Caspian Journal of Environmental Sciences
Online ISSN: 1735-3866
Print ISSN: 1735-3033

Comparison of temporal and spatial patterns of water quality parameters in Anzali Wetland (southwest of the Caspian Sea) using Support vector machine model

Maryam Fallah[1*], Ahmad Reza Pirali Zefrehei [2], Seyyed Aliakbar Hedayati[2], Tahereh Bagheri[3]

1- Department of Natural Resources, Isfahan University of Technology, Isfahan, Iran
2- Department of Aquatics Production and Exploitation, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran
3- Agricultural Research Education and Extension Organization, Iranian Fisheries Science Research Institute, Offshore Research Center, Tehran, Iran

** Corresponding author's E-mail: maryam.fallah85@gmail.com*

ABSTRACT

Urgent is growing to have reliable information from the country's water resources. In recent years, data mining models such as artificial neural network (ANN), gene expression programming, Bayesian network, machine algorithms, such as a support vector machine (SVM), and Random Forest have found widespread use in the field of simulation and prediction of components in aquatic ecosystems. Variables vary greatly on water quality parameters (due to nonlinear and complex relationships). Therefore, conventional methods are not eligible to solve water resource quality management problems. The aim of this study was to investigate the possibility of simulating the spatial and temporal alterations in water quality parameters during the period 1985-2014 in Anzali Wetland using a SVM model. Based on principal components analysis (PCA), the parameters EC, TDS, pH and BOD_5 were selected for analysis in this study. Spearman correlation was calculated to determine the inputs of the model and the correlation coefficient (CC) between the water quality parameters. According to the results of the correlation table analysis, 8 types of structures including different inputs were used to predict the parameters with machine vector. In the next stage, 70% of the data were used to train, while the rest were used for analyzing the models. Criteria for determination coefficient (R^2) and root mean square error (RMSE) were used for evaluation and model performance. The results revealed that in verification stage among different used models, the pH had the highest accuracy (0.95), while the lowest RMSE (0.20). Trend of alterations for optimal model of each parameter on a time scale, indicated an adequate estimation at most points. In general, the results exhibited the appropriate accuracy and acceptable performance of the SVM model in simulating water parameters.

Keywords: Machine algorithms, PCA, RMSE, simulation, wetland.

INTRODUCTION

In recent years, data mining models such as artificial neural network (ANN), gene expression planning, Bayesian network, machine algorithms such as Support vector machine (SVM) and Random Forest have been widely used in the field of simulation and prediction of components in aquatic ecosystems (Samsudin *et al.* 2011). These data-driven models have a definite mathematical structure and can detect nonlinear complex relationships between input and output data without the physical process governing the phenomenon (Nayak *et al.* 2004). The SVM is based on statistical learning theory, which dates back to 1960, and a nonparametric statistical method is monitored (Pao 1989). Classical learning methods, including artificial neural networks, which are the most commonly used, designed to minimize training data set error (empirical error minimization). Contrary to these methods, SVM are based on structural error minimization. In other words, the structure of system, unlike the neural networks, has

DOI:
Received: July 07. 2020 Accepted: Nov. 28. 2020
Article type: Research

not been clear from the outset, in addition to minimizing the experimental error, the structural error is minimized and the optimal system structure is determined during the training process (Bazargan-Lari *et al.* 2010). Despite the speed of computing and the accuracy of SVM and ANN, support vector machines have higher stability and ease of implementation (Xin *et al.* 2010). These machines have more power for non-teaching data (test data) compared to the other learning methods. In classical models such as artificial neural networks, the network structure is clearly defined before training and is not practically optimized. However, in SVM models, network structure is also optimized along with weights (Bazargan-Lari *et al.* 2010). Several studies published in the field of using this model: Liu *et al*. (2013) examined wavelet combinations and support regression vectors to predict groundwater quality. The results showed that with the combination of wavelet and SVM, the effect of inputs on the outputs of the model has increased and it makes accuracy of predictions more reliable (Liu et al., 2013). Saedi *et al.* (2016) reviewed the salinity changes of the Caspian Sea using SVM, exhibiting the accuracy of this model (Saedi *et al.* 2016). Solgi *et al.* (2017) studied the biochemical oxygen demand (BOD_5) in Karun River using SVM based on principal components analysis (PCA) reporting that applying wavelet transforms to input data improves the results. Knowledge of water quality and its changes are of particular importance in future planning and proper management of water resources. Since wetlands are considered to be the most important ecological resources, providing information about their ecological changes helps to find the reasons of changing and adopting appropriate management policies (Rafii *et al.* 2011). Anzali Wetland has undergone so many changes during the last few decades, providing suitable conditions for expansion of aquatic plants. Sedimentation and rising wetland bed accelerated through bed loading extension of floating plants such as *Azolla* sp. and other aquatic plants. In addition, by increasing sediment load in Anzali wetland during the last half century, the depth of the wetland has decreased from 11 m to 2.5 m (Azari 2009). Variables vary greatly on water quality parameters (due to nonlinear and complex relationships). Thereby, conventional methods are not good enough to solve the problem of water resource quality management (Xiang *et al.* 2006). The aim of this study was to investigate the possibility of simulating the spatial and temporal changes in the water quality parameters including electrical conductivity (EC), total dissolved solids (TDS), pH and BOD_5 during the period of 1985-2014 in Anzali Wetland using SVM model.

