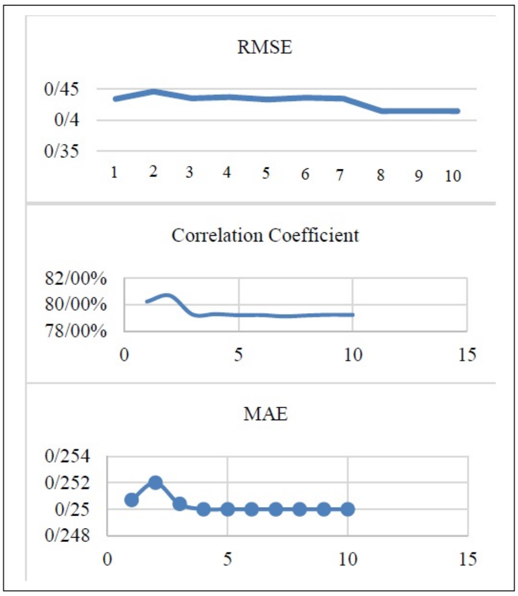

Figure 1: Effect of various parameters C on SVM model

Table 7: Parameter epsilon affects the SVM model

MAE	RMSE	Correlation Coefficient
0.254	0.426	81.71%

the lowest significant root mean square error value of 0.426, and the highest correlation coefficient (r=81.71%); hence it was included in this model. Therefore the parameter performance becomes like in Table 8.

Table 8: Effect of various parameter epsilons on SVM

Parameter Epsilon	RMSE	MAE	Correlation Coefficient
0.001	0.416	0.253	80.65%
0.002	0.416	0.252	80.68%
0.003	0.416	0.254	80.68%

4.6 Model Equation: (SVM) Duration Model

The connection weights for the optimal SVM Brick Model as determined by the Weka algorithm are presented in Table 9.

The estimated duration of the irrigation project can be represented using the levels threshold and linking weights shown in Table 9, as follows:

0.004	0.416	0.255	80.70%
0.005	0.426	0.252	80.68%
0.006	0.426	0.252	80.68%
0.007	0.426	0.253	80.70%
0.008	0.426	0.253	80.69%
0.009	0.427	0.255	80.71%
0.01	0.426	0.253	80.72%
0.02	0.426	0.254	81.71%
0.03	0.255	0.427	79.81%
0.04	0.256	0.426	80.77%
0.05	0.257	0.426	80.75%

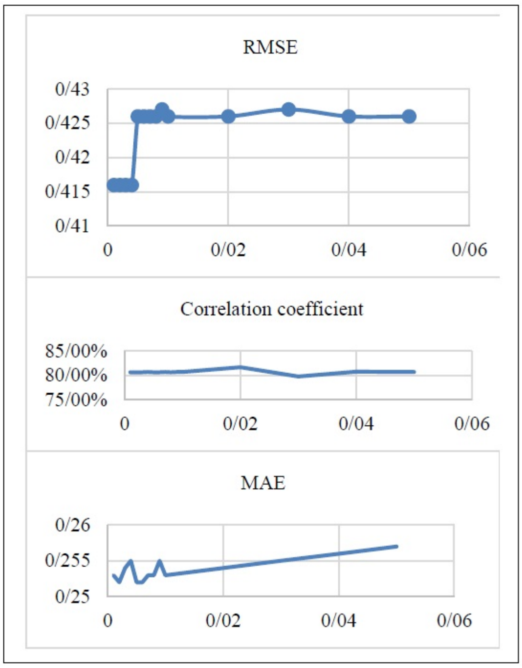

Figure 2: Effect of various parameter epsilons on SVM duration model performance

Table 9: The threshold, levels, and weights for the ideal (SVM) Duration Model

Layer	**wji (weight from node i in the input layer to node j in the hidden layer)**								
Input	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9
Weight	0.3361	0.0192	−0.103	−0.0432	0.421	−0.165	0.0893	−0.0768	−0.0761
Input	V_{10}	V_{11}	V_{12}	V_{13}	V_{14}	V_{15}			
Weight	−0.3534	−0.834	−0.163	0.199	−0.431	−0.0281			
Bais					5.107				

