Journal of Novel Applied Sciences
Available online at www.jnasci.org
©2013 JNAS Journal-2013-2-S3/1054-1063
ISSN 2322-5149 ©2013 JNAS

Simulation and Performance Assessment between hybrid algorithms SVR-CACO and SVR-CGA to more accurate predicting of the pipe failure rates

Moosa Kalanaki[1] and Jaber Soltani[2*]

1. Master Degree in Artificial Intelligence, Aboureyhan Campus, University of Tehran, Iran
2. Assistant Professor of Irrigation and Drainage Engineering Department, Aboureyhan Campus, University of Tehran, Iran

***Corresponding author*: Jaber Soltani**

ABSTRACT: Nowadays many studies have been done for prediction pipe failure rates in urban and other places, each of them with effective parameters has own features. So, managers and those responsible for these systems should have a more accurate and real knowledge of the structural failures and breakage in the main water supply pipes. Several studies and methods have been introduced for predicting failure rates in urban water distribution network pipes by researchers, each of them has some special features regarding the effective parameters and many methods such as Classical and Intelligent methods are used, leading to some improvements. In this paper, effective parameters for predicting water distribution network are taken into two models (Hybrid SVR-CACO and SVR-CGA) and are compared with each other, an analysis and comparison of various types of kernel and loss functions is performed for SVM. This research is aimed at optimizing related parameters to SVM and selecting the optimal model of SVM for better pipe failure rate prediction by CACO and CGA.

***Keywords*:** Support vector regression, Continuous ant colony algorithm, Continuous genetic algorithm, Kernel functions, Pipe failure rates

INTRODUCTION

One of the main objectives of efficient management and optimal operation of urban water distribution network is to present a model for predicting breakages of urban water distribution networks. This leads to achieving some goals such as supplying sanitary potable water of a high quality and quantity as required Accordance with WHO standards, reducing the waste of water caused by breakages as well as lowering the repair, maintenance and rehabilitation costs.

To achieve the mentioned objectives, In the first stage events and factors influence in water distribution networks must be considered and the effects of pipeline characteristics on failures of pipes be determined so as to reduce the number of accidents through correct and targeted policy-making, and as a result more precisely predict the pipe breakage rates in distribution systems and take necessary actions for preventing them.

To show the importance of this issue, in 1998, about one million accidents have occurred, attributing to themselves more than 20 percent of the total Water and Wastewater Companies budget used for repairs and rehabilitation. Of course, 30 percent of the incidents have occurred to the distribution system pipes. In addition, studies show the maintenance costs of the traditional water sector has increased from 3 million dollars in 1999 to 10 million dollars in 2001 (Elahi Panah, 1998). In most cases, accidents and pipe failures occur as a result of several factors some of which being measurable such as age, length, diameter, depth and pressure of the pipes (Tabesh et al., 2009). Hence, order to provide a comprehensive model, all these factors should be considered. Many studies have employed with many methods, including ANN, ANFIS, fuzzy logic, and SVM in related field.

A model introduced that associated breakage factor with age. They presented an exponential model for predicting pipe failure rates (Shamir and Howard, 1979).

In another study with comparing among NLR, ANN and ANFIS methods with some effective parameters results of the comparisons indicated that ANN and ANFIS methods are better predictors of failure rates compared with NLR. The results of the comparison between ANN and ANFIS showed that ANN model is more sensitive to pressure, diameter and age than ANFIS; So, ANN was more reliable (Tabesh et al., 2009).

SVM techniques used non-linear regression for environmental data and proposed a multi-objective strategy, MO-SVM, for automatic design of the support vector machines based on a genetic algorithm. MO-SVM showed more accurate in prediction performance of the groundwater levels than the single SVM (Giustolisi, 2006).

Pressure sensitive and EPANET was used for estimating and Hydraulic modeling. EPANET results were used as SVM inputs. Research results showed that the leakage rate is predictable and the smallest changes are predictable using the employed sensors (Mashford, 2009).

A prediction model of the pipe break rate was first developed using genetic programming. which minimizes the annual cost of break repair and pipe replacement. Finally, the optimal pipe replacement time was determined by the model (Xu et al., 2013).

Rough set theory and support vector machine (SVM) was proposed to overcome the problem of false leak detection. For the computational training of SVM, used artificial bee colony (ABC) algorithm, the results are compared with those obtained by using particle swarm optimization (PSO).Finally; obtained high detection accuracy of 95.19% with ABC (Mandal et al., 2012).