MATERIALS AND METHODS

Study area

Anzali wetland situated in north of Iran in Guilan Province in Bandar-e Anzali County, in 48º45' to 49º42' E and 36º55' to 37º32' N respectively. (Fig. 1). The wetland is divided into four areas of Abkenar (West), East, Central and Siah keshim, which are distinguished for some characteristics. The Anzali wetland complex with an area of 15,000 ha and height of 24 m below the global sea level is located in the southwest coast of the Caspian Sea, its basin is a small part of the southern coastline of the sea. The wetland complex extends northwards to the sea, from the east to Khomam and Pirbazar, from the west to Kapurchal and Abkenar, and southwards to the Sowmeh Sara and part of Rasht (JICA 2005).

Used data

Anzali wetland quality data used in this study belonged to a period of 29 years from 1985 to 2014. Choosing the proper inputs to the model is one of the most important issues in the statistical models. Hence, PCA was employed. Principal component analysis is one of the multivariable statistical methods that is suitable for reducing the complexity of variables and the proper interpretation of information when we have a large amount of information. Further details on PCA are provided in other sources (Tabachnick & Fidell 2001). The statistical status of the quality data of the Anzali Wetland water is shown in Table 1.

Table 1. The statistical properties of the parameters used in the statistical period.

	pH	TDS (Mg L^{-1})	EC (μs cm^{-1})	BOD_5 (Mg L^{-1})
Max	9.23	6221	9720	150.08
Min	6.80	125	151	0.80
Mean ±SE	7.97 ± 0.45	919.64 ± 115.1	1467.4 ± 189.2	22.01 ± 2.7

Support Vector machine model

Support vector machine is a tool based on statistical learning theory, and learning is performed based on a set of educational data (input and output pairs). In this method, searching for discovering objective function (estimation or decision functions) in the modeling process is in conjunction with maximizing the utility and generalizability

Caspian J. Environ. Sci. Vol. 19 No. 1 pp. 95~104
DOI:
Received: July 07. 2020 Accepted: Nov. 28. 2020
Article type: Research

of the model, and in that way, its complexity is minimized. This is done using input and output training data. The procedure is consisted of giving a set of M input and output data pair, (X_i, Y_i), ..., M and i = 1, into the system and the system attempts to assign the function f (x) from input to output to a kind of learning that can accurately estimate y outputs for unobserved X values =Vapnik, 1998). Fig. 2 shows the schema of this model. In the SVM regression model, the function associated with the dependent variable Y is a function of several independent variables x. It is assumed to be similar to other regression issues, including determination of the relationship between independent and dependent variables with an algebraic function such as f (x), plus a disturbance (allowed error, ε) (Vapnik 1995). Equation 1:

$$f(X) = W^T.\emptyset(x) + b$$

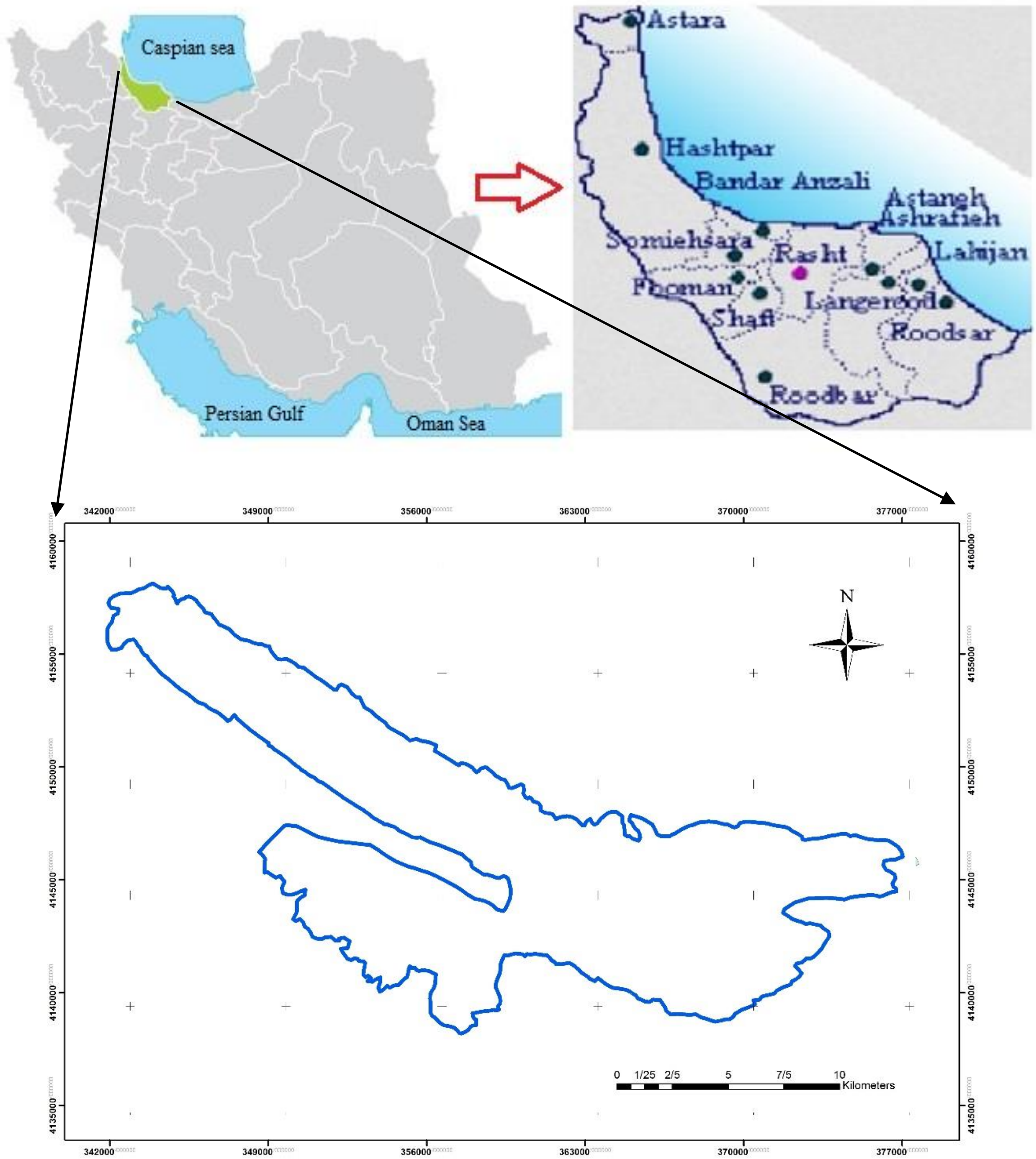


Fig. 1. Geographical location of Anzali Wetland.

If W is the weight vector and b is the constant of characteristic of the regression function (bias term) and Ø is also a kernel function, then the purpose is to find a functional form for f (x). This is accomplished by training the SVM model by a set of examples (training set) (Misra *et al.* 2009). To calculate b and w, the risk function (equation 2)

Caspian J. Environ. Sci. Vol. 19 No. 1 pp. 95~104
DOI:

Received: July 07. 2020 Accepted: Nov. 28. 2020
Article type: Research

in the ε-SVM model should be optimized by considering the conditions (constraints) in equation 3 (Yoon *et al.* 2011).

Equations 2 and 3:

$$W^T.\emptyset(X_i)+\mathrm{b}-y_i \le \varepsilon+\varepsilon_i^*,$$

$$\frac{1}{2}W^T.W+C\sum_{i=1}^{N}\varepsilon_i+C\sum_{i=1}^{N}\varepsilon_i^*$$

$$y_i-W^T..\emptyset(X_i)-\mathrm{b}\le \varepsilon+\varepsilon_i,\varepsilon_i,\varepsilon_i^* \ge 0$$

where C (cost factor) is a regularization constant and positive integer, which determines the fineness when the model training error occurs. Ø, Kernel function, N, number of samples (model training patterns), and two characteristics ε_i^* and ε_i are slack variables. Finally, the regression SVM function can be rewritten in the following form, equation 4: (Yoon *et al.* 2011).

$$f(x)=\sum_{i=1}^{N}\overline{\alpha_i}\ .\emptyset(X_i)^T.\emptyset(x)+b$$

where α_i is the mean of Lagrange multipliers. The calculation of Ø (x) in its characteristic space may be very complicated (Yoon *et al.* 2011). To solve this problem, the common trend in the SVM regression model is to select a kernel function, the general form of the kernel function is presented in equation 5: (Hsu *et al.* 2003).

$$K(X,X_k)=\emptyset({X_i}^T)\sqrt{b^2-4ac}$$

Various kernel functions can be used to make different types of ε-SVM. The types of kernel functions that can be used in the regression SVM model are: a polynomial kernel with three target attributes and kernel of radial base functions (RBF) with a target characteristic, in which the kernel function of RBF was used in study and calculated according to equation 6: (Hsu *et al.* 2003).

$$K(X,X_k)=\exp\left(\frac{-1}{\sigma^2(X-X_k)^2}\right)$$

To solve the problem, the values of the parameters C and ε, γ as well as parameters related to the selected kernel function should be specified. To optimize the three parameters, the Grid Search method is used (Hsu *et al.* 2003). Noteworthy, the R $_{3.4}$ software was used to utilize the SVM model.

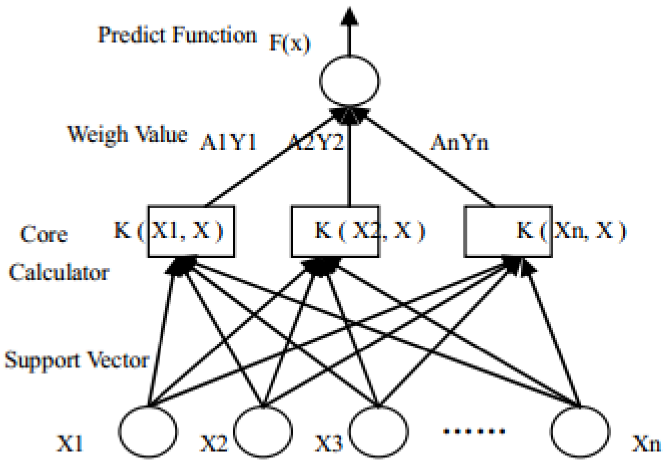


Fig. 2. Schematic model of the support vector machine model.

