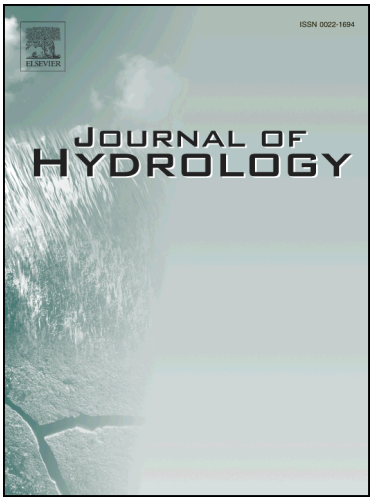
Accepted Manuscript



Review papers

A review of the artificial intelligence methods in groundwater level modeling

Taher Rajaee, Hadi Ebrahimi, Vahid Nourani

PII: S0022-1694(19)30006-X
DOI: https://doi.org/10.1016/j.jhydrol.2018.12.037
Reference: HYDROL 23358

To appear in: *Journal of Hydrology*

Received Date: 19 May 2018
Revised Date: 7 December 2018
Accepted Date: 8 December 2018

Please cite this article as: Rajaee, T., Ebrahimi, H., Nourani, V., A review of the artificial intelligence methods in groundwater level modeling, *Journal of Hydrology* (2019), doi: https://doi.org/10.1016/j.jhydrol.2018.12.037

A review of the artificial intelligence methods in groundwater level modeling

Taher Rajaee [1] Hadi Ebrahimi [2] Vahid Nourani [3]

[1] Associate Professor, Department of Civil Engineering, University of Qom, Qom- Iran (Corresponding Author)
Email: taher_rajaee@yahoo.com; trajaee@qom.ac.ir
Tel: +989126275800

[2] PhD. Student in Civil Engineering, Department of Civil Engineering, University of Qom, Qom- Iran
Email: hadi_e1360@yahoo.com
Tel: +989125526846

[3] Professor, Department of Water Resources Engineering, Faculty of Civil Engineering, University of Tabriz, Tabriz- Iran
Email: nourani@tabrizu.ac.ir
Tel: +989144030332

Abstract:

This study is a review to the special issue on artificial intelligence (AI) methods for groundwater level (GWL) modeling and forecasting, and presents a brief overview of the most popular AI techniques, along with the bibliographic reviews of the experiences of the authors over past years, and the reviewing and comparison of the obtained results. Accordingly, 67 journal papers published from 2001 to 2018 were reviewed in the terms of the features and abilities of the modeling approaches, input data consideration, prediction time steps, data division, etc. From the reviewed papers it can be concluded that despite some weaknesses, if the AI methods properly be developed, they can successfully be used to simulate and forecast the GWL time series in different aquifers. Since some of the stages of the AI modeling are based on the experience or trial-and-error procedures, it is useful to review them in the special application on GWL modeling. Many partial and general results were achieved from the reviewed papers, which can provide applicable guidelines for researchers who want to perform similar works in this field. Several new ideas in the related area of research are also presented in this study for developing innovative methods and for improving the quality of the modeling.

Keywords: Groundwater level forecasting, Artificial intelligence, Neural networks, Review, Wavelet

Contents

1. Introduction

Measurement and analysis of the groundwater level (GWL) in aquifers is an important and useful task in the management of the groundwater resources, and the knowledge about the GWL variations can be used for quantifying the groundwater availability. The GWL variations in wells provide a direct measure of the impact of groundwater development, and important information about aquifer dynamics is often embedded in the continuously recorded GWL time series (Butler et al. 2013). Therefore, the modeling and predicting of GWL is necessary for water managers and engineers to qualify and quantify groundwater resources and to maintain a balance between supply and demands.

For GWL modeling, conceptual or physical based models are traditionally the main tool; however they have some practical limitations, including the need for large amount of data and input parameters. In many cases, data is limited on one hand, and obtaining accurate predictions is more important than understanding underlying mechanisms, on the other hand, and therefore, the black-box artificial intelligence (AI) models can be a suitable alternative. Although there are different methods for modeling and predicting GWL in aquifers such as conceptual, physical, numerical, statistical, etc., however in recent years, AI methods have been used for their simplicity and acceptable results, and many researches have investigated the performance of AI models for GWL modeling in different parts of the world. This study is a review of those papers that have used AI methods for modeling and forecasting GWL. Of course, these methods have some weaknesses, such as overtraining, low generalizability, risk of using unrelated data, incorrect modeling with inappropriate methods, and so on. However, their simplicity use, high speed run and acceptable accuracy without the need to know the problems physics have led many researchers to apply them. It should be noted that it is the nature or perhaps the defect of the AI models that if they were developed for the prediction of a specified time series, the accurate results could not necessarily be derived in the similar ones; but the major advantage of AI models is the nonlinear and complicated phenomena modeling without the need for full understanding underlying mechanisms (Rajaee and Boroumand, 2015). Therefore, the use of AI approaches in GWL modeling has steadily increased and attracted interest of many researchers in the world.

In order to develop new and better AI approaches for GWL modeling, it is important to investigate what has been done with AI models and current researches, and there is a need for researchers to know what other scholars have done in this regard. Many review papers have been recently published that have explored using AI models in hydrology (e.g., Solomatine (2005)), or in different hydrological and water resources fields (e.g., Maier et al. (2010) in the field of river variables modeling, and Wu et al. (2014) in the field of water quality modeling), while, no review paper is find that has centered on the specific use of AI models for GWL modeling and forecasting. Each hydrological phenomenon has its own characteristics, and it is reasonable that the use of AI models in GWL modeling to be reviewed individually. Nourani et al. (2014) have cited and reviewed some wavelet-AI studies in GWL modeling (5 papers); however, in the best knowledge of the authors, there is not yet an individual and comprehensive review paper evaluating the application of AI methods in GWL modeling and forcasting.

The current review study presents and compares the details of the journal papers dealing with the AI methods for GWL modeling and forecasting, in the terms of the features and abilities of the modeling approaches, the input data consideration, the quantity and quality of used data, the study areas and aquifers, the prediction time steps, the data division, etc. 67 papers are reviewed in this study. These papers have been published in the international journals belongs to the famous publications such as Elsevier, Springer, IWA, Wiley, ASCE, etc. during the period from 2001 to 2018. The papers were found from searching the web using the relevant key words, and were chosen because they were published in well-known international journals in the fields of hydrology, water resources and AI sciences. Based on the search, Journal of Hydrology (Elsevier) with 12 papers and Water Resources Management (Springer) with 11 papers are the journals that have been published the most papers in this regard. Also, Hydrological Processes (Wiley), Journal of Hydroinformatics (IWA), Hydrogeology Journal (Springer) and Computers & Geosciences (Elsevier), each one have been published three papers in this regard. The rest of the journals (a total of 29 journals) that had papers in this regard have been published one or two papers so far (Table 1).

Figure 1 shows number of published papers regarding AI in GWL modeling (reviewed in this study) with respect to year of publication. As can be seen, such publications have increased in recent

years. Therefore, due to the interest of researchers in this field and given the difficulty of conceptual/numerical GWL modeling, this review was provided to help new researches in this field.

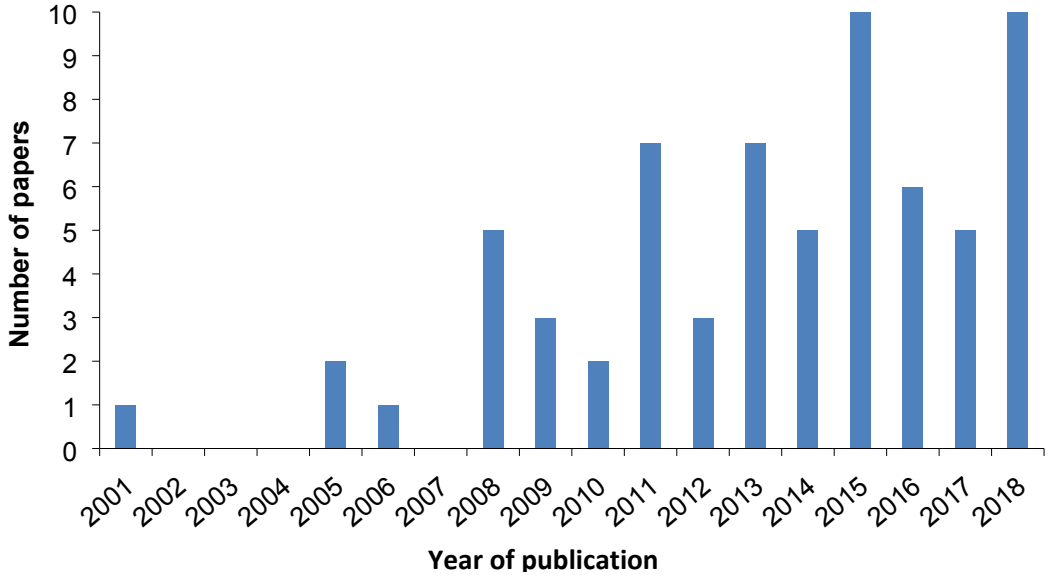


Figure 1. Number of published papers regarding AI methods in GWL modeling (used in this study) with respect to year of publication.

Details of the selected papers are given in Table 1, where the papers on the subject of GWL modeling with AI methods are comparing regarding to the authors and year of publication, journals and impact factor (IF), region of study, type of utilized AI methods, hydrological input variables, time steps and range of total data. The abbreviations used in the Table 1 have been explained in the end of the table.

In the following, some very commonly used AI methods for modeling GWL are addressed. The methods include artificial neural networks (ANN), adaptive neuro-fuzzy inference system (ANFIS), genetic programming (GP), support vector machine (SVM) and some hybrid models such as wavelet-AI models. Firstly, a brief description of each method is presented and thereafter the related conducted studies are cited and reviewed. This is followed by general results and discussion, conclusions and recommendations for future avenues of research.

Table 1.

Details of the reviewed papers, where the AI methods were used to model the GWL.

No.	Author (year)	Journal (2016 IF)	Region of study	Used AI models	Input variables	Time step	Range of total data (Number of data)
1	Coulibaly et al. (2001)	Water Resources Research (4.397)	Gondo aquifer, Burkina Faso	ANN	GWL, P, T	Monthly average	1986-1996 (108 sets)
2	Lallahem et al (2005)	Journal of Hydrology (3.483)	Chalky aquifer of northern France	ANN	GWL, R, mean T, effective R, potential ET	Monthly	1988-1999 (132 sets)
3	Daliakopoulos et al. (2005)	Journal of Hydrology (3.483)	Island of Crete, Greece	ANN	GWL, T, P, Q	Monthly	1988-2002 (160 sets)
4	Nayak et al. (2006)	Water Resources Management (2.848)	Central Godavari Delta System, India	ANN	GWL, neighboring wells GWL, R, canal releases	Monthly average	1981-1989 (108 sets)
5	Krishna et al. (2008)	Hydrological Processes (3.014)	Andhra Pradesh state, India	ANN	GWL, R, ET	Monthly average	1995-2004 (120 sets)
6	Mohammadi (2008)	Practical Hydroinformatics (Book)	Chamchamal plain, Iran	ANN	MODFLOW output parameters	Monthly	1986-1998 (144 sets)
7	Feng et al (2008)	Groundwater (2.067)	Shiyang river basin, northwest China	ANN	GWL, P, E, Q, population, irrigation ratio, irrigation area	Monthly	1980-1997 (216 sets)
8	Tsanis et al. (2008)	Journal of Hydroinformatics (1.634)	Messara Valley, Crete, Greece	ANN	P, T, runoff, GWL, specific yield	Monthly	1981-2002 (264 sets)
9	Nourani et al. (2008)	Hydrological Processes (3.014)	Tabriz aquifer, Iran	ANN	GWL, R, mean T, Q	Monthly	1995-2004 (120 sets)
10	Kholghi and hosseini (2009)	Environmental Modeling & Assessment (1.023)	Qazvin plain, Iran	ANFIS, Kriging	GWL	Spatial modeling	Spatial modeling
11	Banerjee et al. (2009)	Environmental geology (no IF)	Hyderabad, India	ANN	Not mentioned in the paper	Monthly	2005-2007 (23 sets)
12	Yang et al. (2009)	Journal of Arid Environments (1.835)	Western Jilin, China	ANN	GWL	Monthly average	1986-2004 (132 sets)
13	Mohanty et al. (2010)	Water Resources Management (2.848)	Orissa, India	ANN	GWL, R, E, River stage, SWL, Pumping rate	Weekly	2004-2007 (174 sets)

No.	Author (year)	Journal (2016 IF)	Region of study	Used AI models	Input variables	Time step	Range of total data (Number of data)
14	Chen et al. (2010)	Journal of Hydrologic Engineering (1.694)	Southern Taiwan	ANN	GWL, neighboring wells GWL	Monthly average	1997-2003 (63 sets)
15	Chen et al. (2011)	Journal of Water Resources Planning and Management (3.537)	Southern Taiwan	ANN	GWL, neighboring wells GWL	Monthly average	1998-2004 (76 sets)
16	Jalalkamali et al. (2011)	Journal of Hydroinformatics (1.634)	Kerman, Iran	ANFIS, ANN	GWL, T, R	Monthly	1988-2009 (264 sets)
17	Adamowski and Chan (2011)	Journal of Hydrology (3.483)	Quebec, Canada	WANN, ANN	GWL, P, T	Monthly average	2002-2009 (84 sets)
18	Yoon et al. (2011)	Journal of Hydrology (3.483)	Beach of the Donghae city, Korea	SVM, ANN	GWL, P, tide level	Six-hourly	2004-2006 (2370 sets)
19	Sreekanth et al. (2011)	Environmental Earth Science (1.569)	Maheshwaram watershed, India	ANN, ANFIS	GWL, R, E, T, H	Monthly	2000-2006 (84 sets)
20	Nourani et al. (2011)	Environmental Engineering Science (1.426)	Shabestar plain, Iran	ANN-GS	GWL, R, lake level	Monthly	1994-2006 (144 sets)
21	Trichakis et al, (2011)	Water Resources Management (2.848)	Edward's aquifer, Texas, USA	ANN	GWL, P, day number, pumping	Daily	Not mentioned in the paper
22	Rakhshandehroo et al. (2012)	Arabian Journal for Science and Engineering (0.865)	Shiraz plain, Iran	ANN	GWL, P, T, E, Q	Monthly	1993-2004 (138 sets)
23	Taormina et al. (2012)	Engineering Applications of Artificial Intelligence (2.894)	Lagoon of Venice, Italy	ANN	GWL, R, ET	Hourly	2005-2008 (23850 sets)
24	Kisi and Shiri (2012)	Hydrology Research (1.754)	Illinois State, USA	Wavelet-ANFIS, ANFIS	GWL	Daily	2001-2008 (2430 sets)
25	Shirmohammadi et al. (2013)	Water Resources Management (2.848)	Mashhad plain, Iran	ANFIS	P	Monthly	1992-2007 (180 sets)
26	Sahoo and Jha (2013)	Hydrogeology Journal (2.109)	Konan basin, Kochi, Japan	ANN	GWL, R, T, river stage, seasonal dummy variables	Monthly	1999-2004 (72 sets)
27	Shiri et al. (2013)	Computers & Geosciences (2.533)	Hoengchon, south Korea	GP, ANN, ANFIS, SVM	GWL, R, ET	Daily average	2001-2008 (2920 sets)

