Jurnal Sosial Humaniora (JSH)
2022, Volume 15, Ed.1
ISSN Online: 2443-3527
ISSN Print: 1979-5521

Nonlinear EVM Based on Support Vector Regression Growth Model for Predicting Project Completion Time

Tri Joko Wahyu Adi[1], **Arif Abadil Ghullam**[2]
[1]Civil Engineering Department, Sepuluh Nopember Institute of Technology, Surabaya, Indonesia, 60111
[2]Management of Technology Department, Sepuluh Nopember Institute of Technology, Surabaya, Indonesia, 60264, arif.abadil.g@gmail.com

Received: 26/10/2021.
Reviewed: 16/06/2022.
Published: 31/07/2022.

Subject Area: Environmental and Demographic Change

Abstract

Earned Value Management (EVM) is a method used to monitor and to predict project completion time. This method uses a linear approach in predicting project time completion. Unfortunately, most of the projects run in dynamic environments with complex characteristics, causing project progress to require a non-linear approach. That is why the use of EVM in monitoring and predicting non-linear project completion time is less effective. This study proposes a more realistic alternative approach using non-Linear EVM based on the Support Vector Regression (SVR) - Growth Model. The SVR-growth model is used to accommodate the non-linear progress of the project, while the EVM is used to represent the predicted results of the project completion time. For model validation, 5 data on oil and gas field development construction projects in Jawa, Bali and Nusa Tenggara Regions were used as case studies. The results of this study indicate that the results of project completion time prediction using the SVR-Growth Model provide high accuracy and precision compared to the traditional EVM method.

Keywords: *Support Vector Regression; Non-Linear Growth Model; Earned Value Management.*

Introduction

Project managers are responsible for fulfilling project completion according to planned cost and time. To perform those tasks, project managers were assisted with project monitoring tools. The Earned Value Management (EVM) method is commonly used as a monitoring tool for project supervision and control. EVM integrates scope, time, and cost in one monitoring system. EVM has a formula that is used to predict the project final cost and completion time (Vanhoucke M. , 2014).

Along with the development changes, construction activities are increasingly complex. The phases of the project work package can run simultaneously in parallel and become more dynamic. This becomes difficult to observe with traditional EVM approaches that have assumptions and simplification. One of them is a linear approach in predicting the completion time. This results in the accuracy of the prediction results using this traditional EVM method (Warburton & Cioffi, 2016).

To improve these deficiencies, a non-linear EVM approach has been carried out through the concept of a growth model (Narbaev & De Marco, 2013). The growth model is used to formulate the non-linear character of the project through the population growth formulation (Warburton, De Marco, & Sciuto, 2017). However, the growth model has limited prediction on projects with slow start due to the limitations of the population growth formulation in recognizing the non-linear EVM curve profile (Warburton, De Marco, & Sciuto, 2017).

The ability to recognize project characters in the above growth model concept is enhanced with the help of artificial intelligence (Willems & Vanhoucke, 2015) (Vanhoucke M. , 2019). The project data information is too large and complex to be investigated manually. Complex computation of project data sets is accomplished using the aid of algorithms in machine learning (Huang, 2009) (Prieto, Prieto, Ortigosa, Ros, & Pelayo, 2016). One of the learning machine methods is model building using the Support Vector Regression (SVR) method. SVR has the advantage of kernel tricks process that can accommodate non-linear regression modeling (Vanhoucke & Wauters, 2016) (Peško, et al., 2017).