In this paper, combined ANN and GA model has been used to determine the effective parameters in pipe failure rates in water distribution system using the combination of ANN and GA. ANN model was developed in order to related parameters of breakage with pipe failure rates. The results lead to minimize the simultaneous error rates (Soltani and Rezapour Tabari, 2012).

In this paper, effective parameters in predicting pipes failure rate of water distribution network are taken into two models Hybrid SVR-CACO and SVR-CGA compared with them, an analysis and comparison of various types of kernel and loss functions is performed for SVR. This research is aimed at optimizing parameters related to SVR and selecting the optimal model of SVR for better pipes failure rate prediction by CACO and CGA algorithms. So in this research compared hybrid SVR-CGA and SVR-CACO models to predict pipe failure rates in water distribution networks to reduce the number of events. By comparing these results with the other methods such as ANFIS, ANN, and ANN-GA that had been done in the past results show the SVR-CACO model has better performance than the other models especially in time elapse. Also, with combining SVR and CGA obtained better parameters proportional to the data type and the model shows better performance in accuracy.

MATERIALS AND METHODS

Support vector machines (SVMs) are learning machines which implement the structural risk-minimization inductive principle to obtain a good generalization on a limited number of learning patterns (Gonzalez-Abril et al., 2011). Support vector machine is an algorithm to maximize a mathematical function based on data sets. To create maximum margin, at first two adjacent parallel planes and a separator is designed. They get away from each other until they hit the data. The plane farthest from the others is the best separator (Carrizosa& Romero Morales, 2013).

Support vector machine regression (SVR) is a method to estimate the mapping function from Input space to the feature space based on the training dataset (Vapnik, 1992).

In the SVR model, the purpose is estimating w and b parameters to get the best results. w is the weight vector and b is the bias, which will be computed by SVM in the training process.

In SVR, differences between actual data sets and predicted results is displayed by ε. Slack variables are (ξ_i, ξ_i^*) considered to allow some errors that occurred by noise or other factors. If we don't use slack variables, some errors may occur, and then the algorithm cannot be estimated. Margin is defined as margin=$\frac{1}{||w||}$. Then, to maximize the margin, through minimizing $||w||^2$, the margin becomes maximized. These operations give in Equations (1- 3) and these are the basis for SVR (Vapnik, 1992).

$$\text{Minimize } \frac{1}{2}||w||^2+C\sum_{i=1}^{n}\xi_i+\xi_i^* \tag{1}$$

subject to :

$$y_i\left(w^T x_i+b\right)\geq 1-\xi_i \tag{2}$$

$$\xi_i,\ \xi_i^*\geq 0,\quad i=1,2,3,\ldots,n \tag{3}$$

C is a parameter that determines the tradeoff between the margin size and the amount of error in training and ξ_i, ξ_i^* are slack variables.

2.1. Kernel Functions

The basic idea of mapping input variable to the higher dimensional space is for easier separation by linear functions. Because it is difficult and more costly to work with high dimensional feature space, so we use feature space. A kernel function is a linear separator based on inner vector products and is defined as Equation. (4):

$$k(x_i, x_j) = x_i^T x_j \tag{4}$$

If the data points are moved using $\varphi: x \to \varphi(x)$ to the feature space (higher dimensional space), their inner products turn into Equation. (5) (Qi et al., 2013).

$$k(x_i,x_j)=\varphi(x_i)^T.\varphi(x_j) \tag{5}$$

x_i is the support vectors and x_j is the training data.

Kernel functions are equivalent to inner product in the feature space. Thus, instead of doing costly computing in feature space, we use kernel functions (Hofmann et al.,2008). Nonlinear kernel functions give more linear separable ability to the feature space; in the other words, we add so much to the dimensions that data are separated in a linear form.

The most important kernel functions and the related parameters are included in Table 1, where U and V are training and test data respectively.

