Modeling of Annual Peak Flows in Pakistan by Using Artificial Neural Network

By

Hameed Ullah

Reg. No. 55-FBAS/MSST/F14

Department of Mathematics & Statistics

Faculty of Basic and Applied Sciences

International Islamic University, Islamabad

Pakistan

2017

Neural networks.

Artificial intelligence.

Modeling (Mathematics)

Modeling of Annual Peak Flows in Pakistan by Using Artificial Neural Network

By

Hameed Ullah

Supervised By

Dr. Ishfaq Ahmad

Department of Mathematics & Statistics

Faculty of Basic and Applied Sciences

International Islamic University, Islamabad

Pakistan

2017

Modeling of Annual Peak Flows in Pakistan by Using Artificial Neural Network

By

Hameed Ullah

A Thesis
Submitted in the Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
In
STATISTICS

Supervised By

Dr. Ishfaq Ahmad

Department of Mathematics & Statistics

Faculty of Basic and Applied Sciences

International Islamic University, Islamabad

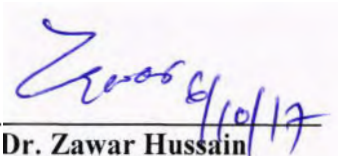**Pakistan**

2017

Certificate

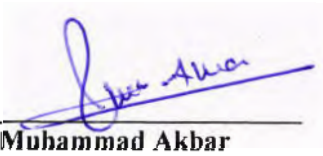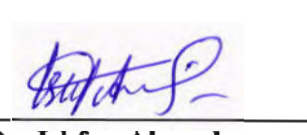## Modeling of Annual Peak Flows in Pakistan Using Artificial Neural Network
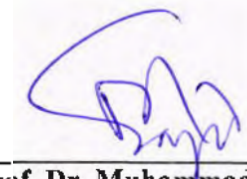
By

Hameed Ullah

A DISSERTATION SUBMITTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF THE MASTER OF SCIENCE IN STATISTICS

We accept this dissertation as conforming to the required standard.

1. _____
Dr. Zawar Hussain
External Examiner

2. _____
Dr. Muhammad Akbar
Internal Examiner

3. _____
Dr. Ishfaq Ahmad
Supervisor

4. _____
Prof. Dr. Muhammad Sajid, T.I
Chairman

**Department of Mathematics & Statistics
Faculty of Basic and Applied Sciences
International Islamic University, Islamabad
Pakistan
2017**

Dedicated

To

My beloved

Parents
and
Respected teachers

With their efforts and support I would not be able to complete this task, without their Love, Prayers, and Support, I am nothing.

DECLARTION

I hereby declare that this thesis, neither as a whole nor as a part thereof, has been copied out from any source. It is further declared that I have prepared this thesis entirely on the basis of my personal efforts made under the sincere guidance of my supervisor. No portion of the work, presented in this thesis, has been submitted in the support of any application for any degree or qualification of this or any other institute of learning.

Signature: -----------------------

Hameed Ullah

MS (STATISTICS)

Registration No: 55-FBAS/MSST/F14

Department of Mathematics & Statistics

Faculty of Basic and Applied Sciences

International Islamic University,

Islamabad, Pakistan

Forwarding Sheet by Research Supervisor

The thesis entitled **"Modeling of Annual Peak Flows in Pakistan by Using Artificial Neural Network"** submitted by **Hameed Ullah** (Registration # 55-FBAS/MSST/F14) in partial fulfillment of M.S degree in Statistics has been completed under my guidance and supervision. I am satisfied with the quality of his research work and allow him to submit this thesis for further process to graduate with Master of Science degree from Department of Mathematics and Statistics, as per IIU Islamabad rules and regulations.

Dated: _____

<div align="right">

Dr. Ishfaq Ahmad

Assistant Professor

Department of Math & State,

International Islamic University,

Islamabad.

</div>

Acknowledgements

TABLE OF CONTENTS

List of Figures

List of Tables

List of Abbreviations

NDMA	National disaster Management Authority
UNESO	United Nations Educational, Scientific and Cultural Organization
WMO	World Meteorological Organization
FFC	Federal Flood Commission
NDMA	National Disaster Management Authority
GIS	Geographical Information System
ICT	Information and Communications Technology
QFF	Quantitative Flood Forecasting
ANN	Artificial Neural Network
W-ANN	Wavelet Artificial Neural Network
WAPDA	Water and Power Development Authority
FFN	Feed-Forward Network
FFMLPs	Feed-Forward Multilyer Perceptron
MLP	Multilyer perceptron
LMS	Least Mean Square
CWT	Continues Wavelet Transform
DWT	Discrete Wavelet Transform

CA1	Approximation Coefficient
CD1	Detail coefficient
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
ME	Maximum Error
ESD	Error Standard Deviation
NRMSE	Normalized Root Mean Square Error
E	Nash-Sutcliffe Coefficient
IOA	Index of Agreement
MLR	Multiple Linear Regression
WNM	Wavelet Network model
TAR	Threshold Auto-regressive Model
TNC	Total Nitrogen Concentration

Abstract

The data of annual maximum peak flows into stream or river system is not straight forward but rather complex function of hydrology and geography of the area. Well recognized statistical ways have been used to forecast the flood occasions with their frequency and magnitude. In such composite situations, the advancement of models based on sequential observations may increase understanding the hidden hydrological processes. The discrete wavelet transform (DWT) is utilized to decompose annual peak flow time series information into wavelet coefficients. The coefficients of wavelet are then utilized as inputs to ANN models to forecast peak flow. Regular Artificial Neural Network (T-ANN) and Wavelet-Ann conjunctional models (db1 and db2) are being used in this study to compare its performance for one year ahead forecasting of annual maximum peak flows (AM) of two gauging sites (Mangla and Khanki) on river Jhelum in Pakistan. At last the result demonstrate that Wavelet-ANN conjunctional models is the most reasonable model for the annual maximum flows (AM) of two gauging sites (Mangla and khanki) on Jhelum in Pakistan as compare to the regular artificial neural network models (T-ANN).

CHAPTER 1

INTRODUCATION

1.1 Introduction to Floods

Floods can be defined as, a situation when excess water covers land which is usually dry Miklas and Yang (2010). The floods usually destroy property including agricultural land, houses, livestock's and human losses, while infrastructural damages and salinity is the most common losses occur during floods. Floods also cause diseases and contamination in drinking water. World has experienced numerous floods throughout its history. World ancient civilizations live near streams and rivers in their era to irrigate fields and collect drinking water easily.

Mostly the water channels that do not have the capacity to take heavy rain water become cause of floods. But heavy rainfall is not always the cause of floods. They can also result from other causes, especially in seaside where flood can be produced by a hurricane, a tsunami or a high tide concurring with unusually high river levels. There is another reason of floods even in the dry weather environment when the earthquake damage water dams and the water cover dry areas in the downstream areas. It is a hazardous situation which can result in heavy material and human losses.

Flooding can happen in any area, in the rural or urban areas. Flooding can also happen if the flow rate is greater than the size of the river. People all over the world are facing many problems because of abnormal environmental occurrences. Flooding, downpours, droughts and fast winds that cause cyclones and destroy anything that comes in their way. Floods are natural calamities that occur not only in Pakistan, but in other parts of the world as well. Flood prediction can be defined as the use of actual

1

rainfall and data of stream flow in rainfall-runoff and stream current models to predict current rates and level of water for phases starting from some hours to days forward, dependent on the magnitude of the crisis or river basin. In order to make the most precise flood prediction, it is good to have a lengthy time-series of historical information that connects stream current to measured previous rainfall happenings. Flood prediction is important to warn the people of the coming floods. The purpose of flood warning is to decide whether flood warnings should be given to the public or past warnings should be declared invalid.

1.2 Floods in Pakistan

Pakistan has been suffered from floods many times in the history, the size of population affected, extent of area, damage to crops and other economic damages. Pakistan have also experienced such undesirable floods situations and the first flood was occurred in 1950, second in in 1955 and third and fourth flood were occurred in the preceding year 1956 and 1957 respectively as reported by the Federal Flood Commission (FFC) of Pakistan.

The FFC also witnessed five floods in seventies, four in eighties, three in nineties and the remaining are thereafter. Till date 599,459 square km' area has been damaged because of these floods 11,239 costly human lives lost and the national economy suffered losses of over Rs 39 billion. In the last 63 years, flood has affected 180,234 villages. But preventing floods is very critical. According to FFC the flood in 1950 submerged 10,000 villages, extending over an area of 17,920 square km and killed 2,190 people and on the basis of large number of human losses it is considered the most dangerous flood of the country history. According to FFC more than thirteen thousand villages were flooded and a thousand people were killed during 1992 flood.

2

Almost thousand lives were lost and more than two thousand villages were destroyed by the floods of 1977 reported by FFC. Pakistan faced the second most dangerous flood in 2010, which took the lives of 1,985 people and destroyed more than seventeen thousand villages and 16 million hectares' valuable land. It has been also observed by FFC that 516 and 571 people lost their lives during 2011 and 2012 floods respectively. Summing up, the floods of 2010, 2011 and 2012 took the lives of more than three thousand people and the monetary losses measured were about $16 billion.

According to the National Disaster Management Authority (NDMA) reports, in 2013, monsoon floods killed 69 people, in which most of the people 22 persons lost their lives in Sindh provinces, 18, 15 and 14 people lost their lives in Baluchistan, Punjab and in Khyber Pakhtunkhwa province respectively. The NDMA reports that the flooding has up till now affected the lives of more than 81 thousand people and more than 300 villages, destroyed about 135,000 acres of crops, 2,533 houses has been fully destroyed and 1700 houses found partially damaged. In September 2014, heavy monsoon rains and floods in the catchment zones of eastern Indian rivers Chenab, Ravi, Sutlej, and Jhelum, caused flooding in Punjab, Northern areas and Azad Kashmir areas. They affected the lives of more than 2.5 million people and approximately 129,000 families were suffered. In many situations causing the loss of available food and cash crops. The sources of non-farm income and services affected contain many small businesses, industry and loss of employment because of disturbance of the economy. The projected cost of the rehabilitation effort was US$439.7 million.

In 2015 heavy rains affected district Chitral particularly in Khyber Pakhtunkhwa, Gilgit-Baltistan, low lying parts near River Indus in Sindh, some

areas of province Punjab and some mountainous areas of Baluchistan. Reasonable to heavy rain in higher catchments of main rivers and their branches caused flood flows, which generated damages to people and public and private infrastructure. In 2015 flooding affected above 1.933 million people and 4,634 villages (destroying 10,716 families), in addition to, killing 238 people. Nearly thirteen hundred people have been wounded and 45 lacs were affected by the flash floods in the three consecutive years 2014, 2015 and 2016. The Details are given below in table 1.1, which is reported by Ministry of water and power of Pakistan (Annual floods report 2015).

Fig 1.1 damages of houses by the floods of 2016

Fig 1.2 Damages caused by the floods of 2013

Fig 1.3 show the floods area of 2015

Fig.1.4 damages caused by floods 2014

Table: 1.1 Historical Flood Events Experienced in Pakistan

S.No	Year	Direct losses (USS Million) @IUSS=PKR86	Lost Lives (No)	Effected Villages (No)	Flood Area Sq.-km
1	1950	489	2191	19998	17918
2	1955	377	678	6947	20481
3	1956	319	161	11610	74407
4	1957	300	82	4497	16004
5	1959	233	86	3901	10423
6	1973	5135	475	9720	41473
7	1975	685	127	8627	34932
8	1976	3484	46	18391	81921
9	1977	336	847	2184	4618
10	1978	2225	394	9198	30596
11	1981	300	81	2072	4192
12	1983	134	40	644	1881
13	1984	76	41	252	1091
14	1988	859	509	99	6145
15	1992	3011	1007	13210	38757
16	1994	842	432	1620	5569
17	1995	737	590	6851	16687
18	2010	10001 @IUSS-PKR86	1986	17554	169998
19	2011	3728 @IUSS-PKR94	515	38699	27582

20	2012	2639 @IUSS=PKR95	572	14160	4745
21	2013	2002 @IUSS=PKR98	332	8298	4481
22	2014	439 @IUSS=PKR101	369	4064	9780
23	2015	171 @IUSS=PKR105.0	237	4635	2878
24	2016	6 @IUSS=PKR104.8	152	44	-
Total	-	38528	11950	207275	631177
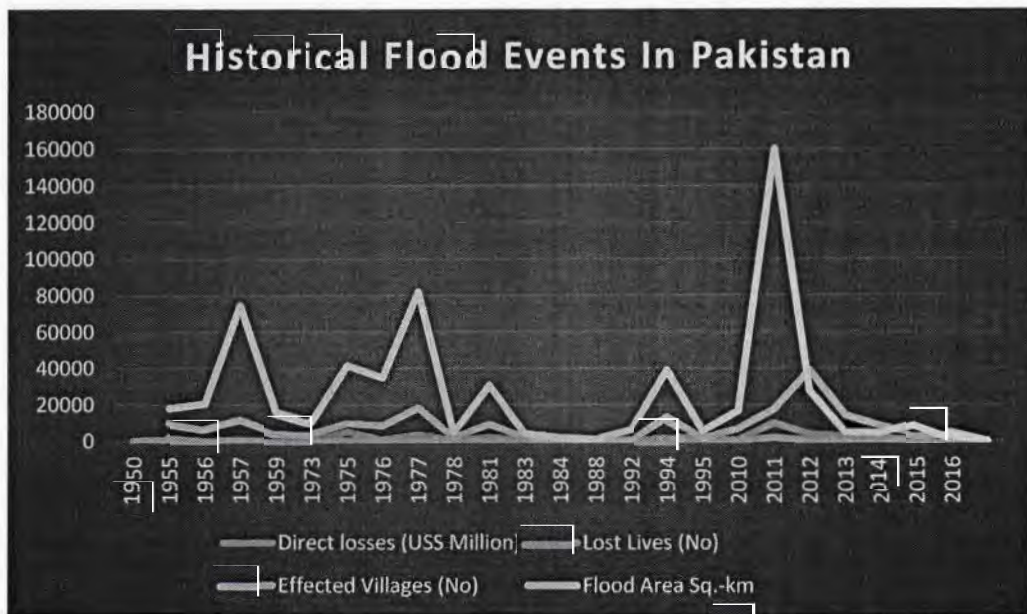
Figure: 1.5 Show the flood events in Pakistan since 1950-2016

The NDMA reported that economic losses in the last five year floods were more than the losses in previous 63 years collectively. Because of unavailability of a

6

proper disaster management system Pakistan faced heavy losses in floods, the approximate monetary value of those losses is $800 million every year.

