Maximum design flood discharge by statistical downscaling data for the spillway of Surqawshan dam

By

Asos Kaikhasraw Abdulkarem

Table of content:

CONTENTS **Page**

- Abstract ………………………………………………………… 1
- INTRODUCTION …………………………………………… 2
- METHODOLOGY ………………………………………. 5
- RESULT AND DISCUSSION …………………………... 9
- CONCLUSIONS AND RECOMMENDATIONS ……… 21
- References ………………………………………………. 22

Abstract:

Design flood discharge of spillway of dams is one of the important criteria for dam safety analysis and one of the main principles that the designer engineer depends on during design of dams. There is a great variety of methods used around the world to estimate of design flood discharge. In this research, new design flood discharge for spillway of dams by using statistical downscaling data and general circulation method was investigated and compare new design flood discharge with historical design flood discharge were calculated from the observed historical data, using soil conservation service (SCS) for checking safety of dam. Surqawshan Dam was used as a case study. The dam is located in Sulaimania Governorate. The dam is approximately 65 km northwest of Sulaimania City. It is under construction on the Chamyrezan valley. Surqawshan Dam is a rock-fill dam. Linear genetic programming was worked on the maximum rainfall data be taken from a gauge station nearby Surqawshan Dam and also the downscaled input variables. Eight LGP models were developed. The overall results show that using Rmax by its reciprocal gives better results rather than taking its natural logarithm (Ln (Rmax)). The result shows that Surqawshan Dam is under safety since the design flood discharge, Q_{10000} was obtained to be 987 m³/s although it was computed as 1091 m³/s by using the observed peak discharges. The underlying reason for this result is that, according to the downscaled rainfall data, there is a decreasing trend in annual maximum rainfall data value.

Keywords: Rain-fall data, design flood discharge, statistical downscaling, General circulation model, Linear genetic programming, Soil Conservation Service.

INTRODUCTION

Water is one of the essential thing for our life and renewable matter which has directly and indirectly effects on the life of humanity. In the past human was searching for a place near sources of fresh water like rivers, and this case important up to date. They construct dams for multi-purpose such as flood control, irrigation, electric power production, etc…

The determination of design flood discharges is one of the important key for the design of hydraulic structures like spillway of the dam and to determine the dimension of spillway. Maximum flood discharge depending on many factors like topography, land type, catchment area (watershed), rainfall data and various other factors. Climate change effects of proposed actions and design projects, especially the design of a hydraulic structures such as spillway of dams which is a structure used to transfer flood water and normal water coming from upstream of the dam to downstream of the dam in order to prevent overflowing water over the dam body, causing damage, risk, collapsing and create disaster in this place or this town.

The goal of this study is to estimate the design flood based on climate general circulation method and to compare with the value of the design floods that were calculated from the observed historical data. In this research, Surqawshan Dam was chosen as case study.

There are various probable reasons of dam failure, like structural, seismic geologic and hydrologic. Hydrological failure occurs when the designer estimate and use less value of design flood discharge if compare with that water coming during flood, emergency spillway or bottom outlet is provided to help the service (main) spillway during flood season and increase freeboard of the dam with considering water force on the dam body.

In general, for this case earth dam type is more dangerous than concrete dam because of erosion and scour when occurring on the dam body during rainfall and flood may reason partial or full failure of the dam. As well as in concrete dam failure may occur due to seepage of water through abutments and foundations.

When spillway capacity is increases, the cost of building spillway is directly increased, but the potential of overflowing failure decreases, this technique is known as "cost to save a life and infrastructure".

It is necessary to calculate the design flood discharges by more than one method and then make an engineering judgment to use the maximum value that was determined.

Around the world Climate change in different ways happened, scientists have become concerned about global warming, due to humanity effect on the climate system, through the increases of the natural greenhouse impact and human activity.

Few change in weather and climate may have the probable to bring significant increases in damages to existing infrastructure.

A global climate models also known as a general circulation models, both term are abbreviated as GCM are programs of computer that made of more than hundred thousand lines of code. They calculate the interactions between the land, atmosphere and ocean using factors such as heat, carbon dioxide, water vapor, and the Earth's rotation as inputs. Global climate models are used for forecasting climate change, understanding the climate and forecasting weather.

Statistical downscaling, is a statistical relationship between the output data of the climate model and the historic observed climate data for the same historical period.

In artificial intelligence, genetic programming (GP) is a process of evolving programs, starting from a number of unfit (usually random) programs, fit for a particular task by applying operations analogous to natural genetic processes to the population of programs. This study focused on Linear Genetic Programming (LGP).

Artificial intelligence (AI) refers to computer systems able of analysis hard and complex tasks that historically only a human could do, such as reasoning, making decisions, or solving problems in few time.

Surqawshan Dam one of the dam is under construction is selected as a case study. Linear genetic programming (LGP) is used to estimate new flood design discharge based on statistical downscaling variable data provided from Coupled Global Climate Model A2 scenario (CGCM2A) and to compare with that flood design discharge based on observed historical rainfall data taking from a metrological station. The design flood discharge is calculated by using soil conservation service (SCS) method. The scope of this work is to check whether the spillway of Surqawshan Dam is in safety or not, (capable of transferring maximum flood discharge or not).

METHODOLOGY

Statistical downscaling, is a statistical relationship between the output data of the climate model and the historic observed climate data for the same historical period.statistical downscaling contain two component predictors and predictands, where predictors are the input data and the predictand is the output data.

Predictand =f (predictors).

Figure 1.1 provides the relationship between the predictand and the predictor.

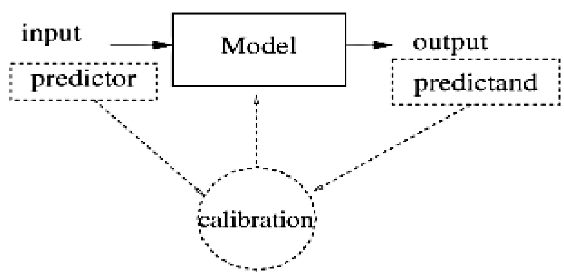


Figure 1.1. A schematic illustration between predictand and predictor

The characterise value groups are used as input data (predictor) taking from the Canadian Centre for Climate Modeling and Analysis. In our research , the box 13X-14Y was downloaded for the location of Surqawshan dam.

At first the rainfall data taked from metrological gauge station of Dukan town and Secondly the downscaled rainfall data downloaded from CGCM3A2. Using linear genetic programming for this work.