Table 1. kernel functions used in SVR

Kernel function Type	Formula	Related variables
Gaussian RBF	$k = e^{-\frac{(u-v)(u-v)'}{2p_1^2}}$	P1 defines RBF function width ,like as $.\delta$
Exponential RBF	$k=e^{-\sqrt{\frac{(u-v)(u-v)'}{2p_1^2}}}$	P1 defines ERBF function width like as RBF.
Polynomial	$k = (UV + 1)^{p1}$	P1 determines Polynomial degree.
Spline	Z=1+UV+ $\left(\frac{1}{2}\right)$UV min(u,v) - $\left(\frac{1}{6}\right)\min(u,v)^3$ K=Prod(z)	

α_i is the vector of Lagrange multipliers and represent support vectors. If these multipliers are not equal to zero, they are multipliers; otherwise, they represent support vectors (Vapnik and Chapelle, 2000).
Here, w is equal to Equation. (6).

$$w=\sum_i(\alpha_i-\bar{\alpha}_i)x_i \tag{6}$$

Accordingly, the SVR function *F*(**x**) becomes the following function.

$$F(x)=\sum_{i=1}(\bar{\alpha}_i-\alpha_i)K(x_i,x)+b \tag{7}$$

Equation. (7) can map the training vectors to target real values, while allowing some errors. To minimize errors and minimize risks, the goal is to find a function that can minimize risks, Equation. (8).

$$R_{emp}[f]=\frac{1}{l}\sum_{i=1}^{l} c(x_i,y_i,f(x_i)) \tag{8}$$

$c(x_i, y_i, f(x_i))$, denotes a cost function determining how the estimation error will be penalized based on empirical data X. R_{emp} represents empirical risks. While dealing with a few data in very high dimensional spaces, this may not be a good idea, as it will lead to over-fitting and thus bad generalization properties. Hence, one should add a capacity control term, which in the SV results is to be $\|w\|^2$ and leads to the regularization of risk function (Smola and Scholkopf, 1998).

$$R_{reg}[f]=R_{emp}[f]+\frac{\lambda}{2}\|w\|^2 \tag{9}$$

In Equation. (9), $\lambda > 0$ and is used for regularization. Loss function determines how to penalize the data while estimating. In this study, epsilon and quad loss functions were used and compared with each other. As noted, the loss function can determine how to penalize SVM errors. ε-insensitive loss function shown in Fig.1.

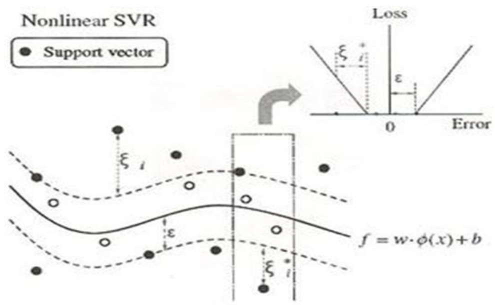

Figure 1. ε-insensitive loss function

A Loss function implies to ignore errors associated with points falling within a certain distance. If ε-insensitive loss function is used, errors between $-\varepsilon$ and $+\varepsilon$ are ignored. If C=Inf is set, regression curve will follow the training data inside the margin determined by ε (Smola and Scholkopf, 1998). The related equation is shown in Equation. (10).

$$|\xi|_\varepsilon = \begin{cases} 0 & \text{if } |\xi| \le \varepsilon \\ |\xi| - \varepsilon & \text{otherwise.} \end{cases} \tag{10}$$

The quadratic loss function Fig.2 drives the proximal hyper plane close enough to the class itself, and it penalizes every error (Qi et al., 2013). If quadratic loss function is used, memory requirements will be four times less than ε-insensitive loss function (Gunn, 1998).

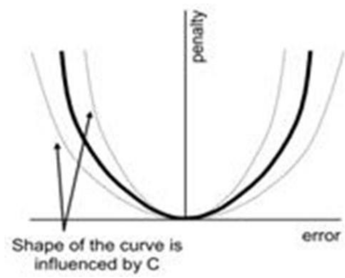

Figure 2. quadratic loss function

3. Searching Algorithms

Generally so far, the previous researches have appropriated SVM parameters obtained by trial and error. In the trial and error method, each parameter is tested to approach the appropriated values. However, this method has very time-consuming and is not sufficiently accurate.

Therefore, in this study integrated models compare and proposed to search for the possible solutions. As it noted, using the trial and error method is suitable for simpler and easier problems, but is not prone for complex ones of higher dimensional space, because it takes more time and ultimately might not converge to the solution. Therefore, due to the higher dimensions of the problem and the number of data, the integrated models were considered. Also, intelligent searching models have a high capability and suitable performance related to this problem; so, these algorithms were selected. In this study, search doing in practical solution with intelligent searching algorithms and the best of them selected for choosing an optimal SVR structure. Considering the fact that the parameters used in SVR and corresponding parameters are continuous in the solution Space, Continuous GA and Continuous ACO had been used, because when the variables are continuous, it is more logical to represent them by floating-point numbers.

