"Environmental Engineering" 10th International Conference
Vilnius Gediminas Technical University
Lithuania, 27–28 April 2017
http://enviro.vgtu.lt

eISSN 2029-7092 / eISBN 978-609-476-044-0

Article ID: enviro.2017.093
DOI: https://doi.org/10.3846/enviro.2017.093

Groundwater Level Prediction Using Support Vektor Machines and Autoregressive (AR) Modelss

Fatih Üneş[1], Mustafa Demirci[2], Yunus Ziya Kaya[3], Eyup Ispir[1], Mustafa Mamak[4]

[1,2]Civil Engineering Department, Iskenderun Technical University, Iskenderun, Turkey
[3,4]Civil Engineering Department, Osmaniye Korkut Ata University, Osmaniye, Turkey
E-mails: [1]fatih.unes@iste.edu.tr, [1]mustafa.demirci@iste.edu.tr (corresponding author),
[3]yunuszkaya@osmaniye.edu.tr, [4]mmamak@osmaniye.edu.tr

Abstract. Water resources managers can benefit from accurate prediction of the availability of groundwater. Ground water is a major source of water in Turkey for irrigation, water supply and industrial uses. The ground water level fluctuations depend on several factors such as rainfall, temperature, pumping etc. In this study, Hatay Amik Plain, Kumlu region was evaluated using Autoregressive (AR) and Support Vektor Machines (SVMs) methods. The monthly groundwater level was used the previous years data belonging to the Kumlu region.

Keywords: groundwater level, prediction, Amik plain, Support Vektor Machines (SVMs).

Conference Topic: Water engineering.

Introduction

Total water consumption with population increase is continuously increasing. In order to meet the need for water, overshoots made from underground water are causing significant falls in groundwater level. Prediction of groundwater level is important for effective planning and sustainable groundwater management. Application of SVMs in developing a reliable groundwater level fluctuation forecasting system to generate trend forecasts is being discussed.

SVMs have been recently introduced relatively new statistical learning technique. Due to its strong theoretical statistical framework, SVM has proved to be much more robust in several fields, especially for noise mixed data, than the local model which utilizes traditional chaotic techniques (Xinying *et al.* 2004). Jin *et al.* (2009) proposed SVM based dynamic prediction of groundwater level. Mohsen *et al.* (2010) made an attempt with SVMs and ANNs for predicting transient groundwater levels in a complex groundwater system under variable pumping and weather conditions. Heesung *et al.* (2011) developed two nonlinear time-series models for predicting the groundwater level (GWL) fluctuations using ANNs and SVMs.

In this study, groundwater level measured in the previous years belonging to the Kumlu region was performed using Autoregressive (AR) and Support Vektor Machines (SVMs) methods. Monthly total rainfall and monthly average temperature data measured at the Antakya Meteorological Station and the static underground water level monthly measurement data of the observation well No. 474 belonging to DSI in Kumlu region between 2000 and 2015 were used. The DSI observation well is located at 36.21981 latitude and 36.29114 longitude, with a depth.

Methodology

Autoregressive (AR) Model:

The autoregressive (AR) model of an order p can be written as AR(p) and is defined as

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} \ldots\ldots\ldots + \phi_{t-p} y_{t-p} + z_t \tag{1}$$

where: z_t is a purely random process; and $E(z_t) = 0$, $Var(z_t) = \sigma_z^2$. The parameters $\phi_1, \ldots, \phi_p$ are called the AR coefficients. The name "autoregressive" comes from the fact that X_t is regressed on the it's past values. In this paper, model AR1 has been applied to groundwater level data by using MATLAB. The Yule-Walker equation was used to estimate AR coefficients.

