To anticipate flood in basins that is lacked of statistics with use of regression table and artificial-nervous networks

[Case study-southern basins in Caspian lake]

Afshin Jahangirzadeh, Ayandegan Higher Educational Institute, Iran, e-mail: afshin.jk@gmail.com

Keyvan Kimiaei, Mahab Ghodss Consulting Expert, Iran, e-mail: keivan326@gmail.com

Abstract:

Artificial nervous networks, (ANN) in basis of hydrological modeling have been used increasingly. In spite of, less attention than pay to use of this tools for estimating flood in basins that is lacked of statistics, is one of the complicated problems of hydrologists. In this article, the abilities of this two kinds of nervous network in anticipating T years floods in southern basins in the Caspian lake with use of MATLAB 7.0.4 software considered and with new compared. For this reason, in first, nervous networks with use of physiography date and climatic, selected by multiple regression model, trained and the best structure of network. For estimating the torrents of T years of the similar basins of statistics lack, in terms of coefficient correlation among observational and calculated discharges have been selected. The obtained results, the ability of ANN in anticipation of T year floods and the impact of the type of selected network in careful anticipation have been proved.

Keywords:

Nervous network, MLP, Nervous network Elman, Regression model, Anticipate flood, The basins of statistics lake, Caspian lake, IRAN.

1. Introduction

Chronological and local changeable a river system has needed the anticipating of discharge. Anticipating of discharge is an important necessary to manage of optimum water resources, prevent of draught, anticipate of flood for maintaining human safety, water construction and to preserve of environment. Draught and flood will bring a great deal of damages and financial and human injuries. As a result a sever require to systems with ability to anticipate of levels of water or discharge in river is necessary. Anticipating of flood in basins that lack of statistic -[in other words in basins that statistics and information about torrent isn't at all]- is very difficult and is one of the greatest disturbances of hydrologists. In this basins and local climate, anticipated the flood, with respect to FEH (UK Flood Estimation Handbook), in related to these cases, anticipating must be on the basis of Hydrological data such as basin area, rain and the type of soil, as similar to data of basins that have statistics[3]. Therefore, several methods must be applied for anticipating of flood in basins that lack of statistics and have a little need to hydrological statistics.

Artificial nervous network (ANN) at first, were introduced in 1940s. In 1960's this tools have advanced increasingly. After of half 1980's, nervous networks have regarded to researchers. Rumel hart and mcclelland

(1986) have found an algorithm that was suitable for teaching (instruction) of these networks with each size and complexity. Research in ANN has counting to a number of these various networks, and instructional algorithms were introduced[3]. In Artificial nervous networks try to find a similarity between of biological structure of human brain and nervous network, and have an ability to power of learning, generalization, and decision making. These networks in order to have special features such as, parallel processing. Tolerable against of error, power of learning and etc. So far, we are able to solve the problems that in view of perception and definition, have been difficult, for example, economical, medicine, engineering problems and in approximation of functions, have shown a successful.

The first application of water-nervous network in 1991 have suggested by Daniel. After 1991, this proposal have applied for anticipating of discharge in various basins. In this research , ANN for modeling of relations between physical features of basin and the amount of discharge with periods of return of various teaching is that process that don't need a complete definition of internal relationships (Physical process) between physical features in basin and amount discharge, so that these relations during of learning processes will be recognized.

2. Materials and Methods

1-2. Multiple Regression Model

Multiple regression model is a new model that used in this research have defined as a nonlinear model $Q_{T_r} = \alpha T_r^{\beta}$ that Q_{T_r} is a discharge for periods of different return, T_r is a time of return and α, β is a stable coefficient of equations that in terms of physiography data of basins and climate have been defined

$$\begin{aligned}\alpha = a_0 + a_1 A + a_2 P + a_3 L_r + a_4 S_r \\ + a_5 S_b + a_6 F_g + a_7 F_b + a_8 T_c \\ + a_9 H + a_{10} R_2\end{aligned} \qquad (1)$$

