3rd International Conference on
Hydrology & Meteorology
September 15-16, 2014 HICC, Hyderabad, India

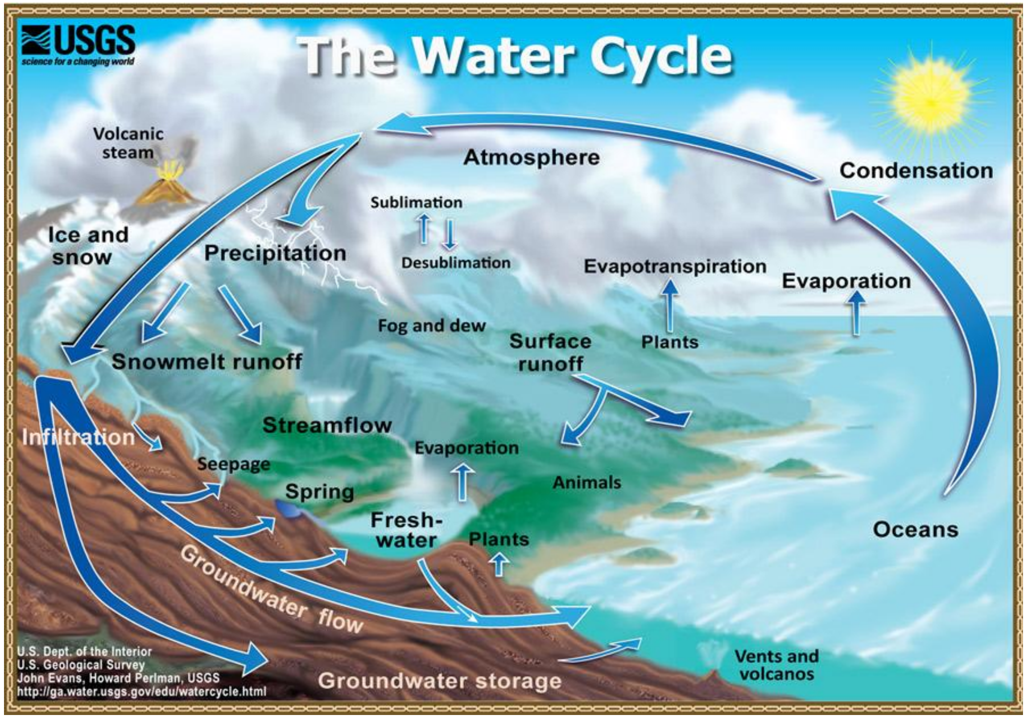
Investigation of Sensitivity of Popular Training Methods to Initial Weights in ANN Rainfall-Runoff Modeling

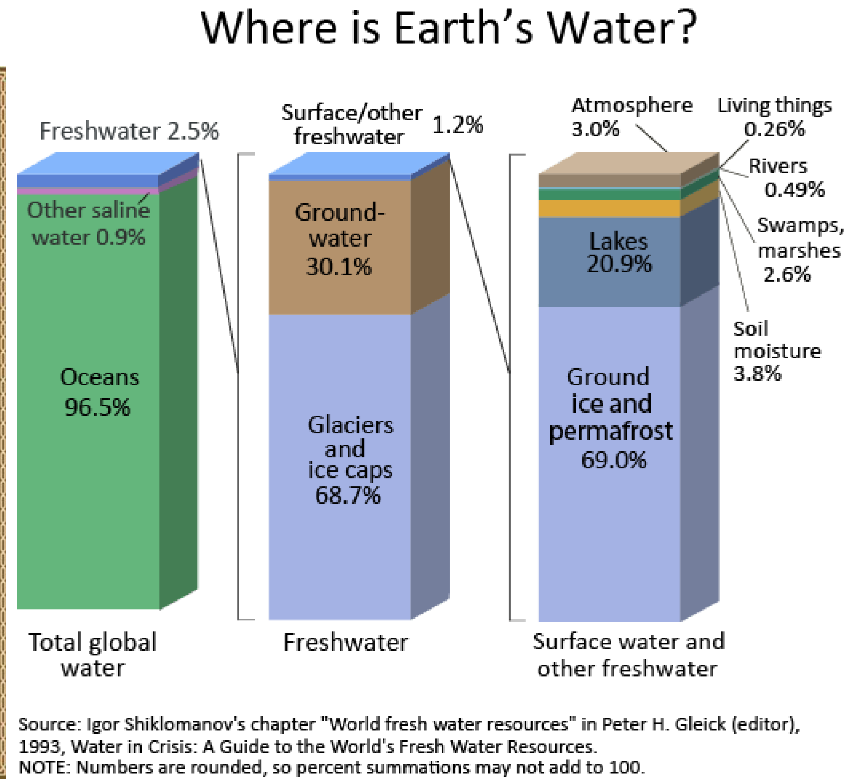
Vikas Kumar Vidyarthi
Ph.D. Scholar
Department of Civil Engineering
IIT Kanpur