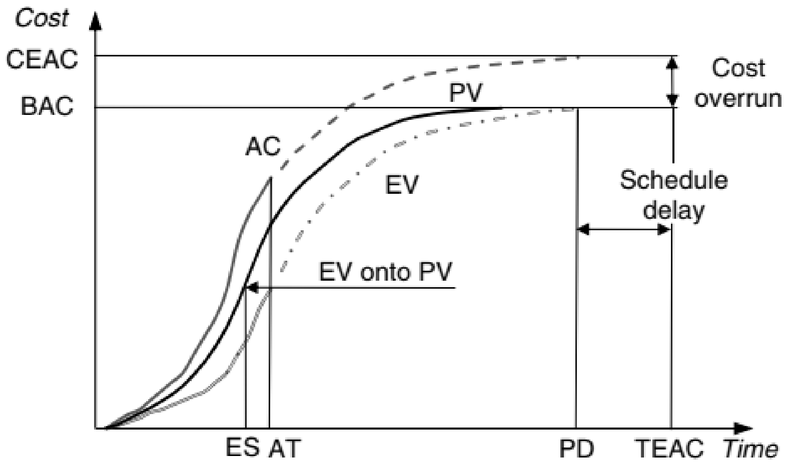
This study proposes a non-linear prediction model of project completion time by combining SVR with the growth model EVM, to increase the ability on recognizing non-linear project progress patterns. Besides being able to provide more accurate prediction results, this proposed model can also be used as an Early Warning System for project managers to meet the target project completion time.

Literature Review

Earned Value Management

The project is supervised and controlled by evaluating current achievements and anticipating upcoming deviance events. Various tools developed to closely control a project, one of them is the Earned Value Management methods (EVM) (Vanhoucke M. , 2014). EVM is a tool that is able to measure project progress that have been achieved and integrating time, cost and scope of works in one monitoring system. During its improving method, the terms of cost-based project completion are enhanced by the time-based terms called Earned Schedule (ES) (Lipke, 2003). The EVM concept with Earned Schedule improves the time-based prediction formulation as described in Figure 1. EVM Concept Illustrationbelow (Narbaev & De Marco, 2014):

Figure 1. EVM Concept Illustration



Earned Schedule as the achievement of the planned project phase based on time is measured by projecting the actual achievements of the current project phase to the planning curve (PV). ES can be used to see how much time deviation (SV) becomes the actual progress against project planning outcomes.

Index Based Approach

In traditional EVM, the prediction of project completion time is done by using the current schedule performance index (SPI) formulation in completing the remaining project plan (planned duration (PD) minus ES) (Willems & Vanhoucke, 2015; Vanhoucke M. , 2012) described as follows:

$$EAC_{(time)} = AT + \frac{(PD-ES)}{SPI} \quad (1)$$

$$SPI = \frac{ES}{AT} \quad (2)$$

This formulation applies the assumption of simplifying project performance that is used until the project is completed according to the slope between the project planned time achievement and the actual project time achievement (ES/AT). Meanwhile in a dynamic project, the slope between planned achievement and actual achievement changes throughout time until project is completed. Making predictions using the traditional EVM approach on dynamic projects will give less accurate prediction results, especially during the early stages of the project (Narbaev & De Marco, 2013).

Regression Based Approach

To reduce its limitation, several studies have developed EVM-based predictions with non-linear regression formulations and S-curve fitting using a growth model formulation (Narbaev & De Marco, 2014). The regression formula will describe an ongoing project through the shape of a plan curve (PV) or an actual curve (EV). After getting the regression formulation that fits on the PV curve, the formulation is used to extrapolate the EV curve to get the total cost or completion time (Warburton, De Marco, & Sciuto, 2017). Those growth model formulation has limitations in recognizing the dynamic pattern of a project because it is limited by the flexibility of the growth model regression formulation, so analyzing from this situation requires more realistic modeling formulation (Willems & Vanhoucke, 2015).

Support Vector Regresion

Artificial intelligence has ability to recognize, filter information, generalize and learn from previous works. Artificial intelligence is built by machine learning perceptron. Perceptron is an encoding algorithm architecture for decision boundary functions. Perceptron consists of an aggregation function and a transfer function. The aggregation function collects the weights of all related input signal components. The transfer function filters the aggregation results and classifies them to form decision responses. Perceptron architecture allows independent search to obtain a decision boundary based on input data (Hamel, 2009).

The predictor variable is transformed into a vector form. A vector will have a unit of length (vector unit) and direction, making it easier to measure the similarity score between vectors through the dot product. In multidimensional vectors, the decision boundary will be obtained in the form of a convex field (hyperplane). Hyperplane on the support vector is built through a dual approach. The dual approach encourages the algorithm to look for a hyperplane shape from the trainer's data input using normal direction search and use offset terms as a constraint when constructing the hyperplane. The concept of forming a hyperplane using a vector is called a support vector (Vapnik, 1998; Hamel, 2009).

