MODELING AND FORECASTING OF MONTHLY RAINFALL USING MATHEMATICAL MODELS AND MACHINE LEARNING MODELS: A CASE STUDY IN MORPHOU, NORTHERN CYPRUS

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF APPLIED SCIENCES

OF

NEAR EAST UNIVERSITY

By

JULIA ALJAMAL

In Partial Fulfilment of the Requirements for the Degree of Master of Science in Civil Engineering

NICOSIA, 2020

JULIA ALJAMAL
MODELING AND FORECASTING OF MONTHLY RAINFALL USING MATHEMATICAL MODELS AND MACHINE LEARNING MODELS: A CASE STUDY IN MORPHOU, NORTHERN CYPRUS
NEU 2020

MODELING AND FORECASTING OF MONTHLY RAINFALL USING MATHEMATICAL MODELS AND MACHINE LEARNING MODELS: A CASE STUDY IN MORPHOU, NORTHERN CYPRUS

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF APPLIED SCIENCES

OF

NEAR EAST UNIVERSITY

By
JULIA ALJAMAL

In Partial Fulfilment of the Requirements for the Degree of Master of Science in Civil Engineering

NICOSIA, 2020

Julia ALJAMAL: MODELING AND FORECASTING OF MONTHLY RAINFALL USING MATHEMATICAL MODELS AND MACHINE LEARNING MODELS: A CASE STUDY IN MORPHOU, NORTHERN CYPRUS

Approval of Director of Graduate School of Applied Sciences

Prof.Dr. Nadire ÇAVUŞ

We certify this thesis is satisfactory for the award of the degree of Masters of Science in Civil engineering

Examining Committee in Charge:

Prof.Dr. Hüseyin Gökçekuş	Committee Chairman, Supervisor, Civil Engineering, NEU
Assist. Prof. Dr. Youssef KASSEM	Co-Supervisor Mechanical Engineering Department, NEU
Prof.Dr. Gözen Elkıran	Civil Engineering Department, NEU
Assoc.Prof. Dr. Fidan Aslanova	Civil Engineering Department, NEU
Assoc. Prof. Dr. Hüseyin Çamur	Mechanical Engineering Department, NEU

We sincerely state that every academic rule and ethical conduct are stretched to every detail obtained and presented in this research work. In addition, every material and result that are not attached to this study are well cited and referenced as it is expected by the governing rules and regulation.

Name, Last Name: **Julia ALJAMAL**

Signature:

Date: 17/8/2020

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and thanks to my supervisors Prof. Dr. Hüseyin GÖKÇEKUŞ and Assist. Prof. Dr. Youssef KASSEM for their guidance, suggestions and much good advice and his patience during the correction of the thesis. They have been my mentor and my support at all the times. I am very thankful to them for giving me an opportunity to work on interesting projects I am immensely grateful for your kindness, patience, time and professional contributions to the success of my study. Thanks for always pushing me for more.

Huge appreciation goes to my family who not only have been my back bone through the phases of life but have also been my source of motivation in my pursuit of excellence. For the financial, moral, spiritual and psychological assistance.

I also appreciate the support of my good friend, Humphrey Adun for his contribution and encouragement during the course of this thesis. Finally, I give sincere thanks to my colleagues and everyone who have helped in anyway in my journey to this point in life.

This research was generously supported by the Department of Civil Engineering of Near East University. I am also grateful to all supporters**.**

To my parents ...

ABSTRACT

One of the major challenges of a hydrological cycle is forecasting rainfall. It is very difficult, because, given the complexity and unforeseen variability, it is still not possible to create an ideal model. Therefore, the main idea behind this research work is for utilizing Artificial Neural Network (ANN) to determine the occurrence of rainfall. Two types of ANN models were utilized; the Radial basis neural network (RBFNN) and cascade neural network (CFNN). Ninety-nine models of both ANN are established in this study by varying the weather parameters. The ANN simulations made use of a historical data (1985 – 2017) of meteorological parameters which is known to affect rainfall occurrence. The models output is measured using the R-squared value and root mean squared error. Of the 99 ANN models, the best prediction was provided by ANN-68 RBFNN and ANN-92 CFNN neural network. The ANN models was also compared with empirical response surface methodology (RSM) models. Based on that fact that previous studies have shown that machine learning algorithms are less efficient prediction models, the RSM analysis was done using the optimal parameters that yielded the most accurate prediction models for the ANN models. This study also compared the prediction accuracy between the RSM model and multiple linear regression model, and the coefficient of correlation result shows that the RSM model gives a better accuracy. The result confirmed that the ANN models had better prediction accuracy. The proposed approach also shows how to use the ANN modeling technique to classify the best meteorological variables needed for the most important rainfall-affecting parameters that are concerned with climatology. Finally, forecasting of future rainfall was done using the winters' method in the Minitab software environment and ARIMA model using python programming language. Comparison between both models showed that the ARIMA model gave a more accurate forecasting result.

Keywords: Artificial neural network; temperature; global solar radiation; sunshine duration; rainfall; wind speed; response surface methodology

ÖZET

Hidrolojik bir döngünün en büyük zorluklarından biri yağış tahminidir. Çok zordur, çünkü karmaşıklık ve öngörülemeyen değişkenlik göz önüne alındığında, ideal bir model oluşturmak hala mümkün değildir. Bu nedenle, bu araştırma çalışmasının arkasındaki ana fikir, yağış oluşumunu belirlemek için Yapay Sinir Ağını (YSA) kullanmaktır. İki tip YSA modeli kullanılmıştır; Radyal temel sinir ağı (RBFNN) ve kaskad sinir ağı (CFNN). Bu çalışmada hava parametrelerini değiştirerek her iki YSA'nın doksan dokuz modeli oluşturulmuştur. YSA simülasyonları, yağış oluşumunu etkilediği bilinen meteorolojik parametrelerin geçmiş verilerinden (1985 - 2017) yararlanmıştır. Model çıktısı R kare değeri ve kök ortalama kare hatası kullanılarak ölçülür. 99 YSA modelinden en iyi tahmin ANN-68 RBFNN ve ANN-92 CFNN sinir ağı tarafından sağlandı. YSA modelleri ampirik tepki yüzey metodolojisi (RSM) modelleri ile de karşılaştırıldı. Önceki çalışmaların, makine öğrenme algoritmalarının daha az verimli tahmin modelleri olduğunu gösterdiği gerçeğine dayanarak, RSM analizi YSA modelleri için en doğru tahmin modellerini veren optimal parametreler kullanılarak yapıldı. Bu çalışma aynı zamanda RSM modeli ile çoklu lineer regresyon modeli arasındaki tahmin doğruluğunu karşılaştırmış ve korelasyon sonucunun katsayısı RSM modelinin daha iyi bir doğruluk sağladığını göstermektedir. Sonuç, YSA modellerinin daha iyi tahmin doğruluğuna sahip olduğunu doğruladı. Önerilen yaklaşım ayrıca klimatolojiyle ilgili en önemli yağış etkileyen parametreler için gerekli en iyi meteorolojik değişkenleri sınıflandırmak için YSA modelleme tekniğinin nasıl kullanılacağını göstermektedir. Son olarak, gelecekteki yağış tahminleri Minitab yazılım ortamında kışın yöntemi ve python programlama dili kullanılarak ARIMA modeli kullanılarak yapıldı. Her iki model arasındaki karşılaştırma ARIMA modelinin daha doğru bir tahmin sonucu verdiğini gösterdi.

Anahtar Kelimeler: Yapay sinir ağı; sıcaklık; küresel güneş radyasyonu; güneş ışığı süresi; yağış; rüzgar hızı; tepki yüzeyi metodolojisi

TABLE OF CONTENTS

CHAPTER 3: METHODOLOGY

CHAPTER 4: RESULT

CHPATER 5: CONCLUSION

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANFIS:	Adaptive neuro-fuzzy inference
ANN:	Artificial neural network
AI:	Artificial Intelligent
ARIMA:	Autoregressive integrated moving average
BOM:	Bureau of Meteorology
CFNN:	Cascade Forward Neural Network
CLR:	Cluster wise Linear Regression
GSR:	global solar radiation
IEE:	irrigation economic efficiency
IUE:	irrigation use efficiency
KNN:	k-Nearest Neighbors Method
MLP:	multi-layer perceptron
MSG:	Meteosat Second Generation
RBFNN:	Radial Basis Function Neural Network
RF:	Random Forests
RSM:	Response surface methodology
SD:	sunshine duration
SVM:	Support vector machine

CHAPTER 1

INTRODUCTION

1.1 Overview of Rainfall and Neural Network Prediction Models

Considering that the Cyprus is in the Mediterranean Sea region, its climate is similar to the Mediterranean climate. However, variations relating to climate of Cyprus is dependent on geographical conditions. The surface area of Cyprus is 9,250km^2. North Cyprus is a geographical region in Cyprus whose surface area is 3,355km^2 **(**Mehmet & Bicak, 2002; Hobbs 2016**).** The land areas in the environs of Northern Cyprus are diversifies; some of which are agriculture which constitute 56.7%, forestry which takes up about 19.5%, 5% of grass area **(**Phillips-Agboola & Egelioglu, 2012**).** The land area of Northern island is also made of 10.7% covered with towns, villages, rivers, reservoirs and about 8.2% of bare land. The irrigable land makes up 87km^2.

Rainfall plays an important role to humans, plants and animals. Its significance is seen in agriculture and farming; water resource is one of the most natural resource on earth. There is a limitation of water resources in the Island. Rainfall is very vital to the generation of water resource towards the Northern side of Cyprus, and two-third of rainfall starting from October happened within a period of about 5months **(**Song et al., 2018**).** Also coupled with increasing population and exponential growth in economic activities on the Island, the availability of water resource is important in the study area. Climate changes is also a global issue, which affects water availability **(**Seino et al., 2018**).** This thesis proposes to analyze the impact of meteorological parameters on rainfall, because awareness of availability of rainfall is important for agricultural activities, urban domestic water usage, and other economic activities like production **(**Iizumi & Ramankutty, 2015**).** The prediction of rainfall helps in managing macro-level problems like food and agriculture, due to the growing poor rainfall in the globe as a result of climate changes. (Lima & Guedes, 2015).

The prediction of rainfall have been studied over the years using machine learning algorithms. Machine learning is a component of computer science which utilizes the concept of artificial intelligence in computing mathematical or statistical relationship of data. The advantage of machine learning is that it can compute algorithms which cannot be represented by conventional mathematical methods. In other words, machine learning algorithms find complex non-linear relationship between input and output, and this has gathered practical applications for pattern recognition, classification, and prediction analysis in several fields of study. Artificial neural network is a one of the most utilized machine learning tools in academic researches. It consist of input, hidden and output layers, designed to simulate the operation of the human brain. Upon that, it contains neurons in each layers, which stores specific information about the relationship between input and output layers. Previous studies have developed artificial neural networks for rainfall prediction.

In a study by (Mislan et al., 2015**),** back propagation ANN is used in predicting monthly rainfall in Tenggarong station in Indonesia. The study found that a high accuracy was gotten in the prediction process. A similar study was carried out **(**Kashiwao et al., 2017**)** for Japan. The study utilized multilayer perception (MLP-ANN) and radial basis function neural network (RBFNN) in analyzing different meteorological parameters like atmospheric pressure, temperature, humidity, vapor pressure as input variables for rainfall prediction. In conclusion, RBFNN was less effective in predicting compared to MLP-ANN. **(**Dash et al., 2018**)** worked on Kerala state to check how rainfall is predicted in that region. (ELM), K-nearest neighbor (KNN) and artificial neural network (ANN). The study proved that the ELM architecture gave the most accurate prediction. (Hashim et al., 2016**)** used adaptive neuro fuzzy inference (ANFIS) to verify the best meteorological variables for rainfall prediction. The study showed that amongst the meteorological factors considered, the most significant variable for rainfall prediction was wet day frequency. A study by **(**Bagirov et al., 2018**)** compared support vector machine, MLP-ANN, k-nearest neighbor, and multiple linear regression in rainfall prediction of 24 stattions in Australia. The result showed that SVM gave the best prediction accuracy.

Based on literature, it is seen that the most utilized meteorological parameters in developing machine learning models are wet day frequency, vapor pressure, concentrated and lowest temperature, cloud cover, dew point, humidity and wind speed. In this thesis, the meteorological parameters utilized are minimum temperature, maximum temperature, average temperature, global solar radiation, sunshine duration, and wind speed for prediction of regular rainfall. This thesis used extra input variables of sunshine duration and global solar radiation. This study used Guzalyurt as the main research region. This choice was made because, according to Phillips Agboola & Egelioglu (2012), Güzelyurt has the biggest dams in Northern Cyprus with capacity of 1470 m3 for Akdeniz dam and 4120 m3 for Gemikonağı dam. Also, according to State Hydraulic Works (2003), the biggest coastal aquifer is Güzelyurt aquifer.

1.2 Scope of Study

The literature review covers most continents of the world, including Africa, Asia and so on. This present research work is a case study of a location called Morphou, in the Northern part of Cyprus. In summary, the concentration of the rainfall was never the same, it ranges from 300mm towards the plains to an amount of about 1200mm around Troodos found within the southwestern region of the Island. Güzelyurt aquifer known as the most important supplying aquifer receives water from the remaining drainage area of the southwest of the island. Regarding the literature review, it reveals a clear lack of investigation of the link among rainfall and meteorological parameters particularly global solar radiation and sunshine duration. The idea behind the inclusion of these factors to the models is summarized based on the several studies related to drought variation and amount of groundwater.

- There is a significant relationship between the amount of rainfall, solar radiation, sunshine duration and drought amount at a specific region.

- The factors that affect the occurrence of drought are rainfall, solar irradiation, cloud cover surface roughness etc.

Regarding the literature review, it reveals a clear lack of predicting the amount of monthly rainfall at a specific region using mathematical models except Multiple Linear Regression. Therefore, the objectives of this study are then:

- Investigate the link among rainfall and meteorological parameters as related to Morphou region, Northern Cyprus.
- Develop mathematical equations to evaluate the once-a-month rainfall through the use of a method called a response surface methodology (RSM). These equations depend on meteorological parameters.
- Evaluate the accuracy of RSM method and Machine learning models including CFNN and RBFNN with different combination of input parameters.
- Forecast the future amount of rainfall of 2015-2025

1.3 Significance of the Study

This thesis gains its significance in the role of rainfall to agriculture, flood prediction and management, water reserve management and rainfall prediction. Rainfall prediction is important for agricultural planning, in order to ensure optimum production of seasonal plants. Also, this thesis will be instrumental for government planning for flood management, especially in the flood-prone region as Cyprus, in order to eradicate or minimize human and property loss. Finally, rainfall prediction will help in daily planning of social activities.

1.4 Problem Statement

Despite the presence of machine learning models for rainfall prediction in literature, there is still a lack of accurate prediction accuracy for rainfall. This constitutes a huge concern, as more accurate rainfall prediction is important for agriculture, flood management and water resource reserves. Also, Cyprus is a region which is prone to flooding, hence, accurate rainfall prediction is crucial for better flood management strategies.

1.5 Methodology

The input variables of meteorological data are used in predicting monthly rainfall, using RBFNN, CFNN and RSM models. Seven available input elements including number of month (NM), minimum temperature (Tmin), maximum temperature (Tmax), average temperature (tavg), wind speed (WS), global solar radiation (GSR) and sunshine duration (SD) will be considered to assess their influence on prediction of monthly rainfall. Several models will be built which include functions for various possible combinations of the used inputs. The software environment used for the analysis is Matlab R17b. The study area for which the meteorological data is retrieved is Morphou, in northwestern part of Cyprus. The data retrieved is for between 1985 and 2017. The drainage area of the region of Morphou flanks towards the direction of the northern region of North Cyprus, specially renewing the ground water resource of Guzelyurt aquifer, that constitute the primary water entering the aquifers.

1.6 Overview of the Thesis

The information provided below illustrates how this research work is done:

Chapter 1 states an outline of the significance of water resource concerning the subject matter, rainfall prediction, and some literature regarding ANN modelling for rainfall prediction. Also, the scope and aim of the study are explained. The significance and limitation of this thesis is also highlighted in details.

Chapter 2 reviews relevant literature on water resource and management in Cyprus. Also literature regarding prediction of rainfall is also discussed.

Chapter 3 gives detailed explanation of the methodology used in this thesis for rainfall prediction. The study area is also discussed, and the data for carrying out the analysis in this study shown in this chapter.

Chapter 4 focuses on the result of the machine learning and mathematical models used in this thesis. The comparison of the results is also made to ascertain how every model created handle prediction.

Chapter 5 discusses every detail in conclusion concerning the work and the practical implications of the results obtained in this thesis. Also, further study regarding this thesis is also stated.

CHAPTER 2

LITERATURE REVIEW

This section gives a detailed review of previous studies on water resources in Cyprus. Also the prediction modelling of rainfall globally and specifically for Cyprus is also discussed in the section.

Rainfall plays a huge role in global economy, as water resource is needed for several economic activities like agriculture, production and domestic uses. Hence, the prediction is rainfall is an important activity to the global economy as it helps in efficient planning, either for its efficient utilization or to avoid risks like flooding. It is known that rainfall prediction is not an easy task, especially when high prediction accuracy is needed. Empirical formulas have been previously designed for this purpose, however due to their low accuracy, machine learning tools have been utilized for predictions, and they have shown better performance prediction accuracy.

2.1 Related Works on ANN Prediction of Precipitation

Abbot et al. (2014) made use of artificial neutral networks to study the Input selection and optimization for monthly precipitation prediction in Queensland, Australia. They used three ways in predicting the medium-term once-a-month precipitation in Queensland: first of all, they made use of classier statistical modeling techniques in definite ANNs than historically used by the BOM (Bureau of Meteorology). Secondly, forecasting precipitation as a constant function deployed by a model instead of a small number of distinct categories with an assigned probability. Finally, a model, that can easily adapt to test and include extra input data series, simultaneously develops as climatic knowledge grows.

Kashiwaoet al.(2017) used data they got from Japan Meteorological Agency to develop and test a local precipitation prediction system using artificial neural networks (ANNs). They concluded that this method is sufficient enough to predict the precipitation in Japan and that MLP (multi-layer perceptron) model's precipitation performance is better than that of RBFN (radial basis function network) model.

Bagirov rt al. (2017) worked on forecasting monthly precipitation in Victoria, Australia by developing and using Cluster wise Linear Regression (CLR) technique for the process. Virtually all the regions in the study area concurred that is more suiTable for predicting precipitations than CREM (cluster wise regression method based on EM algorithm), MLR (multiple linear regression), SVM regression (support vector machine) and ANNs. It is said that CLR method can substitute well for other models for predicting precipitation. Also the study showed that New policy, planning and management decisions are needed for more sustainable operation of water resources systems can be got when CLR method is applied in hydrological study.

Lazri et al. (2013) developed ANN models using the retrieved temperature from the SEVIRI radiometer, to estimate precipitation. The result showed a high prediction accuracy using MLP algorithm.

Nastos et al. (2013) utilized Artificial Neural Networks in Athens (Greece) to examine the precipitation concentration forecast. The result showed that minimum once-a-month rain concentration is related to the next four months in a row (ANN#3 model; MBE= −0.01) and the worst one is related to the maximum once-a-month rain concentration for the next four months in a row (ANN#2; MBE=+1.5). In conclusion, one can never tell maybe in the future ANNs could be trusted for forecasting rain concentration.

Bagirov et al. (2018) studied the monthly precipitation in Australia based on climatic conditions. The monthly precipitation was predicted by studying data driven models efficiency like regression's SVM, the multiple linear regression, the k-nearest neighbors and the artificial neural networks. They discovered that SVM (support vector machine) and ANN (1) are the best among others for predicting precipitation. They would have considered kNN (k-Nearest Neighbors Method) and ANN (0) models too the two did not show excellent accuracy in many locations but showed in few.

Ramana et al. (2013) developed Artificial Neural Network (ANN) using wavelength method to forecast the once-a-month precipitation in Darjeeling rain gauge station. They were able to discover that wavelet neural network models have better results than ANN models.

Bisht et al. (2015) prediction of average once-a-month precipitation of Nainital town concerning wet bulb, dry bulb, minimum temperature, maximum temperature and wind speed in Nainital town using the back propagation neural network model. Not too many errors are found in SVM and ANN models in once-a-month precipitation assessment. In conclusion, precipitation can be predicted for years to come. The study showed that subtle Rainfall-Runoff interactions concerning all cases can be perfectly absorbed by just a neutral network model.

Alhashimi et al. (2014) used Artificial Neural Network and Time Series Models to predict precipitation in Kirkuk. Testing data set have a higher correlation coefficient R2 value of (0.91) and lower Root Mean Square Errors RMSE values of (27.278) by just using ANN model with four inputs. ARIMA (autoregressive integrated moving average) and MLR (Multi Linear Regression) models are not in any way better than ANN model in forecasting tool for predicting precipitation from available data.

Dubey (2015) utilized Artificial Neural Network to forecast precipitation in Pondicherry. The study explained that trained neutral networks by feed forward distributed time delay network where MSE (Mean Square Error) ranged from 0.0083 to 0.0120 provided the best performance.

Abbot and Marohasy (2013) studied the once-a-month rainfall prediction for the Bowen Basin by simply artificial intelligence (Queensland, Australia) and the potential benefits you can derive from it. Neutral networks as modern mathematical techniques can be deployed to attain greater seasonal precipitation forecasts. The study showed that the non-linear relationships in rainfall data can be handled by neural network algorithms. The practical implication of the study was that accurate precipitation prediction would create for a decreased economic loss relating to flood occurences in the region.

Pour et al. (2020) utilized SVM, Random Forests (RF) and Bayesian Artificial Neural Networks (BANN) to make forecast of precipitation and its limits for the period of Northeast Monsoon (NEM) in Peninsular Malaysia from synoptic predictors. From the outcome got from all the Machine Learning models, BANN was more efficient than others. Bayesian Artificial Neural Networks (BANN) provides the following results: a normalized root mean square error of 0.04–0.14, Nash-Sutcliff Efficiency of 0.98–1.0, and modified agreement index of 0.97–0.99 and Kling-Gupta efficient index 0.65–0.96 lead period prediction of approximately a month. In terms of 95% confidence interval (CI) band got from all the Machine Learning models, BANN was relatively better than the ones provided by the remaining models used. Bayesian Artificial Neural Networks (BANN)'s p-factor for forecasting rainfall indices ranges between 0.95-1.0 while that of rfactor goes within 0.25–0.49. Furthermore, the use of higher lead time in applying BANN to forecast rainfall indices gave a positive remark. Lastly, the main reason for having NEM rainfall and rainfall extremes in Peninsular Malaysia was as a result of SLP in the northern part of the South China Sea.

Ali et al. (2020) designed a hybrid CEEMD-RF-KRR model for predicting rainfall in a Pakistan region. The study combined two empirical models, which are the Kernel ridge regression and the RF models. The study was set out to find solution to the non-stationary challenges tackled by rainfall forecasting models. The result of the study displayed that CEEMD-RF-KRR model was more efficient than the comparative models at all three sites. The CEEMD-RF-KRR model provides correlation coefficient values within the range of 0.97–0.99, index of Willmott between 0.94–0.97, Nash-Sutcliffe coefficient of 0.94–0.97 and index of LegatesMcCabe within the range of 0.74–0.81. The use of index of Legate-McCabe as the main yardstick and also the results got from the lowest magnitudes of RMSE = 2.52 mm and MAE = 1.98 mm showed that CEEMD-RF-KRR provided the most accurate results in relation to Gilgit station. The recommended hybrid CEEMD-RF-KRR model provides a more accurate result in predicting rainfall. The practical implication of the study was that the proposed hybrid machine learning is efficient for accurate prediction of rainfall in the region, in order to avoid drought, and also improve agricultural activites.

Pham et al. (2020) carried out a rainfall prediction analysis using ANFIS, SVM and ANN and the result relating to the study hinted that all the Artificial Intelligence models relatively gave out a good performance. In addition, Support Vector Machines (SVM) was arguably the most efficient method for forecasting rainfall among other methods. Furthermore, input variability through the utilization of the Monte Carlo approach also confirmed that SVM is the most powerful and efficient prediction model in relation to others. For quickly and accurately predicting day-to-day rainfall, the study involving artificial intelligence can be very useful.

Velasco et al. (2019) presented an implementation of a weak-head rainfall prediction which make use of multilayer perceptron neural network (MLPNN) to develop past rainfall data. Models concerned with Multilayer Perceptron Neural Network models gave out MAE value of 0.01297 and MAE value of 0.1388. These models also provided RMSE value of 0.01512 and RMSE value of 0.01557. The use of Multilayer Perceptron Neural Network in predicting rainfall is to hint the required time space needed for the best ways of planning activities and causes of rainfall to organizations and every individual.

The Table 2.1 gives a summary of the studies which have analyzed the prediction of precipitation using ANN.