3.1. Continuous ANT colony

The Ant colony optimization algorithm is an optimized technique for resolving computational problems which can be discovered good paths. The process by which ants could establish the shortest path between ant nests and food. Initially, ants leave their nest in random directions to search for food.

This technique can be used to solve any computational problem that can be reduced to finding better paths in a graph these formulas have been shown in Equations. (11) and (12). This method had been chosen from (Hong et al., 2011) paper.

$$P_k(i,j) = \begin{cases} \underset{s \in M_k}{\arg\max} \{[\tau(i,s)]^\alpha [\eta(i,s)]^\beta\} & \text{, if } q \le q0 \\ \text{Eq. (12)} & \end{cases} \tag{11}$$

$$P_k(i,j)=\begin{cases} \frac{[\tau(i,s)]^{\alpha}[\eta(i,s)]^{\beta}}{\sum_{s\epsilon M_k}[\tau(i,s)]^{\alpha}[\eta(i,s)]^{\beta}}, j\notin M_k \\ 0 \qquad \text{O.w} \end{cases} \tag{12}$$

where τ(i, j) is the pheromone level between node i and node j, μ(i, j) is the inverse of the distance between nodes i and j. In this study, the forecasting error represents the distance between nodes. The α and β are parameters determining the relative importance of pheromone level and M_k is a set of nodes in the next column of the node matrix for ant k. q is a random uniform variable [0, 1] and the value q_0 is a constant between 0 and 1, i.e., $q_0 \epsilon$[0, 1]. The local and global updating rules of pheromone are expressed as Equation. (13) and (14), respectively

$$\tau(i,j)=(1-\rho)\,\tau(i,j)+\rho\tau_0 \tag{13}$$

$$\tau(i,j)=(1-\delta)\tau(i,j)+\delta\Delta\tau(i,j) \tag{14}$$

The δ is the global pheromone decay parameter, 0 <δ< 1, and, based on authors' experiments.

The Δ^{τ} (i, j), expressed as Equation. (15), is used to increase the pheromone on the path of the solution

$$\Delta\tau(i,j) =\begin{cases} \frac{1}{L} & \text{, if (i,j)}\epsilon\text{globalbest route} \\ 0 & \text{O.W} \end{cases} \tag{15}$$

where L is the length of the shortest route.

At first to get SVR related parameters, each parameters show by 10 nodes, so the range of numbers limited between [0, 9]. For getting more accurate computed parameters 5 numbers considered.

The values ρ of and τ0 are set to be 0.2 and 1, respectively. Assume the limits of parameters σ, C, and ε are 1, 100,000, and 1, respectively. Numbers of nodes for each ant set to 50, so total nodes are equal to 150.

3.2. Continuous Genetic Algorithm

The continuous GA is inherently faster than the binary GA, because the chromosomes do not have to be decoded prior to the evaluation of the cost function (Haupt and Haupt, 2004). Thus, using the aforementioned variables like kernel parameters as decision variables in a population-based optimization strategy may be a way of constructing an optimal SVR. To cover the entire search space, the initial population was considered randomly, commensurate with the best fitness function Equation. (16) of each population; the best of them has been selected. Some properties of GA, such as the ability of solving hard problems, noise tolerance, easy to interface and hybridize, make them a suitable and quite workable technique for parameter identification of fermentation models (Angelova and Atanassov, 2012).

$$\text{Minimize cost:}\left|\bar{Y}_{pred}-Y_{train}\right| \tag{16}$$

where $\bar{Y}_{pred}$ and Y_{train} are predicting and training output, respectively. In this algorithm, the parameters must be optimized and determined by GA that includes iterations to reach convergence. At first, Number of chromosomes, mutation rates, and crossovers must be correctly determined to reach the best results. The next step, Objective function, decision variables and their constraints must be determined in this model. To start the optimization process, initial GA variables like mutation, crossover and selection rates must be determined. Also required parameters for SVR such as kernel and loss functions must be determined. These parameters vary according to the kernel functions and data types. Fig.3 shows SVR-CGA and SVR-CACO hybrid algorithms which predict the pipe failure rates.