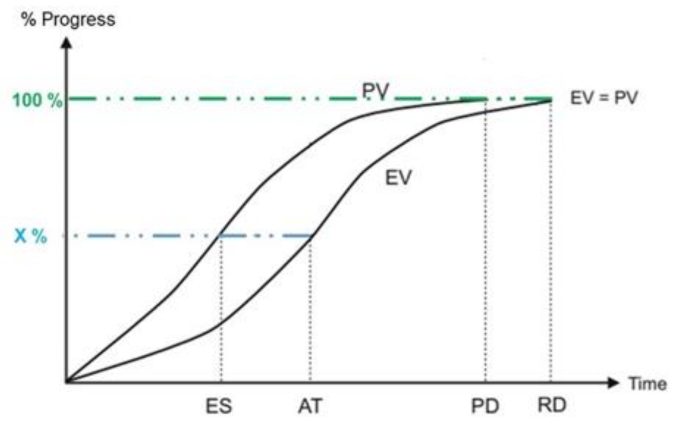
Support Vector usually used in classification and regression operations. Support Vector Machine in classification is used to separate classes through a decision surface while Support Vector Regression formulates the decision surface as a regression function between response variables and predictor variables (Drucker, Burges, Kaufman, Smola, & Vapnik, 1997; Gunn, 1998; Vapnik, 1998). Pattern recognition of non-linear regression is accommodated through the kernel trick approach using the Radial Basis Function (RBF) kernel function (Schölkopf & Smola, 2002; Pingping, Bin , Haihui , & Hu , 2019).

Machine learning has been widely used in the field of project management (Vanhoucke & Wauters, 2016; Peško, et al., 2017). Among several artificial intelligence methods, Support Vector Regression (SVR) has the advantage of performing pattern recognition of a project from the experience of other similar project data (Cheng, Peng, Wu, & Chen, 2010; Cheng, Hoang, Roy, & Wu, 2012; Vanhoucke & Wauters, 2014).

Methodology

This research is a quantitative study to observe the causal relationship between two or more variables. Project supervision and control variables were observed using the earned value management method. The growth model approach shows that each PV and EV curve has its own relationship between the project phase achievement against time (Narbaev & De Marco, 2013; Warburton, De Marco, & Sciuto, 2017; Narbaev & De Marco, 2014), this relationship is illustrated in Figure 2. Growth Model Concept in EVMbelow.

Figure 2. Growth Model Concept in EVM



From the illustration above, ES has its own suitability relationship with the achievement of the project planning phase curve (PV) and AT has its own suitability relationship with the achievement of the actual project phase curve (EV). When the project is runs (grows) to a certain time (X%) where the actual project achievement (EV) is the same at the project plan is achieved (PV). It will show the point in time that should have been achieved (ES) and the actual point in time where the project has just been achieved (AT). When the project stage is completed or the progress reaches 100%, the project according to the project plan will be completed on planned duration (PD) and in reality the project is completed in Real Duration (RD).

The model proposed in this study adopts the growth model concept above combined with the non-linear regression learning machine formulation from Support Vector Regression, to formulate a model of the relationship between the EVM variables on each curve. Modeling data is made from historical oil and gas field development project data in the Jawa, Bali and Nusa Tenggara regions, as shown in Table 1. Historical Project Databelow:

Table 1. Historical Project Data

No	Project Name	Production Capacities (boepd)	Start	Completed Plan
1	Proyek A	± 4.000	April 2018	August 2019
2	Proyek B	± 36.000	March 2017	March 2019
3	Proyek C	± 3.500	January 2018	June 2020
4	Proyek D	± 8.500	September 2018	September 2020
5	Proyek E	± 3.000	April 2018	October 2020

In each project, the progress (%) of the actual project stage achievements is used as EV and the project phase plan is used as PV at each point t. Data is processed as initial data preparation before being processed as learning process data. ES is measured by projecting the current achievements AT to the PV curve using the following formula:

$$SV_{(t)} = t + \left(\frac{EV - PV_t}{PV_{t+1} - PV_t}\right) - AT \tag{3}$$

$$ES_{(t)} = SV_{(t)} + AT \tag{4}$$

The AT and ES values that have been calculated above then normalized to the planned duration (PD) in Figure 2. Growth Model Concept in EVMso the learning machine process can precisely represents the project conditions. The results of the initial data preparation are tabulated in Table 2. SVR Training Databelow:

Table 2. SVR Training Data

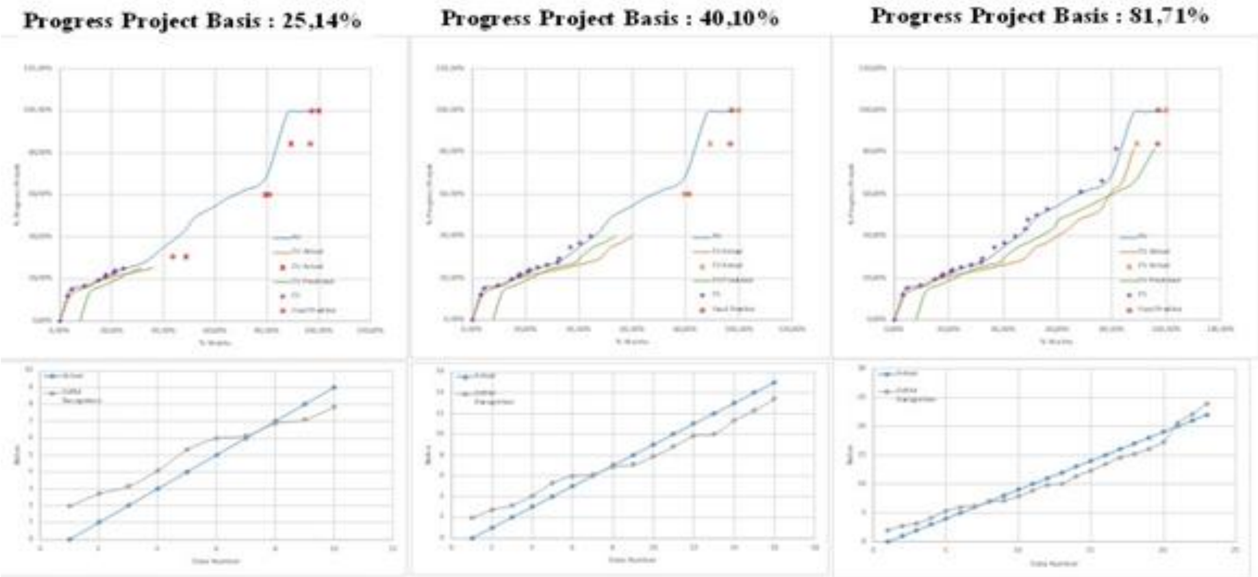
No. Data	Project Name	Earned Schedule	Progress Plan	Actual Progress	Actual Time	No. Data	Project Name	Earned Schedule	Progress Plan	Actual Progress	Actual Time	No. Data	Project Name	Earned Schedule	Progress Plan	Actual Progress	Actual Time
Variabel		Xa	Xb	Xc	Y	Variabel		Xa	Xb	Xc	Y	Variabel		Xa	Xb	Xc	Y
Categorial		Prediktor	Prediktor	Prediktor	Response	Categorial		Prediktor	Prediktor	Prediktor	Response	Categorial		Prediktor	Prediktor	Prediktor	Response
1	Project A	0,00%	0,00%	0,00%	5,88%	30	Project C	28,57%	9,59%	9,59%	30,00%	59	Project D	32,00%	70,00%	70,00%	36,00%
2	Project A	14,29%	0,21%	0,21%	11,76%	31	Project C	31,79%	11,55%	11,55%	33,33%	60	Project D	32,00%	70,00%	70,00%	40,00%
3	Project A	16,81%	1,17%	1,17%	17,65%	32	Project C	36,19%	13,96%	13,96%	36,67%	61	Project D	32,00%	70,00%	70,00%	44,00%
4	Project A	23,53%	3,09%	3,09%	23,53%	33	Project C	38,25%	14,38%	14,38%	40,00%	62	Project D	32,00%	70,00%	70,00%	48,00%
5	Project A	27,17%	4,99%	4,99%	29,41%	34	Project C	43,22%	15,17%	15,17%	43,33%	63	Project D	32,00%	70,00%	70,00%	52,00%
6	Project A	30,94%	7,33%	7,33%	35,29%	35	Project C	46,67%	18,55%	18,55%	46,67%	64	Project D	32,00%	70,00%	70,00%	56,00%
7	Project A	31,88%	9,14%	9,14%	41,18%	36	Project C	51,50%	24,75%	24,75%	50,00%	65	Project D	32,00%	70,00%	70,00%	60,00%
8	Project A	35,37%	10,67%	10,67%	47,06%	37	Project C	56,27%	28,15%	28,15%	53,33%	66	Project D	32,00%	70,00%	70,00%	64,00%
9	Project A	35,82%	12,19%	12,19%	52,94%	38	Project C	59,79%	31,79%	31,79%	56,67%	67	Project D	32,00%	70,00%	70,00%	68,00%
10	Project A	41,28%	13,60%	13,60%	58,82%	39	Project C	60,00%	36,29%	36,29%	60,00%	68	Project D	60,33%	82,00%	82,00%	72,00%
11	Project A	41,80%	16,96%	16,96%	64,71%	40	Project C	63,33%	41,04%	41,04%	63,33%	69	Project D	60,06%	82,00%	82,00%	76,00%
12	Project A	47,62%	23,56%	23,56%	70,59%	41	Project C	66,67%	46,24%	46,24%	66,67%	70	Project D	60,04%	82,00%	82,00%	80,00%
13	Project A	52,97%	29,05%	29,05%	76,47%	42	Project C	70,00%	49,55%	49,55%	70,00%	71	Project D	60,04%	82,00%	82,00%	84,00%