Table 2.1: Summary Table of ANN modelling of precipitation prediction

Reference	Model	Input	Output	Location
Abbot et. Al (2014)	Artificial neural network	Minimum Temperature Maximum temperature Atmospheric pressure SOI, IPO,DMI,NINO	Monthly Rainfall	Queensland
Kashiwao et al.(2017)	Multi-layer perception(MLP) Radial basis function network(RPFN)	Atmospheric pressure (on-site) Atmospheric pressure (sea-level) Precipitation Temperature Open-air temperature Vapor pressure Humidity Wind velocity	Total precipitation	Sapporo, Matsuyama,naha in japan
Bagirov et. Al(2017)	Cluster wise Linear Regression (CLR)	Maximum temperature Minimum	Monthly Rainfall	Victoria, Australia

		temperature Evaporation Vapour pressure Solar radiation		
Lazri et al. (2013)	Artificial neural network using multilayer perceptions algorithms	Brightness temperature difference between the IR3.9 lm and the IR10.8 lm channels ,Brightness temperature difference between the IR3.7 and the WV7.3 lm channel ,temperature at the 10.8-lm channel ,Brightness temperature difference between the IR10.8 and the IR12.0-lm channels ,Brightness temperature difference between the WV7.3 and the IR12.0 lm channel ,Brightness temperature difference between the WV6.2 and the IR10.8 lm channel ,Brightness temperature	Daily and monthly rainfall	North of Algeria

		difference between the IR8.7 and the IR10.8 lm channel, Average temperature at IR10.8 channel of cloud H Vertical extension of cloud ,Brightness temperature difference between the IR3.7 and the WV7.3 lm channel of ,Brightness temperature difference between the IR3.7 and the IR10.8 lm channel of cloud ,Average temperature difference of cloud between two consecutive images at IR10.8 channel, Vertical extension difference of cloud between two consecutive images ,Cloud water path difference of cloud between two consecutive images		
Nastos et. al.(2013)	Artificial neural network using multilayer	Mean monthly rain intensity of	Forcast rain intensity (mm/day) for	Athens, Greece

	perceptions algorithms	the four previous months ,Maximum monthly rain intensity of the four previous months ,Minimum monthly rain intensity of the four previous months ,consecutive months Cumulative PC of the four previous months, cumulative incidence rate with rain intensity ≥14.4 mm/day of the four previous months, Cumulative PC of the four next months ,Cumulative expected incidence rate with rain intensity ≥14.4 mm/day for the next four months	next four month	
Bagirov et. Al. (2018)	Support Vector Machines designed for Regression, Multiple Linear	Maximum temperature (Tmax), Minimum temperature (Tmin),	Monthly rainfall	Australia

	Regression, k-Nearest Neighbours Method, Artificial Neural Networks	Evaporation (Evap), Vapour pressure (VP), and Solar radiation (Rad).		
Ramana et. Al. (2013)	wavelet technique, Artificial Neural Network	minimum and maximum temperature	monthly rainfall	Darjeeling in Himalayan foot hills
Bisht et. Al. (2015)	Artificial Neural Network (ANN) and Support Vector Machine (SVM	wet bulb, dry bulb, minimum temperature, maximum temperature and wind speed	average monthly rainfall	Nainital Region
Alhashimi (2014)	autoregressive integrated moving average (ARIMA), (MLR) and the artificial neural network (ANN) techniques	air mean temperature, relative humidity and wind speed	monthly rainfall	Kirkuk station in Iraqi
Dubey(2015)	Artificial neural network	minimum temperature, maximum temperature, water vapor pressure, potential evapotranspiration and crop evapotranspiration	rainfall	Pondicherry

Abbot & Marohasy (2013)	Artificial neural network	Southern Oscillation Index (SOI), Inter-decadal Pacific Oscillation (IPO), Niño 3.4 (Nino), maximum atmospheric temperature (MaxT), and minimum atmospheric temperature (MinT)	Monthly rainfall	Queensland, Australia
Pour et al. (2020)	Support Vector Machines (SVM), Random Forests (RF) and Bayesian Artificial Neural Networks	Rainfall amount, average rainfall intensity, days with rainfall more than 95-th percentile rainfall, and dry days	Seasonal rainfall and rainfall extremes	Peninsular Malaysia
Ali et al. (2020)	Complete ensemble empirical mode decomposition (CEEMD) combined with Random Forest (RF) and Kernel Ridge Regression (KRR) algorithms in designing a	Rainfall data were collected from three stations	Monthly rainfall	Gilgit, Muzaffarabad, Parachinar in Pakistan

	hybrid CEEMD-RF-KRR model			
Pham et al. (2020)	Adaptive Network based Fuzzy Inference System optimized with Particle Swarm Optimization, Artificial Neural Networks and Support Vector Machines	Maximum temperature, minimum temperature, wind speed, relative humidity and solar radiation	Day-to-day rainfall	Hoa Binh province, Vietnam
Velasco et al. (2019)	Multilayer Perceptron Neural Network	Average temperature Maximum temperature Minimum temperature Average wind speed Relative humidity Total rainfall Visibility Day Month Year	Week-ahead rainfall	Mindanao

2.2 Related Studies on Water Resources and Management in Cyprus

Beste and Akün (2019) examined the groundwater quality and the necessary ways to manage the irrigation water system in Güzelyurt province in North Cyprus. The Figures got from Sodium Adsorption Ratio (SAR) showed that the groundwater in the study area is good. values indicated excellent groundwater quality. In addition, Magnesium Adsorption Ratio (MAR) is not good enough for the study. The implementation of groundwater for carrying out activities relating to irrigation was emphasized at the later end of this study.

Elkiran and Ergil (2006) examined water budget in North Cyprus. It was seen from the result that in order to prevent the accumulation of water, there needs to be construction of new dams. Most importantly, if there will be blockage of surface water into the sea or storing of water for use, the urgent need for a new dam and underground storage facilities are highly important. Furthermore, examined and usable rainwater must be used instead of water from aquifers in irrigating gardens in every home. Also the creation of detoxified plants is most likely to generate extra means of source of water just to counterbalance the growing water deficit.

Suner and Kırval (2017) worked on how to effectively bring out the Water Resources and the Global Warming concerning the northern region of Cyprus (Water of Peace Project). According to the results got from this study, the habitat of the study area will benefit from the increased water resource of the island that is improved by the research work done here. Rainfall and humidity will be on the rise because of this. The good use of porTable water on natural features is very likely to mitigate the adverse influence of global warming in the study area. In conclusion, the control of global warming will be achieved if there is no conversion of energy to generate porTable water and water for irrigation.

Elkiran and Ergil (2006) Integrated Water Resources Planning and Management of North Cyprus on supply and demand quantities. According to the results shown, the prevailing poor irrigation system was the cause of losing approximately half of the irrigation water. An approximate value of close to 8.6 MCM was derived for the total annual conveyance in pipelines and pipe networks. By simply installing new pipes, there is a chance of storing about 5.7 MCM of water inside the conveyance system. A Figure of about 93.2 MCM got from groundwater resources added to the 101.8 MCM of water generated during the period of 2002. There is a need for better planning for the 3.6 MCM water treated by the sanitary plants and is diverted to the sea without any utilization to come up with such value. Serious attention in educating and training the farmers must be given. Floating pipeline system between Turkey and TRNC are still under the supervision of technical and economic studies with the hope making an average diameter of pipe ranging between 0.8-1.0m to convey 75 MCM per annum, 1.0 USD per m3 is the roughly guessed value for the cost of water. The medium size desalination plant built in Gazi Magosa for Eastern Mediterranean University needs (1000 m3 /day) is 0.55 USD and from this value it can be said that the cost for the systems are not the same. In conclusion, water insufficiency that has been a big issue to overcome towards the northern part of Cyprus can be corrected by building similar plants that was built during the late period of 1990s in the southern part of Cyprus. It is a known fact that this is very good and ideal because it is reliable.

Elkiran and Ongul (2009) analyzed the current and past water budget of the country under standard condition and the condition is not favourable. The irrigation water contains roughly 70% of the total water utilized as it is shown by the past analysis of NC pertaining to water budget. It was observed that there was a decline in the capacity of the aquifers any time the drought season comes. In conclusion, the much required water in the irrigation sector was reduced from the resources during summer seasons.

A Jamal and T¨urker (2015) worked on the limestone sub aquifers in Cyprus to access its yearly water equilibrium. According to the outcome of the study, the yearly groundwater recharge that entered the sub aquifers is 1126 mm. Furthermore, the northern foothills supplying the sub aquifers is never as susceptible to effects from the climate compared to southern foothills. The data accumulation estimated the yearly removal from the 11 sub-aquifers to be 13.34 MCM, with the storage capacity being 214 cubic meter.

Elkiran and Ergil (2004) analyzed the Water Budget of Girne Region in north of Cyprus. The outcome of the study reviewed that approximately half of the irrigation water is lost as a result of prevailing poor irrigation system. An estimated value of approximately 1.6 MCM happens to be all the yearly conveyance losses in pipelines and pipe networks. The once-a-month water withdrawal from the aquifer reaches to the maximum value of about 3 MCM any time the summer season occurs. Knowing that the safe yield capacity of Girne aquifer is about 10.5 MCM annually, during drought period nearly 65 % of this safe yield amount is extracted.

Elkiran et. al. (2019) made use of laboratory experiment results in figuring out the discharge water reuse potentials as a component of integrated water resource management in Northern Cyprus. According to the outcome of the study area, about 20MCM of water contribution will most likely be enough for the water budget and there are chances it will serve the vulnerable environment well. It is relatively cheaper to implement tertiary treatment because it is not expensive at $0.2/m3 compared to desalination of water process that cost more at $1/m3

Elkiran and Aysen (2008) worked on the northern side of Cyprus to examine substitute mitigation approaches and the effects of water scarcity. According to the most relevant suggestion from the study in this region, it is advisable to transport water from Turkey to the northern side of Cyprus through the use of pipelines just to simply decrease the over-usable of the underground water. Furthermore, it is also advised that they use the wastewater reuse on some certain crop patterns. In conclusion, they suggested that the limited amount of water in Turkey can be reduced by desalinating the sea.

Yukselen (2002) analyzed how heavy metal contaminates the surface sea water quality in Northern Island. The seawater in the area of smelting got polluted by chromium, nickel and copper. During precipitation, the concentration of Copper and iron increased by the multiple of ten. The site needs an important remediation.

Gokcekus et al. (2006) used Acid Mine Drainage in Morphou Bay, Northern Cyprus to estimate Pollution of Coastal Region Impacted. The study enables us to know that acid mine drainage is the reason why ground and surface water are contaminated. They concluded that an urgent assessment is needed on both urgent attention should be given to minning area and water quality in the region.

ŞAHİN et al. (2013) utilized Modeling of Morphou (Güzelyurt) Flood and Remedial Measures. The viable ones are concerned with building a detention basin for storing water and a lateral channel for diverting extra flow from Zodia Creek to Potami Creek. In conclusion, there is an improvement in the flow carrying capacities of the creeks.

2.3 Related Studies on Implementation of ANN for Precipitation Prediction in Cyprus

Karafistan et al. (2019) checked for the occurrence of many contaminants in fish in copper mining-impacted Morphou through the utilization of artificial neural networks (ANN). High ppm values like arsenic (0.19–7.91), copper (0.00–9.41), mercury (0.02–1.06), Lead (0.04–1.29), Cadmium (0.00– 1.28), and Chromium (0.00–1.08) are found in the results. Feeding red mullet and surface feeders, like the horse mackerel, detected variations. They warm against too much consumption of fishes from these polluted areas and a more elaborated sampling strategy is suggested to rework on the study area.

Ergil (1999) used Guzelyurt aquifer in Cyprus to carry out salinity problem. The result showed that 8.5MCM/year value is what is needed to extract water from the southern part of the aquifer. 17.4 MCM/year is what is needed to be calculated to lead to the reduction of the volume of the aquifer as a result of pumping an average amount of water above the safe yield capacity. They concluded that 1565MCM is the only obtainable freshwater in the aquifer and an estimated lifespan of below 90 years is left for the aquifer.

Abdullahi et al. (2017) used ANN to evaluate the future climatic change on the basis of evapotranspiration in Cyprus. They concluded that the incomplete climate parameters have no effect on how ANN proficiently forecast future ETo (evapotranspiration) in the areas.

When you include more inputs, the performance of R2 in the study areas changes from 0.8959 – 0.9997 and 0.8633 – 0.9996

Nourani et al. (2019) were able to use artificial neural network (ANN) models for one-step ahead and three-step ahead coupled with the efficiency of three Markovian in predicting once-a-month precipitation. Most importantly in three-step ahead prediction, EANN (Emotional-ANN) model performed more in a better way than both FFNN (feed forward neural network) and WANN (Wavelet-ANN) and the reason is because it was able to handle magnifying error in multi-step ahead prediction than the other two.

Kahramanoğlu et. al. (2020) aimed at bringing out the best way of sustaining water by simply checking out the irrigation use efficiency (IUE), irrigation economic productivity (IEP), irrigation economic efficiency (IEE) and irrigation dietary efficiency (IDE) of a few number of crops in Northern Cyprus. According to the outcome of this research work in the study area, it is stated that carob and fig crops will need a minimum of 24 litres of water to yield 1kg of fruit. Furthermore, open field lettuce and greenhouse egg-plant will also need a minimum about 10 – 16 litres of water to yield 1kg of vegeTables. The irrigation economic productivity (IEP) and irrigation dietary efficiency (IDE) of crops were categorically not the same and they were also very important for lots of crops. Furthermore it is seen that appropriating crops specifically on their water intake will cut down on water consumption. In conclusion, the implementation of how crops absorb water to regroup them is most likely to cause a fall in the use of water resources as it carries on to supply the country.

CHAPTER 3

METHODOLOGY

The information pertaining the area of study, the meteorological parameters concerning meteorology, and the RBFNN, CFNN, and RSM transformed methods are fully explained in this part of this thesis. Furthermore, every model that is utilized in predicting the once-a-month rainfall are systematically shown too. Figure 3.1 shows the sequence taken in the study.

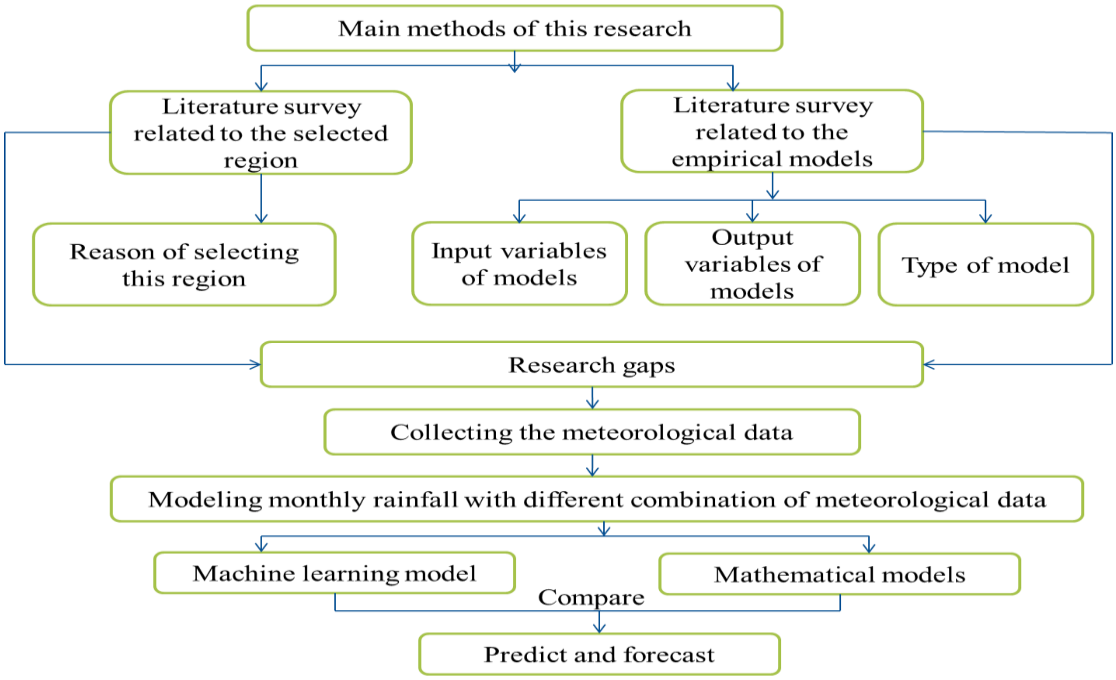


Figure 3. 1: Sequence of the research

3.1 Description of the Study Area

Morphou can be found towards the northwest region of Cyprus. Figure 3.2 depicts the location while Table 3.1 displays the area-specific information. Seven meteorological parameters can be found in the dataset in the years that falls within 1985 and2017. Rainfall ranges from 300mm towards the plains and mounted to about 1200mm within the range of Troodos placed totally towards the southwestern region of the island. A section of the drainage renewed the the water underneath the earth in the Güzelyurt aquifer. Güzelyurt aquifer is referred to as probably the major source of water (Gökçekuş, 2019; Alsalibi, 2019). In conclusion, with the all said earlier, the rainfall prediction done in this region is very useful.

Figure 3. 2: Cyprus

Table 3. 1: Morphou

Region location	
Latitude (°N)	35° 12' 3.528"
Longitude (°E)	32° 59' 26.808"
Elevation (m)	49

3.2. Simulation Using RBFNN and CFNN

Every model utilized in predicting the once-a-month rainfall in study area are shown below:

3.2.1. Radial Basis Function Neural Networks (RBFNNS)

The function of human's brain was stimulated through the use of many created algorithms as a result of the achievements in neuroscience during mid-1980s. The result for non-linear mapping between inputs and outputs were got using Artificial neutral networks (ANNs). These units are commonly called neuronss (Najafi-Marghmaleki et. al., 2009; Esmaeili-Jaghdan, 2016). Every neurons is like a coding system that works in similar with other neurons and this fact alone regard all neuronss as the most vital part of ANNs. The output of a neuron is regarded as the processed information and it will definitely be the input of the next neuron while in conclusion, it will be the output of the network. The adjusted weights in interconnection layers can be used to train the ANNs. There is never a need for a statistical training process due to the fact that every network can adequately learn from preceding trainings. These networks have become very common in every sector as a result of issues with errors in information gathered, or lack of preprocessing of data. The application of ANN span areas that involve classification or pattern recognition. The architecture of RBFNN consist of layers which are interconnected in storing and transferring information about the relationship between input and output variables. This is shown in Figure 3.3. Information is received by the input layer, and this is applied to a transfer function that is already defined. The values for model inputs are much similar to that of nodes of input layer. RBF-NN's key section of structure is known as the hidden layer (Panda et.al., 2008; Haykin, 2004).

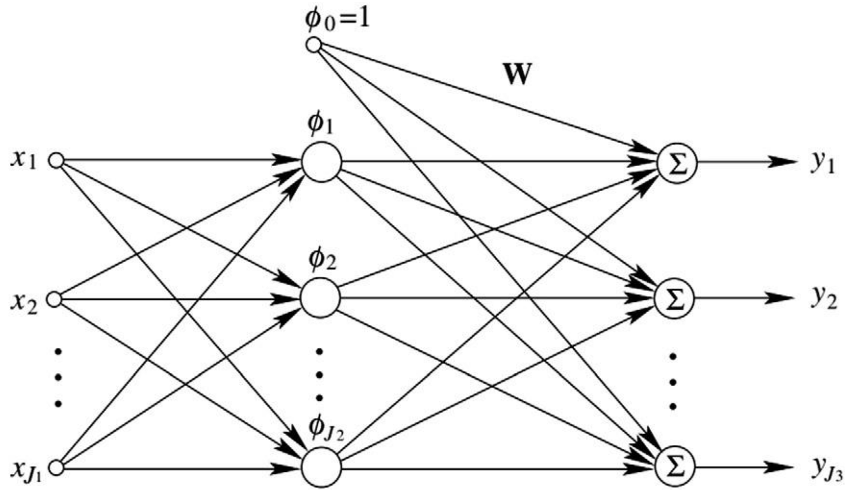


Figure 3. 3: Schematic representation of RBF-NN (Barati-Harooni & Najafi-Marghmaleki 2016)

3.2.2. Cascade forward neural networks (CFNNS)

ANN is also called an artificial intelligence system. The biological nervous systems is the inspiration behind the ANN to work as a processing information system (Artrith & Urban, 2016). It is very much possible for ANN to recognize patterns and pick up information whenever it is in connection its environment (Jafar et al., 2010). Artificial Neural Network (ANN) consists of layers; and every of these layers contains neuron (Dharma et al., 2017). Massive weight links join these neurons together; this permtis the conversion of data amid the layers. This architecture relies so much on the training stage and the testing stage to forecast every system's reaction. The non-linear relationship between the input and output variable is computed and this is stored in the neurons; which is done to discover the important values of the weight values in the training process. In predicting the new data, there occurs a recording of the connection regarding input and output variables. But, on the other hand, the performance of the system is tested using a part of the input data and the data got from prediction is put against the real data in the testing stage. ANN model can be used to find solutions to complex problems that look relatively difficult to unravel through the utilization of conventional models. It can also be used to find solution to problems that are without algorithmic solution. It could also be that the algorithmic solution requires higher technical approach to solving the issues (Shabanpour et al., 2017). The CFNN topology's structure begins with the input and output neurons. The neural network formerly contained every output neurons; therefore, network contains the new neurons. The network makes an attempt to make the most of the relationship of the inputs and output variables by constantly updating the errors in the network. Not until the lesser error value within the network gained, the process carried out here will not stop. This network is labelled as a cascade due to this. CFNN comprises the input, hidden and output layers. The output value can only be represented when the subsequent value got from the process involved in this part undergoes a transfer function (Khatib et al., 2012b).

The neurons access weights and biases' values. Learning data and the controlled training process perform this assessment. The adjustment done by Artificial Neural Networks (ANNs) to biases and weights' values is for decreasing error rate amid outputs and targets whenever the training process occurs.

The network's performance depends on the amount of hidden neurons. Though, there has not been an official method in literatures in discovering highest neurons' amount. A lot of related works in connection to research done here (Blum, 1992; Boger & Guterman, 1997) find answer to the question about neurons' amount through the use of different unofficial ways. High training and generalization errors are very likely to happen any time the optimal numbers become higher than the hidden neurons numbers. But, on the other hand, over-fitting and high variance will only come about when optimal numbers' number become lower than the hidden neurons. Therefore, the finest performance of the network will only be got the calculation involving the optimal number of hidden neurons is done (Geman et al., 1992).

The performance of the model will be ascertained by using statistical methods like RMSE, and this will be computed from iterations. The lowest RMSE value indicate high prediction accuracy.

Thus, Fig.4 shows the flowchart of the sequence of training and testing of the data for accurate rainfall prediction. The explanation on how to predict the productivity in the sequence of the suggested CFNN method can be seen below:

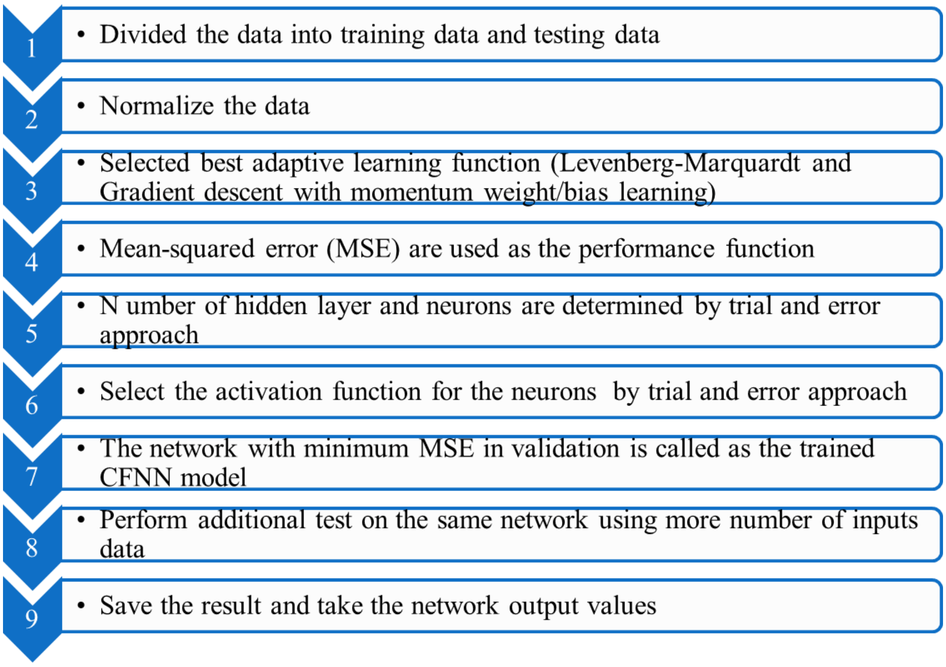


Figure 3. 4: Steps of CFNN method

3.2.3 Training and testing

Based on Levenberg-Marquardt (LM) optimization, TRAINLM becomes visible as a training function providing informs about neutral connections' values concerning weight and bias. The back propagation algorithm, known as a gradient descent algorithm, can be utilized as a learning algorithm. Linear and non-linear become activation functions on behalf of the neurons. The logistic-sigmoid (logsig) and tangent-sigmoid (tansig) become activation functions. The amount produced falls within the range of 0 – 1. It is expressed below:

$$\text{logsig} = \frac{1}{1 + e^{-x}} \tag{3.1}$$

$$\text{tansig} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3.2}$$

Equation (3.3) can be utilized for normalizing the value of data between 0 and 1. Eq. (3.4) is used in retrieving the actual data from the normalized values.

$$x_n = \frac{x_{actual} - x_{min}}{x_{max} - x_{min}} \tag{3.3}$$

$$x_{actual} = x_n(x_{max} - x_{min}) + x_{min} \tag{3.4}$$

3.3 Response Surface Methodology (RSM)

RSM becomes vital in modelling, problem and prediction analysis when a response (dependent variable) is affected by quite a lot of variables (independent) (Ghoreishi & Heidari, 2013). RSM joins mathematical and statistical technique in determining empirical model building and optimizations. Response surface methodology technique involves developing optimum functional relationship between response and a set of input parameters as shown in equation 3.5 (Gupta, 2010). In most cases, a low order first and second order polynomial in certain regions of the independent variables are employed. The first and second order models is expressed in equation 3.5 and 3.6 (Khidhir et al., 2015).

$$y = (x_1, x_2, x_3 \ldots . x_k) \tag{3.5}$$

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon \tag{3.6}$$

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i} \sum_{j} \beta_{ij} x_i x_j + \varepsilon \qquad (3.7)$$

Where β_0 is the constant, the coefficients$\beta_1, \beta_2, . . ., \beta_k$ and $\beta_{11},\ \beta_{22}, . . ., \beta_{kk}$ are the linear and the quadratic terms, respectively, while $\beta_{12}, \beta_{13}, . . ., \beta_{k1}$ are interacting terms.

3.4 Multiple Linear Regression

Multiple linear regression is a model that analyses the relationship between how an independent variable, y, relates with two or more dependent variables, x. It shows the correlation analysis of variables in order to find out if there is a significant relationship between variables (Preacher, and Bauer, 2016). The multiple linear regression model describes a quadratic, curved, relationship between different x variables and a single y-variable. $E(Y \parallel X) = \mu + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \ldots + \beta_P X_P$

The equation shows the multiple linear regression model. In the model, $\beta_1, \beta_2, \beta_3, \beta_p$ are fixed and unknown variables. The μ variable is a systematic coefficient.

3.5. Input and Output Variables

Past studies pertaining to this research have assessed the rainfall through the use of numerous meteorological parameters (Hashim et.al.,2016 ; Mohd-Safar et.al.,2018; Mohammad pour et.al.,2018). Seven weather data were used to predict the once-a-month rainfall in this current work. Table 2 below gives more information about this present work. Additional input parameters like global solar radiation and sunshine duration are suggested by this present study. The connection between solar radiation and weather data play a big role in selecting these parameters. A number of scientific researchers have studied sunshine duration and rainfall (Díaz-Torres et. al.,2017; Miguntanna and Jayasinghe,2015; Kumar and Kaur.2016; Reddy and Ranjan,2003).

Table 3. 2: Input Parameters and output parameters

Parameters	Parameter description	Abbreviation
Input 1	Monthly minimum temperature	Tmin
Input 2	Monthly maximum temperature	Tmax
Input 3	Monthly averaged temperature	Tav
Input 4	Monthly global solar radiation	GSR
Input 5	Monthly sunshine duration	SD
Input 6	Monthly wind speed	W
Input 7	Month	NM

3.6. Rainfall Prediction with Selected Inputs

Figure 3.5 contains the various methods used in accessing the once-a-month rainfall. The CFNN and RBFNN method make use of weather data as inputs. Trial and error is used in determining the optimum ANN architecture. Different activation function, hidden layers, and input combinations are varied to ascertain high prediction performance. Several CFNN and RBFNN models are created just to get the optimum performance results.