Outline:

- Introduction
  - Hydrologic cycle
  - Rainfall-Runoff modeling
- Artificial Neural Network
  - Gradient-Descent algorithm
  - Levenberg-Marquardt algorithm
- Sensitivity of optimization methods on Initial weights
- Results and discussion
- Conclusion

Introduction



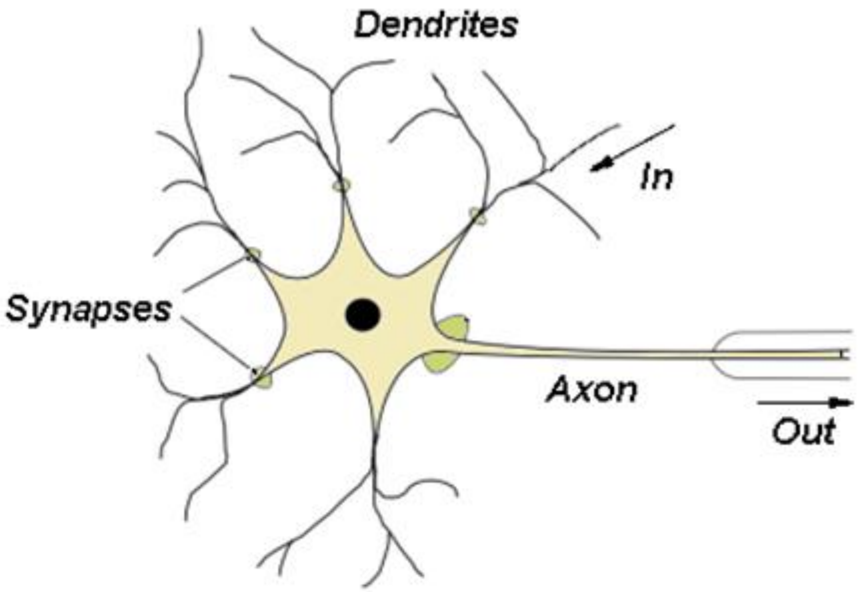


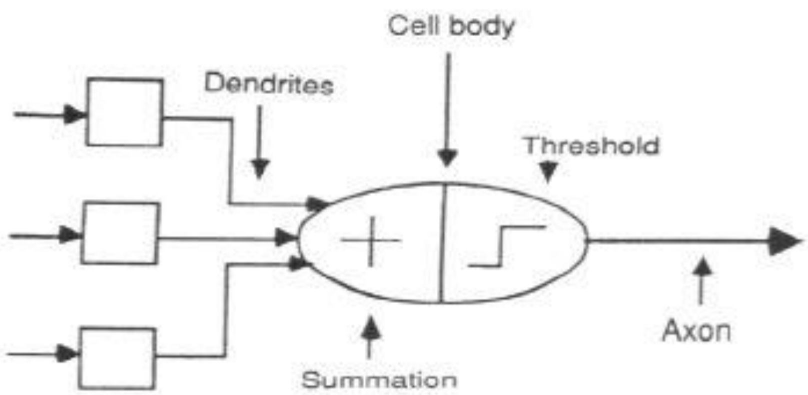
- Runoff estimation is key component of any water resources project, planning or management.

- There are several methods for Rainfall-Runoff modeling broadly divided into two categories: Conceptual and Data-driven techniques.

