EVAPORATION WATER BALANCE IN ARID AREA ANBAR GOVERNORATE - IRAQ

AHMED SAUD MOHAMMED

UNIVERSITI SAINS MALAYSIA

2018

EVAPORATION WATER BALANCE IN ARID AREA ANBAR GOVERNORATE - IRAQ

by

AHMED SAUD MOHAMMED

Thesis submitted in fulfillment of the

requirements for the degree of

Doctor of Philosophy

March 2018

بسم الله الرحمن الرحيم

" يَرْفَعِ اللَّهُ الَّذِينَ آمَنُوا مِنكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ "

سورة المجادلة (11)

In the Name of Allah, the most Beneficent, the most Merciful

" Allah will raise those who have believed among you and those who were given knowledge, by degree. And Allah is acquainted with what you do "Surah Al-Mujaadila (11)

DEDICATION

ToAllah (My Lord) to gain satisfaction and the achievement of pardon and forgiveness.

To the spirit of my father and God rest his soul who was strong support me throughout his life

To beloved my mother for her patience and great support to me

To my dear wife for her unlimited love, support, patience and encouragement;

To my three sons: Mustafa, Hala , and Saad for their understanding, patience, helping and bearing my absence at home;

my brothers and my sister who embrace me with their love,

Ahmed

ACKNOWLEDGEMENTS

First and foremost, I would like to thank and praise Allah, the almighty Merciful, who bestowed upon me the successful accomplishment of this PhDresearch work. Respectfully. I would like to take this opportunity in extending my humble gratitude to my supervisor,**Prof. Dr. Md Azlin Md Said** for his guidance and support for this research – without his dedication, I would not have been able to overcome all the difficulties and further render success to this research. I also extend my gratitude to**Associate Prof.Dr. Rozi Abdullah** for his constant support and kind advice. Also, thank you to**Associate Prof.Dr Ammar Hatem** (field supervisor)and **Prof Dr. Issam Mohammed Abdel-Hamid Al-Hadithy**for their wonderful and timely assistance.

I also would like to express my appreciation to the University of Anbar and to the College of Engineering for giving me the opportunity to complete my study and achieve my dream and my ambition to serve my country and my university.

A personal thank you to Dr. Ahmed Hazem Abdul Karim for his efforts in the research and analysis of soil samples, as well as to the engineer Ammar in the soil laboratory in Engineering college – University of Anbar for his dedication and assistance. Also worthy of mention are Dr Halem Qadarman, Dr Faisal Mahmod, for their help and support. Also great thanks to my uncle Maddah Hassan Al-Alousi for his supprot and constant encouragement in my academic journey. Finally, I would not be here without my mother who always showers me with support, love and prayers.Also, I would like express my sincere love, affection and gratitude to my wife, whom I admire and treasure for her outstanding patience, assistance and support, as well as to my children who have endured my absence for quite a while.

TABLE OF CONTENTS

CHAPTER THREE : STUDY AREA

CHAPTER FOUR: METHODOLOGY

CHAPTER FIFE: RESULTS AND DISCUSSION

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Description
ANN	Artificial Neural Network
AET	Actual EvapoTranspiration
AI	Artificial Intelligent
BCM	Blanny Criddle Model
CA	Cluster Analysis
CE	Coefficient of Efficiency
CN	Curve Number
DEM	Digital Elevation Model
DTM	Digital Terrain Model
DSCAU	Desert Studies Center in Al-Anbar University
ET	Evapotranspiration
ETM+	Enhanced Thematic Mapper +
GIS	Geographical Information System
IMOS	Iraqi Meteorological Organization Seismology
lnSR	Natural logarithm of observe Surface Runoff value
lnRa	Natural logarithm of observe Rainfall value
lnlosses	Natural logarithm of observe water losses value
lnSlope	Natural logarithm of observe catchment slope value
lnRncoff	Natural logarithm of observe Runoff coefficient value
lnlength	Natural logarithm of observe catchment length value
lnArea	Natural logarithm of observe catchment Area value
lnMult	Natural logarithm of multiplying all independent variables of SR except Rainfall.

MAPE	The Mean Average Percentage Error
MLR	Multiple Linear Regression
MWBM	Monthly Water balance Modelling
MSS	Multi Spectral Scanner
MLP	Multilayer Perceptron
NAE	Normalized Absolute Error
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NSE	Nash-Sutcliffe efficiency
NGIA	National Geospatial Intelligence Agency
PET	Potential EvapoTranspiration
PI	Performance Indicator
Ra	Rainfall
RS	Remote Sensing
R^2	Coefficient of determination
r	Correlation Coefficient
RMSE	Root Mean Square Error
Rncoff	Runoff Coefficient
RBF	Radial Basis Function
RH	Relative Humidity
SR	Surface Runoff
SLR	Simple Linear Regression

SVM	Support Vector Machine
SRTM	Shuttle Radar Topography Mission
SS	SunShine
STA	Study Area
STD	Standard Deviation
SNN	Simple Neural Network
SOM	Self-Organizing Map
SLT	Statistical Learning Theory
T	Temperaure
TIN	Triangulated Irregular Networks
ThM	Thornthwaite Model
UH	Unit Hydrograph
UTM	Universal Transverse Mercator
USGS	United States Geological Survey
VIF	Variance inflation factor
WBM	Water Balance Modelling
WSUR	Water Surplus
WDIF	Water Deficit
WS	Wind Speed
WRDAG	Water Resources Directorate in Anbar Governorate

LIST OF SYMBOLS

Symbole	Description
a	Constant
°C	Degree of Celsius
cm	Centimeter
M	Meter
Km	Kilometer
mm	Millimetre
Km2	Square Kilometer
hr	Hour
Sec	Second
M2	Square meter
M3	Cubic Meter
K	Correction Factor
ρ	Percentage of number of sunshine hour
I	Annual heat index
i	Monthly heat index
ΔS	Change in soil moisture storage
ω	Vector of coefficient
R	Runoff
T	Temperature
V	Volume m^3
%	Natural logarithm of multiplying all independent variables of SR

