A DESIGN OF EDUCATIONAL COMPUTER GAME

Ali Ozdemir
Manisa Celal Bayar University, Turkey

Aysegul Alaybeyoglu
Izmir Katip Celebi University, Turkey

Kadriye Filiz Balbal
Manisa Celal Bayar University, Turkey

ABSTRACT: Nowadays, computer-assisted games have become a popular activity, especially for young people and children. It is expected that using computer games for educational purposes will make learning fun and amusing for every age of individuals. In this study, an educational computer game is designed especially to help students learn amusingly. The designed game can be used either individually or by teachers in the classroom environment and can easily be adapted to each course and topic. The questions in the game can be easily updated. In the study, questions are designed for mathematics lesson.

Keywords: educational game, educational computer game, mathematics game, game design

INTRODUCTION

Mathematics, which is indispensable in every field of daily life from past to present, is also a scientific discipline that originates from all other sciences. For this reason, great importance has been attached to the learning and teaching of mathematics throughout history. Effective teaching of mathematics is very important for understanding, learning, interpreting and using mathematics (Fırat, 2011).

In general, there are many difficulties in learning and teaching of abstract mathematics consisting of concepts and rules (Ekinözü and Sengül, 2007). Increasing opportunities through the development of technology enable the development of new methods and techniques for overcoming these difficulties. By using computer technology, abstract concepts can be embodied and learning and teaching activities can be more permanent, effective, meaningful and fun.

The vast majority of today's young people pass their time by playing computer games (Kafai, 2001). It is expected that applying game to the learning activity will provide more effective, meaningful and permanent learning. There are many studies in literature which use computer games in education and researchers attract more attention to this field day by day.

In this study, an educational game design which can be played in both class environment or individually is developed. It has been applied to mathematics lesson and at the end of the study, students' opinions are evaluated. In following sections of the paper, related works, educational computer games, the designed educational computer games, methods and findings are stated.

RELATED WORKS

There are many studies about educational computer games in the literature. In this section, related studies carried out in Turkey and abroad are summarized.

Nuhoğlu, Tüzün, Kaya and Cınar (2011) developed an educational game design model. They pointed out that using the model they developed in the educational game design process as a result of the summer examination of the area they were doing with the educational game.

Bakar, Tüzün and Cağıltay (2008) determined the opinions of the 6th grade students in the study they have conducted in the social studies course related to the use of educational computer games in the lessons. As a result of their study conducted on 24 number of students for 9 weeks, they found that the educational environment is found usefull by the students.

Offenbach (1964) designed a game about probability and attempted to determine its effects on learning ability of the pre-school and primary 4th grade students. He has found that both age groups have different approaches and older children are more normative.

Brown (2000) examined the effect of computer-assisted instruction on primary and middle school students in a two-year study. At the end of the study conducted on 214 students, it was determined that they could make a great contribution to the learning of computer-assisted instruction when it was used in an appropriate way.

Oztürk (2005) has implemented computer aided instruction design and has studied on the topic of permutation and probability. The software, which is provided with activities and simulations, also includes multiple-choice questions at the end of the topic.

Gürbüz and Birgin (2012) investigated the effects of computer-based instruction on the students' perceptions of probability. At the end of the study, they found that computer-assisted instruction is more effective than traditional teaching.

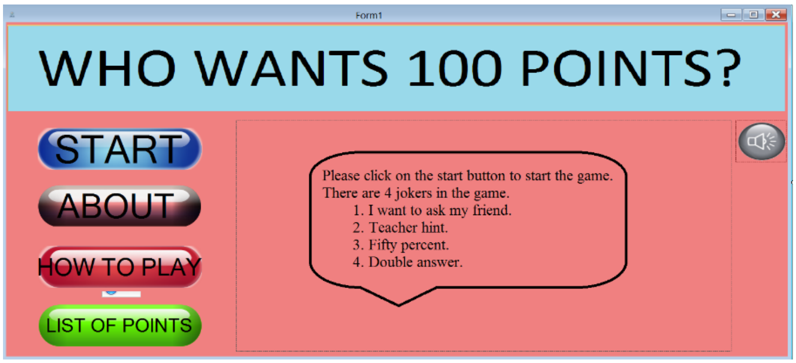
EDUCATIONAL COMPUTER GAMES

It is inevitable that computer technologies that effect every area of our life are also used in the field of education. Especially, courses like mathematics, in which abstract concepts are taught must be enriched with technology support. The visualization and concretization of abstract concepts by using computer technologies provide more effective, permanent and meaningful learning for students (Hewson, 1985; Novak and others, 1983; Thornton and Sokoloff, 1990 and 1998). It is supposed that student who actively participates in the learning environment learns the knowledge better (Bayraktar, 2002).
Today's young people spend a lots of their time on the computer. A large part of this time is spent by playing games. The use of highly enjoyable and motivating computer games for educational purposes may be beneficial in increasing the efficiency and quality of educational activities (Squire, 2011). Students who are bored in the classical learning environment and are reluctant to lessons can enjoy learning by participating in learning environments which are enriched by educational computer games. In particular, if students are organized as groups and participate in the game, learning can be encouraged in a collaborative manner. Indeed, in learning environments supported by an educational game, students are more concerned with the problem and concentrate better on the lesson (Çakmak, 2000).

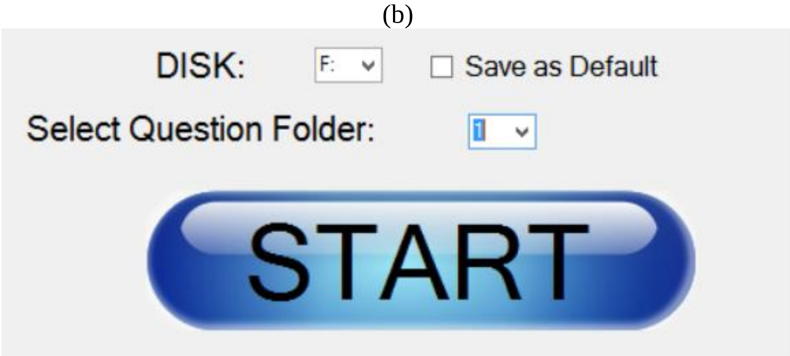
DESIGNED EDUCATIONAL COMPUTER GAME

When the educational computer games used in the learning environment are designed appropriately, they enable the students to participate in the learning process in an active way (Whelan, 2005). In the teacher-centered learning environment, the student who tries to learn only by listening passively, is active, authoritative and motivated during playing the game (Whelan, 2005).
The purpose of the designed educational computer game is to enable student to participate actively to the learning process in the classroom environment. The designed educational computer game is named "Who Wants to Get 100 Points". Figure 1 shows the screenshots of the educational computer game.

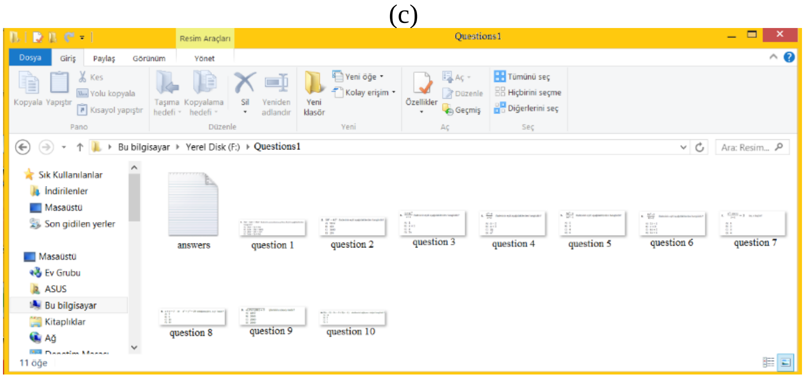
(a)


Form1
WHO WANTS 100 POINTS?
START
ABOUT
HOW TO PLAY
LIST OF POINTS
Please click on the start button to start the game.
There are 4 jokers in the game.
1. I want to ask my friend.
2. Teacher hint.
3. Fifty percent.
4. Double answer.

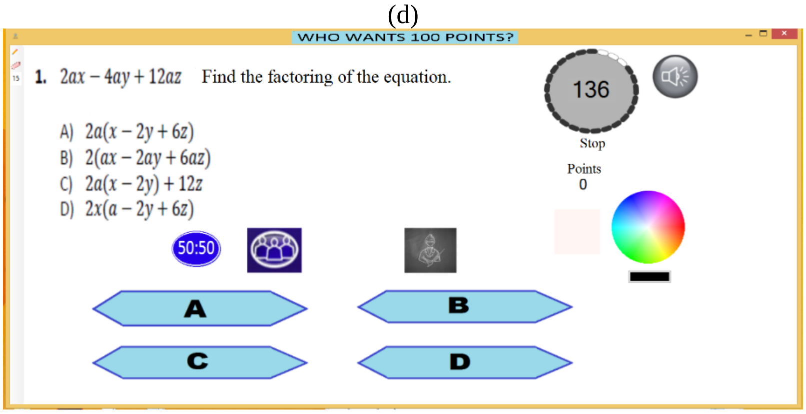
(b)


DISK:
F:
Save as Default
Select Question Folder:
1
START

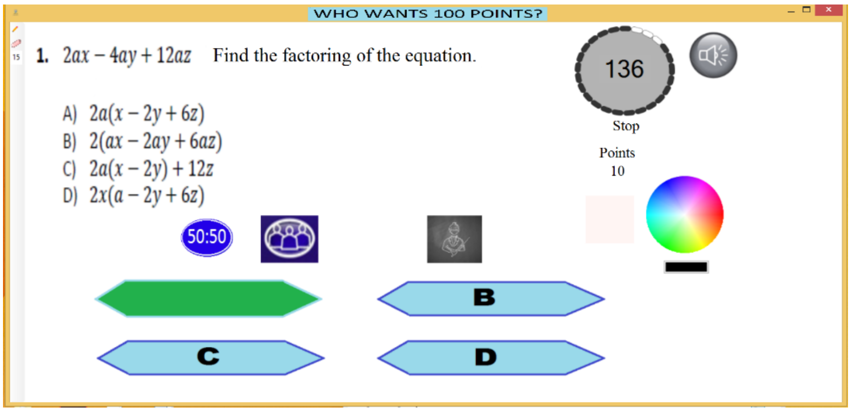
(c)