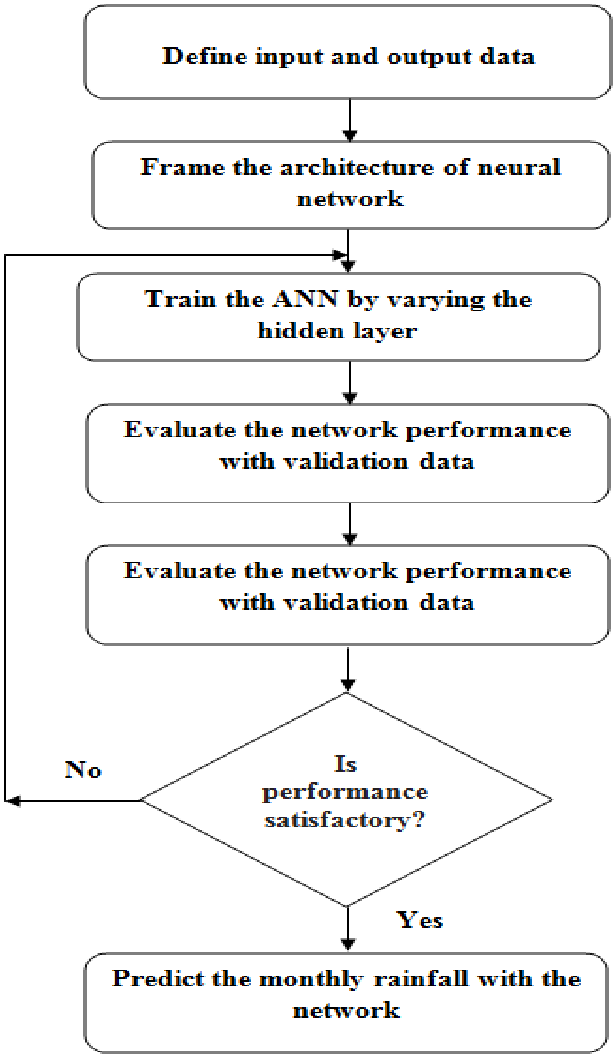


Figure 3. 5: Proposed CFNN and RBFNN model to predict the once-a-month rainfall at thestudy area

The data division utilized in this study is that of a conventional method, where the data is split on arbitrary basis (Bowden et.al. 2002). About three-quarter of the data (1985-2010) was utilized for training, whereas the left over one-quarter (2011-2017) was used for testing. CFNN and RBFNN models were trained using the training data. The trained model is used for the testing data. The accuracy of the testing result is hinged on how good the training process was. The effect of the testing data is not significant on training. Testing data gives a self-regulating network performance throughout training.

The range of values of the climate variables can be seen in Table 3.3. The inputs numbers for developing CFNN and RBFNN models are shown in Table 4.4. The performance of the CFNN model can mostly be influence by hidden layers' numbers and neuronss' numbers. The used construction for the CFNN model in this study is displayed in Figure 3.6. The number of epochs numbers consist are 100000 while that of performance goal is 0.001 in this research. The hidden layers range from 1 – 10 neuronss. The number of neurons range between 5 and 50 neurons.

Table 3. 3: Range of minimum and maximum values of variables

Input	Parameter description	Limit		Unit
		Minimum	Maximum	
Tmin	Monthly minimum temperature	22.0	1.5	℃
Tmax	Monthly maximum temperature	36.3	14.1	℃
Tav	Monthly averaged temperature	29.0	7.3	℃
GSR	Monthly global solar radiation	703.8	0	Cal/cm^2 -day
SD	Monthly sunshine duration	12.7	0	h/day
W	Monthly wind speed	4.8	0.4	m/s
NM	Months			
Output				
R	Monthly rainfall	159.0	0	Mm

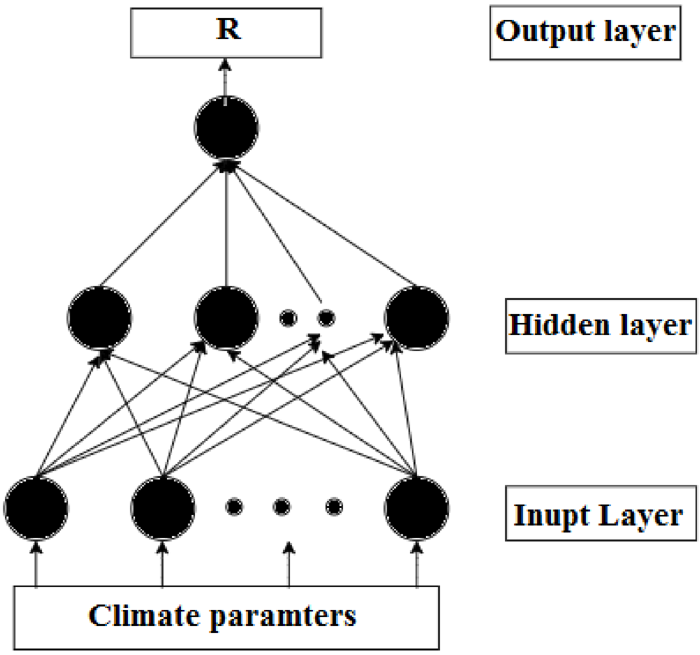


Figure 3. 6: The CFNN model structure

Table 3. 4: ANN models developed in this thesis

Model	**1-input**		
Model 1.	T_{min}	Model 2.	T_{max}
Model 3.	T_{ave}	Model 4.	WS
Model 5.	GSR	Model 6.	SD
Model 7.	NM		
	2-inputs		
Model 8.	T_{min}, T_{max}	Model 9.	T_{min}, T_{avg}
Model 10.	T_{min}, WS	Model 11.	T_{min}, GSR
Model 12.	T_{min}, SD	Model 13.	T_{min}, NM
Model 14.	T_{max}, T_{avg}	Model 15.	T_{max}, WS
Model 16.	T_{min}, T_{avg}	Model 17.	T_{max}, SD
Model 18.	T_{max}, GSR	Model 19.	T_{max}, NM
Model 20.	T_{avg}, WS	Model 21.	T_{avg}, GSR

Model 22.	T_{avg}, SD	Model 23.	T_{avg}, NM
Model 24.	WS, GSR	Model 25.	WS, SD
Model 26.	WS, NM	Model 27.	GSR, SD
Model 28.	GSR, NM	Model 29.	SD, NM
		3-inputs	
Model 30.	T_{min}, T_{max}, WS	Model 31.	T_{min}, T_{max}, SD
Model 32.	T_{min}, T_{max}, GSR	Model 33.	T_{min}, T_{max}, NM
Model 34.	T_{min}, T_{avg}, SD	Model 35.	T_{min}, T_{avg}, WS
Model 36.	T_{min}, T_{avg}, GSR	Model 37.	T_{min}, T_{avg}, NM
Model 38.	T_{max}, T_{avg}, WS	Model 39.	T_{max}, T_{avg}, GSR
Model 40.	T_{max}, T_{avg}, NM	Model 41.	T_{ave}, WS, SD
Model 42.	T_{ave}, WS, GSR	Model 43.	T_{ave}, WS, NM
Model 44.	T_{ave}, SD, NM	Model 45.	T_{ave}, SD, GSR
Model 46.	T_{ave}, GSR, NM	Model 47.	T_{min}, WS, SD
Model 48.	T_{min}, WS, NM	Model 49.	T_{min}, SD, NM
Model 50.	T_{min}, SD, GSR	Model 51.	T_{min}, GSR, NM
Model 52.	T_{min}, WS, GSR	Model 53.	T_{max}, WS, SD
Model 54.	T_{max}, WS, NM	Model 55	T_{max}, SD, GSR
Model 56.	T_{max}, SD, NM	Model 57.	T_{max}, GSR, NM
Model 58.	T_{max}, WS, GSR	Model 59.	WS, GSR, SD
Model 60.	WS, GSR, NM	Model 61.	WS, SD, NM
Model 62.	SD, GSR, NM		
		4-inputs	
Model 63.	T_{min}, T_{max}, WS, GSR	Model 64.	T_{min}, T_{max}, WS, SD
Model 65.	T_{min}, T_{max}, WS, NM	Model 66.	T_{min}, T_{max}, GSR, SD
Model 67.	T_{min}, T_{max}, GSR, NM	Model 68.	T_{min}, T_{max}, SD, NM
Model 69.	T_{min}, T_{avg}, WS, SD	Model 70.	T_{min}, T_{avg}, WS, GSR
Model 71.	T_{min}, T_{avg}, WS, NM	Model 72.	T_{min}, T_{avg}, GSR, NM
Model 73.	T_{min}, T_{avg}, GSR, SD	Model 74.	T_{min}, T_{avg}, SD, NM

Model 75.	T_{max}, T_{avg}, WS, GSR	Model 76.	T_{max}, T_{avg}, GSR, NM
Model 77.	T_{max}, T_{avg}, WS, NM	Model 78.	T_{max}, T_{avg}, SD, NM
Model 79.	T_{max}, T_{avg}, GSR, SD	Model 80.	T_{min}, WS, GSR, SD
Model 81.	T_{min}, WS, GSR, NM	Model 82.	T_{min}, GSR, SD, NM
Model 83.	T_{max}, WS, GSR, NM	Model 84.	T_{max}, GSR, SD, NM
Model 85.	T_{avg}, WS, GSR, SD	Model 86.	T_{max}, WS, GSR, SD
Model 87.	T_{avg}, WS, GSR, NM	Model 88.	T_{avg}, GSR, SD, NM
Model 89.	WS, GSR, SD, NM		
		5-inputs	
Model 90.	T_{min}, T_{max}, WS, GSR, SD	Model 91.	T_{min}, T_{max}, WS, GSR, NM
Model 92.	T_{min}, T_{avg}, WS, GSR, NM	Model 93.	T_{max}, T_{avg}, WS, GSR, NM
Model 94.	T_{min}, WS, GSR, SD, NM	Model 95.	T_{max}, WS, GSR, SD, NM
Model 96.	T_{avg}, WS, GSR, SD, NM		
		6-inputs	
Model 97.	T_{min}, T_{max}, WS, GSR, SD, NM	Model 98.	T_{min}, T_{avg}, WS, GSR, SD, NM
		7-inputs	
Model 99.	$T_{min}, T_{max}, T_{avg}$, WS, GSR, SD, NM,		

3.7. Appraisal of the Developed Models

The advanced CFNN and RBFNN models were assessed expansively to predict the once-a-month rainfall. The used statistical indicators include MSE and RMSE. The formula used can be seen in the equations 3.8, 3.9, 3.10.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(a_{a,i} - a_{p,i}\right)^2}{\sum_{i=1}^{n}\left(a_{p,i} - a_{a,ave}\right)^2} \tag{3.8}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(a_{a,i} - a_{p,i}\right)^2 \quad (3.9)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(a_{a,i} - a_{p,i}\right)^2} \quad (3.10)$$

where n repsents the data count, $a_{p,i}$ represents the values that are predicted, $a_{a,i}$ stands for the retrieved values, $a_{a,ave}$ denotes the average values of the retrieved data.

3.8. Forecasting Analysis

In this section, a month ahead rainfall forecasting is computed using winters' method (ın Minitab software) and ARIMA model usıng python software. In the forecastıng models, the data is split into training and validation set to analyze the forecasting ability of the models, and then used to forecast monthly rainfall between 2017 and 2025.

3.8.1 Winters methods

The winter's method, which is a statistical forecasting method, was used in computing the time series rainfall prediction. The period of January 1985 to December 2014 was used as the past dataset for the time series. In the forecasting, the method type was varied between multiplicative and additive type. The winters' method uses past and recent data in making forecast for time series data. The choice of the winters' method for making forecast of rainfall in this thesis was because the rainfall occurrence is a seasonal time series data (Rahman et al, 2016).

The smoothing constants were also varied to get the best forecasting values. The smoothing constants are the level, trend and seasonal smoothing variables. The comparison of the accuracy of the forecasting was made using the data points of January 2015 to December 2017. The Figure 3.7 shows the time series plot of rainfall.

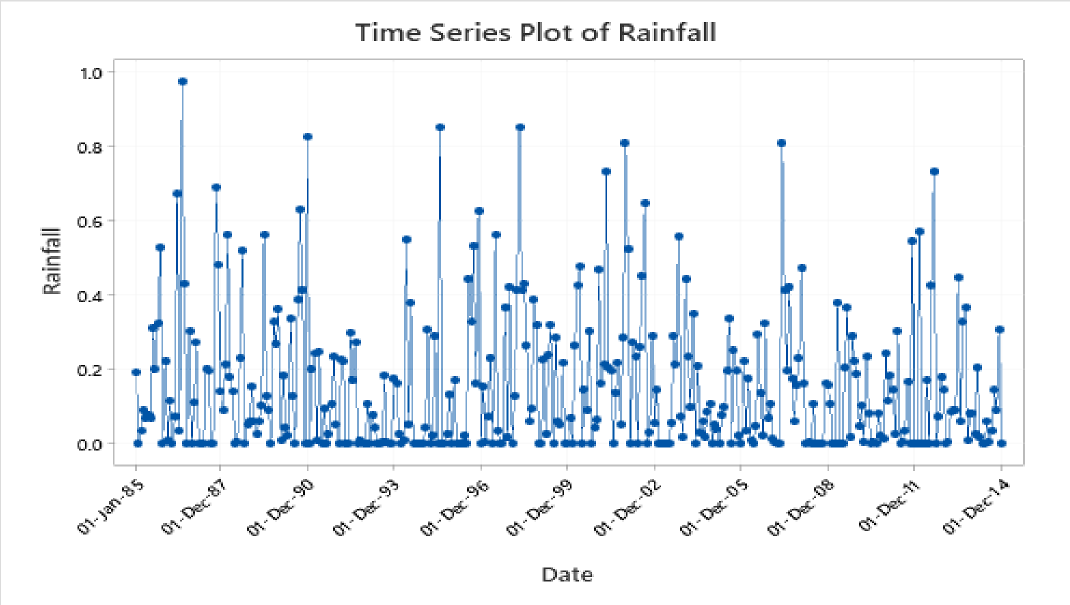


Figure 3. 7: Plot of rainfall in time series

3.8.2. ARIMA model

ARIMA is a synonym for Autoregressıve integrated moving average. It is a statistical method that is used for capturing structure of data in time series pattern. The ARIMA model uses an observation of dependent relationship of an observed data with some number of lagged observation. A standard notation is used of ARIMA (p, d, and q) where the parameters are substituted with integer values to quickly indicate the specific ARIMA model being used. The p denotes the number of lag observations included in the model, also called the lag order, the number of times that the raw observations are differenced, also called the degree of differencing is denoted as the d. The q is the size of the moving average window, also called the order of moving average. Adopting an ARIMA model for a time series assumes that the underlying process that generated the observations is an ARIMA process. This may seem obvious, but helps to motivate the need to confirm the assumptions of the model in the raw observations and in the residual errors of forecasts from the model. In this thesis, the python programming language is used in constructing the ARIMA model. The python language is used because it is a robust platform where the data can be preprocessed and iterative functions can be carried out.

CHAPTER 4
RESULTS

In this section, the ANN models will be discussed and in terms of their performance. In this study, 99 models was developed based on different combination of input parameters for rainfall prediction. Also, for excellent prediction, different architecture of ANN models was developed; with different hidden layers, neurons and transfer function. The best architecture is selected based on the least MSE and high R-square values.

Aslso, in this study, two types of ANN models was utilised. The RBFNN and CFNN. The result of both ANN models will be separately discussed.

4.1 CFNN Models

4.1.1 CFNN models with one input

Each of the input parameter was used in predicting the output variable, which is the rainfall. The training was carried out several times, and the average performance was recorded. Hidden layers from 1-3 was used, and different transfer function was tested to check the best accurate ANN architecture for prediction. A section of the data which was unused during the data known as the test data, is used on the developed model. The accuracy of the models was analysed using the MSE and R-square values. The Table 4.1 shows the different ANN architecture of each model (ANN-1 – ANN-7) with their corresponding statistical performance values (RMSE and R-squared). The result shows that the ANN-2 and ANN-3 have the best prediction accuracy in terms of the testing performance. This shows that T_{max} and T_{avg} are more significant parameters in rainfall prediction.

Table 4.1: Evaluation of the one-input model and statistical performance (CFNN)

Model	Transfer function	Hidden layer	Neurons	Epochs	MSE (training)	R^2 (Training)	R^2 (Testing)	RMSE (Testing)
ANN-1	TANSIG	2	30	0.0250	0.0198	0.4813	0.1569	0.1772
ANN-2	TANSIG	2	10	0.0092	0.0176	0.5367	0.3915	0.1384
ANN-3	TANSIG	2	20	0.0242	0.0208	0.4660	0.3825	0.1436
ANN-4	TANSIG	3	20	0.0149	0.0283	0.2562	0.0038	0.2158
ANN-5	LOGSIG	1	20	0.0204	0.0160	0.5783	0.3239	0.1484
ANN-6	LOGSIG	2	10	0.0086	0.0146	0.61811	0.1617	0.2047
ANN-7	TANSIG	2	30	0.0489	0.0294	0.2339	0.0862	0.2190

4.1.2 CFNN models with two input

22 models (ANN-8 to ANN-29) was developed for two input combinations of the input variables. Different transfer function and hidden layers was developed to test the most accurate ANN-model. The RMSE value and R-squared values was used in checking how accurate the developed models are. It is seen as shown in the Table 4.2 that the ANN models and the statistical performance of each model. The result showed that based on the testing performance, ANN-14 gives the most accurate performance, followed by ANN-8 with R-sqaured value of 0.7059 and 0.6696 respectvely. This shows that the combination of (T_{max}, T_{avg}) and (T_{min}, T_{max}) are the most efficient models for rainfall prediction, if two input variales are to be combined.

Table 4.2: Evaluation of the two-input model and statistical performance (CFNN)

Model	Transfer function	Hidden layer	Neurons	Epochs	MSE (training)	R^2 (Training)	R^2 Testing)	RMSE (Testing)
ANN - 8	LOGSIG	1	20	0.0151	0.0125	0.6719	0.669615	0.1035
ANN - 9	TANSIG	2	10	0.0104	0.0113	0.7027	0.5941	0.1182

ANN - 10	LOGSIG	3	10	0.0290	0.0213	0.4435	0.2225	0.1673
ANN - 11	LOGSIG	1	20	0.0173	0.0162	0.5718	0.4563	0.1360
ANN - 12	LOGSIG	1	20	0.0075	0.0137	0.6403	0.4897	0.1299
ANN - 13	LOGSIG	1	20	0.0200	0.0196	0.4821	0.4189	0.1414
ANN - 14	TANSIG	2	20	0.0098	0.0132	0.6590	0.7059	0.0965
ANN - 15	TANSIG	1	10	0.0212	0.0175	0.5553	0.4824	0.1349
ANN - 16	LOGSIG	3	10	0.0099	0.0116	0.692	0.6285	0.1098
ANN - 17	LOGSIG	1	20	0.0072	0.0128	0.6658	0.5505	0.1224
ANN - 18	LOGSIG	1	10	0.0090	0.0157	0.5850	0.4730	0.1265
ANN - 19	LOGSIG	3	10	0.0185	0.0148	0.6146	0.4006	0.1403
ANN - 20	TANSIG	1	10	0.0251	0.0199	0.4749	0.4356	0.1315
ANN - 21	TANSIG	2	10	0.0075	0.0159	0.5800	0.4387	0.1337
ANN - 22	LOGSIG	1	20	0.0138	0.0150	0.6038	0.5450	0.1213
ANN - 23	TANSIG	3	10	0.0159	0.0156	0.593516	0.3876	0.1457
ANN - 24	TANSIG	1	30	0.0125	0.0161	0.5741	0.4536	0.1300
ANN - 25	LOGSIG	2	10	0.0098	0.0123	0.6758	0.4199	0.1381
ANN - 26	TANSIG	2	10	0.0421	0.0327	0.1586	0.0169	0.1824
ANN - 27	TANSIG	2	10	0.0116	0.0127	0.6663	0.3892	0.1537
ANN - 28	TANSIG	3	10	0.0181	0.0170	0.5551	0.4563	0.1276
ANN - 29	LOGSIG	1	20	0.0039	0.0122	0.7035	0.3211	0.1796

4.1.3 CFNN models with three input

The Table 3 shows the different combination of three input variables for rainfall prediction, using different hidden layers, neurons and transfer functions. The performance of the models from ANN-30 to ANN-62 was checked using the RMSE values and R-squared values. The performance analysis was carried out for both training and testing model results. The epoch values are also presented in the Table 4.3. The result showed that the best accurate prediction based on the highest R-squared value and lowest RMSE vlaues in the testing model are ANN-33 and ANN-32 respectively. This shows that the combination of (T_{min}, T_{max}, and GSR) and (T_{min}, T_{max}, and NM) have the best prediction. This justifies the result of two input combination which showed that (T_{min}, T_{max}) gave accurate prediction when combined. Also, the result shows that Logsig trainig function gives the most efficient transfer function in modelling three input combinations of the most accurate prediction models. The worst model in three input combination are ANN-54 (T_{max}, WS, NM).

Table 4.3: Evaluation of the three-input model and statistical performance (CFNN)

Model	Transfer function	Hidden layer	Neurons	Epochs	MSE (training)	R^2 (Training)	R^2 (Testing)	RMSE (Testing)
ANN - 30	LOGSIG	2	20	0.0102	0.0096	0.7463	0.6371	0.1093
ANN - 31	LOGSIG	1	10	0.0076	0.0104	0.7278	0.5940	0.1123
ANN - 32	LOGSIG	1	10	0.0086	0.0115	0.6964	0.6511	0.1037
ANN - 33	LOGSIG	2	10	0.0044	0.0115	0.7201	0.6833	0.1185
ANN - 34	LOGSIG	1	10	0.0286	0.0125	0.6703	0.6463	0.1103
ANN - 35	LOGSIG	1	10	0.0124	0.0135	0.6430	0.5884	0.1150
ANN - 36	TANSIG	2	10	0.0167	0.0123	0.6755	0.6352	0.1113
ANN - 37	LOGSIG	2	10	0.0101	0.0110	0.7091	0.6563	0.1053
ANN - 38	LOGSIG	1	20	0.0188	0.0124	0.6737	0.5840	0.1138
ANN - 39	LOGSIG	3	10	0.0102	0.0115	0.6957	0.6382	0.1059
ANN - 40	LOGSIG	1	10	0.0135	0.0124	0.6919	0.5917	0.1276
ANN - 41	LOGSIG	1	10	0.0175	0.0139	0.6341	0.4593	0.1361
ANN - 42	TANSIG	3	20	0.0104	0.0134	0.6477	0.3223	0.1547
ANN - 43	LOGSIG	3	10	0.0176	0.0148	0.6104	0.2733	0.1591

ANN - 44	LOGSIG	1	20	0.0094	0.0139	0.6344	0.5819	0.1225
ANN - 45	LOGSIG	2	20	0.0134	0.0155	0.6089	0.4733	0.1579
ANN - 46	TANSIG	1	10	0.0160	0.0182	0.5820	0.3022	0.1882
ANN - 47	TANSIG	1	10	0.0065	0.0137	0.6556	0.4351	0.1595
ANN - 48	TANSIG	1	10	0.0294	0.0228	0.4002	0.2755	0.1526
ANN - 49	LOGSIG	1	10	0.0158	0.0143	0.6279	0.5518	0.1162
ANN - 50	TANSIG	1	20	0.0085	0.0128	0.6618	0.4563	0.1369
ANN - 51	LOGSIG	3	10	0.0121	0.0158	0.5857	0.4647	0.1381
ANN - 52	LOGSIG	1	30	0.0121	0.0144	0.6192	0.4432	0.1393
ANN - 53	TANSIG	3	20	0.0096	0.0138	0.6432	0.4619	0.1324
ANN - 54	LOGSIG	1	20	0.0104	0.0158	0.5858	0.2094	0.1725
ANN - 55	LOGSIG	3	10	0.0169	0.0141	0.6309	0.5938	0.1145
ANN - 56	LOGSIG	1	20	0.0095	0.0145	0.6192	0.5712	0.1201
ANN - 57	LOGSIG	1	20	0.0078	0.0130	0.6579	0.4650	0.1413
ANN - 58	TANSIG	2	10	0.0134	0.0142	0.6262	0.4746	0.1278
ANN - 59	LOGSIG	3	30	0.0051	0.0110	0.7112	0.4807	0.1386
ANN - 60	LOGSIG	3	10	0.0183	0.0156	0.5926	0.2548	0.1845
ANN - 61	LOGSIG	1	30	0.0137	0.0138	0.6347	0.4193	0.1377
ANN - 62	TANSIG	2	10	0.0112	0.0136	0.6424	0.5074	0.1230

4.1.4 CFNN models with four inputs

A total of 26 models was developed for the four input combination of the input variables for rainfall prediction, which is shown in Table 4 as ANN-63 to ANN-88. The performance of each model is checked using the statistical performance criterias. The conclusion of the most accurate and worst model is verified using the the statistical performance criterias of the testing results of each model. The result showed based on the performance criteria of the testing results that ANN-76 whose input combinations are (T_{min}, T_{max}, GSR and NM) have the most accurate predictions. This corroborates the result gotten in the three input combination also which concluded the each of the combination of T_{min}, T_{max}, GSR and NM yielded high performance prediction compared to other input combinations. Also, in this 4 input combination result, it is shown in Table 4.4 that Logsig training function provided the best traning sequence for accurate prediction. Furthermore, the worst prediction was obtained for ANN-72 model (combinations of T_{min}, T_{avg}, GSR and NM).