IMBANGAN AIR PENYEJATAN DALAM KAWASAN GERSANG ANBAR GOVERNORATE – IRAQ

ABSTRAK

Anggaran imbangan air untuk lembangan tidak mempunyai tolok dalam persekitaran kawasan gersang adalah satu cabaran utama. Masalah utama ialah tiada persamaan tepat untuk menentukan penyejatan dan air larian permukaan dalam kawasan gersang disebabkan kekurangan data. Regresi lelurus berganda kaedah regresi (MLR) berlangkah dan mundur digunakan untuk membangunkan model penyejatan dan model SR (airlarian permukaan). Keputusan menunjukkan bahawa model penyejatan dengan kaedah regresi lelurus dapat dibuktikan mempunyai kecekapan dan keupayaan untuk meramal penyejatan dan lebih baik daripada model lain yang utama yang digunakan untuk menggangar penyejatan. Keputusan menunjukkan bahawa model penyejatan yang dibangunkan mendapat R^2 (0.923), NAE (0.134) dan, NSE (0.91) lebih baik daripada model Thornthwaite and Blanny criddle dengan keputusanmasing-masing, R^2 (0.884), NAE (0.583) and, NSE (0.278) and R^2 (0.91), NAE (0.324) and NSE (0.611). Faktor-faktor pengaruh penting ialah suhu, kelajuan angin dan jam pancaran suria. Untuk mengenal pasti parameter yang airlarian permukaan dan memilih kelompok penting untuk faktor utama model airlarian permukaan dalam kawasan tadahan, tiga kumpulan pembolehubah tak bersandar telah dipilih dan dimasukkan dalam analysis MLR. Keputusan menunjukkan bahawa model air larian terbaik adalah Kumpulan 2 dengan kaedah regresi mundur mendapat R^2 (0.744), NAE (0.146) dan NSE (0.722) dimana faktor pengaruh penting lain ialah hujan, cerun kawasan tadahan, luas kawasan tadahan dan pekali airlarian. Untuk meningkatkan ketepatan model anggaranairlarian, tiga kumpulan sama yang digunakan dalam model airlarian MLR telah digunakan dalam analysis dua jenis model ANNs dan teknik AI (SVM). Keputusan menunjukkan bahawa MLP

mengatasi kaedah RBF dan SVM dalam meningkatkan proses ramalan, dimana model airlarian MLP Kumpulan 2 mempunyai keputusan terbaik jika dibandingkan dengan model yang lain. Keputusan model airlarian MLP kumpulan 2; dalam fasa latihan (R^2= 0.846, RMSE = 0.160, NAE = 1.251, NSE = 0.846); manakala dalam fasa ujian (R^2= 0.788, RMSE = 0.182, NAE = 0.628, NSE = 0.775). Persamaan regressi penyejatan telah dintegrasikan dengan perisian GIS (ArcGis 10) telah digunakan untuk menyediakan peta taburan beruang penyejatan bulanan dan bermusim, dan peta lebihan air di dalam kawasan kajian. Model regresi airlarian telah digunakan untuk menentukan airlarian permukaan di dalam tadahan kawasan kajian menggunakan data yang ditransposisi dari kawasan tadahan berdekatan. Proses transposisi data telah dijalankan untuk menggangarkan isipadu airlarian di dalam tadahan kawasan kajian (yang tidak mempunyai tolok cerapan). Isipadu airlarian berada diantara 1,321,732 m^3dan 2,488,979 m^3. Peta taburan ruangan isipadu airlarian telah dijalankan menggunakan persekitaran GIS didalam seluruh kawasan kajian.

EVAPORATION WATER BALANCE IN ARID AREA ANBAR GOVERNORATE - IRAQ

ABSTRACT

Estimation of water balance for ungauged basin in arid area environment is a major challenge.The main problem is there are no precise equations to estimate evaporation and surface runoff in arid area due to lack of data in these regions. Multiple linear regression (MLR) stepwise and backward regression methods were used to develop evaporation and surface runoff models. The results showed that the evaporation developed model in linear regression method has proven its efficiency and its ability to predict evaporation and the superiority against most important models that used for estimating the evaporation. The results for evaporation developed model were R^2 (0.923), NAE (0.134) and NSE (0.91) better than Thornthwaite and Blanny Criddle models with results of R^2 (0.884), NAE (0.583) and, NSE (0.278) and R^2 (0.91), NAE (0.324) and NSE (0.611) respectivily. The significant influence factors are temperaure, wind speed and sunshine. To identify the parameters for surface runoff and to select the significant groups for main factors of runoff prediction model in catchments, three groups of independent variables have been established for MLR analysis.The results showed that the best surface runoff model for Group 2 backward regression method with R^2 (0.744) and NAE (0.146) and NSE (0.722) where the significant influence factors were rainfall, catchment slope, catchment area and runoff coefficient. To improve the accuracy of runoff prediction model, similiar three groups of MLR surface runoff model were analysis for two ANNs models and AI techniques(SVM).The results indicate that MLP showed better results compare to RBF and SVM methods for the predictive process, where the surface runoff MLP Group 2 produced the best results compared to other models. The results of surface runoff MLP Group 2 were in Training Phase (R^2= 0.846, RMSE 0.160, NAE =

1.251, NSE = 0.846) while in Testing Phase (R^2= 0.788, RMSE 0.182, NAE = 0.628, NSE = 0.775). Regression equation for evaporation model was integrated in GIS software (ArcGIS 10) to map the spatial distribution for monthly and seasonal evaporation, water surplu for whole catchment study area. Runoff regression equation was used to estimate the sub-catchments runoff in the study area using transposition approach. Transposition of surface runoff data process was carried out to estimate runoff volume in sub-catchment study area (ungauged area). The runoff volume ranged between 1,321,732 $m^3$to 2,488,979 m^3.Spatial distribution for runoff volume were carried out using GIS environment on the entire study area.