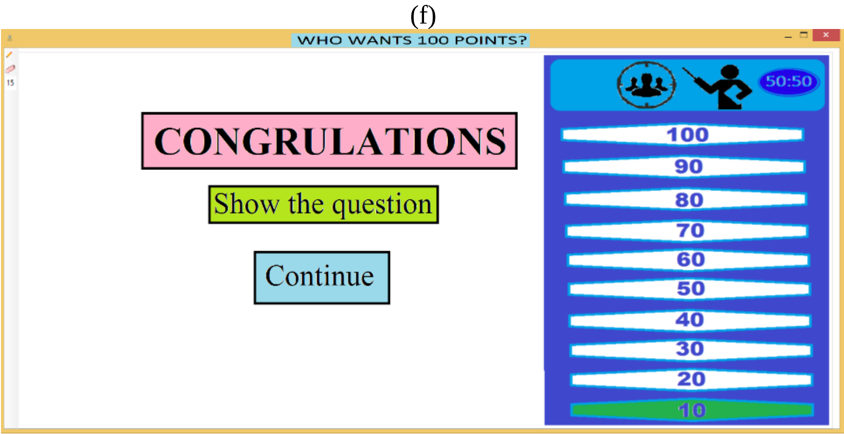

Questions1
Bu bilgisayar › Yerel Disk (F:) › Questions1
Sık Kullanılanlar
İndirilenler
Masaüstü
Son gidilen yerler
Masaüstü
Ev Grubu
ASUS
Bu bilgisayar
Kitaplıklar
Ağ
answers
question 1
question 2
question 3
question 4
question 5
question 6
question 7
question 8
question 9
question 10
11 öğe

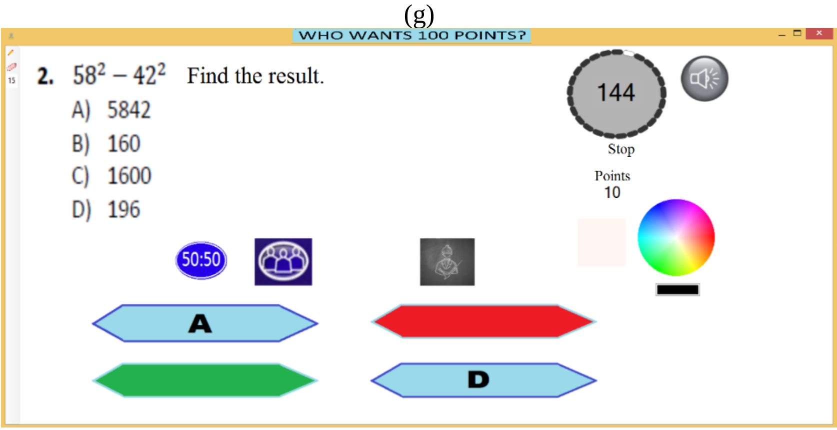
(d)

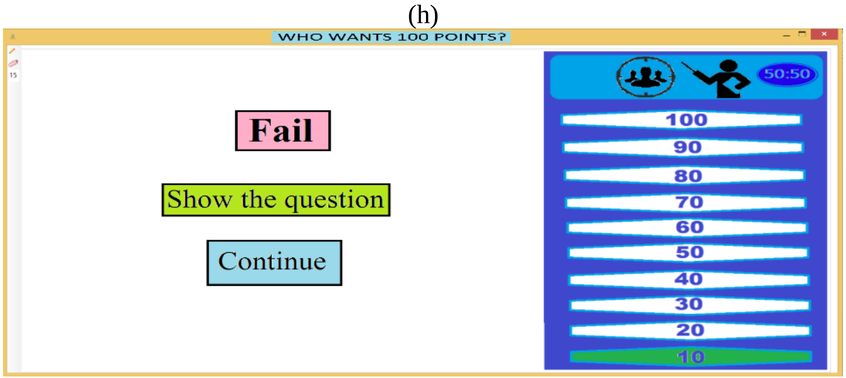

WHO WANTS 100 POINTS?
1. $2ax - 4ay + 12az$ Find the factoring of the equation.
A) $2a(x - 2y + 6z)$
B) $2(ax - 2ay + 6az)$
C) $2a(x - 2y) + 12z$
D) $2x(a - 2y + 6z)$
136
Stop
Points
0
50:50
A
B
C
D

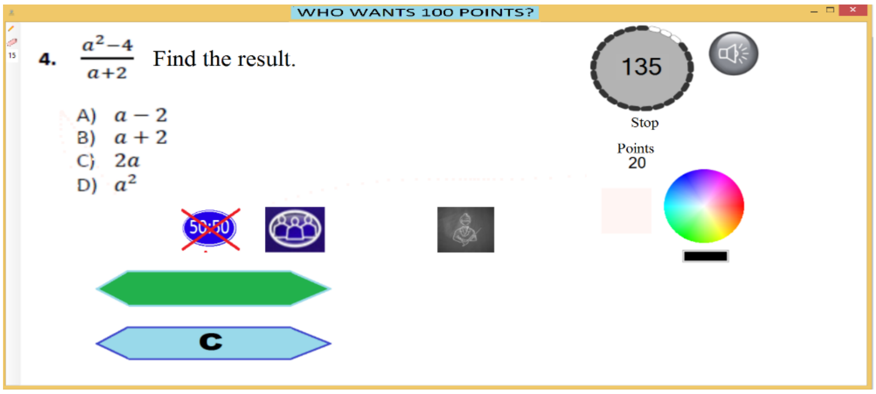
(e)


WHO WANTS 100 POINTS?
1. $2ax - 4ay + 12az$ Find the factoring of the equation.
A) $2a(x - 2y + 6z)$
B) $2(ax - 2ay + 6az)$
C) $2a(x - 2y) + 12z$
D) $2x(a - 2y + 6z)$
136
Stop
Points
10
50:50
B
C
D
(f)

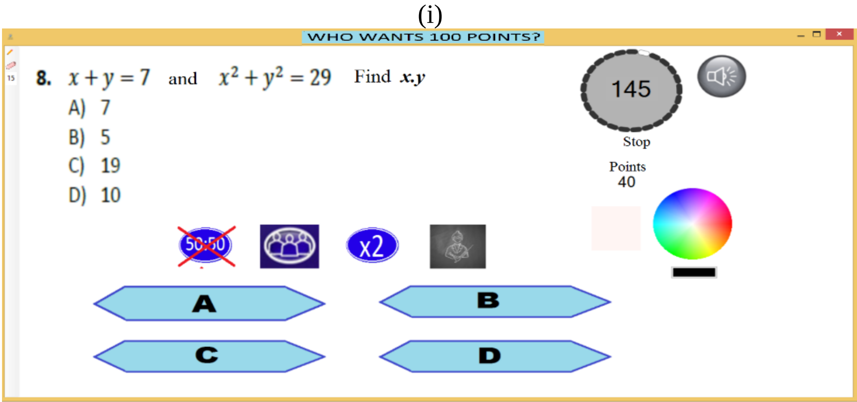

WHO WANTS 100 POINTS?
CONGRULATIONS
Show the question
Continue
50:50
100
90
80
70
60
50
40
30
20
10
(g)


WHO WANTS 100 POINTS?
2. $58^2 - 42^2$ Find the result.
A) 5842
B) 160
C) 1600
D) 196
144
Stop
Points
10
50:50
A
D
(h)

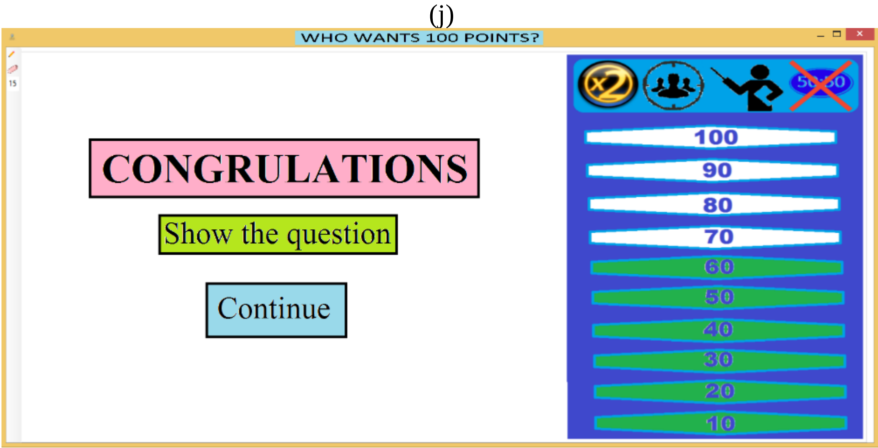

WHO WANTS 100 POINTS?
Fail
Show the question
Continue
50:50
100
90
80
70
60
50
40
30
20
10
(ı)



(i)



(j)



(k)

Figure 1. Screen Views of the Developed Educational Computer Game

Figure 1.a shows main page of the designed educational computer game. Figure 1.b shows screenshot of the information displayed when the "How to Play" button is clicked. When the "Start" button is clicked, a window is opened to select the drive and folder where the questions are stored during the game is being played (Figure 1.c). The questions were prepared in advance by the teacher and collected in picture format as shown in Figure 1.d and collected in a folder. By this way, several folders consisting of different questions can be prepared and different games can be played on top of each other. As it can be seen in Figure 1.e, there are options of jokers namely fifty percent, “I want to ask my friend” and teacher hint joker; one of which can be selected by the learner. Since the eighth question, there is also a double answer joker (Figure 1.j). The learner has 140 seconds to answer the question. But if necessary, the teacher can stop the time counter. Designed game also includes the status of the answer, whether it is correct or incorrect, pen and the eraser to use when solving the question, color palette for the pen, and sound effects to increase excitement (Figure 1.e,f,g,h,k).

METHOD

The aim of this study is to help students learn in an effective and permenant way by designing an educational computer game. In the scope of this study, an educational computer game design was realized. Semi-structured interviews were held with the students participating in the study at the end of the game, which was played with participation of all the students in the classroom environment and in mathematics lesson. These interviews were recorded and analyzed respectively.