Model evaluation criteria

To evaluate the model used in this research, the determination coefficient (R^2), the high values of this coefficient, indicate better performance of the model and also the Root Mean Square Error (RMSE) was used.

$$R2=\left(\frac{(n\sum(y)-(\sum x)(\sum y)}{\sqrt{n(\sum x^2)-(\sum x)^2\sqrt{n(\sum y^2)-(\sum y)^2}}}\right)^2$$

where x is the estimated value, y is the measured value and n is the number of samples.

Caspian J. Environ. Sci. Vol. 19 No. 1 pp. 95~104
DOI:
Received: July 07. 2020 Accepted: Nov. 28. 2020
Article type: Research

$$RMSE = \sqrt{\frac{\sum^{N}{}_{i=1}(x_{est,i} - x_{meas,i})^2}{N}}$$

where x_{est}, is the estimated value, x_{meas}, is the measured value, and N, is the number of samples. In addition to the above criteria, the distribution and timing diagrams of observational values-computational time-versus-time comparison and analysis are used.

RESULTS AND DISCUSSION

The existence of a large number of input variables is a major problem in the development of SVM models, which may prevent SVM from finding the optimal model. Increased number of input variables due to the elevated number of weights relative to the number of inputs, complicates the structure of the model and its instability. In some cases, the variables are also highly correlated, and all of them will have a kind of duplicate information entry. These issues may prevent finding optimal models. In order to reduce inputs, although some information may be lost (Noori *et al.* 2009), we also used the proposed PCA method (Noori *et al.* 2011). Based on PCA, the EC, TDS, pH and BOD_5 were selected for analysis in the structure of this study. In order to determine the inputs of the model and the correlation among water quality parameters, we calculated spearman correlation among the water quality parameters using SPSS, 22 software in Table 2.

Table 2. Correlation coefficient (CC) between water quality parameters.

	pH	EC	TDS	BOD_5
pH	1			
EC	0.146	1		
TDS	0.158	0.881**	1	
BOD_5	-0.009	0.282**	0.295**	1

**Significant at the level of 0.001.

According to the results of the analysis of correlation coefficient (CC) in Table 2, 8 types of structures including different inputs were used to predict the water quality parameters by SVM. The characteristics of these structures are presented in Table 3.

Table 3. Selected compositions input parameter models, SVM.

Structure	Input	Output
1	pH, TDS, BOD_5	EC
2	pH, TDS	EC
3	EC, TDS, BOD_5	pH
4	EC, TDS	pH
5	pH, EC, BOD_5	TDS
6	pH, EC	TDS
7	pH, EC, TDS	BOD_5
8	EC, TDS	BOD_5

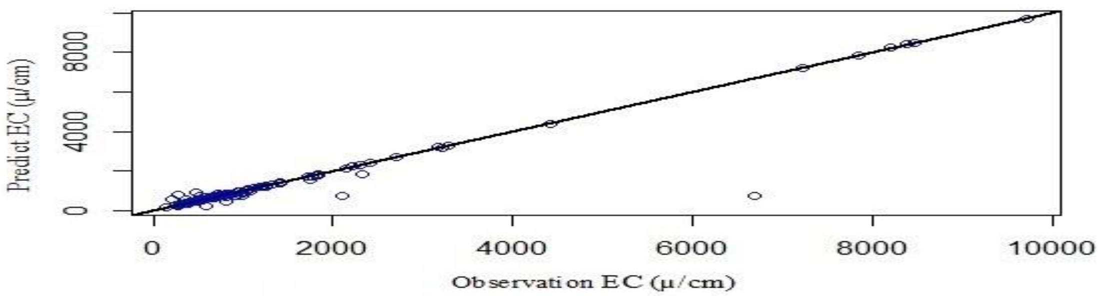
In the next stage, 70% of the data were employed to train, while the rest were used for testing and model analyzing. To determine the optimal combination of input variables in the SVM, the optimal Grid Search function of the network, the most suitable characteristics of the parameters C and ε, gamma were selected for each model. In Table 4, the model results are presented based on the optimal function for each structure.

Table 4. The results of SVM model training and validation in Anzali wetland.

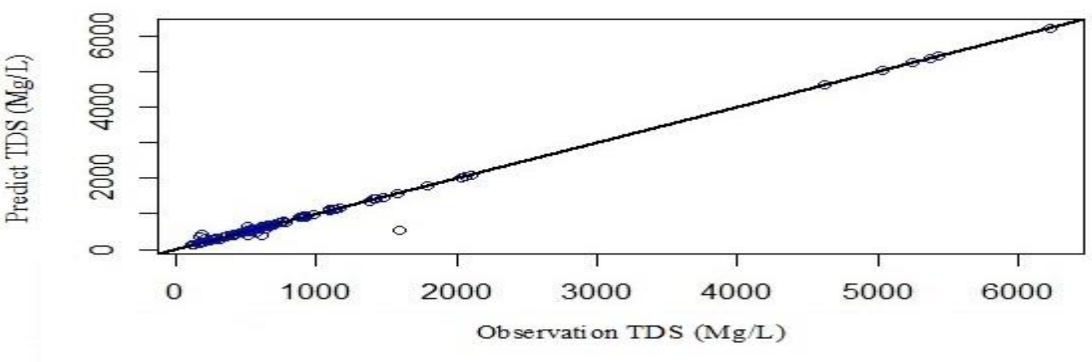
Pattern No.	Training		Validation		Model variables		
	RMSE	R^2	RMSE	R^2	C	ε	δ
1	2424.053	0.81	605.86	0.89	20	0.01	1
2	2625.35	0.73	700.68	0.87	1	0.1	0.5
3	7.991	0.85	0.210	0.95	15	0.01	10
4	8.53	0.88	0.426	0.9	1	0.1	0.5
5	1491.43	0.76	113.18	0.93	10	0.01	5
6	1250.25	0.89	270.11	0.91	1	0.1	0.5
7	35.305	0.75	5.65	0.88	30	0.2	25
8	45.33	0.86	28.24	0.94	1	0.1	0.5