Table 4.4: Evaluation of the four-input model and statistical performance (CFNN)

Model	Transfer function	Hidden layer	Neurons	Epochs	MSE (training)	R2 (Training)	R2 (Testing)	RMSE (Testing)
ANN - 63	LOGSIG	1	10	0.0059	0.0101	0.7344	0.5148	0.1287
ANN - 64	LOGSIG	1	20	0.0119	0.0096	0.7494	0.6107	0.1109
ANN - 65	TANSIG	2	10	0.0073	0.0101	0.7348	0.5345	0.1254
ANN - 66	LOGSIG	1	10	0.0092	0.0109	0.7150	0.6295	0.1048
ANN - 67	TANSIG	1	10	0.0051	0.0093	0.7550	0.6005	0.1204
ANN - 68	TANSIG	1	20	0.0065	0.0090	0.7618	0.6219	0.1167
ANN - 69	TANSIG	1	10	0.0109	0.0161	0.5781	0.1570	0.1772
ANN - 70	TANSIG	1	10	0.0187	0.0157	0.5903	0.3915	0.1384
ANN - 71	LOGSIG	1	10	0.0085	0.0180	0.5370	0.3825	0.1436
ANN - 72	LOGSIG	1	20	0.0224	0.0231	0.3919	0.0038	0.2158
ANN - 73	TANSIG	1	20	0.0224	0.0150	0.6033	0.3240	0.1484
ANN - 74	TANSIG	1	20	0.0061	0.0130	0.6571	0.1618	0.2047
ANN - 75	LOGSIG	1	10	0.0066	0.0264	0.3109	0.0862	0.2190
ANN - 76	LOGSIG	1	20	0.0122	0.0114	0.6999	0.6696	0.1035
ANN - 77	LOGSIG	1	20	0.0065	0.0118	0.6869	0.4772	0.1261
ANN - 78	TANSIG	3	20	0.0099	0.0099	0.7386	0.5322	0.1211

ANN - 79	TANSIG	1	20	0.0127	0.0114	0.6991	0.5237	0.1276
ANN - 80	LOGSIG	1	20	0.0138	0.0120	0.6848	0.3430	0.1687
ANN - 81	LOGSIG	1	20	0.0204	0.0125	0.6798	0.3618	0.1533
ANN - 82	LOGSIG	1	20	0.0082	0.0119	0.6876	0.4583	0.1385
ANN - 83	TANSIG	2	20	0.0239	0.0124	0.6727	0.3108	0.1631
ANN - 84	LOGSIG	1	10	0.0062	0.0122	0.6826	0.4830	0.1300
ANN - 85	LOGSIG	1	10	0.0090	0.0147	0.6132	0.5088	0.1240
ANN - 86	LOGSIG	1	30	0.0216	0.0140	0.6282	0.4743	0.1281
ANN - 87	LOGSIG	3	10	0.0114	0.0159	0.5806	0.2967	0.1491
ANN - 88	LOGSIG	1	20	0.0112	0.0128	0.6672	0.5197	0.1272
ANN - 89	LOGSIG	1	10	0.0076	0.0131	0.6598	0.3276	0.1532

4.1.5 CFNN models with five inputs

For the ANN modelling of five inputs used in this study, 7 models was developed. The result of the model for the five input combinations is shown in Table 4.5. The models were evaluated using the performance criteria of MSE, RMSE and R-squared values. Different hidden layers from 1-3 and training fuctions was modelled for the traning models. Based on the testing results, ANN- 92 and ANN-91 gave the highest R-sqaured values of 0.7059 and 0.5446 respectively. Also both models gave the least MSE values of 0.0930 and 0.1194. The input combination for the best prediction accuracy in the five input combination was $T_{min}, T_{avg}, WS, GSR\ and\ NM.$ The result also showed that the worst prediction was gotten from ANN-90.

Table 4.5: Evaluation of the five-input model and statistical performance (CFNN)

Model	Transfer function	Hidden layer	Neurons	Epochs	MSE (training)	R^2 (Training)	R^2 (Testing)	RMSE (Testing)
ANN - 90	TANSIG	1	10	0.0105	0.0080	0.7905	0.4436	0.1434

ANN - 91	LOGSIG	1	20	0.0095	0.0113	0.7009	0.5446	0.1194
ANN - 92	TANSIG	2	10	0.0119	0.0106	0.7188	0.7059	0.0930
ANN - 93	TANSIG	3	20	0.0105	0.0117	0.6947	0.4908	0.1217
ANN - 94	LOGSIG	1	20	0.0189	0.0139	0.6343	0.4841	0.1299
ANN - 95	TANSIG	1	10	0.0094	0.0137	0.6390	0.4934	0.1300
ANN - 96	TANSIG	2	10	0.0092	0.0136	0.6402	0.4762	0.1254

4.1.6 CFNN models with six inputs

Two models was developed for the six input combinations for rainfall prediction. This is shown in Table 4.6. The accuracy of the ANN-97 and 98 models is evaluated using the RMSE, MSE and R-squared values. The most accurate model is based on the testing results, and the Table 6 shows that the ANN-98 model gives the most accurate prediction with an R-sqaured value of 0.6744 and the lower RMSE value of 0.1023. The result concludes that the combination of $T_{min}, T_{avg}, WS, GSR\ SD, and\ NM$ gives the best prediction as compared to $T_{min}, T_{max}, WS, GSR\ SD, and\ NM$. This can also be interpreted to mean that the input variable of T_{avg} is a more important input variable as compared to T_{max}.

Table 4.6: Evaluation of the six-input model and statistical performance (CFNN)

Model	**Transfer function**	**Hidden layer**	**Neurons**	**Epochs**	**MSE (training)**	**R^2 (Training)**	**R^2 (Testing)**	**RMSE (Testing)**
ANN - 97	LOGSIG	1	20	0.0107	0.0105	0.7235	0.5603	0.1166
ANN - 98	TANSIG	3	10	0.0115	0.0111	0.7068	0.6744	0.1023

4.1.7 CFNN models with seven inputs

All the input parameters was used in the moddling of the 7 input variables for rainfall prediction. The result is shown in Table 4.7. The result shows a good prediction of R-squared value of 0.5848 and RMSE value of 0.1148 in the testing result.

Table 4.7: Evaluation of the seven-input model and statistical performance (CFNN)

Model	Transfer function	Hidden layer	Neurons	Epochs	MSE (training)	R^2 (Training)	R^2 (Testing)	RMSE (Testing)
ANN - 99	TANSIG	2	10	0.0193	0.0096	0.7470	0.5848	0.1148

4.1.8 Observation from CFNN developed model

99 models was developed in this study, using one to seven combinations of input variables for rainfall prediction. The Table 4.8 shows the performance criteria which is the R-squared and RMSE for the testing result of each model. The Table 4.8 shows that the best prediction is seen for ANN-92 which has the highest R-squared value of **0.7059** and least RMSE value of 0.0930. The selected optimum architecture is shown in bold font in the R-squared and RMSE column in Table 4.8.

Table 4.8: R^2 value and RMSE value for the optimum ANN architecture of 99 models developed (in bold) (CFNN)

Input number	Model	R^2	RMSE
	ANN - 1	0.1570	0.1772
	ANN - 2	0.3915	0.1384
	ANN - 3	0.3825	0.1436
1	ANN - 4	0.0038	0.2158
	ANN - 5	0.3240	0.1484

	ANN - 6	0.1618	0.2047
	ANN - 7	0.0862	0.2190
	ANN - 8	0.6696	0.1035
	ANN - 9	0.5941	0.1182
	ANN - 10	0.2226	0.1673
	ANN - 11	0.4563	0.1360
	ANN - 12	0.4897	0.1299
	ANN - 13	0.4190	0.1414
	ANN - 14	0.7059	0.0965
	ANN - 15	0.4825	0.1349
	ANN - 16	0.6285	0.1098
	ANN - 17	0.5506	0.1224
2	ANN - 18	0.4731	0.1265
	ANN - 19	0.4007	0.1403
	ANN - 20	0.4356	0.1315
	ANN - 21	0.4388	0.1337
	ANN - 22	0.5451	0.1213
	ANN - 23	0.3876	0.1457
	ANN - 24	0.4536	0.1300
	ANN - 25	0.4199	0.1381
	ANN - 26	0.0170	0.1824
	ANN - 27	0.3893	0.1537
	ANN - 28	0.4563	0.1276
	ANN - 29	0.3211	0.1796
	ANN - 30	0.6371	0.1093
	ANN - 31	0.5940	0.1123

	ANN - 32	0.6511	0.1037
	ANN - 33	0.6833	0.1185
	ANN - 34	0.6463	0.1103
	ANN - 35	0.5884	0.1150
	ANN - 36	0.6352	0.1113
	ANN - 37	0.6563	0.1053
	ANN - 38	0.5840	0.1138
	ANN - 39	0.6382	0.1059
	ANN - 40	0.5917	0.1276
3	ANN - 41	0.4593	0.1361
	ANN - 42	0.3223	0.1547
	ANN - 43	0.2733	0.1591
	ANN - 44	0.5819	0.1225
	ANN - 45	0.4733	0.1579
	ANN - 46	0.3022	0.1882
	ANN - 47	0.4351	0.1595
	ANN - 48	0.2755	0.1526
	ANN - 49	0.5518	0.1162
	ANN - 50	0.4563	0.1369
	ANN - 51	0.4647	0.1381
	ANN - 52	0.4432	0.1393
	ANN - 53	0.4619	0.1324
	ANN - 54	0.2094	0.1725
	ANN - 55	0.5938	0.1145
	ANN - 56	0.5712	0.1201
	ANN - 57	0.4650	0.1413
	ANN - 58	0.4746	0.1278
	ANN - 59	0.4807	0.1386

	ANN - 60	0.2548	0.1845
	ANN - 61	0.4193	0.1377
	ANN - 62	0.5074	0.1230
	ANN - 63	0.5148	0.1287
	ANN - 64	0.6107	0.1109
	ANN - 65	0.5345	0.1254
	ANN - 66	0.6295	0.1048
	ANN - 67	0.6005	0.1204
	ANN - 68	0.6219	0.1167
	ANN - 69	0.1570	0.1772
	ANN - 70	0.3915	0.1384
	ANN - 71	0.3825	0.1436
	ANN - 72	0.0038	0.2158
	ANN - 73	0.3240	0.1484
	ANN - 74	0.1618	0.2047
4	ANN - 75	0.0862	0.2190
	ANN - 76	0.6696	0.1035
	ANN - 77	0.4772	0.1261
	ANN - 78	0.5322	0.1211
	ANN - 79	0.5237	0.1276
	ANN - 80	0.3430	0.1687
	ANN - 81	0.3618	0.1533
	ANN - 82	0.4583	0.1385
	ANN - 83	0.3108	0.1631
	ANN - 84	0.4830	0.1300
	ANN - 85	0.5088	0.1240
	ANN - 86	0.4743	0.1281

	ANN - 87	0.2967	0.1491
	ANN - 88	0.5197	0.1272
	ANN - 89	0.3276	0.1532
	ANN - 90	0.4436	0.1434
	ANN - 91	0.5446	0.1194
	ANN - 92	**0.7059**	**0.0930**
5	ANN - 93	0.4908	0.1217
	ANN - 94	0.4841	0.1299
	ANN - 95	0.4934	0.1300
	ANN - 96	0.4762	0.1254
6	ANN - 97	0.5603	0.1166
	ANN - 98	0.6744	0.1023
7	ANN - 99	0.5848	0.1148

Table 4.9: Ranking of the ANN models in terms of the R^2 and RMSE values of the testing results (CFNN)

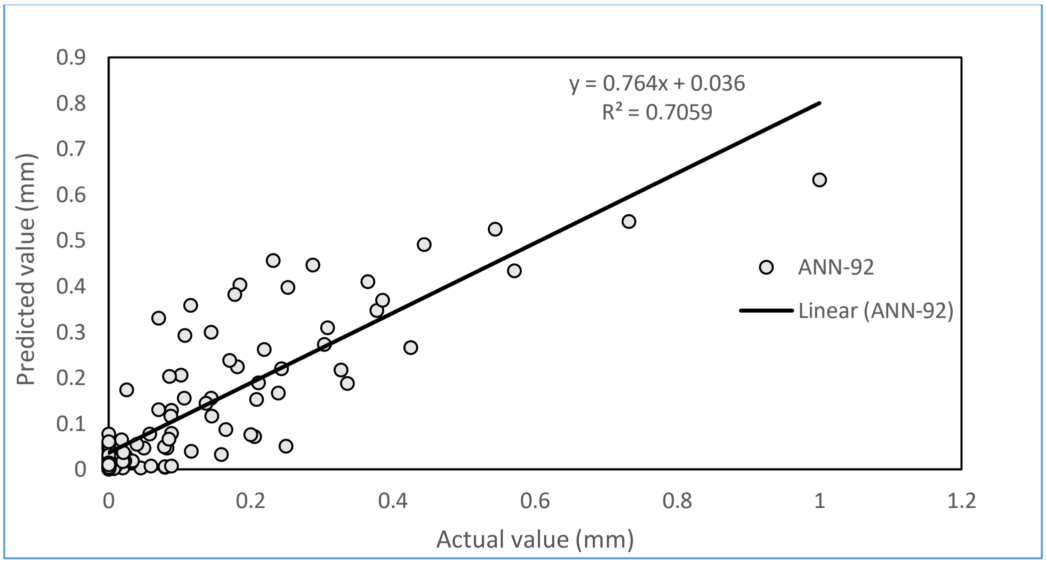
Statistics		**Rank of ANN models**								
	Rank	1	2	3	4	5	6	7	8	9
	Model	ANN - 92	ANN - 14	ANN - 33	ANN - 98	ANN - 8	ANN - 76	ANN - 37	ANN - 32	ANN - 34
R^2	Rank	10	11	12	13	14	15	16	17	18
	Model	ANN - 39	ANN - 30	ANN - 36	ANN - 66	ANN - 16	ANN - 68	ANN - 64	ANN - 67	ANN - 9
	Rank	19	20	21	22	23	24	25	26	27

	Model	ANN - 31	ANN - 55	ANN - 40	ANN - 35	ANN - 99	ANN - 38	ANN - 44	ANN - 56	ANN - 97
	Rank	28	29	30	31	32	33	34	35	36
	Model	ANN - 49	ANN - 17	ANN - 22	ANN - 91	ANN - 65	ANN - 78	ANN - 79	ANN - 88	ANN - 63
	Rank	37	38	39	40	41	42	43	44	45
	Model	ANN - 85	ANN - 62	ANN - 95	ANN - 93	ANN - 12	ANN - 94	ANN - 84	ANN - 15	ANN - 59
	Rank	46	47	48	49	50	51	52	53	54
	Model	ANN - 77	ANN - 96	ANN - 58	ANN - 86	ANN - 45	ANN - 18	ANN - 57	ANN - 51	ANN - 53
	Rank	55	56	57	58	59	60	61	62	63
	Model	ANN - 41	ANN - 82	ANN - 11	ANN - 50	ANN - 28	ANN - 24	ANN - 90	ANN - 52	ANN - 21
	Rank	64	65	66	67	68	69	70	71	72
	Model	ANN - 20	ANN - 47	ANN - 25	ANN - 61	ANN - 13	ANN - 19	ANN - 2	ANN - 70	ANN - 27
	Rank	73	74	75	76	77	78	79	80	81
	Model	ANN - 23	ANN - 3	ANN - 71	ANN - 81	ANN - 80	ANN - 89	ANN - 5	ANN - 73	ANN - 42
	Rank	82	83	84	85	86	87	88	89	90
	Model	ANN - 29	ANN - 83	ANN - 46	ANN - 87	ANN - 48	ANN - 43	ANN - 60	ANN - 10	ANN - 54
	Rank	91	92	93	94	95	96	97	98	99
	Model	ANN - 6	ANN - 74	ANN - 1	ANN - 69	ANN - 7	ANN - 75	ANN - 26	ANN - 4	ANN - 72

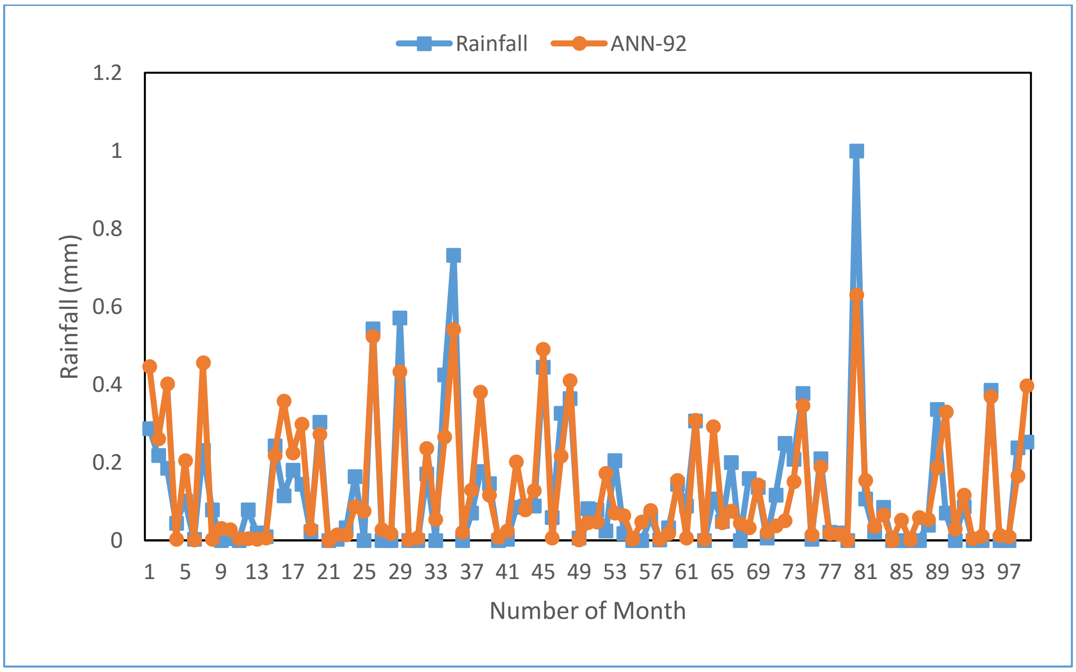
Statistics		**Rank of ANN models**								
RMSE	Rank	1	2	3	4	5	6	7	8	9
	Model	ANN - 92	ANN - 14	ANN - 98	ANN - 8	ANN - 76	ANN - 32	ANN - 66	ANN - 37	ANN - 39

Rank	10	11	12	13	14	15	16	17	18
Model	ANN - 30	ANN - 16	ANN - 34	ANN - 64	ANN - 36	ANN - 31	ANN - 38	ANN - 55	ANN - 99
Rank	19	20	21	22	23	24	25	26	27
Model	ANN - 35	ANN - 49	ANN - 97	ANN - 68	ANN - 9	ANN - 33	ANN - 91	ANN - 56	ANN - 67
Rank	28	29	30	31	32	33	34	35	36
Model	ANN - 78	ANN - 22	ANN - 93	ANN - 17	ANN - 44	ANN - 62	ANN - 85	ANN - 96	ANN - 65
Rank	37	38	39	40	41	42	43	44	45
Model	ANN - 77	ANN - 18	ANN - 88	ANN - 40	ANN - 79	ANN - 28	ANN - 58	ANN - 86	ANN - 63
Rank	46	47	48	49	50	51	52	53	54
Model	ANN - 12	ANN - 94	ANN - 24	ANN - 95	ANN - 84	ANN - 20	ANN - 53	ANN - 21	ANN - 15
Rank	55	56	57	58	59	60	61	62	63
Model	ANN - 11	ANN - 41	ANN - 50	ANN - 61	ANN - 51	ANN - 25	ANN - 2	ANN - 70	ANN - 82
Rank	64	65	66	67	68	69	70	71	72
Model	ANN - 59	ANN - 52	ANN - 19	ANN - 57	ANN - 13	ANN - 90	ANN - 3	ANN - 71	ANN - 23
Rank	73	74	75	76	77	78	79	80	81
Model	ANN - 5	ANN - 73	ANN - 87	ANN - 48	ANN - 89	ANN - 81	ANN - 27	ANN - 42	ANN - 45
Rank	82	83	84	85	86	87	88	89	90
Model	ANN - 43	ANN - 47	ANN - 83	ANN - 10	ANN - 80	ANN - 54	ANN - 1	ANN - 69	ANN - 29
Rank	91	92	93	94	95	96	97	98	99
Model	ANN - 26	ANN - 60	ANN - 46	ANN - 6	ANN - 74	ANN - 4	ANN - 72	ANN - 7	ANN - 75

The best prediction of rainfall using the input variables retrieved in this study is checked by using the R-squared value and the RMSE value of the testing results. The Table 4.9 shows that the best model that gave the most accurate prediction based on its highest R-squared value and lowest RMSE value is ANN-92, which has the input combination of $T_{min}, T_{avg,} WS, GSR\ and\ NM$. The next best prediction is ANN-14, as it shows the second highest and lowes R-squared and RMSE values respectively. The input combination of ANN-14 model is T_{max}, T_{avg} . This shows that accurate prediction is not solely dependent on many input parameters, however, the right input combination can yield accurate rainfall predictions as depicted in the performance of the two input combination model of ANN-14.



(a)



(b)

Figure 4.1 : Predicted vs Experimental values for ANN-92 algorithm

4.2 RBFNN Models

The RBFNN model is designed with spread constant ranging from 10 to 0.1 as shown in Table 10. The performance of the model is checked using the statistical performance criteria values of the testing results. Section 3.4.1 to section 3.4.7 shows the output of the model of ANN-1 to ANN-99 based on number and uniform input combinations. The input variables are combined based on one input to seven input combinations.

4.2.1 RBFNN models with one input

The Table 4.10 shows the result of one input of the RBFNN model. The statistical performance criteria values value shows that the ANN-6 gives the best accuracy as it has the highest R-squares value. The spread constant which gives the highest accuracy is terms of lowest RMSE value is 0.1

Table 4.10: Evaluation of the three-input model and statistical performance of RBFNN model

	Spread constant	**10**	**5**	**1**	**0.5**	**0.1**	**0.05**
ANN-1	R^2	0.3539	0.3539	0.3650	0.3520	0.1184	-
	RMSE	0.1407	0.1407	0.1407	0.1420	0.1857	-
ANN-2	R^2	0.4951	0.4951	0.4414	0.4600	0.2936	-
	RMSE	0.1232	0.1232	0.1301	0.1280	0.1487	-
ANN-3	R^2	0.4128	0.4128	0.4202	0.4043	0.4043	-
	RMSE	0.1338	0.1338	0.1331	0.1350	0.1350	-
ANN-4	R^2	0.0934	0.0932	0.0688	0.0527	-	-
	RMSE	0.1646	0.1646	0.1669	0.1703	-	-
ANN-5	R^2	0.404369	0.404369	0.416283	0.424583	-	-
	RMSE	0.1368	0.1368	0.1344	0.1333	-	-
ANN-6	R^2	0.5356	0.5358	0.5459	0.5584	0.5542	-
	RMSE	0.1190	0.1190	0.1186	0.1193	**0.1176**	-
ANN-7	R^2	0.0003	0.0003	0.0046	0.0125	0.0021	0.0081

RMSE	0.1734	0.1734	0.1727	0.1719	0.1803	0.1901

4.2.2 RBFNN models with two input

The performance of two input variables is shown in Table 4.11. 22 models was developed with spread constant ranging from 20 to 0.5 used based on trial and errors to get the most accurate prediction. The result in Table 4.11 shows that the best model based on highest R-squared value is ANN-8 using spread constant of 10. This is the input combination ofT_{min}, T_{max}. Also, the least RMSE value was obtained for ANN-14 (T_{max}, T_{avg}).

Table 4.11: Evaluation of the two-input model and statistical performance of RBFNN model

Model	**Spread constant**	**20**	**15**	**10**	**5**	**1**	**0.5**
ANN-8	R^2	-	-	0.6905	0.6709	0.4780	-
	RMSE	-	-	0.0972	0.1004	0.1410	-
ANN-9	R^2	-	0.6443	0.6443	0.6497	0.5763	0.5148
	RMSE	-	0.1064	0.1064	0.1058	0.1192	0.1286
ANN-10	R^2	0.1216	0.1216	0.1314	0.1244	0.0396	-
	RMSE	0.1411	0.1411	0.1406	0.1419	0.1654	-
ANN-11	R2	0.4160	0.4160	0.4209	0.4373	0.4716	-
	RMSE	0.1384	0.1384	0.1372	0.1367	0.1304	-
ANN-12	R2	0.5379	0.5408	0.5399	0.5388	0.5504	-
	RMSE	0.1195	0.1192	0.1193	0.1197	0.1209	-
ANN-13	R2	0.3557	0.3557	0.3670	0.3648	0.3958	-
	RMSE	0.1401	0.1401	0.1389	0.1397	0.1389	-
ANN-14	R2	0.6806	0.6846	0.6729	0.6567	0.5838	-
	RMSE	0.0976	0.0970	0.0985	0.1007	0.1117	-
ANN-15	R2	0.1341	0.1316	0.1316	0.1316	0.1316	-
	RMSE	0.4078	0.4078	0.4078	0.4078	0.4078	-
ANN-16	R2	-	0.6443	0.6443	0.6498	0.5764	0.5148
	RMSE	-	0.1064	0.1064	0.1058	0.1192	0.1286

ANN-17	R2	0.5703	0.5700	0.5676	0.5654	0.5095	-
	RMSE	0.1144	0.1147	0.1149	0.1150	0.1219	-
ANN-18	R2	0.5024	0.5024	0.4979	0.4805	0.4073	-
	RMSE	0.1250	0.1250	0.1254	0.1270	0.1376	-
ANN-19	R2	0.4997	0.4997	0.5214	0.5235	0.3839	-
	RMSE	0.1228	0.1228	0.1192	0.1190	0.1377	-
ANN-20	R2	0.4027	0.4027	0.4091	0.4224	0.1410	-
	RMSE	0.1349	0.1349	0.1346	0.1332	0.1846	-
ANN-21	R2	0.4282	0.4282	0.4349	0.4416	0.4453	-
	RMSE	0.1354	0.1354	0.1341	0.1341	0.1328	-
ANN-22	R2	0.5405	0.5405	0.5418	0.5363	0.5433	-
	RMSE	0.1185	0.1185	0.1188	0.1206	0.1202	-
ANN-23	R2	0.4267	0.4267	0.4373	0.4390	0.4551	-
	RMSE	0.1318	0.1318	0.1303	0.1300	0.1282	-
ANN-24	R2	0.4276	0.4276	0.4261	0.4234		-
	RMSE	0.1339	0.1339	0.1340	0.1343	-	-
ANN-25	R2	0.5342	0.5342	0.5306	0.5259	-	-
	RMSE	0.1199	0.1199	0.1201	0.1206	-	-
ANN-26	R2	0.2220	0.2220	0.1452	0.0255	-	-
	RMSE	0.1534	0.1534	0.1591	0.1726	-	-
ANN-27	R2	0.5301	0.5317	0.5317	0.5287	0.5079	-
	RMSE	0.1199	0.1199	0.1199	0.1206	0.1320	-
ANN-28	R2	0.4034	0.4034	0.4132	0.4151	0.4609	-
	RMSE	0.1375	0.1375	0.1364	0.1365	0.1328	-
ANN-29	R2	0.5506	0.5506	0.5646	0.5630	0.5007	-
	RMSE	0.1171	0.1171	0.1154	0.1150	0.1249	

4.2.3 RBFNN models with three input

The performance of three input variables is shown in Table 12. 33 models was developed with spread constant ranging from 25 to 1 used based on trial and errors to get the most accurate prediction. The result in Table 4.12 shows that the best model based on highest R-squared value is ANN-33 using spread constant of 10. This is the input combination ofT_{min}, T_{max} ,NM. Also, the least RMSE value was obtained for ANN- 40 with spread constant of 15 (T_{max}, T_{avg}, NM).