The changing world climate is considered the most important factor affecting natural calamities such as flood and tropical hurricanes. Along with other regions of the world Pakistan is also sensitive climate change and faced its outcomes, which includes the happening of natural calamities such as storms, fire, floods, food insecurity, droughts, and epidemic diseases. The effects of floods can be divided into primary, secondary and tertiary effects. Direct contact with flood water are the primary effects, population who suffers from the destruction of infrastructure destroyed by floods are the secondary effects, whereas tertiary effects are considered the long-term effects of floods such as changes in the location of river ways. To reduce the consequences of flood several measures to be taken, warning about floods is considered a useful measure to reduce losses during floods. Information and Communications Technologies (ICT), Geographical Information Systems (GIS) and remote sensing can be used to inform people in advance about flood. Forecasting floods in the future is a composite process because of the variation in geographic positions, rain and type of soil and magnitude of catchments that influence the levels of river water. Modeling like Wavelet-Artificial Neural Network, Quantitative Flood Forecasting (QFF) and Artificial Neural Networks (ANN) have been established and put into practice in different places. Linear statistical models, non-linear time series analysis and prediction, radar information Systems Stream, rain-gauge networks and hydrograph analysis have been utilized as data sources till now in different regions of the world.

The current study will utilize the non-linear time series analysis and prediction methods to forecast the floods in Pakistan and the models of non-linear time series Wavelet-Artificial Neural Network (W-ANN) and Artificial Neural Network (ANN) was used.

1.3 Objectives of the Study

The specific objectives of the study are given as follow

- ➢ To display that Artificial Neural Network and Wavelet-ANN conjunctional models can be used as effective approach to forecast floods from hydrological data.

- ➢ To regulate the architecture of Artificial Neural Network and Wavelet-ANN conjunctional model that will provide the most affective forecasting performance for flood.

- ➢ To make a more precise forecasting of the floods by using Artificial Neural Network and Wavelet-ANN conjunctional model, to mitigate the flood losses.

- ➢ To compare the performance of Artificial Neural Network and Wavelet-ANN for flood prediction.

CHAPTER 2

LITERATURE REVIEW

2.1 Review of Methodology

This chapter briefly shows the previous studies about the topic. Flood forecasting using Artificial Neural Network have been conducted with some journal from other researchers. All the journals are made as a reference for my study to complete this thesis. However, the different are only the scope of study and different modeling approaches use in each study. Current studies have shown that a parallel computing model has an ability to predict river flow which called Artificial Neural Network.

Tokar, A. S, & Johnson, P.A. (1999) applied the methodology of ANN to predict everyday overflow as a function of daily rain, meeting of snow and temperature, for the little Patuxent River crisis in Maryland. For the rainfall-runoff Artificial Neural Network (ANN) were compared with the simple theoretical model and current techniques including statistical regression. It was observed that the ANN model delivers more organized approach; the calibration data dimension may be decreased and reduces the time consumed in the calibration of the model. The results also indicate that it represents the flexibility of current methods and on improvement upon the prediction accuracy.

Zealand et al. (1999) concluded that the methodology of ANN modeling is a perfect substitute to the conventional time series methods for making input output simulations and hydrologic models that are generally used in modeling the interior

processes of a watershed. ANN models have the capability of learning and summarizing the basic features from inputs that may include irrelevant and unnecessary information. ANN are helpful in the solutions of problems without current algorithmic or in situations where the algorithms are complex to implement.

Campolo *et al.* (1999) developed ANN to examine and predict the actions of the river Tagliamento in Italy, during heavy rainy seasons. The model uses the information received from different rain gauges in mountainous regions about the rainfall distribution and forecast the water level. For the prediction of model accuracy, they used mean square error. They concluded that increasing the time perspective, thus making the model best suited for flood foretelling, reduces the precision of the model.

Dawson and Wilby (1999) analyzed the performance of two ANN models Multilayer Perceptron (MLP) and radial basis function network(RBF). They used the gradual multiple linear regression models and zero order forecast of river current. The results conclude that ANNs are well fitted to rainfall-runoff modeling and actual forecasting.

Imrie *et al.* (2000) concluded that artificial neural networks (ANNs) provide a quick and elastic way of creating models for river current forecasting, and showed that the ANN performs well in comparison with traditional methods. They present a way for better generalization during training by the addition of a guidance system to the cascade-correlation learning architecture. In the United Kingdom Two case studies were developed for the validation of data that comprises values that are deviated from the calibration data. With the standard error backpropagation, the capability of the developed algorithm to generalize on new data was compared. The

capability of ANNs trained with unrelated results activation functions to infer beyond the calibration data was observed.

Loing *et al.* (2000) analyses the ANN model very appropriate for current forecasting tool with a high ratio of water-level precision at Dhaka Bangladesh. The effectiveness of fit test R^2, root mean square and mean absolute error to be used in this study. They concluded that the results are valid for the policymakers to decrease needless data collection of gauging stations in lower cost.

Thirumalaiah, K, & Deo, M. C. (2000) study the ANN model for the planning of water resources system such as lakes and power plants requires actual or on-line forecasting of runoff and river stage. The study demonstrates the use of NNs to actual forecasting of hourly flood runoff and everyday river stage along with rainfall sufficiency for India.

Coulibaly *et al.* (2000) used multi-layer feed-forward neural networks for real-time lake in flow prediction. Multivariate hydrological time series methodology from chute-du-Diable hydro system was used in Canada. Model function was compared to statistical model and operational conceptual model. The results indicate that this method is useful for improving predictive precision and a substitute to dynamic adaptive forecasting. (ASCE Task Committee on Artificial Neural Networks in Hydrology 2000a,) concluded that ANN has been effectively used in hydrology-related areas such as rainfall–runoff modeling, stream flow forecasting, groundwater modeling, water quality, water management policy, rain foretelling, hydrologic time series and reservoir operations.

Wei *et al.* (2002) proposed artificial neural network (ANN) predicted a way for flood calamity problem in which the neural network model and its fundamental

principles were discussed. They utilized the example of flood disaster area in china from 1949 to 1994 is used for display. The result proved that the utilization of ANN with the capability of mapping or function similar to the flood disaster prediction problem is a suitable approach which is not only simple but also dependable.

Campolo *et al.* (2003) use the ANN model for flood forecasting to decrease the economic losses and the hazard of people. In this study flood forecasting model is described that uses the real-time information available basin to forecast the water level growth. The model is based on ANN, which were effectively used in previous works to foretell floods in a loose basin and to forecast water level growth in the Arno basin water current conditions. They concluded that as the flow rate increase it increase the authenticity of the forecast model each time ahead. It has been also suggested that the model is quite useful for the prediction of floods and its severity.

Rajurkar *et al.* (2004) studied the daily flows during floods using ANN model in India. The data of two large catchments in India and the data applied by WMO earlier of five catchments were used in India and for five other catchments applied earlier by the WMO for the comparison of the operational hydrological models. The results are estimated the rainfall runoff from a linear model and ANN model and concluded that ANN is valuable in modeling the rainfall runoff association in the updating coming.

Shresta *et al.* (2005) used ANN to provide a flexible way of developing flood flow simulation models. This study analyzes methods for developing generalization of the ANN using three different flood events data sets from the Neckar River in Germany. The results of this study indicate that the ANN in a suitable structure can increase forecasting ability to certain degree before the calibrated data sets.

12

Dawson *et al.* (2005) used ANN for flood forecasting for the ungagged catchments and found the results of ANN more authentic and acceptable than multiple regression models.

Dawson *et al.* (2006) applied the Artificial Neural Network (ANN) with in the area of hydrological modeling for over ten years but comparatively little attention has been paid to the use of these tools for flood assessment in ungagged catchments. The data used for this study from the canter of Ecology and Hydrology flood estimation Handbook (FEH) for 850 catchments across the UK. They compared the results of ANN model with multiple linear failures and conclude that ANNs provide better flood estimates that can be utilized by engineers and hydrologists.

Nayak *et al.* (2006) analyzed the potential of several ANNs techniques in predicting the ground water level variations in a free coasted aquifer in India. The results indicate that the model prediction is quite correct as assessed by various statistical indices. The results are also valid to predict the water level up to 4 months in advance fairly well.

Anter *at al.* (2006) for the Blue Nile catchment analyzed rainfall-runoff model based on ANN for the daily based data from 1992 to 1999. The data period was divided into two sets, for the model calibration the data from 1992 to 1996 was used the data from 1997 to 1999 was used for model validation. ANN model and physical based distributed model were compared and it was concluded that for simulation ANN technique for rain fall-runoff is best technique.

Adamowski and Jan F (2008) used the recently developed wavelet forecasting method compared to Multiple Linear Regression analysis (MLR, Autoregressive Integrated Moving Average analysis (ARIMA), and Artificial Neural Network (ANN)

examination for forecasting utilizing data from the radian River Watershed in Ontario, Canada. They concluded that the finest forecasting model for 1 day, 2 day, 3-day lead-time was a wavelet analysis model as compared to (MLR), (ARIMA), and (ANN).

Wang *et al.* (2009) developed the Wavelet Network Model (WNM) to forecast the inflow in three Chinese dams. They prediction was made for both short and long term input, runoff of the three dams by using Wavelet Network Model (WNM) such as mean annual discharge, 10 days' average seasonal discharge, mean of daily discharge data and annual maximum peak flows discharge. At the same time, Threshold Auto-regressive model (TAR) was also used for that forecast and it was revealed that Wavelet Network Model (WNM) is generally acceptable as compare to the Threshold Autoregressive model (TAR).

Islam (2010) developed the river stage neural network model to forecast the water level of Dhaka city. The data from five stations out of which three the Ganges, Brahmaputra, and Meghna rivers are chosen as input nodes and Dhaka on the Buriganga river were selected as output node for the neural network. In this study the estimated R^2, root mean square and mean absolute error are used for the evaluation of model. The study concluded that the results are suitable for actual flood prediction by decreasing calculation time, improving water resources organization and decreasing the avoidable cost of field data collection.

Adamowski *et al.* (2010) proposed Wavelet-Artificial Neural Network (W-ANN) and Artificial Neural Network (ANN) models for flow prediction use for two dissimilar rivers in Cyprus. They compared the relative performance of couple Wavelet-Artificial Neural network (W-ANN) to the regular Artificial Neural Network

14

(ANN) models for flow forecasting. They concluded that in both cases the two (W-ANN) models were found to provide more correct flow forecasting than the ANN.

Shamsedin A.Y. (2010) used an ANN foretelling the Blue Nile river currents in Sudan. He developed the four ANNs rainfall runoff-models based on the formation of the multilayer perceptron. The results of the study indicate that the model which utilizes the both seasonal expectations of the observed outflow and the rainfall index has the best presentation. The result also shows that the forecast updating has considerably improved the quality of the discharge flow in developing countries.

Adamowski et al. (2011) used a fresh method based on combining Discrete Wavelet Transform (DWT) and Artificial Neural network (ANN) for ground water level forecasting. They used the variables; monthly total rainfall, normal temperature and normal groundwater level data at two places in the Chateauguay watershed in Cuebec, Canada to developed and validate the model. They compared Wavelet-Artificial Neural Network (W-ANN) model to the regular Artificial Neural Network (ANN) models and Autoregressive Integrated Moving average (ARIMA) models for monthly groundwater level prediction and found that the W-ANN model provide more correct monthly average groundwater level predictions as compared to the ANN and ARIMA models.

He et al. (2011) applied ANN model to examine the relationships among the use of land, manure, and hydrometer logical conditions in 59 river basins over Japan and then used to guess the monthly river total nitrogen concentration (TNC). The results indicate that the ANN model gave acceptable forecasts of stream TNC and looks to be helpful tool for forecasts of TNC in Japanese rivers. The results also show that the ANN model could give correct estimation of nitrogen concentration in rivers.

Singh, R.M. (2012) used Wavelet-ANN conjunction model for chronological patterns taken from temporal observations of annual peak flow series of 70 years at Okhla Weir in New Dehli. He proposed the methodology to illustrated with real data, and concluded that the limited performance evaluation of the methodology show potential application of the developed methodology.

Valipour *et al.* (2013) applied three techniques Auto Regressive Moving Average (ARMA), Auto Regressive Integrated Moving Average (ARIMA) and Artificial Neural Network (ANN) models for predicting the inflow of Dez Dam Reservoir. They used the data monthly outflow from 1960 to 2007. For training the model they used initially 42 years' data and for forecasting purpose only 5 years past data. The study concluded that the ANN model with sigmoid Activation function and 17 neurons in the hidden layer was selected as the best fitted model for forecasting inflow of the Dez Dam Reservoir.

Raman *et al.* (2013) used the Wavelet-Artificial Neural Network model (W-ANN) to compare with Artificial Neural Network models (ANN) for monthly rainfall data forecast of Darjeeling district in the extreme north of west Bengal state in the east of India. The results indicated that the performance of Wavelet-Artificial Neural Network models (W-ANN) are more useful than the Artificial Neural Network models (ANN).