$$\begin{aligned}\beta = b_0 + b_1 A + b_2 P + b_3 L_r + b_4 S_r \\ + b_5 S_b + b_6 F_g + b_7 F_b + b_8 T_c \\ + b_9 H + b_{10} R_2\end{aligned} \qquad (2)$$

that A is an area, P is environment, L_r is a length of flood way, S_b is a middle of coefficient of basin, S_r is middle of coefficient flood way, F_g is a Gravilius coefficient, F_b shape coefficient, T_c is the time of concentration, R_2 is two years rain fall, H is an altitude of average the basin, and. a_i, b_i are a stable coefficients of equations.

2.2. Multilayer Perceptron Nervous Network

The (Fig. 1) shows a MLP of network structure that have a three layers interance layer, secret layer, and, outlet layer. So from this name of network is known, this network could have been one or several secret layer. MLP network is a type of received networks that each neuron, receive its enterances from out of network or from previous layer and own outputs towards of another layer or from outlet of network have sent. Output of each neuron as a weight of every inputs and passing them from a stimuli function have calculated. In this type of networks the direction of connection is toward of front and closed way isn't in this network.

2.3. Elman nervous network

Elman nervous network is a type of feedback networks. In feedback networks, connections are in direction that next layers have been connected to preview layers. In other words, in feedback networks, the path is closed. Elman network, is a three layers network with state of feedback, is from outlet of secret layer to inlet of secret layer. Secret layers are as a set of inlets, so, the methods of learning diffiusion back and standards of mistake, can be used in this network, but for another feedback networks isn't possible. Teaching(instruction) of Elman network, on the basis of new inlets and outlets in plus of stored quantity in secret layer with a time step of delay is accomplished. Elman network merely need a secret layer that sufficient neurons have been[4].

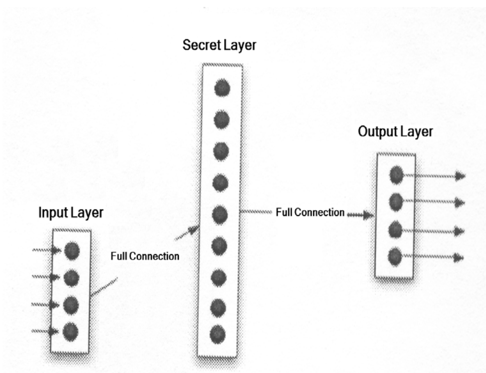

Fig. 1. Multilayer Perceptron Structure

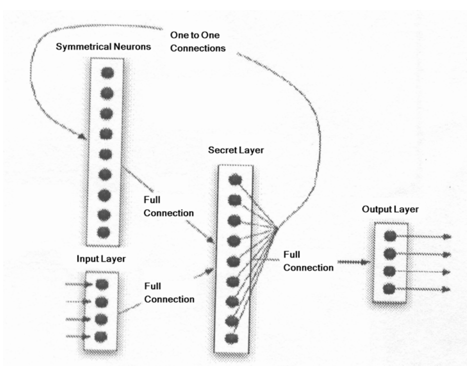

Fig. 2. Elman Network Structure

3. Analysis, argument and survey

1.3. Regression model

The best variables have been determined on basis of the largest correlation coefficient, according to table (1) after of calculating of α, β quantities of each station with Excel software and Linest function, with applications of one by one of variables in model and calculations of coefficient correlations between observable and calculating discharges.

Just as considered, in south basin and state 3 the presence of shape coefficient of variable inside of Granilus Coefficient and in west basin, state 5, the presence of middle gradient basin inside of middle gradient of flood way, is the same as a index with other form. Therefor, in western basin of south Caspian sea, 4 state and in eastern basin of south of Caspian sea, state 2, as the best states with minimum data and excellent precise, have proposed.

