ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

CLIMATE CHANGE IMPACTS ON CATCHMENT-SCALE EXTREME RAINFALL VARIABILITY

Ph.D. THESIS

Ali DANANDEH MEHR

Department of Civil Engineering

Hydraulics and Water Resources Engineering Program

APRIL 2016

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

CLIMATE CHANGE IMPACTS ON CATCHMET-SCALE EXTREME RAINFALL VARIABILITY

Ph.D. THESIS

Ali DANANDEH MEHR
(501112506)

Department of Civil Engineering

Hydraulics and Water Resources Engineering Programme

Thesis Advisor: Prof. Dr. Ercan KAHYA

APRIL 2016

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

İKLİM DEĞİŞİKLİNİN HAVZA ÖLÇEĞİNDE EKSTREM YAĞIŞLAR DEĞIŞKENLİĞİNE ETKİLERİ

DOKTORA TEZİ

Ali DANANDEH MEHR
(501112506)

İnşaat Mühendisliği Anabilim Dalı

Hidrolik ve Su Kaynaklari Mühendisliği Programı

Tez Danışmanı: Prof. Dr. Ercan KAHYA

NİSAN 2016

Ali DANANDEH MEHR, a Ph.D. student of ITU Graduate School of Science Engineering and Technology student ID 50112506, successfully defended the thesis/dissertation entitled "CLIMATE CHANGE IMPACTS ON CATCHMENT-SCALE EXTREME RAINFALL VARIABILITY", which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Prof. Dr. Ercan KAHYA** Istanbul Technical University	.....
Jury Members :	**Prof. Dr. Necati AĞIRALİOĞLU** Istanbul Technical University	.....
	Prof. Dr. Melek KAZEZYILMAZ ALHAN Istanbul University	.....
	Assoc. Prof. Dr. İbrahim CAN Atatürk University	.....
	Assist. Prof. Dr. Ali UYUMAZ Istanbul Technical University	.....

Date of Submission : 01 March 2016
Date of Defense : 12 April 2016

To my respected professor: Ercan KAHYA,

FOREWORD

With the praise of Allah, I am ready to defence my Ph.D. thesis now and I would like to thank the following persons and institutions that contributed to make this dissertation become true.

Prof. E. Kahya, Istanbul Technical University, Turkey,

Dear professor, I sincerely thank you for all your supports and guidances that has benefited me a lot in all aspects of my life.

Prof. V. Nourani, University of Tabriz, Iran,

I would like to express my sincere appreciation to you for your inspiring advice about studying in Istanbul Technical University. Definitely, those times that I attended in your master's courses are of the priceless times of my life.

Prof. M.R.M. Tabatabai, Shahid Beheshti University,

I sincerely thank you for all you have done for me. If I am ready now to be honoured with a PhD degree, I am sure that one of the main reasons behind is your serious efforts to train me during and after my masters' period.

Dr. S. Soltani, Ministry of Energy, Iran,

Thank you very much for all your supports.

Olgay Şen, Istanbul Technical University, Turkey,

My wonderful classmate, I am glad that I came here and found great friends like you. Thanks for all your helps.

My dear parents, lovely wife, and beloved son,

Definitely who I am now is because of your accompanying and patience. Thanks a million. You will always be in my heart.

Last in sequence but not least in importance, I am greatly thankful to my jury members for their constructive comments and the Scientific and Technological Research Council of Turkey (TUBITAK) for the financial support of this thesis. I also acknowledge Turkish State Meteorological Service and Turkish State Hydraulic Works for kindly providing the data used in the thesis.

March 2016 Ali DANANDEH MEHR

TABLE OF CONTENTS

Page

ABBREVIATIONS

AMR : Annual Maximum Rainfall
ANN : Artificial Neural Network
BP : Backpropagation
CCSM3 : Community Climate System Models-generation 3
CDF : Cumulative Distribution Function
DEM : Digital Elevation Model
DM : Distribution Model
DSc : Dynamic scaling
DSI : Turkish State Hydraulic Works
ECHAM : European Centre Hamburg Model
FFBP : Feed-Forward Backpropagation
GBA : Grid-Based Analysis
GIS : Geographic Information Systems
GCM : General Circulation Model
GP : Genetic Programming
HadGEM : Hadley Centre Global Environmental Model
IPCC : Intergovernmental Panel on Climate Change
ICTP : International Centre for Theoretical Physics
KGE : Kling-Gupta Efficiency
LN-2 : Two-Parameter Lognormal
LN-3 : Three-Parameter Lognormal
LP-3 : Log-Pearson Type 3
LSh : Linear Shifting
MAB : Mean Absolute Bias
MDS : Model Design Storm
MMR : Monthly Maximum Rainfall
MPI : Model Performance Index
NCAR : National Center for Atmospheric Research
NRD : Number of Rainy Days
NSE : Nash-Sutcliffe Efficiency
ODS : Observational Design Storm
PDF : Probability Density Functions
PDS : Projected Design Storm
PDS_{BC} : Bias Corrected Projected Design Storm
PI : Persistence Index
PR-R : Projected Rainfall-Runoff
RCM : Regional Climate Model
RCP : Representative Concentration Pathway
RegCM3 : Regional Climate Model Version 3
RegCM4 : Regional Climate Model Version 4
RMSE : Root Mean Squared Error
SRES : Special Report on Emissions Scenarios

SWAT : Soil and Water Assessment Tool
TAR : Total Annual Rainfall
TMR : Total Monthly Rainfall
TSMS : Turkish State Meteorological Service
VE : Volumetric Efficiency

LIST OF TABLES

Page

LIST OF FIGURES

Page

CLIMATE CHANGE IMPACTS ON CATCHMENT-SCALE EXTREME RAINFALL VARIABILITY

SUMMARY

A changing climate alters regional hydrology and thereby virtually affects every aspect of human life. The local impacts of climate change may significantly vary in different locations even within the borders of a country. Turkey, as a part of the Mediterranean region, is located in one of the most vulnerable regions to climate change. In order to figure out the local impacts of climate change on extreme rainfalls across Rize Province, located in the north east of Turkey, this doctoral thesis provides a grid-based performance evaluation of two general circulation models (GCMs) for rainfall reproduction over the province and investigates the variability of extreme rainfall events at three future periods (i.e., 2013-2040; 2041-2070; and 2071-2099) with respect to high emission/concentration greenhouse gas scenarios. In addition, a data-driven projected rainfall-runoff model is provided for projected streamflow prediction concerning the limited data availability in the province.