Artificial Neural Network

- The first artificial neuron was produced in 1943 by the neurophysiologist Warren **McCulloch** and the logician **Walter Pits**.





From Human Neurons to Artificial Neurons

ANN cont…..

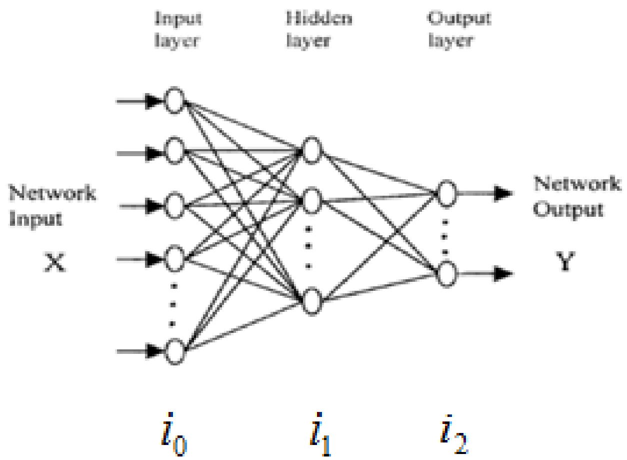
Feed-Forward steps:

i_1^{th} neuron input : $h_{i_1} = \sum W_{i_1 i_0} x_{i_0}$

i_1^{th} neuron Output : $V_{i_1} = \dfrac{1}{1+e^{-h_{i_1}}}$

i_2^{th} neuron input : $h_{i_2} = \sum W_{i_2 i_1} v_{i_1}$

i_2^{th} neuron Output : $y_{i_2} = \dfrac{1}{1+e^{-h_{i_2}}}$



Learning steps:

Error: $E(t) = \dfrac{1}{2}\sum (y_{i_2}^d - y_{i_2})^2$ ---------------- (1)

Objective function (1) can be minimized by Gradient descent algorithm, Levenberg-Marquardt algorithm or any other optimization method.

- **Gradient Descent method** :
  - Uses first-order derivative to create search direction: $s^k = -\nabla f\left(x^{(k)}\right)$
  - Works well when initial point is far away from optimum.

- Newton's method :
  - Uses 2nd order derivatives to create search directions:

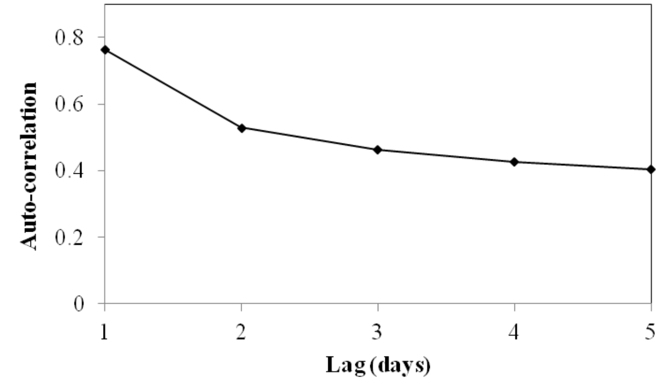
$$s^k = -\left[\nabla^2 f\left(x^{(k)}\right)\right]^{-1} \nabla f\left(x^{(k)}\right)$$

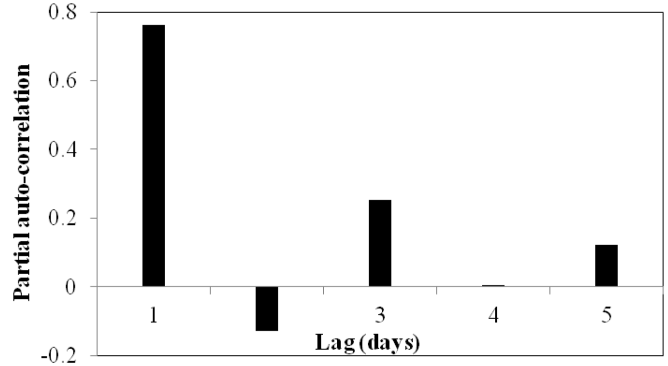
  - Suitable & efficient when initial point is close to optimum.