CHAPTER ONE

INTRODUCTION

1.1 Outline brief

Water resources serve to be one of the most crucial criteria for societies building and their development. Aspects like evaluation, planning, and management of water resources have stood out to be one significant in the humans life, particularly in arid regionscharacterizing extremely limited amount of rainfalland uneven spatial distribution.The rainfall rates are one of the most substantial natural resources in arid environments and it is weighted upon as a distinctive source of surface runoff and groundwater recharge. Moreover, it was reported that water is insufficient especially in arid areas, therefore attention should be given to water studies in these areas (Al-Maliki and Dairi, 2005; Hekmat, 2010).

By definition arid and semi-arid regions are those areas where water is at its most scarce, the hydrological regime in these areas is naturally extreme and highly variable, these areas have been subjected to the high pressures in freshwater resources due to the growth of the population, increasing water use for domestic, agriculture and climate change (Wheater *et al.*, 2008).

The water crisis in Iraq is affected by the policies of neighboring countries,such as Turkey and Syria.Turkey is building 14 dams on the Euphrates river and its tributaries, 8 dams on the Tigris river and its tributaries to control the river water.The Euphrates river which runs through Syria before it goes into Iraq.TheSyrian Govermentestablished 5 dams on the Euphrates river.

The study area constitutes one of the important areas in the western region of Iraq, since it has main and secondary dry valleys that receive large amounts of rainwater when it rains during the rainy season (Kamel *et al*., 2011). The development of natural resources and land uses and farming can be manipulated for agriculture (Al-Alusy, 2011). Therefore, hydrological studies of the dry valleys in the study area are of special significance as they are connected to the development of water resources, agricultural and pastoral areas (Mohammed, 2012).

In the rainfall months, these valleys are reported to collect large amounts of rainfall that are unexploited and not managed properly.In some occasions, rainstorms lead to flooding, as proven in Haditha city when Haditha meteorological station recorded 118.6.mm on 11/16/1994 (Al-Alusy, 2011) .

One fundamental components of the hydrologic cycle is evaporation and it plays a very important role in hydrological studies, water resource management andirrigation system design (Sammen, 2013). Evaporation is considered to be a one major requirement in the planning and designing of any irrigation project in arid and semiarid regions. The average annual evaporation in the western region of Iraq, is more than 3000 mm(Kamel *et al*., 2011).A precise estimation of evaporation would lower the risk of wasting great quantities of water (Abdullah *et al*., 2014).

The runoff estimation process is an important step in a hydrological studies especially in arid regions. The water surplus is the term given to the excess of rainfall over the potential evapotranspiration values during specific months of the year. The annual potential evaporation in the western region of Iraq are approxmitely ranging btween 1150-2000 mm (Al-Fatlawi and Jawad, 2011; Al-Maliki and Dairi, 2005). Rainfall–runoff relationships in arid regions have an important role to play in understanding the dynamic aspects of the hydrological processes taking place in arid regions (Pilgrim *et al*., 1988;Ye *et al*., 1997; McIntyre *et al*., 2007; Wheater *et al*., 2008; McIntyre and Al-Qurashi, 2009).

Of late, various artificial intelligence (AI) techniques have been adopted to fulfill hydrological modelling purposes. These techniques have notably demonstrated a satisfactory performance especially when the available data are not enough to apply numerical and physical models (El-Shafie *et al*., 2011; Rajurkar *et al*., 2004).

Water balance studies have been considered to be significant studies that address many kinds of problems because they highlight the spatial relationship between the incident rainfall in a given area and amount of rainfall that return to the atmosphere caused by evaporation and evapotranspiration in order to assess the amount of water surplus and water deficit in this area (Al-Maliki and Dairi, 2005).

The use of the remote sensing (RS) techniques and geographical information system (GIS) shows high capabilities in dealing clarifying and showingclimate and hydrological data in thematic maps to assist the decision-makers when it comes to making sound decisions.

1.2 Problem statement

Water shortage and limited water resources especially in arid region can severely affect the sustainable development of this region. The water shortage in arid regions are caused by low rainfall and uneven distribution throughout the wet seasons such as Iraqi western desert that has an average annual rainfall of about 115 mm, which puts a risk to the agriculture (Ministry of Agriculture, 1977).

Water balance is a good method to approach the problems associated with the region.Water resource assessment studies rely heavily on the water balance, depending on the wide range of meteorological and hydrological variables. A simple error in calculating these variables will give incorrect results especially in the dry areas due to the remarkable increase of temperatures and low rainfall in water balance estimation. Evaporation plays a vital role particularly for thesmall dry valleys in the arid regions. Evaporation is a climate component that has a significant role in determining the amount of runoff, water balance and water surplus distribution of the arid catchment area. Evaporation and evapotranspiration processes leave a negative effect which can lead to lost water resources. Therefore, stepsare taken to measure and estimate the evaporation in the arid region which are important toidentify the surplus water for optimal management of water that leads toan increase agriculture production.

In arid regions, the evaporation from bare soil is shown to havea greater importance in relation to the transpiration from plants due to the larger area of bare soil, lack of vegetation and the frequency of small rainfall events allowing the bare soil to return the water to the atmosphere (Pilgrim *et al*., 1988). Consequently the estimation of the evaporation in dry areas is more significant than that of evapotranspiration. Because of this, the initiative to evaluate evaporation developed model is a significant step in arid regions using Multiple Linear Regression (MLR)based on the climatic data from meteorological stations around study area, which further facilitates the estimation of the water balance.