Table 4.12: Evaluation of the three-input model and statistical performance of RBFNN model

Model	Spread constant	25	20	15	10	5	1
ANN30	R^2	-	-	-	0.6254	0.6293	0.1897
	RMSE	-	-	-	0.1071	0.1087	0.1845
ANN31	R^2	-	-	-	0.6977	0.6504	0.1614
	RMSE	-	-	-	0.0965	0.1034	0.2020
ANN32	R^2	-	-	-	0.6831	0.6610	0.2465
	RMSE	-	-	-	0.1004	0.1038	0.1987
ANN33	R^2	-	-	-	0.6997	0.6747	-
	RMSE	-	-	-	0.0950	0.1005	-
ANN34	R^2	-	0.6480	0.6483	0.6483	0.6373	0.5349
	RMSE	-	0.1062	0.1061	0.1060	0.1080	0.1285
ANN35	R^2	-	0.6277	0.6603	0.6537	0.6749	-
	RMSE	-	0.1088	0.1046	0.1060	0.1067	-

ANN36	R^2	-	0.6176	0.6238	0.6290	0.5960	-
	RMSE	-	0.1120	0.1109	0.1101	0.1160	-
ANN37	R^2	-	0.6368	0.6336	0.6265	0.6445	-
	RMSE	-	0.1077	0.1082	0.1096	0.1058	-
ANN38	R^2	-	0.6282	0.6134	0.6087	0.5349	-
	RMSE	-	0.1054	0.1070	0.1077	0.1179	-
ANN39	R^2	-	0.6527	0.6522	0.6535	0.6205	0.3492
	RMSE	-	0.1034	0.1036	0.1035	0.1085	0.1866
ANN40	R^2	-	0.6907	0.6951	0.6750	0.6571	-
	RMSE	-	0.0959	0.0950	0.0978	0.1004	-
ANN41	R^2	-	0.5361	0.5360	0.5213	0.5328	-
	RMSE	-	0.1193	0.1189	0.1211	0.1206	-
ANN42	R^2	-	0.4500	0.4500	0.4563	0.4594	-
	RMSE	-	0.1323	0.1323	0.1315	0.1301	-
ANN43	R^2	-	0.4371	0.4369	0.2416	0.1466	-
	RMSE	-	0.1307	0.1307	0.1542	0.1704	-
ANN44	R^2	-	0.5380	0.5446	0.5643	0.5418	-
	RMSE	-	0.1197	0.1184	0.1153	0.1182	-
ANN45	R^2	-	0.5285	0.5279	0.5335	0.5112	-
	RMSE	-	0.1211	0.1213	0.1204	0.1256	-
ANN46	R^2	-	0.4225	0.4377	0.4349	0.4268	-
	RMSE	-	0.1380	0.1350	0.1359	0.1403	-
ANN47	R^2	-	0.5386	0.5429	0.5421	0.5380	-

	RMSE	-	0.1193	0.1188	0.1188	0.1209	-
ANN48	R^2	-	0.3760	0.3760	0.2819	0.1578	-
	RMSE	-	0.1382	0.1382	0.1493	0.1688	-
ANN49	R^2	-	0.5442	0.5455	0.5637	0.5341	-
	RMSE	-	0.1188	0.1186	0.1156	0.1199	-
ANN50	R^2	-	0.5364	0.5333	0.5197	0.0042	0.3654
	RMSE	-	0.1202	0.1209	0.1230	0.2262	0.1670
ANN51	R^2	-	0.4158	0.4206	0.4184	0.4275	-
	RMSE	-	0.1394	0.1384	0.1396	0.1429	-
ANN52	R^2	0.0055	0.0289	0.0283	0.2095	-	-
	RMSE	0.2773	0.2796	0.2793	0.2088	-	-
ANN53	R^2	-	0.5636	0.5592	0.5309	0.5168	-
	RMSE	-	0.1155	0.1162	0.1188	0.1203	-
ANN54	R^2	-	0.4816	0.4816	0.3617	0.1906	-
	RMSE	-	0.1251	0.1251	0.1388	0.1636	-
ANN55	R^2	-	0.5628	0.5652	0.5740	0.5515	-
	RMSE	-	0.1165	0.1163	0.1146	0.1185	-
ANN56	R^2	-	0.5550	0.5761	0.5657	0.5848	-
	RMSE	-	0.1172	0.1143	0.1155	0.1125	-
ANN57	R^2	-	0.4816	0.4859	0.5041	0.4894	-
	RMSE	-	0.1297	0.1291	0.1250	0.1276	-
ANN58	R^2	0.5082	0.5037	0.5040	0.4650	0.4666	-
	RMSE	0.1233	0.1249	0.1248	0.1289	0.1272	-

ANN59	R^2	0.5432	0.5426	0.5376	0.5235	0.4323	-
	RMSE	0.1185	0.1192	0.1202	0.1219	0.1409	-
ANN60	R^2	0.4124	0.4124	0.4123	0.2924	0.1878	-
	RMSE	0.1363	0.1364	0.1364	0.1522	0.1731	-
ANN61	R^2	0.5473	0.5472	0.5472	0.4149	0.3861	-
	RMSE	0.1175	0.1176	0.1176	0.1359	0.1390	-
ANN62	R^2	0.5546	0.5576	0.5670	0.5723	0.5586	-
	RMSE	0.1160	0.1155	0.1142	0.1134	0.1150	-

4.2.4 RBFNN models with four input

The performance of three input variables is shown in Table 12. 27 models was developed with spread constant ranging from 25 to 5 used based on trial and errors to get the most accurate prediction. The result in Table 4.13 shows that the best model based on highest R-squared value and least RMSE value is ANN-68 using spread constant of 25. This is the input combination ofT_{min}, T_{max} , $SD, NM.$

Table 4.13: Evaluation of the four-input model and statistical performance of RBFNN model

Model	**Spread constant**	**25**	**20**	**15**	**10**	**5**
ANN63	R^2	-	0.6610	0.6559	0.6059	0.3005
	RMSE	-	0.1034	0.1040	0.1112	0.1729
ANN64	R^2	0.6767	0.6785	0.6848	0.6338	0.5737
	RMSE	0.1011	0.1012	0.0999	0.1097	0.1185

ANN65	R^2	-	0.6045	0.5834	0.4416	0.1884
	RMSE	-	0.1101	0.1128	0.1322	0.1860
ANN66	R^2	0.6767	0.6785	0.6848	0.6338	0.5737
	RMSE	0.1011	0.1012	0.0999	0.1097	0.1185
ANN67	R^2	0.6719	0.6866	0.6848	0.6834	0.5643
	RMSE	0.1028	0.1015	0.1010	0.1008	0.1245
ANN68	R^2	0.7140	0.6904	0.6729	0.6688	0.5697
	RMSE	0.0934	0.0976	0.1005	0.1014	0.1181
ANN69	R^2	0.5994	0.5929	0.6181	0.5718	0.5706
	RMSE	0.1132	0.1141	0.1116	0.1177	0.1194
ANN70	R^2	0.5991	0.5935	0.6092	0.5977	0.4087
	RMSE	0.1151	0.1159	0.1141	0.1148	0.1429
ANN71	R^2	0.5932	0.6477	0.5875	0.5449	0.4242
	RMSE	0.1137	0.1057	0.1149	0.1210	0.1396
ANN72	R^2	0.6287	0.6287	0.6269	0.6650	0.6095
	RMSE	0.1110	0.1112	0.1116	0.1045	0.1174
ANN73	R^2	0.6483	0.6556	0.6490	0.6222	0.5723
	RMSE	0.1057	0.1046	0.1065	0.1113	0.1191
ANN74	R^2	0.6394	0.6512	0.6410	0.6410	0.5960
	RMSE	0.1066	0.1053	0.1072	0.1072	0.1127
ANN75	R^2	0.6056	0.5851	0.5935	0.5791	0.4851
	RMSE	0.1099	0.1128	0.1113	0.1136	0.1282
ANN76	R^2	0.6632	0.6514	0.6432	0.6485	0.5115

	RMSE	0.1011	0.1048	0.1063	0.1050	0.1293
ANN77	R^2	0.6216	0.6241	0.5842	0.3669	0.1104
	RMSE	0.1058	0.1054	0.1103	0.1399	0.1960
ANN78	R^2	0.6902	0.6777	0.6501	0.6424	0.4879
	RMSE	0.0958	0.0980	0.1028	0.1045	0.1282
ANN79	R^2	0.6598	0.6587	0.6456	0.6293	0.5610
	RMSE	0.1022	0.1024	0.1050	0.1075	0.1178
ANN80	R^2	0.5282	0.5320	0.5249	0.5131	0.4840
	RMSE	0.1219	0.1217	0.1231	0.1257	0.1358
ANN81	R^2	0.4001	0.3808	0.3752	0.2697	0.3631
	RMSE	0.1408	0.1460	0.1463	0.1635	0.1579
ANN82	R^2	0.5503	0.5429	0.5408	0.5458	0.4987
	RMSE	0.1175	0.1195	0.1200	0.1198	0.1281
ANN83	R^2	0.4371	0.4372	0.4646	0.2756	0.3505
	RMSE	0.1350	0.1349	0.1317	0.1565	0.1517
ANN84	R^2	0.5467	0.5347	0.5657	0.5797	0.5424
	RMSE	0.1195	0.1227	0.1173	0.1153	0.1203
ANN85	R^2	0.5336	0.5194	0.5177	0.5038	0.4733
	RMSE	0.1197	0.1223	0.1233	0.1250	0.1335
ANN86	R^2	0.5522	0.5580	0.5364	0.5226	0.4559
	RMSE	0.1176	0.1171	0.1209	0.1218	0.1366
ANN87	R^2	0.4160	0.3886	0.4028	0.2692	0.3054
	RMSE	0.1372	0.1432	0.1406	0.1607	0.1627

ANN88	R^2	0.5304	0.5311	0.5429	0.5501	0.4973
	RMSE	0.1216	0.1215	0.1194	0.1186	0.1288
ANN89	R^2	0.5272	0.5209	0.5211	0.4050	0.3241
	RMSE	0.1205	0.1212	0.1212	0.1384	0.1564

4.2.5 RBFNN models with five inputs

The performance of three input variables is shown in Table 13. 7 models was developed with spread constant ranging from 25 to 5 used based on trial and errors to get the most accurate prediction. The result in Table 4.14 shows that the best model based on highest R-squared value and least RMSE value is ANN-90 and ANN-92 using spread constant of 25. This is the input combination of $T_{min},\ T_{max}\ ,WS\ \ GSR,\ SD$and $T_{min},\ T_{svg}\ ,WS\ \ GSR, NM$ respectively.

Table 4.14: Evaluation of the five-input model and statistical performance of RBFNN model

Model	**Spread constant**	**25**	**20**	**15**	**10**	**5**
ANN90	R^2	0.5947	0.5712	0.5636	0.5196	0.4034
	RMSE	0.1143	0.1185	0.1190	0.1285	0.1517
ANN91	R^2	0.5420	0.5271	0.4789	0.3653	-
	RMSE	0.1213	0.1248	0.1330	0.1502	-
ANN92	R^2	0.5940	0.5750	0.4996	0.4901	0.2636
	RMSE	0.1144	0.1193	0.1312	0.1321	0.1860
ANN93	R^2	0.5693	0.4833	0.4111	0.2824	0.1736
	RMSE	0.1154	0.1298	0.1406	0.1649	0.2179

ANN94	R^2	0.5018	0.5037	0.4640	0.2982	0.3234
	RMSE	0.1262	0.1261	0.1320	0.1616	0.1725
ANN95	R^2	0.4901	0.5023	0.5030	0.2868	0.2671
	RMSE	0.1270	0.1268	0.1272	0.1599	0.1829
ANN96	R^2	0.4922	0.4936	0.5069	0.2950	0.3347
	RMSE	0.1266	0.1268	0.1246	0.1575	0.1867

4.2.6 RBFNN models with six inputs

The six input combination consist of two models. How accurate the ANN-97 and ANN-98 are checked using the statistical performance criteria values of the testing result. The Table 4.15 shows that ANN-98 model developed with a spread constant of 25 gives the best prediction accuracy based on its highest R-squared value and lowest RMSE value. The result concludes that the combination of $T_{min}, T_{avg}, WS, GSR\ SD, and\ NM$ gives the best prediction as compared to $T_{min}, T_{max}, WS, GSR\ SD, and\ NM.$

Table 4.15: Evaluation of the six-input model and statistical performance of RBFNN model

Model	**Spread constant**	25	20	15	10	5
ANN-97	R^2	0.5187	0.4434	0.4473	0.3493	-
	RMSE	0.1265	0.1398	0.1435	0.1536	-
ANN-98	R^2	0.5367	0.5147	0.4452	0.2824	0.3278
	RMSE	0.1238	0.1309	0.1421	0.1713	0.1901

4.2.7 RBFNN models with seven inputs

All the input parameters are used in the testing of this model, with different trials of spread constant, to get the most accurate prediction model. The statistical result of the developed models for $T_{min}, T_{max}, T_{avg}, WS, GSR\ SD, and\ NM$ input variables is shown in Table 4.16. The R-squared value of the model is 0.7061 and the lowest RMSE value obtained was 0.1303. The spread constant of 25 was the ideal value for obtaining this performance.

Table 4.16: Evaluation of the six-input model and statistical performance of RBFNN model

Model	**Spread constant**	**25**	**20**	**15**	**10**	**5**
ANN-99	R^2	0.4986	0.4508	0.3345	0.1853	0.0626
	RMSE	0.1303	0.1389	0.1616	0.2012	0.2620

4.2.8 Observation from RBFNN developed model

In the RBFNN models, different spread constant was used based on trial and error to get the most excellent model for testing the input paramaters. 99 models was developed based on different combination of input parameters. Table 4.17 shows the optimum spread constant that gave the best R-squared and RMSE of each of the model. The best model of the 99 models is also shown to be ANN-68 with highest R squared value of 0.7140 and lowest RMSE value of 0.0933.

Table 4.17: R^2 value and RMSE value for the optimum ANN architecture of 99 models developed (in bold) (RBFNN)

Input number	**Model**	**R^2**	**RMSE**
	ANN - 1	0.3651	0.1407
	ANN - 2	0.4952	0.1232
	ANN - 3	0.4203	0.1331

1	ANN - 4	0.0935	0.1646
	ANN - 5	0.4246	0.1333
	ANN - 6	0.5585	0.1176
	ANN - 7	0.0125	0.1719
	ANN - 8	0.6906	0.0972
	ANN - 9	0.6496	0.1058
	ANN - 10	0.3624	0.1406
	ANN - 11	0.4714	0.1304
	ANN - 12	0.5504	0.1192
	ANN - 13	0.3956	0.1389
	ANN - 14	0.6844	0.0970
	ANN - 15	0.1341	0.4078
	ANN - 16	0.6496	0.1058
	ANN - 17	0.5703	0.1144
2	ANN - 18	0.5024	0.1250
	ANN - 19	0.5233	0.1190
	ANN - 20	0.4222	0.1332
	ANN - 21	0.4452	0.1328
	ANN - 22	0.5432	0.1185
	ANN - 23	0.4551	0.1282
	ANN - 24	0.4275	0.1339
	ANN - 25	0.5341	0.1199
	ANN - 26	0.2220	0.1534
	ANN - 27	0.5317	0.1199
	ANN - 28	0.4609	0.1328
	ANN - 29	0.5645	0.1150

	ANN - 30	0.6293	0.1071

	ANN - 31	0.6977	0.0965
	ANN - 32	0.6831	0.1004
	ANN - 33	0.6997	0.0950
	ANN - 34	0.6483	0.1060
	ANN - 35	0.6749	0.1046
	ANN - 36	0.6290	0.1101
	ANN - 37	0.6445	0.1058
	ANN - 38	0.6282	0.1054
	ANN - 39	0.6535	0.1034
	ANN - 40	0.6951	0.0950
3	ANN - 41	0.5361	0.1189
	ANN - 42	0.4594	0.1301
	ANN - 43	0.4371	0.1307
	ANN - 44	0.5643	0.1153
	ANN - 45	0.5335	0.1204
	ANN - 46	0.4377	0.1350
	ANN - 47	0.5429	0.1188
	ANN - 48	0.3760	0.1382
	ANN - 49	0.5637	0.1156
	ANN - 50	0.5364	0.1202
	ANN - 51	0.4275	0.1384
	ANN - 52	0.2095	0.2088
	ANN - 53	0.5636	0.1155
	ANN - 54	0.4816	0.1251
	ANN - 55	0.5740	0.1146
	ANN - 56	0.5848	0.1125
	ANN - 57	0.5041	0.1250
	ANN - 58	0.5082	0.1233

	ANN - 59	0.5432	0.1185
	ANN - 60	0.4124	0.1363
	ANN - 61	0.5473	0.1175
	ANN - 62	0.5723	0.1134
	ANN - 63	0.6610	0.1033
	ANN - 64	0.6848	0.0998
	ANN - 65	0.6044	0.1101
	ANN - 66	0.6848	0.0998
	ANN - 67	0.6864	0.1007
	ANN - 68	0.7139	**0.0933**
	ANN - 69	0.6180	0.1116
	ANN - 70	0.6092	0.1140
	ANN - 71	0.6475	0.1057
	ANN - 72	0.6650	0.1045
	ANN - 73	0.6556	0.1046
	ANN - 74	0.6512	0.1052
4	ANN - 75	0.6056	0.1099
	ANN - 76	0.6631	0.1010
	ANN - 77	0.6239	0.1054
	ANN - 78	0.6902	0.0957
	ANN - 79	0.6598	0.1022
	ANN - 80	0.5320	0.1216
	ANN - 81	0.3999	0.1407
	ANN - 82	0.5503	0.1175
	ANN - 83	0.4646	0.1317
	ANN - 84	0.5796	0.1152
	ANN - 85	0.5336	0.1197

	ANN - 86	0.5580	0.1171
	ANN - 87	0.4160	0.1372
	ANN - 88	0.5500	0.1186
	ANN - 89	0.5271	0.1205
	ANN - 90	0.5947	0.1143
	ANN - 91	0.5420	0.1213
	ANN - 92	0.5940	0.1144
5	ANN - 93	0.5693	0.1154
	ANN - 94	0.5037	0.1261
	ANN - 95	0.5030	0.1268
	ANN - 96	0.5069	0.1246
6	ANN - 97	0.5185	0.1265
	ANN - 98	0.5367	0.1238
7	ANN - 99	0.4986	0.1303

Table 4.18: Ranking of the ANN models in terms of the R^2 and RMSE values of the testing results (RBFNN)

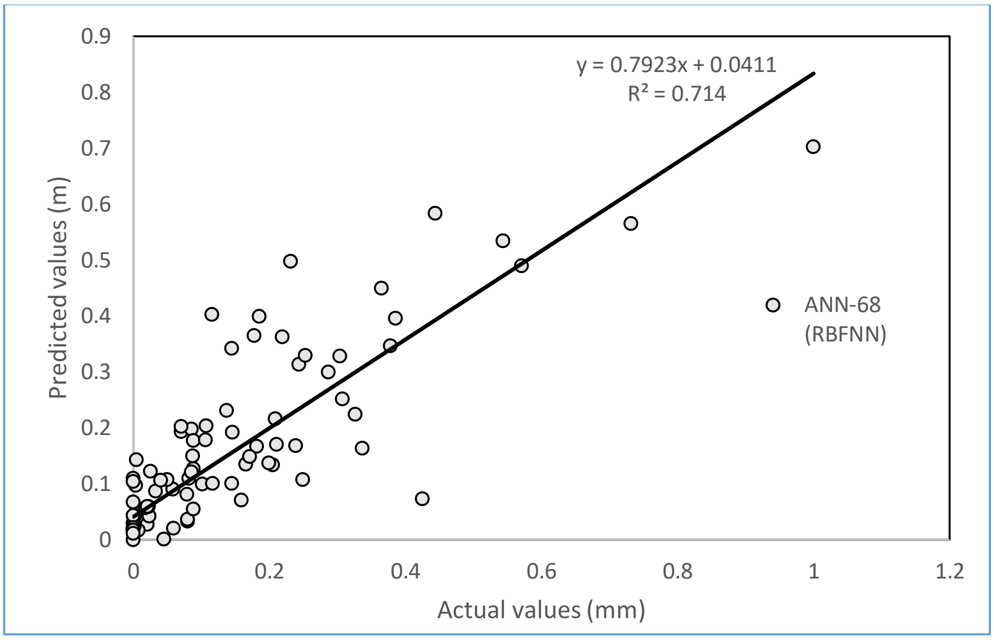
Statistics		**Rank of ANN models**								
	Rank	1	2	3	4	5	6	7	8	9
	Model	ANN - 68	ANN - 33	ANN - 31	ANN - 40	ANN - 8	ANN - 78	ANN - 67	ANN - 64	ANN - 66
R^2										
	Rank	10	11	12	13	14	15	16	17	18
	Model	ANN - 14	ANN - 32	ANN - 35	ANN - 72	ANN - 76	ANN - 63	ANN - 79	ANN - 73	ANN - 39

Rank	19	20	21	22	23	24	25	26	27
Model	ANN - 74	ANN - 9	ANN - 16	ANN - 34	ANN - 71	ANN - 37	ANN - 30	ANN - 36	ANN - 38
Rank	28	29	30	31	32	33	34	35	36
Model	ANN - 77	ANN - 69	ANN - 70	ANN - 75	ANN - 65	ANN - 90	ANN - 92	ANN - 56	ANN - 84
Rank	37	38	39	40	41	42	43	44	45
Model	ANN - 55	ANN - 62	ANN - 17	ANN - 93	ANN - 29	ANN - 44	ANN - 49	ANN - 53	ANN - 6
Rank	46	47	48	49	50	51	52	53	54
Model	ANN - 86	ANN - 12	ANN - 82	ANN - 88	ANN - 61	ANN - 22	ANN - 59	ANN - 47	ANN - 91
Rank	55	56	57	58	59	60	61	62	63
Model	ANN - 98	ANN - 50	ANN - 41	ANN - 25	ANN - 85	ANN - 45	ANN - 80	ANN - 27	ANN - 89
Rank	64	65	66	67	68	69	70	71	72
Model	ANN - 19	ANN - 97	ANN - 58	ANN - 96	ANN - 57	ANN - 94	ANN - 95	ANN - 18	ANN - 99
Rank	73	74	75	76	77	78	79	80	81
Model	ANN - 2	ANN - 54	ANN - 11	ANN - 83	ANN - 28	ANN - 42	ANN - 23	ANN - 21	ANN - 46
Rank	82	83	84	85	86	87	88	89	90
Model	ANN - 43	ANN - 24	ANN - 51	ANN - 5	ANN - 20	ANN - 3	ANN - 87	ANN - 60	ANN - 81
Rank	91	92	93	94	95	96	97	98	99
Model	ANN - 13	ANN - 48	ANN - 1	ANN - 10	ANN - 26	ANN - 52	ANN - 15	ANN - 4	ANN - 7

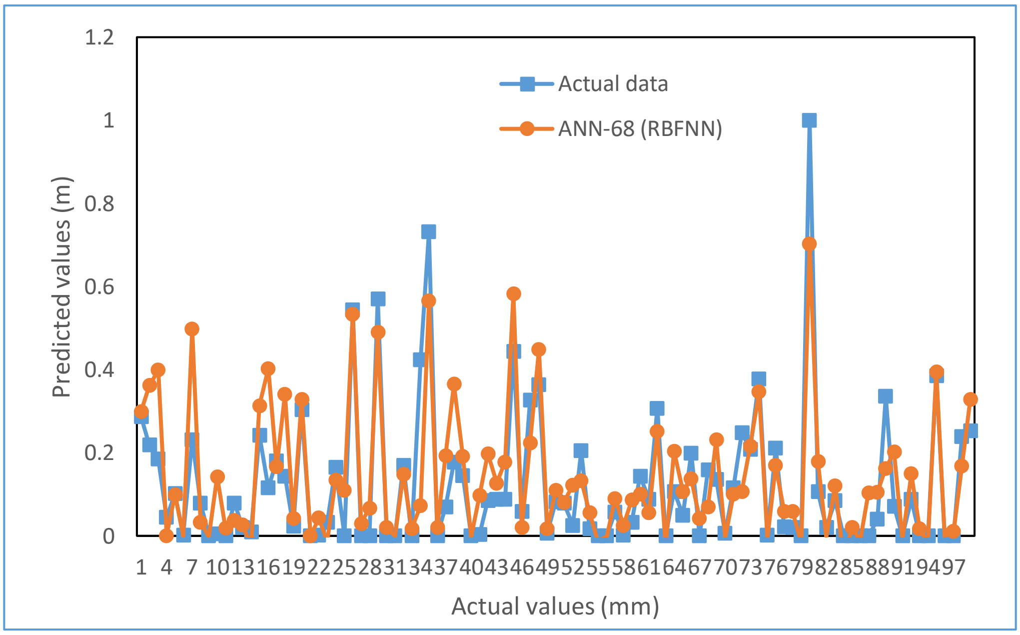
Statistics		Rank of ANN models								
RMSE	Rank	1	2	3	4	5	6	7	8	9
	Model	ANN - 68	ANN - 33	ANN - 40	ANN - 78	ANN - 31	ANN - 14	ANN - 8	ANN - 64	ANN - 66
	Rank	10	11	12	13	14	15	16	17	18
	Model	ANN - 32	ANN - 67	ANN - 76	ANN - 79	ANN - 63	ANN - 39	ANN - 72	ANN - 35	ANN - 73
	Rank	19	20	21	22	23	24	25	26	27
	Model	ANN - 74	ANN - 38	ANN - 77	ANN - 71	ANN - 9	ANN - 16	ANN - 37	ANN - 34	ANN - 30
	Rank	28	29	30	31	32	33	34	35	36
	Model	ANN - 75	ANN - 36	ANN - 65	ANN - 69	ANN - 56	ANN - 62	ANN - 70	ANN - 90	ANN - 17
	Rank	37	38	39	40	41	42	43	44	45
	Model	ANN - 92	ANN - 55	ANN - 29	ANN - 84	ANN - 44	ANN - 93	ANN - 53	ANN - 49	ANN - 86
	Rank	46	47	48	49	50	51	52	53	54
	Model	ANN - 61	ANN - 82	ANN - 6	ANN - 22	ANN - 59	ANN - 88	ANN - 47	ANN - 41	ANN - 19
	Rank	55	56	57	58	59	60	61	62	63

Model	ANN - 12	ANN - 85	ANN - 25	ANN - 27	ANN - 50	ANN - 45	ANN - 89	ANN - 91	ANN - 80
Rank	64	65	66	67	68	69	70	71	72
Model	ANN - 2	ANN - 58	ANN - 98	ANN - 96	ANN - 18	ANN - 57	ANN - 54	ANN - 94	ANN - 97
Rank	73	74	75	76	77	78	79	80	81
Model	ANN - 95	ANN - 23	ANN - 42	ANN - 99	ANN - 11	ANN - 43	ANN - 83	ANN - 21	ANN - 28
Rank	82	83	84	85	86	87	88	89	90
Model	ANN - 3	ANN - 20	ANN - 5	ANN - 24	ANN - 46	ANN - 60	ANN - 87	ANN - 48	ANN - 51
Rank	91	92	93	94	95	96	97	98	99
Model	ANN - 13	ANN - 10	ANN - 1	ANN - 81	ANN - 26	ANN - 4	ANN - 7	ANN - 52	ANN - 15

The most accurate prediction of rainfall using the input variables gotten and used in this study is checked using the R-squared value and the RMSE value of the testing results. The Table 4.18 shows that the best model that gave the most accurate prediction based on its highest R-squared value and lowest RMSE value is ANN-68, which has the input combination of T_{min}, T_{max} , SD, NM. The next best prediction is ANN-14, as it shows the second highest and lowes R-squared and RMSE values respectively. The input combination of ANN-33 model is T_{min}, T_{max} , NM.