Parameter used to find design flood discharge by SCS method is described as below:

$$Qp = \frac{0.207 * A * Q}{tp}$$

where Qp is peak discharge (m^3/s), A is catchment area in km^2, Q= daily Run-off from the catchment (mm) and tp is time to peak.

tp = 0.7 tc

tc is concentration time in minute.

tc = 14.6 *L*A - 0.1 *S′- 0.2

where L is length of the mainstream from the starting point to dam site in km, S′=slope of the mainstream in the catchment area.

$$Q = \frac{(P - 0.2\ S)}{(p + 0.8\ S)}$$

where Q is daily runoff from catchment (mm).

$$S = \left(\frac{25400}{CN} - 254\right)$$

where CN is a curve number of runoff that is a function of antecedent soil moisture, land use and other factors affecting runoff in a catchment. A curve number is a dimensionless number defined such that $0 \leq CN \leq 100$. For impervious and water surface CN=100, for natural surface $CN \leq 100$.

The range of antecedent moisture conditions for each class is shown in Table 1

Table (1) Runoff curve numbers (Normal watershed condition).

Description of Land Use	Hydrology Group of Soil			
	A	B	C	D
Cultivated land :				
Without conservation treatment	72	81	88	91
With conservation treatment	62	71	78	81
Range land or pasture:				
Poor Cover	68	79	86	89
Good Cover	39	61	74	80
Good Condition, Meadow	30	58	71	78
Woods or forest land:				
poor cover, thin stand	45	66	77	83
Good Cover	25	55	70	77
Farmsteads	59	74	82	86
Roads	74	84	90	92

Where each group is defined as:

Group A:, Deep loess, aggregated silts and deep sand.

Group B: sandy loam, Shallow loess.

Group C: Shallow sandy loam, Clay loam, soil low in organic content, and soil usually high clay.

Group D: Soil that swells significantly when wet, heavy plastic clays, and certain saline.

$$P = p' + Sd * Kt$$

where P' is maximum average daily precipitation during this month, Sd is standard deviation of the maximum yearly rainfall data

$Kt = (\frac{yt - y'n}{sn'})$

where Kt is the frequency factor, yt is reduced mean, a function of sample size N and is given in tables, y´n is reduced standard deviation, a function of sample size N

$Yt = -(\ln(\ln(\frac{Tr}{Tr-1})))$

where Tr is return period.

availability of rainfall data from (1961-2012).

Table 2 Daily rainfall data collected from Dukan dam.

Type of data	Available	missing
Daily Rainfall	1/1/1961 to 31/12/2012	From (March to JUN) 1974, January 1980 and April 1991.

For available rainfall data from 1961 to 2012 draw a chart to Maximum annual rainfall as shown in figure (4.2)

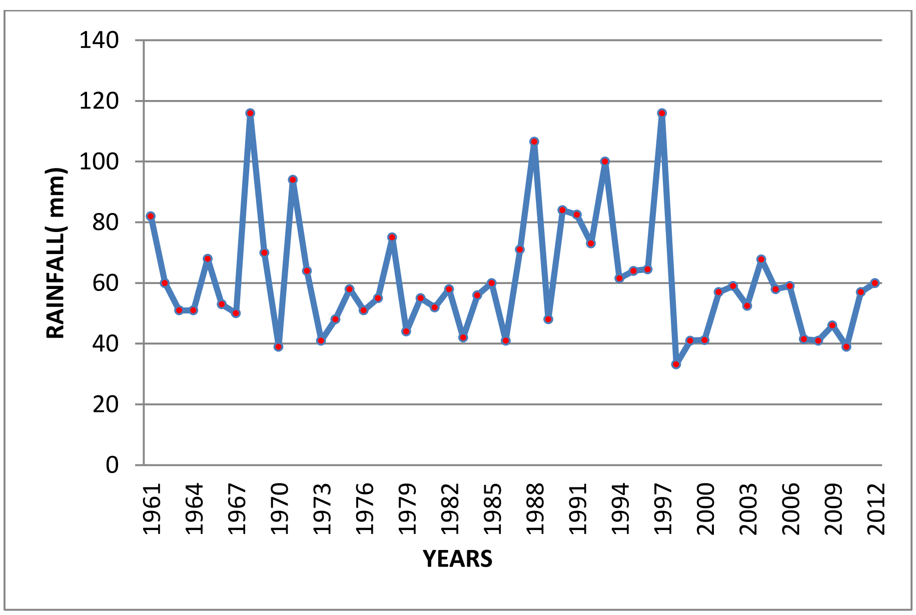


Figure 2 Maximum yearly rainfalls in ferry gauge of Dukan da

RESULT AND DISCUSSION

The maximum monthly rainfall (R_{max}) is used in all LGP models only the rainy months, January, February, March, April, May, October, November and December for each year. In LGP2, LGP3, LG4, LGP6, LGP7 and LGP8 models, take reciprocal R_{max} data range it between 0.01 to 2 (see Table 3). This result has a good correlation between CGM3 data input sets and R_{max} . In LGP1, LGP5 models, natural logarithm of R_{max} is used.

Table 3 shows the correlation between CGM3 variables 12 input sets with daily rainfall

CGM Variables	LN (R_{max}(t))	1/(R_{max} (t-2))	1/(R_{max} (t-1))	1/ (R_{max}(t))
D1	0.132	-0.130	-0.122	-0.128
D2	-0.118	0.119	0.123	0.115
D3	-0.101	0.163	0.153	0.114
D4	0.103	-0.101	-0.106	-0.104
D5	0.130	-0.089	-0.104	-0.123
D6	-0.134	0.098	0.110	0.135
D7	0.112	-0.067	-0.078	-0.109
D8	-0.336	0.259	0.262	0.266
D9	-0.289	0.160	0.181	0.208
D10	-0.312	0.182	0.173	0.173
D11	-0.355	0.218	0.210	0.209
D12	-0.373	0.248	0.245	0.245

Table 4 Correlation matrix among the CGM3 and output variables.

CGCM3 Variables	LN (R_{max}(t))	1/(R_{max} (t-2))	1/(R_{max} (t-1))	1/ (R_{max}(t))
D1	0.157	-0.081	-0.083	-0.083
D2	0.069	-0.087	-0.107	-0.074
D3	0.132	-0.130	-0.122	-0.128
D4	-0.118	0.119	0.123	0.115
D5	-0.084	0.025	0.024	0.026
D6	-0.101	0.163	0.153	0.114
D7	0.103	-0.101	-0.106	-0.104
D8	-0.058	0.030	0.011	0.002
D9	-0.059	0.049	0.034	0.033
D10	-0.111	0.038	0.043	0.064
D11	0.130	-0.089	-0.104	-0.123
D12	0.078	-0.049	-0.051	-0.073
D13	0.105	-0.033	-0.037	-0.054
D14	0.020	-0.079	-0.085	-0.030
D15	0.088	-0.092	-0.089	-0.092
D16	-0.134	0.098	0.110	0.135
D17	-0.057	0.017	0.014	0.023
D18	-0.011	0.064	0.010	0.027
D19	0.112	-0.067	-0.078	-0.109
D20	-0.336	0.259	0.262	0.266
D21	-0.045	0.082	0.076	0.068
D22	-0.061	0.051	0.071	0.075
D23	-0.289	0.160	0.181	0.208
D24	-0.312	0.182	0.173	0.173
D25	-0.355	0.218	0.210	0.209
D26	-0.373	0.248	0.245	0.245