- **Levenberg-Marquardt method** :
  - Combination of Gradient Descent and Newton's method
  - Start with Descent method when required to search in a large space and later with Newton's method when required to search near optimum point.

ANN Model Development for Bird Creek Basin

- Study area : Bird Creek Basin, USA (Total Area: 2344 Km^2)
- Data available: Rainfall and Runoff data (8/1/1995 to 31/10/2008)
- Data division: 60% for training and 40% for testing.
- Inputs selection: ACF $\geq$ 0.7, CCF $\geq$ 0.25 and PACF $\geq$ 0.7
- Inputs selected: P(t), P(t-1), P(t-2) and Q(t-1)
- GD Algorithm Parameters: Learning Rate = 0.01 and Momentum constant = 0.9
- LM Algorithm Parameters: μ = 0.001
- ANN-Architecture selection: AARE, R^2, RMSE, NRMSE and Threshold Statistics
- Best architecture for BPA : 4_12_1
- Best architecture for LMA : 4_12_1





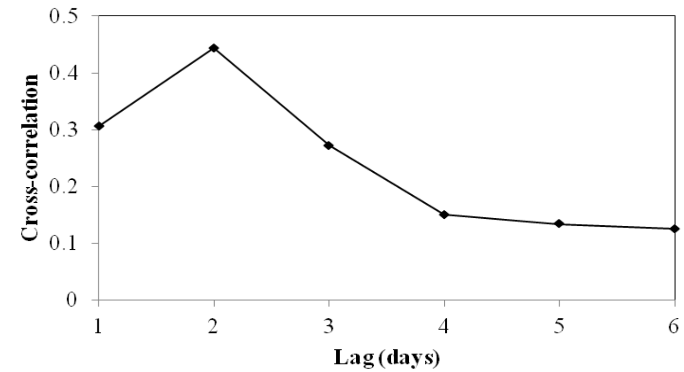


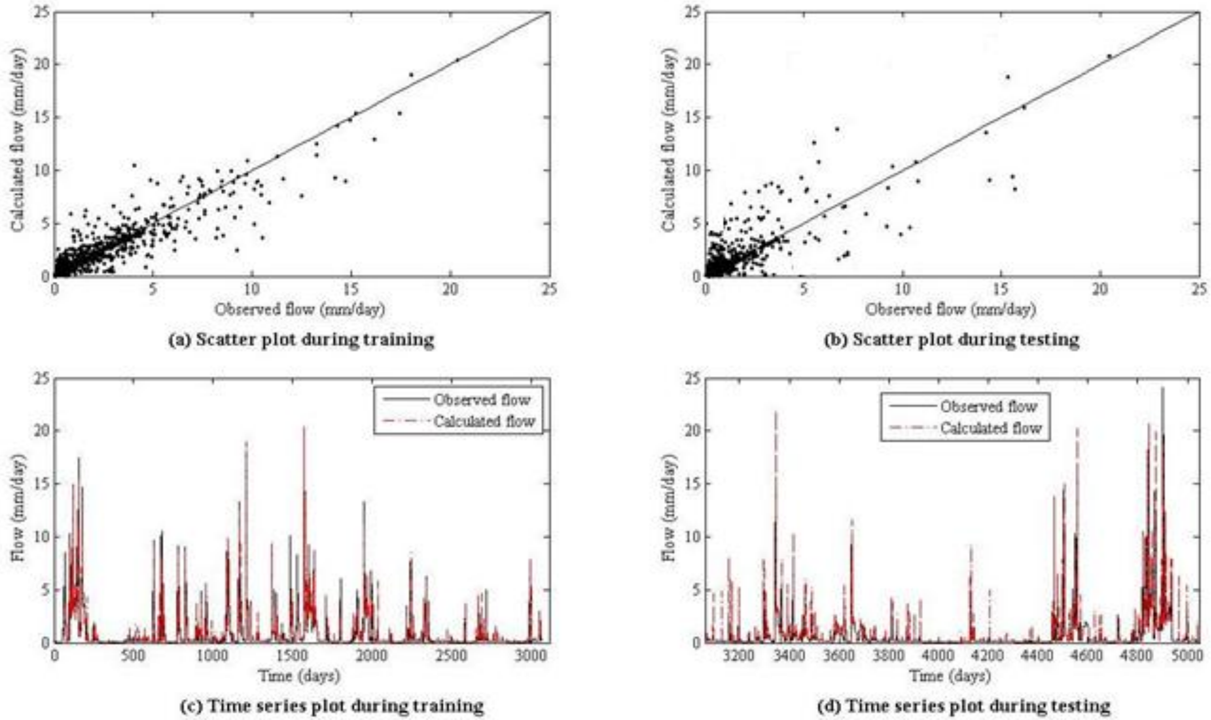
Table1: Performance comparison of ANN models trained using GD and LM Algorithms