(a)



(b)

Figure 4.2 : Predicted vs Experimental values for ANN-68 algorithm

4.3 Response Surface Methodology

The RSM result of the model 92 input variables is shown in equation ------

$$
\begin{aligned}
\mathrm{RF} = {} & 0.281032 + 6.20538 * \mathrm{T_{min}} - 6.3230 * \mathrm{T_{avg}} + 0.0770 * \mathrm{WS} + 0.233901 * \mathrm{GSR} \\
& - 0.073942 * \mathrm{NS} + 0.4122 * \mathrm{T_{min}} * \mathrm{T_{avg}} - 1.37486 * \mathrm{T_{min}} * \mathrm{WS} \\
& - 6.93449 * \mathrm{T_{min}} * \mathrm{GSR} + 0.8175 * \mathrm{T_{min}} * \mathrm{NM} + 1.3804 * \mathrm{T_{avg}} * \mathrm{WS} \\
& + 6.1034 * \mathrm{T_{avg}} * \mathrm{GSR} - 1.0155 * \mathrm{T_{avg}} * \mathrm{NM} - 0.2714 * \mathrm{WS} * \mathrm{GSR} \\
& - 0.2122 * \mathrm{WS} * \mathrm{NM} + 0.4531 * \mathrm{GSR} * \mathrm{NM}
\end{aligned}
$$

The R^2 of this correlation formula is 0.6396. The appendix A shows the interactions of the input variables. Also the Figure 4.3 shows the confidence index of each input parameter in the correlations. It is seen that the most relevant input parameter are the month, wind speed and solar radiation. The p value is < 0.0001 and the F-value is 5.43. The mean square error is 0.0781.

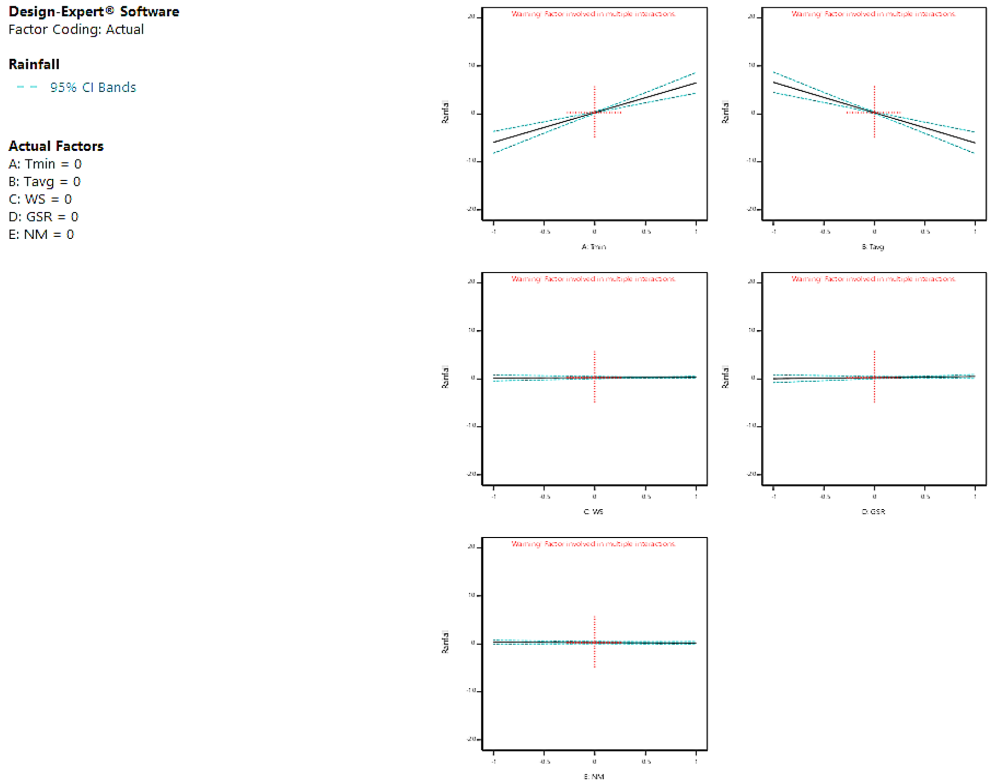


Figure 4.3 : Interalated effects of input variables

The RSM result of the model 68 input variables is shown in equation ------

$$RF = 0.2063 + 4.2462T_{min} - 4.0466T_{max} + 0.1247 * SD - 0.001652 * NM + 0.3342 * T_{min} * T_{max} - 4.9729 * T_{min} * SD - 0.3310T_{min} * NM + 4.03771 * T_{max} * SD + 0.2458 * T_{max} * NM + 0.0651 * SD * NM$$

The R^2 of this correlation formula is 0.6645. The appendix A shows the interactions of the input variables. Also the Figure 4.4 shows the confidence index of each input parameter in the correlations. It is seen that the most relevant input parameter are the month and solar radiation. The p value is < 0.0001 and the F-value is 10.46. The mean square error is 0.1377.

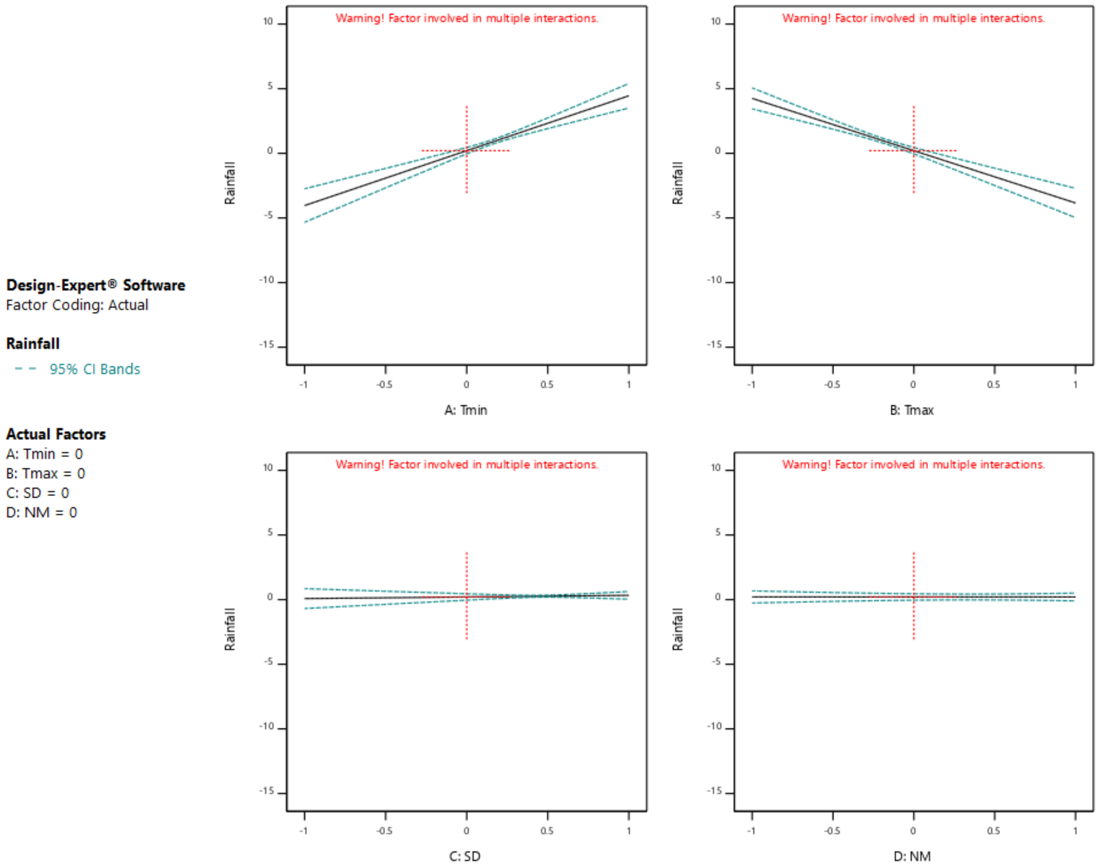
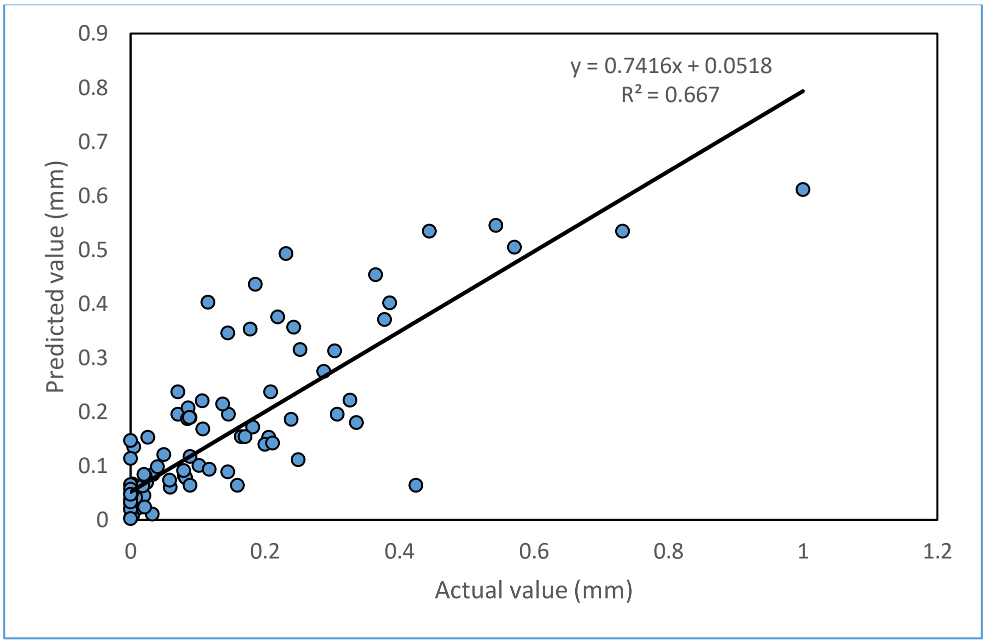
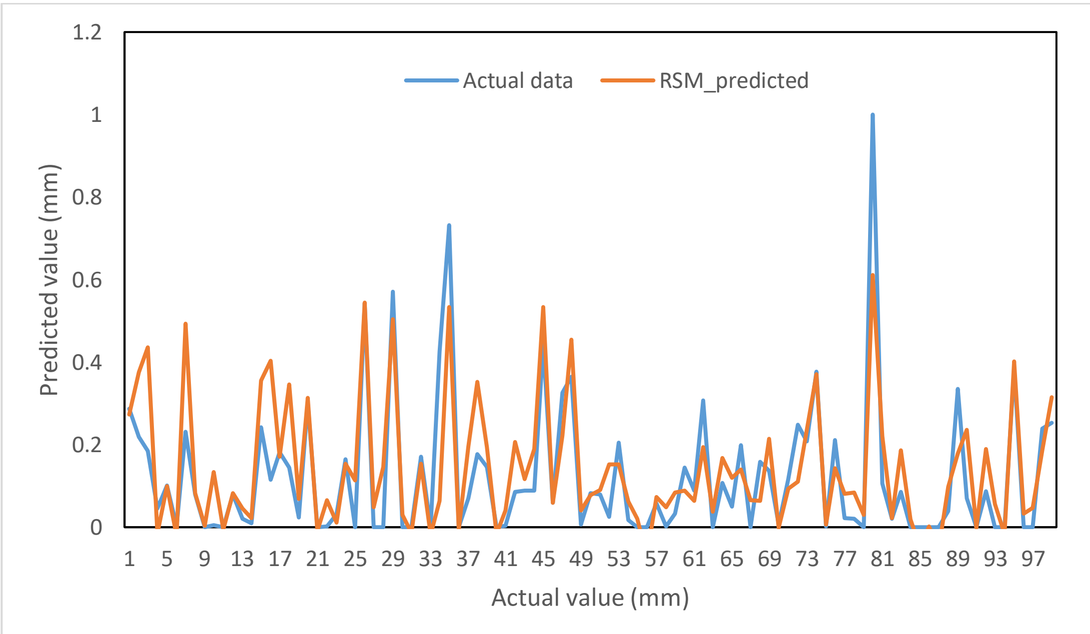


Figure 4.4: Confidence index



(a)



(b)

Figure 4.5: Predicted vs Experimental values for RSM

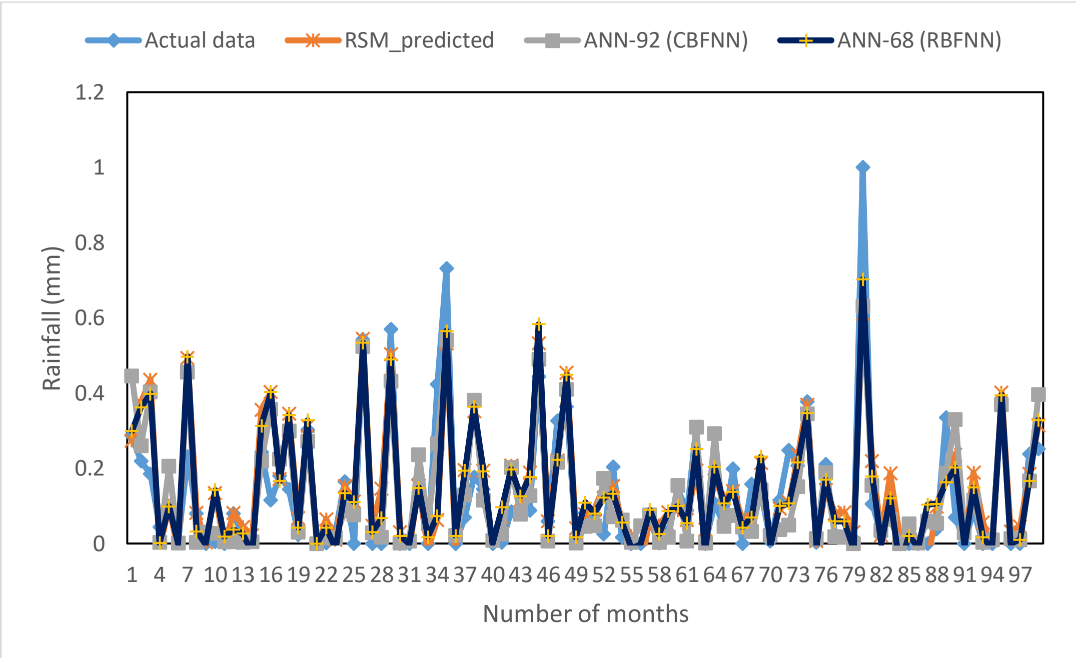


Figure 4.6: Comparative illustration of the models used in this thesis

Table 4.19: Statistiscal comparison of models used in this thesis

Models	ANN-92 (CFNN)	ANN-68 (RBFNN)	RSM
R^2	0.7059	0.7140	0.6670
RMSE	0.0933	0.0930	0.1140

The Figure 4.6 and Table 4.19 gives a comparison of the ANN algorithms and the RSM statistical performance. The accuracy of the models is analysed using the correlation of coefficient and RMSE values. The Table 4.19 shows that the ANN 68 (RBFNN) has the best performance and thus is the most accurate predictive model developed in this thesis.

4.4 Multiple Linear Regression

MLR-model 92

For the Model 92 input combination, the regression equation gotten was:

$$RF = -0.0296NM + 1.99438T_{min} - 2.27776T_{avg} + 0.034861GSR - 0.06221WS + 0.38208$$

The R^2 of the regression equation of the model 92 input combination is 0.57002 while the adjusted R^2 is 0.5626.

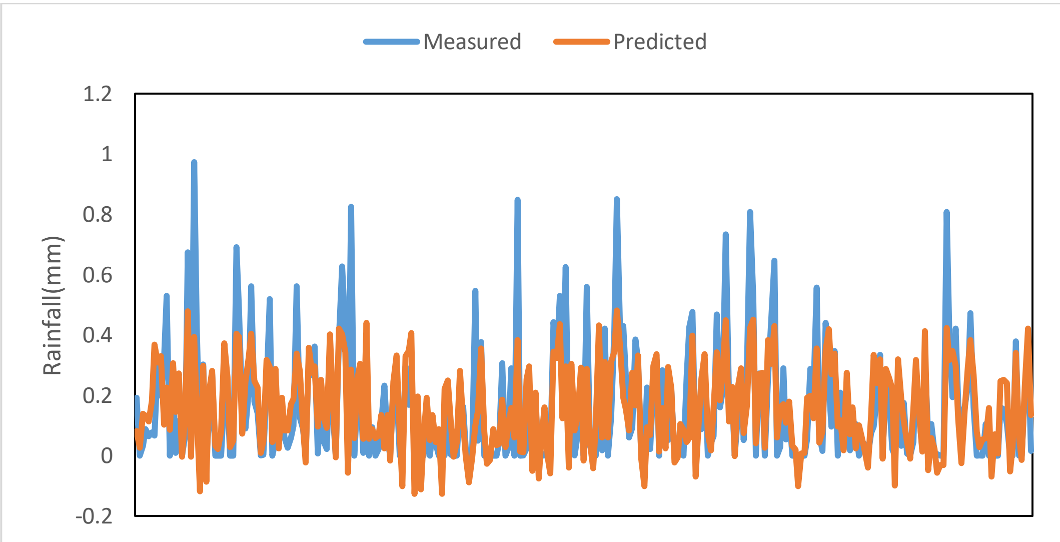


Figure 4.7: Predicted vs Experimental values for MLR (model 92) for the training dataset

The Figure 4.7 shows the measure vs predicted value for the training dataset using model 92.

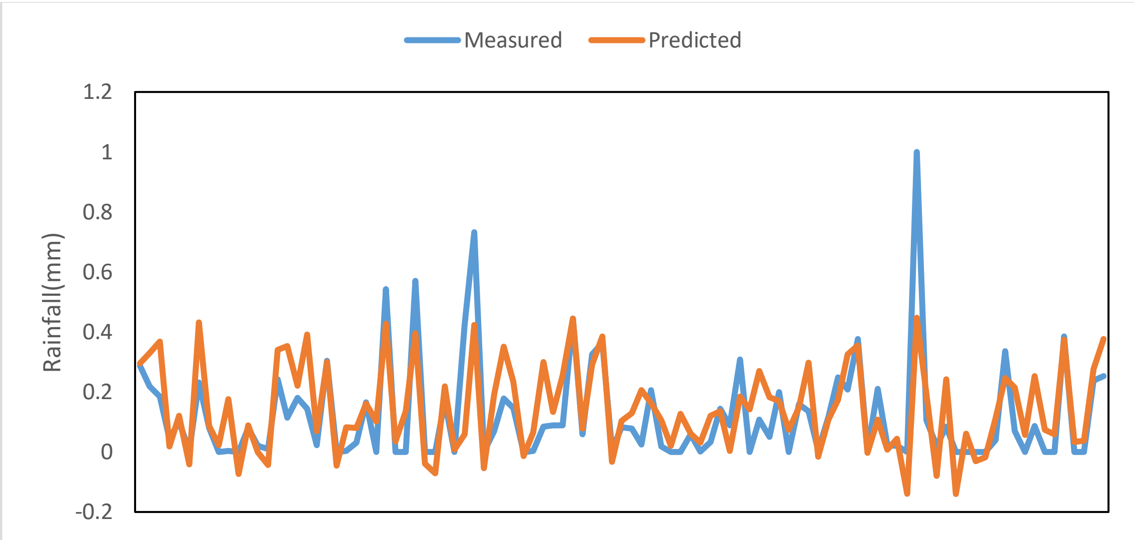


Figure 4.8: Predicted vs Experimental values for MLR (model 92) for the testing dataset

The Figure 4.8 shows the predicted vs measured graph of rainfall for the predicted and measured dataset. The R^2 value for the testing dataset is 0.5359.

MLR-model 68

For the model 68 input combinations, the regression equation gotten is:

$$\mathrm{RF} = -0.00535\mathrm{NM} + 0.854435\mathrm{T}_{\mathrm{min}} - 1.01992\mathrm{T}_{\mathrm{max}} - 0.31974\mathrm{SD}$$

The R^2 value for the training dataset is 0.590824 and the adjusted R^2 value is 0.585219.

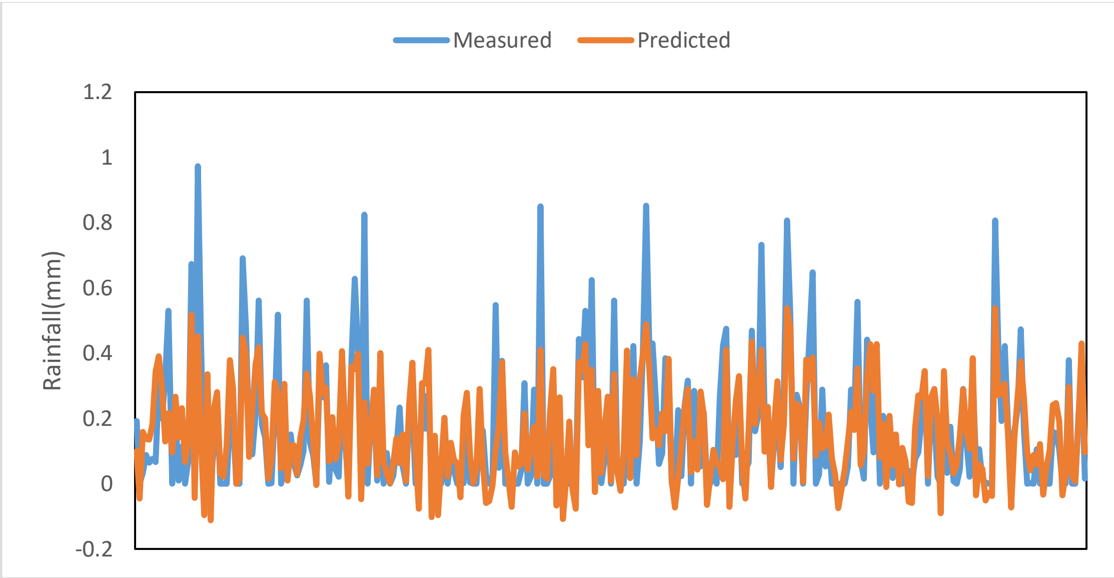


Figure 4.9: Predicted vs Experimental values for MLR (model 68) for the training dataset

The Figure 4.9 shows the measure vs predicted value for the training dataset using model 92.

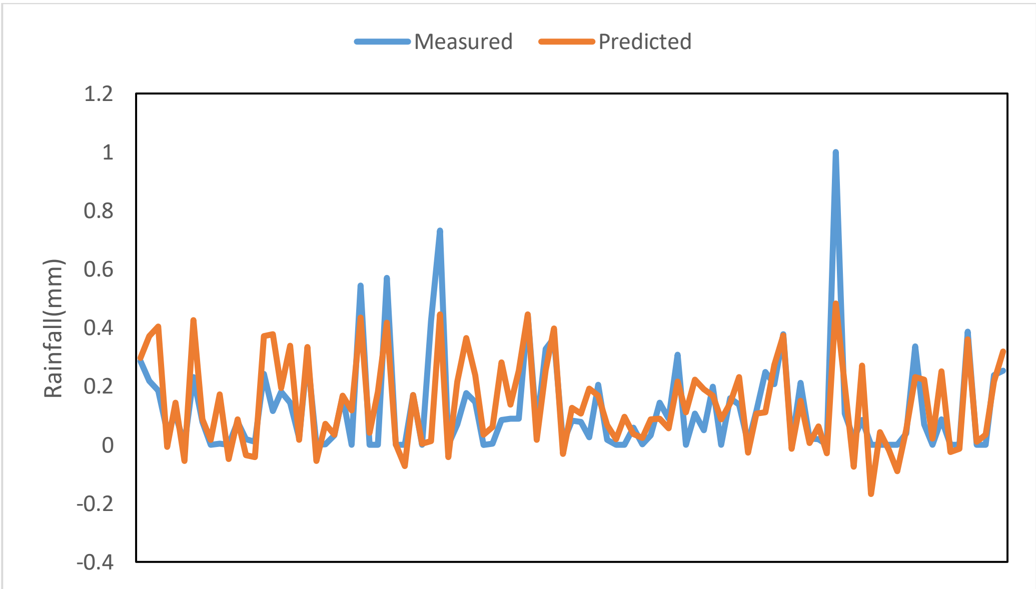


Figure 4.10: Predicted vs Experimental values for MLR (model 68) for the testing dataset

The Figure 4.10 shows the predicted vs measured graph of rainfall for the predicted and measured dataset for the testing dataset. The R^2 value for the testing dataset is 05537.

Table 4.20: Statistiscal comparison (R^2) of MLR and RSM models used in this thesis

Models	Model 92	Model 68
RSM	0.6396	0.6645
MLR	0.5339	0.5337

The Table 4.20 shows the statistical comparison between the RSM and MLR for model 92 and 68 input combinations. The Table shows that the RSM gives a more accurate prediction accuracy for both models. This is due to the combinations of input parameters in the RSM model which portrays a more practical approach to prediction analysis, as different parameters affects rainfall at a particular time.

4.5 Forecasting Result

4.5.1 Winters method

The pattern of rainfall is seen to be right based on the peaks and troughs of the seasonal change. The statistical performance result for this forecast is 0.13 and 0.22 for the MSD and MAD respectively.