RESULTS

In this part of the study, the findings related to the students' views on the practice in the light of the teacher's observations on the play line, records and interview results are explained.

In the interviews, all students stated that they enjoyed the educational computer game they played. For example, one of the students answer is "*I enjoyed the game very much. Because when you're solving a question, it's like you're in a real race. It is very nice to have fun with sweet excitement."* Another student stated his opinion as "*It was a very pleasant play. Solving a mathematical question has become pleasurable, fun and exciting."*. Many students stated that they were happy when they solved the questions and won points and they also stated that they never bored and they wanted to play again and again. Even a student said that "*Mathematics has never been so enjoyable. I always want to play games. I don't want you to stop asking questions.".*

A large number of students stated that educational play is contributing to the success of the course. While one of the student stated as "I *solved questions at home to earn more points in the rating*.", another student stated his opinion as "*I did not like math lessons, I could not. But I want to do it now. Because I want to score points*".

Most of the students agree that playing the game in a class environment with whole students participate made it very enjoyable. In this subject, one of the students stated his opinion as "*It is very nice to solve each question in groups of 2 or 3 people. I have solved the question with my friend. When I get stuck in a place, I ask my friend's opinion. By this way I both learn and proud of myself as I earned points. My friends are congratulating me."*.

According to observations made by the teacher during the play of the designed educational computer game, positive findings were reached. We observed that there was a collaborative learning environment based on solidarity and assistance manner between students and an exciting and entertaining lesson with active participation of students during the game. He also pointed out that students repeatedly play games and therefore insist on solving more questions.

CONCLUSIONS AND RECOMMENDATIONS

In this study, an educational computer game design was proposed and the opinions of the students related to the game played in the classroom environment were analyzed. When the findings in the study are examined, the following conclusions are obtained.

When both the observations of teacher and the results of the interviews made with the students participating in the application are examined, it is obtained that educational computer games play a positive role in the mathematics lesson. It is very important for the student-centered learning that the designed educational computer game enables students to participate actively in the classroom.

Especially in mathematics lessons it is very difficult for students to solve questions without bothering and even enjoying solving questions. The students who enjoyed playing, having fun and solving the questions with pleasure during the game stated that they were not bored and wanted to play again and again. The whole class act as a as a single team to get points, and they are congratulated by their friends when they solved the question, and they are supported when they could not solve it.

As a result, it was observed that students preferred a game-based course instead of a classical mathematics course. The teacher was happy to see that students solved the questions in amusement and found the application positive. In this study, educational game has made a great contribution to students' interest in mathematics lesson. Moreover, students have stated that they want to be taught in a similar way in all subjects.

According to findings from this study, the following subjects can be suggested for new studies.

1. The research was conducted by high school students. Other work can be done at different class levels.
2. The designed educational computer game was applied in mathematics lesson. By applying in other lessons, the effects on students can be determined.
3. The research data were obtained as semi-structured interviews and observation results. Other studies can be collected using different data collection methods and techniques.

4. More studies involving educational computer game design can be done.
5. While the game design is implemented, more interviews can be made with the educators. By this way educators participation can be provided while designing the game.

REFERENCES

Akgün, E., Nuhoğlu, P., Tüzün, H., Kaya, G. & Çınar M. (2011). *Bir eğitsel oyun tasarımı modelinin geliştirilmesi*. Eğitim Teknolojisi Kuram ve Uygulama, 1(1), 41- 61.

Bakar, A., Tüzün, H., & Çağıltay, K. (2008). *Öğrencilerin Eğitsel Bilgisayar Oyunu Kullanımına İlişkin Görüşleri: Sosyal Bilgiler Dersi Örneği*. Hacettepe Üniversitesi Eğitim Fakültesi Dergisi (H. U. Journal of Education) 35: 27-37

Bayraktar, B. B. (2002). Bilgi Sistemleri ve Yönetim Bilgi Sistemi Olarak Yüksek Performans Yönetim Modeli. Endustri İliskileri ve İnsan Kaynakları Dergisi. Cilt 4, Say. 2.

Brown, F., 2000. Computer Assisted Instruction in Mathematics Can Improve Students' Test Scores: A Study, Nabse Journal, p. 18.

Çakmak, M. (2000). İlköğretimde matematik öğretimi ve aktif öğrenme teknikleri. Gazi Üniversitesi Eğitim Fakültesi Dergisi, 20(3), 111-118.

Ekinözü, İ. ve Şengül, S. 2007. Permütasyon ve olasılık konusunun öğretiminde canlandırma kullanılmasının öğrenci başarısına etkisi. Kastamonu Eğitim Dergisi, 15(1), 251–258.

Fırat, S. (2011). *Bilgisayar destekli eğitsel oyunlarla gerçekleştirilen matematik öğretiminin kavramsal öğrenmeye etkisi.* Yüksek Lisans Tezi, Adıyaman Üniversitesi Fen Bilimleri Enstitüsü, Adıyaman.

Gürbüz, R. ve Birgin, O. 2012. The effect of computer-assisted teaching on remedying misconceptions: The case of the subject "probability". Computers & Education 58, s. 931–941

Hewson, P.W. (1985). "Diagnosis and Remediation of an Alternative Conception of Velocity Using a Microcomputer Program". American Journal of Physics. 53,684-690.

Kafai, Y. B. 2001. The educational potential of electronic games: From games-to-teach to games-to learn. Playing by the rules. https://culturalpolicy.uchicago.edu/sites/culturalpolicy.uchicago.edu/files/kafai.pdf Erişim Tarihi: 10.01.2017

Novak, J.D., Gowin, D.B. ve Johansen, G.T., (1983), "The Use of Concept Mapping and Knowledge Vee Mapping with Junior High Science Students". Science Education, 67(5), 625-645.

Offenbach, S. I. 1964. Studies of Children's Probability Learning Behavior: I. Effect of Reward and Punishment at Two Age Levels. Child Development, 35(3), 709- 771.

Öztürk, G. 2005. İlköğretim 8. Sınıf Düzeyinde Permutasyon ve Olasılık Ünitesinin Bilgisayar Destekli Öğretim Tasarımı. Yüksek Lisans Tezi. Balıkesir Üniversitesi, Balıkesir.

Squire, K. (2011). Video Games and Learning: Teaching and Participatory Culture in the Digital Age. Technology, Education--Connections (the TEC Series). Teachers College Press. 1234 Amsterdam Avenue, New York, NY 10027.

Thornton, R.K.; Sokoloff, D.R. (1998), "Assesing Students Learning of Newton's Laws: The Force and Motion Conceptual Evaluation of Active Learning Laboratory and Lecture Curricula". American Journal of Physics. 66,338-352.

Thornton, R.K., Sokoloff, D.R. (1990), "Learning Motion Concepts Using Real-Time Microcomputer-Based Laboratory(MBL) Activities to Help Students Overcome Some Common Conceptual Difficulties in Kinematics". American Journal of Physics. 58,858-867.

Ural, M. N. (2009). Eğitsel bilgisayar oyunlarının eğlendirici ve motive edici özelliklerinin akademik başarıya ve motivasyona etkisi. *Doktora Tezi, Anadolu Üniversitesi Eğitim Bilimleri Enstitüsü, Eskişehir*.

Whelan, D.L. (2005). Let the games begin. *School Library Journal*, *51*(4), 40-43.

INVESTIGATION OF METHODS USED IN MODELING AND FORECASTING OF STREAM FLOWS AND LOCATION OF MACHINE LEARNING METHODS IN THIS AREA

Osman Selvi
Istanbul Aydin University, Turkey

Ilham Huseyinov
Istanbul Aydin University, Turkey

ABSTRACT: Machine learning is used today to produce different kinds of results in many areas of engineering. In our work, studies on modelling hydrological processes of machine learning methods have been investigated. The researchers make short or long term forecasts by modeling the hydrological processes of a selected basin. Researchers thanks to the estimates made can be created early warning systems in the short term or can be provide knowledge to decision makers in the long term for sustainable water resource management. In the literature survey made in this context, the answers to the questions about the modeling methods of streamflows in hydrology, the place of machine learning methods in this area, input parameters used in modeling, input methods used for selecting models and performance evaluation methods of developed models were searched. Most commonly used methods for modeling stream flows are artificial neural network methods (ANN), support vector machine(SVM), bayesian methods, fuzzy logic methods and AR & MA derivative methods. While stream flow data is used as the basic input parameter, it is seen that parameters such as the amount of precipitation and the evaporation rate are used in some studies. The most frequently used methods for efficiency evaluation are MSE, MAE, RMSE, r, R^2 and NSE.

Keywords: stream flow forecasting, machine learning, hydrological models

INTRODUCTION

Water is an important compound that supply viability in the biosphere which have an important task like blood circulation. The continuation or sustainability of biophysical and social processes is closely related to the existence of water (Falkenmark, 2017).

In our world with limited water resources, population growth, rapid urbanization, the increase in the standard of living, the need for renewable energy sources resulted in an increase in the demand for water increase with each passing day. In addition, hazardous waste involved in water resources, climate change and other adverse effects increase the stress on water resources. Therefore, effective management of water resources both locally, regionally and globally is important for the welfare of humanity, economic development and protection of the environment. ((WWAP) U. N., 2017) ((WWAP) U. N., 2016) (KANSAL, 2005) (Rockström and others, 2014) Human as a dominant force behind the changes in water resources, deals with the optimization of using water for different purposes in the management of water resources for the continuation of its existence (Falkenmark, 2017).

There are many research topics such as stream flow forecasting, change in lake water level or storage volume prediction in the literature for the purpose of effective management of water resources. Stream flow forecasting, which is the subject of our work, is a demanding process due to the complexity and non-linear structure of the factors affecting the formation of stream flow (Maheswaran & Khosa, 2012) (Yaseen and others, 2016).

Researchers use various modelling tools (SWAT, TOPNET, GR4J, GR2M etc.) or machine learning tools (Neural Network, Support Vector Machine, Fuzzy Lojic, Bayesian Approaches, ARIMA or ARMAX etc.) or hybrit modelling in order to effective predictions about the future. And they compare them with various performance criteria such MSE, MAE, RMSE, NSE, r, R^2. Researchers which using machine learning methods are often used lagged days stream flow values for their model. The precipitation and evaporation parameters, however, used with the stream flow parameter at various delay times.