No.	Author (year)	Journal (2016 IF)	Region of study	Used AI models	Input variables	Time step	Range of total data (Number of data)
28	Fallah-Mehdipour et al. (2013)	Journal of Hydro-environment Research (1.429)	Karaj plain, Iran	GP, ANFIS	GWL, P, E	Monthly	7 years (84 sets)
29	Maheswaran and Khosa (2013)	Computers & Geosciences (2.533)	Northern Saanich Peninsula, Canada	Wavelet-Volterra, Wavelet-ANN, ANN	GWL	Monthly average	1975-2002 (324 sets)
30	Moosavi et al. (2013a)	Water Resources Management (2.848)	Mashhad plain, Iran	Wavelet-ANFIS, Wavelet-ANN, ANFIS, ANN	GWL, P, E, average Q	Monthly average	1992-2007 (180 sets)
31	Moosavi et al. (2013b)	Arabian Journal for Science and Engineering (0.865)	Mashhad plain, Iran	Wavelet-ANFIS, Wavelet-ANN	GWL, P, E, average Q	Monthly average	1992-2007 (180 sets)
32	Emamgholizadeh et al. (2014)	Water Resources Management (2.848)	Bastam plain, Iran	ANFIS, ANN	R recharge, irrigation returned flow, pumping rates	Monthly	2002-2011 (108 sets)
33	Suryanarayana et al. (2014)	Neurocomputing (3.317)	Visakhapatnam, India	Wavelet-SVR, SVR, ANN	GWL, P, max T, mean T	Monthly	2001-2012 (129 sets)
34	Tapoglou et al. (2014)	Journal of Hydrology (3.483)	Bavaria, Germany	ANN-NeuroFuzzy- GS	GWL, SWL, T, R	Daily	2008-2012 (1460 sets)
35	He et al. (2014)	Water Resources Management (2.848)	Ganzhou region, China	Wavelet-ANN, ANN	GWL	Monthly	1995-2004 (120 sets)
36	Ying et al. (2014)	Journal of Water Supply Research and Technology-Aqua (0.824)	Jilin Province, China	ANN	GWL	Monthly	1986-2013 (about 336 sets)
37	Jha and Sahoo (2015)	Hydrological Processes (3.014)	Konan basin, Kochi, Japan	ANN-GA	GWL, R, T, river stage, seasonal dummy variables	Monthly	1999-2004 (72 sets)
38	Yang et al. (2015)	Arabian Journal of Geosciences (0.955)	Dongshan Island, Fujian, China	Wavelet-ANN, ANN	GWL	Monthly average	2000-2011 (144 sets)
39	Khalil et al. (2015)	Hydrogeology Journal (2.109)	Manitou mine site, Quebec, Canada	Wavelet-ensemble-ANN, Wavelet-ANN, ANN	Recharge, P, T	Daily	2009-2011 (900 sets)
40	Mirzavand et al. (2015)	Natural Hazards (1.833)	Kashan plain, Iran	ANFIS, SVR	R, E, Q, Aquifer discharge	Monthly	1990-2010 (240 sets)
41	Juan et al. (2015)	Journal of Hydrology (3.483)	Qinghai-Tibet Plateau, China	ANN	GWL, T, P	Daily	2010-2012 (653 sets)

No.	Author (year)	Journal (2016 IF)	Region of study	Used AI models	Input variables	Time step	Range of total data (Number of data)
42	Gholami et al. (2015)	Journal of Hydrology (3.483)	Caspian southern coasts, Iran	ANN	P, tree-rings	Annually	1970-2013 (44 sets)
43	Khaki et al. (2015)	Environmental Earth Sciences (1.569)	Langat basin, Malaysia	ANFIS, ANN	R, H, E, min and max T	Monthly average	2007-2013 (79 sets)
44	Nourani et al. (2015)	Journal of Hydrology (3.483)	Ardabil plain, Iran	Wavelet-ANN, ANN	GWL, R, runoff	Monthly	1988-2012 (300 sets)
45	Mohanty et al. (2015)	Water Resources Management (2.848)	Mahanadi delta of Odisha, India	ANN	GWL, R, E, river stage, SWL, pumping rates	Weekly	2004-2007 (174 sets)
46	Gong et al. (2015)	Water Resources Management (2.848)	shore of Lake Okeechobee, Florida, USA	ANFIS, SVM, ANN	GWL,SWL, P, T	Monthly	1998-2009 (144 sets)
47	Sun et al. (2016)	Hydrology and Earth System Sciences (4.437)	Nee Soon swamp forest , Singapore	ANN	SWL, P	Daily	2012-2013 (730 sets)
48	Chang et al. (2016)	Journal of Hydrology (3.483)	Zhuoshui River basin, Taiwan	ANN (SOM-NARX)	GWL, Q, R	Monthly average	2000-2013 (168 sets)
49	Han et al. (2016)	Journal of Environmental Management (4.010)	western Hexi Corridor, northwest China	ANN (SOM-statistical model)	GWL, climate conditions, well extractions, Q, reservoir operations	Monthly	1998-2010 (156 sets)
50	Hosseini et al. (2016)	Arabian Journal of Geoscience (0.955)	Shabestar plain, Iran	ANN-Ant colony	GWL, R, E, Q, T, annual time series	Monthly	2000-2009 (108 sets)
51	Nourani and Mousavi (2016)	Journal of Hydrology (3.483)	Miandoab plain, Iran	Wavelet-ANFIS, Wavelet-ANN	GWL, P, Q	Monthly	2001-2011 (132 sets)
52	Yoon et al. (2016)	Computers & Geosciences (2.533)	South Korea	SVM, ANN	GWL, P	Daily	2003-2008 (About 2000 sets)
53	Ebrahimi and Rajaee (2017)	Global and Planetary Change (3.915)	Qom plain, Iran	Wavelet-ANN, Wavelet-SVR, ANN, SVR	GWL	Monthly	2002-2013 (132 sets)
54	Barzegar et al. (2017)	Science of the Total Environment (4.900)	Azarbaijan, Iran	Wavelet-ANN	GWL	Monthly	1985-2016 (312 sets)
55	Nie et al. (2017)	Journal of Water Supply Research and Technology-Aqua (0.824)	Jilin Province, China	SVM, ANN	P, E, T	Monthly	2003-2014 (144 sets)

No.	Author (year)	Journal (2016 IF)	Region of study	Used AI models	Input variables	Time step	Range of total data (Number of data)
56	Huang et al. (2017)	Journal of Hydroinformatics (1.634)	Three Gorges Reservoir Area, China	SVM, ANN	GWL	Daily, Weekly, Monthly	Daily: 2013-2014 Weekly & monthly: 2007-2010
57	Wen et al. (2017)	Hydrology Research (1.754)	Northwestern China	Wavelet-ANN, ANN	GWL, P, E, T	Monthly	2003-2010 (91 sets)
58	Yu et al. (2018)	Water Resources Management (2.848)	Northwest China	Wavelet-ANN, Wavelet-SVR	GWL, ET, Q	Monthly	2000-2010 (132 sets)
59	Wunsch et al. (2018)	Journal of Hydrology (3.483)	Southwest Germany	ANN	P, T	Weekly	1948-2008 (3081 sets)
60	Zare and Koch (2018)	Journal of Hydro-environment Research (1.429)	Kermanshah, Iran	Wavelet-ANFIS	GWL, P	Monthly	1991-2013 (261 sets)
61	Mukherjee and Ramachandran (2018)	Journal of Hydrology (3.483)	India	SVR, ANN	GRACE satellite data, P, min and max T, H, wind	Multi-monthly	2005-2013 (35 to 67 sets)
62	Ghose et al. (2018)	Groundwater for Sustainable Development (no IF)	Odisha, India	ANN	P, T, H, ET, Q	Monthly	1988-2007 (about 170 sets)
63	Lee et al. (2018)	Hydrogeology Journal (2.109)	South Korea	ANN	SWL, pumping rates	Hourly	2016-2017 (8712 sets)
64	Rakhshandehroo et al. (2018)	Journal of Hydrologic Engineering (1.694)	Florida and Arkansas, USA	Wavelet-ANN, ANN	GWL	Daily	2002-2007 (about 1800 sets)
65	Guzman et al. (2018)	Environmental Modeling & Assessment (1.023)	Mississippi delta region, USA	SVR, ANN	GWL, P, ET	Daily	1985-1994 (about 3400 sets)
66	Tang et al. (2018)	Geotechnical and Geological Engineering (no IF)	Northern United Kingdom	SVM, ANN, random forest, k-nearest neighbor	GWL	Hourly	2016-2017 (likely about 8000 sets)
67	Kouziokas et al. (2018)	Water Resources Management (2.848)	Pennsylvania, USA	ANN	P, T, H	Daily	2014 (365 sets)

Abbreviations: P, precipitation; T, temperature; R, rainfall; H, humidity; E, evaporation; ET, evapotranspiration; Q, river flow/discharge/runoff; SWL, surface water level; GS, geostatistics.

2. Artificial intelligence Methods for GWL modeling

2.1. Artificial Neural networks (ANN) for GWL modeling

2.1.1. Introductory

ANNs are computational models inspired by biological neural networks. They can be used to approximate functions that are generally unknown, or to predict future values of possibly noisy time series based on past histories. ANNs are composed of simple elements operating in parallel. As in nature, the connections between elements largely determine the network function (Beale et al., 2010). A common ANN comprised of multiple elements, called neurons (processing elements), and connection pathways that link them. The neurons having similar properties are grouped in one single layer. Typically, three separate layers exist in an ANN, namely input, hidden and output layers. The input layer takes input variables, which in the case of GWL forecasting are usually the precipitation, temperature, GWL, etc. time series. In the hidden and output layers, each neuron passes its weighted and biased input through a desired transfer (activation) function to produce a result. ANNs are trained with a sample data, so that a particular input leads to a specific target output. Training means tuning the adjustable network parameters (called weights and biases) to optimize the network performance. The training process can be done with various training (learning) algorithms. The Levenberg-Marquardt (LM) algorithm, the back-propagation (BP) algorithm, the Bayesian regularization (BR) algorithm and the gradient descent with momentum and adaptive learning rate back-propagation (GDX) algorithm are examples of most used training algorithms in the literature.

Different ANN types have been widely described in the literature; however several types of them are briefly presented here. The feed-forward neural networks (FNNs) propagate input signal through the network in a forward direction, layer by layer. The multilayer perceptron (MLP) network as a historical FNN consists of an input layer, one or more hidden layers, and one output layer. The recurrent neural networks (RNN) feed the outputs of the hidden layer back to itself. In the RNNs, an additional layer is interconnected with the hidden layer that plays the role of the network history. The radial basis function (RBF) networks are also feed-forward, but have only one hidden layer that uses

Gaussian transfer function and a standard Euclidean distance to measure how far an input vector is from a specific center vector. The amount of Euclidean distance is transferred by the Gaussian function that determines the output of the layer. RBF networks tend to learn much faster than a FNN.

The self-organizing map (SOM) network as a kind of ANNs consists of one input layer and one output layer called 'Kohonen' layer. The input layer is fully connected to the output layer. The SOM is trained using an unsupervised competitive training algorithm. The n-dimensional input vector is sent through the network, and the Euclidean distance between the weight vector and the input vector is computed. The training process will be continued to select best neurons that reduce the distance between the weights and inputs. An advantage of the SOM is to map high-dimensional input space into low dimensional space.

Regardless of the type of utilized ANN, they have some common modeling stages. Figure 2 shows the typical stages of using ANNs for GWL simulation and forecasting.

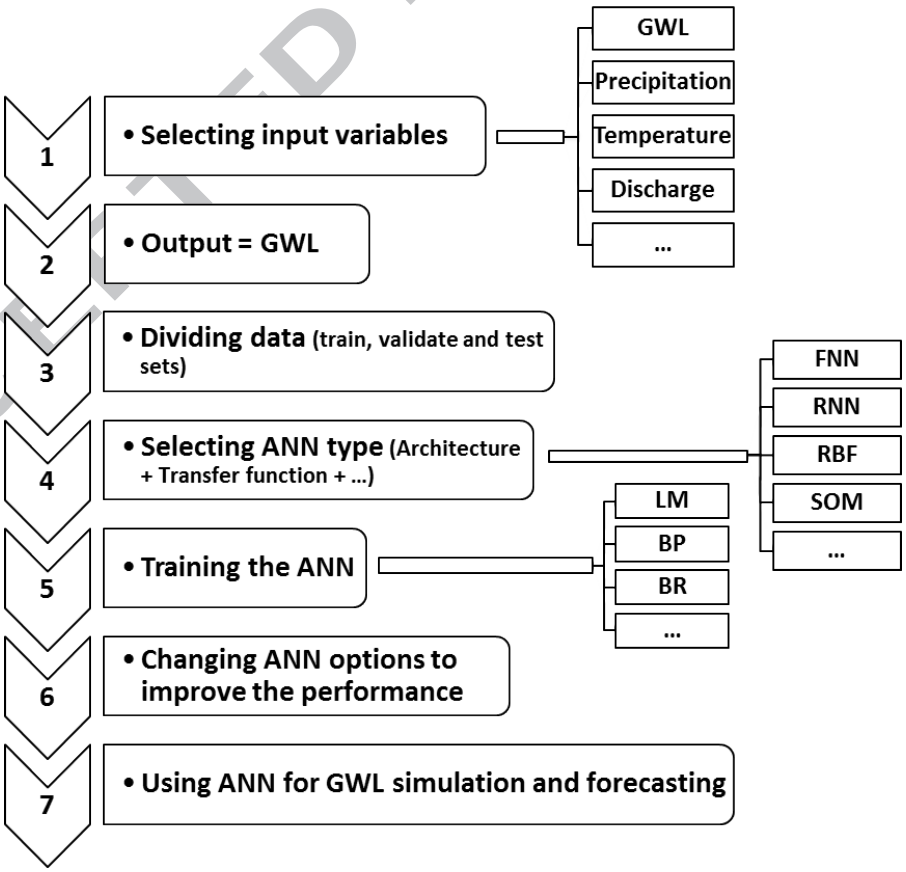


Figure 2. The stages of using ANNs for GWL forecasting.