Guimaraes Santos *et al.* (2014) proposed Wavelet-Artificial Neural Network and Artificial Neural Network hybrid models for everyday stream flow prediction at 1,3,5 and 7 days in advance on the low frequency components of the original signal (approximation). The data utilized in this study relate to the denaturalized everyday stream current time series at the Sobradinho hydrologic plant, in the SaoFrancisc

River, Brazil. The results indicate that the projected W-ANN Hybrid models provide considerably better results than the classical ANN models for all tested situation.

CHAPTER 3

MATERIALS AND METHODS

3.1 Introduction

This chapter will explain about the methods used in this study. The data collection and study area is also explained in this chapter. The modeling details are being described in next section.

3.2 Modeling Details

Artificial Neural network and Wavelet-Artificial Neural Network will be used to forecast floods in river Jhelum. The performance of W-ANN and ANN data training and testing will be determined using Correlation Coefficient (R), determination coefficient, and coefficient of Nash Sutcliffe (E), Index of Arguments (IOA), MAE and RMSE. To adjust the joining weight, Back propagation (BP) process and Multilayer Perceptron (MP) process are used.

3.3 Study Area and Data

The annual peak flows data of two sites (Khanki and Mangla) situated on river Jhelum in Pakistan have been used in this study. The data of these sites have been collected from the hydrology department of Water and Power Development Authority (WAPDA) and Federal Flood Commission (FFC). The data of Annual Peak flows at each site have been noted in the peak of wet season (from July to September). The

values are the maximum value in 24 hours during whole year at a particular site called Annual Peak flows.

The river system of Pakistan (Figure 3.3.1) begins from the snow-covered Himalayan and the Karakoram series. There are mainly five rivers that mostly flows mostly in the province of Punjab; Sutlej, Ravi, Jhelum, Indus and Chenab. River Jhelum originates from south eastern part of Kashmir and its width is approximately 774 kilometers and it meets with Chenab River at Trimmu. Jhelum also passes through Srinagar before entering Pakistan. Number of barrages and dams has been constructed on River Jhelum one in them is Mangla dam constructed in 1967 and is considered the largest earth-fill barrier of the world. It has stowage dimensions of approximately 5.9 million acre-feet. River Sutlej flow over the famous areas of Pakistan and Indian Punjab, it flows in northern of the Vindhya Series, and east of the Central Sulaiman Series in Pakistan. River Sutlej is approximately 550 kilometers extensive; it is also called Red River. River Chenab irrigate Jammu and Kashmir, and then it joins River Jhelum at Trimmu, and it is approximately 960 kilometers extensive.

River Ravi, similar several other rivers of the area, originates from the Himalayas. Later it flows in the south-west area of the Indian Punjab. Ravi irrigates the boundary lands of India and Pakistan and its length is about 700 kilometers. Lahore city is situated at Ravi's eastern bank therefore it is called the River of Lahore. The longest river of Pakistan is the River Indus which is creating from Himalayan area. On the base of annual flows of water, it is the world's 21st largest River. The overall length of the river is about 3,180 kilometers; also it is called a lifeline of Pakistan's. River Indus creates from the Tibetan plateau nearby Lake Mansarovar, which is in China. And then it runs over Jammu and Kashmir, come in Gilgit-

19

Baltistan area and movements over the whole distance of the country and combines with the Arabian Sea. The key support of agriculture and water necessities of Pakistan are satisfying by river indus. Figure 3.3.1 shows the River system and flood routing model of Pakistan.
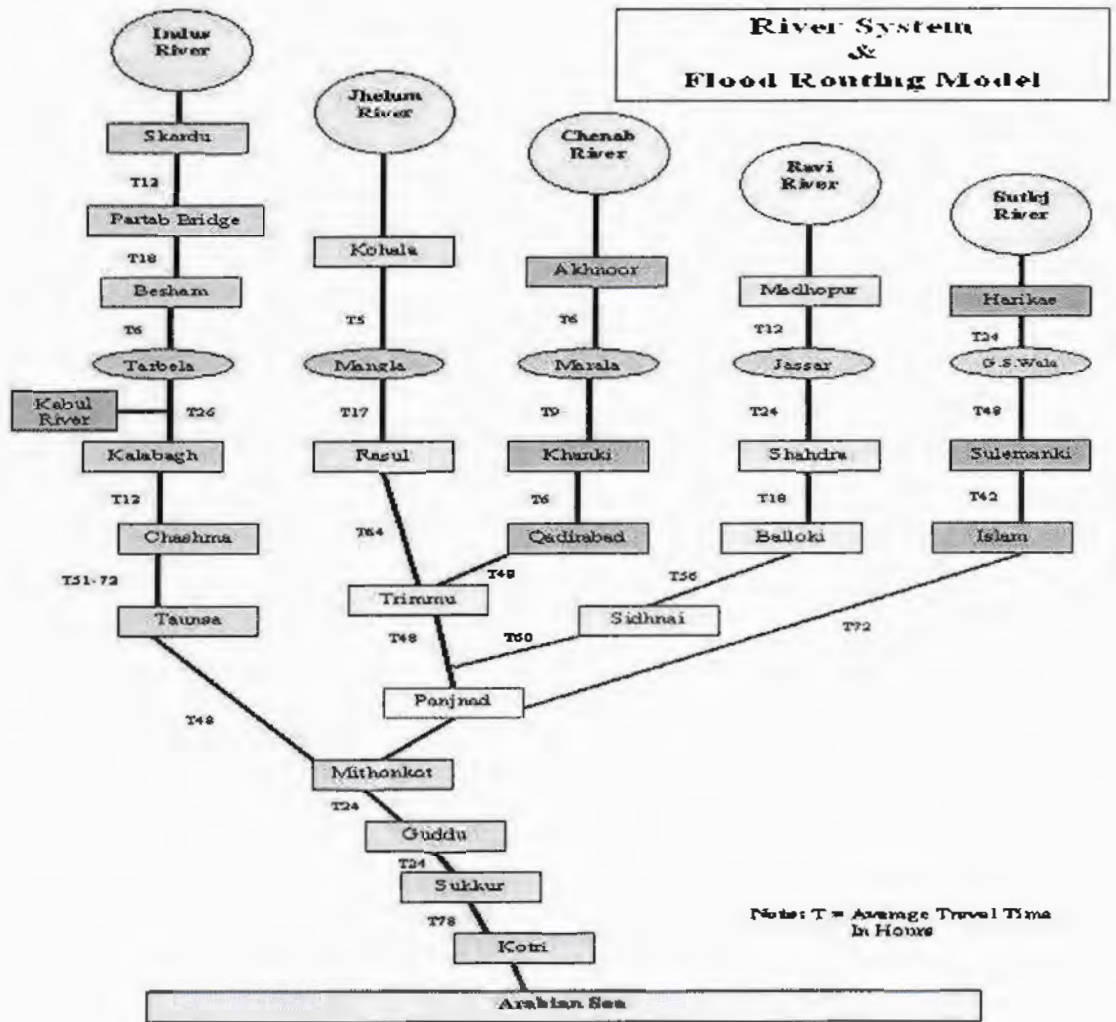
Figure 3.3.1: River system and flood routing model of Pakistan

20

3.4: Artificial Neural Network

Artificial Neural Networks (ANNs) have been influence accurate from their beginning by the appreciation that the human brain calculates in a completely different system from the normal digital computer. The brain is an extremely composite, non-linear, and similar computer. It has the skills to establish its physical constituents, identified as neurons in a tremendously distributed and equivalent network, to achieve certain calculations several times faster than the quicker digital computer system in presence nowadays. ANNs contain of various simple processing fundamentals called nodes or neuron. Every neuron is then associated to other neurons through direct links. Every link is then connected with a weight that denotes the magnitude of outward signal. The processing of every neuron is follow out in dual steps, first the weighted summation of the inputs is bringing and is trailed by the importance of activation function.

Artificial Neural Network are commonly applied by using electrical mechanisms or are imitation in a software on computer system. They are categorized by the designs of networks amongst the nodes (named as architecture), the procedures of defining weights on the networks (called their learning procedure), the activation function, and the number of layers: single layer, bilayer, and multilayer. When overall signals of a network move in only one direction, the ANN is known as Feed-Forward Network (FFN). Other than, they are known as Recurrent Networks. Generally, the frequently used artificial neural networks in hydrological models are Feed-Forward Multilayer Perceptron's (FFMLPs), which include three layers, and employ monotonic activation functions in all layers. Their weights are generally found by organized training by using back-propagation process. The performance and learning capability of artificial neural network depend upon the fitness of its architecture,

which is needed to be pre-specified, more precisely, the number and configuration of its hidden nodes (Imrie *et al* 2000). Composite non-linear problems such as time series data prediction can be solved through ANN that is a non-linear mathematical model. Multilayer Perceptron (MLP) is the type of Artificial Neural Network which is used to adjust the connection weight. MLP comprises three layers: input layer; hidden layer; and output layer. MLP has only one input and output layer but more than one hidden layer. Each of these layers comprises nodes. These nodes are associated to each other.

Input layer can contain that number of nodes as having by the data input study. Output layer, however, must have only one node to get output. It is up to user selection how to use Hidden layer. There must be an additional dummy neuron known by the name of bias nodes in input layer. Bias node acts as one of the players in the network. The activation function is done to calculate the arriving value and gives the value in output layer. So, the actual computational process may occur. Input layer applies linear transfer function. However, a sigmoid function is normally used in hidden and output layer. Following the activation function, we can determine the connection weight by data training.

The procedure for the adjustment of weight begins in the output layer and carries backward towards the input layer. Until the performance of target is obtained the process of feed-forward and back propagation algorithm continues. The weights of the values are determined when the data training process completes. Following this a single feed forward computation is employed. Results are then assessed by conducting performance measures to the actual versus the predicted values. Figure 3.4.1 shows the Multilayer Perceptron with three layers as stated before.

22

3.4.1: Multilayer Perceptron

A typical three-layer feed forward network is shown in the figure 3.4.1
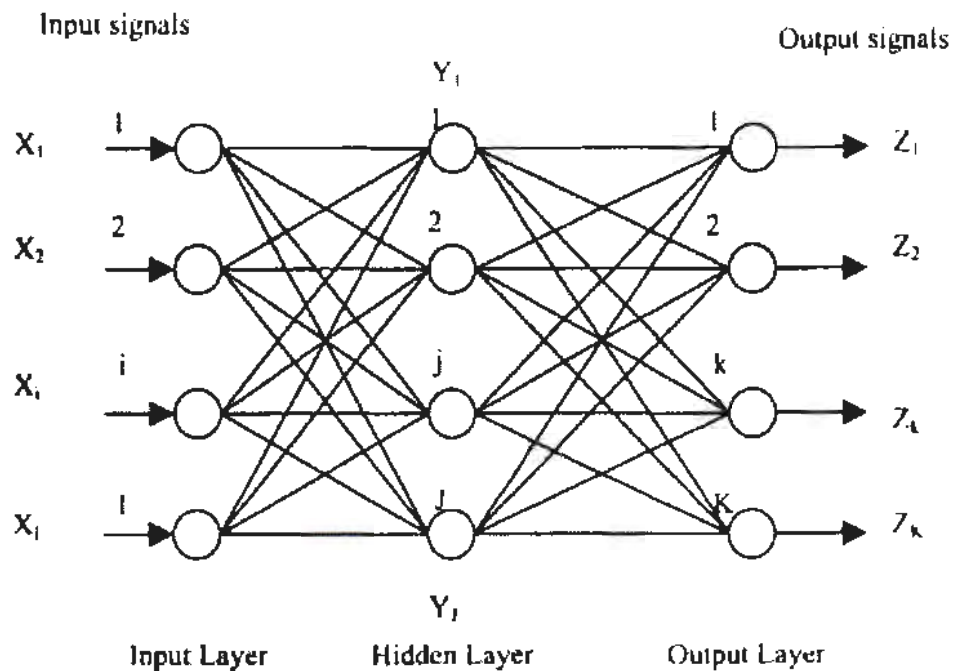
Fig. 3.4.1 Multilayer Perceptron with three layers

The above figure is a three-layer feed forward artificial neural network. The first layer, called the input layer, comprises I nodes and connects with the input variables. This layer performs no computation but is used to distribute i inputs into the network. The last layer is connected to the output variables, and is called the output layer. This layer consists of K output nodes.

One or more layers of processing units that have no direct relation to the outside variable and are located between the input and output layers are called hidden layers. An MLP may have many hidden layers. However, only one hidden layer will be considered here. This layer consists of J hidden nodes. In each layer, the processing elements are called nodes or neurons. Each of these neurons connected to all the neurons of adjoining layers. The parameters linked with these connections are

known as weights. Information used by the net to solve a problem is represented by weights. Generally, all connections are 'feed forward' that allow information to transfer from an earlier layer to the next consecutive layer. The neurons in a layer are not interconnected, and neurons in non-adjoining layers are not connected, however, for recurrent neural networks this is not the case.