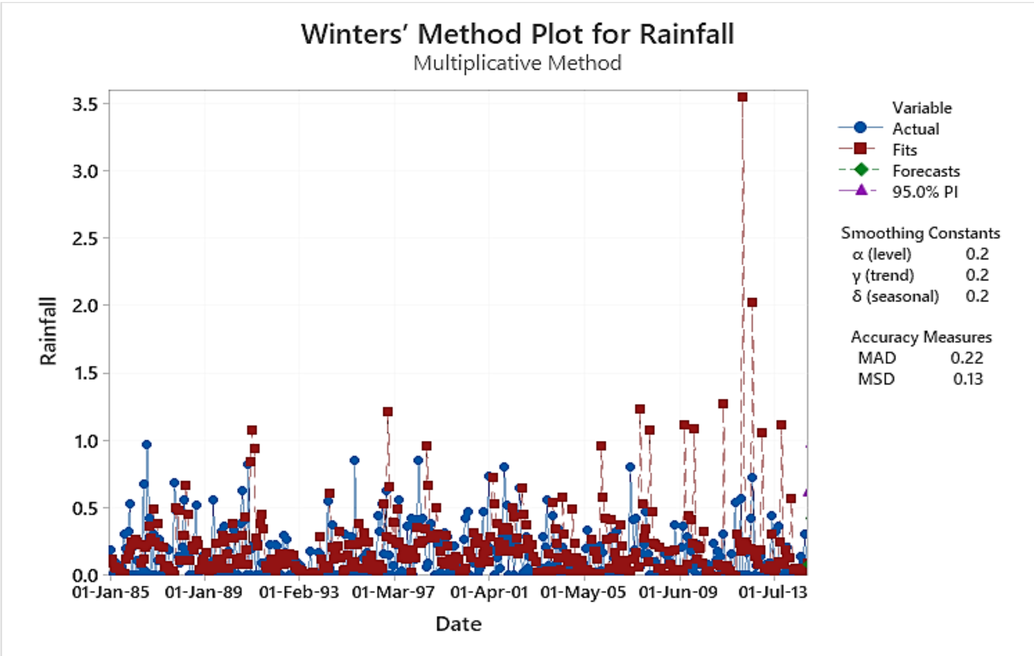


Figure 4.7: Winters' method plot for rainfall (multiplicative method)

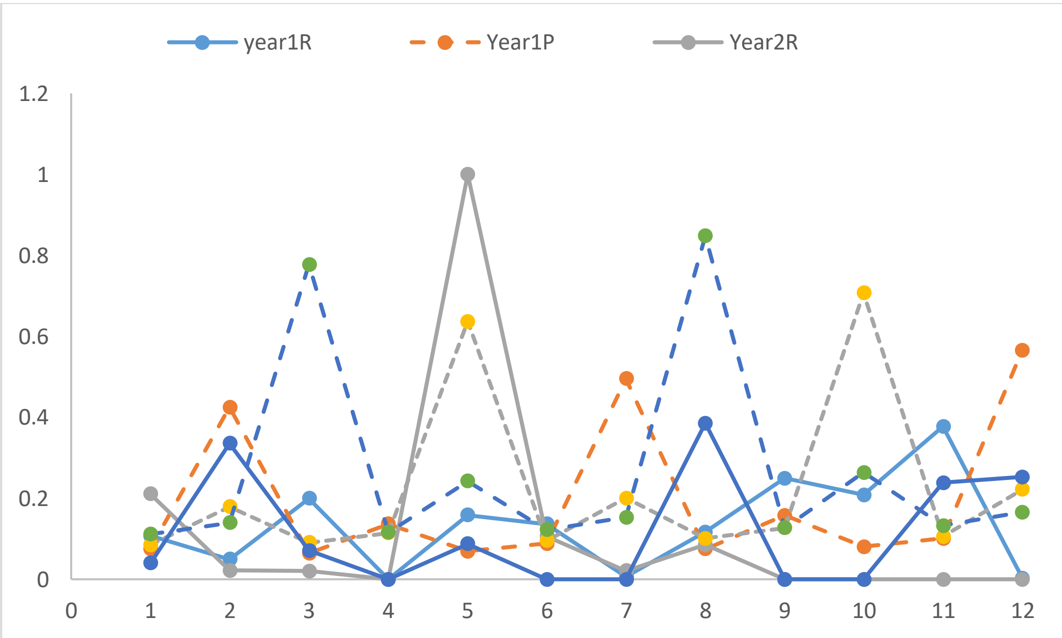


Figure 4.8 : Forecasted vs actual rainfall values for multiplicative method (1/2015-12/2017)

The Figure 4.9 and 4.10 shows the additional model type of the winters' method. The Figure shows that the forecasted values for the period of 2015 – 2017 shows a better relationship with the actual values as compared with the multiplicative model type.

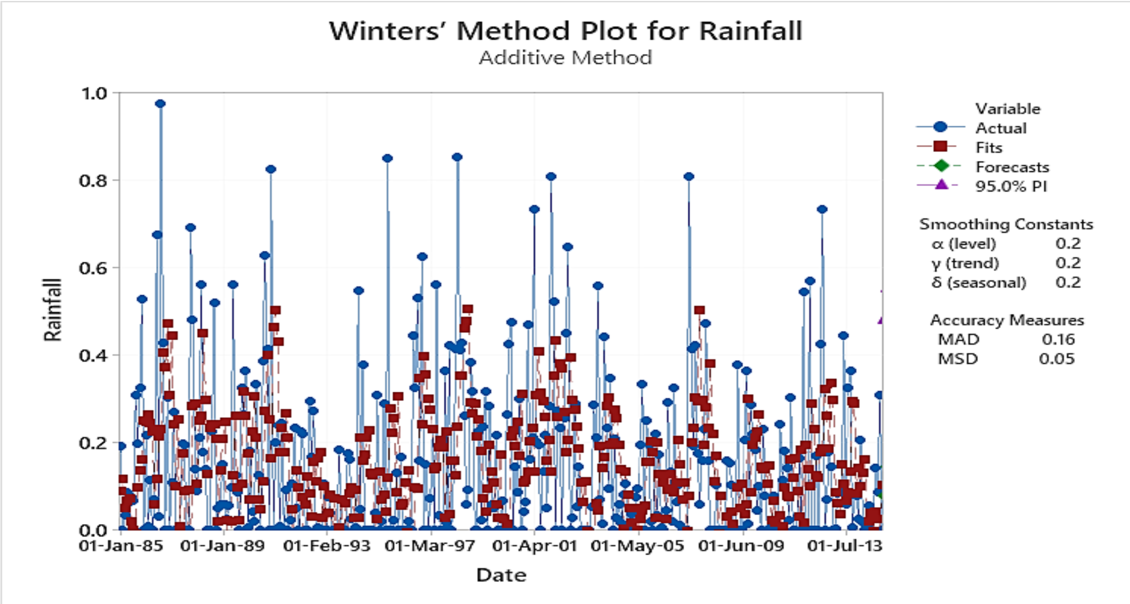


Figure 4.9 : Winters' method plot for rainfall (additive method)

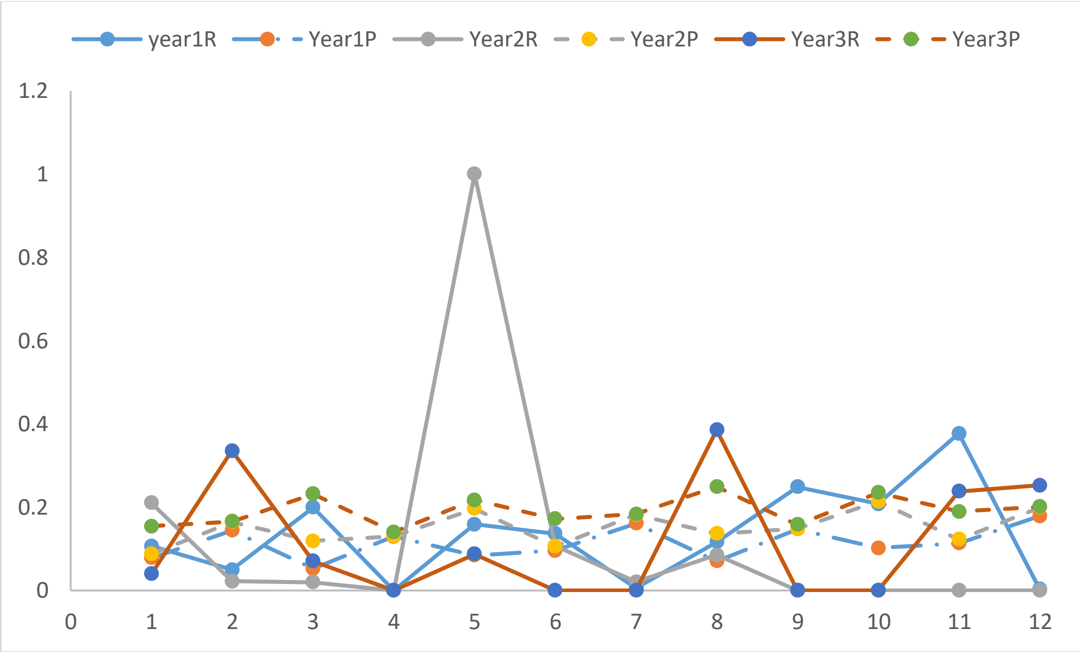


Figure 4.10 : Forecasted vs actual rainfall values for additive method (1/2015-12/2017)

The additive model type is used in making forecast from 2015 to 2025. The Figure 4.11 shows the forecasted values.

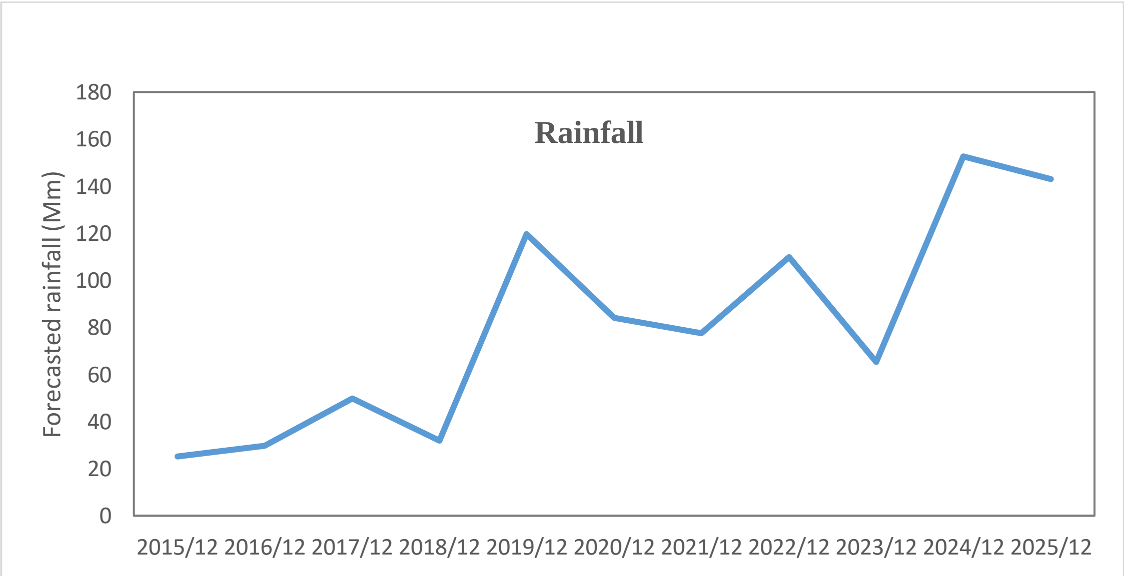


Figure 4.11 : Forecasted rainfall for December across 2015-2025

The Figure 4.11 shows the forecasted rainfall from 2015-2025 for the wettest month in North Cyprus. The forecasted result shows that there would be decrease in the rainfall value in 2020. And this was noticeable in this year as there was low rainfall. The graph shows that in 2021, there will be drop in rainfall value. This can be attributed to the issue of global warming. However, there is a rise in rainfall in 2024 and 2025, and this can be attributed to the efforts to cut down climate change, which studies have shown to be on a positive path. The RMSE value for the forecasted rainfall between 2017 and 2025 was 0.375.

Table 4.20: Forecasted rainfall values In Mm for 2015-2020

Dates	Forecast	Dates	Forecast	Dates	Forecast	Dates	Forecast
01-01-15	14.65099	01-01-18	49.75166	01-01-21	157.0602	01-01-24	97.27278
01-02-15	15.34369	01-02-18	27.825	01-02-21	69.50497	01-02-24	155.82
01-03-15	89.67842	01-03-18	41.73716	01-03-21	144.1092	01-03-24	124.977
01-04-15	13.96935	01-04-18	87.18614	01-04-21	89.87329	01-04-24	123.483
01-05-15	30.26306	01-05-18	54.76955	01-05-21	89.07505	01-05-24	68.211
01-06-15	19.66581	01-06-18	54.66708	01-06-21	49.449	01-06-24	101.2396
01-07-15	20.25911	01-07-18	30.21	01-07-21	73.4718	01-07-24	157.41
01-08-15	116.6775	01-08-18	45.70399	01-08-21	152.2411	01-08-24	129.9919
01-09-15	17.93618	01-09-18	95.31801	01-09-21	94.88811	01-09-24	128.3984
01-10-15	38.39493	01-10-18	59.78437	01-10-21	93.99048	01-10-24	151.05
01-11-15	24.68063	01-11-18	59.5825	01-11-21	157.2351	01-11-24	105.2064
01-12-15	25.17453	01-12-18	31.8	01-12-21	77.43863	01-12-24	152.64
01-01-16	143.6765	01-01-19	49.67082	01-01-22	158.8951	01-01-25	135.0067
01-02-16	21.90301	01-02-19	103.4499	01-02-22	99.90293	01-02-25	133.3139
01-03-16	46.5268	01-03-19	64.79919	01-03-22	98.9059	01-03-25	73.14
01-04-16	29.69545	01-04-19	64.49793	01-04-22	54.855	01-04-25	109.1733

01-05-16	30.08996	01-05-19	155.1204	01-05-22	81.40546	01-05-25	154.23
01-06-16	170.6756	01-06-19	53.63765	01-06-22	155.979	01-06-25	140.0215
01-07-16	25.86984	01-07-19	111.5817	01-07-22	104.9178	01-07-25	138.2293
01-08-16	54.65867	01-08-19	69.81401	01-08-22	103.8213	01-08-25	76.32
01-09-16	34.71027	01-09-19	69.41335	01-09-22	57.24	01-09-25	113.1401
01-10-16	35.00538	01-10-19	38.637	01-10-22	85.37229	01-10-25	127.2
01-11-16	157.2033	01-11-19	57.60448	01-11-22	150.255	01-11-25	145.0363
01-12-16	29.83667	01-12-19	119.7136	01-12-22	109.9326	01-12-25	143.1447
01-01-17	62.79054	01-01-20	74.82883	01-01-23	108.7367		
01-02-17	39.72509	01-02-20	74.32878	01-02-23	158.9523		
01-03-17	39.92081	01-03-20	41.34	01-03-23	89.33912		
01-04-17	155.2317	01-04-20	61.57131	01-04-23	158.6502		
01-05-17	33.8035	01-05-20	127.8455	01-05-23	114.9474		
01-06-17	70.92241	01-06-20	79.84365	01-06-23	113.6522		
01-07-17	44.73991	01-07-20	79.2442	01-07-23	158.7933		
01-08-17	44.83623	01-08-20	150.3027	01-08-23	93.30595		
01-09-17	23.85	01-09-20	65.53814	01-09-23	155.5656		
01-10-17	37.77033	01-10-20	135.9774	01-10-23	119.9622		
01-11-17	79.05427	01-11-20	84.85847	01-11-23	118.5676		
01-12-17	49.75473	01-12-20	84.15963	01-12-23	65.19		

The Table 4.20 shows the forecasted rainfall for 2015 to 2025.

4.5.2 ARIMA model

The Figure 4.12 is presented to show the pattern of rainfall through the months. The Figure shows that there is seasonality in the data.

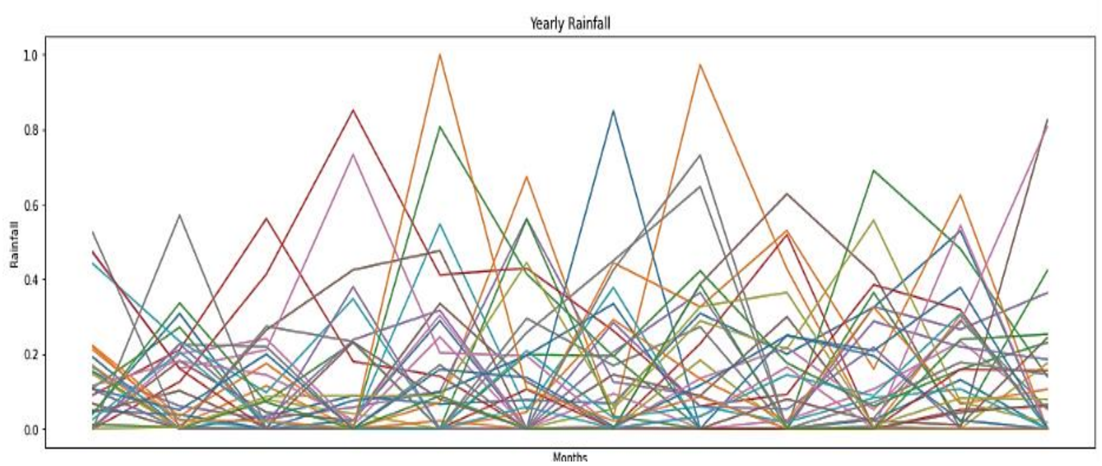


Figure 4.12: Plot of monthly rainfall through the years

The dataset is split into three, training, testing and validation to model and test the ARIMA model. The Table 3.5 shows some of the result of the result of the model and the error of each forecast. The Table shows that the highest error recorded was 0.311. The best forecast was gotten for the month of march.

Table 4.21: Predicted and Measured rainfall value for the validation dataset

Month	Rainfall	Predicted	Error
2013-01-01	0.000063	0.124109	-0.124046
2013-02-01	0.003145	0.091952	-0.088808
2013-03-01	0.085535	0.075498	0.010037
2013-04-01	0.088050	0.119209	-0.031159
2013-05-01	0.088050	0.138146	-0.050096
2013-06-01	0.444025	0.129537	0.314488
2013-07-01	0.059119	0.178272	-0.119153
2013-08-01	0.326415	0.189134	0.137282

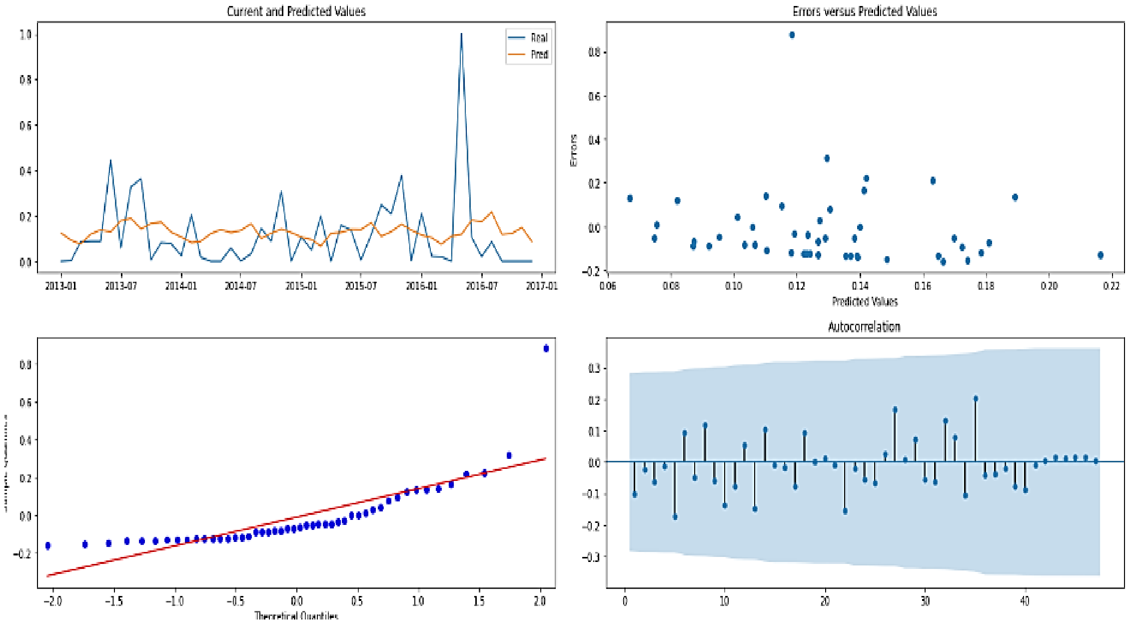


Figure 4.13: Plot of test and predicted rainfall values, with the erros and autocorrelations

The Figure 4.13 shows the test and the predicted rainfall values. Also, the erros are presented.

The python code used for developing the ARIMA is shown as:

```
Model = sm.tsa.statespace.SARIMAX (future, order= (3, 0, 0), seasonal
order =(0,1,1,12), trend='c')

result = model.fit (disp=False)
```

while the forecast was made using:

```
forecast = result. Predict (start = len(future),

                          end = (len(future)-1) + 10 * 12,

                          typ = 'levels').rename('Forecast')
```

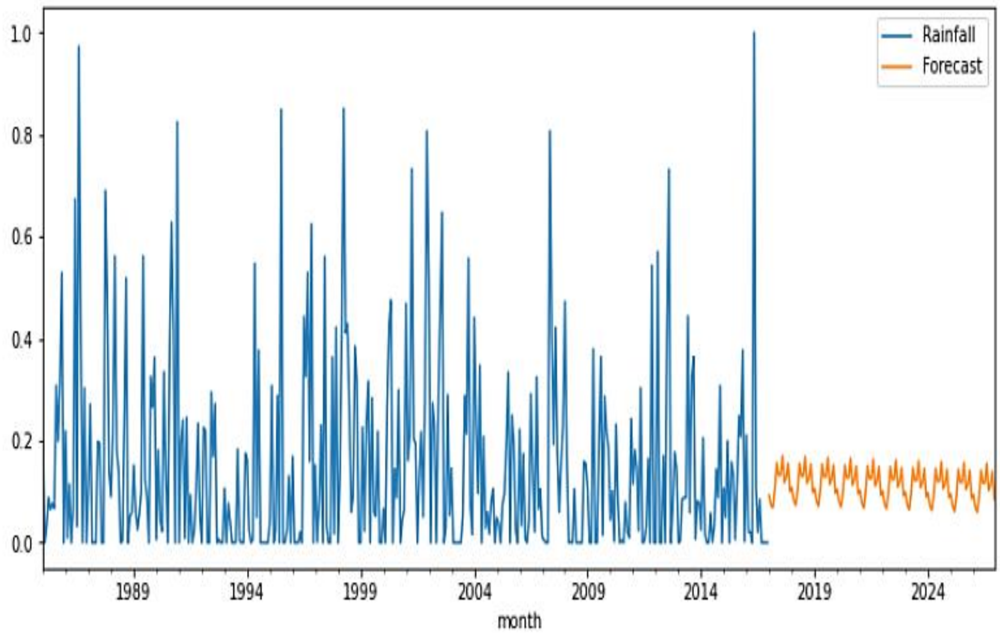


Figure 4.14: Plot of future forcast of rainfall between 2017 and 2025

The Figure 4.14 shows that the top peaks of rainfall were not totally efficiently retrieved by the ARIMA model, as was shown from the error retrieved from the model development. However, the forecast shows the continuous pattern of rainfall over the years and reduction in rainfall values. The ARIMA also showed a reduction of the RMSE value from 0.21 to 0.14 which represents a 34.38% reduction.

Table 4.20: Forecasted rainfall values In Mm for 2017-2025 using ARIMA

Dates	Forecast	Dates	Forecast	Dates	Forecast	Dates	Forecast
01-01-17	44.29537	01-01-18	67.02542	01-01-21	59.54083	01-01-24	52.05536
01-02-17	16.10126	01-02-18	30.70277	01-02-21	23.21768	01-02-24	15.73221
01-03-17	11.47656	01-03-18	17.46788	01-03-21	9.982565	01-03-24	2.497099
01-04-17	62.54916	01-04-18	62.27313	01-04-21	54.78772	01-04-24	47.30226
01-05-17	138.6015	01-05-18	137.299	01-05-21	129.8136	01-05-24	122.3281

01-06-17	103.0985	01-06-18	101.1263	01-06-21	93.64089	01-06-24	86.15543
01-07-17	104.258	01-07-18	101.9341	01-07-21	94.44864	01-07-24	86.96317
01-08-17	158.9722	01-08-18	156.5586	01-08-21	149.0731	01-08-24	141.5877
01-09-17	82.94445	01-09-18	80.48326	01-09-21	72.99779	01-09-24	65.51233
01-10-17	97.20417	01-10-18	94.72148	01-10-21	87.23602	01-10-24	79.75055
01-11-17	137.0209	01-11-18	134.5313	01-11-21	127.0458	01-11-24	119.5604
01-12-17	56.86206	01-12-18	54.36918	01-12-21	46.88371	01-12-24	39.39824
01-01-19	64.53114	01-01-22	57.04568	01-01-25	49.56021		
01-02-19	28.20799	01-02-22	20.72253	01-02-25	13.23706		
01-03-19	14.97288	01-03-22	7.48741	01-03-25	0.001943		
01-04-19	59.77803	01-04-22	52.29257	01-04-25	44.8071		
01-05-19	134.8039	01-05-22	127.3185	01-05-25	119.833		
01-06-19	98.6312	01-06-22	91.14574	01-06-25	83.66027		
01-07-19	99.43895	01-07-22	91.95348	01-07-25	84.46801		
01-08-19	154.0634	01-08-22	146.578	01-08-25	139.0925		
01-09-19	77.9881	01-09-22	70.50264	01-09-25	63.01717		
01-10-19	92.22633	01-10-22	84.74086	01-10-25	77.25539		
01-11-19	132.0361	01-11-22	124.5507	01-11-25	117.0652		
01-12-19	51.87402	01-12-22	44.38855	01-12-25	36.90309		
01-01-20	62.03599	01-01-23	54.55052				
01-02-20	25.71284	01-02-23	18.22737				
01-03-20	12.47772	01-03-23	4.992254				
01-04-20	57.28288	01-04-23	49.79741				
01-05-20	132.3088	01-05-23	124.8233				

01-06-20	96.13605	01-06-23	88.65058
01-07-20	96.94379	01-07-23	89.45833
01-08-20	151.5683	01-08-23	144.0828
01-09-20	75.49295	01-09-23	68.00748
01-10-20	89.73117	01-10-23	82.24571
01-11-20	129.541	01-11-23	122.0555
01-12-20	49.37887	01-12-23	41.8934

The Table 4.20 shows the forecasted rainfall for the ARIMA model.

Comparative analysis between the wınters model and the ARIMA model is ascertained using the RMSE value of the forecasted period between 2017 and 2025. The ARIMA model is shown to have a better forecast accuracy as the RMSE value is 0.14 as compared to the 0.37 RMSE value of the winter’s method.

CHAPTER 5
CONCLUSIONS

Rainfall is a vital phenomenon to humanity, as it affects our everyday life. The global climatic changes is affecting rainfall cycles, and this has a direct impact of some activities like agriculture, industrial activities, and events like flood and land slide. This has been made it even more important to keep track of rainfall occurrence, in order to manage these stated activities. The availability of water is likewise crucial, hence the need to be able to make forecast of rainfall. To make rainfall forecast with less cost and high precision, modern technology is utilized. Researches have shown that artificial intelligence is effective for making accurate rainfall prediction.

This thesis was aimed at developing a prediction using artificial neural network for effectively predicting rainfall in North Cyprus. The thesis utilized the Radial basis neural (RBFNN) network and the cascade neural network (CFNN) in describing different ANN architecture having different hidden layers, neurons and transfer functions. Also, an important process in this prediction model development is the combination of different input variables in each ANN model. The reason for doing this was to find the optimum input variables that will give the most accurate rainfall prediction. Also, this thesis compared the optimum ANN architecture with the response surface methodology (RSM) mathematical model. The thesis found out that the most important input variables based on the best prediction accuracy were ANN-92 for CFNN and ANN-68 for RBFNN. Both ANN models outperformed the RSM mathematical model whose highest prediction accuracy was 0.6670 whereas the model accuracy for ANN-68 and ANN-92 was 0.7140 and 0.7059 respectively. The practical implication of these results is that, in making rainfall predictions in North Cyprus, the Temperature and the month are important parameters. This is so because in both optimum ANN models, there are temperature variables.

The forecasting models used in this thesis was the winter’s method and the ARIMA model. The ARIMA model gave a comparatively more accurate forecasting estimation.