Table 5 Correlation among CGCM variables input sets involved in LGP6 model.

	d3	*d4*	*d6*	*d7*	*d11*	*d16*	*d19*	*d20*	*d23*	*d24*	*d25*	*d26*	*d27*
d3	1.00												
d4	-0.38	1.00											
d6	-0.77	0.54	1.00										
d7	0.30	-0.97	-0.48	1.00									
d11	-0.07	-0.41	-0.04	0.41	1.00								
d16	-0.37	0.85	0.45	-0.85	-0.28	1.00							
d19	0.25	-0.78	-0.35	0.81	0.29	-0.96	1.00						
d20	-0.16	0.49	0.26	-0.48	-0.58	0.37	-0.35	1.00					
d23	0.05	0.31	-0.01	-0.33	-0.30	0.34	-0.36	0.36	1.00				
d24	-0.11	0.35	0.21	-0.31	-0.10	0.27	-0.25	0.47	0.57	1.00			
d25	-0.07	0.36	0.19	-0.33	-0.17	0.31	-0.28	0.58	0.56	0.94	1.00		
d26	0.01	0.37	0.12	-0.35	-0.24	0.33	-0.32	0.66	0.51	0.82	0.93	1.00	
d27	-0.13	0.12	0.11	-0.10	-0.12	0.14	-0.11	0.27	0.21	0.17	0.21	0.25	1.00

Table 6 Statistical showing of LGP models for testing and training sets.

LGP MODEL	Phase	R2	RMSE	Data	Standard Deviation	variance	max	min	mean
LGP1	Training	0.216	0.802	Observed	0.9	0.81	4.75	-0.69	3.16
				Predicted	0.42	0.18	3.84	1.53	3.06
	Testing	0.214	1.003	Observed	1.12	1.26	4.22	-0.69	2.79
				Predicted	0.45	0.2	4.01	1.58	2.92
LGP2	Training	0.701	0.086	Observed	0.14	0.02	2	0.009	0.074
				Predicted	0.14	0.019	2.09	-0.1	0.1
	Testing	0.456	0.207	Observed	0.28	0.078	2	0.015	0.143
				Predicted	0.214	0.046	2.11	-0.07	0.13
LGP3	Training	0.646	0.088	Observed	0.14	0.02	2	0.009	0.074
				Predicted	0.129	0.017	1.96	-0.22	0.08
	Testing	0.182	0.260	Observed	0.28	0.078	2	0.015	0.143
				Predicted	0.097	0.01	0.49	-0.39	0.09
LGP4	Training	0.639	0.091	Observed	0.14	0.02	2	0.009	0.074
				Predicted	0.09	0.007	1.42	-0.03	0.07
	Testing	0.497	0.215	Observed	0.28	0.078	2	0.015	0.143
				Predicted	0.27	0.072	2.65	-0.09	0.1
LGP5	Training	0.247	0.793	Observed	0.9	0.81	4.75	-0.69	3.16
				Predicted	0.56	0.31	4.31	1.12	3.07
	Testing	0.245	0.987	Observed	1.12	1.26	4.22	-0.69	2.79
				Predicted	0.56	0.32	3.98	0.44	2.96
LGP6	Training	0.629	0.093	Observed	0.14	0.02	2	0.009	0.074
				Predicted	0.13	0.02	1.9	0	0.1
	Testing	0.145	0.263	Observed	0.28	0.078	2	0.015	0.143
				Predicted	0.067	0.005	0.27	0	0.11
LGP7	Training	0.642	0.092	Observed	0.14	0.02	2	0.009	0.074
				Predicted	0.14	0.02	2.13	0	0.09
	Testing	0.475	0.209	Observed	0.28	0.078	2	0.015	0.143
				Predicted	0.25	0.06	2.38	0	0.12
LGP8	Training	0.831	0.063	Observed	0.15	0.02	2	0.01	0.072
				Predicted	0.14	0.02	2.02	0.01	0.08
	Testing	0.721	0.094	Observed	0.12	0.013	0.5	0.01	0.082
				Predicted	0.03	0.001	0.22	0.03	0.06

in following sections of this research, LGP8 model is used in modeling studies.