Support Vector Machines (SVMs)

Support Vector Regression (SVR) is a regression method based on Support Vector Machines (SVM) (Vapnik 1995; Schölkopf, Smola 2002). The idea behind SVM is to find a hyperplane that separates two classes in the transformed feature (input) space with a maximum distance. SVR aims to find the optimal regression hyperplane, that all training samples lie within an ε-margin around it and is also as flat as possible (Schölkopf, Smola 2002). A support vector machine is an dimensional vector that divides data into two optimal categories on Hyperplane. SVM models are closely related to artificial neural networks and using a sigmoid kernel function; has a two-layer, forward-feed artificial neural network (Haykin 1998). An interesting feature of SVM is; by minimizing the average error rate on the data set from the empirical risk minimization principle derived, statistical learning. In the theory of structural risk minimization. The basics of SVM One of the assumptions is that all the samples in the training set are independent and similar (Song *et al.* 2012).

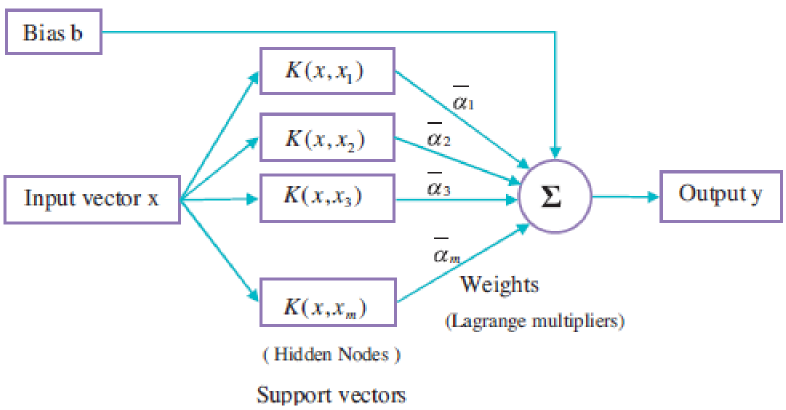


Fig. 1. Network architecture of SVM. Adapted from (Chen and Yu 2007)

Network architecture of SVM, adapted from (Chen and Yu (2007) is shown in Fig. 1. SVM differs from the other classification methods significantly. Its intent is to create an optimal separating hyperplane between two classes to minimize the generalization error and thereby maximize the margin. SVM is an approximate implementation of structural risk minimization approach. Structural risk minimization method described that the error rate of learning machine on test data is bounded by the sum of training error rate and a term that based on Vapnik–Chervonenkis dimension (Haykin 1998).

Results

In this study, the monthly groundwater level data measured in Kumlu, DSI data and the monthly total rainfall and monthly average temperature data measured at the Antakya Meteorological Station were used to determine groundwater level. Modeling was carried out using 192 data of monthly ground water level, monthly total precipitation and monthly average temperature values measured for 16 years between 2000 and 2015.

AR model results

AR1 model was employed for groundwater level to offer a comparison to the bagging-SVM forecast model. The distribution and scatter graphs are shown in Figure 2 and Figure 3.

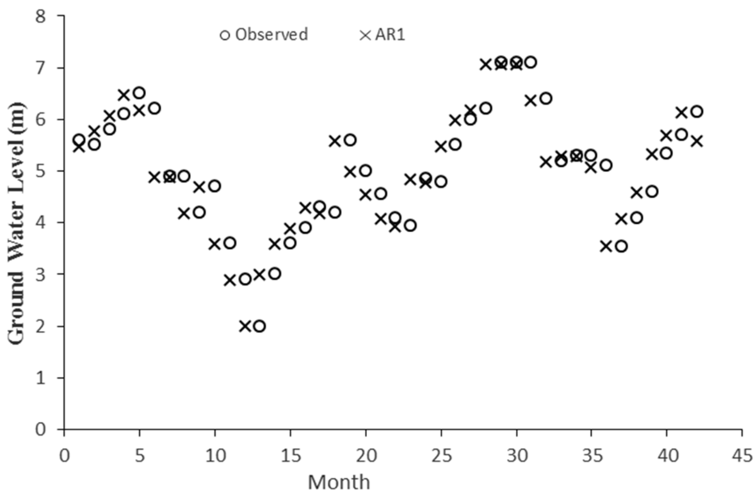


Fig. 2. Measurement and AR1 distribution chart for underground water level for test data