Activation function

Input pattern

X_1

W_1

X_2 W_2

j

Y_j

Output pattern

W_i

X_i

W_I

X_I
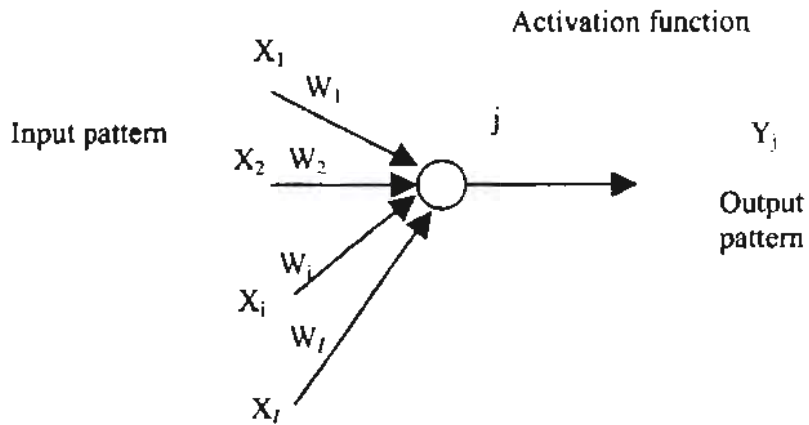
The architecture of a typical neuron is shown in the above Figure. Every hidden neuron j gets I arriving signals (x_i) from each neuron i in the preceding layer (for instance, from Input layer to Hidden layer). Connected with every arriving signal (x_i), a weight (w_{ij}) is connected between layer 1 and J. The effective arriving signal (net_j) to neuron j is the sum of weights of all the arriving signals as follows

$$net_j = \sum_{i=1}^{I} w_{ij}x_i \qquad (3.4.1)$$

This effective arriving signal (net_j) passes over a non-linear activation function (often called a transfer function or threshold function) that produces the outgoing signal which is called activation or activity level (y_j) of the neuron. Usually, a neuron sends its activation as a signal to several other neurons. It is pertinent to note that at

24

one time only one signal can be sent by a neuron, though that signal is programed to numerous other neurons simultaneously.

3.4.2 Delta Learning Rule for Feed-Forward Multilayer Perceptron

The delta learning rule is also called the LMS (Least Mean Square), error back propagation training algorithm. In the feed forward propagation training Input patterns are entered sequentially. If a pattern is entered and its classification or association comes out to be inaccurate, the current least mean square classification error is minimized through the adjustment of the synaptic weights as well as the thresholds. Therefore, the purpose of this algorithm is to adjust the synaptic weights w_{ij} (i=1, I) and w_{jk} {j=1 J), which connect Input to Hidden layer, and Hidden to Output layer respectively that will assure the minimization of the error function {E_k) in equation 3.4.2.1. The instantaneous estimate of the mean-square error calculated at the output layer is defined as follows (Rumelhart et al, 1986)

$$E_k = \frac{1}{2}(z_k - t_k)^2 \qquad (3.4.2.1)$$

Where z_k the output from output is layer and t_k is the target value. The outgoing signal (z_k) is a function of the activation as follows:

$$z_k = f(net_k) \qquad (3.4.2.2)$$

The effective incoming signal net_k, comprising weighted sum signals from the hidden layer (y_j) is calculated as follows:

$$net_k = \sum_{j=1}^{J} w_{jk} y_j \qquad (3.4.2.3)$$

The outgoing signal (y_j) is the result of the incoming signal net_j, being passed through the activation function, which is not necessarily the same as that for z_k, as follows:

$$y_j = f(net_j) \qquad (3.4.2.4)$$

Again, the effective incoming signal net_j, comprising weighted sum signals from input layer (x_i) is calculated as:

$$net_j = \sum_{i=1}^{I} w_{ij} x_i \qquad (3.4.2.5)$$

The derivative of the error functions with respective to w_{jk}, can be expressed after applying the chain rule, as follows:

$$\frac{\partial E_k}{\partial w_{jk}} = \frac{\partial E_k}{\partial z_k} \frac{\partial z_k}{\partial net_k} \frac{\partial net_k}{\partial w_{jk}} \qquad (3.4.2.6)$$

Finally, the updated weights are

$$w_{ij} = w_{ij} + \Delta w_{ij} \qquad (3.4.2.7)$$

3.5 Activation Function

Artificial neural network is a procedure to sum the product of the connected weights, and the input signals, and produce an output or activation transfer function. The activation function is the specific function for the input units. The nodes of a specific layer get the equivalent type of activation function. Generally, in every case a non- linear activation transfer function is used. There are several activation transfer functions employed in ANNs, the most commonly used ones being the following the unipolar binary function or sigmoid function (S), the bipolar binary function (B), the hyperbolic tangent function (T) and the linear function (L).

3.5.1 The unipolar binary function or sigmoid function (S)

The sigmoid flinction, as implemented by Hall & Minns (1993), Dawson & Wilby (1998), and Campolo et al (1999), is defined as

26

$$S(\text{net}_{j,}) = \frac{1}{1+\exp(-\lambda \text{net}_j)} \qquad (3.5.1)$$

Where $\lambda > 0$ in above sigmoid function. λ is in proportion to the nodes acquire defining the steepness of the function, which in general is set equal to one. The term $\lambda net_{j,}$ will lies between ($-\infty$ and $+\infty$), whereas S ($net_{j,}$) will lies between (0 and 1).

3.5.2 The bipolar binary function (B)

The bipolar function is denoted by $B(\text{net}_j)$ and is given by

$$B(\text{net}_j) = \frac{1-\exp(-\lambda \text{net}_j)}{1+\exp(-\lambda \text{net}_j)} \qquad (3.5.2)$$

Where $\lambda > 0$. Again, λ is in proportion to the nodes gain defining the steepness, of the function, which in generally set equal to one. The term $\lambda net_{j,}$ will lies between ($-\infty$ and $+\infty$). While B ($net_{j,}$) is bounded among (-1 and 1).

3.5.3 The hyperbolic tangent function (T)

The hyperbolic tangent function, as implemented by Varoonchotikul *et al.* (2002b), is defined as follows:

$$T(net_j) = tanh(\lambda net_j) \qquad (3.5.3)$$

Where $\lambda > O$. Once more λ is in proportion to the nodes acquire defining the steepness, of the function, which in general is set equal to one. The term λnet_j will lies between ($-\infty$ and $+\infty$), wheras $T(net_j)$ is bounded among (-1 and 1).

3.5.4 The Linear function (L)

The linear function, as implemented by Karunanithi *et al.* (1994), is defined as follows:

$$L(net_j) = \lambda net_j \qquad (3.5.4)$$

where $\lambda > 0$. λ is in proportion to the nodes acquire defining the steepness of the function, which in general is set equal to one. The term λnet_j will lies between $(-\infty$ and $+\infty)$, wheras $(T(net_j)$ is bounded among (-1 and 1).

3.6: Wavelet

3.6.1 Introduction

A class of functions called Wavelets are used for a given function localization both in its position and scaling. Signal processing and time series are the domain of Wavelets function. It divides the data functions or operators into various parts, which bases upon Wavelets, and studies every part with a resolution matched to its scale (Daubechies [1]). Scale and time are two variables on which the Wavelet transform depends while studying signal processing. Two primary types of wavelet transforms are continuous (CWT) and discrete Wavelet Transform (DWT).

The CWT works with functions defined over the whole real axis while the DWT deals with functions which are defined over a range of integers. Wavelets are mathematical processes which represent a time-scale of the time series and their relationships to analyze time series when the data is non-stationary. Wavelet analysis is helpful in the long-time intervals where the frequency information's are low and in the shorter intervals where the frequency information is high.

3.6.2. Definition

Wavelet is a small wave function denoted by the notation ψ ($\cdot$). It is a small wave which is in contrast with waves that are large wave for example sinc wave that usually grows and decays in an infinite period of time and similarly a wave that is small, grow and decay in stipulated period of time. The range of a wavelet mother

28

function (ψ ($\cdot$)) lies between $-\infty$ and ∞, and it must fulfill the following given properties:

(1) The value of the integral of ψ ($\cdot$) is always equal to zero:

$$\int_{-\infty}^{+\infty} \Psi(u)du = 0 \tag{3.6.2.1}$$

(2) The integral of the square of ψ ($\cdot$) is unity:

$$\int_{-\infty}^{+\infty} \psi^2(u)du = 1 \tag{3.6.2.2}$$

(3) Admissibility Condition: $\Psi(f)|$

$$C_\Psi(t) = \int_0^\infty \psi^* \frac{[\Psi(f)]^2}{f} df \quad \text{satifies} \quad 0 < C_\Psi < \infty \tag{3.6.2.3}$$

Equation (3.6.2.1) shows that the value of the wavelet function ψ above zero must be cancelled out by journey below zero. Visibly, the function 3.6.2.1 fulfill the property 1, but equation (3.6.2.2) tells that mother wavelet function (ψ) must have some value different from zero. If this law is satisfied, then from the continuous wavelet transform we may construct again the signal under analysis.

Haar created a wavelet function in 1910 (Figure 3.6.2.1), which was named Haar wavelet function and considered one of the oldest wavelet functions:

$$\psi^{(H)}(u) = \begin{cases} +1 & if \quad 0 \le u < \frac{1}{2} \\ -1 & if \frac{1}{2} \le u < 1 \\ 0 & else \end{cases} \tag{3.6.2.4}$$

29

Wavelet analysis
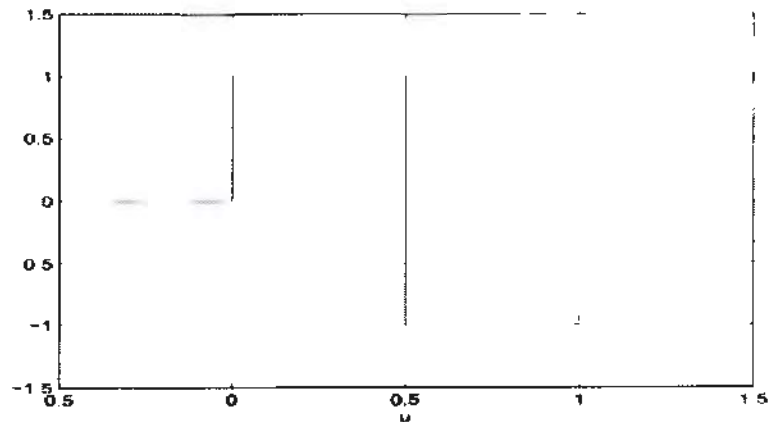
Figure 3.6.2.1: The Haar wavelet function

Haar wavelet is the earliest wavelet. To discuss any wavelets Haar's name comes first. It is a discontinuous wavelet and looks like a step function. Harr wavelet and Daubechies db1 are the same. In wavelet research Ingrid Daubechies is well known, he developed efficiently supported orthonormal wavelets which give high support to discrete wavelet analysis and make it feasible. Daubechies family wavelets are written as dbN, where N stands of the order of Daubechies family and db stands for the sure name of the wavelet.

3.6.3 Wavelet Analysis

Although we define a wavelet function but we see how it functions. For the Fourier analysis, we look first, which is the structure of the function under analysis in terms of sinusoidal waves of different frequencies. This condition is sufficient that the function we are using is stationary.

In case, there are singularities and changes in frequency as often the case, the Fourier analysis ceases to function. Here we get the average values of the changing frequencies that are not useful to us. When we match a wavelet function of varying scales and positions than we know that how a function that is given differ from one

30

time period to another. There is more flexibility in wavelet analysis due to which we can choose a specific Wavelet for matching the type of function under analysis.

The function ψ ($\cdot$) is the mother wavelet and a double wavelet index family can be constructed by translating and dilating this mother wavelet.

$$\psi_{\lambda,t}(u) = \frac{1}{\sqrt{\lambda}} \, \psi \left(\frac{u-t}{\lambda}\right) \qquad (3.6.3.1)$$

where $\lambda > 0$ and t is finite.

The standardization on the right-hand side of equation (3.6.3.1) is

$$\| \, \psi_{\lambda,t} \, \| = \|\psi\| \text{ for all } \lambda, \text{ t.}$$

Functions are considered linear combination where the information is secure and make no losses to those functions that are in discrete form or as an integral in the continuous form of the selected wavelet family.

3.6.4 Continuous Wavelet Transform

Over the continuous time if for the transformation function or signal x($\cdot$) is defined than the continuous wavelet transform (CWT) can be used. The parameters λ and t continuously vary which are used for making the wavelet family. The transform is to compute the amplitude coefficients for dilation (λ) and for translation (t) of the mother wavelet ψ(.), is the best fit to signal x($\cdot$) which makes $\psi_{\lambda,t}$, such that to integrate the product of the signal:

$$\left(x, \psi_{\lambda,t}\right) = \int_{-\infty}^{+\infty} \psi_{\lambda,t} \, (u)x(u)du \qquad (3.6.3.2)$$

We can make a shape of the wavelet function how does it fit the signal when we change the λ from one dilation to another. If we want to know variation in the nature

31

of the signal over time we have to change t. Here we have these coefficients $\{< x,$

$\psi_{\lambda,t} > | \lambda > 0, -\infty < t < \infty\}$ and can be called CWT of x ($\cdot$).

To keep all necessary information from original signal is a useful property of the CWT. If the admissibility condition (equation 3.6.2.3) is fulfilled by the wavelet function ψ ($\cdot$) and the signal x ($\cdot$) satisfies

$$\int_{-\infty}^{+\infty} x^2\,(t)dt < \infty$$

Using below inverse transform, x(.) can be recovered from its CWT:

$$x(t) = \frac{1}{C_\psi} \int_{0}^{+\infty} \left[\int_{-\infty}^{+\infty} \{x, \psi_{\lambda,u}\} \psi_{\lambda,u}\,(t)du \right] \frac{d\lambda}{\lambda^2} \tag{3.6.3.3}$$

In equation (3.6.2.3) $C\psi$ is defined above.

The signal x ($\cdot$) and its CWT belongs to the same family with different illustrations. The CWT is used to get x ($\cdot$) in a unique way, through this process we can make access to the hidden insights of the signals.

3.6.5 Discrete wavelet transform

To use the CWT under the given analysis of a signal can gives plenty of information. With the help of dilations and translations of the mother wavelet the signal is analyzed. Clearly, in the CWT redundancy will be many more. Considering only the subsamples of the CWT we will be able to keep intact the key properties of the transform and this can lead us to the DWT. To work on a discrete sample information or time series x ($\cdot$) the DWT is used, where t = 0, 1, ... N–1 time to be finite.