14	Project A	59,12%	37,61%	37,61%	82,35%	43	Project C	73,33%	58,19%	58,19%	73,33%	72	Project D	60,04%	82,00%	82,00%	88,00%
15	Project A	65,05%	50,87%	50,87%	88,24%	44	Project C	76,54%	74,84%	74,84%	76,67%	73	Project D	60,04%	82,00%	82,00%	92,00%
16	Project A	66,44%	63,54%	63,54%	94,12%	45	Project C	79,34%	86,26%	86,26%	80,00%	74	Project D	60,04%	82,00%	82,00%	96,00%
17	Project A	71,67%	70,75%	70,75%	100,00%	46	Project C	79,39%	87,11%	87,11%	83,33%	75	Project E-1	5,47%	20,76%	20,76%	8,33%
18	Project A	84,29%	85,25%	85,25%	105,88%	47	Project C	79,46%	87,91%	87,91%	86,67%	76	Project E-1	5,52%	26,73%	26,73%	16,67%
19	Project A	85,47%	89,67%	89,67%	111,76%	48	Project C	83,46%	89,74%	89,74%	90,00%	77	Project E-1	4,74%	26,73%	26,73%	25,00%
20	Project A	89,87%	92,92%	92,92%	117,65%	49	Project C	89,85%	90,51%	90,51%	93,33%	78	Project E-1	4,61%	30,38%	30,38%	33,33%
21	Project A	89,97%	93,08%	93,08%	123,53%	50	Project C	93,33%	90,52%	90,52%	96,67%	79	Project E-1	8,83%	33,34%	33,34%	41,67%
22	Project C	3,09%	1,01%	1,01%	3,33%	51	Project D	3,71%	7,71%	7,71%	4,00%	80	Project E-1	17,26%	42,29%	42,29%	50,00%
23	Project C	6,28%	2,46%	2,46%	6,67%	52	Project D	7,75%	18,54%	18,54%	8,00%	81	Project E-2	6,01%	30,70%	30,70%	43,75%
24	Project C	9,53%	4,47%	4,47%	10,00%	53	Project D	9,73%	25,54%	25,54%	12,00%	82	Project E-2	47,97%	36,28%	36,28%	50,00%
25	Project C	11,28%	5,07%	5,07%	13,33%	54	Project D	13,02%	35,20%	35,20%	16,00%	83	Project E-2	52,04%	43,68%	43,68%	56,25%
26	Project C	11,29%	5,20%	5,20%	16,67%	55	Project D	16,90%	39,76%	39,76%	20,00%	84	Project E-2	52,25%	48,10%	48,10%	62,50%
27	Project C	11,54%	5,46%	5,46%	20,00%	56	Project D	20,31%	44,15%	44,15%	24,00%	85	Project E-2	52,37%	51,90%	51,90%	68,75%
28	Project C	17,65%	6,24%	6,24%	23,33%	57	Project D	23,32%	65,22%	65,22%	28,00%	86	Project E-2	57,26%	59,72%	59,72%	75,00%
29	Project C	26,67%	7,74%	7,74%	26,67%	58	Project D	26,48%	69,89%	69,89%	32,00%	87	Project E-2	62,54%	65,58%	65,58%	81,25%
												88	Project E-2	63,83%	72,05%	72,05%	87,50%