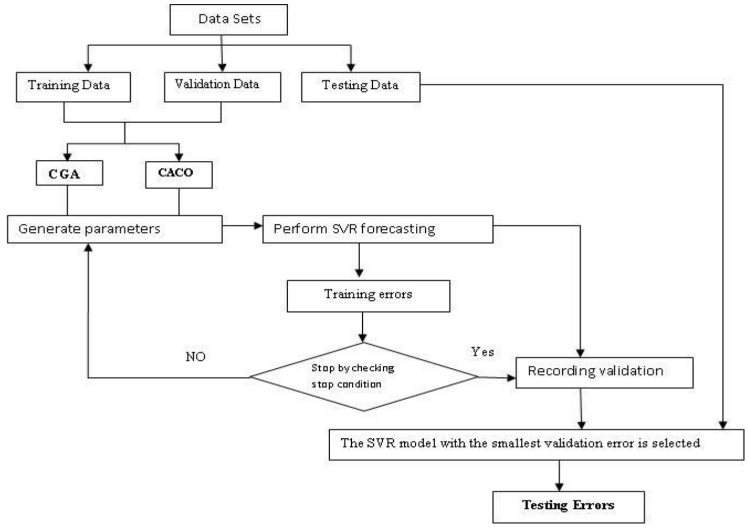

Figure 3. The proposed SVR-CGA and CACO-SVR hybrid system which predict the pipe failure rates

In this research, the selection rate is considered at 0.5; thus, the chromosomes must be firstly sorted by their fitness functions and half of the best population is selected for the next generation. The single point Crossover rate has been considered and mutation rate selected between 0.1 and 0.2 based on its kernel function type. This research has been developed by MATLAB (version 7.12(R 2011a)) and SVM Toolbox and parameters were localized by Continues GA to solve these problems. Equation. (17) is used for normalizing the Input values to the models.

$$x_n=0.8\frac{(x-x_{min})}{(x_{max}-x_{min})}+0.1 \qquad (17)$$

x is the original value ,x_{min} is the minimum value and x_{max} is the maximum value between input values, and x_n shows normalized values. So that, input results are between [0.1, 0.9].

Also, In this paper, the root of mean squared error (RMSE), normal root of mean squared error (NRMSE) and coefficient of determination(R^2) are used as assessment criteria of the reliability of the model.

$$R^2=\frac{\left(\sum_{i=1}^{n}(y_{actual}-\bar{y}_{actual})(y_{pred}-\bar{y}_{pred})\right)^2}{\sum_{i=1}^{n}(y_{actual}-\bar{y}_{actual})^2\sum_{i=1}^{n}(y_{pred}-\bar{y}_{pred})^2} \qquad (18)$$

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{actual_i}-y_{pred_i}\right)^2} \qquad (19)$$

$$NRMSE=\frac{RMSE}{var(y_{actual})} \qquad (20)$$

Where y_{actual} is the observed data, $y_{prediction}$ is the predicted data, $y_{average}$ is the average of data and n is the number of observations. Also, var(y_{actual}) is the variance of actual data.

4. Case Study

In this study, in addition to the quality of data, accuracy of them has been considered. To evaluate the proposed hybrid algorithm a part of a water distribution network of a city in Iran is considered as the study area this city is one of the most travelled cities Fig.4. The area of this district is 2,418 hectares and covers 93,719 properties with 579,860 m of distribution pipes including steel pipes 800, 700 and 600 mm in diameter, and asbestos cement and cast iron pipes 400, 300, 250, 200, 150, 100 and 80 mm in diameter. The installation and execution of the network pipelines in this area generally started in 1981. According to statistical records, this region has the greatest failure rate especially on asbestos cement. In this study, due to incomplete data on steel and cast iron pipes, asbestos cement pipes are only used in the modeling process.