Researches mostly are used trial and error method or corelation analysis methods (ACF, PACF etc.) for appropriate feautere selection. According authors, in the studies made under the article on stream flow forecasting, the least studied input parameter is the selection topic. Stream flow forecasting is used for many purposes depending on the forecast period. Hourly or daily forecasting models approaches often involve floods or disasters are aimed to be avoided. By weekly, monthly or seasonal forecasting models, the aim is to ensure effective use of water resources for various areas. By long term prediction, it is aimed to assist decision makers in future planning.

In this context, 32 studies between 1998 and 2017 were examined. The methods used in the researchs are examined under the headings such as estimation ranges, performance evaluation criterias and countries of implementation.

METHODS

In the literature, it is possible to examine the methods used for stream flow forecasting in three groups.
These;

1. Hydrological modelling tools,
2. Machine learning tools,
3. Create hybrid models from hydrological modeling tools and machine learning tools.

When the studies are examined, it seems that there is not a single solution that can produce the right result under all conditions. When the performance of the same method is examined in terms of the input parameters, it is seen that there are differences in the field from the applied area.

While (Demirel, Venancio, & Kahya, 2009) using stream flow and precipitation as the input parameters in their studies, (XU, ZHU, ZHANG, XU, & XIAN, 2009) using stream flow, precipitation and evaporation values for the best result.

1. Hydrological Modeling

In literature review most used hydrological model is SWAT (Soil Water Assesment Tool). And others TOPNET, TOPMODEL, GR4J and GR2M models. We can be examined hydrological models based on two perspective. First based on spatial representation and second based on internal processes (Toro, Meire, F.Gálvez, & Riverola, 2013).

1.1. Examining Based on Spatial Representation

1.1.1. Aggregated Hydrological Models

In aggragated hydrological models, rain and other hydrological variables for modelled basin are assumed to have global and constant properties for the whole basin.

1.1.2. Distributed Hydrological Models

In distributed models a selected basin divided into cells. And rainfall and other hydrological variables used for simulating each cells. The hydrological models in this structure are more data and process intensive modeling methods than the other modelling methods. Ant they require more expretise than others. TOPNET model and TOPMODEL model are two examples used in this respect.

1.1.3. Semi Distributed Hydrological Models

At this modelling approaches, a basin divided in sub basins. And hydrological variablles like rainfall are divided for subbasin and hydrological process is calculated for each sub basin.

1.2. Examining Based on Internal Process

When the hydrological model is analyzed in terms of hydrological processes, it is possible to distinguish three categories. First one based on physically based, second on conceptual based and third metric based models.
The hydrological models in all three categories show differences in the manner in which hydrological processes are handled as follows.

1.2.1.Physical Models

Such models are models in which physical, chemical or biological processes are represented mathematically. They use measurable state variables of time and space. Physical models are useful methods to understand the physical operation of the region where modeling is to be performed. However, one of the challenges is the multitude of the needed parameters.

1.2.2. Conceptual Models

In this modeling approach, hydrological processes all trying to be processed through a conceptual model. Calculation of hydrological processes takes place on the created conceptual model. GR4J model and GR2M model are two examples used in this respect.

1.2.3. Empirical Models

Models that work in this modeling approach are models that do not take into account hydrological processes related to the relevant basin. This modeling approach aims to establish various relationships among existing data and to model the behavior of the basin. And at these models phisical process don't care. Their main advantage is that they require minimum data. But it can be disadvantage to don't care physical process.

This modeling approach will be examined in more detail below under the heading of black box methods. Readers who wish to examine the detailed comparison of hydrological models can benefit from (Devi, P, & S, 2015) or from (Kauffeldt, Wetterhall, Pappenberger, Salamon, & Thielen, 2016).

2. Black Box Modeling

Unlike hydrological modeling approaches, black-box modeling approaches are not concerned with how the hydrological process takes place. Black box approaches are known as "data driven". The aim here is to establish a relationship between the target data for the relevant hydrological structure (Stream Flow, Precipation, Evatoration, etc.). In this way hydrological modeling of the relevant region can be carried out and forecasts for the future can be made.

Most of the studies for stream flow forecasting in the literature are composed from Neural Networks. And other frequently used modelling approaches are Support Vector Machine(SVM), Fuzzy Logic approaches, Bayesian approximations, AR & MA derivative approaches.

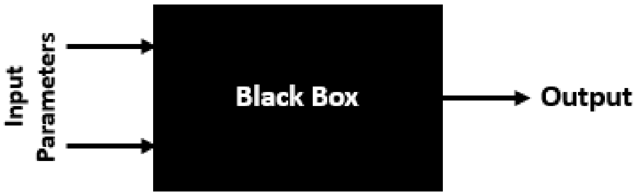


Figure 3 - Black box methods

In the black-box approach, it is seen that the input parameters are mostly used for the forecasting of stream flow and precipitation parameters. Selecting the correct input parameters is another important topic in the Black Box methods, which have the approach of entering a certain set and associating the parameters with the output parameters. A more detailed explanation will be given in Chapter-4.

The stream flow or precipitation parameters used are not based solely on the previous time slot data from the time slot to be estimated. Some studies show that they are used to estimate values from more than one previous time period. In the literature, the number of days of the past is indicated by lag. Lag (1) represents the previous day, Lag (2) represents the data of the past two days.

2.1. Neural Network

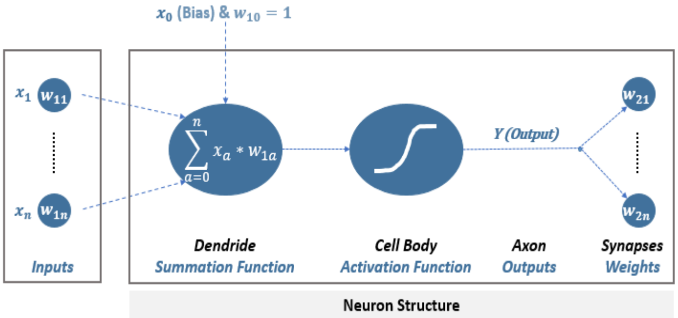


Figure 4 - The logic of working an artificial neuron

As can be seen, a simple artificial neural network uses input parameters (outputs and weights from other neurons) to produce output of the kind targeted with the aid of two functions. These are the summationt function and the activation function. The task of the summation function is to calculate the sums of the multiplied results of weight and produced output by previous neurons. The activation function is used to produce the desired results from the neuron's collection function. A linear function can be used when it is considered that there is a linear relationship between input and output data, while a step function can be used when two different results like 0 or 1 are desired to be produced. In artificial neural networks, each neuron will produce results depending on the nature of the selected activation function.

An artificial neural network consists of three different types of layers. Input layer, hidden layer and output layer. When studies done in the literature are examined, it can be seen that the activation function of artificial neural networks varies in the selection of the backpropagation algorithm, the selection of the number of hidden layers or the selection of the number of neuron in the hidden layers. The researcher specifies the number of hidden layers, the number of neurons in each these layers, the activation function to be used in neuron, and the backpropagation algorithms, so that the ideal result can be obtained. When the activation functions used in the hidden layers are examined, it is seen that the following activation functions are used. (Prasad, Deo, Li, & Maraseni, 2017) is used tangent sigmoid ve log sigmoid, (Yaseen and others, 2016) logarithmic sigmoid, (Noori and others, 2011), (Demirel, Venancio, & Kahya, 2009), (Ballini, Soares, & Andrade, 1998) and (Uysal, Lorman, & Lensoy, 2016) sigmoid, (Noori & Kalin, 2016) log-sigmoid and hyperbolic tangent sigmoid, (B.Humphrey, Gibbs, Dandy, & Maier, 2016) Logistic transfer function, (Singh A. , 2016) Gaussian transfer function and (XU, ZHU, ZHANG, XU, & XIAN, 2009) ise used logistic function. When Linear transfer function used by (Prasad, Deo, Li, & Maraseni, 2017), (Demirel, Venancio, & Kahya, 2009), and (B.Humphrey, Gibbs, Dandy, & Maier, 2016) at output layer, purline functions used by (Prasad, Deo, Li, & Maraseni, 2017) and positive linear function used by (Prasad, Deo, Li, & Maraseni, 2017). Another properties that should be known in an artificial neural network is the choice of the back propagation algorithm.

The task of backpropagation algorithms is to change the weights between the neurons so that the difference between the net result and the actual result can be minimized. If selected backpropagation is to be specially examined, (Prasad, Deo, Li, & Maraseni, 2017) and (Uysal, Lorman, & Lensoy, 2016) use the Newton's method and (Kalteh, 2013) use the Levenberg-Marquardt algorithm. It is seen that 32 studies that were investigated in the literature survey developed 18 indent ANN and derivative models. Researchers are evaluating the results of hydrological models or other machine learning tools with ANN in their work. Some researchers used ANN hydrological models in conjunction with performance evaluation. Let's start by examining the hydrological models and models created with ANN. (Singh A. , 2016) compared SWAT with the Radial Basis Neural Network (RBNN) and showed that RBNN gave better results. (XU, ZHU, ZHANG, XU, & XIAN, 2009) Compared the ANN with the XXT hydrological model and showed that ANN gave better results. (XU, ZHANG, & ZHAO, Stream Flow Forecasting by Artificial Neural Network and TOPMODEL in Baohe River Basin, 2009) Compared the performances of the ANN with the TOPMODEL hydrological modeling tool and found that ANN gave better results. (Demirel, Venancio, & Kahya, 2009) Show that ANN can model peak flow better than SWAT in their study.

Now let's examine some of the results of ANN modeling with other machine learning tools.(Sun, Wang, & Xu, 2014) compared ANN with Gaussian Process Regression (GPR) and ARMAX in their study. When the results produced by the models are examined, they stated that GPR is the best end result model. ANN produced better results than ARMAX. (Baratti and others, 2003) compared ANN and ARMAX results in the study they performed and stated that ANN produced better results. We will also discuss further work on ANN in the following issues, as the topic title is compared to ANN in other topics.