2.1.2. Bibliographic review

Recent experiments in GWL modeling have reported that ANNs may offer a promising alternative for conceptual methods. In one of the first studies, Coulibaly et al. (2001) compared three types of ANN models using GWL, precipitation and temperature time series as the inputs of models to simulate average monthly GWL in the Gondo aquifer, Burkina Faso. Simulation results showed that the RNN is most efficient compared to the static structure input delay ANN and RBF-ANN. Lallahem et al. (2005) evaluated ANN for estimating the monthly GWL in an unconfined chalky aquifer in northern France. The input data was the GWL of 13 piezometers, rainfall, mean temperature, precipitation and potential evapotranspiration, and the main objective was to simulate the GWL in a selected piezometer. The simulations revealed the merit of using MLP models. Daliakopoulos et al. (2005) tested seven different ANN models with various architectures and training algorithms for monthly GWL forecasting in the island of Crete, Greece. The input variables were the past GWL, temperature, precipitation and river discharge. The FNN trained with the LM algorithm had the best results. Nayak et al. (2006) investigated the potential of MLP trained with BP algorithm in forecasting the monthly GWL in an unconfined coastal aquifer in India. The input variables were selected as precipitation, canal releases and GWL of the observation well and two neighboring wells. The performance was good for 1 and 2-month ahead forecasting, but was deteriorated after 2-month.

Krishna et al. (2008) applied several ANN training algorithms to predict monthly GWL in an urban coastal aquifer in Andhra Pradesh state, India. It was found that the FNN trained with LM algorithm is a good choice, compared to BR and RBF algorithms. In their study, GWL were also predicted in neighboring wells using model parameters from the best network of a well. Mohammadi (2008) tested MODFLOW and two types of ANN, i.e., MLP and RNN to simulate the monthly GWL of a karstic aquifer, located in Iran. He used data sets generated by MODFLOW for training of the ANNs. The results indicated that ANN models needed less input data and took less time to run, compared to MODFLOW. Nourani et al. (2008) compared six different types of ANNs for spatiotemporal GWL forecasting in Tabriz aquifer, Iran. The monthly GWL in central well,

precipitation, mean temperature and average discharge were selected as the inputs. The optimal ANN was a FNN trained with LM algorithm, which was then applied to forecast GWLs of selected wells, as the spatial model.

Feng et al (2008) applied FNN to investigate the effects of 7 factors i.e.: initial GWL, precipitation, evaporation, water reservoir inflow, population, synthesis irrigation ratio, and irrigation area, on monthly GWL in shiyang river basin, China. Sensitivity analysis with the models demonstrated that groundwater extraction for irrigation is the predominant factor responsible for declining GWL. Tsanis et al. (2008) developed a FNN, trained with the LM algorithm with five input variables, i.e., precipitation, temperature, runoff, GWL and specific yield for forecasting the monthly GWL in Messara Valley, Crete, Greece. They used a deterministic component, which linked precipitation with the seasonal recharge of the aquifer and projected the seasonal average precipitations. Results showed that the specific yield marginally improved the forecasting but the linearly projected precipitation component drastically increased the forecasting.

Banerjee et al. (2009) used FNN model trained with LM algorithm to predict the monthly GWL of four diversified wells in Kurmapally watershed, Hyderabad, India. They have not mentioned the used input variables but forecasted the GWL considering varying recharge and pumping conditions. Yang et al. (2009) applied the BP-ANN and the integrated time series (ITS) models to forecast monthly average GWL in the western Jilin province of China. The input variables were only the past GWLs at different intervals of time. The simulation results indicated that both ANN and ITS models were accurate in reproducing the GWLs, but in the test phase, the ANN was superior to the ITS.

Mohanty et al. (2010) developed three different training algorithms, viz., LM, BR and GDX algorithms for weekly GWL forecasting in a tropical humid region, eastern India. The inputs to the models consisted of precipitation, pan evaporation, river stage, water level in the drain, pumping rate and GWL in the previous week. The BR algorithm was found slightly superior to the two other algorithms. Chen et al. (2010) combined the theory of SOM and RBF. The proposed model could decide the number of RBF-ANN hidden units with using the two-dimensional feature map which is constructed by SOM. The inputs were the monthly average GWLs of six wells in southern Taiwan, while the output was the monthly average GWL of an individual well. The results showed that the

four-site input model was more competent compared to the single-site model and six-site model. One year later, Chen et al. (2011) combined of the SOM and BP-ANN for the same study area. Here, the model inputs were the monthly average GWLs of ten wells, while the output was the GWL of an individual well. It was found that the multi-site SOM-BP-ANN model provided the most accurate predictions in comparison to the autoregressive integrated moving average (ARIMA) and single ANN models.

Trichakis et al, (2011) simulated daily GWL by MLP at a well located in the karstic artesian Edward's aquifer in Texas, USA. The input variables were the day number, precipitation, pumping and GWL. The testing data were used to check the ability of the MLP to interpolate or extrapolate in other wells in the region. The results showed that there was a need for exact knowledge of pumping from each well in karstic aquifers as it was difficult to simulate the sudden drops and rises. Sreekanth et al. (2011) compared the FNN trained with LM algorithm and ANFIS for estimation of the GWL of the Maheshwaram watershed, India. The inputs included the monthly GWL in 22 wells along with rainfall, temperature, evaporation and relative humidity. The results showed that the FNN provided better accuracy compared to ANFIS.

Rakhshandehroo et al. (2012) used FNN, RBF, RNN and a generalized regression neural network for monthly GWL prediction in Shiraz plain, Iran. The precipitation, GWL, temperature, evaporation and runoff were utilized as the input data. Best performances were achieved by FNN and RNN networks, respectively. Taormina et al. (2012) applied FNN for long period simulations of hourly GWLs in a coastal unconfined aquifer sited in the Lagoon of Venice, Italy. The FNN was first trained to perform one-hour-ahead predictions using past GWL, rainfall and evapotranspiration data. After the training, simulations were produced by feeding back the computed outputs in place of past observed data. The FNN reconstructed accurate GWL for long periods, at least six months, relying only on the rainfall and evapotranspiration data. Sahoo and Jha (2013) compared MLP trained with LM algorithm and multi linear regression (MLR) approach in monthly GWL forecasting considering rainfall, temperature, river stage, GWL and 11 seasonal dummy variables as inputs. The study area was Konan basin, located in Kochi, Japan. They concluded that MLP models have better results;

however, considering the practical advantages of the MLR, it was recommended as a cost-effective GWL modeling tool.

Ying et al. (2014) compared the RBF-ANN, ARIMA and ITS models for GWL forecasting of two wells in Jilin, China. Monthly GWL was the only variable used to develop the models. They concluded that for forecasting the dynamics of the GWL, the RBF-ANN is preferable, but for analyzing GWL variation, the ITS and ARIMA may be more appropriate.

Juan et al. (2015) developed two FNN models, one with three inputs (previous GWL, temperature and precipitation) and another with two inputs (temperature and precipitation only) to forecast the daily variations of the supra-permafrost GWL in the Qinghai-Tibet plateau, China. The FNNs were trained with LM algorithm, and the results indicated that the three inputs model produced better accuracy performance. However, if there are no field observations of the GWL, the models developed using only two inputs also have good accuracy. Gholami et al. (2015) used a MLP trained with LM algorithm to simulate annual GWL fluctuations of two wells located in an alluvial aquifer of the Caspian Sea southern coasts, Iran, for the period from 1912 to 2013. The tree-ring diameter and the precipitation during the growing season were the input parameters for the MLP, and the GWL during the growing season was the output. The results showed that the integration of dendrochronology and ANN renders a high degree of accuracy in the simulation of annual GWL. Mohanty et al (2015) applied FNN for simultaneous forecasting of the weekly GWL in 18 wells located over a river basin in India. The inputs were selected as rainfall, pan evaporation, river stage, water level in the surface drain, pumping rates of 18 sites and GWLs of 18 sites in the previous week, which led to 40 input nodes and 18 output nodes. Comparison between the LM, BR and GDX training algorithms showed that the GDX was the most suitable algorithm for the study area.

Sun et al. (2016) applied an MLP trained with LM algorithm to forecast the daily GWL in a freshwater swamp forest of Singapore. The inputs to the model were the surrounding reservoir levels and rainfall. The results revealed that MLP produced better prediction with a leading time of 1 day compared to MLR.

Wunsch et al. (2018) used the nonlinear autoregressive with exogenous inputs neural network (NARX) for GWL forecasting of several wells in southwest Germany. Precipitation and temperature

were chosen as input variables. All input and target time series were decomposed using the seasonal trend based on loess algorithm to detect significant time lags and determine input and feedback delays needed for NARX application. The results showed that NARX is suited to perform GWL predictions for uninfluenced observation wells, even though the number of input variables is limited. Ghose et al. (2018) developed the RNN model to forecast monthly GWL of a well in Odisha, India as a function of rainfall, temperature, humidity, runoff and evapotranspiration. From the results, evapotranspiration and runoff were the influencing parameters which affect the GWL, and inclusion of them improved the model efficiency.

Lee et al. (2018) applied the FNN to predict hourly GWL of 8 observation wells located in Yangpyeong riverside area, South Korea. They investigate the relative impacts of the input variables, and as a result used the river level and pumping rates from two extraction wells as input variables, while the precipitation was found to be a weak influencing factor, and therefore it was not used as an input variable. Kouziokas et al. (2018) used multiple FNN with various network structures and training algorithms to forecast the daily GWL of a well located in Montgomery County, Pennsylvania, USA. Using the humidity, precipitation, and temperature as input variables the FNN with the LM training algotithm was the best model.

2.1.3. Results

An assessment of the various studies on ANN modeling of the GWL revealed the following issues:

1) The ANN models can be extended easily from univariate to multivariate cases compared to the conceptual models, and the model complexity can be varied simply by altering the transfer function, training algorithm or network architecture. Similar to the regression models, the input variables can be considered based on an empirical proof or a correlation analysis. The results of the reviewed papers also indicated that ANNs capture the complex non-linear behavior of the GWL time series relatively better than the regular regression models such as ARIMA and MLR.

2) The reviews reveal that the LM algorithm is the most popular training algorithm used to train ANNs for GWL modeling. The LM algorithm is a modification of the classic Newton algorithm used for finding an optimum solution to a minimization problem. The LM algorithm is often characterized as more stable and efficient, and some researchers point out that it is faster and less easily trapped in local minima than other training algorithms (Daliakopoulos et al., 2004). Zounemat-kermani et al. (2013) in a study of comparison the performance of RBF and LM feed-forward ANNs for predicting daily watershed runoff, concluded that LM algorithm is superior to the RBF in prediction of one day ahead base and high flows, but the RBF algorithm outperformed the LM in predicting flood events. The GWL time series do not possess a characteristic such as flood in runoff time series, therefore it seems that the superiority of LM in GWL modeling correspond to the results of the study of Zounemat-kermani et al. (2013).

3) The three layers FNN with the sigmoid transfer function in the hidden layer and linear transfer function in output layer is the most common structure of ANN for GWL modeling. The sigmoid function is differentiable, continuous, and monotonically increasing in its domain and it is the most frequently employed function in modeling (Ravansalar and Rajaee, 2015). It should be mentioned that in the majority of reviewed papers the structure of ANN and number of hidden neurons were achieved by a trial-and-error procedure.

2.2. Adaptive neuro-fuzzy inference system (ANFIS) for GWL modeling

2.2.1. Introductory

The adaptive neuro-fuzzy inference system is a combination of an adaptive neural network (AN) and a fuzzy inference system (FIS), thus it has potential to capture the benefits of two methods in a single framework. Jang (1993) introduced architecture and a learning procedure for the ANFIS that uses a neural network learning algorithm for constructing a set of fuzzy if-then rules with appropriate

membership functions (MFs) from the specified input-output pairs. The FIS corresponds to a set of fuzzy if-then rules that have learning capability to approximate nonlinear functions. There are two approaches for FIS, namely Mamdani and Sugeno. The differences between these two approaches arise from the consequent part. Mamdani's approach uses fuzzy MFs, whereas Sugeno's approach uses linear or constant MFs. The ANFIS is an AI method with flexible mathematical construction which is capable of identifying complex nonlinearity and uncertainties due to randomness and imprecision between variables, without attempting to reach an understanding as to the nature of the phenomena. This approach is capable of approximating any real continuous function on a compact set to any degree of accuracy. Thus, in parameter estimation/forecasting, where the given data are such that the system associates measurable system variables with an internal system parameter, a functional mapping may be constructed by ANFIS that approximates the process of estimation of the internal system parameter. More information on ANFIS can be found in Jang (1993).

2.2.2. Bibliographic review

In the area of GWL modeling with ANFIS, Kholghi and hosseini (2009) applied the ordinary kriging and ANFIS for spatial interpolation of GWL in an unconfined aquifer in Qazvin, Iran. They use the GWL data of 95 wells for training and testing the models. The Gaussian MF was used in the ANFIS models. The results showed that the contour plot of isopieze lines estimated by ANFIS was more efficient than those by kriging.

Jalalkamali et al. (2011) investigated the abilities of ANFIS and ANN with various combinations of monthly temperature, rainfall and GWLs in two neighboring wells as the inputs to predict the GWL of another well, located in Kerman plain, Iran. The results showed that applying the GWLs of the current and one month before of the well and the neighboring wells was the best input combination to predict GWL, and the ANFIS models using Gaussian MF had better results compared to the ANNs.