Estimating the amount of surface runoff as the primary factor in the subject of developing any region and identifying factors that influence and contribute to gaining runoff is a very important topic but not studied deeply for arid areas. One of the essential factors is toidentify the spatial distribution of surface runoff and its effectiveto the study area. Predicting surface runoff in ungauged catchment areas is vital to practical applications like runoff forecasting and for water resources management.

The runoff prediction is one of the most useful hydrological systems, it may be used to predict aspects of flooding, to run an assessment on the water harvest projects to develop the arid areas, to help run the reservoir operation, or to be used in the prediction of the water born contamination transport (Jain, 1996; El-Shafie *et al.*, 2011).Different methods have been used in runoff prediction that involve conceptual and empirical models. Nevertheless, it is cannot consider any of them as a single superior model. As the hydrological process is very complex, the accurate runoff is hard to be predicted using the linear recurrence relations or physical-based watershed.

The use of the multiple linear regression (MLR) and Artificial Neural Network (ANN) technique for rainfall-runoff modeling has given the system theoretical modeling approach a new dimension and as a successful tool it has been applied in recent years, to settle a lot of problems concerned with hydrology and water resources engineering (ASCE, 2000). In spite of the wide strides and the increasing trends during the recent years with regard to the utilization of AI techniques for SR prediction modeling, there remain some areas that require further investigation like the hydrological and geomorphological variables in predicting runoff and the use of diverse AI technique to get the best prediction. Therefore, ANNs and AI techniques represent the recent trends in the field of hydrological modelling for improving the modelling prediction ability without the need to have extra data and effort.

1.3 Research objective

The research objectives are:

1. To determine the evaporation model using multiple linear regression (MLR).
2. To determine the surface runoff model using multiple linear regression (MLR).
3. To assess the appropriate surface runoff prediction using ANN and AI techniques for arid region, and compare the results.
4. To estimate the spatial inputs of evaporation and surface runoff for water balance modeling.
5. To evaluate spatial distribution of volume surface runoff in catchment area.

1.4 Research Scope

Monthly and daily climate data for these years (1975-2012) and (2006-2012) respectively are provided from the Iraqi Meteorological Organization Seismology (IMOS) Baghdad in February 2013 to be adopted in the evaporation developed model and water balance calculation.Rainfall runoff actual data were provided from theIraqi Ministry water resources through the national program for the development of the Euphrates river basin from (1998 – 2002).The runoff generation are influenced several factors like catchments losses, catchments slope, annual runoff coefficient percent, drainage length and catchment area. Therefore, the data for these factors were added to the rainfall runoff actual data to develop the runoff prediction model with help of the MLR.

1.5 Thesis Organization

The thesis consists of 6 chapters whereChapter One gives a brief description of the water scarcity and problem statement, the research objective and research scope. Chapter Two supplies a relevant background of literature review on the water scarcity, hydrological process, evaporation, surface runoff, MLR, AI techniques (ANNs and SVM) and water balance modeling calculations. Chapter Three describes the study area, hydrological details of the study area and database operation. Chapter Four represent the research methodology, which will include data collection and analysis, hydrological data, meteorological data, runoff data, soil data. In Chapter Five the results and discussion are discussed the evaporation model using MLR, potential developed surface runoff model using MLR and ANNs and AI techniques, estimation of the water balance calculation using three models with help of GIS and remote sensing. The conclusion of the thesis and recommendation for the future work are given in chapter Six.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

Arid and semi-arid regions are areas with scarcity of water and data considered as a major obstacle for any development or exploitation for these regions. Thus, optimum and efficient use of scarce fresh water resources has become a major concern these days. As the population increases in a dramatic way, it led to a significant pressure on the vital resource. The process of transformation of rainfall into runoff over a catchment is very intricate highly non-linear, and it shows of both temporal and spatial variability. Therefore, this chapter sheds light on a general review about the most significant hydrological processes that occur in arid regions, evaporation and surface runoff regression model, artificial intelligence (AI) background theory andwater balance modeling calculation with help of remote sensing and GIS.

2.2 Hydrological Processes of Arid Area.

The hydrological behaviors in arid regions are not the same as the humid areas (Sen, 2008; Simmers, 2003). The ephemeral valleys streams in arid regions are characterized by poor amount of rainfall data and surface runoff, soil properties and initial conditions (Nouh, 2006) and the scarcity of accurate observations (Pilgrim *et al*., 1988).Although it is undeniable that water is essential in arid regions, hydrological data have historically been very scarce. It has been widely stated that the major limitation of the development of arid zone hydrology is the fact that high quality observations are insufficient (McMahon, 1979; Nemec and Rodier, 1979;

Pilgrim *et al*., 1988). There are many good reasons for this as populations are usually sparse and economic resources limited, the climate is harsh and hydrological events infrequent. However, without the reliable long-term data and experimental research, there has been an inclination to depend on humid zone experience and modelling tools, and data from other regionsWheater, (2010). At best, such results will be very inaccurate. At worst, there is a real hazard of adopting inappropriate management solutions which ignore the specific features of dryland response (Wheater*et al*., 2008). Seeking to ensure the availability of hydrological data in arid areas it is insignificant step to develop methodologies for hydrological studies and hence to define priorities for both the research and hydrological data.

2.2.1 Rainfall

Rainfall is the most important climate elements affecting both the arid and semi-arid regions, despite the fluctuation and scarce rainfall, sometimes suddenly it fall in a high quantity leading to flood occurrence on the earth surface (Al-Maliki and Dairi, 2005). It is considered to be a key variable in the hydrological cycle regardless of the climate region. In arid regions, rainfall events generally have a short duration and high intensity and they are often characterized by a large degree of spatial heterogeneity, (Pilgrim *et al*., 1988; Wheater *et al*., 1991; Martinez-Mena *et al*., 1998; Lazaro *et al*., 2001; Wheater *et al*., 2008; Andersen, 2008).