In SVR, machine learning will transform the predictor variable into a vector with as many dimensions as the number of predictor variables. In this study, there are 3 dimensions called EV and ES. Support Vector Regression will use each of these data as vectors to form a hypertube on the hyperplane.

Result and Discussion

The modeling above uses the Support Vector Regression tools available in Matlab 2018b. The modeling results obtained are used to test project B data at various stages of the project which can be summarized in Figure 3. Model Result Databelow:

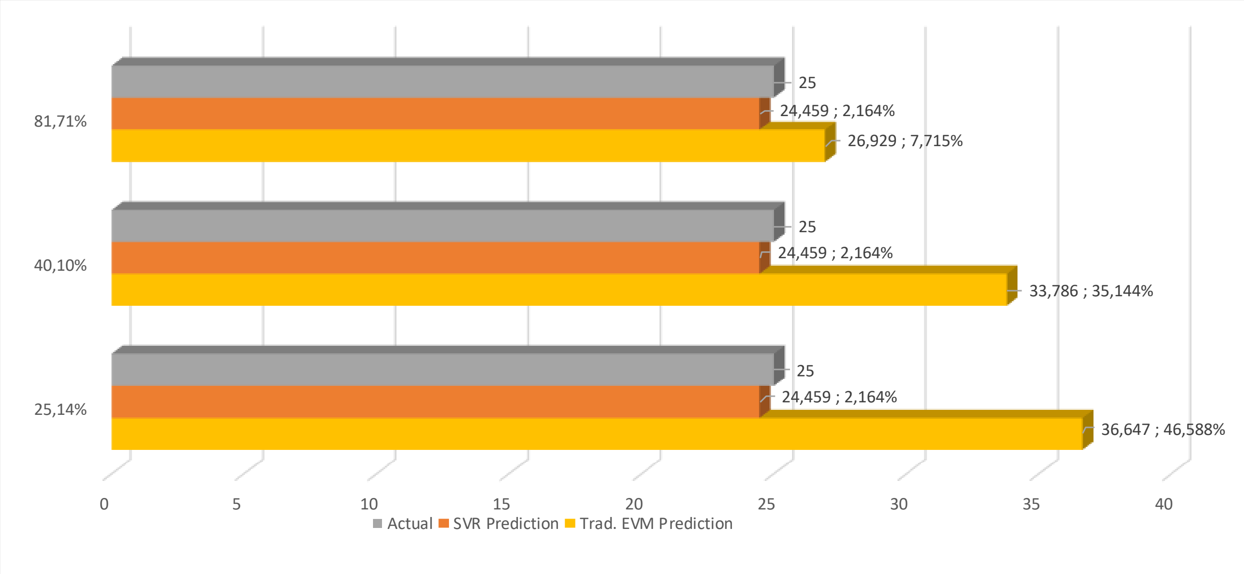
Figure 3. Model Result Data



% Data Actual Basis	25,14%	40,10%	81,71%
MAE of Pattern Recognition	4,198%	4,945%	5,266%

From the three groups of graphs above, it can be seen that the ability of SVR modeling in recognizing patterns to actual progress has a stable error ranging from 4% to 5%. The SVR Model that has been obtained then used to predict the project completion time based on the various levels of the actual project phases. These results are compared with the prediction results using traditional EVM method as shown in the Figure 4. Prediction Comparison Results between SVR Model and Traditional EVMbelow:

Figure 4. Prediction Comparison Results between SVR Model and Traditional EVM



From the data above shows that the SVR Model provides prediction results with a better accuracy and precision based on various level of actual project achievements.