Prior to hydrological assessment of climate change at catchment-scale, an applied methodology is necessary to evaluate the performance of climate models available for a given catchment. Therefor, this thesis, in the first phase, presents a grid-based performance evaluation approach as well as an intercomparison framework to evaluate the uncertainty of climate models for rainfall reproduction over Rize Province. For this purpose, outputs of two general circulation models (GCMs), namely ECHAM5 and CCSM3, downscaled by a regional climate model (RCM), namely RegCM3, are used. To this end, five rainfall-borne climatic statistics are computed from the outputs of ECHAM5-RegCM3 and CCSM3-RegCM3 combinations in order to compare with those of observations in the province for the reference period 1961-1990. Performance of each combination is tested by means of scatter diagram, bias, mean absolute bias, root mean squared error, and model performance index (MPI) measures. The results of this phase of study indicated that ECHAM5-RegCM3 overestimates the total monthly rainfall observations whereas CCSM3-RegCM3 tends to underestimate. In terms of maximum monthly and annual maximum rainfall reproduction, ECHAM5-RegCM3 shows higher performance than CCSM3-RegCM3, particularly in the coastland areas. In contrast, CCSM3-RegCM3 outperforms ECHAM5-RegCM3 in reproducing the number of rainy days, especially in the inland areas. The results also revealed that if a GCM-RCM combination performs well for a portion (statistic) of a catchment, it is not necessarily appropriate for the other portions (statistics). Moreover, the MPI measure demonstrated the superiority of ECHAM5-RegCM3 to CCSM3-RegCM3 up to 33% excelling for annual rainfall reproduction in Rize Province.

In the second phase of the thesis, a grid-based methodology is developed and applied to estimate future extreme rainfalls over the province, compatible with adaptive

catchment-scale water resources management. Extreme value theory (EVT) is applied to analyse observational and projected extreme rainfall data including dynamically downscaled climate models simulations from two different GCM-RCM combinations including ECHAM5-RegCM3 and HadGEM2-RegCM4 forced by A2 special report on emissions scenarios (SRES-A2) and representative concentration pathway-8.5 (RCP8.5) greenhouse gas scenarios. A new rapid and effective bias correction method is also developed and applied to adjust the climate models simulations. The EVT analysis results demonstrated significant differences between the model runs for both reference and future periods with considerable spatial variability in rainfall extremes. Based upon ensemble mean results, approximately 30% decrease in the median value of extreme rainfall events is projected over the study region for the near future 2013-2039 and middle of the century. This change dramatically decreases down to 15% of its historic value at the end of the century. Results from the implemented scenarios revealed that more intense rainfalls produced by the GCM forced by RCP8.5. In addition, the results from goodness-of-fit tests among different distribution models showed that general extreme value distribution can be used appropriately to characterize the behavior of projected extreme rainfalls over the study region.

It would be very helpful to hydrologists if data-driven technology incorporating the existing hydrologic observations and forthcoming climatological projections can be integrated into the catchment-scale hydrological modelling. In the last phase of the study, This thesis delineates the development and implementation of two advancements of state of the art genetic programming (GP) approach, namely linear GP (LGP) and multigene GP (MGGP), and an artificial neural network method, namely feed-forward backpropagation (FFBP), for projected rainfall-runoff (PR-R) modelling at one of the catchments in Rize Province by employing observational streamflow together with projected rainfall and temperature data as model inputs. The proposed models have been evaluated in terms of five efficiency criteria including Nash-Sutcliffe efficiency, persistence index, Kling-Gupta efficiency, volumetric efficiency, and root mean squared error. The results show promising role of GP-based models for both bivariate and multivariate PR-R modelling. An explicit bivariate LGP model exhibited the best overall behaviour for the study region, closely followed by MGGP. The lagged prediction has also been realized as the main drawback of the implemented FFBP algorithm.

İKLİM DEĞİŞİKLİNİN HAVZA ÖLÇEĞİNDE EKSTREM YAĞIŞLAR DEĞIŞKENLİĞİNE ETKİLERİ

ÖZET

İklim değişikliği bölgesel hidrolojiyi değiştirerek, insan hayatını birçok alanda etkilemektedir. İklim değişikliğinin etkileri, bir ülkenin sınırları içinde yer alan farklı bölgelerde önemli ölçüde değişkenlik gösterebilir. Türkiye, Akdeniz Bölgesi'nin bir parçası olarak, iklim değişikliğine karşı en hassas bölgeler içinde yer almaktadır. Bu doktora tez çalışmasında, Türkiye'nin kuzeydoğusunda yer alan Rize ili genelinde, iklim değişikliğinin ekstrem yağışlar üzerinde etkilerini lokal olarak anlamak amacıyla, farklı genel sirkülasyon modellerinin (GSM) grid tabanlı performans değerlendirmesi yapılarak, üç farklı gelecek zaman periyodu için (2013-2040, 2041-2070 ve 2071-2099) farklı iklim değişikliği senaryoları altında ekstrem yağışların değişkenliği incelenmiştir. Buna ek olarak, Rize ili içinde sınırlı yağış ve akış verisi bulunan bir akarsu için, veri esaslı projeksiyonlu bir yağış-akış modeli oluşturularak gelecekteki akımları tahmin edilmiştir.