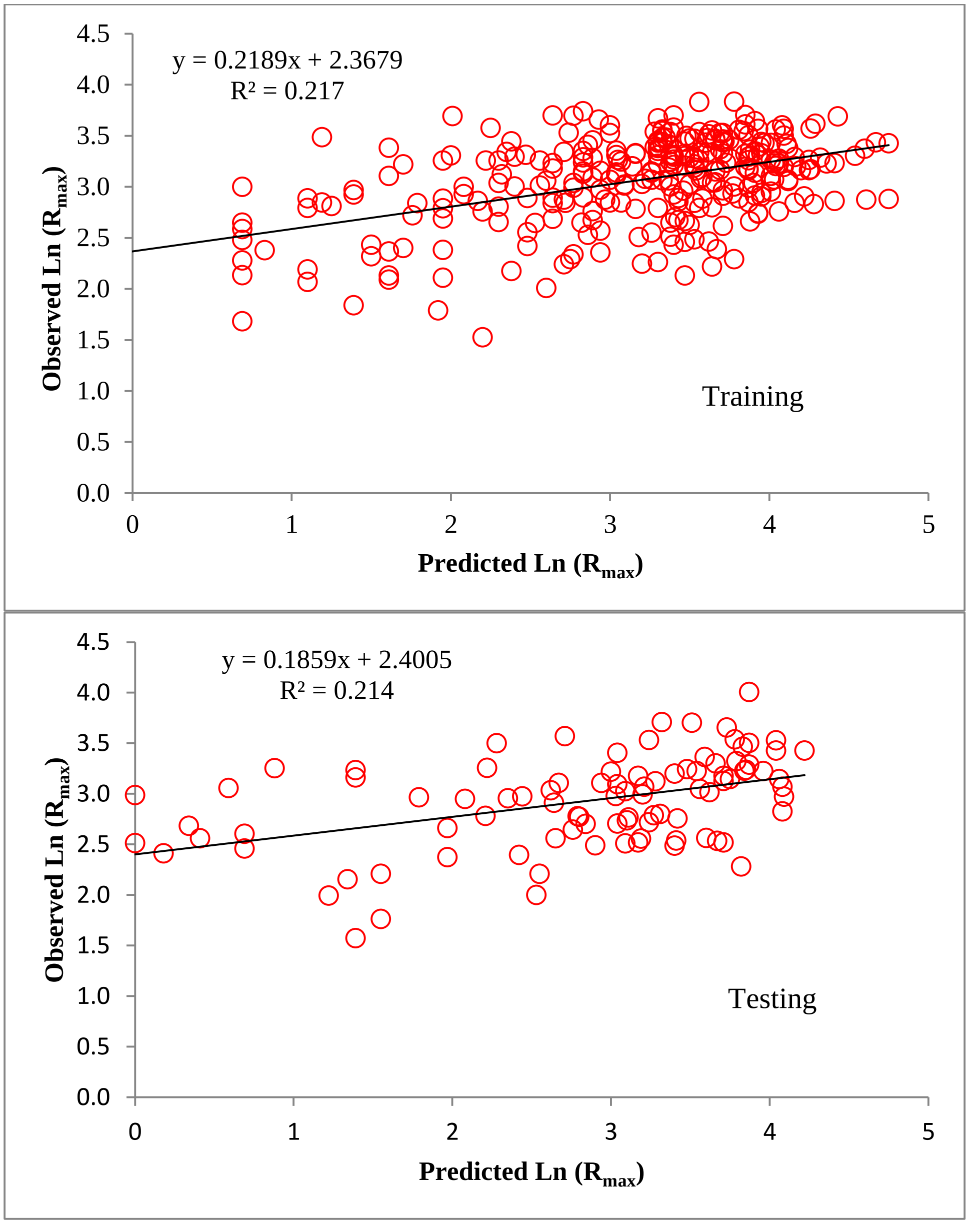


Figure 3 Scatter plot of LGP1.

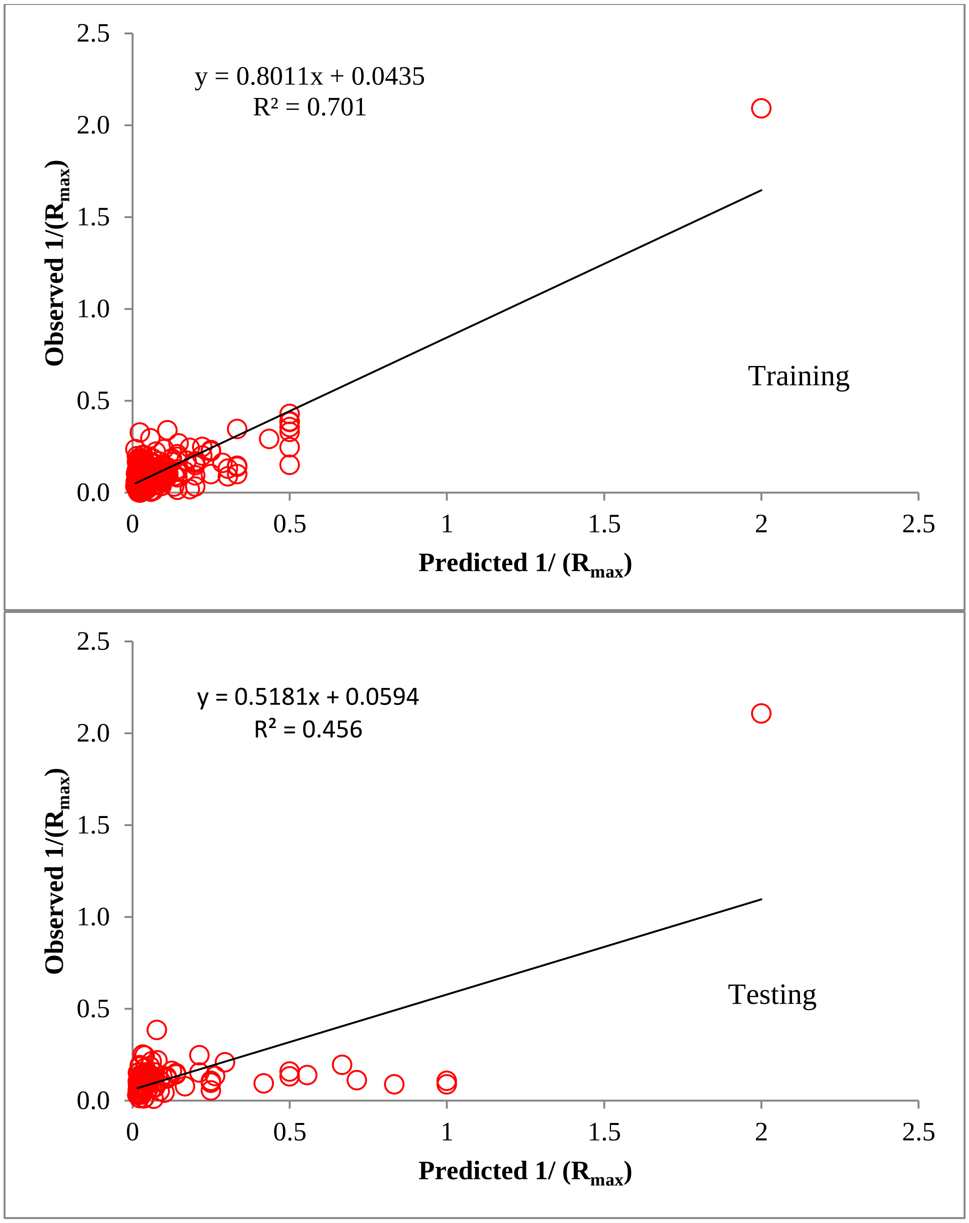


Figure 4 Scatter plot of LGP2.