Conclusion

The new approach method proposed in this study uses a learning machine Support Vector Regression Growth Model combined with the Earned Value Management method can be applied to model non-linear patterns during project implementation. The experimental results show that the resulting modeling is able to recognize the project's progress profile and provide better and more accurate prediction results since the early stages of the project compared to the traditional EVM-ES method. Further studies can be related to parameters that can affect the prediction results of the Support Vector Regression Growth Model. These parameters can come from internal risk factors or external risk factors.

References

Cheng, M.-Y., Hoang, D.-N., Roy, A. F., & Wu, Y.-W. (2012). A Novel time-depended evolutionary fuzzy SVM inference model for estimating construction project at completion. *Engineering Application of Artificial Intelligence*, 744 - 752.

Cheng, M.-Y., Peng, H.-S., Wu, Y.-W., & Chen, T.-L. (2010). Estimate at Completion for construction projects using Evolutionary Support Vector Machine Inference Model. *Automation in Construction, 19*, 619-629.

Drucker, H., Burges, C., Kaufman, L., Smola, A., & Vapnik, V. (1997). Support Vector Regression Machine. *Advances in Neural Information Processing Systems 9: Proceedings of the 1996 Conference* (pp. 155 - 161). MIT Press.

Gunn, S. (1998). *Support Vector Machines for Classification and Regression.* Southampton, UK: University of Southampton.

Hamel, L. (2009). *Knowledge discovery with support vector machines.* Hoboken, New Jersey: John Wiley & Sons, Inc.

Huang, Y. (2009). Advances in Artificial Neural Networks – Methodological Development adn Application. *Algorithms 2*, 973-1007.

Lipke, W. (2003). Schedule is different. *The Measurable News*, 31-34.

Narbaev, T., & De Marco, A. (2013). An Earned Schedule-based regression model to improve cost estimate at completion. *Internataional Journal of Project Management*, 1 - 12.

Narbaev, T., & De Marco, A. (2014). Combination of Growth Model and Earned Schedule to Forecast Project Cost at Completion. *Journal of Construction Engineering and Management*.

Peško, I., MuIenski, V., Šešlija, M., Radovic, N., Vujkov, A., Bibic, D., & Krklješ, M. (2017). Estimation of Costs and Durations of Construction of Urban Roads Using ANN and SVM. *Complexity*.

Pingping, X., Bin , L., Haihui , Z., & Hu , J. (2019). Optimization and control of the light environment for greenhouse crop production. *Scientific Reports*.

Prieto, A., Prieto, B., Ortigosa, E. M., Ros, E., & Pelayo, F. (2016). Neural networks : an overview of early research, current frameworks, and new challenge. *Neurocomputing*, 242-268.

Schölkopf, B., & Smola, A. (2002). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond.* London, England: The Massachusetts Institute of Technology.

Vanhoucke, M. (2012). Measuring the efficiency of project control using fictitious and empirical project data. *International Journal of Project Management, 30*, 252-263.

Vanhoucke, M. (2014). *Integrated Project Management and Control.* Switzerland: Springer International Publishing.

Vanhoucke, M. (2019). Tolerance limits for project control: An overview of different approaches. *Computers & Industrial Engineering 127*, 467 - 479.

Vanhoucke, M., & Wauters, M. (2014). Support Vector Machine Regression for project control forecasting. *Automation in Construction, 47*, 92-106.

Vanhoucke, M., & Wauters, M. (2016). A comparative study of Artificial Intelligence methods for project duration forecasting. *Expert Systems With Applications 46*, 249 - 261.

Vapnik, V. N. (1998). *Statistical learning theory.* New York: Wiley.

Warburton, R., & Cioffi, D. (2016). Estimating a project's earned and final duration. *International Journal of Project Management*, 1493 - 1504.

Warburton, R., De Marco, A., & Sciuto, F. (2017). Earned schedule formulation using nonlinear cost etimates at completion. *Journal of Modern Project Management*, 75 - 81.

Willems, L., & Vanhoucke, M. (2015). Classification of articles and journal on project control and earned value management. *International Journal of Project Management* , 1610 - 1634.