With the help of DWT, it is easy to analyze the time series for discrete dilations and translations of ψ ($\cdot$). Generally, to get the dilation values λ, dyadic scales are used, such that λ is in the form of 2^{j-1}, $j = 1, 2, 3 \ldots$). Analyzing the translation values t within a dilation of 2^{j-1}, is then sampled at 2^j interval. Scaling and wavelets are the two sets of functions that is viewed as high-pass and low-pass filters respectively are operated by Discrete wavelet transform (DWT). In the analysis of the high frequency and low frequency the original time series are passed through high-pass and low-pass filters respectively and are separated at different scales.

The low frequency signal is of greater importance than the high frequency signal. As an input for the next decomposition stages the output of the low-pass filter is used whereas at the time of signal reconstruction the output of high-pass filter is used. By using a series filtering and down sampling processes the wavelet coefficients are calculated.   According to Pour Arian *et al* (2016), two trends of deconstruction of DWT, one is the approximation that is passes on low frequencies and the other is details passes on high frequencies (see in figure 3.6.5.1).
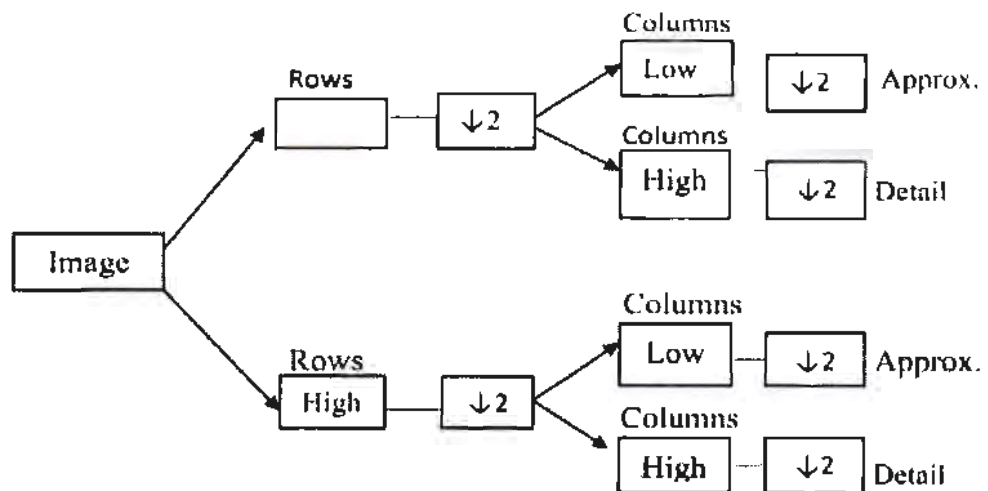
Figure 3.6.5.1: Low-Pass & High-Pass filters in DWT

33

Beginning from signal s with a length of N, the sets: approximation coefficients (CA_l) and detail coefficients (CD_l) are calculated. The vectors (CA_l) and (CD_l) are gained by involving, with low-pass filter (Lo_D) for approximation and with high-pass filter (Hi_D) for detail respectively, which is followed by dyadic decision.

More specifically, the initial step is presented by S.G Mallat (1989) is given below
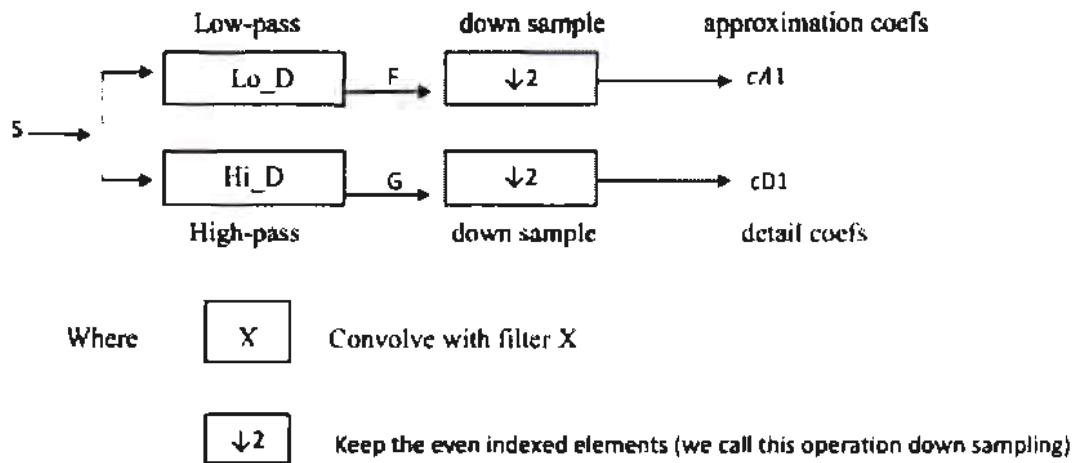
Figure; 3.6.5.2 Low-Pass (Approx.) and High-Pass (Details) coefficients of DWT

Each filter is a length of 2L. For a signal with length of N, the length $N + 2L - 1$ are of signals F and G, and then the length of the coefficients CA_l and CD_l are as below.

$$\left[\frac{N-1}{2} + L\right]$$

3.7: Wavelet-ANN Conjunction Model for Flood Events Prediction

The wavelet analysis is connected to the Artificial Neural Network strategy for annual maximum peak flow prediction in this research, which is shown in figure 3.7.1. First of all, the measured annual maximum peak flow time series data, Q were decayed into several multi-frequencies, time series including of details (high frequency, low scale), which is $Q_{D1}(t)$; $Q_{D2}(t)$; ......$Q_{Di}(t)$, and approximate (high scale, low frequency) – $Q_{Da}(t)$ by Desecrate Wavelet Transform (DWT). The symbol

34

Di shows the Wavelet-ANN Model for Flood prediction. Level i decayed details time series, where a denote approximation time series.

However, the decayed Q(t) time series data were used as inputs to the Artificial Neural Network (ANN) model. The original data of time series in the next step, Q(t+1) is output to the Wavelet-ANN conjunctional model. The graphical representation of above methodology is shown as in Figure 3.7.1.
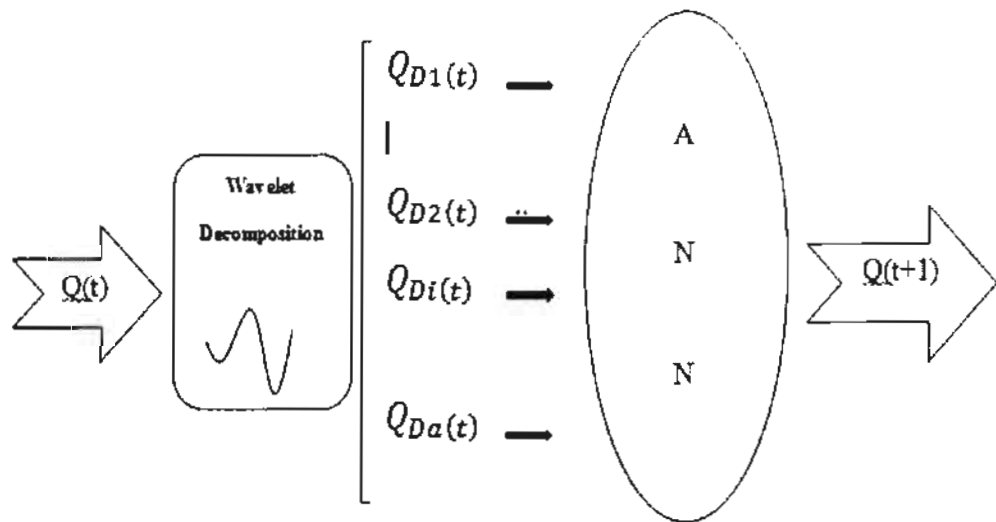
Fig. 3.7.1: Schematic diagram of Wavelet-ANN conjunction Model

3.8: To obtained the optimal architecture the following parameters will be used

3.8.1: Learning Rate

There are many of the learning algorithms but, learning rate is a commonly known parameter in the learning algorithms. Where the ANN arrives at minimum solutions and this speed is affected by the Learning rate. The learning rate is same to the step-size parameter from the gradient-descent algorithm is used in backpropagation. If the learning rate is very high the system will contrast with the true

35

solution or it will be completely diverged and reverse is the case in a very low learning rate, the system take enough long time in the convergence of final solution.

3.8.2: Momentum Parameter

To avoid the system convergence to a local minimum or saddle point the momentum parameter is generally used it is also useful in the increase of speed of convergence. But keep in mind that the settings of the momentum parameter should not be too much high otherwise the system will be unstable because a high momentum parameter creates the risk of overshooting of the minimum. If the momentum parameter is sufficient low than it become unable to prevent the local minimum and also make slow the training system.

3.8.3: Iteration

Typically, you'll split your test set into small batches for the network to learn from, and make the training go step by step through your number of layers, applying gradient-descent all the way down. All these small steps can be called *iterations*.

3.8.4: Training set

The training set assign and adjust weights to neural network. Training parameter is a solution to the problem in size of weight and also solves the issue of biasedness in learning of the training algorithm.

3.8.5: Validation Set

Over fitting is also a common problem in neural networks, therefore, to minimize the over fitting problem the validation set may be used. It is not possible to adjust the weights of the network with validation set, but we can only verify the

increase in accuracy over the training data set. This set is responsible for the increase in accuracy that is not known to the network before or the system is not trained on it.

During the training data set if the accuracy increases and the accuracy during validation data set remain the same or declining than we are facing with the problem of over fitting and such cases we should stop the training.

3.8.6: Testing Set

To verify the actual predictive neural networks' power in the final solution the testing set is used.

3.9: Goodness of fit of test of the Models

To evolution the capability and consistency of the models, it is compulsory to clever the capabilities of artificial neural networks in both training and estingperiods. Several statistical techniques are generally used to decide the capability of the Neural Network and Wavelet Neural Network (W-ANN).

3.9.1: Root Mean Square Error (RMSE)

The positive square root of the mean/average of the square of all the error. RMSE is a measure of how spread out these residuals is. Root mean square error (RMSE) is a commonly used quantity of the variances among the predicted values. The smaller RMSE shows that the model is best fitted. The formula of RMSE is given by

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2} \qquad (3.9.2)$$

Where x_i represent the errors and $\hat{x}_i$ is average of the errors.

37

3.9.2: Mean Absolut Error

MAE calculates the average size of the errors, in a group of predictions, deprived of think about their direction. The average of mean absolute error is the absolute differences among estimated and observed values where all separate differences, have equivalent weight. The MAE is a statistical term which is used to measure how close actual and predicted outcomes. The minimum mean absolute error indicate that the model is best fitted. The formula of mean absolute error is shown as given below.

$$MAE = \frac{1}{n}\sum_{j=1}^{n}|y_i - \hat{y}_j| \qquad (3.9.3)$$

Where y_i is the actual value and $\hat{y}_i$ is the predicted values.

3.9.3: Pearson correlation coefficient (r)

Correlation – frequently compute as a coefficient of correlation – designates the effectiveness and direction of the linear association among two variables (i.e, model input and output values). There are different coefficients which are applied for various circumstances. The Pearson correlation coefficient is the best one, which is achieved by split the co-variance of the two random variables by the multiplication of their standard deviations. The Pearson correlation coefficient is given by the formula.

$$R = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2 \sum_{i=1}^{n}(y_i-\bar{y})^2}} \qquad (3.9.7)$$

If the association is +1 then there is a perfect positive linear relationship between model and observations, and -1 means that there is perfect negative linear relationship between actual and estimated values. If correlation is 0 then we say that there is no linear relationship amongst observed and predicted values. The square root of

correlation coefficient (r^2), identified as the determination coefficient, explains how many of the variance among the dual random variables is defined by the linear fitted model.

3.9.4: Nash-Sutcliffe coefficient (E)

To evaluate the prediction of hydrological, model this co-efficient is frequently used. However, it can also be used to quantitatively pronounce the correctness of model outputs for additional things than discharge (i.e. Temperature, Absorptions etc.). The Nash-Sutcliffe coefficient (E) is defined by the formula.

$$E = 1 - \frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model})^2}{\sum_{i=1}^{n}(X_{obs,i} - \bar{X}_{obs})^2} \qquad (3.9.8)$$

Where $X_{obs,i}$ indicate the observed values and X_{model} indicate the model or predicted values at time i.

The efficiencies of Nash-Sutcliffe will lie between $-\infty$ to 1. An effectiveness of $E = 1$ indicates the perfect match among observed and predicted values. An effectiveness of 0 shows that the predictions of model are as correct as the mean average of the actual data, however an effectiveness fewer than zero ($i.e - \infty < E < 0$) arises when the experimental mean is better predictor as compare to the fitted model. Basically, the fitted model will be more accurate when the model efficiency (E) is approaches to one.

3.9.5: Index of Agreement (IOA)

It seeks to adapt the coefficient of Nash–Sutcliffe by punishing the variations in the mean of observed values and predicted values. Though, due to further squaring expressions, this index is very tactful to outliers in a set of data.

The term Index of Agreement established by Willmott (1981) as a standard normal quantity of the standard of the prediction error of the model and well lies among 0 and 1. The value of $IOA = 1$ shows an accurate match, where 0 shows no agreement at all (Willmott, 1981). It is defined as

$$IOA = 1 - \frac{\Sigma(x_{a} - x_{p})^{2}}{\Sigma[|(x_{a} - \bar{x}_{a})| + |(x_{p} - \bar{x}_{p})|]^{2}}$$

(3.9.9)

Where x_a is measured and x_p is computed value, where as $\bar{x}_a$ and $\bar{x}_p$ are the average values of x_a and x_p values respectively.