2.2.Support Vector Machine

Is a another commonly used methods for modeling hydrological processes is the SVM. SVM is a statistically based machine learning tool. SVM is used for classification in machine learning methods as well as for regression (Kalteh, 2013). If it is desired to classify with SVM and the classes in the data set are separable, it is desirable to draw a hyperplane between the classes so that they are farthest from class boundary lines. When SVM is used for regression, it is desirable to have such a function that it is within a certain range ($+\varepsilon$, $-\varepsilon$) to find a function that can produce results that will cover the whole data set.

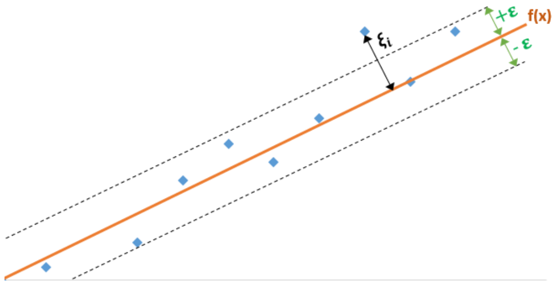


Figure 5. Working logic of SVR

Suppose we have an f (x) function like Eq.-1 (Equation-1) that will cover almost the entire data set, such as the figure.

$\boldsymbol{f(x)} =< \boldsymbol{w}, \boldsymbol{x} > +\boldsymbol{b}$ (Equation-1)

w is used for support vector weight and **b** is bias.

In order to satisfy this equation, it is aimed to minimize **w**.

$minimize\ \frac{1}{2} \parallel w \parallel^2$ (Equation-2)

Subject $\begin{cases} y_i - < w, x_i > -b \leq \varepsilon \\ < w, x_i > +b - y_i \leq \varepsilon \end{cases}$ (Equation-3)

Sometimes it may be desirable to increase the error tolerance of the results produced by the function. In this case $\boldsymbol{\xi_i}, \boldsymbol{\xi_i^*}$ variables are used.

In this case, equations Eq.-2 and Eq.-3 will be updated as follows.

Subject: $minimize\ \frac{1}{2} \parallel w \parallel^2 + C \sum_{i=0}^{N} (\xi_i, \xi_i^*)$ (Equation-4)

$Constraints \begin{cases} y_i - < w, x_i > -b \leq \varepsilon + \xi_i \\ < w, x_i > +b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}$ (Equation-5)

C is constant user defined in the objective function in Eq.-4. And is used as a balance element.

As can be seen from the equation, slack variables are used as an balance element on the equation.

The method used to find the minimum or maximum value of a function bound to a certain constraint is the Lagrange method.

Lagrange multipliers are often used to solve the objective function in Eq.-4.

$$\mathrm{W}(a, a^*) \max -\frac{1}{2} \sum_{i}^{N} \sum_{j}^{N} (a_i - a_i^*)(a_j - a_j^*) K(x_i - x_j) - \varepsilon \sum_{i}^{N} (a_i + a_i^*) + \sum_{i}^{N} y_i (a_i - a_i^*)$$

$$Constraints \begin{cases} \sum_{i}^{N} (a_i - a_i^*) = 0 \\ 0 \leq a_i \leq C \\ 0 \leq a_i^* \leq C \end{cases}$$

The equation for non-linear function is as follows.

$$f(x) = \sum_{i}^{N} (a_i - a_i^x) K(x_i, x_j) + b$$

a represent lagrange coefficients. And K represent selected kernel function.

In SVR application, researchers first point out that a proper kernel function is chosen (Asefa, Kemblowski, McKee, & Khalil, 2006). It is seen that radial basis function is used in all of the kernel functions used in the researches. (Asefa, Kemblowski, McKee, & Khalil, 2006), (Kalteh, 2013), (Sattari, Pal, Apaydin, & Ozturk, 2013)

In order to model with SVM, the researchers must have chosen the appropriate **C**, **ε** and **ξ** values.

When the studies comparing SVM with other models are examined, it is seen that the researchers reached the following results.

(Yaseen and others, 2016) show that ELM produces better results than SVM when they compare the models they created with ELM and SVM in their study.

They show that when comparing (Sattari, Pal, Apaydin, & Ozturk, 2013) models in which they work with ELM and SVM, the ELM produces similar results to SVM.

(Kalteh, 2013) used a 40-year stream flow value for monthly stream flow forecasting. In their work, they are derivated different models from SVR and ANN. They have reported that SVR produces better results when comparing the performances of the models it produces.

2.3. Fuzzy Logic

Fuzzy Logic is a introduced by Lotfi A. Zadeh in 1965. In this approach, it is accepted that all events in nature have certain degrees of uncertainty, which does not include certainty. In fuzzy logic there is a different approach from classical cluster membership. In fuzzy logic, the actual value of any variable is expressed by membership rates between 0 and 1 by each of the fuzzy clusters concerned. Various membership functions are used to determine membership levels. With the help of membership functions, real values can be transformed into fuzzy values or fuzzy values with real values (Defuzzification).

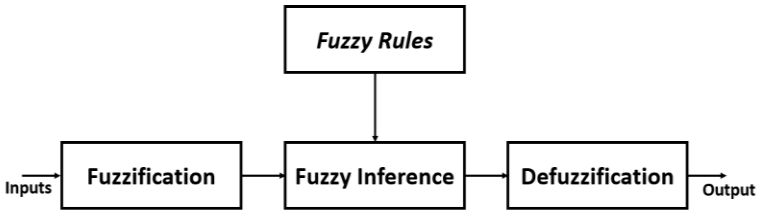


Figure 6 - Working logic of Fuzzy Logic

With the help of the fuzzy rules it is aimed to establish relation between input variables and output value. Fuzzy rules are created by people in the field of expertise that is required to be modeled. The fuzzy rules are similar to the if examples in basic programming information. In the fuzzy inference phase, the results obtained from membership functions for each rule created by the investigator are obtained by logical operations with various results. Subsequently, the results obtained from each rule are combined using methods such as the Tusukamoto method, the Takagi-Sugeno-Kang method, or the mamdami method. In the last stage, the fuzzy results obtained from the fuzzy interference phase are converted into real values.

Literature research in this context shows that fuzzy logic is used as part of hybrid systems. (Ballini, Soares, & Andrade, 1998) Compared the results of a model constructed with ANN using ANFN (Adaptive Neural Fuzzy Network) containing fuzzy logic and ANN. The comparison also shows that ANFN models produce better results from the ANN network. (Valença & Ludermir, 2000) Developed different models with PARMA (Periodic ARMA) and FuNN (Fuzzy Neural Network). When models evaluate the results, FuNN has shown that produces better results. In another research different from before, (Turan & Yurdusev, Fuzzy Conceptual Hydrological Model for Water Flow Prediction, 2016) are used fuzzy system in interior process (Soil Moisture and Routing Storage Process) of GR2M hydrological modeling tool. Modeled soil moisture and routing storage processes using fuzzy systems. They have shown that the F-GR2M models produce better results in the GR2M around 10%.

2.4. AR & MA Derivative Models

Some of the methods used with stream flow forecasting are ARIMA or ARMAX models. Relevant models are derived under the heading AR & MA because they are derived from Autoregressive (AR) and Moving Average (MA) models. They are known in the literature as Box-Jenkins models. Before explaining ARIMA and ARMAX models, AR and MA and ARMA models must be understood. The requirement for three models (AR, MA or ARMA) to be installed is that the data set must be stationary.
In AR model, a dependent variable is a function of lagged days data.

$$AR(p) = b + \sum_{i=1}^{p} a_i Y_{t-i} + \varepsilon_t$$

Y the future value is the desired value to predict, while the $\boldsymbol{a}$ coefficient indicates the relationship between the lagged values and the value to be estimated. $\boldsymbol{\varepsilon_t}$ Represent the white noise.
In the MA model, it is assumed that the function of **Y** is related to past error terms in predicting the future value.

$$MA(q) = \varepsilon_t + \sum_{i}^{q} m_i \varepsilon_{t-i}$$

ε_t Represent the white noise,while the m coefficients indicate the relationship between the error terms and the value to be estimated.

$$ARMA(p,q) = b + \varepsilon_t + \sum_{i=1}^{p} a_i Y_{t-i} + \sum_{i}^{q} m_i \varepsilon_{t-i}$$

ARMA model is a combination of AR and MA models.
The researcher decides which AR, MA, or ARMA model to use by looking at the graph of the autocorrelation function (ACF) and partial autocorrelation fuction's results.

If the ACF graph is slow, and the PACF graph shows a rapid decline, the AR model is preferred. If the ACF graph shows a rapid decline and the PACF graph shows a slow decline, the MA model is preferred, whereas if both graphs show a slow decline, the ARMA model is preferred. Unlike the ARMA model, ARIMA model can be operated with non-stationary data set. In this model, the non-stationary data set is subjected to a certain degree of difference processing until it becomes stationary. For this reason, unlike ARMA model, it also needs d parameter. The d parameter indicates the differencing degree. With stabilization, observation values are made to move within a fixed average. In the ARMAX model, unlike the other Box-Jenkins models, only the history data of the target variable is used as the input parameter. Variables from different types are also used in this model which are thought to affect the outcome of the targeted variable. Unlike the ARMA model, the k variable is used in the equation but represents other variables.

$ARMAX(p, q, k) = Y$

$$Y = b + \varepsilon_t + \sum_{i=1}^{p} a_i Y_{t-i} + \sum_{i}^{q} m_i \varepsilon_{t-i} + \sum_{i}^{k} d_i x_{t-i}$$

(Sun, Wang, & Xu, 2014) Compared the models they created with ANN (MLP), GPR and ARMAX in their study, and showed that GPR and ANN (MLP) produced better results than ARMAX. (Baratti and others, 2003) have created two different models of ANN(MLP) and ARMAX in their studies. They have shown that ANN (MLP) produces better results than ARMAX when comparing two models. (Valença & Ludermir, 2000) Compared the FNN (Neural Fuzzy Network) model with the ARIMA model and showed that the FNN models produce better results than the ARIMA models.