Shirmohammadi et al. (2013) applied system identification, time series, and ANFIS models to predict monthly GWL in Mashhad plain, Iran. The only input variable of the models was the precipitation. In the ANFIS models, they tested several MFs such as Triangular, Gaussian and Bell-

shaped functions. The results showed that the Bell-shaped MF had the best performance, and the ANFIS model outperformed both time series and system identification models.

Emamgholizadeh et al. (2014) compared ANN and ANFIS in forecasting of monthly GWL in Bastam plain, Iran. They considered the rainfall recharge, irrigation returned flow and pumping rates from water wells as input data and found that ANFIS outperformed the ANN. The results showed that applying ANFIS with different structures had the most accuracy when it used with trapezoidal MF.

Mirzavand et al. (2015) investigated the abilities of ANFIS and SVR in estimating monthly GWL fluctuation in the Kashan plain, Iran, by using the inputs of stream flow, evaporation, spring discharge, aquifer discharge and rainfall. The results indicated that the ANFIS model using Bell-shaped MF performed better than the SVR. Khaki et al. (2015) applied ANN and ANFIS to simulate monthly average GWL in the Langat Basin, Malaysia. The GWL, rainfall, humidity, evaporation, minimum temperature and maximum temperature were applied as the input variables of the models. The obtained results of the ANFIS models were superior to those of ANNs, and in the ANFIS models the Bell-shaped MF outperformed the Gaussian MF. Gong et al. (2015) tested the validity of ANN, SVM and ANFIS in the prediction of the monthly GWL for two wells near Lake Okeechobee in Florida, United States. The precipitation, temperature, past GWLs and lake level were used as input data. The results showed that the GWL predictions from ANFIS and SVM were more accurate than that from ANN.

2.2.3. Results

The review of cited studies on ANFIS modeling of the GWL showed that:

1) In the cited papers, applying ANFIS as an alternative approach to predict the GWL leads to more accurate results in comparison with the ANN. Since ANFIS integrates both neural networks and fuzzy logic principles, it is more likely to deal with non-stationary time series more effectively.

2) In three studies (i.e., Shirmohammadi et al. 2013; Mirzavand et al. 2015; Khaki et al. 2015) the Bell-shaped MF was the best in comparison with other MFs, while in two

studies (i.e., Kholghi and hosseini 2009; Jalalkamali et al. 2011) the Gaussian MF yielded higher accuracy, and in the study of Emamgholizadeh et al. (2014) the Trapezoidal MF was the best in comparison of others. In the meanwhile, Gong et al. (2015) have not mentioned anything about the used MF. Generally, there was not any exact method for choosing the MFs in the reviewed papers, and instead, a trial-and-error procedure was used for finding an appropriate MF. So, use of those MFs which do not cause overfitting and give least error can be recommended.

2.3. Genetic programming (GP) for GWL modeling

2.3.1. Introductory

The GP as a generalization of genetic algorithm (GA) is an evolutionary algorithm based on biological evolution inspired by Darwinian theories of natural selection and survival of the fittest. The GP considers an initial population of randomly generated equations, which are achieved from the random variables, numbers and functions. The function involves arithmetic operators (+, –, ×, ÷) and other mathematical functions (e.g., *sin*, *cos*, etc.) or user-defined expressions, which should be chosen based on some understanding of the process. The initial population is then applied to an evolutionary process to evaluate the fitness of the evolved programs by defining a fitness function. In forecasting problems the root mean squared error (RMSE) between forecasted and observed data is often used as the fitness function. The programs that best fit the data are then selected to produce better program through two genetic operators: crossover and mutation. The evolution process is repeated and driven towards to find expressions which describe the data and give the best performance of the model.

2.3.2. Bibliographic review

Shiri et al. (2013) investigated the abilities of GP, ANFIS, ANN, SVM and ARIMA techniques for daily GWL forecasting in Korea. The GWL, rainfall and evapotranspiration data were used as the inputs of the models. For GP models, the root relative squared error was employed as the fitness function. The results showed that GP models were superior compared to other models. Fallah-

Mehdipour et al. (2013) compared the capability of the GP and ANFIS to predict and simulate monthly GWLs in three wells in the Karaj plain of Iran. The precipitation, evaporation and GWLs were used as the inputs of the models. They have noted that the fitness function of GP was considered an error criterion, but they have not mentioned the type of it. Results showed that in the GP models the average errors were less compared to the ANFIS models.

2.3.3. Results

Originally developed for optimization problems, the GP was extended to solve forecasting problems such as GWL forecasting. In this case, the minimum error (e.g. RMSE) between forecasted and observed GWLs has been applied as the fitness function of the GP. Although, among other AI methods, the GP may not be the best way to forecast the GWL, in the two aforementioned studies, this model outperformed other models. Similar to the ANN and ANFIS, in the reviewed GP papers, the input parameters were chosen based on a combination of empirical and trial-and-error analysis. The low number of papers on GWL modeling via GP demonstrates the need to investigate more about application of GP and in GWL modeling.

2.4. Support vector machine (SVM) for GWL modeling

2.4.1. Introductory

The SVM is a statistical machine learning theory. It has not a priori determined structure, but the input vectors supporting the model structure are selected through a model training process (Vapnik, 1998). This machine learning method is based on the extension of the idea of identifying a hyper-plane that separates two classes in classification. A SVM constructs hyper-planes in an infinite dimensional space, which can be used for classification, regression, or other tasks. The mappings used by SVM schemes are designed to ensure that dot products may be computed easily in terms of the variables in the original space, by defining them in terms of a kernel function selected to suit the problem. The SVM can also be used as a regression method. The support vector regression (SVR) method uses the same principles as the SVM for classification, with only a few minor differences. The

SVR generalization performance depends on a good setting of some parameters and the kernel function. The SVR parameters represent some constants like regularization constant and kernel function constant, and control the prediction (regression) model complexity. The kernel function changes the dimensionality of the input space to perform the regression task with more confidence. A full mathematical overview of SVM is presented by Vapnik (1998). Originally developed for classification, it was extended to solve prediction problems, and in this capacity was used in hydrology related tasks.

2.4.2. Bibliographic review

Yoon et al. (2011) developed ANN and SVM models for predicting GWL fluctuations of two wells at a coastal aquifer in South Korea, considering a six-hourly time step. The past GWL, precipitation and tide level were selected as the inputs of the models. It was found that the past GWL was the most effective input variable for the study site, and tide level was more effective than precipitation. The results showed that the performance of the SVM was better than the ANN. Yoon et al. (2016) utilized a weighted error function approach to improve the performance of ANN and SVM models for the prediction of daily GWL in response to rainfall. The input variables were GWL and rainfall data in South Korea. The comparison of the models showed that the recursive prediction performance of the SVM was superior the ANN.

Huang et al. (2017) used the chaos theory to select the best input lags of GWL time series, and developed the SVM and BP-ANN models. Using the particle swarm optimization method to obtain the parameters of SVM, the models were applied to predict the daily, weekly and monthly GWL in China. The chaotic SVM model had higher accuracy than the linear SVM and chaotic BP-ANN models. Nie et al. (2017) employed precipitation, evaporation, and temperature as the inputs of SVM and RBF-ANN models to forecast monthly GWL in Jilin province, China. The SVM model was more accurate and had fewer uncertainties caused by errors in the measurements of the inputs and outputs.

Mukherjee and Ramachandran (2018) applied the GRACE satellite terrestrial water storage (TWS) data along with meteorological variables precipitation, min and max temperature, humidity

and wind to predict GWL with the SVR, ANN and linear regression models. The results showed that TWS is a highly significant variable to model GWL, and the SVR was the best model. Guzman et al. (2018) compared SVR and NARX-ANN models for GWL prediction of an irrigation well located in the southeastern USA. They evaluated the best combination from three input variables, i.e., daily GWL, precipitation and evapotranspiration data for each model. The GWL + precipitation scenario provides the optimal combination for model inputs, and the SVR was superior to the ANN. Tang et al. (2018) concluded that the least square SVM perform better than classical SVM and some other AI models in GWL forecasting. The only input variable was the hourly GWL of four observation wells located in northern UK.

2.4.3. Results

The SVMs/SVRs are powerful machine learning methods that have been developed and applied for many classification/prediction problems over past years. Although the number of published papers considering GWL modeling via SVM is low, however it should be noted that SVM has been used for predicting of many time series for a myriad of practical applications in the world.

In the SVM modeling, the appropriate selection of the kernel function and parameter values is critical. In the five of seven aforementioned papers, the RBF kernel function was selected, whereas in the two other ones (i.e., Yoon et al., (2011) and Mukherjee and Ramachandran (2018)) the utilized kernel function was not mentioned. Over period of years, the RBF function has become the choice of many researchers as the kernel function for SVR because of its accuracy and reliable performance (Suryanarayana et al., 2014).

For selecting the optimum parameters of SVM model, most of the papers have employed a procedure like trial-and-error, except Huang et al. (2017) that have been used the particle swarm optimization method to obtain the optimum parameters of SVM.

2.5. Hybrid AI techniques for GWL modeling

2.5.1. Introductory

Since it has been revealed that the AI models have some limitations with the nonlinear and non-stationary processes, some hybrid modeling approaches which include certain data-preprocessing and/or combine different AI techniques have been also developed in the recent years to increase the capabilities of the AI methods. Combining different AI methods in different stages of the modeling, and applying efficient methods for input data pre-processing make the developing of these models more effective. For example, the GP technique can be used to optimize the AI input variables and/or AI regulation parameters. In another example, the geostatistical techniques such as Kriging can be combined with the AI methods for spatiotemporal GWL modeling. According to the capability of geostatistics tools in spatial estimation, hybrid AI-geostatistic models have been applied in some papers to use their potential for spatiotemporal simulation of GWL.

The wavelet analysis is an example for the data pre-processing, which has been widely used in GWL modeling. Wavelet analysis is applied for de-noising, compression and decomposition of input data time series. Wavelet is a time-dependent spectral analysis that unravels time series in the time-frequency space to provide a time-scale description of the processes and their relationships (Daubechies 1990). The Wavelet analysis can be performed continuously or discretely. The continuous wavelet transform (CWT) can operate at every scale; but it requires a lot of computational time, and generates a large amount of data. In many studies the discrete wavelet transform (DWT) was used, where only a subset of scales and positions are chosen to make the calculations. In the wavelet-AI models cited in scientific papers, the decomposed sub-time series were used as the inputs of AI models, instead of the main time series. The schematic structure of some of hybrid AI models for GWL modeling is shown in Figure 3.

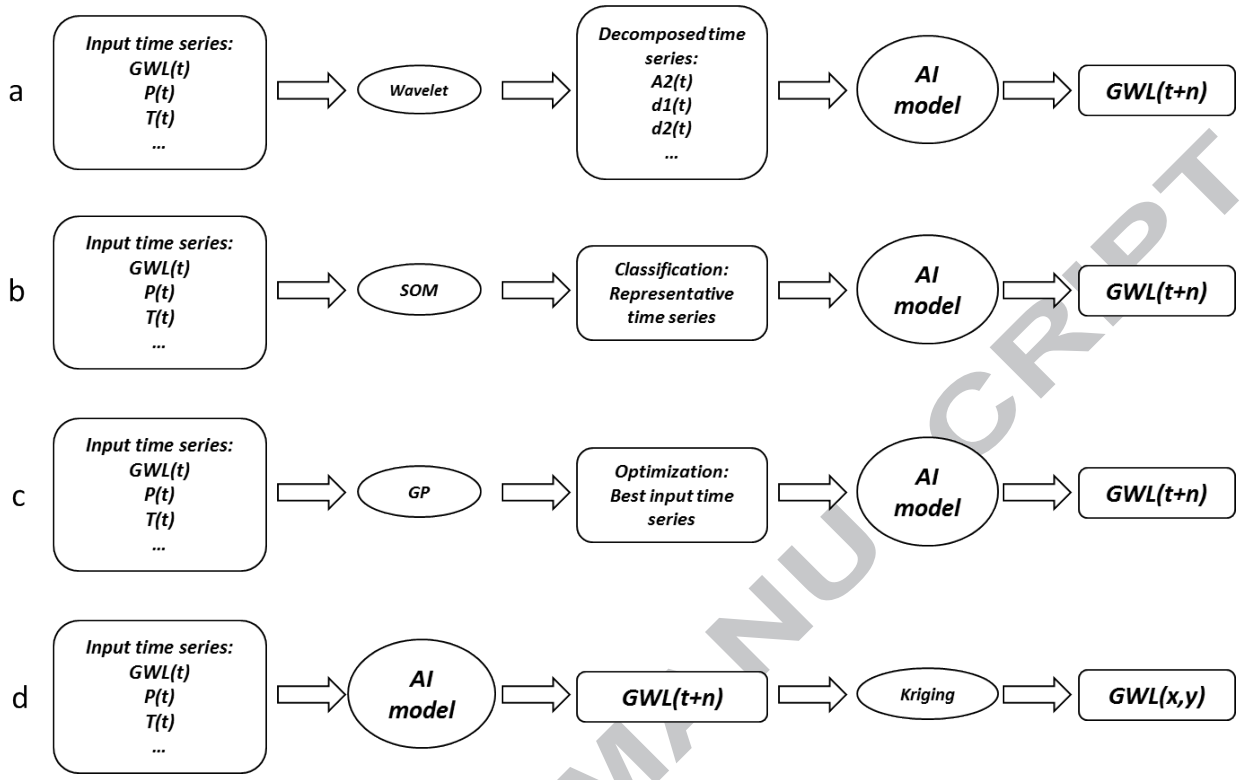


Figure 3. Schematic structure of some hybrid AI models for GWL modeling. a) Wavelet-AI model b) SOM-AI model c) GP-AI model d) AI-Kriging model.

2.5.2. Bibliographic review

Afterwards the AI methods were developed for prediction problems, the researchers tried to combine different type of them to overcome the shortcomings and increase their accuracy. Almost since 2011, there has been an interest in application of the wavelet analysis in combination with different AI methods. Adamowski and Chan (2011) used a Wavelet-ANN model for GWL forecasting at two sites in Quebec, Canada. The monthly total precipitation, average temperature and average GWL were decomposed at two levels by wavelets and imposed to the ANN. The model was found to provide more accurate GWL forecasts compared to the ANN and ARIMA models. Nourani et al. (2011) presented an ANN-geostatistics methodology for spatiotemporal prediction of GWL in Shabestar plain, which adjoins to Urmieh Lake as a coastal aquifer in Iran. Monthly GWLs data from 11 piezometers, rainfall, and lake water levels were the inputs of ANN. The ANN was trained for

each piezometer to predict the GWL of the next month. Then Kriging was applied to the outputs from ANN models in order to estimate GWL at any desired point in the plain.