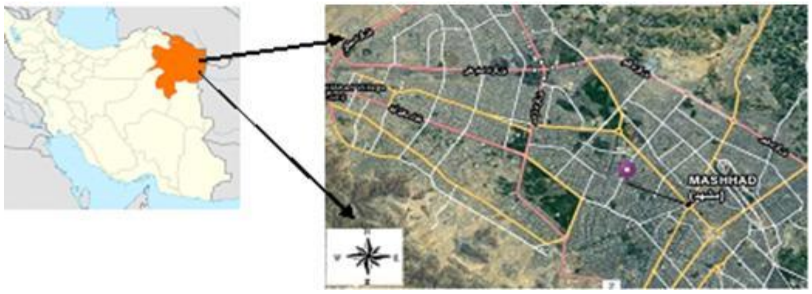

Figure 4. Schematic of study area and pressure measurement points

In order to failure rate modeling of the asbestos cement pipes, the daily events recorded in the 2005 to 2006 years over 2438 record data such as diameter, year of implementation ,installation depth, total accident happens and the average of hydraulic pressure. These data have been collected from local water and water waste company.

RESULTS AND DISCUSSION

Input parameters to the combined SVR-CACO and SVR-CGA models included installation depth, pressure, age, length, and diameter and the output parameter model base on predicted output values from the training input values. Calculated variables to appropriate the best values of the kernel functions based on the best type of loss is as follows:

1- According to ε-insensitive function that consists of the effective parameters of kernel function (ε, C).
2- According to quadratic loss function that consists of the effective parameters of kernel function (C).

The results obtained by mentioned algorithms and appropriate kernel and loss functions have been shown in Table 2.

Table 2. Comparing results between CACO and CGA

	Kernel Type	Algorithm Type	Time(s)	R^2	RMSE	NRMSE	C	P1	ε
ε-insensitive	RBF	CACO- SVR	7970.8	0.9998929	0.00723	0.01829	124.56	3.125	0.00348
		SVM-CGA	10416	0.9998341	0.00795	0.02202	34.0906	1.498	0.000114
	ERBF	CACO-SVR	7896.5	0.9908480	0.07701	0.20325	115.79	0.712	0.03156
		SVM-CGA	8798	0.9739223	0.01407	0.17277	176.021	1.1031	0.000726
	Polynomial	CACO-SVR	7919.4	0.4876852	0.44232	1.24770	54.754	6.315	0.02346
		SVM-CGA	8516	0.4954612	0.73222	1.90157	78.7583	9.278	0.842413
	Spline	CACO-SVR	7311.7	0.6629965	0.36977	1.01165	43.57	-	0.12680
		SVM-CGA	9880	0.6867411	0.33909	0.90198	4.2508	-	0.142851
quadratic	RBF	CACO-SVR	317.23	0.9999767	0.00286	0.00812	72.739	5.415	--------
		SVM-CGA	549.4	0.9999841	0.00275	0.00691	344.49	3.496	---------
	ERBF	CACO-SVR	259.82	0.9859318	0.07599	0.20145	344.58	3.83	--------
		SVM-CGA	314.9	0.9989343	0.99893	0.04175	0.08333	1.771	--------
	Polynomial	CACO-SVR	711.35	0.9623311	0.11415	0.30398	1456.89	5.62	--------
		SVM-CGA	992.35	0.9068322	0.90683	0.19916	0.55049	8.546	--------
	Spline	CACO-SVR	206.03	0.7856149	0.31790	0.81007	356.8	-	--------
		SVM-CGA	391.9	0.7841079	0.78410	0.27634	0.74645		--------
	linear	CACO-SVR	198.56	0.7877169	0.27653	0.73376	224.57	-	--------
		SVM-CGA	5182.4	0.7330571	0.73305	0.93543	0.34324	−	--------

According to the above results, quadratic loss function shows better results in time and other parameters than ε-insensitive in Table 2.

Comparing results between Hybrid SVR-CACO and SVR-CGA models has been shown in Fig.5, time considered as a comparison parameter. By over looking at Table 2 and Fig.5 can be found CACO searching algorithm has a better performance in time-consuming than the CGA.

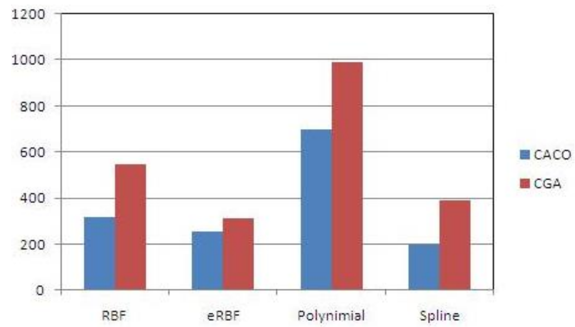

Figure 5. comparing between Hybrid SVR-ACO and SVR-CGA models time

Results in Table 2 shows the correlation and estimation of error in the RBF kernel has an excellent performance relative to the other kernel functions also by comparing the results in Table 2, shows quadratic loss function present a better result than ε-insensitive loss function.