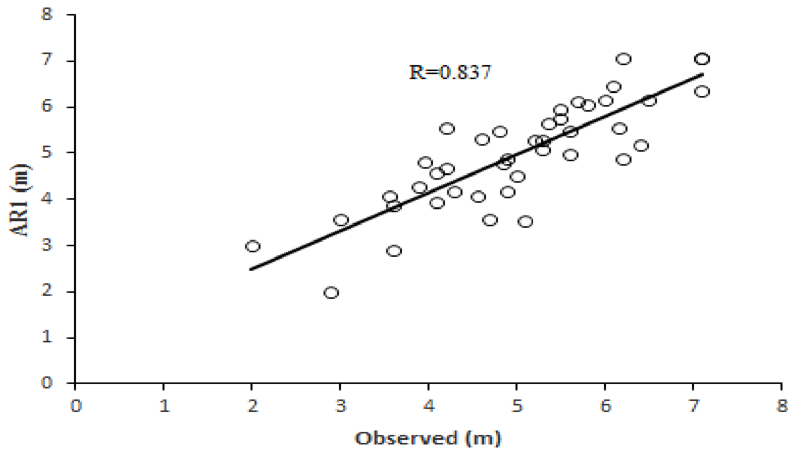
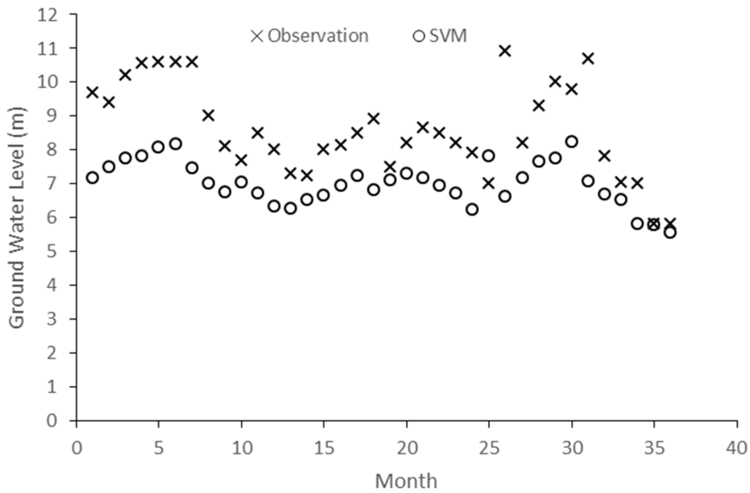


Fig. 3. Measurement and AR1 model scatter graph for ground water level for test data

AR1 model results show that the correlation coefficient is high and the groundwater level estimate is close to the actual values shown in Figure 2. The correlation coefficient R = 0.837 was obtained as seen the Figure 3.

SVM Model Results

In the support vector machines (SVM) model, 150 data of 192 were used training and 42 data were analyzed for the test. Monthly Mean Precipitation (MP), Monthly Average Temperature (MT), Monthly Ground Water Level (GWL+1) were used for the Ground Water Level Estimates. Estimated testing results are shown in Fig. 4 and 5 as, respectively, the distribution and scatter plots.