3. Hybrid Approaches

When investigating the preferred methods of this method, it is seen that the modeling of the related basin using hydrological modeling tools. Researchers who prefer this approach aimed to make forecasting by outputting output values as input parameters to black box approaches (mostly ANN) without calibrating output obtained from the modeling program (Noori & Kalin, 2016) or after calibrating (B.Humphrey, Gibbs, Dandy, & Maier, 2016)**.**

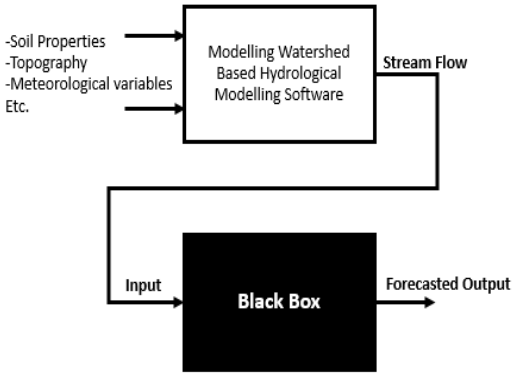


Figure 7 - Working logic of hybrit approaches

Researchers therefore aim to achieve more robust results by combining the power of hydrological modeling programs to better model the hydrological process and the ability of black box approaches to make better predictions. (Noori & Kalin, 2016) and (B.Humphrey, Gibbs, Dandy, & Maier, 2016) we shall try to explain the reasons why the hydrological modeling programs and the ANN method from black box approaches are preferred together;

As an advantage of the use of hydrological modeling programs, stream flow data can be obtained better than ANN methods, since the modeling takes into account the hydrological processes. The disadvantage of using hydrological modeling programs is that if this approach is preferred, a number of temporal and spatial variables will be needed. However, in this modeling approach, it is stated that the calibration process is much more time consuming and laborious than ANN. The advantage of modeling with ANN is that the relationship between limited and well-chosen input parameters and output data can be established without having to evaluate the hydrological processes of the relevant basin.

In addition to the advantage described above, it is stated that ANN approaches can produce better results in some cases than the forecasting data obtained with hydrological modeling programs (B.Humphrey, Gibbs, Dandy, & Maier, 2016). Modeling with ANN has advantages as well as possible disadvantages. These are: The input parameters may be selected incorrect, backpropagation algorithm may be stick around to the local minimum or overfitting. (Noori & Kalin, 2016) and (B.Humphrey, Gibbs, Dandy, & Maier, 2016) have worked on three different scenarios. Forecasting with the hydrological modeling tools, forecasting with the black box approaches (both preferred different variants of ANN) and the third model hybrid models in both models. The results show that when evaluated by researchers, hybrid approaches produce better results than the other two modeling approaches.

FEATURES AND SELECTION CRITERIA

In hydrological modeling approaches, the parameters required by the model used are specific. However, in the black box approach, the choice of input parameters is decided by the researchers. The input parameters and input parameter selection criteria used in the black box approaches under this topic are examined. (XU, ZHU, ZHANG, XU, & XIAN, 2009) Indicated that Stream Flow, Precipication, Evaporation data as input parameters are parameters that should be used in the basin. (XU, ZHANG, & ZHAO, Stream Flow Forecasting by Artificial Neural Network and TOPMODEL in Baohe River Basin, 2009) Have indicated that past and current precipitation and stream flow data are important input parameters for stream flow forecasting.

(XU, ZHANG, & ZHAO, Stream Flow Forecasting by Artificial Neural Network and TOPMODEL in Baohe River Basin, 2009) Have observed that the results obtained from the stream flow and precipitation data as ANN input parameters produced better results than the studies using only stream flow or only precipitation data. Also there was a limited decrease in the performance of the ANN structure when the evaporation data was used as the input parameter. If the black-box approaches are evaluated in general, it is seen that the Stream Flow parameter is commonly used in various lagged numbers. The precipitation and evaporation parameters, however, appear to be used with the stream flow parameter at various delay times. Here, it is seen that only inputting the model data to be used is not enough and it is seen that the selection of the appropriate input parameters is an important topic title in order to increase the performance of the model. One of his tasks is to be able to select the correct input parameters so that the model he wants to establish can produce the closest accuracy. (Yaseen and others, 2016)

An unnecessarily selected parameter can weaken the working speed of the model or the resultant power it produces. Or correctly selected parameters can increase the degree of closeness to the model's accuracy. When the approaches used to select input parameters are examined, it seems that basically one of the two approaches is preferred. First is trial and error method and second is correlation analysis.

1.Trial and Error Method

In this method, the researcher uses different combinations of input parameters or different lagged day’s values in his hands to produce the results closest to the truth.

2. Correlation Analysis Methods

Another approach used to select input parameters is to analyze the correlation between input parameters and output parameters. For this purpose, ACF (Auto-Corelation Function) and PACF (Partial Auto Corelation Function) methods are widely used. (Yaseen and others, 2016) aimed to increase the performance of the model by using ACF and PACF functions with different lagged days of the stream flow parameter. In each of the five models created by them, they used the stream flow value for different days of the past to find the value of the day that wanted to estimate.

3. Another Feature Selection Methods

(Prasad, Deo, Li, & Maraseni, 2017) using the IIS(Iterative Input Selection) algorithm and (B.Humphrey, Gibbs, Dandy, & Maier, 2016) using the IVS(Input variable selection) algorithm to use the most ideal input parameters for modeling. In the studies made under the article on stream flow forecasting, the least studied input parameter is the selection topic.

FORECAST RANGE

The data on stream flows and other hydrological variables belonging to many rivers in our country and in the world have been measured regularly for many years. From this point of view, forecasting can be done by considering stream flow forecasting time series. When studies done in the literature are examined, it is seen that forecasting is used to make short-term estimates (daily, hourly) or long-term estimates (seasonal, monthly, weekly). In short-term estimates, the aim is to create early warning systems to protect against natural disasters such as floods. (Yaseen and others, 2016) (Keshtegar, Allawi, Afan, & El-Shafie, 2016) In long-term studies, the goal is to be able to provide highly accurate estimators of decision makers in order to effectively manage water resources

(Keshtegar, Allawi, Afan, & El-Shafie, 2016) (Yaseen and others, 2016). This enables effective sharing between stakeholders (Agriculture, Industry, Nature, People) using water resources. It can also be used to set up early warning systems to protect against various natural disasters.

MODEL EVALUATION CRITERIAS

Evaluating the performance of a generated model can be more than one purpose. These (P. Krause, 2005), assessing how closely the model created can model the related basin to the truth, improving the performance by changing the input parameters or the internal processes of the model by looking at the results of the generated model and comparison of performances of studies done by different researchers in the same basin as the generated model. It is expected that the results of the generated model will produce the closest values to the truth. Various methods should be used to evaluate the performance of the model created for this purpose. Performance evaluation is made by comparing estimated value with actual value. More than one performance evaluation criterion appears to be used in the literature review. For this reason, a single method for evaluating performance is not considered adequate. Each of the methods used can make the comparison from a different angle. To this end, researchers have evaluated the results of their work, often using different combinations of the approaches outlined below. Researchers evaluate their work under two headings.

1. Subjective Evaluation

One of the first methods used by researchers who want to work in any field is to examine the data of the basin they want to work with various kinds of graphs. They can evaluate the methods they can use to model the basin.
Another reason for the demand of the researchers to use the graphics may be that they want to examine the real values of the basin related to the model with visual materials. In this context, nearly all of the studies reviewed show that scatter plots are used.

2. Objective Evaluation

Another method of evaluation used by researchers is objective assessment. The researchers who want to use this method are evaluating the results of the models and the results with various mathematical methods. The researchers basically compare the following two properties between the model and the real values. First the amount of error between two models and second investigation of the change of consequences together. There are many methods to evaluate the performance in the examined articles. You can access them from the Table-5. In our topic, we will briefly describe the five most commonly used methods.

2.1. Methods Used to Determine Fault Quantities

Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|Q_m^t - Q_o^t|$$

m is used to represent the model, while **o** is used for the observed value.
MAE is a value found by dividing the totals of errors by the number of errors. Producing near-zero results indicates that the generated model produces near-true results. It produces the same result in error vectors with different error values but with the same average of errors.

Mean Square Error (MSE)

$$MSE = \frac{\sum_{t=1}^{n}(Q_m^t - Q_o^t)^2}{n}$$

it is value that sum of the squares of errors divided by number of elements in vector. Producing near-zero results indicates that the generated model produces near-true results. If the error size increases, it will produce very large results. RMSE can be used instead of this feature when it is desired to evaluate with smaller numbers. It can produce the same result in two vectors with different standard deviation characteristics.

Root Meas Square Error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(Q_m^t - Q_o^t)^2}{n}}$$

it is value that root of sum of the squares of errors divided by number of elements in vector. Producing near-zero results indicates that the generated model produces near-true results.

A Simple Comparison of Three Method

Now let's consider that we have created 4 different models for a related basin. And let's examine the difference between our models and the observed values on the following table.
The table shows the error amounts of 5 models with different standard deviation values.

Table 1 - Error vectors for five models

	Error Amount For Each Observation									
Model-1	3	3	3	3	3	3	3	3	3	3
Model-2	2	2	2	2	2	4	4	4	4	4
Model-3	1	1	1	1	1	1	1	1	2	20
Model-4	5	6	5	6	5	6	5	6	5	6
Model-5	7	0	8	0	8	0	8	0	8	0

Now let's look at a table using 3 methods that we have explained the amount of errors.

Table 2 - Comparison of 5 models against different performance criteria

	Model-1	Model-2	Model-3	Model-4	Model-5
Standart Deviation	0	1.05	5.67	0.50	3.91
MAE	3	3	3	5.50	3.90
MSE	9	10	41.20	30.50	30.50
RMSE	3	3.16	6.42	5.52	5.52

2.2. Investigation of the Change of Consequences Together

There are a variety of tools that are used when it is desired to examine the direction and degree of the relationship between a dependent variable and an independent variable, or for the purpose of interchanging the generated model and the actual values and the direction of the changes. Under this heading, Correlation Coefficient (r), Determination Coefficient (R^2) and Nash-Sutcliffe Efficiency (NSE) will be explained and their results will be shown with a sample application.