Caspian J. Environ. Sci. Vol. 19 No. 1 pp. 95~104
DOI:
Received: July 07. 2020 Accepted: Nov. 28. 2020
Article type: Research

The statistical analysis of RMSE and CC is shown in Table 4 (0.210 and 0.95, respectively). The inputs of structure 3 are more appropriate than the structure 4 of pH models, as well as other compounds that are modeled with other structures at the verification stage, hence definitely are more accurate. Albeit, based on the statistical indicators, this argument can be expressed for this structure. In the same way, for TDS, the structure 5 is the same as 6, best in comparison with other models at the verification stage (respectively, 113.18 and 0.93 respectively). In comparison with the RMSE and CC indices, the EC model is 605.86, 700.68, and 0.89, 0.87 respectively for structures 1 and 2, respectively. Compared to the BOD_5 model, structure 7 exhibits a better accuracy (55.65 and 0.88, respectively). In Fig. 3, the graph of the best-performing models (optimal) is shown for each parameter for the validation data. According to Fig.3, most of the values estimated by the model and the observed values of TDS, EC and BOD_5 are compared with pH, except for a few points on the bisector line, which implies the equality of observation values and its line prediction. In Fig. 4, the process of variation for the optimal model of each parameter was given on a time scale, which indicates suitable estimation in most of the points.
An examination of the optimal charts of each studied parameters (Figs. 3 and 4) revealed that the SVM model exhibits a more acceptable performance in predicting the EC, TDS, pH and BOD_5 values. By examining the input parameters, it is possible to obtain data on the values of statistical indices. The number of inputs can also affect the performance of the models, as shown once comparing the outputs of parameters in this study. According to Table 4, the model of pH has the highest accuracy while RMSE in the verification stage. Due to the nature of each phenomenon, the reason can be searched, and although the data-driven models are often non-linear and complex, however, the impact of environmental processes on this issue should also be considered. Several studies have reported the effectiveness of SVM (Liu *et al.* 2013; Saedi *et al.* 2016; Solgi *et al.* 2017), in accordance with the findings of the present study which also exhibited the good performance of RBF. Kakaei Lafadani *et al.* (2013) used ANN, SVM and also Mike's conceptual hydrologic models to simulate the daily flow of Eskandary basin reporting that the SVM model with the kernel of the radial base function exhibited the lowest error in the daily flow forecast.



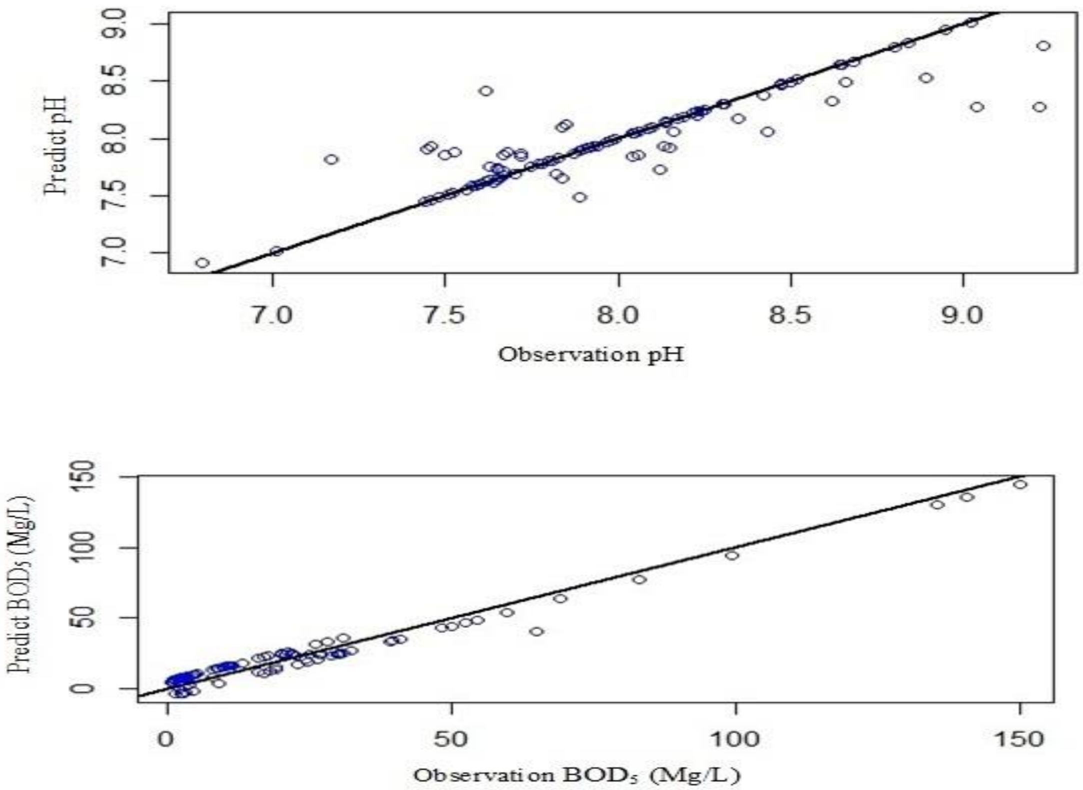
a



b

Caspian J. Environ. Sci. Vol. 19 No. 1 pp. 95~104
DOI:

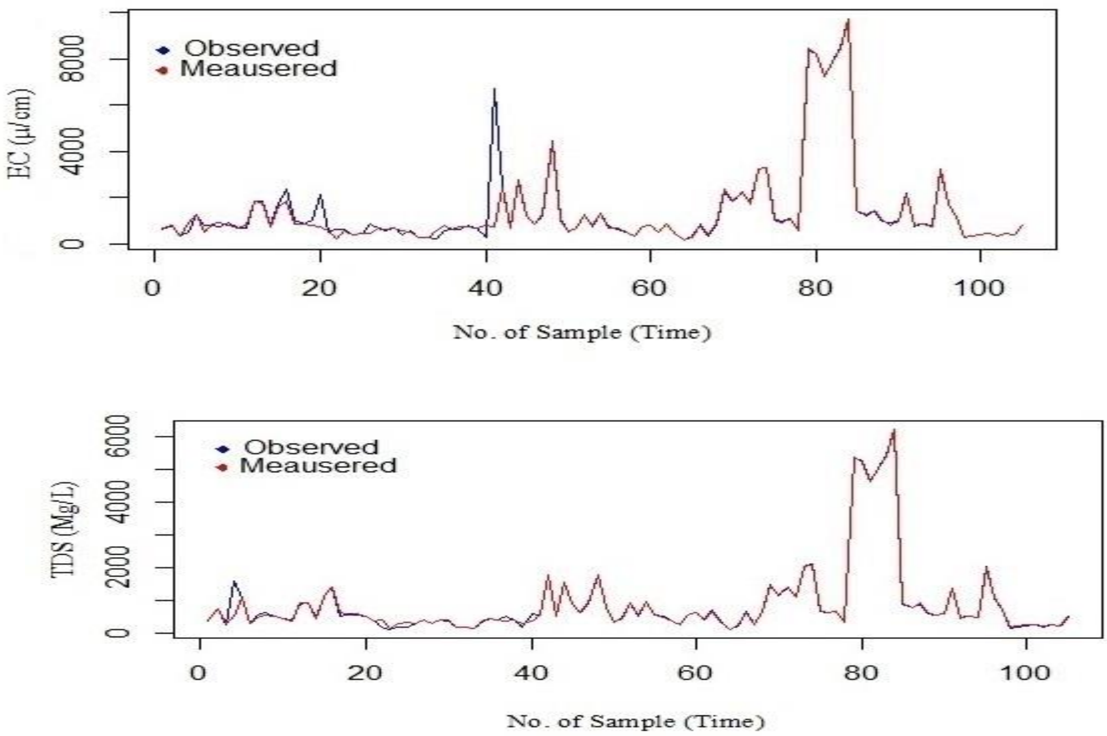
Received: July 07. 2020 Accepted: Nov. 28. 2020
Article type: Research



c

d

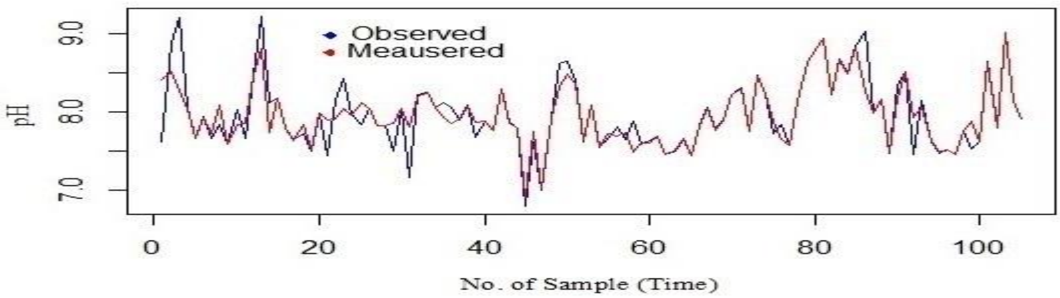
Fig. 3. Scattering diagram of the observational-computational values of SVM in the validation phase; (a) EC; (b) TDS; (c) pH; (d) BOD_5.



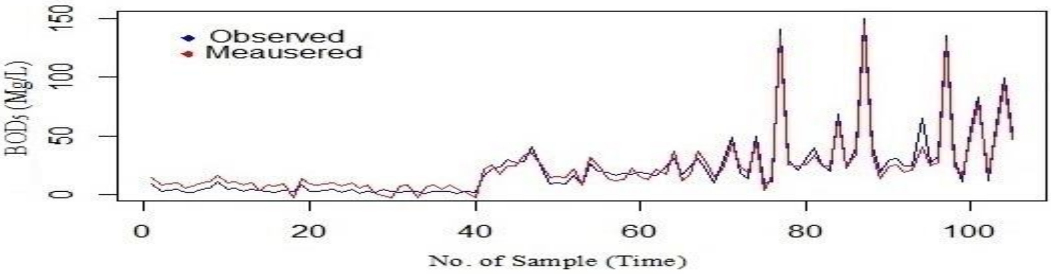
a

b

Caspian J. Environ. Sci. Vol. 19 No. 1 pp. 95~104
DOI:
Received: July 07. 2020 Accepted: Nov. 28. 2020
Article type: Research



c



d

Fig. 4. Time-series graphs of observational-computational values relative to the time taken from the SVM in the validation phase; (a) EC; (b) TDS; (c) pH; (d) BOD_5.