Table 1. Best selected variables in linear regression model

Basins	East basin of the south of the Caspian sea	
State	Best Variables	Correlation Coefficient
1	A	0.874
2	$A - F_g$	0.901
3	$A - F_g - F_b$	0.925
4	$A - F_g - F_b - R_2$	0.930
5	$A - F_g - F_b - R_2 - S_r$	0.948
6	$A - F_g - F_b - R_2 - S_r - L_r$	0.974
Basins	West basin of the south of the Caspian sea	
State	Best Variable	Correlation Coefficient
1	F_g	0.712
2	$F_b - A$	0.838
3	$F_b - A - S_r$	0.886
4	$F_b - A - S_r - H$	0.936
5	$F_b - A - S_r - H - S_b$	0.949
6	$F_b - A - S_r - H - S_b - L_r$	0.953

3.2. Multi layers perspetron of Nervous Network

Selected variables of regression model, because of application of stimuli function of tangent sigmoid in secret layer, with use of MATLAB software, according to below relationship in limited [-1, 1] have been standardized.

$$P_n = \frac{2\left(P_i - P_{\min}\right)}{\left(P_{\max} - P_{\min}\right)} - 1 \qquad (3)$$

That P_i is inlet of i'th, $P_{\min}$ is a minimum of inlets, $P_{\max}$ is maximum inlets, and P_n are standardized inlets, in order to improve of property of generalizing network, entire data; have divided to three instructional set (%75 entire data), evaluating (%33 educational data) and a test (%25 Entire data).

For teaching (instruction) of this network have used 12 instructional algorithms that in briefly: *cgp, cgf, rp, gdx, gda, gdm, gd, oss, bfg, scg, cgb* and *lm* . After instruction stage, network with data from exit of bounded, educational data have been tested, and by using of regressional analysis, coefficient correlations between observational and calculating discharges have been calculated. Finally, the optimum amount of secret neurons, in each algorithm and each state, in terms of the largest coefficient correlation in result of regression analysis in according to table (2) and table (3) have determined as have shown. Algorithms *bfg* and *lm* that are the best algorithms eastern basins in south of Caspian sea and western basin of Caspian sea, with the largest coefficient correlations as average is parallel to 0.973, 0.916 in ration to other algorithm, and each state the better results from regression mode have shown.

Algorithms *gd, gdm, gda* and *gdx* are the weakest algorithms in MLP network. The largest coefficient correlations in average algorithms in every state in eastern basin in the south of Caspian sea is parallel to 0.941 in 6 state is acquired that doesn't have much differences with 2 state and the quantity of 0.933. So, in this basin the state 2 can be suggested as the best state. In basin as small as south Caspian sea, the largest coefficient correlation, average of algorithms in each state is parallel to 0.933 that in four state has acquired so, in this basin the four state is similar to regression model as the best state have suggested.

3.3. Elman Network

Instruction of Elman network is completely like the instruction of multilayer prespetron network, however, with the difference that merely from *gdx* algorithm for instruction have used. The results from this network have presented in tables (2) and (3). Just as you observed, in this network unlike multilayer percepetron network, west basin in the best state, state 6 was 0.906 coefficient correlations that has a great difference with state 4 and 0.888 quantities, and the save as multi layer prespetron network eastern basin from south Caspian sea, with the best state, that is, state 6, with this difference that in this network correlation coefficient of state 6 (0.936) is very much from state 2 (0.899).

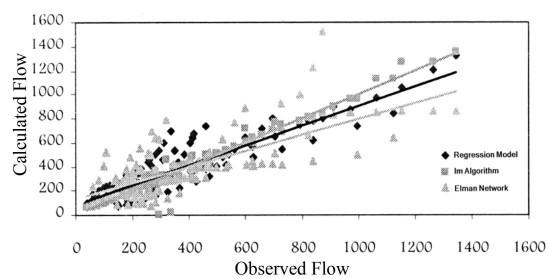

Fig. 3. Comparing the calculated flow from *lm* Algorithm by Elman Model and east basin of south of the Caspian sea regression in state 2

In general Elman network have a weak performance in relative to multilayer perceptron and regression model in two basins. In comparison with the quantity of frequently networks, can find the speed of little convergence of Elman network in ration of MLP network. So, Elman network needs a lot of secret neurons in ration MLP network. In Fig. 3 and Fig. 4 Comparable graph of the best algorithms of MLP with Elman network, and regression model in proposed states in these basins have been drawn.