Havza ölçeğinde iklim değişikliğinin hidrolojik değerlendirme öncesinde uygulanabilir bir performans değerlendirme yöntemi gereklidir. Tezin birinci bölümünde, Rize ili üzerinde yağış üretimi için iklim modellerinin belirsizlik derecesini belirlemek amacıyla grid tabanlı performans değerlendirme yaklaşımı sunulmuştur. Bu amaçla, RegCM3 bölgesel iklim modelinin (BİM), Rize ilini kapsayan beş hücresinin çıktıları kullanılmışır. BİM simülasyonu için başlangıç koşulları iki GSM (ECHAM5 ve CCSM3) tarafından temin edilmiştir. ECHAM5-RegCM3 ve CCSM3-RegCM3 kombinasyonlarının çıktılarından hesaplanan beş farklı yağış kaynaklı iklim istatistiği, 1961-1990 referans dönemi için ildeki gözlem verileri ile karşılaştırılmıştır. Bu istatistikler: (i) toplam aylık yağış (her ayın günlük yağışlarının uzun vadeli ortalamasının toplamı), (ii) aylık maksimum yağış (her aydaki uzun vadeli ortalama günlük yağışların maksimum değeri), (iii) yıllık toplam yağış (uzun vadeli ortalama günlük yağışların toplamı), (iv) yıllık maksimum yağış (her yıl için maksimum günlük yağış), ve (v) yağışlı gün sayısıdır. Her bir kombinasyonun performansı, dağılım şeması, hata, ortalama mutlak hata, hata kareler ortalamasının karekökü ve model performans indeksi (MPİ) kriterleri kullanarak test edilmiştir. Bu aşamanın sonuçlarına göre CCSM3-RegCM3 kombinasyonu gözlenmiş aylık toplam yağış verilerini gerçek değerlerin altında, ECHAM5-RegCM3 kombinasyonu ise gözlem değerlerinin üzerinde tahmin etmektedir. Maksimum aylık ve yıllık maksimum yağış üretimi açısından (özellikle kıyı alanlarda), ECHAM5-RegCM3 daha yüksek bir performans göstermektedir. Buna karşılık, CCSM3-RegCM3 yağışlı gün sayısını ECHAM5-RegCM3'e göre daha doğru (özellikle iç kısmında) tahmin etmektedir. Ayrıca, çalışmanın bu aşamasındaki sonuçları, bir havzanın belirli bir bölümü (istatistiği) için iyi bir performans gösteren GSM-BİM kombinasyonunun, aynı havzanın diğer bölümleri (istatistiği) için uygun olmayabileceğini göstermiştir. Ayrıca, MPİ kriterine göre

ECHAM5-RegCM3 kombinasyonu yıllık yağış üretimi için CCSM3-RegCM3 kombinasyonuna göre % 33'e kadar daha iyi sonuç vermiştir.

Tezin ikinci bölümünde, gelecekteki ekstrem yağışları tahmin etmek için hücre tabanlı bir metodoloji geliştirilmiştir. Bu metodoloji adaptif havza-ölçekli su kaynakları yönetimine uyumlu olması amacı ile bir orijinal hata düzeltme yöntemi ve ekstrem yağışları olasılığının tam tahmini için yeni bir karar kriteri ortaya koyulmuştur. SRES-A2 ve RCP8.5 senaryolarına göre ECHAM5-RegCM3 ve HadGEM2-RegCM4 çıktılarını kullanarak önerilen yöntemler Rize ili için uygulanmıştır. Bu amaçla, üç gelecek dönemdeki tahmini ekstrem yağışları ve 1961-1990 referans dönemindeki gözlemsel ekstrem yağışları arasında bir mukayese çalışması gerçekleştirilmiştir. Bu çalışmada, BİM hücrelerine dayalı düzeyde ekstrem değer teorisi kullanılarak, hem referans dönemi için, hem de gelecek yüz yıl için Rize iline ait 2, 10, 25, 50, 100 ve 200 yıllık yağışlar elde edilmiştir. Pratik açıdan bakıldığında, yerel yönetimler tarafından genellikle her bölge için en uygun dağılım modeli ve dağılım uygunluk testleri belirlenir. Türkiye'de, Meteoroloji Genel Müdürlüğü tarafından ekstrem değer teorisi kullanılması için gözlem yıllık maksimum yağış serisinin en az 10 yıllık kayıtlardan oluşturulması tavsiye edilmiştir. Ekstrem değer analizleri için, sağ çarpık olasılık dağılımları içeren: iki parametre lognormal (LN-2), üç parametre lognormal (LN-3), genelleştirilmiş ekstrem değer (GEV), Gumbel ve log-Pearson tip 3 (LP-3) dağılım fonkisyonlarının kullanılması önerilmiştir (http://www.mgm.gov.tr/). Dolayısıyla, bu tezde, 20. ve 21. yüzyılda ekstrem yağışları değişikliğini tahmin etmek amacıyla, bu olasılık dağılım fonksiyonları kullanılmıştır. Bahsedilen olasılık dağılımlarının matematiksel ayrıntıları bu bölümde detaylı şekilde anlatılmıştır. Dağılım fonksiyonlarının uygunluğunu test etmek için Kolmogorov–Smirnov (K-S) ve Anderson-Darling (A-D) testleri kullanılmıştır.

Bu aşamanın sonuçlarına göre (i) bu tezde önerilen hata düzeltme yöntemi, literatürde mevcut olan klasik yöntemler ile karşılaştırılabilir, (ii) tarihi ekstrem yağış olaylarının desenini en iyi veren dağılım modeli gelecekteki ekstrem olayları için etkili bir model olmayabilir. SRESA2 ve RCP8.5 sera gazı senaryoları göz önüne alındığında, ekstrem değeri analizi Rize ilinin bir çok bölgesinde gelecekteki tahmini tasarım fırtınalar büyüklüğünde önemli bir düşüş göstermektedir. Senaryoların ortalama sonuçlarına dayanarak, Rize ilinin çoğu bölgesinde (özellikle yakın (2013-2040) ve orta (2041-2070) gelecekte) tahmini ekstrem yağışların medyan değerinde 30% azalış vardır. Uzak gelecek (2071-2099) göz önüne alındığında, tahmini ekstrem yağışların medyan değerinde % 15 oranında azalış görülmüştür. Ayrıca, SRESA2 ve RCP8.5 sera gazı senaryolarını karşılaştırıldğında, HadGEM2-RegCM4 kombinasyonu günlük ekstrem yağış üretimi için ECHAM5-RegCM3 kombinasyonuna göre daha şiddetli fırtınalar vermiştir. Ayrıca, K-S ve A-D testlerinin sonuçlarına göre kullanılan dağılım fonksiyonları arasında, GEV modeli çalışma bölgesi için en iyi fonksiyon olarak belirlenmiştir.