Results from ε-insensitive loss function and related kernel functions present in Fig. [6-8].

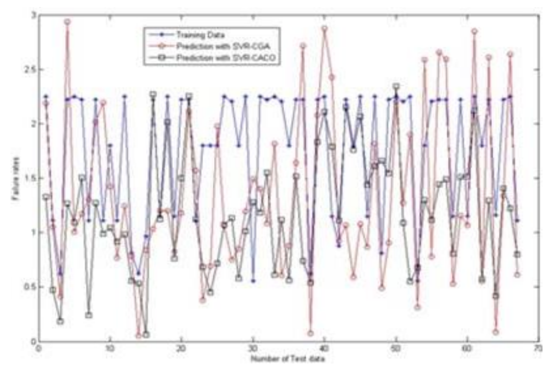

Figure 6. Results by Poly kernel function and loss function

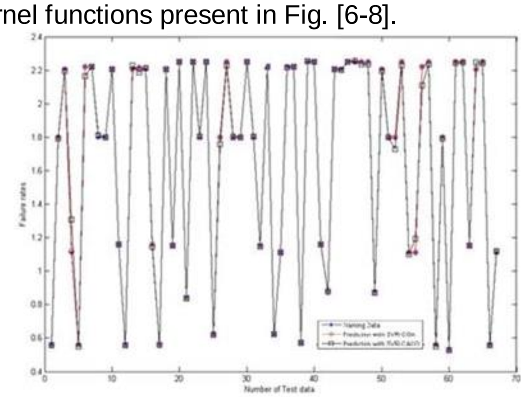

Figure 7. Results by RBF kernel function and loss function

According to the Figure results can be found CACO searching algorithm as same as CGA in accuracy but it has a better performance in time-consuming than the CGA. Fig. 7 shows the best prediction result than the other kernelin upper figures.

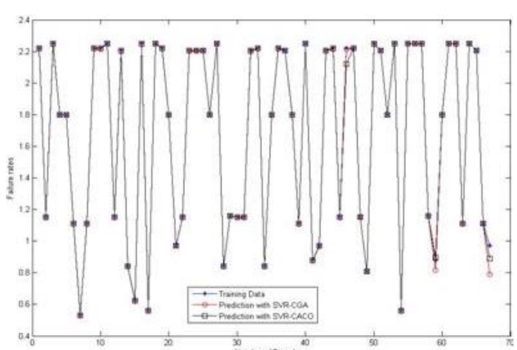

Figure 8. Results by eRBF kernel function and Quadratic loss function

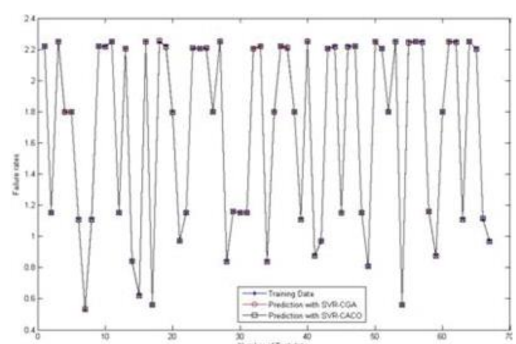

Figure 9. Results by eRBF kernel function and Quadratic loss function

Results from quadratic loss function and related kernel functions present in Fig.9 and Fig.10.

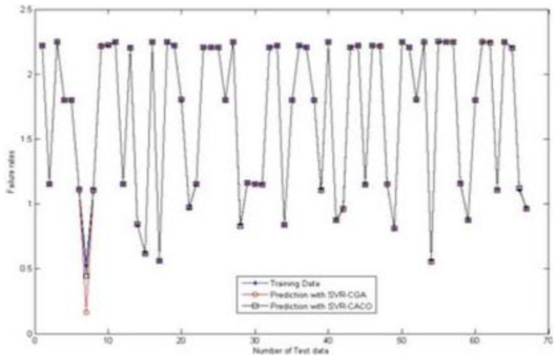

Figure 10. Results by Polynomial kernel function and Quadratic loss function

Upper Figures show the quadratic loss function performances, According to these figures our results have a better prediction and show better results than ε-insensitive loss function.