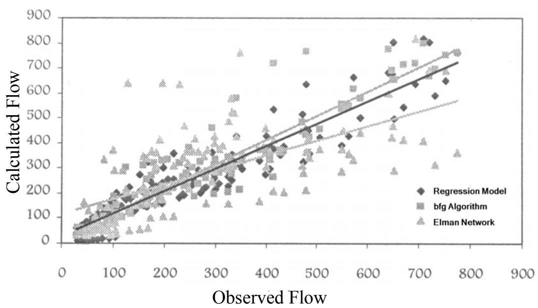

Fig. 4. Comparing the calculated flow from *bfg* Algorithm by Elman Model and west basin of south of the Caspian sea regression in state 4

4. Conclusions

New regression model to estimate of flood way discharges have shown an acceptable performance by increasing of variables this model, coefficient been increased. In eastern basin of south Caspian sea, the shape of under basins and in western basin of south Caspian sea, the altitude of under basins have a great influence in estimating of flood way discharges. MLP network have a better performance than regression model and Elman network and among its algorithms. Algorithms lm and bfg proposed to anticipate of flood way of two basins. In eastern basin of south of Caspian sea, estimating of discharge with high precise and with minimum data, state 2 and western basin of south of Caspian sea, state 4 have suggested. The performance of integrated model that is, regression model in determining of inlet variables of MLP network and MLP network in determining the best variables, to estimate of flood way of discharges and precise anticipating of flood way has been very successful. Elman network, because of weak performance, prosperous of speed of low convergence and need to quantity of great secret neurons, in order to anticipate of flood in western and eastern basins in south of Caspian sea have not proposed.

5. References

1.Lekkas, D.F., Onof, C., Lee, M.J., Baltas, E.A., 2004, "*Application of artificial neural networks for flood forecasting*", J. Vol. 6, No. 3, P.P. 209-205.

2. WMO, 1992, "*Simulated real-time intercomparision of hydrological models*", In: WMO operational hydrology report, No. 38, WMO No. 779, World Meteorological Organization, Geveva

3. Dawson, C. W., Abrahart, R.J., Shamsedin, A.Y., Wilby, R. L., 2006, "*Flood estimation at ungauded sites using artificial neural networks*", J. Hydrology, Vol. 319, No., P.P. 391-409

4. Demuth, H., Beale, M., 2005, "*Neural Network Toolbox User's Guide*", Version 4, The Mathworks Inc.

Table 2. Multilayer perceptron and Elman results in eastern basin of the south of the Caspian sea

Algorithm	States	gd			gdm			gda			gdx			gdx*			rp		
		R	Neuron	Iteration	R	Neuron	Iteration	R	Neuron	Iteration	R	Neuron	Iteration	R	Neuron	Iteration	R	Neuron	Iteration
Gradient Descent of Error Back-propagation	1	0.738	3	110	0.790	13	110	0.845	13	89	0.864	19	110	0.885	18	110	0.891	19	110
	2	0.803	7	110	0.835	1	110	0.930	8	105	0.906	10	110	0.899	15	110	0.965	16	108
	3	0.781	7	110	0.780	10	110	0.892	10	95	0.915	9	110	0.911	12	110	0.960	11	84
	4	0.651	7	110	0.783	7	110	0.896	11	105	0.896	8	80	0.918	9	110	0.951	11	102
	5	0.736	10	110	0.778	8	110	0.935	9	102	0.926	8	110	0.931	6	110	0.971	8	60
	6	0.75	9	110	0.879	1	110	0.960	9	110	0.917	4	110	0.936	7	110	0.975	9	47