Training performance

Model	AARE	R	NRMSE	RMSE	TS1	TS5	TS10	TS25	TS50	TS75	TS100
ANN-GD	72.52	0.863	0.234	0.029	1.14	6.31	15.56	41.96	68.10	79.22	83.40
ANN-LM	**35.45**	**0.938**	**0.161**	**0.020**	**3.79**	**21.63**	**36.47**	**60.49**	**82.65**	**89.90**	**92.97**

Testing performance

Model	AARE	R	NRMSE	RMSE	TS1	TS5	TS10	TS25	TS50	TS75	TS100
ANN-GD	67.06	0.752	0.318	0.039	2.01	8.91	16.57	43.40	70.24	80.06	84.34
ANN-LM	**48.19**	**0.787**	**0.305**	**0.034**	**3.63**	**21.85**	**34.74**	**57.30**	**79.15**	**86.91**	**90.99**



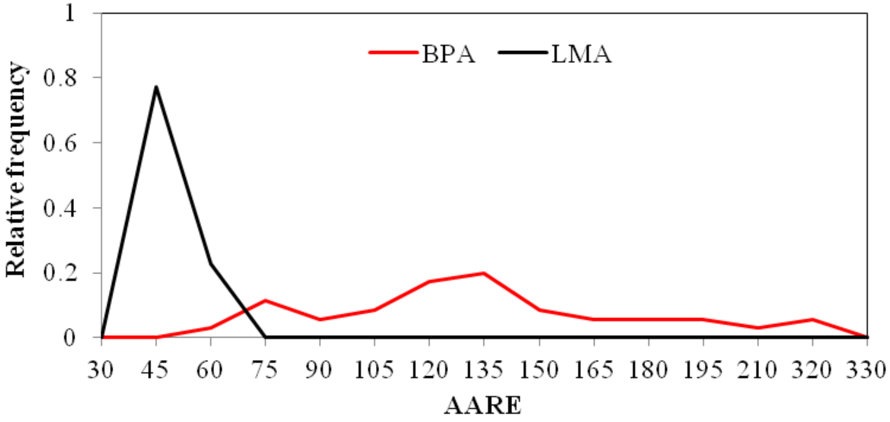
(a) Scatter plot during training

(b) Scatter plot during testing

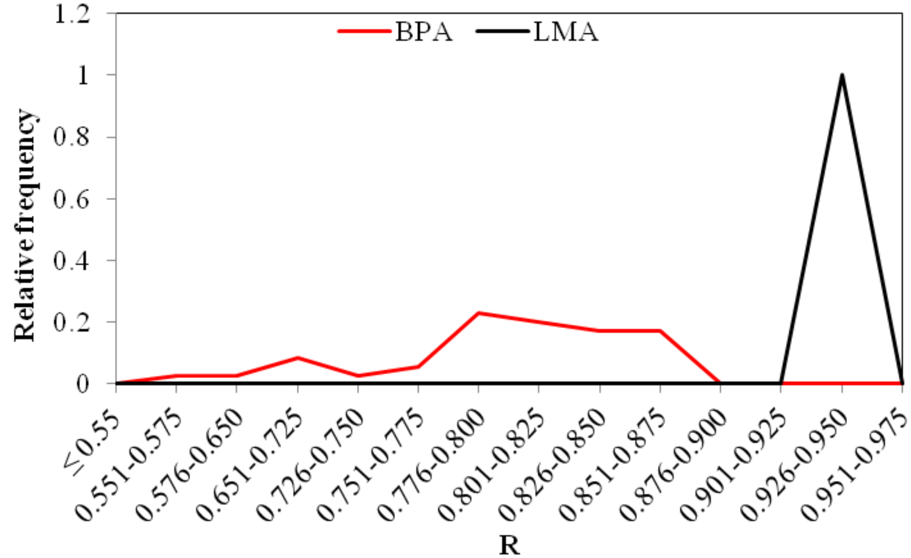
(c) Time series plot during training

(d) Time series plot during testing

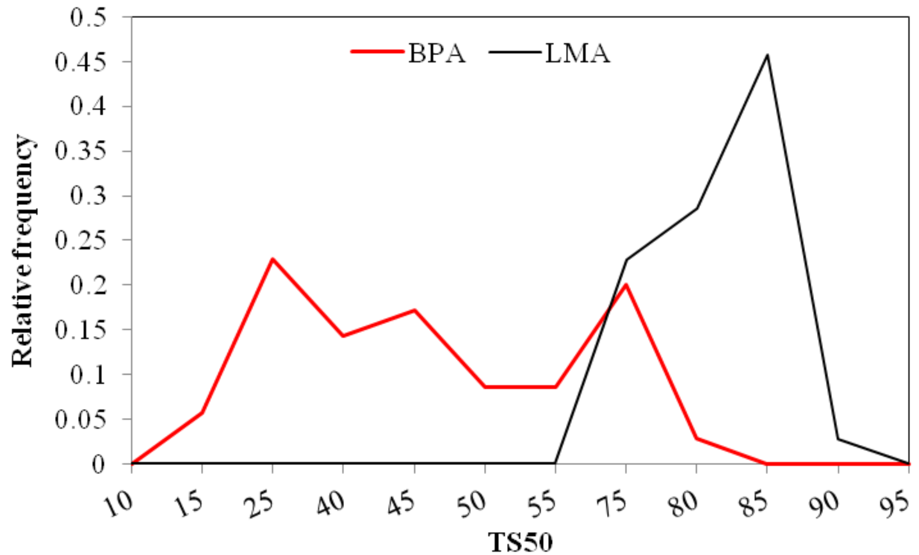
Sensitivity of training methods on initial weights



(a)



(b)



(c)

Relative frequency Vs Error statistics (a) AARE (b) R and (c) TS50

Conclusion

- ANN is capable in forecasting runoff from rainfall data.
- LM method of training performs better than GD in ANN RR modeling due to the least value of AARE(35.45), NRMSE(0.161) RMSE(0.020) and higher values of R(0.938) and all Threshold statistics.