CONCULSION

In the recent decades, public health and health care has been more attention by responsible in each region. One critical infrastructure to achieve the promote public health and better quality in water according to WHO standard is the water distribution systems ready to operate and timeless through correct management and use optimized.

So in this research compared hybrid SVR-CGA and SVR-CACO models to predict pipe failure rates in water distribution networks to reduce the number of events.

The compared SVR models in order to make the relationship between the failure rate parameters in pipes with the number of events and failure of pipes considered as a main component of urban infrastructure, water supply and hygiene and health. Also by using the CGA and CACO optimal kernel function and SVR related parameters has been found.

By comparing these results with the other methods such as ANFIS, ANN, and ANN-GA that had been done in the past, results show the SVR-CACO model has better performance than the other models especially in time elapse. Also, with combining SVR and CGA obtained better parameters proportional to the data type and the model shows better performance in accuracy.

By comparing among the kernel functions RBF and eRBF offered favorable results and in the loss functions regard to quadratic time and accuracy quadratic loss function has been shown more favorable results than ε-insensitive loss function.

REFERENCES

Angelova M, Atanassov K. 2012. Purposeful model parameters genesis in simple genetic ……algorithms, Compt. & Math. App. 64:221–228.

Carrizosa E,Romero Morales D. 2013. Supervised classification and mathematical optimization,Computers & OperationsResearch. 40:150–165.

Elahi Panah NA. 1998. Subsequent water distribution networks in the country until 1400,Journal Water and Environment, 284-15.

Giustolisi O. 2006. Using a multi-objective genetic algorithm for SVM construction, Journal of Hydroinf. 8:125-139.

Gonzalez-Abril L, Velasco F, Ortega JA, Franco L. 2011 .Support vector machines for classification of input vectors with different metrics, Comp. & Math. Appl. 61:2874-2878.

Gunn S. 1998. Support vector machines for classification and regression. Technical report. School of . Electronics and Computer Science, University of Southampton.

Hofmann T, Scholkopf B, Smola A. 2008. Kernel methods in machine learning, The annals of statistics. 36:1171-1220.

Hong WC, Dong Y, Zheng F, Lai CY. 2011. Forecasting urban traffic flow by SVR with continuous ACO. J. Applied Mathematical Modeling. 35:1282–1291.

Haupt R, Haupt S. 2004. Practical Genetic Algorithms, Second Edition, John Wiley and Sons, USA.

Mashford J, Silva DD, Marny D, Burn S. 2009. An approach to leak detection in pipe networks using analysis of monitored pressure value by support vector machine, IEEE Comp. Society. 3: 534 – 539.

Mandal SK, Tiwari MK, Chan FTS. 2012. Leak detection of pipeline: An integrated approach of rough.set theory and artificial bee colony trained SVM. Expert Systems with Applications, 39: 3071–3080.

Qi Z, Tian Y, Shi Y. 2013. Robust twin support vector machine for pattern classification, J. Pattern Recognition. 46 :305-316.

Soltani, J, Mohammad Rezapour Tabari M. 2012. Determination of Effective Parameters in Pipe Failure Rate in Water Distribution System Using the Combination of Artificial Neural Networks and Genetic Algorithm, J. Water and Environment (In Persian). 83:2-15.

Smola AJ, Scholkopf B. 1998. A Tutorial on Support Vector Regression, NeuroCOLT, Royal Holloway College, University of London,.

Shamir U, Howard CDD. 1979. An analytical approach to scheduling pipe replacement. J.AWWA. 71, 24–258.

Tabesh M, Soltani J, Farmani R, Savic DA. 2009. Assessing Pipe failure Rate and Mechanical Reliability of water Distribution Networks Using Data Driven Modeling, Journal of Hydroinf (IWA)11:1-17.

Vapnik VN. 1992. Principles of risk minimization for learning theory. Advances in Neural Information Processing Sys. 4: 831-838.

Vapnik VN, Chapelle O. 2000 . Bounds on error expectation for support vector machines. J. Neural .Computation. 12: 2013-2036.

Xu Q, Chen Q, Ma J, Blanckaert K. 2013. Optimal pipe replacement strategy based on break rate prediction through genetic programming for water distribution network , J. Hydro-environment Research. In press.