CHAPTER 4

RESULTS AND DISSCUSION

4.1: Applications of Wavelet-ANN conjunctional models

Annual peak flows values for 54 years (1960-2013) and 89 years (1925-2013) of Mangla and Khanki sites respectively on River Jhelum in Pakistan are used for the performance Evaluation of Artificial Neural Network and Wavelet-ANN Conjunctional models. Eighty percent (43 years) of annual maximum peak flow time series (1960-2002) for Mangla and eighty percent (71 years) of annual maximum peak flow time series (1925-1995) for Khanki are utilized for training ANN and Wavelet-ANN conjunctional model. Remaining 20 percent (2003-2013) 11years and (1996-2013) 18 years for Mangla and Khanki respectively are utilized for testing the trained models. The characteristics of original annual maximum series values in time series obtained as approximation sub-signal and detail sub-signal by Discrete Wavelet-transform (Haar Wavelet level 1) are inputs to Wavelet-Artificial Neural Network (W-ANN) for the sites Mangla and Khanki respectively.

Fig.4.1.2 and Fig.4.1.3 show approximation and detail Sub-signal of the annual maximum peak flow original series obtained by db1 (Haar wavelet level1) using Discrete Wavelet Transform (DWA) for Mangla site. Also Fig. 4.1.7 and 4.1.8 show approximation and detail Sud-signal of the annual maximum peak flow original series obtained by db1 (Haar wavelet level1) using Discrete Wavelet Transform (DWA) for Khanki site. Approximation and detailed sub-signal for wavelet single level (db2) are shown in Fig. 4.1.4 and Fig.4.1.5 for mangla site. Also, approximation

41

and detail sub-signal for wavelet single level (db2) are shown in Fig.4.1.9 and Fig.4.1.10 for Khanki site. Two Wavelet-Artificial Neural Network (W-ANN) models for Mangla site consistent each of db1 and db2 are existing while the other two Wavelet-Artificial Neural Network (W-ANN) models for Khanki site consistent each of db1 and db2 are presented.

Using db1 transform of Mangla as inputs for first W-ANN model while db2 transform of Mangla considered as inputs for second W-ANN model for Mangla site. And for Khanki site using db1 transform of Khanki as inputs for first W-ANN model while db2 transform of Khanki considered as inputs for second W-ANN model. In the following section the evaluation of good performance for all models are discussed.
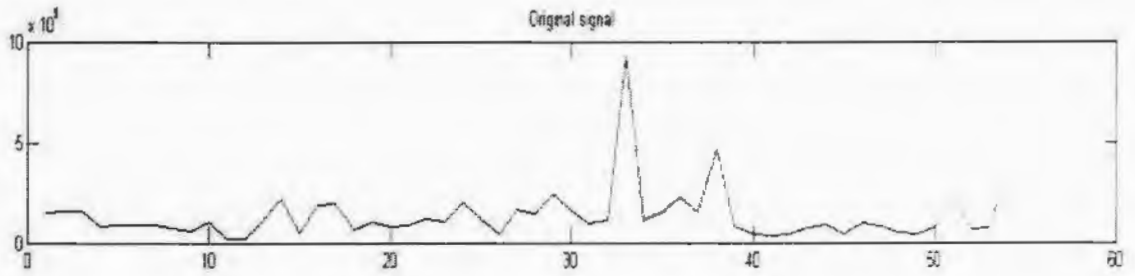
Fig: 4.1.1 Yearly AM flows at Mangla site
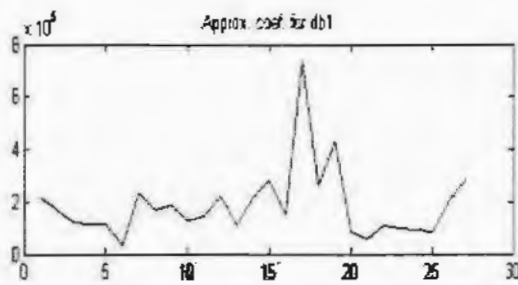

Fig:4.1.2 Approx. sub-signal of yearly
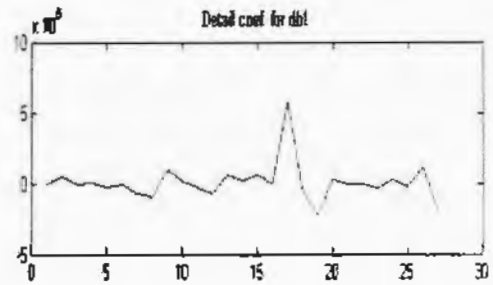
AM flows by DWT db1 at Mangla site.

Fig:4.1.3 Detail sub-signal of yearly AM

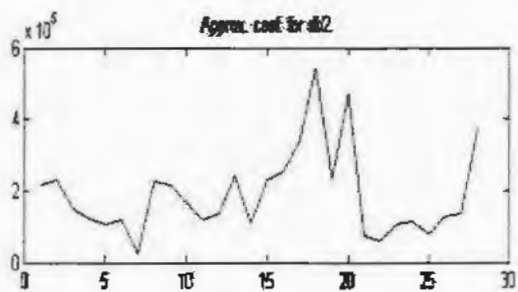Flows by DWT db1 at Mangla site.

Fig:4.1.4 Approx. sub-signal of yearly

AM flows by DWT db2 at Mangla site.
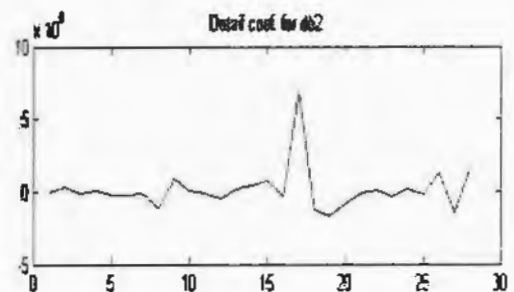
Fig:4.1.5 Detail sub-signal of yearly AM

Flows by DWT db2 at Mangla site.

43

4.2: Comparision of the results for T-ANN and W-ANN conjunctional Models (db1 and db2) for Mangla site

Wavelet-Artificial Neural Network conjunctional models (W-ANN) are nominated employing methodology depicted in Fig 3.6.1. Back propagation neural network training algorithm is employed on MATLAB platform. Internal parameters used are: momentum coefficient = 0.5; number of iteration = 1000 to obtain the ideal architecture.

In addition, optimum combinations of activation functions in the hidden and output nodes are obtained with sigmoid function. Performance of Wavelet-Artificial Neural Network (W-ANN) conjunctional model with two inputs, two hidden and one output is compared with one step ahead prediction by actual time series ANN model (T-ANN) with one input, two hidden and one output. To fit T-ANN and W-ANN conjunctional model (db1), the input and output nodes will be used.

The original annual maximum peak flow values in (Fig.4.1.1) are considered the input for T-ANN model with two hidden nodes and one output node, while the approximation and detail sub-signal coefficients given in (Fig.4.1.2 and Fig 4.1.3) are considered as two input nodes for W-ANN conjunctional model (db1) with two hidden and one output node. Table 4.2.2 show that R, RMSE, E, IOA and MAE are obtained from T-ANN and W-ANN conjunctional modle (db1) for Mangla site.

45

Table: 4.2.1 Training and testing errors for ANN and W-ANN (db1) for Mangla site on Jhelum

Models	Errors					
	Training/ Testing	R	RMSE	E	IOA	MAE
W-ANN sigmoid (2-2-1)	Training	0.6394	1.63E+04	0.4914	0.8965	1.83E+04
	Testing	0.5631	3.85E+04	0.3085	0.7382	2.96E+04
T-ANN sigmoid (1-2-1)	Training	0.4963	2.13E+04	0.2168	0.7317	1.96E+04
	Testing	0.4891	4.26E+04	0.0938	0.5861	3.05E+04

The characteristics of original annual maximum series obtained as approximation sub-signal and detail sub-signal by Discrete Wavelet-transform (Haar Wavelet level 1) are inputs to Wavelet-Artificial Neural Network (W-ANN) for Mangla site. The actual annual maximum values (Fig. 4.1.1) in present time step are the output to W-ANN models for Mangla station. The previous steps actual value of annual peak flow in the time series data is input and the value of annual peak magnitude in present time step is output to T-ANN models for Mangla site. Therefore, the output of both models (W-ANN) and (T-ANN) are same for Mangla.

Experimentation with varying number of hidden nodes and training algorithm are performed. The error statistical terms of best performing W-ANN (2-2-1) and T-ANN (1-2-1) model in training and testing are shown in table 4.2.1.

From table 4.2.1, it is observed that prediction for testing data is not well as compared to the training data based on statistical error terms for both models W-ANN (db1) and T-ANN. From table 4.2.1, the correlation coefficient R (0.6394 and 0.5631)

for training and testing data of W-ANN db1 model is greater than the correlation coefficient R (0.4963 and 0.4891) for training and testing data respectively of T-ANN model. The root mean square error RMSR of W-ANN db1 model for training and testing are (1.63E+04 and 3.85E+04) which is less than the RMSE (2.13E+04 and 4.26E+04) of T-ANN model for training and testing data respectively.

Nash-Sutcliffe coefficient E (0.4914, 0.3085) and Index of Agreement IOA (0.8965, 0.7382) of W-ANN db1 model for both training and testing data are greater than the E (0.2168, 0.0938) and IOA (0.7317, 0.5861) of T-ANN model for both training and testing data respectively. The MAE (1.83E+04 and 2.96E+04) for training and testing data of W-ANN db1 model are less than the MAE (1.96E+04 and 3.05E+04) for training and testing data of T-ANN model respectively. Therefore, due to goodness of fit tests (R, RMSE, E, IOA, & MAE) W-ANN conjunctional model (db1) with two input, two hidden, and one output nodes representing as 2-2-1 is selected as best performing model as compare T-ANN model with (1-2-1).

Also, Fig.4.2.1 given below compares the actual and predicted peak flow from T-ANN model. Low flows were accurately forecasted with low flows while high flows being somewhat less accurate. Fig. 4.2.2 given below compares the actual and predicted peak flow from W-ANN (db1) model. Low flows, Medium flows and high flows were approximately accurately forecasted. W-ANN (db1) model forecasts low, medium and high flows for Mangla site is well for a lead-time of 1 year. Therefore Wavelet-Artificial Neural Network (W-ANN db1) model with two input, two hidden and one output nodes is selected as best performing model as compare to T-ANN model with one input, two hidden and one output node.
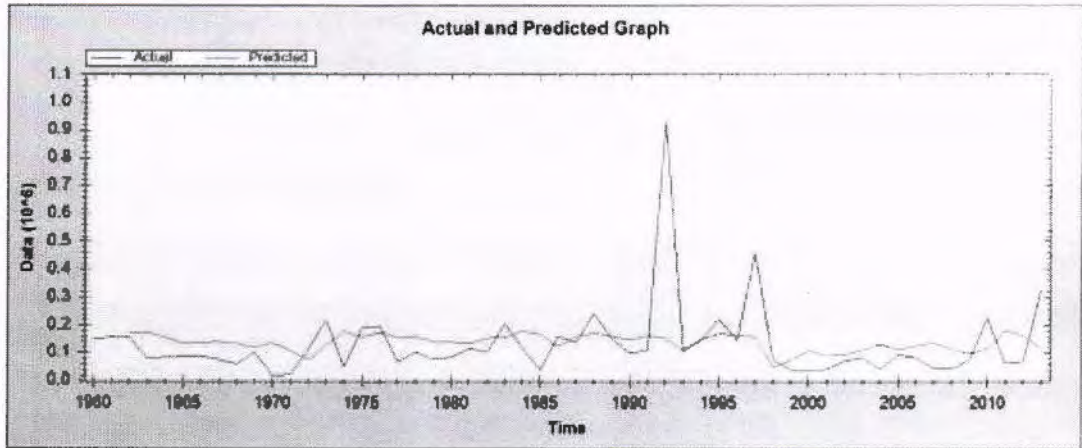
Fig 4.2.1: Actual vs predicted graph of T-ANN model for Mangla site
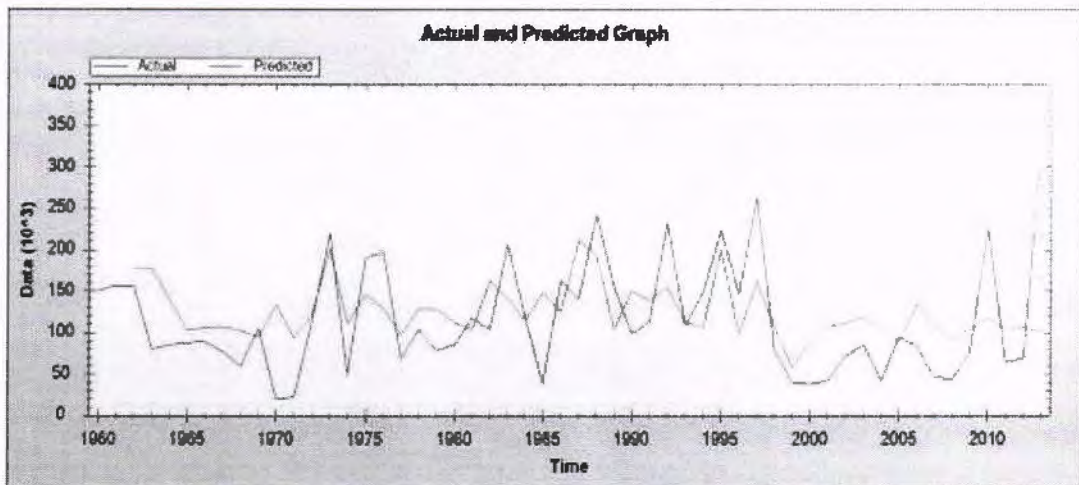
Fig 4.2.2: Actual vs predicted graph of W-ANN (db1) at level1 for Mangla site.