In addition, the prediction of the daily flow in Ghareh Souz basin showed the efficiency of the SVM model (Mohammadpour *et al.* 2012). Moreover, Huang *et al.* (2014) predicted the monthly flow of China's Huaxi River using SVM, reporting that the model carries a high degree of accuracy in predicting the monthly flow of the river. Sedighi *et al.* (2016) also predicted the runoff precipitation process using ANN and SVM in the Roodak basin of Tehran, Iran using MODIS sensor, concluding that SVM has an acceptable ability to estimate runoff.

CONCLUSION

The EC, TDS, pH and BOD_5 parameters are among the most important factors in environmental and fishery studies. Therefore, the simulation of accurate estimation is of great importance. This study also proved the efficiency of the supporting vector machine model. Hence, it is recommended that due to its high accuracy and speed, compared to the other intelligent learning methods, this model is suitable for estimation and evaluation in other aquatic ecosystems.

REFERENCES

Azari, F 2009, Determine the area of the protection (shield) of Anzali wetland, The Iranian Department of Environment, Tehran.105 p. [In Persian].

Bazargan-Lari, MR, Kerachian, R, Sedghi, H, Fallahnia M, Abed-Elmdoust A, Nikoo, MR, 2010, Developing Probabilistic Operating Rules for Real-time Conjunctive Use of Surface and Groundwater Resources: Application of Support Vector Machines, *Journal of Water and Wastewater*, 21:54-69. [In Persian].

Hsu, CW, Chang, CC & Lin, CJ 2003, A practical guide to support vector classification, www.csie.ntu.edu.tw, 4 May 2009.

Huang, S, Chang, J, Huang, Q & Chen, Y 2014, Monthly stream flow prediction using modified EMD-based support vector machine. *Journal of Hydrology*, 511: 4. 764-775.

Japan International Cooperation Agency (JICA) 2005, Integrated management for Anzali wetland. Department of Environment of Iran. Tehran, 182 p. [In Persian].

Kakaei Lafadani, E, Moghaddam Nia, A, Ahmadi, A, Jajarmizadeh, M & Ghafari, M 2013, Stream flow simulation using SVM, ANFIS and NAM models (A case study). *Caspian journal of applied sciences research* 2: 4. 86-93.

Liu, W, Wang, GY, Fu, JY, & Zou, X 2013, Water Quality Prediction Based on Improved Wavelet Transformation and Support Vector Machine. *Advanced Materials Research*, 726: 3547-3553.

Caspian J. Environ. Sci. Vol. 19 No. 1 pp. 95~104
DOI:
Received: July 07. 2020 Accepted: Nov. 28. 2020
Article type: Research

Misra, D, Oommen, T, Agarwal, A, Mishra, SK & Thompson, AM 2009, Application and analysis of support vector machine based simulation for runoff and sediment yield. *Biosystems Engineering*, 103: 3. 527-535.

Mohammadpour, M, Mehrabi, A & Katouzi, M 2012, Daily discharge forecasting using support vector machine. *Journal of Information and Electronics Engineering* 2: 769-772.

Nayak, PC, Sudheer, KP, Rangan, DM & Ramasastri, KS 2004, A neuron-fuzzy computing technique for modeling hydrological time series. *Journal of Hydrology,* 291: 52-66.

Noori R, Morid S, Riahi, H 2009, Effect of Input Variables Preprocessing in Artificial Neural Network on Monthly Flow Prediction by PCA and Wavelet Transformation, *Journal of Water and Wastewater*, 20:13-22.

Noori R. Khakpoor a. Dehghani M. Farohknia A. 2011.Monthly Stream Flow Prediction Using Support Vector Machine Based on Principal Component Analysis, *Journal of Water and Wastewater*, 22: 118-123.

Pao, Y.1989."Adaptive pattern recognition and neural networks,"

Rafii, Y, Malekmohammadi, B, Abkar, A, Yavari, A, Ramezani Mehrian, M, Zohrabi, H 2011, Environmental Change Detection of Wetlands and Protected Areas Using Multi Temporal Images of TM Sensor (Case Study: Neyriz Wetland, Iran), *Journal of environmental studies*, 37:65-76.

Saedi, S, Soleimani, K, Nikmehr, S, Kavian, A 2016, Evaluation of the performance of the support vector machine model in predicting salinity changes aquifer in the Caspian Sea coastal, second national congress in the field of development of agricultural science and natural resources, 11p.

Solgi, A, Pourhaghi, A., Zarei, H, Ansari, H 2017, Modeling and Forecast Biological Oxygen Demand (BOD) using Combination Support Vector Machine with Wavelet Transform, *Journal of Water and Soil*, 31: 86-100.

Samsudin, R, Saad, P & Shabri, A 2011, River flow time series using least squares support vector machines. *Hydrology and Earth System Sciences*, 15: 1835-1852.