Although the rainfall shows low density, the spatial estimates of rainfall are usually calculated from point measurements using well established spatial interpolation techniques such as Thiessen polygons and kriging interpolation (Andersen, 2008) and also the nearest neighbourhood, spline and inverse distance weighting (Meher, 2014).

Regardless of how dense the stations network of rainfall measure in arid region, there will always be areas without a rain station. Therefore, it is important that the values of rainfall in these locations are estimated using the values at surrounding sites (Goovaerts, 1999). There are many methods that have been used in the past for the spatial interpolation of rainfall such as inverse distance weighted method, splines and kriging methodsd. Spatial interpolation process are techniques used to predict the value at a given location based on rainfall values for adjacent areas. For each estimation, the measured values points are weighted depending on where they are located.

There is a major difference among these methods which is the criterion used to estimate the sample point values. Criteria may comprise simple distance relationships (e.g., inverse distance weigh methods), variance minimization criteria (e.g., kriging and co-kriging), curvature minimization criteria of and strengthening of smoothness criteria (splining) (Hartkamp *et al*., 1999). Abushandi(2011) reported that there is maximum and minimum extreme yearly rainfall variability in Wadi Dhuliel arid region in Jordan whereone rain gauge measured the annual rainfall to be 275.7, 93.1, 111.1, 230.4, 194.8, 63.1, and 209.5 mm over seven years. On one single day, 62 mm of rainfall occurred, even though the total annual rainfall in the same year was 100 mm (Sukhnah rain gauge). Thus, these kinds of rainfall events can produce a significant surface runoff, causing a serious soil erosion. Weather behavior and topographical characteristics certainly play important roles here.

2.2.2 Evapotranspiration

Evapotranspiration also manifests itself as the main requirement in planning and designing any irrigation project in arid and semiarid regions. A precise estimation of evapotranspiration would reduce the waste of great quantities of water (Abdullah *et al*., 2014). The water seeps from the Earth's surface to the atmosphere and this is called evapotranspiration (ET) which is the most important water balance computation in a arid and semi-arid climates (Deus *et al*., 2013; Jinxia *et al*., 2012; Wilcox*et al*., 2003). Potential evapotranspiration (PET) that serves to be the average evaporation-transpiration of surface soil planted by growing green plants does not suffer from the lack of water. Approximately, the annual potential evaporation in the Iraq arid areas is approximately 1150 mm (Al-Fatlawi and Jawad, 2011; Al-Maliki and Dairi, 2005).

Vegetation patterns carry a vital role in gauging about evapotranspiration, infiltration and surface runoff (Templeton *et al*., 2014; Hugo *et al*., 2006; Mueller *et al*., 2007). In addition, ascertaining the rainfall losses via evapotranspiration (ET) is a vital step because of the direct relationship between agricultural crop production and the improved water resources investment (Al-Rijabo *et al*., 2008). Many researchers have considered the various methods to estimate PET compared to the actual monthly ET with several empirical equations,which include Thornthwaite 1948, FAO-56 Penman Monteith, Modified Blaney-Criddle, Hargreaves and Samani(Al-Rijabo *et al*., 2008; Ali, 2008; Mohammed, 2008). Evaporation is controlled by several meteorological conditions factors which include the air temperature, air humidity, air pressure, wind speed and precipitation.

Mughrabi(2011) developed an evaporation model of high aridity region using evaporation data for one year. This study has sought to find a correlation between evaporation with climate variables and found that using statistical modeling to estimate the rate of evaporation in dry and acceptable approach to results of field measurements.

Ali (2008) mentioned that there is increasing in evaporation rates in the Najef governorate in Iraq arid area, especially during the hot dry summer months which can reach 548.8, 607.7 and 546.9mm in the months June, July and August respectively. This represents about 47% from total annual evaporation rate 3655.47 mm due to high temperature that recorded in the these months above, while evaporation rates are lower during the months December, January and February, with about 88.02, 82.7 and 117.1 mm respectively.

2.2.3 Surface Runoff

Surface runoff is the rainfall part that reaches the main valley stream when rainfall rate increases on rainfall losses (leakage, storage of depressions, evaporation). This rainfall is also known as effective rainfall (Kamel *et al.*, 2011). Moreover, the additional rainfall over the potential evapotranspiration values during specific months of the year is termed as water surplus (Al-Fatlawi and Jawad, 2011) and generates the surface runoff and groundwater recharge. Rainfall–runoff relationships in arid regions as it plays a great role in understanding the dynamic aspects of the hydrological processes that take place in arid regions, and this is evident from several reviews (Pilgrim *et al*., 1988; Ye *et al*., 1997; McIntyre *et al*., 2007; McIntyre and Al-Qurashi, 2009).

Generally, surface runoff is categorized into two groups following the mechanisms accountable for its generation. The Hortonian-type of runoff or infiltration excess overland flow refers to the situation, where high rainfall events surpass the infiltration capacity of the soil and this lead to the surface runoff. This mechanism is broadly considered to be the dominate runoff generation in region of arid and semi-arid conditions (Pilgrim *et al*., 1988; Hughes, 1995; Andersen, 2008). The other type of overland flow is referred to as saturation excess overland flow and occurs when the rainfall falls on land where the saturation of subsurface would occur, this situation happens when it rains in the bottom of the valley (Alkhoury, 2011).The arid catchment is represented by ephemeral wadis, where a stream runs fully for a short period of time, in the series of heavy rainfall events. To add, these events fill desert dams and thus this leads to the recharge of aquifers (Abushandi, 2011).

Huassain and Ahmed(2008) mentioned that the flood or runoff in the dry area of Iraq depends on the duration and intensity of the rainfall and the volume of runoff at the Wadi al-Ghadaf estuary did not surpass 1.5 million m^3 for 1975-1976, while 3.5 million m^3 in the Nukhayb area for the same year.