For further study, time series forecasting can be done with shorter time intervals for better rainfall accuracy. Also, for improved accuracy, other machine learning models can be used combined with more database for more precise rainfall prediction.

REFERENCES

Abbot, J., & Marohasy, J. (2013). The potential benefits of using artificial intelligence for monthly rainfall forecasting for the Bowen Basin, Queensland, Australia. *WIT Transactions on Ecology and the Environment*, *171*, 287–297.

Abbot, John, & Marohasy, J. (2014). Input selection and optimisation for monthly rainfall forecasting in queensland, australia, using artificial neural networks. *Atmospheric Research*, *138*, 166–178.

Abdullahi, J., & Elkiran, G. (2017). Prediction of the future impact of climate change on reference evapotranspiration in Cyprus using artificial neural network. *Procedia Computer Science*, 120, 276–283.

Ali, M., Prasad, R., Xiang, Y., & Yaseen, Z. M. (2020). Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts. *Journal of Hydrology*, 584, 124647.

Alsalibi B. (2010). Long-term ground water data breakdown and future predictions: Yeşilköy (Agios Andronikos) case study. *Master dissertation*, *Near East University*, Nicosia, Turkey.

Arslan, B., & Akün, E. (2019). Management, contamination and quality evaluation of groundwater in North Cyprus. *Agricultural Water Management*, 222, 1–11.

Artrith, N., & Urban, A. (2016). An implementation of artificial neural-network potentials for atomistic materials simulations: Performance for TiO2. *Computational Materials Science*, *114*, 135-150.

Bagirov, A. M., & Mahmood, A. (2018). A comparative assessment of models to predict monthly rainfall in Australia. *Water resources management*, 32(5), 1777-1794.

Bagirov, A. M., & Mahmood, A. (2018). A comparative assessment of models to predict monthly rainfall in Australia. *Water Resources Management*, *32*(5), 1777–1794.

Bagirov, A. M., Mahmood, A., & Barton, A. (2017). Prediction of monthly rainfall in Victoria, Australia: clusterwise linear regression approach. *Atmospheric Research*, *188*, 20–29.

Beaudin, M., & Zareipour, H. (2017). Energy solutions to combat global warming. *Lecture Notes in Energy*, 33, 753–793.

Bellerby, T., Todd, M., Kniveton, D., & Kidd, C. (2000). Rainfall estimation from a combination of TRMM precipitation radar and GOES multispectral satellite imagery through the use of an artificial neural network. *Journal of applied Meteorology*, *39*(12), 2115-2128.

Berry, M. J., & Linoff, G. S. (2004). *Data mining techniques: for marketing, sales, and customer relationship management*. John Wiley & Sons.

Bisht, D., Joshi, M. C., & Mehta, A. (2015). Prediction of monthly rainfall of Nainital region using artificial neural network (ANN) and support vector machine (SVM). *Ijariie*, *1*(3), 400–406.

Blum, A. (1992). *Neural networks in C++: an object-oriented framework for building connectionist systems*. John Wiley & Sons, Inc.

Boger, Z., & Guterman, H. (1997). Knowledge extraction from artificial neural networks models. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics* (Vol. 4, pp. 3030–3035).

Bowden, G. J., Maier, H. R., & Dandy, G. C. (2002). Optimal division of data for neural network models in water resources applications. *Water Resources Research*, *38*(2), 2-1.

Caudill, M., & Butler, C. (1994). *Understanding neural networks: computer explorations: a workbook in two volumes with software for the macintosh and pc compatibles*. MIT press.

Chiteka, K., & Enweremadu, C. C. (2016). Prediction of global horizontal solar irradiance in Zimbabwe using artificial neural networks. *Journal of Cleaner Production*, *135*, 701-711.

Dash, Y., Mishra, S. K., & Panigrahi, B. K. (2018). Rainfall prediction for the Kerala state of India using artificial intelligence approaches. *Computers & Electrical Engineering*, 70, 66-73.

DDubey, A. (2015). Artificial Neural Network Models for Rainfall Prediction in Pondicherry. *International Journal of Computer Applications*, *120*(3), 30–35.

Dharma, S., Hassan, M. H., Ong, H. C., Sebayang, A. H., Silitonga, A. S., Kusumo, F., & Milano, J. (2017). Experimental study and prediction of the performance and exhaust emissions of mixed Jatropha curcas-Ceiba pentandra biodiesel blends in diesel engine using artificial neural networks. *Journal of cleaner production*, *164*, 618-633.

Díaz-Torres, J. J., Hernández-Mena, L., Murillo-Tovar, M. A., León-Becerril, E., López-López, A., Suárez-Plascencia, C., & Ojeda-Castillo, V. (2017). Assessment of the modulation effect of rainfall on solar radiation availability at the E arth's surface. *Meteorological Applications*, *24*(2), 180-190.

Elkiran, G. (2004). The Water Budget Analysis of Girne Region , North Cyprus. In *Proceedings of the 6th International Congress on Advances in Civil Engineering, 6-8 October 2004 Bogazici University*, Istanbul, Turkey, 6–8.

Elkiran, G., & Ergil, M. (2006). The assessment of a water budget of North Cyprus. *Building and Environment*, 41(12), 1671–1677.

Elkiran, Gozen, & Ergil, M. (2006). Integrated water resources planning and management of North Cyprus. In *Proceedings of Where Waters Meet: Auckland 28 November–2 December 2005: New Zealand Hydrological Society*: Wellington, New Zealand.

Elkiran, Gozen, & Ongul, Z. (2009). Implications of excessive water withdrawals to the environment of Northern Cyprus. *Water and Environment Journal*, 23(2), 145–154.

Elkiran, Gozen, Aslanova, F., & Hiziroglu, S. (2019). Effluent water reuse possibilities in Northern Cyprus. *Water (Switzerland)*, 11(2), 1–13.

Ergil, M. E. (2000). The salination problem of the Guzelyurt aquifer, Cyprus. *Water Research*, 34(4), 1201–1214.

Esmaeili-Jaghdan, H., Shariati, A. & Nikou, M. (2016). A hybrid smart modeling approach for estimation of pure ionic liquids viscosity. *Journal of Molecular Liquid*, 4(2), 14–27.

Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural computation*, *4*(1), 1-58.

Ghoreishi, S. M., & Heidari, E. (2013). Extraction of epigallocatechin-3-gallate from green tea via supercritical fluid technology: Neural network modeling and response surface optimization. *The Journal of Supercritical Fluids, 74,* 128-136.

Gökçekuş, H. (2014). TRNC's water law and policies. Retrieved August 4, 2019 from http://www.oicvet.org/Presentations/Water_Management_Symposium/Turkey/TRNC

Gokcekus, H., Kabdasli, S., Kabdasli, I., Turker, U., Tunay, O., & Olmez, T. (2003). Pollution of coastal region impacted by acid mine drainage in Morphou bay, Northern Cyprus. *Journal of Environmental Science and Health - Part A Toxic/Hazardous Substances and Environmental Engineering*, 38(8), 1445–1457.

Gozen, E., & Turkman, A. (2009). Water scarcity impacts on Northern Cyprus and alternative mitigation strategies. *Journal of Environmental problems of Central Asia and their economic, social and security impacts*, 23(2), 241–250.

Gupta, A. K. (2010). Predictive modelling of turning operations using response surface methodology, artificial neural networks and support vector regression. *International Journal of Production Research*, *48*(3), 763-778.

Hashim, R., Roy, C., Motamedi, S., Shamshirband, S., Petković, D., Gocic, M., & Lee, S. C. (2016). Selection of meteorological parameters affecting rainfall estimation using neuro-fuzzy computing methodology. *Atmospheric research*, *171*, 21-30.

Hashim, R., Roy, C., Motamedi, S., Shamshirband, S., Petković, D., Gocic, M., & Lee, S. C. (2016). Selection of meteorological parameters affecting rainfall estimation using neuro-fuzzy computing methodology. *Atmospheric research*, 171, 21-30.

Haykin, S., & Network, N. (2004). Neural network a comprehensive foundation. *Neural Network*, 2, 41-51

Iizumi, T., & Ramankutty, N. (2015). How do weather and climate influence cropping area and intensity?. *Global Food Security*, 4, 46-50.

Jamal, A. A. A., & Türker, U. (2015). Assessment of the regional water balance of the limestone subaquifers of Cyprus. *Journal of Earth System Science*, 124(3), 541–551.

Kahramanoğlu, İ., Usanmaz, S., & Alas, T. (2020). Water footprint and irrigation use efficiency of important crops in Northern Cyprus from an environmental, economic and dietary perspective. *Saudi Journal of Biological Sciences*, 27(1), 134–141.

Kang, Y., Khan, S., & Ma, X. (2009). Climate change impacts on crop yield, crop water productivity and food security–a review. *Progress in natural Science*, 19(12), 1665-1674.

Karafistan, A., & Gemikonakli, E. (2019). Contaminant evaluation in fish from the mining-impacted Morphou bay, Cyprus, using statistical and artificial neural network analysis. *Mine Water and the Environment*, 38(1), 178–186.

Kashiwao, T., Nakayama, K., Ando, S., Ikeda, K., Lee, M., & Bahadori, A. (2017). A neural network-based local rainfall prediction system using meteorological data on the Internet: A case study using data from the Japan meteorological agency. *Applied Soft Computing Journal*, *56*, 317–330.

Kashiwao, T., Nakayama, K., Ando, S., Ikeda, K., Lee, M., & Bahadori, A. (2017). A neural network-based local rainfall prediction system using meteorological data on the Internet: A case study using data from the Japan Meteorological Agency. *Applied Soft Computing*, 56, 317-330.

Kashyap, Y., Bansal, A., & Sao, A. K. (2015). Solar radiation forecasting with multiple parameters neural networks. *Renewable and Sustainable Energy Reviews*, *49*, 825-835.

Khatib, T., Mohamed, A., Sopian, K., & Mahmoud, M. (2012). Assessment of artificial neural networks for hourly solar radiation prediction. *International journal of Photoenergy*, *2012*.

Khidhir, B. A., Al-Oqaiel, W., & Kareem, P. M. (2015). Prediction models by response surface methodology for turning operation. *Australian journal of Mechanical Engineering*, *3*, 1-6.

Kuligowski, R. J., & Barros, A. P. (1998). Localized precipitation forecasts from a numerical weather prediction model using artificial neural networks. *Weather and forecasting*, *13*(4), 1194-1204.

Kumar, S., & Kaur, T. (2016). Development of ANN based model for solar potential assessment using various meteorological parameters. *Energy Procedia*, *90*, 587-592.

Kundu, S., Khare, D., & Mondal, A. (2017). Future changes in rainfall, temperature and reference evapotranspiration in the central India by least square support vector machine. *Geoscience Frontiers*, 8(3), 583-596.

Kutzbach, J. E. (1967). Empirical eigenvectors of sea-level pressure, surface temperature and precipitation complexes over North America. *Journal of Applied Meteorology*, *6*(5), 791-802.

Lazri, M., Ameur, S., & Mohia, Y. (2014). Instantaneous rainfall estimation using neural network from multispectral observations of SEVIRI radiometer and its application in estimation of daily and monthly rainfall. *Advances in Space Research*, *53*(1), 138–155.

Lima, P. M., & Guedes, E. B. (2015). Rainfall Prediction for Manaus, Amazons with Artificial Neural Networks. *Latin America Congress on Computational Intelligence,* 4(1), 1-5.

Maqsood, I., Khan, M. R., Huang, G. H., & Abdalla, R. (2005). Application of soft computing models to hourly weather analysis in southern Saskatchewan, Canada. *Engineering Applications of Artificial Intelligence*, *18*(1), 115-125.

Mehmet, O., & Biçak, H. A. (2002). *Modern and traditional irrigation technologies in the Eastern Mediterranean*. IDRC.

Mislan, M., Haviluddin, H., Hardwinarto, S., Sumaryono, S., & Aipassa, M. (2015). Rainfall monthly prediction based on artificial neural network: a case study in Tenggarong Station, East Kalimantan-Indonesia. *In the International Conference on Computer Science and Computational Intelligence (ICCSCI 2015)*-Procedia Computer Science 59.

Mohammadpour, R., Asaie, Z., Shojaeian, M. R., & Sadeghzadeh, M. (2018). A hybrid of ANN and CLA to predict rainfall. *Arabian Journal of Geosciences*, *11*(18), 533.

Mohd-Safar, N. Z., Ndzi, D., Sanders, D., Noor, H. M., & Kamarudin, L. M. (2016, September). Integration of fuzzy c-means and artificial neural network for short-term localized rainfall forecasting in tropical climate. In *Proceedings of SAI Intelligent Systems Conference* (pp. 325-348). Springer, Cham.

MuttalebAlhashimi, S. A. (2014). Prediction Of Monthly Rainfall In Kirkuk Using Artificial Neural Network And Time Series Models. *Journal of Engineering and Sustainable Development*, 18(1), 129–143.

Najafi-Marghmaleki, A., Khosravi-Nikou, M. R., & Barati-Harooni, A. (2016). A new model for prediction of binary mixture of ionic liquids+ water density using artificial neural network. *Journal of Molecular Liquids*, *220*, 232-237.Z.

Nastos, P. T., Moustris, K. P., Larissi, I. K., & Paliatsos, A. G. (2013). Rain intensity forecast using Artificial Neural Networks in Athens, Greece. *Atmospheric Research*, *119*, 153–160.

Nastos, P. T., Paliatsos, A. G., Koukouletsos, K. V., Larissi, I. K., & Moustris, K. P. (2014). Artificial neural networks modeling for forecasting the maximum daily total precipitation at Athens, Greece. *Atmospheric Research*, *144*, 141-150.

Nourani, V., Molajou, A., Uzelaltinbulat, S., & Sadikoglu, F. (2019). Emotional artificial neural networks (EANNs) for multi-step ahead prediction of monthly precipitation; case study: northern Cyprus. *Theoretical and Applied Climatology*, 138(3-4), 1419-1434.

Ortiz-García, E. G., Salcedo-Sanz, S., & Casanova-Mateo, C. (2014). Accurate precipitation prediction with support vector classifiers: A study including novel predictive variables and observational data. *Atmospheric research*, *139*, 128-136.

Panda, S. S., Chakraborty, D., & Pal, S. K. (2008). Flank wear prediction in drilling using back propagation neural network and radial basis function network. *Applied soft computing*, *8*(2), 858-871.S.

Pham, B. T., Le, L. M., Le, T. T., Bui, K. T. T., Le, V. M., Ly, H. B., & Prakash, I. (2020). Development of advanced artificial intelligence models for daily rainfall prediction. *Atmospheric Research*, 23(7), 104-110.

Phillips Agboola, O., & Egelioglu, F. (2012). Water scarcity in North Cyprus and solar desalination research: a review. *Desalination and Water Treatment*, 43(1-3), 29-42.

Pour, S. H., Wahab, A. K. A., & Shahid, S. (2020). Physical-empirical models for prediction of seasonal rainfall extremes of Peninsular Malaysia. *Atmospheric Research*, 21(7), 106-115.

Qazi, A., Fayaz, H., Wadi, A., Raj, R. G., Rahim, N. A., & Khan, W. A. (2015). The artificial neural network for solar radiation prediction and designing solar systems: a systematic literature review. *Journal of cleaner production*, *104*, 1-12.

Rahman, M. H., Salma, U., Hossain, M. M., & Khan, M. T. F. (2016). Revenue forecasting using holt–winters exponential smoothing. Research & Reviews: *Journal of Statistics*, 5(3), 19-25.

Ramana, R. V., Krishna, B., Kumar, S. R., & Pandey, N. G. (2013). Monthly rainfall prediction using wavelet neural network analysis. *Water resources management*, *27*(10), 3697-3711.

Reddy, K. S., & Ranjan, M. (2003). Solar resource estimation using artificial neural networks and comparison with other correlation models. *Energy conversion and management*, *44*(15), 2519-2530.

Seino, N., Aoyagi, T., & Tsuguti, H. (2018). Numerical simulation of urban impact on precipitation in Tokyo: How does urban temperature rise affect precipitation?. *Urban Climate*, 23, 8-35.

Sendanayake, S., Miguntanna, N. P., & Jayasinghe, M. T. R. (2015). Predicting solar radiation for tropical islands from rainfall data. *Journal of Urban and Environmental Engineering*, *9*(2), 109-118.

Shabanpour, H., Yousefi, S., & Saen, R. F. (2017). Forecasting efficiency of green suppliers by dynamic data envelopment analysis and artificial neural networks. *Journal of Cleaner Production*, *142*, 1098-1107.

Song, M., Wang, R., & Zeng, X. (2018). Water resources utilization efficiency and influence factors under environmental restrictions. *Journal of Cleaner Production*, 184, 611-621.

Swingler, K. (1996). *Applying neural networks: a practical guide*. Morgan Kaufmann Publisher, California.

Velasco, L. C. P., Serquiña, R. P., Zamad, M. S. A. A., Juanico, B. F., & Lomocso, J. C. (2019). Week-ahead Rainfall Forecasting Using Multilayer Perceptron Neural Network. *Procedia Computer Science*, 161, 386-397.

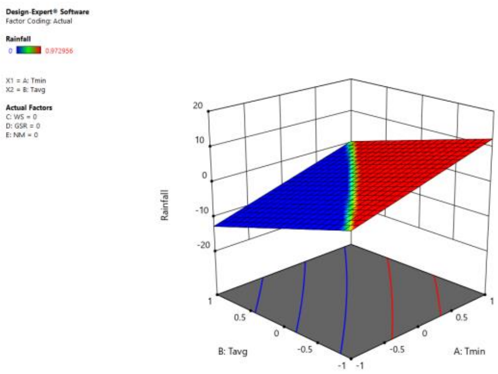
Yanmaz, M. (2013). Modeling of Morphou (Güzelyurt) Flood and Remedial Measures. *Teknik Dergi*, 24(120), 6447–6462.

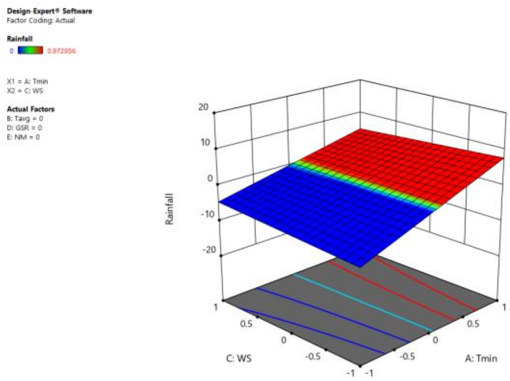
Yukselen, M. A. (2002). Characterization of heavy metal contaminated soils in Northern Cyprus. *Environmental Geology*, 42(6), 597–603.

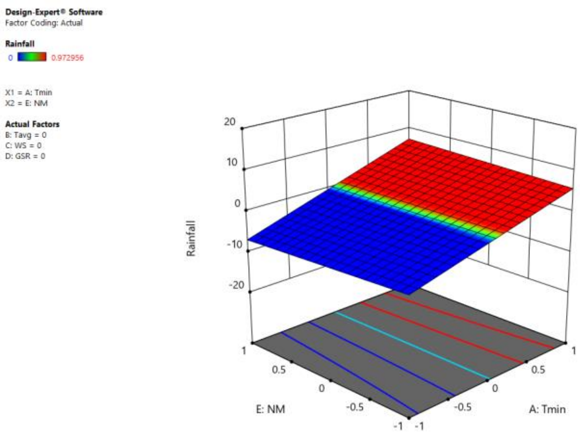
APPENDICES

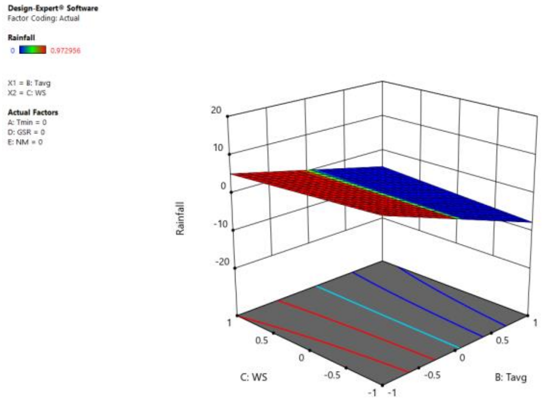
APPENDIX 1

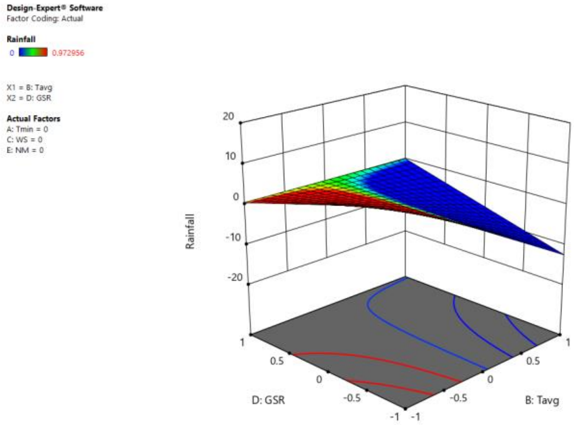
CONTOUR PLOT OF INTERACTION BETWEEN THE INPUT VARIABLES SEQUENTIAL MODEL SUM OF SQUARES

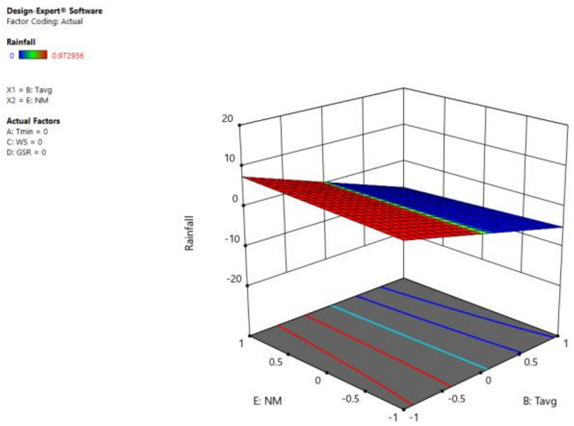


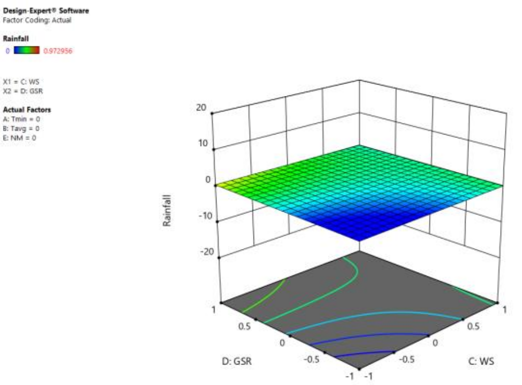


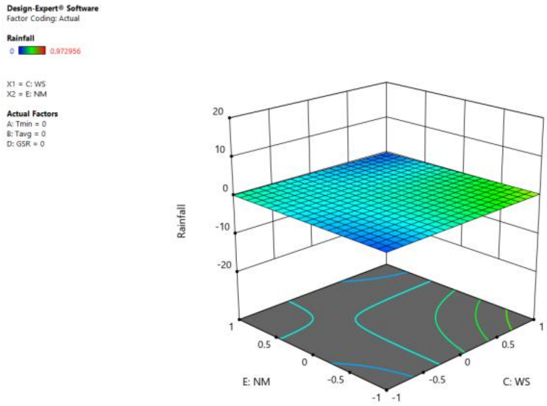


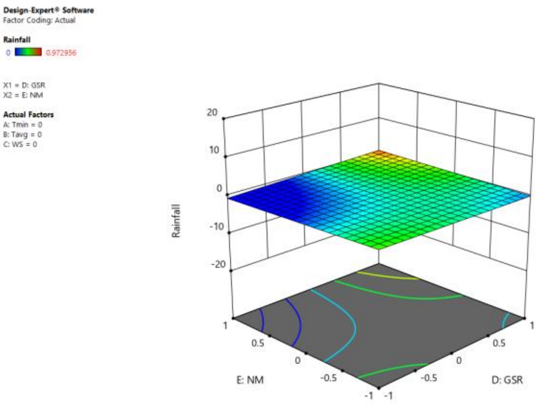

Design-Expert® Software
Factor Coding: Actual
Rainfall
0
0.972956
X1 = B: Tavg
X2 = C: WS
Actual Factors
A: Tmin = 0
D: GSR = 0
E: NM = 0
Rainfall
C: WS
B: Tavg


Design-Expert® Software
Factor Coding: Actual
Rainfall
0
0.972956
X1 = B: Tavg
X2 = D: GSR
Actual Factors
A: Tmin = 0
C: WS = 0
E: NM = 0
Rainfall
D: GSR
B: Tavg


Design-Expert® Software
Factor Coding: Actual
Rainfall
0
0.972956
X1 = B: Tavg
X2 = E: NM
Actual Factors
A: Tmin = 0
C: WS = 0
D: GSR = 0
Rainfall
E: NM
B: Tavg







Squares

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Mean vs Total	7.42	1	7.42			
Linear vs Mean	6.39	5	1.28	77.16	< 0.0001	
2FI vs Linear	**0.7805**	**10**	**0.0781**	**5.43**	**< 0.0001**	**Suggested**
Quadratic vs 2FI	0.1470	5	0.0294	2.08	0.0677	

APPENDIX 2

ETHICAL APPROVAL LETTER

YAKIN DOĞU ÜNİVERSİTESİ

ETHICAL APROVAL DOCUMENT

Date: 17/08 /2020

To the **Graduate School of Applied Sciences**

The research project titled '' MODELING AND FORECASTING OF MONTHLY RAINFALL USING MATHEMATICAL MODELS AND MACHINE LEARNING MODELS: A CASE STUDY IN MORPHOU, NORTHERN CYPRUS '' has been evaluated. Since the researcher(s) will not collect primary data from humans, animals, plants or earth, this project does not need through the ethics committee.

Title: Prof. Dr

Name Surname: Hüseyin Gökçekuş

Signature:

Role in the Research Project: Supervisor

Title: Assist. Prof. Dr.

Name Surname: Youssef Kassem

Signature:

Role in the Research Project: Co-Supervisor

APPENDIX 3

SIMILARITY REPORT

AUTHOR	TITLE	SIMILARITY	GRADE	RESPONSE	FILE	PAPER ID	DATE
Julia Aljamal	ABSTRACT	0%	--	--		1370725854	17-Aug-2020
Julia Aljamal	chapter_4_results docx	0%	--	--		1370722726	17-Aug-2020
Julia Aljamal	CONCLUSIONS	0%	--	--		1370725873	17-Aug-2020
Julia Aljamal	CHAPTER 3	2%	--	--		1370725870	17-Aug-2020
Julia Aljamal	ALL THESIS	5%	--	--		1370725879	17-Aug-2020
Julia Aljamal	CHAPTER 1	6%	--	--		1370725857	17-Aug-2020
Julia Aljamal	CHAPTER 2	11%	--	--		1370725867	17-Aug-2020

Near East University
Civil and Environmental Engineering Faculty

Prof Dr. Huseyin Gokcekus
Dean, Civil and Environmental Engineering Faculty