Mevcut hidrolojik gözlemler ve gelecek klimatolojik projeksiyonlar, veri tabanlı modelleme teknolojisi desteği ile havza ölçekli hidrolojik modelleme çalışmalarına entegre edilebilirse hidrologlara büyük ölçüde yardımcı olacaktır. Bu tez çalışmasının son bölümünde, genetik programlamanın (GP) iki gelişmiş yaklaşımı (lineer GP - LGP ve multigen GP - MGGP), ve bir yapay sinir ağı yöntemi (feed-forward backpropagation - FFBP) geliştirilmiş ve uygulanmıştır. Modelleme

çalışmalarında projeksiyon yağış ve sıcaklık verileri ile birlikte gözlenmiş akım değerleri model girdileri olarak kullanılmıştır. Sonuç olarak, Rize İli havzalarının birinde (Şenöz Deresi) projeksiyonlu yağış-akış (PR-R) modelleme metodu ortaya konulmuştur. Önerilen modeller şu beş farklı performans kriteri açısından değerlendirilmiştir: (i) Nash-Sutcliffe verimliliği, (ii) sebat (persistence) indeksi, (iii) Kling-Gupta verimliliği, (iv) hacimsel verim, ve (v) hata kareler ortalamasının kareköküdür. Sonuçlar hem GP tabanlı çift değişkenli, hem de çok değişkenli PR-R modellerin umut verici rolünü göstermektedir. Bir belirgin çift değişkenli LGP modeli çalışma bölgesi için en iyi genel davranışı sergilemiştir. Ayrıca, gecikmeli tahmin, uygulanan FFBP algoritmanın ana dezavantajı olarak tespit edilmiştir.

1. INTRODUCTION

1.1 Motivation

There is no doubt that changing climate alters regional hydrology and thereby virtually affects every aspect of human well-being, from municipal and industrial water supply to agricultural productivity and wildlife management. The impact of climate change can vary significantly from a region to another region or even from a catchment to another catchment in the region of interest. Thus, analysis using catchment-scale climate projections and further watershed modelling are required to determine their effects on streamflow and flood hazards in different parts of a region. To assess and simulate impacts of such potential change, hydrologists require reliable meteorological variables for both current and future climate conditions. General circulation models (GCMs) provide such information, but their spatial scale is too coarse for regional impact studies. Thus, GCM outputs are commonly downscaled to a finer scale either through statistical downscaling or dynamic regional climate models (RCMs). After downscaling process, either a multi model approach (e.g., Van der Linden and Mitchell 2009, Najafi et al. 2011b) or an intercomparison strategy (e.g., vaze et al. 2011; Bozkurt et al. 2012) can be used to find out the best projections for a given catchment. In both, a catchment-scale performance evaluation of GCM driven RCM outputs at a certain period of time, commonly referred as reference period, would be crucial to realize inevitable biases of the ensemble or the best GCM-RCM combination for the study region.

Turkey is located in the eastern parts of the Mediterranean Region that identified as one of the most vulnerable regions to climate change by the fourth Assessment Report of Intergovernmental Panel on Climate Change (IPCC, 2007). Likewise, Giorgi (2006) defined the Mediterranean Region as one of the climate change "hotspots" in future projections. Therefore, in order to be ready to adapt possible hydrological change and to make wise decision, watershed managers need to know that what types of change will be occured at each catchment in the near an far future.

Thus, impact assessment of climate change in the catchment-scale baseline is an essential research topic for Turkey. With respect to the recently available RCM outputs in the country (Bozkurt et al. 2012; Demircan et al. 2014), this doctoral thesis aims to implement these downscaled outputs to assess the possible change in the rainfall intensity of Rize Province in catchment-scale. To the best of the author's knowledge, there is only a few research in the country linking regional climatological outcomes and a catchment-scale hydrological application.

1.2 Purpose and Scope of the Study

Considering the outputs of different GCMs recently downscaled over the country and historical climate conditions of Rize Province, the main (first) objective of this doctoral thesis is to identify extreme rainfall magnitudes over the province in 21th century. In order to achieve the goal, firstly, based upon rain gauge records in a distinguished period of time (i.e., reference period), a catchment-scale performance evaluation between the available GCM-RCM combinations is accomplished. Then, a new bias correction method is proposed to correct the selected GCM-RCM rainfall projections. Finally, using different probability distribution functions, extreme rainfall intensities are estimated for three future periods and compared to those of current period. It is hoped that the results will be useful for the prospective water resources projects in the province.

The other (second) objective of this research is to apply state-of-the-art genetic programming (GP) techniques to catchment-scale projected rainfall-runoff (PR-R) modelling in order to find improved models for future runoff prediction in poorly gauged catchments. A clear focus of this research is on making better use of the information contained in both observations and RCM model outputs. Two advancements of GP technique including multigene GP (MGGP) and Linear GP (LGP) are explored for this purpose. It is expected that the corresponding resuls would make a significant contribution in the hydrological modelling of poorly or ungauged catchments for the impact assessment studies.