Kisi and Shiri (2012) investigated the ability of a Wavelet-ANFIS model to perform one-, two- and three-day-ahead GWL forecasting of two wells located in Illinois State, USA, using only past daily GWL data. They found that excluding the detail coefficients from the inputs and using only approximation components significantly increase the accuracy of ANFIS models. The hybrid model outperformed ANFIS, particularly for two- and three-day-ahead forecasts.

Moosavi et al. (2013a) applied a number of different structures for ANN, ANFIS, Wavelet-ANN and Wavelet-ANFIS models to evaluate their performances to forecast GWL with 1, 2, 3 and 4 months ahead under two case studies in Mashhad plain, Iran. It was demonstrated that wavelet transform can improve the accuracy of forecasting. It has been also shown that the forecasts made by Wavelet-ANFIS models are more accurate than those by other models. They found that the decomposition level in wavelet transform should be determined according to the periodicity and seasonality of data series. Moosavi et al. (2013b) also investigated the optimum structures of Wavelet-ANN and Wavelet-ANFIS models for GWL forecasting in the same case studies. They used the optimization Taguchi method to assess different factors affecting the performance of models. It was revealed that transfer functions of ANN, membership function types of ANFIS and the mother wavelet type are the most important factors. Comparison of optimal models demonstrated the better performance of Wavelet-ANFIS. Maheswaran and Khosa (2013) showed that wavelet based nonlinear as wavelet-Volterra model performed better than Wavelet-ANN and wavelet-linear regression models for GWL forecasting. The study area was northern Saanich Peninsula, Canada, and the inputs of the models were the level five decomposed monthly average GWL time series.

Suryanarayana et al. (2014) predicted monthly GWL of three observation wells in the city of Visakhapatnam, India, using wavelet-SVR modeling. The monthly data of precipitation, maximum temperature, mean temperature and GWL for the period 2001–2012 are used as the input variables. Results indicated that wavelet-SVR model gives better accuracy compared with SVR, ANN and ARIMA models. He et al. (2014) linked wavelet and fractal theory methods to ANN for GWL forecasting of three sites located in Ganzhou region, northwest China. The fractal dimension was

convenient for quantitatively describing the irregularity or randomness of time series data. The results showed that this model is suitable for sites at which the fractal dimension of the wavelet decomposition detail components is large. Tapoglou et al. (2014) combined ANN, fuzzy logic and Kriging in order to simulate the spatial and temporal distribution of GWL in an area across the Isar River in Bavaria, Germany. The daily data including the GWLs in 64 wells, the surface water elevation at five observation points across the river, temperature and rainfall were used as input variables to the 64 ANNs. Different ANN architectures and variogram models were tested together with the use or not of a fuzzy logic system. The isocontour maps were presented for the hydraulic head. The best results were achieved with the use of the fuzzy logic system and by utilizing the power-law variogram.

Yang et al. (2015) developed a wavelet-ANN and an ITS model to predict GWL of a shallow coastal aquifer in Fujian province, China. The input was only the monthly GWL time series of two representative wells. The wavelet-ANN models provided more accurate results compared to the ITS models. Khalil et al. (2015) compared MLR, ANN, wavelet-MLR, wavelet-ANN, and a wavelet-ensemble ANN model for the forecasting of GWLs as a result of recharge via tailings from an abandoned mine in Quebec, Canada. The wavelet- ensemble ANN consisted of a group of wavelet-ANN members, where each of these members was trained for the same problem, and then combined to produce the output. The daily tailing recharge, total precipitation and mean air temperature were used as inputs, while the output was GWL for lead times of 1-day, 1-week and 1-month. The wavelet-ensemble ANN model performed best for each of the three lead times. Nourani et al. (2015) proposed a wavelet-entropy data pre-processing approach for ANN-based GWL modeling. They used the SOM-based clustering technique to identify spatially homogeneous clusters of GWL data and the wavelet transform to extract the non-stationary GWL, runoff and rainfall time series. The results indicated that the SOM method decreased the dimensionality of the input variables and the wavelet analysis increased the performance of the ANN model. Jha and Sahoo (2015) developed five hybrid ANN-GA models for simulating spatio-temporal GWL in Konan basin, Japan. The relevant input variables such as rainfall, max and min temperature, river stage and GWL have been considered to simulate GWL at 17 sites. The inputs and parameters of the ANN were optimized using GA optimization technique.

The GA was superior to the commonly used trial-and-error method for determining optimal ANN architecture and inputs.

Chang et al. (2016) combined the SOM, the Nonlinear Autoregressive with Exogenous Inputs (NARX) network and the kriging for predicting monthly GWL in Zhuoshui River basin, Taiwan, based on hydrologic data such as rainfall, stream flow and GWL. The SOM was used to classify the spatiotemporal patterns of regional GWL, the NARX was used to predict the mean of regional GWL, and the kriging was used to interpolate the predictions into finer grids of locations. Consequently the prediction of a GWL map was obtained. Han et al. (2016) coupled SOM and a statistical method to predict spatiotemporal monthly GWL in an arid irrigation district in the western Hexi Corridor, northwest China. The SOM was applied to identify spatially homogeneous clusters of wells, and the GWL forecasting was performed through developing a stepwise cluster multisite inference model with various predictors including climate conditions, well extractions, surface runoffs, reservoir operations and GWL measurements at previous steps. Hosseini et al. (2016) combined ANN and ant colony optimization (ACO) to simulate the GWL in Shabestar plain, Iran. The back-propagation ANN was utilized to reproduce GWL variations using the input variables including: rainfall, average discharge, temperature, evaporation, and some annual time series. Then, ACO was used to optimize and find initial connection weights and biases of a BP algorithm during the training phase. They found that the hybrid model could reduce overtraining.

Nourani and Mousavi (2016) presented a hybrid Wavelet-AI-meshless model for spatiotemporal GWL modeling in Miandoab plain, Iran. In this way firstly monthly GWL in different wells were de-noised using threshold-based wavelet method and the impact of de-noised and noisy data was compared in temporal GWL modeling by ANN and ANFIS. Then, both ANN and ANFIS models were calibrated using GWL data of each well, rainfall and runoff to predict the GWL at one month ahead. Finally, the simulated GWLs were considered as interior conditions for the multi-quadric RBF based solve of governing partial differential equation of groundwater flow to estimate GWL at any desired point within the plain. The results showed that the wavelet de-noising approach can enhance the performance of the modeling.

Ebrahimi and Rajaee (2017) investigated the effect of wavelet analysis on the training of the ANN, MLR and SVR approaches in simulating GWL. The only input variable was the monthly GWL data of two wells in the Qom plain, Iran. The results showed that for both wells, the Meyer wavelet produced better results compared to the other wavelets, and the wavelet-MLR and wavelet-SVR were the best models for the wells 1 and 2 respectively. Barzegar et al. (2017) combined wavelet with ANN and group method of data handling (GMDH) models for forecasting the monthly GWL in Azarbijan, Iran. The GWL time series were decomposed with different wavelets at two levels, and the stepwise selection was used to select appropriate lag times as the inputs of the models. To combine the advantages of different wavelets, a least squares boosting algorithm was applied. The boosting multi-wavelet-ANN models provided the best performances. Wen et al. (2017) applied wavelet-ANN with three different input combinations, i.e., (1) GWL only, (2) climatic data, and (3) GWL and climatic data to forecast the monthly GWL of two wells in Zhangye basin, China. The model with only GWL as its input yielded the best performance for one-month forecasts. However for two- and three-monthly forecasts, the model with GWL and climatic data as inputs was superior.

Rakhshandehroo et al. (2018) used wavelet-ANN trained with improved harmony search algorithm to forecast the long term daily GWL of two wells in southeast USA. The only input variable was the daily GWL, and the one-year-ahead prediction with the proposed model was acceptable. Yu et al. (2018) compared the wavelet-ANN and wavelet-SVR models in forecasting of monthly GWL of 3 wells in northwest China. Four wavelet decomposition levels were employed to decompose input time series discharge, evapotranspiration and GWL. The results showed that the wavelet-SVR performed better than wavelet-ANN. Zare and Koch (2018) used wavelet-ANFIS model with several combinations of GWL and precipitation as the inputs to simulate monthly GWL in the Miandarband plain, Iran. The results indicated that using the Symlet mother wavelet with two decomposition levels outperformed other models.

2.5.3. Results

In the last years, development of hybrid modeling approaches is seen, and in particular, there has been an increasing interest in wavelets-AI approaches for GWL modeling. These studies have shown that the hybrid/coupling models performed better than the regular models. As a downside, however, these models have also been criticized on various aspects and, in particular, the risk posed by overtraining of the model and the difficulties of parameter estimation using heuristic methods (Maheswaran and Khosa, 2013). A review of the various studies on hybrid AI modeling of the GWL revealed the following issues:

1) By using the hybrid models and in particular wavelet analysis to extract the input time series, a greater understanding and ability to simulate GWL can be achieved. The results of the studies explored in this section have revealed a higher degree of efficiency of hybrid models compared with single models in accurately forecasting GWL.
2) In the all reviewed wavelet-AI papers, the DWT has been applied to decompose time series rather than CWT. In addition to the simplicity of using DWT, this can partly be due to the nature of GWL time series, because they are recorded discretely. Furthermore, the GWL is linked with several hydrological phenomena; Thus, use of DWT at specific levels which likely refers to hourly, daily or monthly effects appears to be more useful than application of CWT which generates much more redundant information.
3) The more frequently mother wavelets used for GWL decomposition are db2 and db4, which have been considered as the appropriate mother wavelets. According to the Nourani et al. (2014), similarity in shape between the mother wavelet and the time-series is often the best guideline in choosing a reliable mother wavelet. Therefore, it can be an indication of a relative similarity between the general shape of GWL time series and Daubechies family wavelets.
4) According to the study of Maheswaran and Khosa (2012) in the field of hydrological forecasting, some mother wavelet forms that have a compact support showed better performance in the case of time series that have a short memory with transient features.

In contrast, mother wavelets with a wider support yielded better forecasting efficiencies with regard to the time series that have long-term features. Therefore, in the case of GWL time series, it does not seem that compact wavelets to be suitable for decomposition, because the GWL time series have long-term features rather than transient features, and therefore the wavelets with a wider support are more compatible with the time series.

5) In the aforementioned wavelet-based papers, five papers (Adamowski and Chan, 2011; Moosavi et al., 2013a; Moosavi et al., 2013b; Nourani et al., 2015; Ebrahimi and Rajaee, 2017) have used the decomposition level 2, two papers (Suryanarayana et al., 2014; Wen et al., 2017) have used the decomposition level 4 and one paper (Maheswaran and Khosa, 2013) has used the decomposition level 5 as the optimum decomposition levels. In the meanwhile, in Kisi and Shiri (2012), Yang et al. (2015) and Rakhshandehroo et al. (2018) the decomposition levels were not mentioned. According to Nourani et al. (2014), decomposition level *l* contains *l* details, and as an example in the case of monthly modeling denotes 2^n-month mode where $n = 1, 2, \ldots, l$, so 2^2-month mode is nearly seasonally mode. When multilevel sub-signals are entered in the wavelet-ANN models as input nodes, their assigned weights by the ANN method will be different at different decomposition levels; therefore, high weights will be applied to the high levels of the time series (Rajaee, 2011).

3. General results and discussions

In this section, some general results derived from the 67 reviewed papers such as the results related to the considering time steps, input variables, data set size, data division, study areas and type of aquifers, etc. have been mentioned and discussed.

3.1. Time step selection

In the case of utilized time steps, the majority of AI models reviewed in this study have been considered the monthly time steps for GWL modeling. The distribution of the utilized time steps is given in Figure 4. As can be seen, the monthly time step was used in 46 of the 67 papers reviewed, followed by daily (11 papers), daily (4 papers) and weekly (4 papers) time steps. A number of different time steps (i.e., 6-hourly, multi-monthly and annually) were used in some of the papers reviewed as well. The high use of the monthly time steps can be related to the high availability of monthly recorded GWL data compared to other time steps. In the most parts of the world, the GWLs do not have often significant hourly, daily or even weekly variations; however in some areas like coastal aquifers (Yoon et al., 2011; Taormina et al., 2012) or areas near the lake of dams (Lee et al., 2018), the GWLs are under influence of tidal/lake effects, and may have hourly or daily variations.

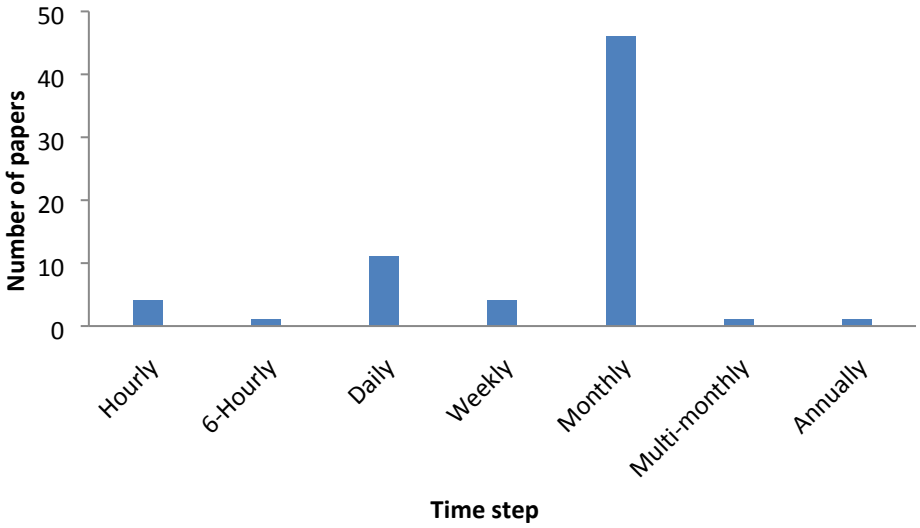


Figure 4. Number of times various time steps have been used for GWL modeling.