Masoud (2015)tried to estimate the relationship between rainfall and runoff and also to provide flash flood hazard warnings for arid area ungauged basins wadi Rabigh - Saudi Arabia based on the hydrological characteristics with the help of the geographic information system (GIS). Wadi Rabigh was suffering from scarce real rainfall runoff data. In this study two storms of 15 and 22 mmof rainfall that generated an amount of surface runoff $15.2\times106\ m^3$ and $7.7\times106\ m^3$ respectively.

2.3 Rainfall Runoff Modelling

The relationship between rainfall and runoff is the nucleus or fulcrum surface water hydrology as the runoff makes the final outcome of rainfall (Al-Sheblaq and Al-Najar, 1995). Accurate predictions of floods are a daunting task and different techniques had been employed with a number of improvements to getaccurate flood estimation (Ghumman *et al*., 2011; Bahremand and De Smedt,2008; Bahremand and De Smedt, 2010; Bhadra *et al*., 2010). It is quite difficult to perform the rainfall runoff analysis due to of the existence of the complex nonlinear relationships in the transformation of rainfall into runoff. That said, the runoff analyses are very important to predict natural calamities. It plays a vital role in the design and operation of multiple components of water resources projects like hydraulic structures (barrages and dams), water supply schemes, etc. Runoff analyses are also required in water resources planning, development and flood mitigations (Ghumman *et al*., 2011). Various types of modelling tools had served to gauge the runoff like lumped conceptual models, distributed physically based models, stochastic models and black box (time series) models (Bahremand and De Smedt, 2008; Bahremand and De Smedt, 2010; Tingsanchali and Gautam, 2000). Conceptual and physically based models attempt to account for all the physical processes involved in the rainfall - runoff process, but their successful use is not widely applicable mainly because of the need for the catchment specific parameters and simplifications involved in the governing equations (Tingsanchali and Gautam, 2000). Moreover, due to the non-stationary behavior and nonlinearity in the data the used of time series stochastic models boasts off higher complexity (Ghumman *et al.*, 2011). These models often necessitate a lot of experience and expertise of the modeler (Güldal and Tongal, 2010).

The unit hydrograph is a simple event model that has limited performance capability (Chow *et al.*, 1988). However , methods of time-series analysis can be used to identify more complex model structures for event or continuous simulation. Synthetic unit hydrographs can readily be generated based on default model parameters, and this is very helpful when the data is scarce (Chow *et al.*, 1988); Duggal and Soni, 2005). However, relatively little work has been carried out to assess the associated uncertainty with these estimates (Wheater *et al.*, 2008). Hydrological models are more efficient and accurate by the day, and they are now used as powerful tools by decision makers concerned with the more pressing demand on water resources. In arid regions, these models are crucial. The problem encountered with hydrological modelling in arid regions is that great data required to support a hydrological model is limited (Wheater *et al*., 2008). Another problem in arid regions is that hydrological responses are far from consistent, but they vary spatially within the catchment of interest and occur over irregular time intervals (Alkhoury, 2011). The relationship established between rainfall and runoff is determined by several factors, like catchment slope, evapotranspiration,land cover and soil texture. Moreover, the urbanization and agricultural activity affect the runoff quality and quantity (Abushandi, 2011).

2.4 Statistical Regression Method

Linear regression seeks to model the relationship between two variables by fitting a linear equation to observed data. One variable is considered to be an explanatory variable, and the other a dependent variable. The simplest models that can predict evaporation or surface runoff from its independent variables are statistical regression methods. These whole empirical techniques are quick and easy to apply since they do not require complexparameter input (Meher, 2014). In statistical modelling, the regression analysis stand out as a statistical process for estimating the relationships among variables. It makes use of many techniques for modelling and analysing several variables, when the focus rests on the relationship between a dependent variable and one or more independent variables. A large number of models in hydrology and climate sciences have to depend on the multiple linear regression to justify the link between key variables. The relationship in the physical world may experience abrupt changes following the climatic, environmental, or anthropogenic perturbations(Seidou, *et al*., 2007).

A particular advantage of using multiple linear regression model is limited hydrological data required without considering the physics of hydrological process of the catchment (Patel*et al*., 2016).Statistical regression or regression analysis is used widely for prediction, forecasting and the understanding of which among the independent variables are connected to the dependent variable. Statistical regression approach is weighted upon as a quick and easy application empirical techniques and it dose not need any complex factors(Patel *et al*., 2016).

2.4.1 Evaporation Regression Model

The simplest models namely the multiple linear regression (MLR) are available to predict evaporation from its independent variables. The models predict evaporation as a function of temperature, wind speed, relative humidity and sunshine. Evaporation data are not always readily available for a certain climatic region. The prediction models for evaporation are often used. An accurate estimation of evaporation is not easy to administer because of the complex interaction between the components of the land-plan-atmosphere system (Ibrahim *et al.*, 2012).

Hanson (1989) investigated the daily evaporation on three sites on the watershed in southwest Idaho India. The study pointed to daily pan evaporation estimated by mean temperature and solar radiation, the correlation coefficients (r) obtained between 0.84 to 0.90. The study summed up that the daily evaporation varied between 7.5 mm/day at the mid elevation site (1,649 m) and 6.5 mm/day at the highest elevation site (2,097 m). Total summer evaporation was 1,255, 1,082, and 795 mm for the low, mid, and high elevation sites, respectively.

Singh (1995) examined the relationship between evaporation measured from US Class A and diverse meteorological parameters.The correlation coefficient (r) was (0.85) between evaporation and max temperature, and 0.82 with wind speed. The coefficient of determinates (R^2) for the relationships for minimum air temperature was 0.7, relative humidity was - 0.56 and bright sunshine hours was 0.15, the correlation results of this study are low between the evaporation and these parameters according to another study.