1.3 Literature Summary

Research into the potential impact of climate change on regional hydrology has received extensive attention in hydrological researches. Different GCM outputs incorporating with different RCMs and the IPCC's greenhouse gas emission/concentration scenarios are commonly used in these studies. A number of regional climate simulation studies have focused on the different regions of Europe (e.g. Christensen et al. 2002; Ekström et al. 2005; Lionello et al. 2008; May 2008; Di Luca et al. 2011; Grossi et al. 2013), Africa (e.g. Afiesimama et al. 2006, Konare et al. 2008, Solmon et al. 2008, Sylla 2010) and Asia (e.g. Chen et al. 2003, Dash et al. 2006; Ozturk et al. 2012) in recent years. There are also some regional climate change simulations and assessments in Turkey (e.g. Krichak et al. 2007, Gao and Giorgi 2008, Önol and Semazzi 2009, Zanis et al. 2009, Ozkul 2009, Önol et al. 2014; Bozkurt and Sen 2013). Regional climate change simulations based on the SRESA2 scenario over the Eastern Mediterranean have been performed by Krichak et al. (2007) and Zanis et al. (2009) for the period 2071–2099. In these studies, the highest seasonal temperature increase for the entire region has been simulated in summer, in range of 4–7°C. The high-resolution RCM hindcast simulation of Önol (2014), which was forced with NCEP-Reanalysis over the Mediterranean–Black Sea domain, revealed significant warming trends in the summer temperatures during the period 1990–2008.

A broad performance evaluation of the GCMs driven RCM outputs using three GCMs including ECHAM5, CCSM3 and HadCM3 and one RCM (i.e., RegCM3) has been performed by Bozkurt et al. (2011). The authors utilized the dynamically downscaled outputs of three GCMs for the eastern Mediterranean–Black Sea region for the period of 1960–1990 and suggested that these GCMs can be used in the climate change impact assessment studies for the region. Following the mentioned study, Önol et al. (2014) examined the precipitation and temperature variations for daily and monthly time scales over the Mediterranean–Black Sea region during the twenty-first century based on IPCC-SRES A2, B1, and A1FI scenarios. A review of recent modelling strategies on the regional climate models impact studies at the catchment-scale was presented by Teutschbein and Seibert (2010). Fu et al. (2013) developed a score-based multi-criteria method to assess and rank the regional performance of 25 GCMs performance for the southeastern Australia. The results

indicated that two GCMs including CGCM3.1-T47 developed by Canadian Centre for Climatic Modelling and Analysis and INM:CM30 developed by Institute for Numerical Mathematics in Russia are the two best, whilst Beijing Climatic Centre CM1, Bjerknes Centre for Climatic Research BCM20, and Atmospheric Research CSIRO:MK35 resulted in relatively crude representations of monthly rainfall climate. Although some catchment-scale hydrological impact studies have been carried out in recent years (e.g., Graham et al. 2007a, b; Bastola et al. 2011; Mondal and Mujumdar 2012; Halmstad et al. 2013; Demirel et al. 2013), to the best of our knowledge, there is only a few studies in the relevant literature focused on the performance evaluation of GCM-RCM combinations in catchment-scale (e.g., van Vliet et al. 2012 and Deidda et al. 2013). More review on the methods that is applied in the current thesis will be present later in the following chapters.

1.4 Thesis Outline

The section gives a brief overview of this thesis. Based on the thesis objectives, the flow diagram of the study is presented in Figure 1.1. The study is completed in five chapters each of which is summerised as flows.

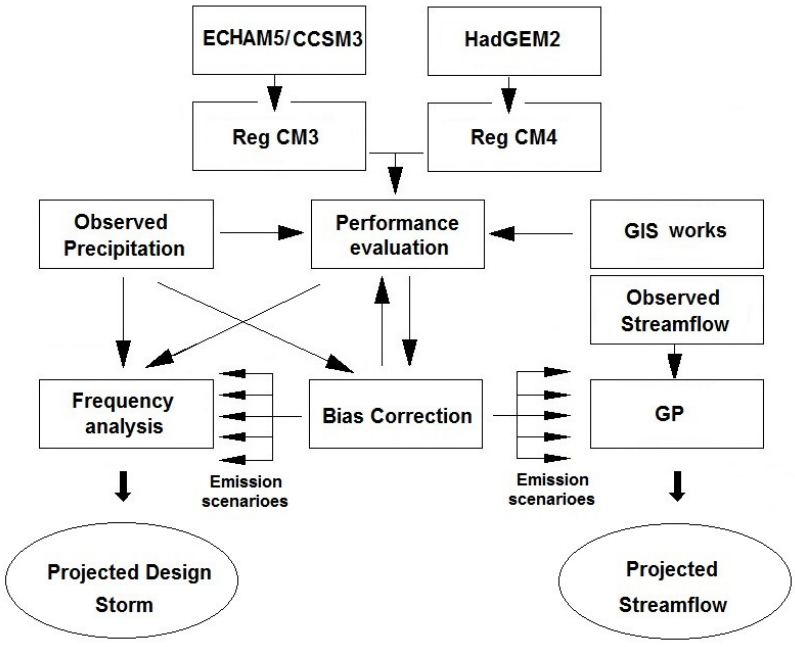

Figure 1.1 : Flow diagram of the thesis.

Chapter 1 presents a short introduction to the thesis, describes the purpose and scopes of the thesis, provides a summary of literature review, and outlines the chapters in the thesis.

Chapter 2 presents a grid-based performance evaluation of GCM-RCM combinations for rainfall reproduction comprising specific introduction and literature review, description of implemented climate models, rainfall climatology of the study area, performance evaluation criteria, and results and concluding remarks of the chapter.

Chapter 3 presents the application of extreme value theory in order to investigate the variability of extreme rainfall events over Rize Province concerning the high emission/concentration greenhouse gas scenarios. The chapter includes specific introduction and literature review, description of implemented models and study area, simulation and observational records, extreme value analysis at grid-based level, definition of future periods and bias correction procedure at each RCM grid-cells, results and discussions, and concluding remarks of the chapter.