Sedighi, F, Vafakhah, M & Javadi, MR 2016, Rainfall–Runoff modeling using Support Vector Machine in snow-affected watershed. *Arabian Journal for Science and Engineering*, 41: 10. 4065-4076.

Tabachnick, BG & Fidell, LS 2001, Using multivariate statistics, 3rd Ed., Allyn and Bacon, Boston, London.

Vapnik, VN 1998, Statistical learning theory. Wiley, New York: 250-320.

Vapnik, VN 1995, The nature of statistical learning theory. Springer, New York: 250-320.

Xin, S, Qing, X, Lei, Y & Ning, L 2010, "A comparative study of eutrophication Evaluation Models Based on SOM Neural Network and SVM", *Journal of Chongqing University (Natural Science Edition)*, 33:119-123.

Xiang, SL, Liu, ZM, Ma, LP 2006, Study of multivariate linear regression analysis model for ground water quality prediction. *Guizhou Science*, 24:60-2.

Yoon, H, Jun, SC, Hyun, Y, Bae, GO & Lee, KK 2011, A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology* 396: 128-138.

Received: July 07. 2020 Accepted: Nov. 28. 2020
Article type: Research

مقایسه الگوهای زمانی و مکانی فراسنجه‌های کیفی آب تالاب انزلی با استفاده از مدل ماشین بردار پشتیبان

مریم فلاح*[۱]، احمدرضا پیرعلی زفره‌ئی[۲]، سیدعلی اکبرهدایتی[۲]، طاهره باقری[۳]

۱- گروه محیط زیست، دانشگاه صنعتی اصفهان، ایران

۲- گروه تولید و بهره‌برداری آبزیان، دانشگاه علوم کشاورزی و منابع طبیعی گرگان، ایران

۳- مرکز تحقیقات شیلاتی آبهای دور (چابهار)، موسسه تحقیقات علوم شیلاتی کشور، ایران

(تاریخ دریافت: ۹۹/۰۴/۱۷ تاریخ پذیرش: ۹۹/۰۹/۰۸)

چکیده

نیاز به اطلاعات مکانی دقیق و بهنگام از منابع آبی کشور همواره رو به رشد است. در سال های اخیر مدل‌های داده کاوی مانند شبکه عصبی مصنوعی، برنامه ریزی بیان ژن، شبکه بیزین، الگوریتم ماشین مانند ماشین بردار پشتیبان وجنگل تصادفی کاربرد گسترده ای در زمینه شبیه سازی و پیش بینی مؤلفه‌ها در بوم‌سازگان آبی پیدا کرده اند. متغیرهای زیادی بر روی فراسنجه‌های کیفیت آب مؤثرند (به دلیل روابط غیرخطی و پیچیده). لذا روش‌های معمول به خوبی قادر به حل مساله مدیریت کیفی منابع آبی نیستند. این مطالعه با هدف بررسی امکان شبیه‌سازی تغییرات مکانی و زمانی فراسنجه‌های کیفی آب (هدایت الکتریکی، جامدات محلول کل، اسیدیته و BOD_5) در بازه زمانی ۱۳۶۴-۱۳۹۳ در تالاب بین المللی انزلی بااستفاده از مدل ماشین بردار پشتیبان است. براساس نتایج تجزیه مولفه های اصلی، فراسنجه های هدایت الکتریکی، جامدات محلول کل، اسیدیته و BOD_5 برای سنجش در ساختارهای این مطالعه انتخاب شدند. به منظور تعیین ورودی های مدل و میزان همبستگی(CC) بین فراسنجه های کیفیت آب به دست آمده، میزان همبستگی اسپیرمن بین فراسنجه‌های کیفی آب محاسبه شد. باتوجه به نتایج سنجش مربوط به جدول همبستگی، ۸ نوع ساختار شامل ورودی‌های متفاوت به منظور پیش بینی فراسنجه‌های کیفی آب، توسط ماشین بردار اجرا شد. در مرحله بعد از ۷۰ ٪ داده ها به منظور آموزش و از بقیه داده ها برای آزمون مدل‌ها استفاده شد. معیارهای ضریب تعیین (R^2) و ریشه میانگین مربعات خطا (RMSE) برای ارزیابی و عملکرد مدل استفاده شد. نتایج نشان داد از بین الگوهای مورد استفاده در مدل های مختلف، مدل اسیدیته، بیشترین دقت (۰/۹۵) و کمترین ریشه مربعات خطا (۰/۲۱۰) در مرحله صحت‌سنجی داشت. بررسی روند تغییرات به ازای مدل بهینه هر کدام از فراسنجه‌ها در مقیاس زمانی، حاکی از تخمین مناسب در بیشار نقاط است. در مجموع، نتایج مطالعه دقت مناسب و عملکرد قابل قبول مدل پشتیبان در شبیه سازی فراسنجه‌های آب را نشان داد.

*مولف مسئول

Bibliographic information of this paper for citing:

Fallah, M, Pirali Zefrehei, AR, Hedayati, SA, Bagheri, T 2021, Comparison of temporal and spatial patterns of water quality parameters in Anzali Wetland (southwest of the Caspian Sea) using Support vector machine model, Caspian Journal of Environmental Sciences, 19: 95-104

Caspian J. Environ. Sci. Vol. 19 No. 1 pp. 95~104
DOI:
Received: July 07. 2020 Accepted: Nov. 28. 2020
Article type: Research