- GD method is highly sensitive to initial weights as the standard deviation of all the error statistics such as AARE(55.39), R(0.064), NRMSE(0.169), RMSE(0.133) and all Threshold statistics are high.
- LM method is nearly insensitive to initial weights as the standard deviation of all the error statistics such as AARE(6.74), R(0.005), NRMSE(0.028), RMSE(0.022) and all Threshold statistics are less.

References

- Maier, H.R., Jain, A., Dandy, G.C., and Sudheer, K.P. (2010), Methods used for the development of neural networks for the prediction of water resources variables in river systems: Current status and future directions, *Env. Mod. Software*, 25, 891-909

- Srinivasulu, S. and Jain, A. (2006), A comparative analysis of training methods for artificial neural network rainfall-runoff modeling, *J. Applied Soft Computing*, 6(3), 295-306.

- Govindaraju, R.S. and A.R. Rao - editors (2000), Artificial Neural Networks in Hydrology, Kluwer *Academic Publishers*, Amsterdam.

- Kalyanmoy Deb, Optimization for Engineering Design: Algorithm and Examples, PHI publications, New Delhi.

- http://water.usgs.gov/edu/watercycle.html

Acknowledgement

- My mentor **Prof. Ashu Jain**, Department of Civil Engineering, IIT Kanpur.

- Computer centre, IIT Kanpur.

- Organizing committee,"HYDROLGY-21014".

Thank you for your attention !

Questions ?