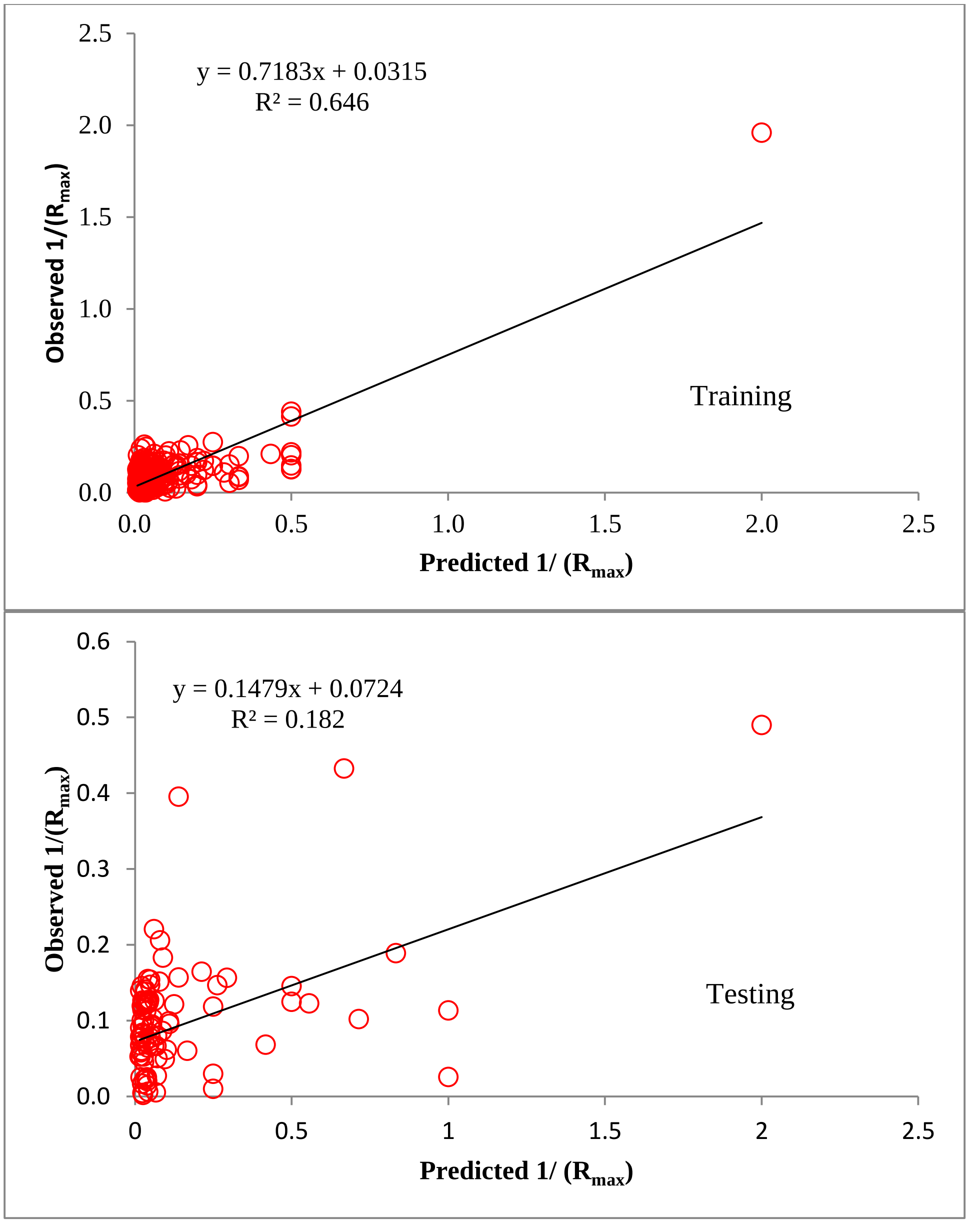
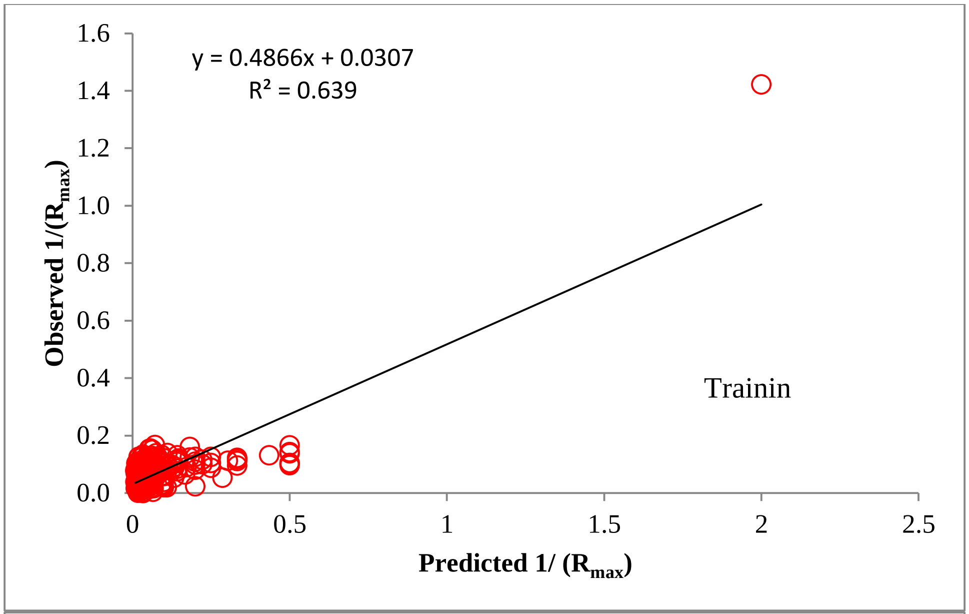


Figure 5 Scatter plot of LGP3.



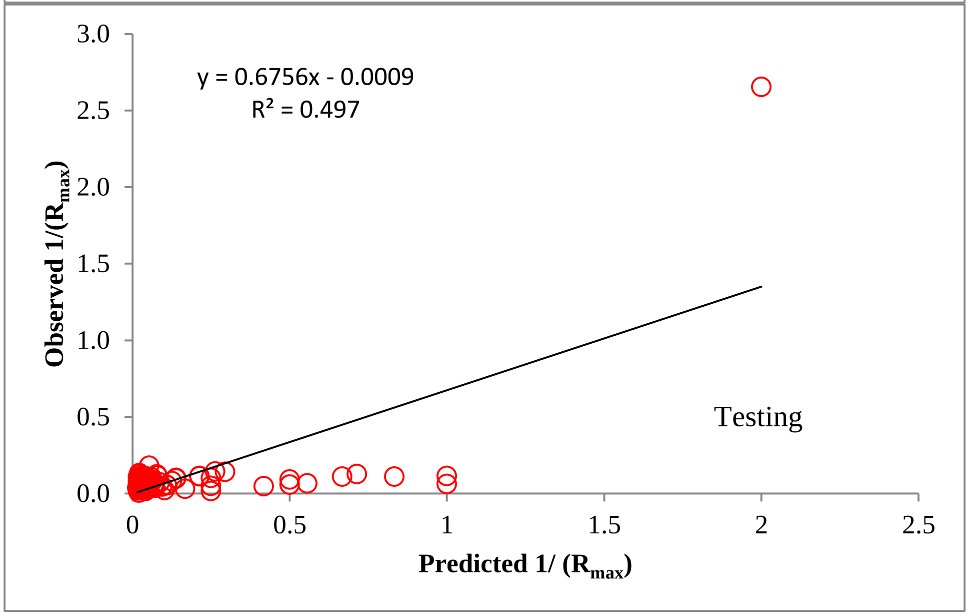


Figure 6 Scatter plot of LGP4.