3.2. Input data consideration

Figure 5 shows the input variables that have been employed in AI-GWL modeling according to the reviewed papers. From Figure 5, it can be found that the past steps of the GWL time series is the most frequently used input variable for AI models to forecast the GWL. Among 67 papers, 52 papers

have been employed the GWL as an input variable. Even 12 papers have been considered the GWL as a single auto-correlated input variable without any other exogenous input variable. As well as the GWL, the precipitation has been frequently used (48 times) as an input variable. Furthermore, some hydrological time series such as temperature, river discharge (surface runoff), evapotranspiration, surface water (lake) level, pumping rates (extraction from wells) and humidity have been also used as the input variables in the reviewed papers. Other employed input variables such as irrigation patterns, population, day number, seasonal dummy variables, tree-rings, etc. have been used to a lesser extent in the reviewed papers, and it seems that some of them cannot be easily accommodated at the stage of input consideration. Although in the stage of input consideration some of the hydrological time series have been used more than the others, however it should be noted that the input data selection has been mostly based on data availability in the study area rather than a physical analysis for the required data. In the meanwhile, this cannot be considered as a weakness of these studies because in many regions data is limited, and also it is the nature of AI models that they can work with any data. However it is better that a statistical analysis and in particular a correlation analysis be done with different data before employing them for modeling in order to obtain suitable input pattern for AI models.

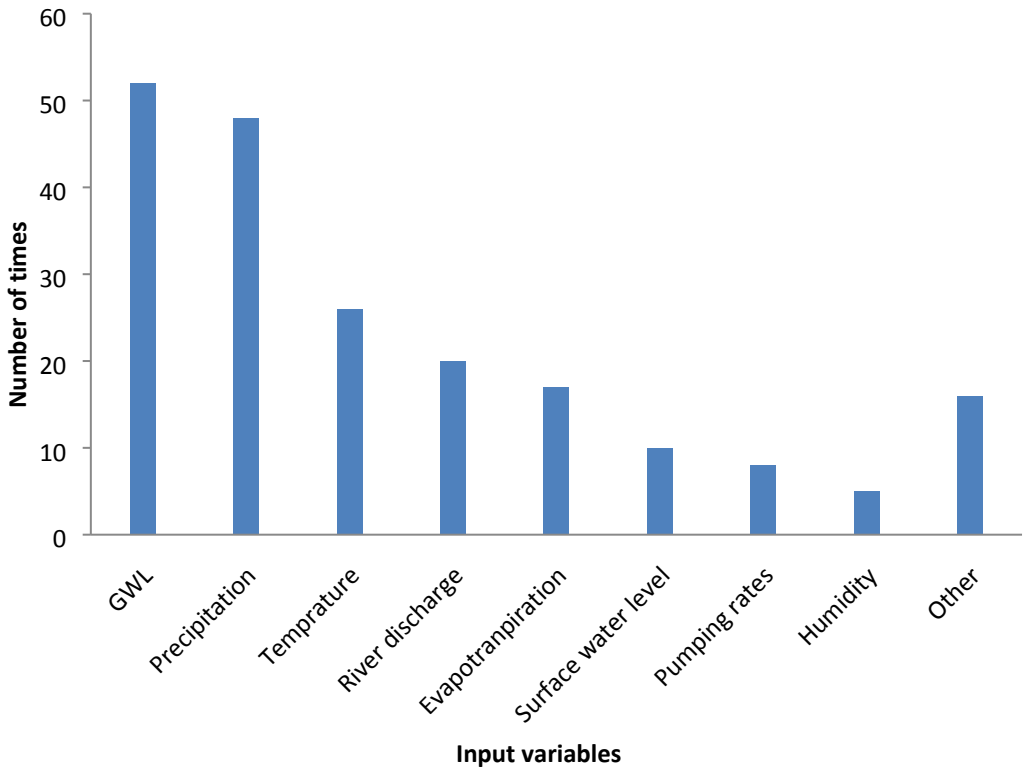


Figure 5. The input variables that have been employed for AI-GWL modeling.

3.3. Data set size

According to the Table 1, the number of total sample data sets applied for GWL modeling varies from 23 sets (Banerjee et al., 2009) to 23850 sets (Taormina et al., 2012). Generally the more samples especially for training can ensure better performance of model giving a better chance for locating global minimum of the error function, provided that an overtraining does not happen during training. However there are some cases that we may not be able to even collect 40 samples for training the model like the data of annual tree-rings in Gholami et al., (2015). The quality of the available data and the relevance of the input data with the desired output are also important since a large amount of irrelevant data can hinder the model performance by confusing the training process (Tsanis et al., 2008). There therefore has to be a balance between the quantity of data and the relevancy to the output.

In the all 67 reviewed papers there was not any fixed rule that say how to get an optimum data set size required for AI modeling. It seems that considering the available data, experimental or perhaps trial-and-error tools were used here. From Table 1 it can be seen that the majority of studies have been applied a data set size between 100-200 sets, and perhaps this can be considered as a suitable data set size. In the meanwhile it can be found that AI models are capable to deal with different size of data set, but there was not any certain comment in the reviewed papers about that in each sample size (i.e. big or small) what should we do for optimizing the model performance (e.g. which training algorithm is better for small sample size in ANN?). It seems trial-and-error procedures have been used here.

3.4. Data division

In the case of data division for training, validation and testing tasks, there was not a specific rule in the reviewed papers which explain how to consider an optimum amount for each sub-data set. In some of the reviewed papers, the total data set were divided into three parts and in some others into two parts (Figure 6). In the three part data division, the first part was used as a training or calibration set; the second part as a validation set to ascertain that the model is generalizing and to stop the training before overfitting, and the third part for a testing of the model in the prediction stage. The names of these three parts, i.e., training, validation and test parts, of course, may be different in some papers. For example in Wunsch et al. (2018), the word “validation” has been used for “testing” set and vice versa. But according to the reviewed papers, two parts data division i.e., using only the training and testing sets is also acceptable in the modeling of GWL time series, considering the fact that some researchers do not mention the validation step. As can be seen from Figure 6 most of papers have used two parts data division (training and testing sets), while some papers have included the validation set, too. Among 46 papers that have used two parts data division, the training-testing sets respectively vary from 56%-44% (Juan et al., 2015) to 90%-10% (Maheswaran and Khosa, 2013; Khalil et al., 2015) of the total data with an average of approximately 70%-30%. In the remaining 20 papers that have added the validation set, the training, validation and testing sets are averagely 60%, 18% and 22% of total data, respectively. It should be noted that in Banerjee et al., (2009) there was

not any explanation about the validation or testing sets, and the performance criteria has been only mentioned for the training data.

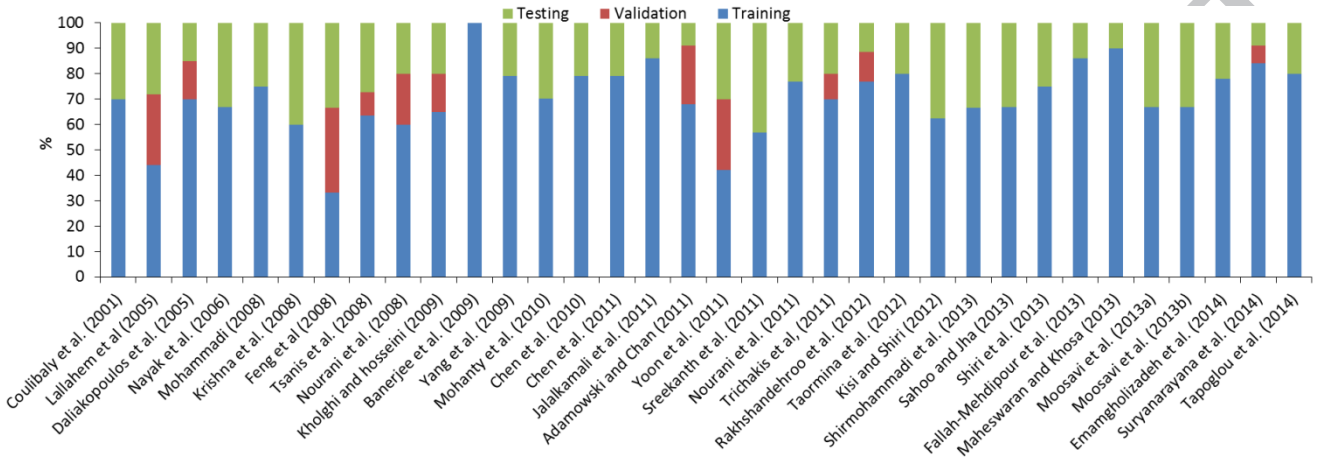


Figure 6. Percentage of the training, validation and testing sets used in the studies related to AI-GWL modeling.

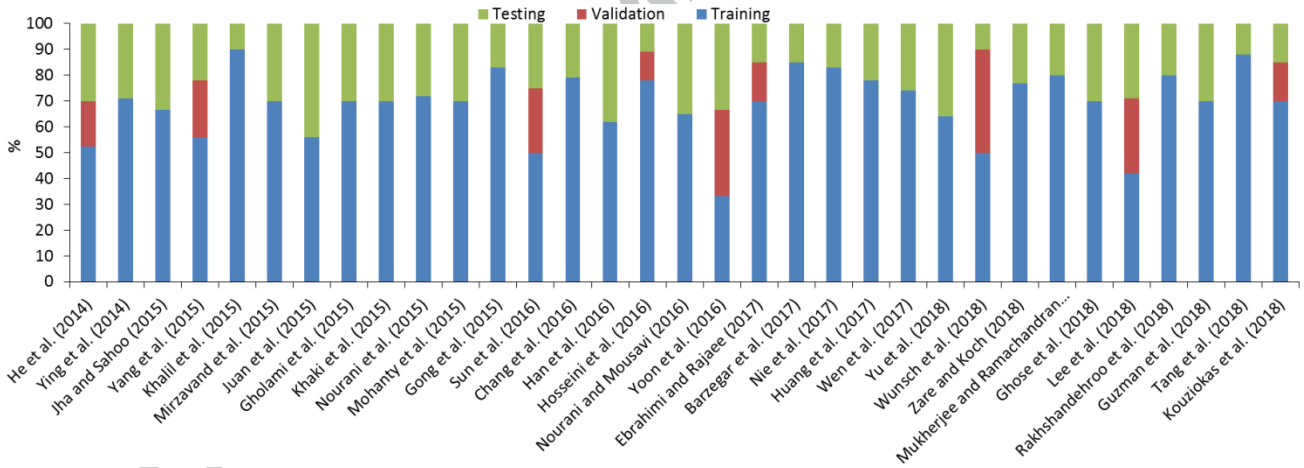


Figure 6. Continued.

3.5. Study areas and type of aquifers

Figure 7 shows the number of reviewed papers with respect to the countries where the study areas are located. A large number of the study areas are located in Iran (19 out of 67 cases). This point maybe shows the interest of Iranian researchers in this field, but it can also be due to the aridity/semi-aridity of regions like Iran, such that the surface water resources are low and the groundwater is the most available water resource, and therefore the GWL data are more available than the surface water

data. China with 11 and India with 9 case studies are placed in the next categories. In this regard, rest of the world can also be seen from the Figure 7. It should be noted that the types of aquifers under study, i.e., whether they were confined, unconfined, karstic, sandy, etc. were briefly explained in the most papers. According to the descriptions about the study areas in the reviewed papers, the most of aquifers were unconfined with alluvial materials like sand, silt, clay, gravel, etc. and only a few of them were semi-confined or karstic, chalky, coastal, etc. It is known that the black-box AI techniques are useful for prediction and forecasting, but they are not built using insights on the physical processes involved. In this type of modeling, the knowledge about the underlying mechanisms is not necessary and the main purpose is obtaining accurate forecasts.

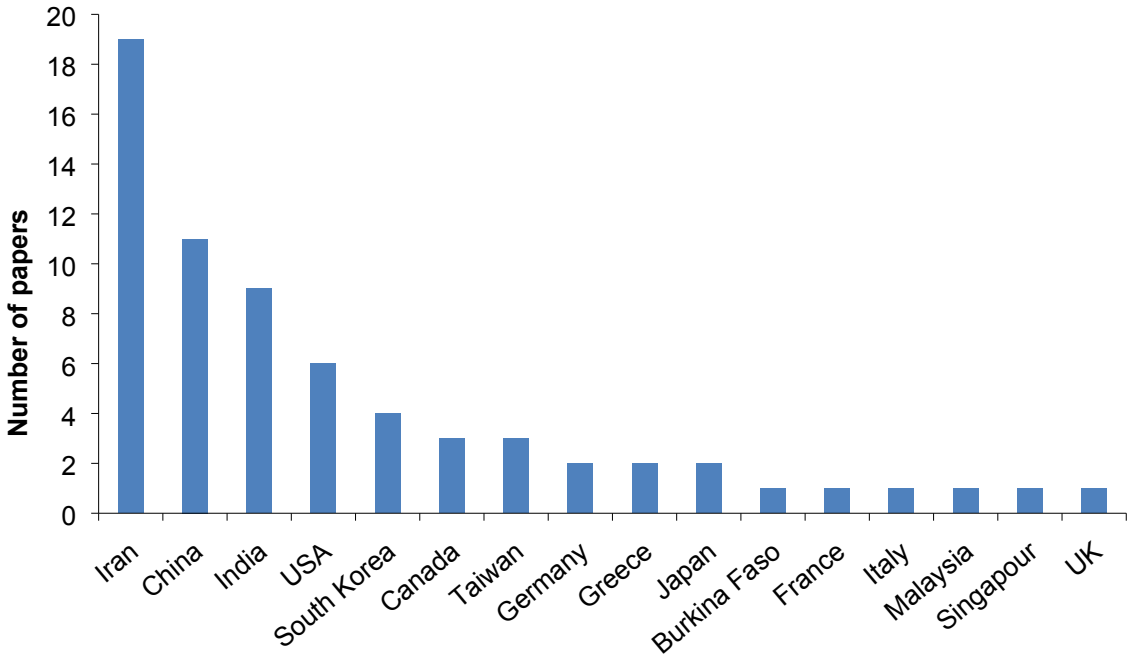


Figure 7. Number of published papers with respect to the countries where the study areas are located.