$$
\begin{aligned}
\text{SVM duration Model} = \{5.107 \; &+(0.3361 \times V_1) + (0.0192 \times V_2) + (-0.103 \times V_3) + (-0.0432 \times V_4) + (0.421 \times V_5) \\
&+(-0.165 \times V_6) + (0.0893 \times V_7) + (-0.0768 \times V_8) + (-0.0761 \times V_9) + (-0.353 \times V_{10})\} \\
&+(-0.834 \times V_{11}) + (-0.163 \times V_{12}) + (0.199 \times V_{13}) + (-0.431 \times V_{14}) + (-0.0281 \times V_{15})\}
\end{aligned} \tag{4.1}
$$

4.7 Model validity of the (SVM) duration

The results of the comparison study are represented in Table 10, the SVM duration model's MAPE and average accuracy percentage were found to be (17%) and (83%) respectively. As a result, the SVM duration model and the real value correspond well.

Table 10: Result of the SVM duration model

Description	**(NO. of observation 23)**
MAPE	17
RMSE	0.85
AA%	83
R	78.31%
R^2	61.32%
MPE	2.77

The MAPE Table 10 was used to determine the error class of this model, by giving this table, the MAPE of the SVM model is excellent. As a result, training the network and checking the complexity of the SVM model takes more time for high prediction accuracy for a validation data set, forecast values are displayed alongside actual values, to assess the reliability of the SVM model's capabilities. R for 23 observations equals (78.31%).

5 Conclusions

1. The study explored the capabilities of machine learning algorithms (Prediction of the duration for channel irrigation using a Support Vector Machine model).
2. Fifteen variables were adopted to develop the model. The influencing factors in estimating durations in Iraq, include (Bill of quantity accuracy) with the first and main (84%) importance place and then followed by (Availability of cash flow) with the relative importance of (82%).
3. The findings illustrate that the predicting in support of vector machine was more precise with average accuracy (83%) and correlation coefficient (R) equal (0.6132).

The study showed that most of the factors that affect the delay of irrigation channel projects are in the planning stage and that the use of proper planning and forecasting of duration helps the completion of projects within the specified period

References

[1] R. Abd Karim, A.G. Abu and F.N.M. Khaja, *Brainstorming approach and mind mapping in writing activity*, Proc. English Educ. Inte. Conf. **1** (2016), no. 2, 423–429.

[2] M. Abdullah and N. Al-Ansari, *Irrigation projects in Iraq*, J. Earth Sci. Geotech. Eng. **11** (2021), no. 2, 35–160.

[3] J.A. Al-Somaydaii, *Development mathematical model for brick works productivity by using support vector machine*, Int. J. Appl. Eng. Res. **11** (2016), no. 23, 11126–11131.

[4] T.M.-A.-A. Al-Suwaidan and M.A. Al-Adlouni, *Principles of creativity*, Cordoba for Publishing and Distribution, 2004.

[5] D. Gough, J. Thomas and S. Oliver, *Clarifying differences between review designs and methods*, Syst. Rev. **1** (2012), no. 1, 1–9.

[6] A.A. Hameed and F.M.S. Al-Zwainy, *Statistical evaluation of the planning and scheduling management process for irrigation and drainage projects*, J. Algebra Statist. **13** (2022), no. 2, 259–282.

[7] S. Ray, *Understanding support vector machine (SVM) algorithm from examples (along with code)*, https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/, 2017.

[8] Saher, *Forecasting of factors affecting brickwork productivity estimation by using artificial neural network*, Thesis Master of Science in Civil Engineering, Baghdad University, 2014.

[9] Z.M. Yaseen, Z.H. Ali, S.Q. Salih and N. Al-Ansari, *Prediction of risk delay in construction projects using a hybrid artificial intelligence model*, Sustainability **12** (2020), no. 4, 1514.