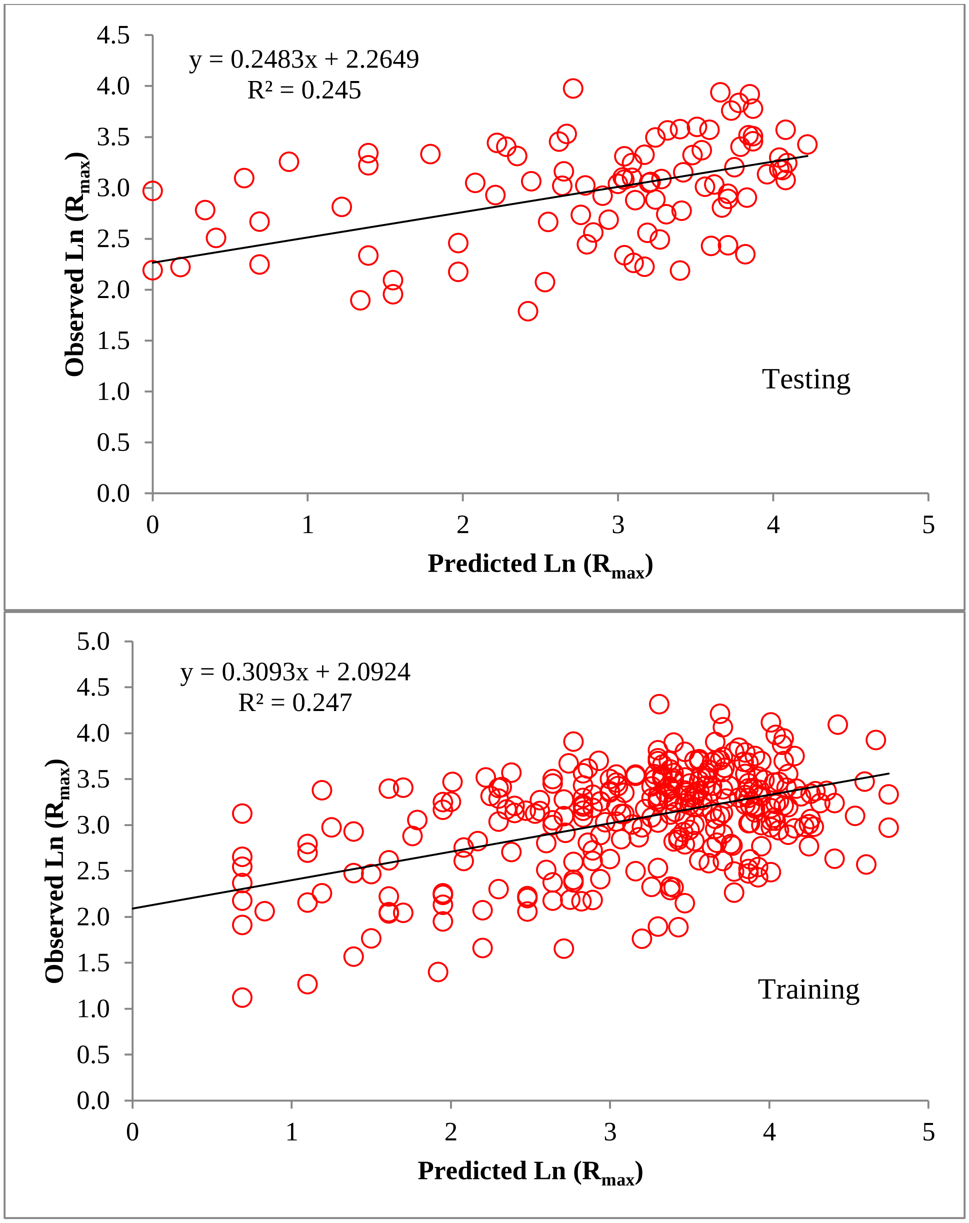


Figure 7 Scatter plot of LGP5.