3.6. Used software programs

More than half of the papers reviewed in this study have mentioned the software programs used for AI modeling, while the rest have preferred not to mention the used software program. Figure 8 shows number of times that different software programs were used to develop ANN, ANFIS, GP and

SVM models for GWL forecasting. It should be noted that in Figure 8, the hybrid models are also considered. As can be seen, the MATLAB is the most used software program. The MATLAB software program has different AI toolboxes that allow the user to easily apply them for the desired purpose with the least needs for coding. Other software programs have been also used. For example the NeuroSolutions (Mohammadi, 2008; Jha and Sahoo, 2014; Gholami et al., 2015) , Qnet (Emamgholizadeh et al., 2014) and R (Mukherjee and Ramachandran, 2018) software programs have been used in some cases for developing the ANN. Even the programming languages such as Visual Basic (Tapoglou et al., 2014) and C (Yoon et al., 2011; Shiri et al., 2013) have been used in several papers. The GeneXpro is a software program in the field of GP and evolutionary computation that has been used by Shiri et al. (2013). Details regarding these software programs can be found on the web, and we do not discuss about them here. Although many papers have not mentioned the used software, it seems that the MATLAB software program is a good choice for development of the AI models.

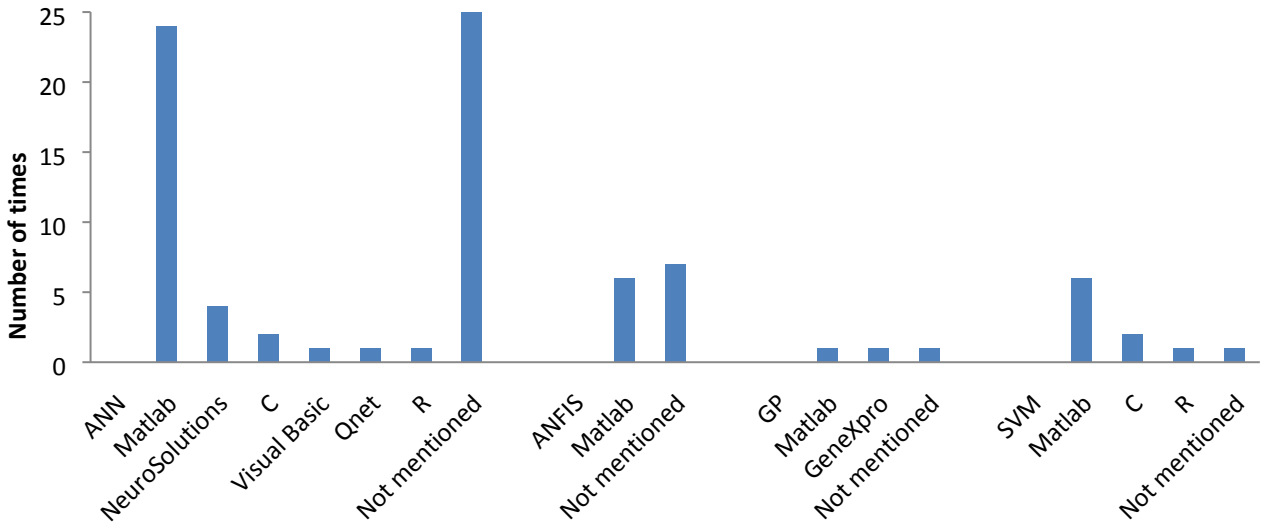


Figure 8. Number of times different software programs have been used to develop ANN, ANFIS, GP and SVM models for GWL forecasting.

3.7. Incorrect development of AI models for GWL forecasting

It is true that in the reviewed papers, the models performed well, but in the most papers, the models performances were measured in a sample testing period, and the models were not tried in a real-world forecasting. Therefore, one cannot be sure about the suitable performance of these models in the real-world GWL forecasting, and they may produce incorrect results in real-world conditions. In the other words, the model may works proper in the training, validation and even testing periods, but in fact it is incorrectly developed. In fact, to ensure the correct performance of the model, it should be tested in a real-world prediction period, while this is overlooked in most papers.

The incorrect development of AI models for GWL forecasting can be occurred in different stages of the modeling. It may be occurred during the input data consideration. If the data are insufficient, incorrect or irrelevant, we should not expect the model to have correct forecasts. When importing the inputs to the model, it is also important whether the inputs are average or related to a specific time. For example, in the monthly time steps, it is important to know whether the input data are related to the monthly average or to a specific day in each month, and whether they are recorded in the same day of each month or not (i.e., whether the record period is 30 days or longer or shorter). Use of too many inputs is also caused by input redundancy, where they may provide redundant information, and cause overfitting, and therefore the real-world forecasted GWL to be incorrect. The incorrect development can also be during the data division in training, validation and testing sub-sets, when the data have not been appropriately divided. The training, validation and testing sub-sets should have the same statistical properties in order to develop the best possible model (Maier et al., 2010). A number of best ways for considering the input data, and input data division can be found in Maier et al. (2010).

One of the most common mistakes occurs when developing hybrid wavelet-Al models. Many recent wavelet-based hydrological (including GWL) forecasting models have been incorrectly developed and cannot properly be used for real-world forecasting problems (Quilty and Adamowski, 2018). According to the Quilty and Adamowski (2018), the incorrect development of wavelet-based forecasting models occurs during wavelet decomposition and as a result imports error into the model

inputs. The origin of this error is due to the boundary condition that is linked to the wavelet decomposition in three main issues, i.e., using future data, inappropriately selecting decomposition levels and wavelet filters, and not properly partitioning training, validation and testing data. The future data issue occurs when a given wavelet requires data from the future of the time series to calculate a wavelet or scaling coefficient in the present. For solving this problem, the causal wavelet algorithms such as *a trous* and maximal overlap DWT should be used since they do not use future data. In addition to the not using the future data, the causal algorithms reduce the number of wavelet and scaling coefficients affected by the boundary condition, which must be removed from the input sub-time series to have a real-world forecasting model. The partitioning issue is also solved when using causal wavelet algorithms, but the wavelet must be applied to the testing/predicting set one record at a time, and then the forecast must be calculated through the model for each testing/predicting record and so on (Quilty and Adamowski, 2018).

In the current review study several AI methods for GWL modeling were investigated by surveying the recent published researches in this field. Here, one of the important issues is exploring which AI method works better and can best simulate the GWL. It seems that the answer to this question can be different in different studies. According to the Table 1, among 67 papers, the ANN, ANFIS, GP, and SVM were respectively declared as the most appropriate models by 28, 6, 2, and 7 papers; while 17 papers used hybrid wavelet-AI models and 7 papers applied other hybrid AI models, and reported that hybrid models led to better modeling. It appears that in the last few years more attention has been paid to apply hybrid models, so that application of hybrid models leads to better results in comparison with single AI models. In particular the pre-processing of input data by common tools such as wavelet analysis has frequently been used in this area to achieve better modeling performance.

4. Conclusions and recommendations

The AI methods have been used for GWL modeling as well as other hydrological and environmental modeling. In this study, 67 papers dealing with AI methods in GWL modeling which were published in 29 international journals from 2001 to 2018 were reviewed. From these papers it was found that AI methods can successfully be used to simulate and predict the GWL time series in different aquifers. This kind of modeling is based on an AI effort to find natural relationships between GWL and different hydrological variables without the need for constructing any conceptual model. The AI models can be useful when it is difficult to build an adequate knowledge driven simulation model due to the lack of the ability to satisfactorily construct a mathematical/physical model of the underlying processes. These models have several key stages including input data consideration, input data division, regulation of the model features, training, testing, etc. which if all the stages carefully be developed, it is expected that the model performance to be good. However, it should be noted that there was not a fixed rule for these stages, such that different studies performed each stage based on an empirical manner and/or trial-and-error procedure considering available data and existing conditions. The obtained results from this review study that were embedded in two separated parts (i.e., the results of each AI method and the general results and discussion) can provide many guidelines for researchers to perform similar works in the related field, develop innovative methods and improve the quality of modeling. For this purpose, the following recommendations can also be suggested:

1) The AI methods can be linked to conceptual-numerical models such as MODFLOW to develop integrated modular models such that each method covers the weak points of the other method. For example, if an AI model generates accurate GWL forecasts in a special aquifer, it can be used to prepare and complete GWL data required for MODFLOW as the input. According to Mohammadi (2008) the ANNs needed less input data and took less time to run, compared to MODFLOW, therefore using ANNs (and other AI methods) can decrease the computations of MODFLOW which are very time-consuming. In another

example the GWL data sets estimated by MODFLOW can be used to train AI models, if there was not enough real data.

2) More attention should be given in the stage of input consideration in order to select appropriate input variables and lag times. In the reviewed papers, the input variables were often selected based on data availability or using simple user-defined relationships. More analytical methods or model-based approaches can be applied to determine input significance, as suggested by Wu et al. (2014). In particular, utilizing the GWL time series as the most widely used and most important input variable for AI GWL forecasting, should be more investigated. The GWL fluctuations provide a direct measure of the impact of groundwater development, and important information about the aquifer dynamics is embedded in GWL time series, so it can be said that the future of GWL is predictable form the past GWL. Furthermore, in the stage of input consideration, the non-causal wavelets such as *a trous* and maximal overlap DWT can be explored to unravel the component features of different input variables in order to determine the lags, correlation and interaction between the hydrological variables and GWL.

3) Regarding different AI methods to simulate the GWL, it can be said that it is not practically possible to recommend one particular type of AI model for a given problem. However it is clear that a hybrid/coupled model likely perform better than a single AI model. Different types of AI techniques can be tested at the different stages of the GWL modeling to select the best AI method in each stage and then combine them to have an optimum modeling performance.

4) In the wavelet decomposition of the GWL, border effects as well as the caution of causality which occurs in the beginning and end of the decomposed sub-time series, is an area that has received a little attention in the papers which have used wavelet-AI models for GWL modeling, so this topic can be raised for the new researches. The decomposition of the total data set at once or each sub-data set (i.e., training, validation and testing sets)

separately, and the ways to prepare the decomposed sub-time series for applying them as the model input is an interesting subject deserving further investigation.

5) According to the Quilty and Adamowski (2018), many wavelet-based hydrological models have been incorrectly developed, and the solution is the use of non-causal wavelet algorithms such as *a trous* and maximal overlap DWT algorithms. Since this has not been done so far in GWL forecasting, the use of these wavelet algorithms should be addressed in a new study.

References

Adamowski, J., Chan, F.H., 2011. A wavelet neural network conjunction model for groundwater level forecasting. J. Hydrol. 407, 28–40.

Barzegar, R., Fijani, E., Asghari Moghaddam, A., Tziritis, E., 2017. Forecasting of groundwater level fluctuations using ensemble hybrid multi-wavelet neural network-based models. Sci. Total Environ. 599–600, 20–31.

Banerjee, P., Prasad, R.K., Singh, V.S., 2009. Forecasting of groundwater level in hard rock region using artificial neural network. Environ. Geol. 58 (6), 1239–1246.

Beale, M.H., Hagan, M.T., Demuth, H.B., 2010. Neural Network Toolbox 7 User's Guide. The Mathworks, Inc., Natick, Massachusetts, USA.

Butler, J.J., Stotler, R.L., Whittemore, D.O., Reboulet, E.C., 2013. Interpretation of water level changes in the high plains aquifer in western Kansas. Groundwater 51(2), 180–190.

Chang, F.J., Chang, L.C., Huang, C.W., Kao, I.F., 2016. Prediction of monthly regional groundwater levels through hybrid soft-computing techniques. J. Hydrol. 541, 965–976.

Chen, L.H., Chen, C.T., Pan, Y.G., 2010. Groundwater level prediction using SOM-RBFN multisite model. J. Hydrol. Eng. 15 (8), 624–631.

Chen, L.H., Chen, C.T., Lin, D.W., 2011. Application of integrated back-propagation network and self-organizing map for groundwater level forecasting. J. Water Res. Plan. Man. 137 (4), 352–365.

Coulibaly, P., Anctil, F., Aravena, R., Bobee, B., 2001. Artificial neural network modeling of water table depth fluctuations. Water Resour. Res. 37 (4), 885–896.

Daliakopoulos, I.N., Coulibaly, P., Tsanis, I.K. 2005. Groundwater level forecasting using artificial neural networks. J. Hydrol. 309, 229–240.

Daubechies, I., 1990. The wavelet transform, time–frequency localization and signal analysis. IEEE Trans. Inf. Theory 36 (5).

Ebrahimi, H., Rajaee, T., 2017. Simulation of groundwater level variations using wavelet combined with neural network, linear regression and support vector machine. Global Planet. Change 148, 181–191.

Emamgholizadeh, S., Moslemi, K., Karami, G., 2014. Prediction the groundwater level of bastam plain (Iran) by artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS). Water Resour. Manag. 28 (15), 5433–5446.

Feng, S., Kang, S., Huo, Z., Chen, S., Mao, X. (2008). Neural networks to simulate regional groundwater levels affected by human activities. Groundwater 46(1), 80–90.

Fallah-Mehdipour, E., Bozorg Haddad, O., Marino, M.A., 2013. Prediction and simulation of monthly groundwater levels by genetic programming. J. Hydro-environ. Res. 7 (4), 1–8.

Gholami, V., Chau, K.W., Fadaee, F., Torkaman, J., Ghaffari, A., 2015. Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers. J. Hydrol. 529, 1060–1069.

Gong, Y., Zhang, Y., Lan, S., Wang, H. 2015. A comparative study of artificial neural networks, support vector machines and adaptive neuro fuzzy inference system for forecasting groundwater levels near Lake Okeechobee, Florida. Water Resour. Manag. http://dx.doi.org/10.1007/s11269-015-1167-8.

Ghose, D., Das, U., Roy, P., 2018. Modeling response of runoff and evapotranspiration for predicting water table depth in arid region using dynamic recurrent neural network. Groundwater for Sustainable Development 6, 263–269. http://dx.doi.org/6. 10.1016/j.gsd.2018.01.007.

Guzman, S.M., Paz, J.O., Tagert, M.L.M., Mercer, A.E. 2018. Evaluation of seasonally classified inputs for the prediction of daily groundwater levels: NARX networks vs support vector machines. Environ. Model. Assess. https://dx.doi.org/10.1007/s10666-018-9639-x182.

Han, J.C., Huang, Y., Li, Z., Zhao, C., Cheng, G., 2016. Groundwater level prediction using a SOM-aided stepwise cluster inference model. J. Environ. Manage. 182, 308–321.