Corelation Coefficient (r)

$$r=[\frac{\sum_{t=1}^{n}(Q_o^t-\bar{Q}_o)(Q_m^t-\bar{Q}_m)}{\sqrt{\sum_{t=1}^{n}(Q_o^t-\bar{Q}_o)^2(Q_m^t-\bar{Q}_m)^2}}]$$

It is a value between -1 and 1. It shows the relationship and direction between a dependent variable and an independent variable.
In the case of zero-result generation, there is no relationship between the two variables. As the produced result is closer to +1, it is understood that the strength of the similarity in the positive direction is increased. When the result is closer to -1, it is understood that the similarity increases in the negative direction.

Determination Coefficienct (R^2)

$$R^2=[\frac{\sum_{t=1}^{n}(Q_o^t-\bar{Q}_o)(Q_m^t-\bar{Q}_m)}{\sqrt{\sum_{t=1}^{n}(Q_o^t-\bar{Q}_o)^2(Q_m^t-\bar{Q}_m)^2}}]^2$$

It is a value between 0 and 1. Used to express how the dependent variable is affected by a independent variable. A value of 0.7 can be said that which are affected by a 70% independent variable of a dependent variable.

Nash-Sutcliffe Efficiency(NSE)

$$NSE = 1-\frac{\sum_{t=1}^{n}(Q_m^t-Q_o^t)^2}{\sum_{t=1}^{n}(Q_o^t-\bar{Q}_o)^2}$$

Value between +1 and -∞. As the produced result approaches +1, it is understood that the output of the generated model approximates the actual values.
Obtaining the +1 value expresses the excellence of the model. If a value between 0 and +1 is considered to produce an acceptable result, values less than 0 will show how far the model is from reflecting the actual values.

A Simple Comparison of Three Method

Let's say that we have 3 different models and that the output they produce is like Table-3.

Table 3 - Sample observations and sample outputs of three models

	Mon	Tue	Web	Thu	Fri	Sat	Sun
Observed	2	4	3	6	7	5	3
Model-1	1	3	2	5	6	4	2
Model-2	0	2	1	4	5	3	1
Model-3	-2	-4	-3	-6	-7	-5	-3

And let's examine and compare the results with our models r, R^2 and NSE.

Table 4 - Comparison of 5 models against different efficiency criteria

	Model-1	Model-2	Model-3
NSE	0,74	-0,44	-29,47
R^2	1,00	1,00	1,00
r	1,00	1,00	-1,00

When the results produced by **r** are examined, it is seen that there is a negative direction and similarity between the data of Model-3 and the realized values. When the results produced by R^2 are examined, it is seen that although there is a change in the negative direction between, it is understood that the model is strong in explaining the real values. When the results produced by NSE are examined, it can be interpreted that Model-1 produces acceptable results but the other two models are far from producing acceptable results.

COUNTRIES

Under the Stream Flow Forecasting topic, studies from Science Direct, Springer Link, and IEEE Xplore sources are the most studied countries in terms of applied countries, with 4 studies in the USA, 4 in China, 4 in Turkey, 4 in Iran, 3 in India, 3 in Australia.

CONCLUSION

Stream flow forecasting, which is the subject of our work, is a demanding process due to the complexity and non-linear structure of the factors affecting the formation of stream flow (Maheswaran & Khosa, 2012) (Yaseen and others, 2016). Researchers do their work for one of two goals. While short-term estimates are used to establish early warning systems, long-term estimates are made for effective water resources management. In this context, 32 studies between 1998 and 2017 were examined. One of the most used hydrological modeling tools is the SWAT. ANN are the most frequently used modeling tools among machine learning tools. SVM are also one of the commonly used tool. It can be seen that models derived from ANN and SVM give better results. Hydrological models used in conjunction with machine learning models are producing better results. Researchers which using machine learning methods are often used lagged days stream flow values for their model. The precipitation and evaporation parameters, however, used with the stream flow parameter at various lagged days. Researches mostly is used trial and error method or corelation analysis methods (ACF, PACF etc.) for appropriate feautere selection. This section is the least worked part of the works done. The most frequently used methods for efficiency evaluation are MSE, MAE, RMSE, r, R2 and NSE.

Table 5 - Detailed table of literature studies

Authors	Methods	Time	Features	Performance Criteria	Applied Region
(Wang, Wang, Wang, Gao, & Yu, 2017)	Bayesian MCMC (Monte Carlo and Markov Chain)	d	Flow Rates	MLE (Maximum Likelihood Estimation)	Zhujiachuan watershed in China
(Prasad, Deo, Li, & Maraseni, 2017)	IIS-Wavelet – ANN, IIS-W-M5 Tree Model	m	Precipitation, Temperature, Evaporation, Mean solar radiation and Vapor pressure	r, Willmott's Index,NSE,RMSE, MAE	Richmond, Gwydir, and Darling River in Australia
(Noori & Kalin, 2016)	SWAT + ANN	d	ANN: SWAT simulated baseflow and stormflow	NSE, BIAS	Atlanta, USA
(Uysal, Lorman, & Lensoy, 2016)	ANN(MLP) , ANN(RBF)	d	Temperature, Precipitation, Snow Data	R^2,RMSE,MAE, Nash-Sutcliffe Model Efficiency	Karasu River, Turkey
(B.Humphrey, Gibbs, Dandy, & Maier, 2016)	GR4J Model + ANN(MLP)	m	Rain, Evaporation, Ground Water, Antecedent Precipitation Index	RMSE , NSE	South East of South Australia
(Turan & Yurdusev, Fuzzy Conceptual Hydrological Model	GR2M Model + Fuzzy System (Takagi-Sugeno)	m	Precipitation, Potential Evapotranspiration,	R^2, RMSE, Nash-Sutcliffe coefficient	Gediz Basin, Turkey

for Water Flow Prediction, 2016)			Soil Moisture Accounting,...		
(Turan, Fuzzy Systems Tuned By Swarm Based Optimization Algorithms for Predicting Stream flow, 2016)	Fuzzy System (Takagi-Sugeno)	m	Stream Flow	MSE, R^2, Nash-Sutcliffe coefficient	Susurluk Basin, Turkey
(Singh S. K., 2016)	TopNet + Hindcast Method	s	TOPNET: DEM, Soil moisture, Snow water equivalent, Depth to ground water, Land cover	RMSE, Ranked Probability Score, Ranked Probability Skill Score	South Island, New Zealand
(Singh A. , 2016)	SWAT , RBNN(ANN)	m	RBNN : Stream Flow, Temperature, Rainfall	R^2, NSE	Nagwa Sub-Watershed In India
(Keshtegar, Allawi, Afan, & El-Shafie, 2016)	HORS	m	Stream Flow	MAE, Normalized RMSE, RMSE, MSE, R^2, Relative Error	Aswan High Dam, Egypt
(Yaseen and others, 2016)	ELM , SVR , Generalized Regression Neural Network	m	Stream Flow	r,NSE, Willmott's Index, RMSE, MAE	Tigris River, Iraq
(Londhe & Gavraskar, 2015)	SVR	d	Stream Flow + Rainfall	r, RMSE	Krishna, Narmada River Basin, India
(Zhao and others, 2015)	Bayesian Joint Probability	d	Stream Flow	RMSE, Continuous Ranked Probability	Three Gorges Reservoir, China
(Vilaysane, Takara, Luo, Akkharath, & Duan, 2015)	SWAT	m	DEM, Land Use, Soil Types, Hydrlogical and meteorological data	NSE, R^2	Xedone River Basin, Laos
(Dehghani, Saghafian, Rivaz, & Khodadadi, 2015)	Dynamic Linear Spatio Temporal Model	m	Stream Flow	Discrepancy ratio	The Great Karun Basin, Iran
(Sun, Wang, & Xu, 2014)	GPR , ARMAX , ANN(MLP)	m	Stream Flow, Precipitation, Temperature	NSE, Water Balance Error	Basins Across The U.S.
(Toro, Meire, F.Gálvez, & Riverola, 2013)	Statistic + ANN	d	River Flow, Rainfall	MAE, Mean Absolute Percentage Error, MSE	Salvajina Reservoir, Colombia
(Sattari, Pal, Apaydin, & Ozturk, 2013)	M5 Tree Model , SVM	d	Stream Flow, Rainfall, Temperature	r, RMSE	Sohu Creek, Ankara, Turkey
(Rathinasamy, Adamowski, & Khosa, 2013)	Bayesian Model Averaging Based Ensemble Multi Wavelet Volterra Nonlinear Model	d, w, m	River Flow	RMSE, MAE, r, NSC	Selway River, ST Joe River, USA
(Kalteh, 2013)	Wavelet-ANN , Wavelet -SVR	m	River Flow	RMSE, MAE, CC	Kharjegil and Ponel Stationsin

					Northern Iran
(Robertson, Pokhrel, & Wang, 2013)	WAPABA	s		RMSEP, Continuous Ranked Probability	Catchments in Eastern Australia
(Maheswaran & Khosa, 2012)	ANN, Wavelet-ANN, Wavelet Linear Regression, Coupled Wavelet–Volterra	m	Stream Flow, Rainfall	Nash–Sutcliffe model efficiency coefficient, RMSE, MAE, r	Cauvery River Basin, India
(Noori and others, 2011)	ANN - SVM	m	Rainfall, Discharge, Sun Radiation, Temperature	R^2, MAE, RMSE	Sofichay River, Iran
(Demirel, Venancio, & Kahya, 2009)	ANN , SWAT	d	ANN : Stream Flow, Precipation	MSE, RMSE	Pracana Basin, Portugal
(Asefa, Kemblowski, McKee, & Khalil, 2006)	SVM, Transfer Function Noise Model	s	Stream Flow, Snow Water Equivalent, Temperature	R^2, MSE, RMSE	Sevier River Basin, USA
(Baratti and others, 2003)	ANN(MLP), ARMAX	m, d	Runoff, Rainfall, Temperature	r, RMSE, Mean Squared Relative Error, MAE,E, Mean 4th Order Error	Tirso Catchment, Italy
(Adenan, Hamid, Mohamed, & Noorani, 2017)	Local Linear Approximation Method	d	River Flow	r	Klang River, Malaysia
(Valença & Ludermir, 2000)	FNN , ARIMA	m	Inflow	Absolute Average Percentual Error	Sobradinho Hydroelectric Power Plant, Brazil
(Ballini, Soares, & Andrade, 1998)	ANFN	s	Inflow	MSE, Mean Absolute Deviation, MRE	Hydroelectric Plants, Brazilian
(Khadangi, Madvar, & Mehdi, 2009)	ANN(RBF), Adaptive Neuro Fuzzy Inference	d	Stream Flow	R^2, MAPE, Relative RMSE	Mahabad River, Iran
(XU, ZHU, ZHANG, XU, & XIAN, 2009)	ANN, XXT	d	ANN:Stream Flow, Precipitation, Evaporation	NSE	Yingluoxia Basin, China
(XU, ZHANG, & ZHAO, Stream Flow Forecasting by Artificial Neural Network and TOPMODEL in Baohe River Basin, 2009)	ANN , TOPMODEL	d	ANN: Stream Flow, Precipitation	Nash And Sutcliffe Efficiency Coefficient	Baohe River Basin, China

D = Daily, W = Weekly, M = Monthly, S = Seasonal

REFERENCES

(WWAP), U. N. (2016). *WATER AND JOBS.* Paris: UNESCO.