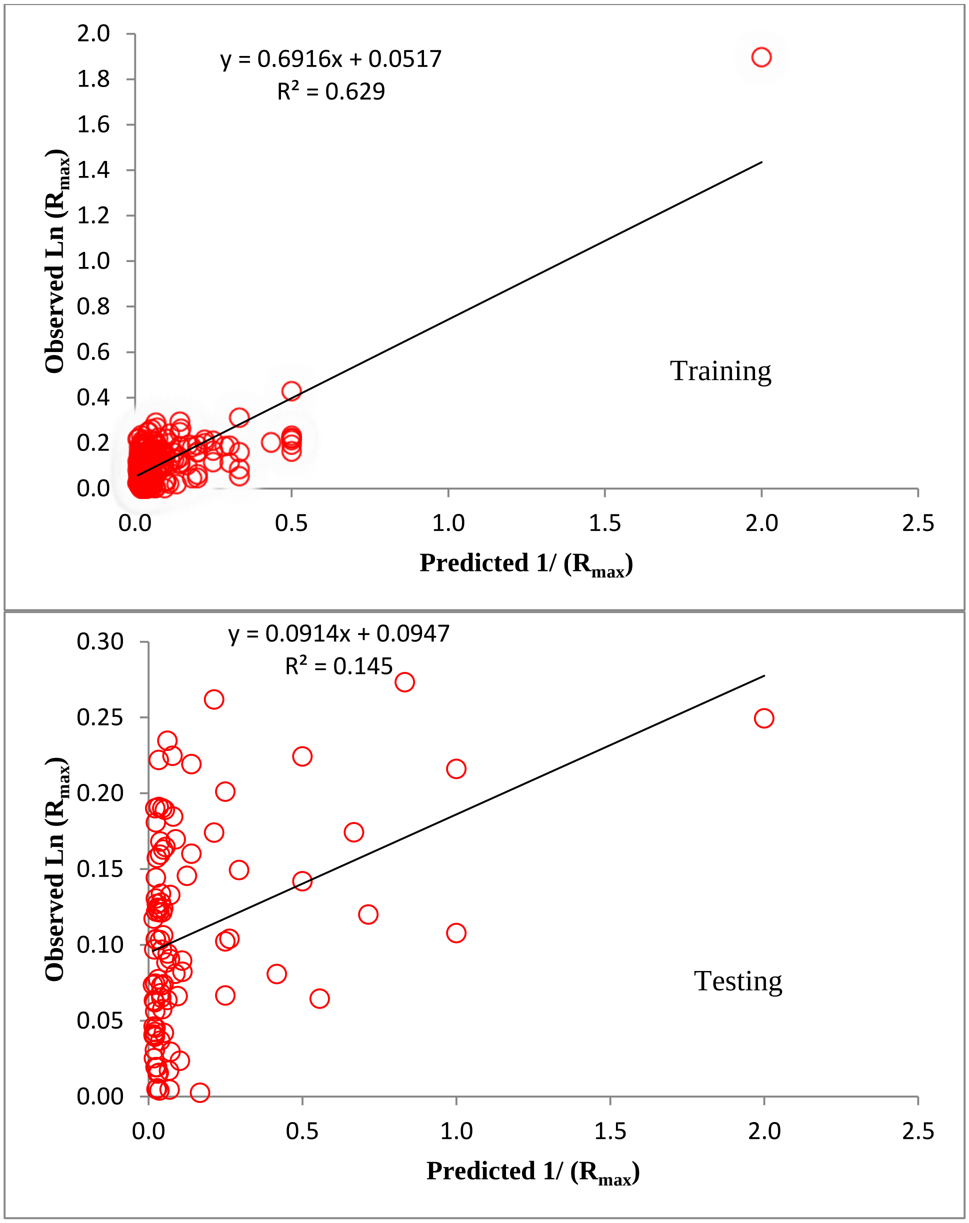


Figure 8 Scatter plot of LGP6.

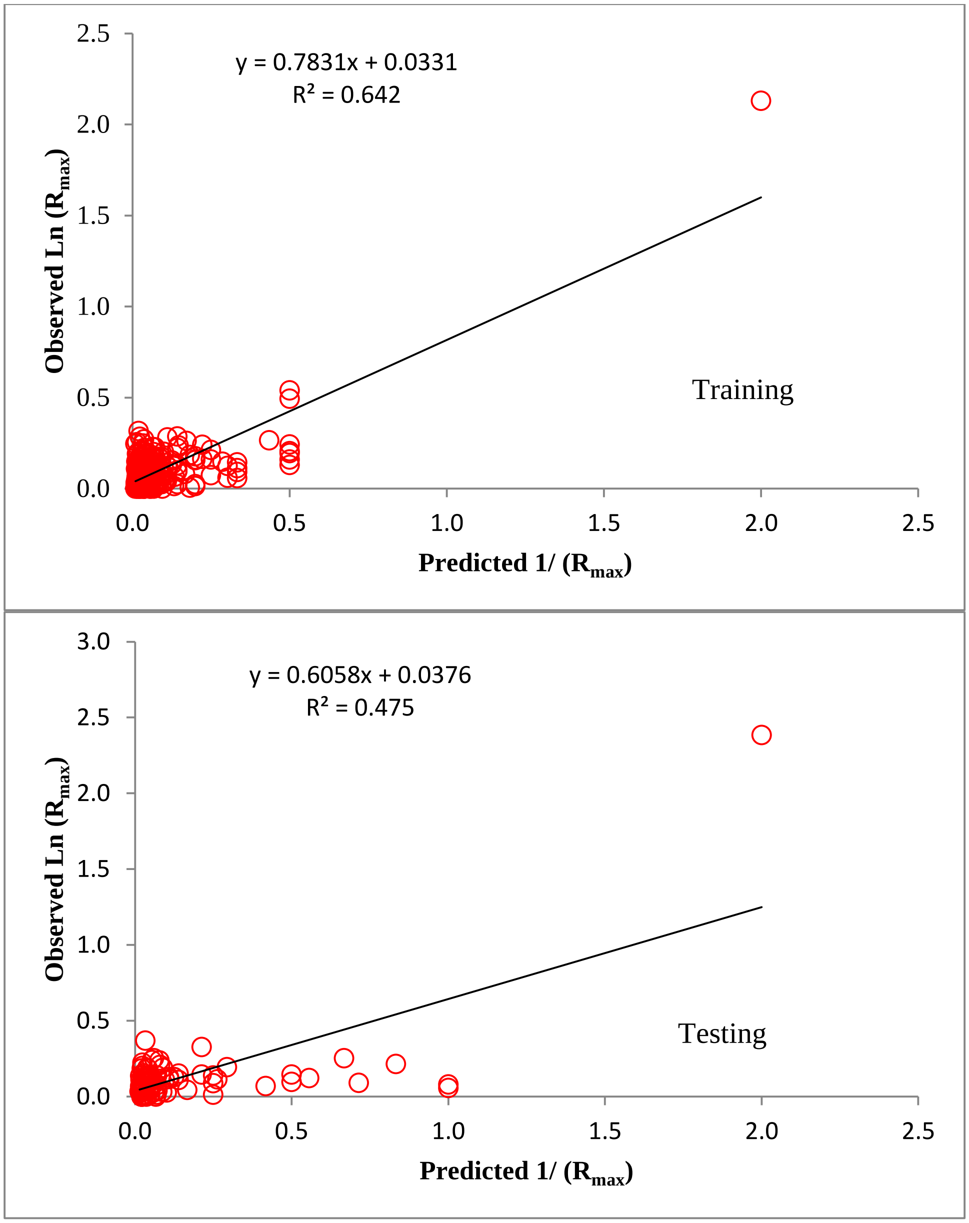


Figure 9 Scatter plot of LGP7.