Khanikar and Nath (1998) studied the relationship between evaporation and meteorological parameters from an open pan evaporimeter. The coefficient of determination (R^2) for minimum temperature was 0.65, maximum temperature 0.64, wind speed was 0.53. The coefficient of determination (R^2) was 0.91 and the conclusion that can be made is that the stepwise regression method serves as the best model to estimate the evaporation at Jorhat. Shrivastava *et al.*, (2000) assessed the relationship between pan evaporation and other climatic factors for Sundrebans, west Bengal. The data that was used an weekly time period basis for 25 years (1963-1987). The linear regression equations, quadratic and cubic have been used for individual parameters and the entire parameters. There is a linear relationship between evaporation and minimum temperature, wind speed, maximum relative humidity, the coefficient of determination (R^2) was 0.92.

Jhajharia (2006) evaluated the correlation relationship between pan evaporation and meteorological parameters in Jorhat. The data used were based on the monthly period during (1970-1998). The statistical regression approach served to correlate the pan evaporation with climatic meteorological parameters, where the stepwise regression approach was applied. The study found that there was a positive significant relationship between evaporation and temperature, wind speed and sunshine duration and non-significant relationship with the relative humidity. Furthermore, it was noted that evaporation was mainly influenced by the combined effect of the maximum air temperature and wind speed at Jorhat. Almedeij (2012) investigated the evaporation in Kuwait state and reported that the correlation of evaporation with temperature was 0.94, RH was -0.92 and wind speed was 0.74. Shirgure (2012) mentioned that multiple correlation coefficients for temperature and wind speed were +0.94 and +0.92, respectively. Balogun (1974) mentioned that from all the climatic factors, wind speed had the weakest correlation with evaporation.Almedeij (2016) proposed an evaporation model in Kuwait - arid region wherethe data used a monthly time period for 23 years (1993-2015). The study showed that that evaporation values, ranged between 0.1 to 40 mm/day, from January – July within this period.

2.4.2 Runoff regression model

Schär*et al*(2004) examined whether the shortage of rainfall observations data can be dealt with by employing the statistical regression approach to establish a rainfall model based on the meteorological data. The data used a monthly time period basis for 15 years (1979 -1993) provided from the European Centre for Medium-Range Weather Forecasts (ECMWF) and from rain gauge networks, to be compared with the observed runoff data for two basins namely Amudarya and Syrdarya in the Aral Sea basin in central Asia. The relationship between two time series data was examined and the correlation coefficient (r) of 0.92.

Zhao *et al.* (2012) examined the relationships between observed monthly runoff data and meteorological data obtained for period (1957-2000). The study performed a multiple linear regression which demonstrated a good performance indicator as the coefficient of determination (R^2) and the Nash-Suttcliffe coefficient (NS) of 0.84 and 0.82 respectively.

Patel *et al.*(2016) looked into an approach that combines rainfall - runoff data for the generation of multiple linear regression rainfall - runoff models for stream flow estimation. Different numbers of input data as rainfall with the previous record of runoff values can help establish a mathematical relationship for estimating the stream flow using the stepwise regression method. The stepwise model is performing well with available data, hence it is selected for the five different scenarios with various range of the data calibration and validation. Furthermore, RMSE and SE were used to study these scenarios. The best stream flow prediction can be drawn by using all independent parameters because of the fact that it gives the highest correlation value among all ten models which is 0.6657.

2.5 Artificial Intelligence (AI) Techniques

The interrelationships between the runoff and the factors affecting the formation is not clear andnon-linear (Dorofki*et al*., 2012).In addition, in most dry areas the data is low and sometimes inaccurate for the formation of runoff, this prompted the researchers to look for other ways to have the ability to deal with this kind of problems including artificial intelligence techniques (Alagha, 2013). The ability of AI in handling and processing complex problems results from the ability to mimic the behavior of the human brain behavior (Rajanayaka *et al*., 2002; Iliadis and Maris, 2007; Chen *et al*., 2008).

To date, several artificial intelligence techniques like the artificial neural networks (ANNs) and support vector machine (SVM)that serve to determine the relationship with rainfall runoff have proven its efficiency as a means of new modeling rather than the traditional models used in this area (Chen *et al*., 2008). AI techniques shows a better performance than other techniques when the phenomenon under study is complex, and with a lack of available data, which is the general situation in most of the rainfall - runoff studies in the dry regions. AI was used extensively and successfully in many complex hydrological studies and with reliable performance and with relatively less effort ,cost and required data, as well, but it is found to be superior when compared with the process based models.

2.5.1 Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANNs) mechanisms simulate the human brain behaves to further develop a form of artificial intelligence system. For the past few years, ANNs have been used in a wide range of applications, including the prediction of problems surrounding water resources (Alkasseh, 2013). In spite of the fact that the concept of artificial neurons was introduced in 1943, these applications have been developed since the back propagation training algorithm was introduced for feed forward ANNs in 1986. Thus, ANNs may be regarded as a fairly new tool in predicting and forecasting (Maier and Dandy, 2000; Palani *et al*., 2008; Gaur *et al*., 2013).

Based on the human brain a neural network can be defined as a model of reasoning. The brain has a densely interconnected set of nerve cells, or basic information-processing units, called neurons. The human brain uses nearly 10 billion neurons and 60 trillion connections, synapses, between them (Shepherd and Koch, 1990). Through a simultaneous use of multiple neurons, the brain can perform its functions much faster than the fastest computers in existence today (Negnevitsky, 2005)

In ANNs, the available data seek to develop empirical relationships (cause–effect or input–output) that express the physical process, and this relationships are applied to new input data to enable the estimation of the output (Sahoo *et al*., 2006; Alagha, 2013). The attractiveness of ANNs marked from their ability to solve highly complex problems characterized by high non-linear relationships among variables also, the ANNs technique is an ideal tool when a large and complex monitoring datasets are present. The power of ANN in addressing complex problems and its powerful tools for computing system for highly complex and nonlinear systems results come from its talent of emulating the behavior of human brain (Rajanayaka*et al*., 2002; Iliadis and Maris, 2007). ANN belongs to the black box time series models and it offers a relatively flexible and quick means of modelling.