Chapter 4 compares the capability of genetic programming (GP) and artificial neural networks approaches for projected Rainfall-Runoff Modelling in on of the catchments located in Rize province. The chapter contains specific introduction and literature review, overview of GP, LGP, MGGP, and feed- forward neural networks, study area and data, different one-day ahead forecasting scenarios, experimental setup and efficiency criteria, LGP, MGGP, and ANN modelling features, and results and concluding remarks of the chapter.

Chapter 5 presents a brief summary of the thesis and concluding highlights. In addition, the chapter provides some suggesstions for further studies based upon the limitations of theis research.

2. GRID-BASED PERFORMANCE EVALUATION OF GCM-RCM COMBINATIONS FOR RAINFALL REPRODUCTION

2.1 Introduction

According to the Intergovernmental Panel on Climate Change (IPCC), climate change is "a long-term, typically decades or longer, change in the state of the climate that can be identified by changes in the mean and/or the variability of its properties". It may be due to natural internal processes, external forcing, or to persistent anthropogenic changes in the composition of the atmosphere or land (IPCC 2007). Based upon significant effects of climate change on water resources (Özdoğan 2011; Groosi et al. 2013; Shen et al. 2014) and flood events (Kundzewicz et al. 2013) general circulation models (GCMs) are increasingly used as main sources of information required to hydrological impact studies specifically in the regions vulnerable to climate change.

GCMs are main tools to generate climate change projections. Progress in GCMs during the recent decades is vast. Their validation with observations for the 20th century has indicated that they are able to simulate the basic features of the climate of the Earth (Şen 2013). However, owing to the inherent uncertainty of GCMs (Foley 2010) they still suffer from spatial and temporal differences in their climate outputs, particularly in precipitation (Varis et al. 2004; Kundzewicz et al. 2008; Şen 2013). In addition, GCMs commonly have relatively low spatial resolution and therefore unable to resolve significant catchment-scale features, including topography and land use, as needed in hydrologic modelling and impact assessment (Moradkhani et al. 2010; Najafi et al. 2011b; Bozkurt et al. 2012). One way to make informed decision associated with a specific region is to obtain fine-scale regional information (i.e., downscaling) using one or more regional climate models (RCMs) that takes boundary conditions from different GCMs. Then either a multi model approach (e.g., Van der Linden and Mitchell 2009; Najafi et al. 2011a) or an intercomparison strategy (e.g., vaze et al. 2011; Bozkurt et al. 2012) can be used to find more appropriate combination(s). In both, a catchment-scale performance evaluation of

GCM driven RCM outputs at a certain period of time, commonly referred as reference period, would be crucial to realize inevitable biases of the ensemble or the best GCM-RCM combination.

In addition to intercomparison studies (e.g. Hofstadter and Bidegain 1997; Errasti et al. 2011; Barfus and Bernhofer 2014; Anandhi and Nanjundiah 2015), there are number of researches about the performance of GCM-RCM combinations (e.g., Jones and Raid 2001; Xuejie et al. 2002; Harvey and Wigley 2003; Jacob et al. 2007; Smith and Chandler 2010; Bozkurt et al. 2012; Roosmalen et al. 2010; Fu et al. 2013; Deidda et al. 2013). Xuejie et al. (2002) using a GCM-RCM combination model, CSIRO driven RegCM2, evaluated the performance of the combination for present-day reproducing and future change of different measures of extreme rainfall and temperature events in China. The authors reported that the CSIRO-RegCM2 combination can reproduce extreme events well in the country. Nieto and Puebla (2006) compared the observational and simulated precipitation variations over the Iberian Peninsula by using empirical orthogonal function and spectral analyses. Their results pointed out that the Geophysical Fluid Dynamics Laboratory Coupled Model (version 2) resulted in the best correlation with the observed annual rainfall cycle. Perkins et al. (2007) evaluated the performance of the IPCC's fourth assessment report models over Australia in terms of simulating daily maximum temperature, daily minimum temperature, and rainfall probability density functions (PDF). The MIROC-M, CSIRO, and ECHO-G models were recommended as more skilful ones for use in climate change impact assessments over the country. In a similar study, Maxino et al. (2008) reported that the CSIRO, IPSL, and MIROC-M are relatively well models to capture the PDFs of the maximum and minimum temperatures and rainfall over the Murray-Darling Basin in Australia. Fu et al. (2013) developed a score-based multi-criteria method to assess and rank the regional performance of 25 GCMs performance for the south eastern Australia. The results indicated that two GCMs including CGCM3.1-T47 developed by Canadian Centre for Climatic Modelling and Analysis and INM:CM30 developed by Institute for Numerical Mathematics in Russia are the two best models, whilst Beijing Climatic Centre CM1, Bjerknes Centre for Climatic Research BCM20, and Atmospheric Research CSIRO:MK35 resulted in relatively crude representations of monthly rainfall climate.

There are also some regional climate change impact assessments and intercomparison studies in Turkey (e.g., Önol and Semazzi 2009; Ozkul 2009; Önol et al. 2014; Bozkurt and Şen 2013). To the best of our knowledge, the only performance evaluation study for the country (so far) has been carried out by Bozkurt et al. (2012). Comparing three GCM-RCM combinations (i.e., ECHAM5-RegCM3, CCSM3-RegCM3, and HadCM3-RegCM3) with those of different sets of observations at eastern Mediterranean–Black Sea region for the period 1960–1990, the authors recommended that all the combinations can be considered for climate change impact assessment in the study region.