Results further upgraded when approximation and detail sub-signal of the annual maximum peak flow original series are obtained using DWT by db2 wavelet (level 1) as presented in table 4.2.2 below. Compare W-ANN db1 model (table 4.2.1) with W-ANN db2 model (table 4.2.2) for further improvement.

To fit W-ANN conjunctional model (db2) the inputs and output used are: The approximation and detail sub-signal of yearly peak measurement values by DWT db2 wavelet level1 given in (fig.4.1.4 and fig 4.1.5) are considered two input nodes for W-ANN conjunctional model (db2) with two hidden and one output node. Table 4.2.2

48

shows that R, RMSE, E, IOA and MAE are obtained from T-ANN and W-ANN conjunctional modle (db2) for Mangla site.

Table: 4.2.2 Training and testing errors for ANN and W-ANN (db2) for Mangla site on Jhelum

Models	Training/Testing	R	RMSE	E	IOA	MAE
Errors						
W-ANN Sigmoid (2-2-1)	Training	0.7984	1.21E+04	0.6362	0.9135	1.04E+04
	Testing	0.6108	2.99E+04	0.4158	0.7916	2.11E+04
T-ANN Sigmoid (1-2-1)	Training	0.4963	2.13E+04	0.2168	0.7317	1.96E+0R
	Testing	0.4891	4.26E+04	0.0938	0.5861	3.05E+04

Table 4.2.2 show the results of W-ANN (db2) model for Mangla site which clearly show that the correlation coefficient (R) is improved from 0.6394 to 0.7984 for training data where it is improved from 0.4963 to 0.6108 for testing data. The value of root mean square errore (RMSE) are reduced from 1.63E+04 to 1.21E+04 for training and from 3.85E+04 to 2.99E+04 for testing data.

The statistical term Nash-Sutcliffe coefficient E are also improved from (0.4914, 0.3085) to (0.6362, 0.4158) for training and testing data correspondingly where the term Index of Agreement IOA are improved from (0.8965, 0.7382) to (0.9135, 0.7916) for training and testing data respectively, while mean absolute error are reduced from (1.83E+04, 2.96E+04) to (1.04E+04, 2.11E+04) for training and testing data respectively. It is clearly observed from table 4.2.2 that the performance

of W-ANN (db2) model is best as compare to the W-ANN db1 model for Mangla station on river Jhelum.
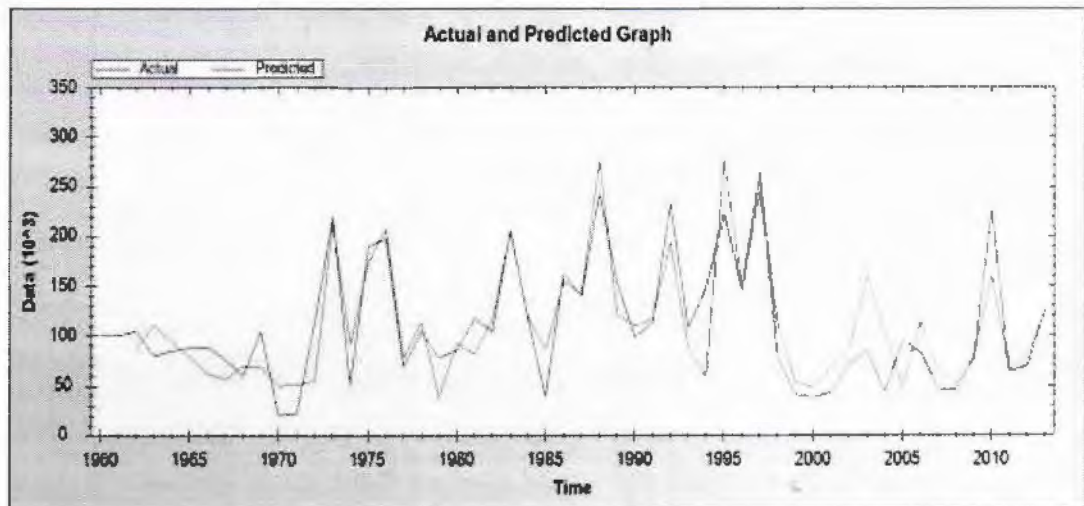
Fig 4.2.3: Actual vs Predicted graph of W-ANN (db2) at level1 for Mangla site

Fig.4.2.3 compares the actual and predicted peak flow from W-ANN (db2) model. Low flows, Medium flows and high flows were accurately forecasted. It can be seen that the W-ANN db1 model less forecast the annual peak flow of Mangla stations, while the WA–ANN db2 model provides closer estimates to the corresponding observed peak flows.

It is clear from Fig 4.2.2 and Fig 4.2.3 above that the W-ANN (db2) model forecasts low, medium and high flows for Mangla site is very well for a lead-time of 1 year as compare to W-ANN (db1) model forecasts for low, medium and high flows. The result improved that when the approximation and detail sub-signal of the annual maximum peak flow original series obtained using DWT by wavelet (level 1) as presented in table 4.2.2.

50

4.3: Comparision of the results for T-ANN and W-ANN conjunctional Models (db1 and db2) for Khanki site

Wavelet-Artificial Neural Network conjunctional models (W-ANN) are nominated employing methodology depicted in Fig 3.6.1. Back propagation neural network training algorithm is employed on MATLAB platform. Internal parameters used are momentum coefficient = 0.5 number of iteration = 1000 to obtain the ideal architecture. In addition, optimum combinations of activation functions in the hidden and output nodes are obtained with sigmoid function. Performance of Wavelet-Artificial Neural Network (W-ANN) conjunctional model with two inputs, two hidden and one output is compared with one step ahead prediction by actual time series ANN model (T-ANN) with one input, two hidden and one output. To fit T-ANN and W-ANN conjunctional model (db1), the input and output nodes will be used are.

The original annual maximum peak flow values in (Fig.4.1.6) are considered the input for T-ANN model with two hidden nodes and one output node, while the approximation and detail sub-signal coefficients given in (Fig.4.1.7 and Fig 4.1.8) are considered as two input nodes for W-ANN conjunctional model (db1) with two hidden and one output node. Table 4.3.1 show that R, RMSE, E, IOA and MAE are obtained from T-ANN and W-ANN conjunctional model (db1) for Khanki site.

Table: 4.3.1 Training and testing errors for ANN and W-ANN (db1) for Khanki site on Jhelum

Models	Errors					
	Training/Testing	R	RMSE	E	IOA	MAE
W-ANN Sigmoid (2-2-1)	Training	0.8021	2.63E+04	0.2991	0.7701	2.01E+04
	Testing	0.6903	4.79E+04	0.1653	0.6013	3.71E+04
T-ANN Sigmoid (1-2-1)	Training	0.7018	3.31E+04	0.1369	0.6003	3.76E+04
	Testing	0.2163	5.98E+04	0.0348	0.5193	4.07E+04

The characteristics of original annual maximum series obtained as approximation sub-signal and detail sub-signal by Discrete Wavelet-transform (Hear Wavelet level 1) are inputs to Wavelet-Artificial Neural Network (W-ANN) for Khanki station. The actual annual maximum values (Fig. 4.1.6) in present time step are the output to W-ANN models for Khanki station. The previous steps actual value of annual peak flow in the time series data is input and the value of annual peak magnitude in present time step is output to T-ANN models for Khanki station. Therefore, the output of both models (W-ANN) and (T-ANN) are same for Khanki.

Experimentation with varying number of hidden nodes and training algorithm are performed. The error statistical terms of best performing W-ANN (2-2-1) and T-ANN (1-2-1) model in training and testing are shown in table 4.3.1. From table 4.3.1, it is observed that prediction for testing data is not well as compared to the training data based on statistical error terms for both models W-ANN (db1) and T-ANN. Table 4.3.1, clearly show that the correlation coefficient R (0.8021 and 0.6903) for training and detail data of W-ANN db1 model is greater than the correlation

coefficient R (0.7018 and 0.2163) for training and testing data respectively for T-ANN model.

The root mean square error RMSR of W-ANN db1 model for training and testing data are (2.63E+04 and 4.79E+04) which is less than the RMSE (3.31E+04 and 5.98E+04) of T-ANN model for training and testing data respectively. Nash-Sutcliffe coefficient E (0.2991, 0.1653) and Index of Agreement IOA (0.7701, 0.6013) of W-ANN db1 model for both training and testing data are greater than the E (0.1369, 0.0348) and IOA (0.6003, 0.5193) of T-ANN model for both training and testing data respectively.

The MAE (2.01E+04 and 3.71E+04) for training and testing data of W-ANN db1 model are less than the MAE (3.76E+04 and 4.07E+04) for training and testing data of T-ANN model respectively. Therefore, due to goodness of fit tests (R, RMSE, E, IOA, & MAE) W-ANN conjunctional model (db1) with two input, two hidden, and one output nodes representing as 2-2-1 is selected as best performing model as compare to T-ANN model with (1-2-1).

Also, Fig.4.3.1 given below compares the actual and predicted peak flows from T-ANN model for Khanki station on Jhelum. The medium flows were accurately forecasted with medium flows while high, and low flows being somewhat less accurate. Fig. 4.3.2 given below compare the actual and predicted peak flow from W-ANN (db1) model. Low flows, Medium flows and high flows were approximately accurately forecasted. W-ANN (db1) model forecasts low, medium and high flows for Khanki site is well for a lead-time of 1 year.

Therefore Wavelet-Artificial Neural Network (W-ANN db1) model with two input, two hidden and one output nodes for Khanki is selected as best performing model as compare to T-ANN model with one input, two hidden and one output node.
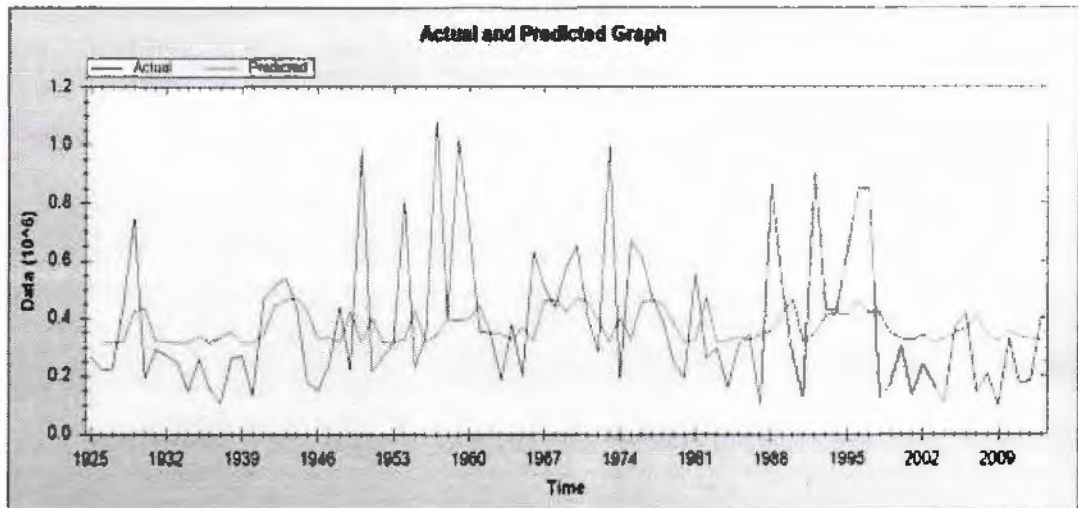
Fig 4.3.1: Actual vs predicted graph of T-ANN model for Khanki site
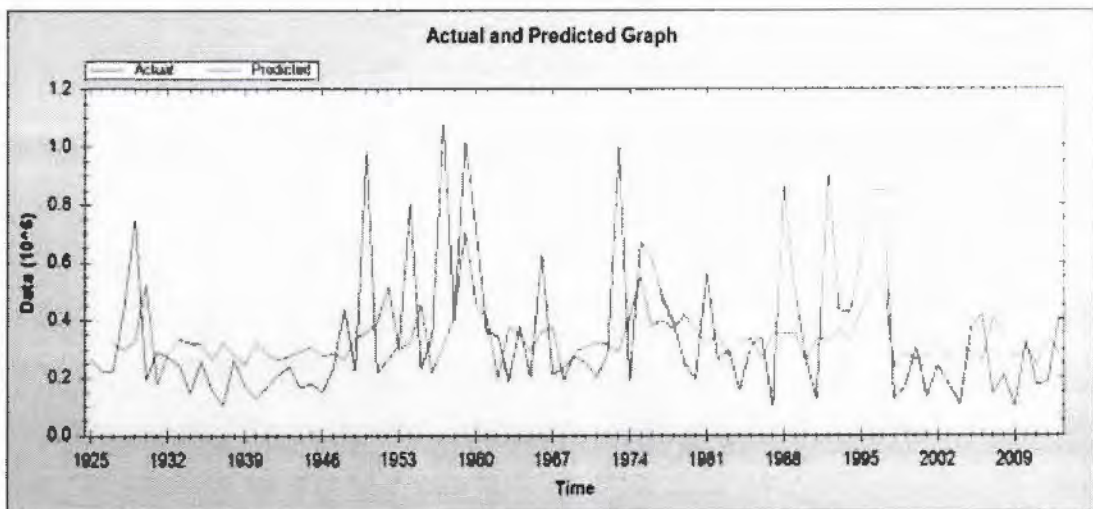

Fig 4.3.2: Actual vs predicted graph of W-ANN (db1) at level for Khanki site

Results further upgraded when approximation and detail sub-signal of the annual maximum peak flow original series are obtained using DWT by db2 wavelet (level 1) as presented in table 4.3.2 below. Compare W-ANN db1 model (table 4.3.1) with W-ANN db2 model (table 4.3.2) for further improvement.