The ANN models have been adopted increasingly in multiple aspects of science and engineering because of their ability to model both linear and nonlinear systems without the need to make any implicit assumptions in a lot traditional statistical approaches (Rajurkar *et al*., 2004). The successful use of ANNs is evident when referring to the river flow prediction, for rainfall-runoff process, for the prediction of water quality parameters and for characterization of soil pollution. ANNs also work well when it comes to predicting evaporation, forecasting rainfall-runoff, predicting flood disaster, and predicting river flow time series(Riad *et al*., 2004).

It has many unique advantages and possesses the capability of representing the arbitrary complex non-linear relationship between system's input and output. ANN can be treated as a universal approximator with an ability to learn from examples without the need to have any explicit physics (Govindaraju, 2000; Rajurkar *et al*., 2004). Previous works by Karunanithi *et al*., (1994); Dawson and Wilby, (1998); Campolo *et al*., (1999); Zealand *et al*., (1999); Imrie *et al*.,(2000) have demonstrated the ANN's capability in stream flow forecasting.Different types of methods are adopted in the runoff prediction involving conceptual and empirical models. Nevertheless, none of these methods can be seen as a single superior model. Because of the complex hydrological process, the prediction of the accurate runoff not easy to do using the linear recurrence relations or physically based watershed. Artificial Neural Network (ANN) models have been used well to model complex non-linear input-output relationships in an extremely interdisciplinary field. The natural behavior of hydrological processes is deemed suitable for the application of ANN method (El-Shafie *et al*., 2011).Generally, forecasting models can be divided into statistical and physical based approaches. Statistical approaches identify the relationships between historical data sets, whereas physical based approaches model the underlying processes in a direct way. MLP networks are closely linked with the statistical models and are the type of ANN that works well with a number of forecasting applications (Rumelhart *et al*., 1988). When using ANNs for forecasting, the modelling philosophy is very much the same with the one used in traditional statistical approaches. In both cases, the unknown model parameters (i.e. the connection weights in the case of ANNs) are adjusted to get the best match between the historical set of model inputs and the corresponding outputs (El-Shafie *et al*., 2011).

ANN techniques show a better performance compared with other techniques when the process concerned is difficult to be described accurately and / or there is a lack of available data which is a common case for a lot of surface runoff prediction problems (Alagha, 2013). Therefore, AI techniques show off a reliable performance in modeling complicated hydrological processes with less cost, effort and data (Almasri and Kaluarachchi, 2005; Trichakis *et al*., 2009). Furthermore, in many cases AI models outperformed the process based models (Banerjee *et al*., 2011; Alagha, 2013).

To examine the suitable AI and ANNs models, there are many performance indicator should be used which is calculated by the monitored data and predicted data. The most common performance indicators used in model evaluation are the coefficient of determination (R^2), root mean square error (RMSE), mean square error(MSE), Nash Sutcliffe Efficiency (NSE), and correlation coefficient (r) (Solaimani, 2009). Sahoo *et al.*, (2006) reported that when comparing the performances efficiency of many models with each other, performance indicators that used will determine which models best performance, because each performance indicator may give different results between models. Therefore, the performance indicator for each model should be evaluated according to all performance indicators(Maier *et al*., 2010).

A neural network consists of at least three or more layers, comprising of an input layer, an output layer and a number of hidden layers, as shown in Figure 2.1. In one layer each neuron is connected to the neurons in the next layer, but there are no connections between the units of the same layer. Depending on the problem the number of neurons in each layer may vary. Figure 2.2 shows the basic component of the node of ANNs.

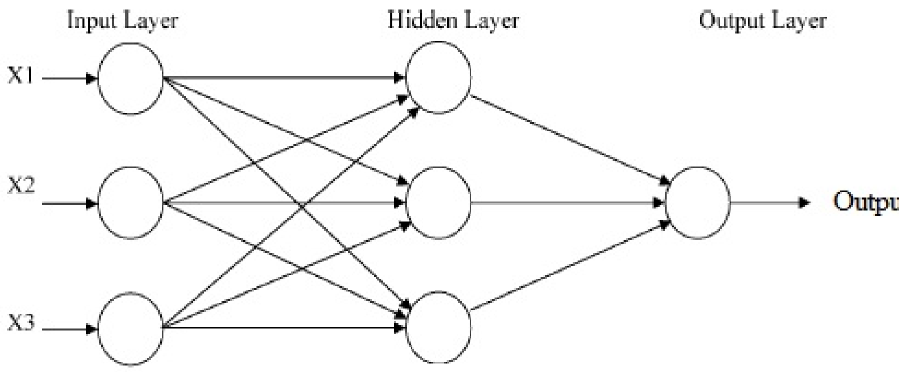

Figure 2.1Structure of feed- forward ANN(from Srinivasulu and Jain, 2006)

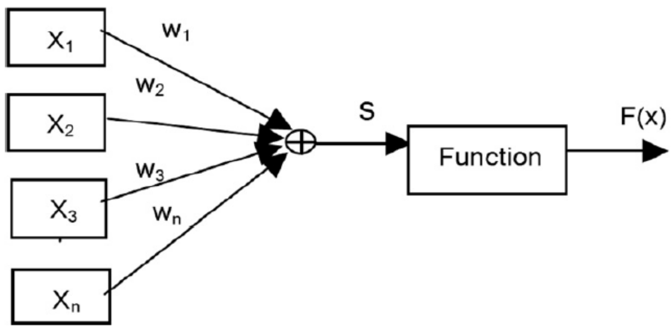

Figure 2.2 Basic component of the node (from Samani *et al*., 2007)

Ghumman *et al* (2011) compared both the ANN and conceptual model.Two performance indicators (coefficient of efficiency(CE) and root mean square error (RMSE)) were used to select the best model.The results showed that at the training