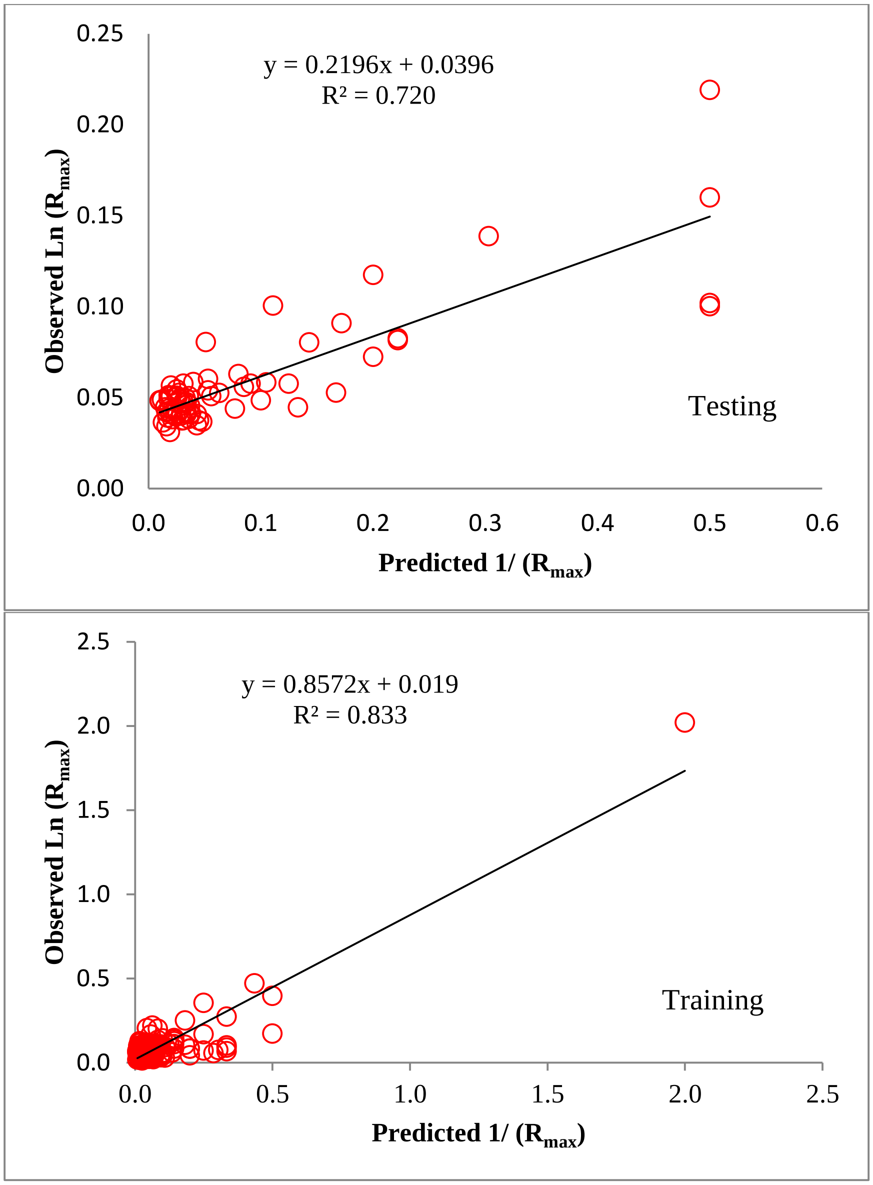


Figure 10 Scatter plot of LGP8.

Select 25% of the data randomly as a (testing) and 75% as (training) set and done the process by using LGP to get the best relationship between (predictors) and (predictand). The output of data of LGP8 model with highest R^2 and the lowest mean square error (MSE) is used to find the peak discharge (Q_{design}) by SCS method for spillway of Surqawshan Dam as shown in Table (8).

The SCS method is adopted using (24-hr) maximum rainfall data and frequency analyses of Gumble distribution. The maximum 24hr rainfall values for return periods of 5, 10,100,200,500, 1000 and 10000 years were obtained for Surqawshan basin. The results of Qp calculation are given in Table 7 and Table 8.

Table 7. shows the calculation of return period of peak discharges by using SCS method. It shows that the design discharge is calculated as 1091 m³/s. It was done by using maximum monthly rainfall data taken from the directorate of Dukan dam.

Table 8. shows the calculation of return period of peak discharges by using SCS method. It shows that the design discharge is calculated as 987 m³/s by using the downscaled rainfall data taken from linear genetic LGP8 model.

Table 7. Result of peak discharge calculation by SCS method from 24 hr max rainfall

TR	5	10	100	200	500	1000	10000
YT	**1.50**	**2.25**	**4.60**	**5.30**	**6.21**	**6.91**	**9.21**
KT	**0.84**	**1.50**	**3.55**	**4.16**	**4.97**	**5.58**	**7.59**
P(MM)	**81**	**94**	**138**	**150**	**167**	**180**	**222**
Q(MM)	**21**	**29**	**60**	**70**	**83**	**94**	**130**
Q(M³/SEC)	**173**	**244**	**501**	**584**	**698**	**787**	**1091**

Table 8. Result of peak discharge calculation by SCS method from 24 hr max rainfall (downscaled)

TR	5	10	100	200	500	1000	10000
YT	**1.50**	**2.25**	**4.60**	**5.30**	**6.21**	**6.91**	**9.21**
KT	**0.84**	**1.50**	**3.55**	**4.16**	**4.97**	**5.58**	**7.59**
P(MM)	**74**	**91**	**122**	**137**	**149**	**178**	**208**
Q(MM)	**17**	**27**	**48**	**59**	**69**	**92**	**118**
Q(M³/SEC)	**142**	**226**	**404**	**498**	**576**	**773**	**987**

CONCLUSIONS AND RECOMMENDATIONS

Conclusions:

In this research, using eight linear genetic programming models, linear genetic programming (LGP) was worked on the input-output data set to find a relationship between the rainfall data and the downscaling variables.

Among all result of (LGP) models, the LGP8 models seem to be the best were R^2=0.831 and MSE=0.063. The overall results show that normalizing Rmax by its reciprocal gives better results rather than taking it′s natural logarithm (Ln (R_{max})). The soil conversation method (SCS) used to determine design flood discharges of Surqawshan Dam. According to the result of statistical downscaling, Surqawshan Dam is under safety since the design flood discharge, Q10000 was equal 987 m^3/s but it was computed as 1091 m^3/s by using the observed peak discharges. The underlying reason for this result is that, according to the downscaled rainfall data, there is a decreasing in annual maximum rainfall data value.

Recommendations:

1- In this research, this way used to find design flood discharge for spillway of dams only for checking safety of the dam not consider like a new method .These process can be used on existing dam or proposed dam.

2- The value of design flood discharges for spillway of the dam should calculate by more than one method and compares the results for more safety of the dam.

3- Risk of climate change should consider during measure or design hydraulic structures.

4- It is possible to make model by small scale for any dam before construction in hydraulic lab to test capacity of spillway (capable for this design flood calculated or not).

5- Early warning system and construction structural for retrofitting can apply on existing dam . A mechanism is known as "cost to save a life and infrastructure ".

References:

- Fang, G., Yang, J., Chen, Y. N., & Zammit, C. (2015). Comparing bias correction methods in downscaling meteorological variables for a hydrologic impact study in an arid area in China. Hydrology and Earth System Sciences, 19(6), 2547-2559.

- Feasibility study of surqawshan dam, 2011. ZHINAB Consulting Eng. company.

- Guven, A. (2009). Linear genetic programming for time-series modelling of daily flow rate. Journal of earth system science, 118(2), 137-146.

- Kwon, H. H., Sivakumar, B., Moon, Y. I., & Kim, B. S. (2011). Assessment of change in design flood frequency under climate change using a multivariate downscaling model and a precipitation-runoff model. Stochastic Environmental Research and Risk Assessment, 25, 567-581.

- Seckin, N., & Guven, A. (2012). Estimation of peak flood discharges at ungauged sites across Turkey. Water resources management, 26(9), 2569-2581.

- Tofiq, F. A., & Guven, A. (2014). Prediction of design flood discharge by statistical downscaling and General Circulation Models. Journal of Hydrology, 517, 1145-1153.

- Tofiq, F. A., & Güven, A. (2015). Potential changes in inflow design flood under future climate projections for Darbandikhan Dam. Journal of Hydrology, 528, 45-51.