(WWAP), U. N. (2017). *Wastewater: The Untapped Resource.* Paris: UNESCO.

Adenan, N. H., Hamid, N. Z., Mohamed, Z., & Noorani, M. S. (2017). A pilot study of river flow prediction in urban area based on phase space reconstruction. *AIP Conference Proceedings, 1870*(1).

Asefa, T., Kemblowski, M., McKee, M., & Khalil, A. (2006, March 1). Multi-time scale stream flow predictions: The support vector machines approach. *Journal of Hydrology, 318*(1-4), 7-16.

B. Humphrey, G., Gibbs, M. S., Dandy, G. C., & Maier, H. R. (2016, September). A hybrid approach to monthly streamflow forecasting: Integrating hydrological model outputs into a Bayesian artificial neural network. *Journal of Hydrology, 540*, 623-640.

Ballini, R., Soares, S., & Andrade, M. G. (1998). An adaptive neural fuzzy network model for seasonal stream flow forecasting. *Neural Networks, 1998. Proceedings. Vth Brazilian Symposium on. IEEE* (s. 215-219). Belo Horizonte, Brazil, Brazil: IEEE.

Baratti, R., Cannas, B., Fanni, A., Pintus, M., Sechi, G. M., & Toreno, N. (2003, October). River flow forecast for reservoir management through neural networks. *Neurocomputing, 55*(3-4), 421-437.

Dehghani, M., Saghafian, B., Rivaz, F., & Khodadadi, A. (2015). Monthly stream flow forecasting via dynamic spatio-temporal models. *Stochastic Environmental Research and Risk Assessment, 29*(3), 861-874.

Demirel, M. C., Venancio, A., & Kahya, E. (2009, July). Flow forecast by SWAT model and ANN in Pracana basin, Portugal. *Advances in Engineering Software, 40*(7), 467-473.

Devi, G. K., P, G. B., & S, D. G. (2015). A Review on Hydrological Models. *Aquatic Procedia, 4*, 1001-1007.

Falkenmark, M. (2017). Water and human livelihood resilience: a regional-to-global. *International Journal of Water Resources Development, 33*(2), 181–197.

Kalteh, A. M. (2013, April). Monthly river flow forecasting using artificial neural network and support vector regression models coupled with wavelet transform. *Computers & Geosciences, 54*, 1-8.

Kansal, M. L. (2005). *System Analysis Techniques In Water Resources Management.* Su Vakfı Yayınları.

Kauffeldt, A., Wetterhall, F., Pappenberger, F., Salamon, P., & Thielen, J. (2016, January). Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environmental Modelling & Software, 75*, 68-76.

Keshtegar, B., Allawi, M. F., Afan, H. A., & El-Shafie, A. (2016, September). Optimized River Stream-Flow Forecasting Model Utilizing High-Order Response Surface Method. *Water Resources Management, 30*(11), 3899-3914.

Khadangi, E., Madvar, H. R., & Mehdi, M. (2009). Comparison of ANFIS and RBF Models in Daily Stream flow Forecasting. *Computer, Control and Communication, 2009. IC4 2009. 2nd International Conference on. IEEE* (s. 1-6). Karachi, Pakistan: IEEE.

Londhe, S., & Gavraskar, S. S. (2015). Forecasting One Day Ahead Stream Flow Using Support Vector Regression. *Aquatic Procedia, 4*, 900-907.

Maheswaran, R., & Khosa, R. (2012, July 11). Wavelet–Volterra coupled model for monthly stream flow forecasting. *Journal of Hydrology, 450-451*, 320-335.

Noori, N., & Kalin, L. (2016, February). Coupling SWAT and ANN models for enhanced daily streamflow prediction. *Journal of Hydrology, 533*, 141-151.

Noori, R., Karbassi, A. R., MOghaddamnia, A., Han, D., Zokaei-Ashtiani, M., Farokhnia, A., & Housheh, M. G. (2011, May 3). Assessment of input variables determination on the SVM model performance using PCA, Gamma test, and forward selection techniques for monthly stream flow prediction. *Journal of Hydrology, 401*(3-4), 177-189.

P. Krause, D. P. (2005). Comparison of different efficiency criteria for hydrological model. *Advances in Geosciences, 5*, 89-97. 09 01, 2017

Prasad, R., Deo, R. C., Li, Y., & Maraseni, T. (2017, November). Input selection and performance optimization of ANN-based streamflow forecasts in the drought-prone Murray Darling Basin region using IIS and MODWT algorithm. *Atmospheric Research, 197*, 42-63.

Rathinasamy, M., Adamowski, J., & Khosa, R. (2013, December 12). Multiscale streamflow forecasting using a new Bayesian Model Average based ensemble multi-wavelet Volterra nonlinear method. *Journal of Hydrology, 507*, 186-200.

Robertson, D. E., Pokhrel, P., & Wang, Q. J. (2013). Improving statistical forecasts of seasonal streamflows using hydrological model output. *Hydrology and Earth System Sciences, 17*, 579-593.

Rockström, J., Falkenmark, M., Allan, T., Folke, C., Gordon, L., Jägerskog, A., . . . Varis, O. (2014, October). The unfolding water drama in the Anthropocene: towards a resilience-based perspective on water for global sustainability. *ECOHYDROLOGY, 7*(5), 1249–1261.

Sattari, M. T., Pal, M., Apaydin, H., & Ozturk, F. (2013, May). M5 model tree application in daily river flow forecasting in Sohu Stream, Turkey. *Water Resources, 40*(3), 233-242.

Singh, A. (2016, June). Modeling Stream Flow with Prediction Uncertainty by Using SWAT Hydrologic and RBNN Models for an Agricultural Watershed in India. *National Academy Science Letters, 39*(3), 213-216.

Singh, S. K. (2016, May). Long-term Streamflow Forecasting Based on Ensemble Streamflow Prediction Technique: A Case Study in New Zealand. *Water Resources Management, 30*(7), 2295-2309.

Sun, A. Y., Wang, D., & Xu, X. (2014, April 16). Monthly streamflow forecasting using Gaussian Process Regression. *Journal of Hydrology, 511*, 72-81.

Toro, C. H., Meire, S. G., F.Gálvez, J., & Riverola, F. F. (2013, August). A hybrid artificial intelligence model for river flow forecasting. *Applied Soft Computing, 13*(8), 3449-3458.

Turan, M. E. (2016, September). Fuzzy Systems Tuned By Swarm Based Optimization Algorithms for Predicting Stream flow. *Water Resources Management, 30*(12), 4345-4362.

Turan, M. E., & Yurdusev, M. A. (2016, January). Fuzzy Conceptual Hydrological Model for Water Flow Prediction. *Water Resources Management, 30*(2), 653-667.

Uysal, G., Lorman, A. A., & Lensoy, A. (2016). Streamflow Forecasting Using Different Neural Network Models with Satellite Data for a Snow Dominated Region in Turkey. *Procedia Engineering, 154*, 1185-1192.

Valença, M., & Ludermir, T. (2000). Monthly streamflow forecasting using an Neural Fuzzy Network Model. 2. P. Neural Networks (Dü.). içinde (s. 117-119). Rio de Janeiro, RJ, Brazil, Brazil: IEEE.

Vilaysane, B., Takara, K., Luo, P., Akkharath, I., & Duan, W. (2015). Hydrological Stream Flow Modelling for Calibration and Uncertainty Analysis Using SWAT Model in the Xedone River Basin, Lao PDR. *Procedia Environmental Sciences, 28*, 380-390.

Wang, H., Wang, C., Wang, Y., Gao, X., & Yu, C. (2017). Bayesian forecasting and uncertainty quantifying of stream flows using Metropolis–Hastings Markov Chain Monte Carlo algorithm. *Journal Of Hydrology, 549*, 476-483.

Xu, J., Zhang, W., & Zhao, J. (2009). Stream Flow Forecasting by Artificial Neural Network and TOPMODEL in Baohe River Basin. *2009 Third International Symposium on Intelligent Information Technology Application Workshops.* Nanchang, China: IEEE.

Xu, J., Zhu, X., Zhang, W., Xu, X., & Xian, J. (2009). Daily streamflow forecasting by artificial neural network in A. *Information, Computing and Telecommunication, 2009. YC-ICT'09. IEEE Youth Conference on. IEEE* (s. 487-490). Beijing, China: IEEE Xplore.

Yaseen, Z. M., Jaafar, O., Deo, R. C., Kisi, O., Adamowski, J., Quilty, J., & El-Shafie, A. (2016, November). Stream-flow forecasting using extreme learning machines: A case study in a semi-arid region in Iraq. *Journal of Hydrology, 542*, 603-614.

Zhao, T., Wang, Q. J., Bennett, J. C., Robertson, D. E., Shao, Q., & Zhao, J. (2015, September). Quantifying predictive uncertainty of streamflow forecasts based on a Bayesian joint probability model. *Journal of Hydrology, 528*, 329-340.