54

To fit W-ANN conjunctional model (db2) the inputs and output used are: The approximation and detail sub-signal of yearly peak measurement values by DWT db2 wavelet level1 given in (Fig.4.1.9 and Fig 4.1.10) are considered two input nodes for W-ANN conjunctional model (db2) with two hidden and one output node. Table 4.3.2 shows that R, RMSE, E, IOA and MAE are obtained from T-ANN and W-ANN conjunctional modle (db2) for Khanki site.

Table: 4.3.2 Training and testing errors for ANN and W-ANN (db2) for Khanki site on Jhelum

Models	Errors					
	Training/Testing	R	RMSE	E	IOA	MAE
W-ANN Sigmoid (2-2-1)	Training	0.9201	1.94E+04	0.3413	0.8961	1.59E+04
	Testing	0.7615	3.59E+04	0.2635	0.7940	3.01E+04
T-ANN Sigmoid (1-2-1)	Training	0.7018	3.31E+04	0.1369	0.6003	3.76E+04
	Testing	0.2163	5.98E+04	0.0348	0.5193	4.07E+04

Table 4.3.2 show the results of W-ANN (db2) model for Khanki site which clearly show that the correlation coefficient (R) is improved from 0.8021 to 0.9201 for training data where it is improved from 0.6903 to 0.7615 for testing data. The value of root mean square errore (RMSE) are reduced from 2.63E+04 to 1.94E+04 for training and from 4.79E+04 to 3.58E+04 for testing data.

The statistical term Nash-Sutcliffe coefficient E are also improved from (0.2991, 0.1653) to (0.3413, 0.2635) for training and testing data correspondingly where the term Index of Agreement IOA are improved from (0.7701, 0.6013) to (0.8961, 0.7940) for training and testing data respectively, while mean absolute error

are reduced from (2.01E+04, 3.71E+04) to (1.59E+04, 3.01E+04) for training and testing data respectively. It is clearly observed from table 4.3.2 that the performance of W-ANN (db2) model is best as compare to the W-ANN db1 model for Khanki station on river Jhelum.
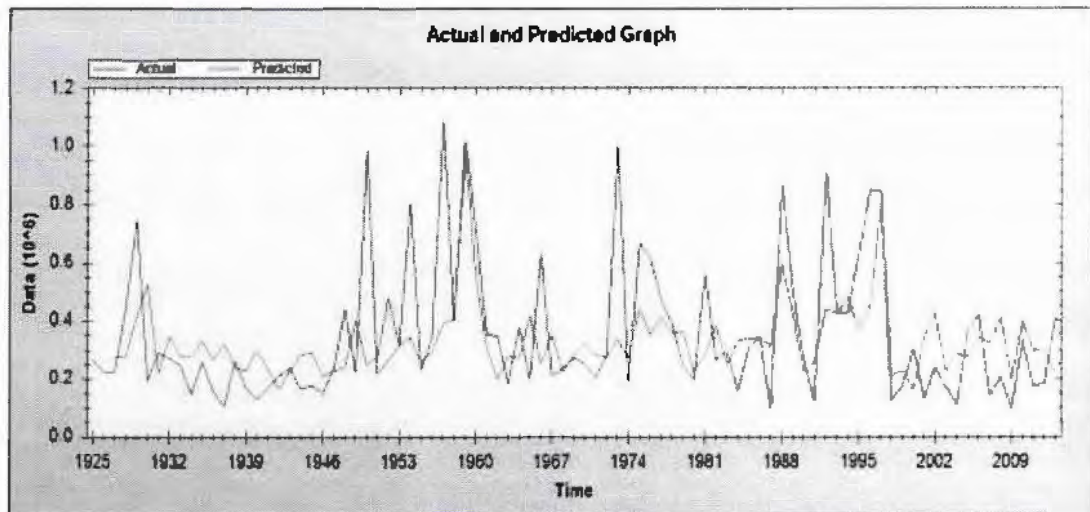

Fig 4.3.3: Actual vs predicted graph of W-ANN (db2) at level1 for Khanki site

Fig.4.3.3 compares the actual and predicted peak flow from W-ANN (db2) model for khanki site. Low flows, Medium flows and high flows were accurately forecasted.

It can be seen that the W-ANN db1 model less forecast the annual peak flow of Khanki stations, while the WA–ANN db2 model provides closer estimates to the corresponding observed peak flows. It is clear from figure 4.3.2 and figure 4.3.3 above that the W-ANN (db2) model forecasts low, medium and high flows for Khanki site is very well for a lead-time of 1 year as compare to W-ANN (db1) model forecasts for low, medium and high flows. The result improved that when the approximation and detail sub-signal of the annual maximum peak flow original series obtained using DWT by wavelet (level 1) as presented in table 4.3.2.

CHAPTER 5

SUMARRY AND CONCLUSION

Regular Artificial Neural Network (T-ANN), Wavelet-ANN conjunctional models (db1) and Wavelet-ANN conjunctional models (db2) have been performed on AM flows of two sites in Pakistan. The AM flows was taken from Federal Flood Commission (FFC) and Water and Power Development Authority (WAPDA) of Pakistan. The data was being measured in cusecs. The length of AM flows has been recorded 54 and 89 years for Mangla and Khanki sites respectively.

In this study the AM flows of 2 sites of Pakistan being considered, are situated on the river Jhelum. In initial stage the qualities of original AM flows consider as approximation and detail sub-signals by wavelet transform (DWT db1 wavelet level 1) and (DWT db2 wavelet level 1) are inputs to Wavelet-ANN (W-ANN) conjunctional model (db1and db2) for Mangla and Khanki sites respectively. Therefore, the original data and wavelet-transform data of 2 sites have been used for further analysis of T-ANN and W-ANN conjunctional models (db1 & db2) respictively. The input, hidden and output neurons were used for T-ANN as (1-2-1) for both stations Mangla and Khanki on river Jhelum in Pakistan, whereas the input, hidden and output nodes were used for W-ANN conjunctional models (db1 and db2) as (2-2-1).

By assigning DWT level 1 (db1) series as input for T-ANN we make W-ANN conjunctional model db1 and observed that by one year ahead forecasting the W-ANN gave more exact result then the T-ANN for both Mangla and Khanki sites respectively. Now the transform of yearly AM flows to approximation and detail sub

signals by DWT level1 (db2) series to the T-ANN, then the W-ANN conjunctional model (db2) are upgraded the results significantly for both sites.

It is understood that the WA–ANN (db1 and db2) models are more exact since wavelet transform (DWT) give valuable decomposition of the original peak flows data, and the wavelet-transform (DWT) data enhances the ability of the T-ANN estimating model by capturing helpful information on different resolution levels.

About the original goals of this study, it is resolved that: (1) the W–ANN (db1 and db2) strategy can be utilized with high accuracy for one year ahead peak flow forecasting for both sites (mangla & Khanki); and (2) the utilization of all wavelet transform decomposed approximation and detail sub-series as inputs to the W-ANN models gives very precise forecasts of peak flows for (Mangla & Khanki) sites on River Jhelum in Pakistan.

The results show that wavelet-ANN conjunctional models are a promising new technique as compare to regular artificial neural network model (T-ANN) for short-term peak flow forecasting in non-perpetual rivers such as, in Pakistan. In general, it can be seen that for one year ahead forecasting for both sites the W-ANN models db1 gave more exact forecasting results than the T-ANN. The results further improved by fitting the W-ANN conjunctional models db2 for both stations (Mangla and Khanki) on river Jhelum in Pakistan.

Recommendations for the Future Study

In this study, there are additionally a few suggestions and recommendations which are given beneath for advancements of this study and future research.

1. For future study the regular artificial neural network (T-ANN), may be compared with Wavelet-ANN conjunctional models (db1 and db2) for other river sites of Pakistan namely Indus, Chenab, Ravi, and Sutlej.

2. It is recommended that the high levels of wavelet deterioration may additionally improve the results.

References

1. Adamowski, Jan F. "Development of a short-term river flood forecasting method for snowmelt driven floods based on wavelet and cross-wavelet analysis." *Journal of Hydrology* 353.3 (2008): 247-266.

2. Adamowski, Jan, and Hiu Fung Chan. "A wavelet neural network conjunction model for groundwater level forecasting." *Journal of Hydrology* 407.1 (2011): 28-40.

3. Adamowski, Jan, and Karen Sun. "Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds." *Journal of Hydrology* 390.1 (2010): 85-91

4. Antar, M. A., Elassiouti, I., & Allam, M. N. (2006). Rainfall-runoff modelling using artificial neural networks technique: a Blue Nile catchment case study. *Hydrological Processes*, *20*(5), 1201-1216.

5. Campolo, M., Andreussi, P. & Soldati, A. 1999 A river flood forecasting with a neural network model. Water Resour. Res. 35 (4), 1191–1197.

6. Campolo, M., Soldati, A., & Andreussi, P. (2003). Artificial neural network approach to flood forecasting in the River Arno. *Hydrological sciences journal*, *48*(3), 381-398.

7. Coulibaly, P., Anctil, F., & Bobee, B. (2000). Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *Journal of Hydrology*, *230*(3), 244-257.

8. Dawson, C. W., & Wilby, R. L. (1999). A comparison of artificial neural networks used for river forecasting. *Hydrology and Earth System Sciences Discussions, 3*(4), 529-540.

9. Dawson, C. W., Abrahart, R. J., Shamseldin, A. Y., & Wilby, R. L. (2006). Flood estimation at ungauged sites using artificial neural networks. *Journal of Hydrology, 319*(1), 391-409.

10. Dawson, C., Abrahart, R., Shamseldin, A., Wilby, R. & See, L. 2005 Flood estimation

11. Guimaraes Santos, Celso Augusto, and Gustavo Barbosa Lima da Silva. "Daily streamflow forecasting using a wavelet transform and artificial neural network hybrid models." *Hydrological Sciences Journal* 59.2 (2014): 312-324

12. He, B., Oki, T., Sun, F., Komori, D., Kanae, S., Wang, Y., ... & Yamazaki, D. (2011). Estimating monthly total nitrogen concentration in streams by using artificial neural network. *Journal of Environmental Management, 92*(1), 172-177.

13. Imrie, C. E., Durucan, S., & Korre, A. (2000). River flow prediction using artificial neural networks: generalisation beyond the calibration range. *Journal of hydrology, 233*(1), 138-153.

14. Islam, A. S. (2010). Improving flood forecasting in Bangladesh using an artificial neural network. *Journal of Hydroinformatics, 12*(3), 351-364.

15. Kamnanithi, N ., Grenney, W . J ., Whitley, D ., & Bovee, K . (1994), Neural networks for river flow prediction, Journal of Computing in Civil Engineering, 8(2), 201-220.

16. Liong, S. Y., Lim, W. & Paudyal, G. N. 2000 River stage forecasting in Bangladesh: neural network approach. J. Comput. Civil Eng. 14 (1), 1–18.

17. Mallat, Stephane G. "A theory for multiresolution signal decomposition: the wavelet representation." *IEEE transactions on pattern analysis and machine intelligence* 11.7 (1989): 674-693.

18. Nayak, P. C., Rao, Y. S., & Sudheer, K. P. (2006). Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water Resources Management, 20*(1), 77-90.

19. PourArian, Mohammad Rasoul, and Ali Hanani. "Blind Steganography in Color Images by Double Wavelet Transform and Improved Arnold Transform." *Indonesian Journal of Electrical Engineering and Computer Science* 3.3 (2016): 586-600.

20. Rajurkar, M. P., Kothyari, U. C., & Chaube, U. C. (2004). Modeling of the daily rainfall-runoff relationship with artificial neural network. *Journal of Hydrology, 285*(1), 96-113.

21. Rumelhart D.E., Hinton E. and Williams J. (1986), Learning internal representation by error propagation, *Parallel Distributed Processing*, 1, 318-362.

22. Scholz, Miklas, and Qinli Yang. "Guidance on variables characterising water bodies including sustainable flood retention basins." *Landscape and urban planning* 98.3 (2010): 190-199.

23. Shamseldin, A. Y. (2010). Artificial neural network model for river flow forecasting in a developing country. *Journal of Hydroinformatics, 12*(1), 22-35.

24. Shrestha, R. R., Theobald, S., & Nestmann, F. (2005). Simulation of flood flow in a river system using artificial neural networks. *Hydrology and Earth System Sciences Discussions*, 9(4),313-321.

25. Singh, R. M. (2012). Wavelet-ANN model for flood events. In *Proceedings of the International Conference on Soft Computing for Problem Solving (SocProS 2011) December 20-22, 2011* (pp. 165-175). Springer India.

26. Thirumalaiah, K., & Deo, M. C. (2000). Hydrological forecasting using neural networks. *Journal of Hydrologic Engineering*, 5(2), 180-189.

27. Tokar, A. S., &Johnson, P. A .(1999). Rainfall-runoff modeling using artificial neural networks. *Journal of Hydrologic Engineering*, 4(3), 232-239.

28. Valipour, M., Banihabib, M. E., & Behbahani, S. M. R. (2013). Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *Journal of hydrology*, 476, 433-441.

29. Varoonchotikul, P., H a l l , M . J . & M i n n s , A . W . (2002b), Flood Forecasting using Jordan Recurrent Artificial Neural Networks, Second International Symposium on Flood Defence, Beijing, Sep. 10-13, 2002

30. Wang, Wensheng, Juliang Jin, and Yueqing Li. "Prediction of inflow at three gorges dam in Yangtze River with wavelet network model." *Water resources management* 23.13 (2009): 2791-2803.

31. Wei *et al.* (2002). Artificial neural network based predictive method for flood disaster Computers & industrial engineering, 42(2), 383-390.

32. Willmott, Cort J. "On the validation of models." *Physical geography* 2.2 (1981): 184-194.

33. Zealand, C. M., Burn, D. H., & Simonovic, S. P. (1999). Short term streamflow forecasting using artificial neural networks. *Journal of hydrology*, *214*(1), 32-48.