Algorithm	States	cgf			cgb			cgp			scg		
		R	Neuron	Iteration	R	Neuron	Iteration	R	Neuron	Iteration	R	Neuron	Iteration
Conjugate Gradient	1	0.883	19	77	0.890	15	81	0.888	20	110	0.880	17	88
	2	0.961	7	70	0.971	15	59	0.964	11	110	0.966	12	110
	3	0.967	13	52	0.974	5	97	0.957	9	110	0.966	13	65
	4	0.964	10	84	0.966	4	104	0.961	4	71	0.964	9	89
	5	0.969	6	88	0.969	7	49	0.969	5	55	0.973	9	70
	6	0.975	7	66	0.973	6	58	0.969	7	41	0.974	2	108

Algorithm	States	oss			bfg		
		R	Neuron	Iteration	R	Neuron	Iteration
Semi-Newton	1	0.882	13	108	0.929	19	110
	2	0.947	16	62	0.972	12	41
	3	0.956	2	84	0.965	10	56
	4	0.960	6	70	0.974	4	68
	5	0.958	8	53	0.974	8	50
	6	0.970	7	85	0.974	3	75

Algorithm	States	lm		
		R	Neuron	Iteration
Marquart-Lounberg	1	0.937	8	55
	2	0.979	9	7
	3	0.981	7	10
	4	0.982	7	6
	5	0.979	9	6
	6	0.980	8	7

gdx* is for Elman network.

Table 3. Multilayer perceptron and Elman results in the western basin of the south of Caspian sea

Algorithm	States	gd			gdm			gda			gdx			gdx*			rp		
		R	Neuron	Iteration	R	Neuron	Iteration	R	Neuron	Iteration	R	Neuron	Iteration	R	Neuron	Iteration	R	Neuron	Iteration
Gradient Descent of Error Back-propagation	1	0.961	13	120	0.606	5	48	0.711	13	101	0.725	17	120	0.698	12	118	0.712	19	110
	2	0.664	11	120	0.687	9	18	0.841	8	110	0.856	17	120	0.788	9	85	0.891	16	108
	3	0.669	3	120	0.719	5	120	0.879	5	112	0.887	13	120	0.870	10	120	0.893	11	84
	4	0.554	1	120	0.847	9	48	0.888	5	120	0.940	9	120	0.888	10	120	0.930	11	102
	5	0.772	6	120	0.673	3	31	0.928	11	101	0.927	1	120	0.875	10	120	0.950	8	60
	6	0.657	9	120	0.658	6	120	0.910	4	105	0.787	3	120	0.906	7	120	0.903	9	47

Algorithm	States	cgf			cgb			cgp			scg		
		R	Neuron	Iteration	R	Neuron	Iteration	R	Neuron	Iteration	R	Neuron	Iteration
Conjugate Gradient	1	0.716	13	17	0.717	5	37	0.714	17	21	0.713	17	28
	2	0.935	9	63	0.913	4	43	0.937	14	40	0.907	10	48
	3	0.910	5	27	0.907	14	29	0.901	12	42	0.893	12	32
	4	0.901	9	41	0.942	9	52	0.922	9	48	0.939	9	49
	5	0.924	1	20	0.946	2	53	0.961	6	30	0.922	1	17
	6	0.929	4	46	0.939	4	43	0.957	6	69	0.924	1	31

Algorithm	States	oss			bfg		
		R	Neuron	Iteration	R	Neuron	Iteration
Semi-Newton	1	0.716	16	32	0.731	5	68
	2	0.924	13	41	0.931	13	52
	3	0.918	13	77	0.947	11	44
	4	0.936	10	58	0.945	12	31
	5	0.944	6	91	0.973	9	34
	6	0.942	8	50	0.972	7	40

Algorithm	States	lm		
		R	Neuron	Iteration
Marquart-Lounberg	1	0.723	7	18
	2	0.947	4	36
	3	0.937	8	12
	4	0.951	5	9
	5	0.937	4	40
	6	0.966	3	8

gdx* is for Elman network.