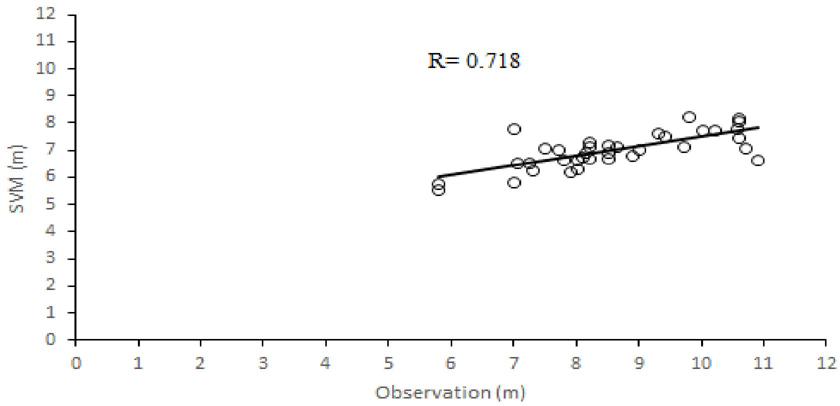


Fig. 5. Scatter graph of SVM and observed values

The SVM estimated values were observed in the test phase and gave worst results than the AR values. The correlation coefficient R = 0.718 was obtained as seen the Figure 5. SVM model results data are close to the actual values shown in Figure 4.

General Evaluation

Using monthly mean precipitation (MP), monthly average temperature (MT), monthly groundwater level (GWL+1) data from Kumlu region, correlation coefficient (R), the lowest mean squared error (MSE) and the total squared error (MAE) are calculated for performance evaluation of AR and SVM models. Results are used to compare the performance of model prediction and the observation data. Comparing parameters of MSE, MAE and R obtained from testing data are shown in Table 1.

Table 1. MSE: Mean square error, MAE: Absolute mean error, R: Correlation coefficient

Model	MSE	MAE	R
AR1	0,433	0,5288	0,837
SVM	3,4932	1,630	0,718

The best model is MSE, the MAE is the smallest, and the R is the biggest model. The AR1 model gave better results than the SVR model for MSE, MAE and R values.

Conclusion

In this study, Autoregressive (AR) model and support vector machines (SVMs) models were investigated in order to improve the methods to estimate the groundwater level. The accuracy of the SVMs model in groundwater level estimation was also investigated, and the results were compared with the AR1 model. Comparisons revealed that the AR1 model had the best accuracy in the groundwater level.

References

Chen, S. T.; Yu, P. S. 2007. Pruning of support vector networks on flood forecasting, *Journal of Hydrology* 347: 67–78. https://doi.org/10.1016/j.jhydrol.2007.08.029

Haykin, S. 1998. *Neural networks: a comprehensive foundation*. 2nd ed. Macmillan, New York.

Heesung, Y.; Jun, S. C.; Yunjung, H.; Bae, G. O.; Kang, K. L. 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. https://doi.org/10.1016/j.jhydrol.2010.11.002

Jin, L.; Chang, J. X.; Zhang, W. G. 2009. Groundwater level dynamic prediction based on chaos optimization and support vector machine, in *Proceedings of the 3rd International Conference on Genetic and Evolutionary Computing*, IEEE Transaction, Zhengzhou.

Schölkopf, B.; Smola, A. J. 2002. *Learning with Kernels*. MIT Press.

Xinying, Y. S.; Liong, V.; Babovic, E. C. 2004. SVM approach for real-time hydrologic forecasting, *Journal of Hydroinformatics* 6(3): 209–223.

Vapnik, V. 1995. *The nature of statistical learning theory*. Springer, New York. https://doi.org/10.1007/978-1-4757-2440-0

Mohsen, B.; Keyvan A.; Emery A., C. Jr. 2010. Comparative study of SVMs and ANNs in aquifer water level prediction, *Journal of Computing in Civil Engineering* 24(5): 408–413. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000043

Song, J.; Desphande, A. S.; Meier, T. B.; Tudorascu, D. L.; Vergun, S.; Nair, V. A.; Biswal, B. B.; Meyerand, M. E.; Birn, R. M., Bellec, P.; Prabhakaran, V. 2012. Age-related differences in test-retest reliability in resting-state brain functional connectivity, PLoS One 7. https://doi.org/10.1371/journal.pone.0049847