Considering spatial resolution degree of current GCMs or even GCM-RCM combinations, none of the abovementioned results obtained from either intercomparison or performance evaluation studies are directly applicable for catchment-scale floodplain engineering works. Additionally, most of the studies compare climate model outputs with reanalysis data that might not be reliable enough for performance evaluation at catchment-scale, specifically in the ungauged or poorly gauged catchments. The reason behind this is that not all reanalysis data are created by observations of the catchment of interest. Some data types, such as precipitation (depending on the reanalysis) are obtained by running (presumably newer) GCM or numerical weather prediction models. Addressing these issues, an applied grid-based GCM-RCM performance evaluation method as well as an intercomparison study between two GCM-RCM combinations are presented in this chapter. The results will be used for selecting scenario-based outputs of the more reliable GCM-RCM model as boundary conditions for extream rainfall analysis over the study area. In addition, in the present chapter, the author provids a grid-based overall performance analysis, whereas most of the previous intercomparison studies did not discuss the overall performance of climate models. Although some catchment-scale hydrological impact studies have been carried out in recent years (e.g., Graham et al. 2007a, b; Bastola et al. 2011; Mondal and Mujumdar 2012; Halmstad et al. 2013; Demirel et al. 2013), to the best of my knowledge, there is only a few studies in the relevant literature focused on the performance evaluation of GCM-RCM combinations in the catchment-scale (e.g., van Vliet et al. 2012 and Deidda et al. 2013). Since such studies make use of small areas to compare the performance of several GCM-RCMs for their ability to reproduce hydrological

variables (such as rainfall, temperature, and others), there is still need for more studies to establish a reference for hydrological applications in general. Information about the implemented models, methods, and observational data together with the obtained results are described in following sections.

2.2 Models Description

The performance analysis in this chapter is restricted to the two GCMs (i.e., ECHAM5 and CCSM3) integrated with an RCM namely RegCM3 due to the fact that they were only available models for the study region providing daily outputs at the time of analysis and therefore the current climate change impact assessment study for Rize Province is based on projections obtained by these combinations. Since the results are planned to be used as a basis for (i) selecting more reliable GCM-RCM combination, (ii) implementing/developing an appropriate bias correction method, and (iii) subsequent projected extreme rainfall modelling, five rainfall-borne statistics (are explained in Section 2.5), derived from GCM-RCM outputs have been considered for performance evaluation in this study.

RegCM3 is the third generation of the Regional Climate Model originally developed at the National Center for Atmospheric Research (NCAR) during the late 1980s and early 1990s. It is supported and maintained in the Earth System Physics (ESP) section of the Abdus salam International Centre for Theoretical Physics (ICTP) in Trieste, Italy. The model is comprised of a dynamical core, physics representing radiative transfer, large-scale or dynamic precipitation, convective precipitation, a planetary boundary layer component, a biosphere component, representation of open ocean and closed water bodies (lakes) and atmospheric chemistry and aerosols. These components are coupled and interactive.

The biosphere component of the model simulates surface processes related to vegetation (for example, evapotranspiration, leaf temperature, and phenology) and hydrology (for example, soil moisture, runoff, and snow) that vary in response to atmospheric conditions. The soil zone in the biosphere component has a depth of three meter, which is subdivided into a 0.1 m surface layer and a vegetation-dependent root zone that ranges from 1 to 2 m. Soil moisture and soil temperature are computed. the component thus provides a large set of variables that are either directly simulated by the model or that can be derived from the model outputs.

Detailed descriptions of physical parameterizations of RegCM3 and step-by-step instructions for downloading and running the model can be found in Pal et al. (2007) and ICTP network (http://users.ictp.it/~regcm/model.html).

The fifth-generation ECHAM model, ECHAM5, is a coupled atmosphere-ocean GCM developed at the Max Planck Institute for Meteorology, Germany. The model is evolved originally from the spectral weather prediction model of the European Centre for Medium Range Weather Forecasts (Simmons et al. 1989). It is used in a number of configurations which differ in the vertical extent of the atmosphere as well as the relevant processes. A detailed description of the ECHAM5 was presented by Roeckner et al. (2003). Compared to the its predecessor, ECHAM4, a number of significant changes have been introduced in both the numeric and physics of the model. These comprise a flux-form semi-Lagrangian transport scheme for positive definite variables like water components and chemical tracers, a new longwave radiation scheme, separate prognostic equations for cloud liquid water and cloud ice, a new cloud microphysical scheme and a prognostic-statistical cloud cover parameterization. The number of spectral intervals is increased in both the longwave and shortwave part of the spectrum. Changes have also been made in the representation of land surface processes, including an implicit coupling between the surface and the atmosphere, and in the representation of orographic drag forces. Also, a new dataset of land surface parameters has been compiled for the new model. On the other hand, horizontal and vertical diffusion, cumulus convection and also the spectral dynamics remain essentially unchanged.

Vertical structure of the model is described by a hybrid coordinate system. The sigma system at the lowest model level gradually transforms into a pressure system in the lower stratosphere. Troposphere lower stratosphere depth reaches to 10 hPa with a total of either 19 or 31 levels. In the middle atmosphere the computation level reaches to 0.01 hPa with either 39 or 90 levels. A truncated series of spherical harmonics is employed to represent vorticity, divergence, temperature and the logarithm of surface pressure in the horizontal (Bozkurt 2013). A flux-form semi-Lagrangian scheme (Lin and Rood, 1996) is used for passive tracer transport, i.e., for the water components (vapor, liquid, solid) and for chemical substances.

CCSM3 is the third generation of the Community Climate System Models that was released to the climate community in 2004. It is a coupled GCM with components

representing the atmosphere, ocean, sea ice, and land surface connected by a flux coupler. The model was designed to produce realistic simulations over a wide range of spatial resolutions, enabling inexpensive simulations lasting several millennia or detailed studies of continental-scale dynamics, variability, and climate change (Collins et al. 2006).

The main components of the CCSM3 system are based on: the Community Atmosphere Model version 3, the Community Land Surface Model version 3, the Community Sea Ice Model version 5, and the Parallel Ocean Program version 1.4.3. The three standard configurations of the model combine the low (T31), intermediate, (T42), and high (T85) resolution. Atmospheric component of CCSM3 is based upon the Eulerian spectral dynamical core with triangular spectral truncation at 31, 42, and 85 wavenumbers. The zonal resolution at the equator ranges from 3.75^{o} to 1.41^{o} for the T31 and T85 configurations. The vertical structure employs 26 levels with a hybrid terrainfollowing coordinate. The developed atmospheric core of the CCSM3 system includes signficant changes to the dynamics, cloud and precipitation processes, radiation processes, and treatment of aerosols. The cloud and precipitation processes contain separate prognostic treatments of liquid and ice condensate; advection, detrainment, and sedimentation of cloud condensate; and separate treatments of frozen and liquid precipitation (Boville et al., 2006). The code, documentation, input datasets, and model simulations are freely available from the model web site (online at http://www.ccsm.ucar.edu/models). Detailed descriptions of numerical parameterizations and physics of CCSM3 can be found in Collins et al. (2006).