He, Z., Zhang, Y., Guo, Q., Zhao, X., 2014. Comparative study of artificial neural networks and wavelet artificial neural networks for groundwater depth data forecasting with various curve fractal dimensions. Water Resour. Manag. 28, 5297–5317.

Hosseini, Z., Gharechelou, S., Nakhaei, M., Gharechelou, S., 2016. Optimal design of BP algorithm by ACOR model for groundwater level forecasting: A case study on Shabestar plain, Iran. Arab. J. Geosci. 9 (6). https://dx.doi.org/ 10.1007/s12517-016-2454-2.

Huang, F., Huang, J., Jiang, S., Shou, C., 2017. Prediction of groundwater levels using evidence of chaos and support vector machine. J. Hydroinform. 19 (4), 586–606.

Jalalkamali, A., Sedghi, H., Manshouri, M., 2011. Monthly groundwater level prediction using ANN and neuro-fuzzy models: a case study on Kerman plain, Iran. J. Hydroinform. 13 (4), 867–876.

Jang, J.S.R., 1993. Adaptive-network-based fuzzy inference system. IEEE, Man, and Cybernetics. 23 (3), 665–685.

Jha, M.K., Sahoo, S., 2015. Efficacy of neural network and genetic algorithm techniques in simulating spatio-temporal fluctuations of groundwater. Hydrol. Process. 29 (5), 671–691.

Juan, C., Genxu, W., Tianxu, M., 2015. Simulation and prediction of suprapermafrost groundwater level variation in response to climate change using a neural network model. J. Hydrol. 529, 1211–1220.

Khaki, M., Yusoff, I., Islami, N., 2015. Simulation of groundwater level through artificial intelligence system. Environ. Earth Sci. 73 (12), 8357–8367.

Khalil, B., Broda, S., Adamowski, J., Ozga-Zielinski, B., Donohoe, A., 2015. Short-term forecasting of groundwater levels under conditions of mine-tailings recharge using wavelet ensemble neural network models. Hydrogeol. J. 23, 121–141.

Kholghi, M., Hosseini, S.M., 2009. Comparison of groundwater level estimation using neuro-fuzzy and ordinary kriging. Environ. Model. Assess. 14 (6), 729–737.

Kisi, O., Shiri, J., 2012. Wavelet and neuro-fuzzy conjunction model for predicting water table depth fluctuations. Hydrol. Res. 43 (3), 286–300.

Krishna, B., Satyaji Rao, Y.R., Vijaya, T., 2008. Modelling groundwater levels in an urban coastal aquifer using artificail neural networks. Hydrol. Process. 22, 1180–1188.

Kouziokas, G.N., Chatzigeorgiou, A., Perakis, K., 2018. Multilayer feed forward models in groundwater level forecasting using meteorological data in public management. Water Resour. Manag., 1–12. http://dx.doi.org/ 10.1007/s11269-018-2126-y.

Lallahem, S., Mania, J., Hani, A., Najjar, Y., 2005. On the use of neural networks to evaluate groundwater levels in fractured media. J. Hydrol. 307, 92–111.

Lee, S., Lee, K., Yoon, H., 2018. Using artificial neural network models for groundwater level forecasting and assessment of the relative impacts of influencing factors. Hydrogeol. J. http://dx.doi.org/ 10.1007/s10040-018-1866-3.

Maheswaran, R., Khosa, R., 2012. Comparative study of different wavelets for hydrologic forecasting. Comput. Geosci. 46, 284–295.

Maheswaran, R., Khosa, R., 2013. Long term forecasting of groundwater levels with evidence of non-stationary and nonlinear characteristics. Comput. Geosci. 52, 422–436.

Maier, H.R., Jain, A., Dandy, G.C., Sudheer, K.P., 2010. Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions. Environ. Modell. Softw. 25, 891–909.

Mirzavand, M., Khoshnevisan, B., Shamshirband, S., Kisi, O., Ahmad, R., Akib., S., 2015. Evaluating groundwater level fluctuation by support vector regression and neuro-fuzzy methods: a comparative study. Nat. Hazards. http://dx.doi.org/10.1007/s11069-015-1602-4.

Mohammadi, K. 2008. Groundwater table estimation using MODFLOW and artificial neural networks. In: Abrahart R.J., See L.M., Solomatine D.P. (eds) Practical Hydroinformatics. Water Science and Technology Library 68 (2), 127–138.

Mohanty, S., Jha, M.K., Kumar, A., Sudheer K.P., 2010. Artificial neural network modeling for groundwater level forecasting in a river Island of eastern India. Water Resour. Manag. 24, 1845–1865.

Mohanty, S., Jha, M.K., Raul, S.K., Panda, R.K., Sudheer K.P., 2015. Using artificial neural network approach for simultaneous forecasting of weekly groundwater levels at multiple sites. Water Resour. Manag. 29 (15), 5521–5532.

Moosavi, V., Vafakhah, M., Shirmohammadi, B., Behnia, N., 2013a. A wavelet-ANFIS hybrid model for groundwater level forecasting for different prediction periods. Water Resour. Manag. 27, 1301–1321.

Moosavi, V., Vafakhah, M., Shirmohammadi, B., Ranjbar, M., 2013b. Optimization of wavelet-ANFIS and wavelet-ANN hybrid models by Taguchi method for groundwater level forecasting. Arab. J. Sci. Eng. 39 (3), 1785–1796.

Mukherjee, A., Ramachandran, P., 2013. Prediction of GWL with the help of GRACE TWS for unevenly spaced time series data in India: Analysis of comparative performances of SVR, ANN and LRM. J. Hydrol. 558, 647–658.

Nayak, P.C., Satyaji Rao, Y.R., Sudheer K.P., 2006. Groundwater level forecasting in a shallow aquifer using artificial neural network approach. Water Resour. Manag. 20, 77–90.

Nie, S., Bian, J., Wan H., Sun, X., Zhang, B., 2017. Simulation and uncertainty analysis for groundwater levels using radial basis function neural network and support vector machine models. J. Water Supply Res. T. 66 (1), 15–24.

Nourani, V., Asghari Mogaddam, A., Nadiri, A.O., 2008. An ANN-based model for spatiotemporal groundwater level forecasting. Hydrol. Process. 22, 5054–5066.

Nourani, V., Ejlali, R.G., Alami, M.T., 2011. Spatiotemporal groundwater level forecasting in coastal aquifers by hybrid artificial neural network-geostatistics Model: A case study. Environ. Eng. Sci. 28 (3), 217–228.

Nourani, V., Hosseini Baghanam, A., Adamowski, J., Kisi, O., 2014. Applications of hybrid wavelet–Artificial Intelligence models in hydrology: A review. J. Hydrol. 514, 358–377.

Nourani, V., Alami, M.T., Daneshvar Vousoughi, F., 2015. Wavelet-entropy data pre-processing approach for ANN-based groundwater level modeling. J. Hydrol. 524, 255–269.

Nourani, V., Mousavi, S., 2016. Spatiotemporal groundwater level modeling using hybrid artificial intelligence-meshless method. J. Hydrol. 536, 10–25.

Quilty J., Adamowski, J., 2018. Addressing the incorrect usage of wavelet-based hydrological and water resources forecasting models for real-world applications with best practices and a new forecasting framework. J. Hydrol. 563, 336-353.

Rajaee, T. 2010. Wavelet and neuro-fuzzy conjunction approach for suspended sediment prediction. Clean-Soil Air Water 38 (3), 275–86.

Rajaee, T., 2011. Wavelet and ANN combination model for prediction of daily suspended sediment load in rivers. Sci. Total Environ. 409, 2917–2928.

Rajaee, T., Boroumand, A. 2015. Forecasting of chlorophyll-a concentrations in South San Francisco Bay using five different models. Appl. Ocean Res. 53, 208–217.

Rakhshandehroo, G.R., Vaghefi, M., Asadi Aghbolaghi, M. 2012. Forecasting groundwater level in Shiraz plain using artificial neural networks. Arab. J. Sci. Eng. 37, 1871–1883.

Rakhshandehroo, G., Akbari, H., Afshari Igder, M., Ostadzadeh, E. 2018. Long-term groundwater level forecasting in shallow and deep wells using wavelet neural networks trained by an improved harmony search algorithm. J. Hydrol. Eng. 23 (2). http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0001591.

Ravansalar, M., Rajaee, T. 2015. Evaluation of wavelet performance via an ANN-based electrical conductivity prediction model. Environ. Monit. Assess. http://dx.doi.org/10.1007/s10661-015-4590-7.

Sahoo, S., Jha, M.K., 2013. Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment. Hydrogeol. J. 21, 1865–1887.

Shiri, J., Kisi, O., Yoon, H., Lee, K.K., Nazemi, A.H., 2013. Predicting groundwater level fluctuations with meteorological effect implications—A comparative study among soft computing techniques. Comput. Geosci. 56, 32–44.

Shirmohammadi, B., Vafakhah, M., Moosavi, V., Moghaddamnia, A. 2013. Application of several data-driven techniques for predicting groundwater level. Water Resour. Manag. 27 (2), 419–432.

Solomatine, D.P., 2005. Data-driven modeling and computational intelligence methods in hydrology. In: Anderson, M. (Ed.), Encyclopedia of Hydrological Sciences. Wiley, New York.

Sreekanth, P.D., Sreedevi, P.D., Ahmed, S., Geethanjali, N., 2011. Comparison of FFNN and ANFIS models for estimating groundwater level. Environ. Earth Sci. 62, 1301–1310.

Sun, Y., Wendi, D., Kim, D.E., Liong, S.Y., 2016. Technical note: Application of artificial neural networks in groundwater table forecasting – a case study in a Singapore swamp forest. Hydrol. Earth Syst. Sc. 20, 1405–1412.

Suryanarayana, C., Sudheer, C., Mahammood, V., Panigrahi, B.K., 2014. An integrated wavelet-support vector machine for groundwater level prediction in Visakhapatnam, India. Neurocomputing 145, 324–335.

Tang, Y., Zang, C., Wei, Y., Jiang, M., 2018. Data-driven modeling of groundwater level with least-square support vector machine and spatial–temporal analysis. Geotechnical and Geological Engineering. https://dx.doi.org/10.1007/s10706-018-0713-6.

Taormina, R., Chau, K., Sethi, R., 2012. Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon. Eng. Appl. Artif. Intel. 25, 1670–1676.

Tapoglou, E., Karatzas, G.P., Trichakis, I.C., Varouchakis, E.A., 2014. A spatio-temporal hybrid neural network-Kriging model for groundwater level simulation. J. Hydrol. 519, 3193–3203.

Trichakis, I.C., Nikolos, I.K., Karatzas, G.P., 2011. Artificial Neural Network (ANN) Based Modeling for Karstic Groundwater Level Simulation. Water Resour. Manag. 25, 1143–1152.

Tsanis, I.K., Coulibaly, P., Daliakopoulos, I.N., 2014. Improving groundwater level forecasting with a feed-forward neural network and linearly regressed projected precipitation. J. Hydroinform. 10 (4), 317–330.

Vapnik, V.N.,1998. Statistical learning theory. Wiley, New York, USA. 736 p.

Wen, X., Feng, Q., Deo, R.C., Wu, M., Si, J., 2017. Wavelet analysis–artificial neural network conjunction models for multi-scale monthly groundwater level predicting in an arid inland river basin, northwestern China. Hydrol. Res. 48 (6), 1710–1729.

Wu, W., Dandy, G.C., Maier, H.R., 2014. Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling. Environ. Modell. Softw. 54, 108–127.

Wunsch, A., Liesch, T., Broda, S., 2018. Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX). J. Hydrol. http://dx.doi.org/10.1016/j.jhydrol.2018.01.045.

Yang, Z.P., Lu, W.X., Lonf, Y.Q., Li, P., 2009. Application and comparison of two prediction models for groundwater levels: A case study in Western Jilin Province, China. J. Arid Environ. 73, 487–492.

Yang, Q., Hou, Z., Wang, Y., Zhao, Y., Delgado, J., 2015. A comparative study of shallow groundwater level simulation with WA–ANN and ITS model in a coastal island of south China. Arab. J. Geosci. 8 (9), 6583–6593.

Ying, Z., Wenxi, L., Haibo, C., Jiannan, L., 2014. Comparison of three forecasting models for groundwater levels: a case study in the semiarid area of west Jilin Province, China. J. Water Supply Res. T. (8), 671–683.

Yoon, H., Jun, S.C., Hyun, Y., Bae, G.O., Lee, K.K., 2011. A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. J. Hydrol. 396 (1), 128–138.

Yoon, H., Hyun, Y., Ha, K., Lee, K.K., Kim, G.,B., 2016. A method to improve the stability and accuracy of ANN- and SVM-based time series models for long-term groundwater level predictions. Comput. Geosci. 90, 144–155.

Yu, H., Wen, X., Feng, Q., Deo, R.C., Si, J., Wu, M., 2018. Comparative study of hybrid-wavelet artificial intelligence models for monthly groundwater depth forecasting in extreme arid regions, northwest China. Water Resour. Manag. 32 (1), 301–323.

Zare, M., Koch, M., 2018. Groundwater level fluctuations simulation and prediction by ANFIS- and hybrid Wavelet-ANFIS/Fuzzy C-Means (FCM) clustering models: Application to the Miandarband plain. J. Hydro-environ. Res. 18, 63–76.

Zounemat-kermani, M., Kisi, O., Rajaee, T., 2013. Performance of radial basis and LM-feed forward artificial neural networks for predicting daily watershed runoff. Appl. Soft Comput. 13, 4633–4644.

Highlights

From the reviewed papers it can be concluded that if the AI methods properly be developed, they can successfully be used to simulate and forecast the GWL time series in different aquifers.

Many partial and general results were achieved from the reviewing of the papers, which can provide applicable guidelines for researchers whom want to perform similar works in this field.

Several new ideas in the related area of research are also presented in this study for developing innovative methods and for improving the quality of the modeling.

The AI methods can be linked to physically/conceptually based models such as MODFLOW to develop integrated modular models such that each method covers the weak points of the other method.

Regarding different AI methods to simulate the GWL, it can be said that it is not practically possible to recommend one particular type of AI model for a given problem. However it is clear that a hybrid/coupled model likely perform better than a single AI model.