In order to downscale the outputs of ECHAM5 and CCSM3 over Turkey, the relevant simulations were performed continuously from 1 January 1960 to 31 December 1990 at Istanbul Technical University through the project of United Nations Development Program entitled “Enhancing the Capacity of Turkey to Adapt to Climate Change”. Detail information about the project can be found in Bozkurt et al. (2012) and Önol et al. (2014). The reference period of GCM-RCM outputs used in this study has been selected identical to the present-day performance evaluation period (i.e., 1960–1990) implemented in the project. Considering the adjustment period (spin-up), we removed the first year outputs from the simulations’ archive before computing climate statistics.

2.3 Rainfall Climatology of Study Area

The study area is Rize Province, located in eastern Black Sea region between the Pontic Mountains and the Black Sea in Turkey, which is considered as the "wettest" corner of the country (Figure 2.1). The province with total area approximately around 3900 km^2 encompasses ten small to mid-size catchments. According to Koppen-Geiger climate classification, it has a borderline oceanic/humid subtropical climate similar to most of the eastern Black Sea region, with warm summers and cool winters. Annual precipitation averaging in the province is around 2500 mm with a maximum rate commonly in late autumn (October to December).

Rainfall climatology (i.e., long-term mean) of the study area at reference period 1961-1990 is investigated in two portions as inland and coastland areas by mainly considering the elevation differences from sea level that depict two primary precipitation regimes. With respect to the well documented significant effect of coastline on rainfall distribution over the coastal area of the eastern Black Sea region (Eris and Agiralioglu 2009), such a separation provides more accurate climatology for each portion due to elimination of error in the averaging procedure. A digital elevation model (DEM) including the borders of the study region and utilized rain gauges locations to derive the province climatology is presented in Figure 2.1.

Coastland is classified as the area where the elevation is lower than 1500m as inland corresponds to the area where the elevation is greater than 1500m, mainly constituting the headwaters of the catchments. As shown in the figure, precipitation data were obtained from 10 rain gauges. Half of them are operated by Turkish State Meteorological Service (TSMS) and the remains of stations are operated by Turkish State Hydraulic Works known as DSI in Turkey. DSI provides only monthly records. Thus, we used the relevant data only for estimating the monthly rainfall climatology of the study area. Homogeneity of the annual rainfall records in the eastern Black Sea region for the period of 1960-2005 has been already proved by Eris and Agiralioglu (2009). Rainfall climatology of each gauge for the reference period is presented in Table 2.1 that indicates notable annual rainfall difference between gauges located in the coastland and inland portions. In contrast to the generally accepted idea of the increase of precipitation with height, the orographic effects cannot be seen for the study area. It means that mean annual precipitation decrease with distance from sea.

Such a pattern was already shown for entire eastern Black Sea region by Eris and Agiralioglu (2009). The authors investigated the precipitation records of 33 rain gauge stations (46-year period) across entire eastern Black Sea region and demonstrated that mean annual precipitation decrease with elevation as well as distance from the sea.

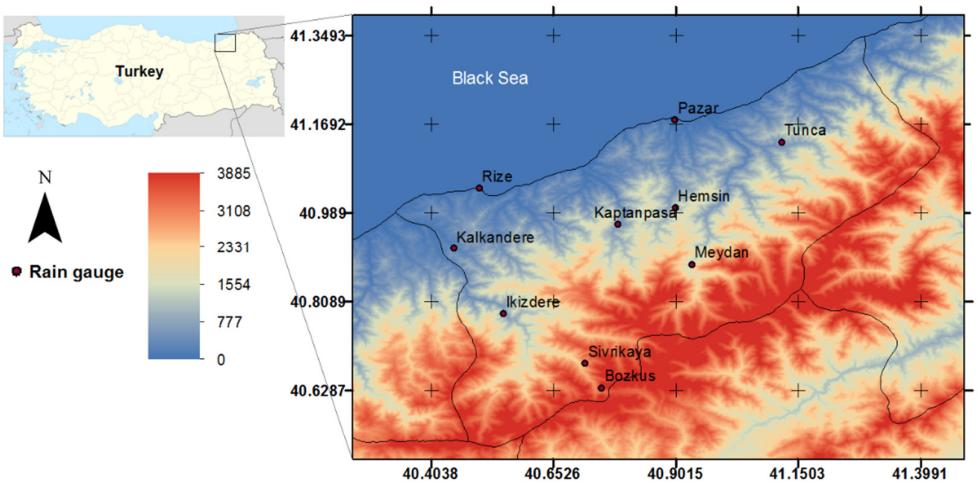

Figure 2.1 : Locations of the study area, Rize Province, and rain gauges used in the thesis.

In order to estimate areal distribution of rainfall climatology, two different averaging methods including arithmetic mean and isohyet lines have been used in this study and the results have been presented in Table 2.2. The Kriging optimal interpolation method (Kitanidis, 1997) was utilized to generate isohyet maps. For instance, isohyets produced for July to December have been presented in Figure 2.2. Every isohyetal map has a contour interval, which implies a difference between the two successive contour lines. This is more accurate than arithmetic mean (Şen 2008), thus, the results of isohyet lines method have been considered to compare with those of climate model outputs at each grid (i.e., grid-based performance evaluation) in this study. Five selected RegCM3 grids, providing good spatial coverage of the province, having the resolution of 27 km are also shown in the figure. In addition, with respect to the total area of grids overlaid on each portion, the area weighted average method has also been applied to portion-based performance evaluation.

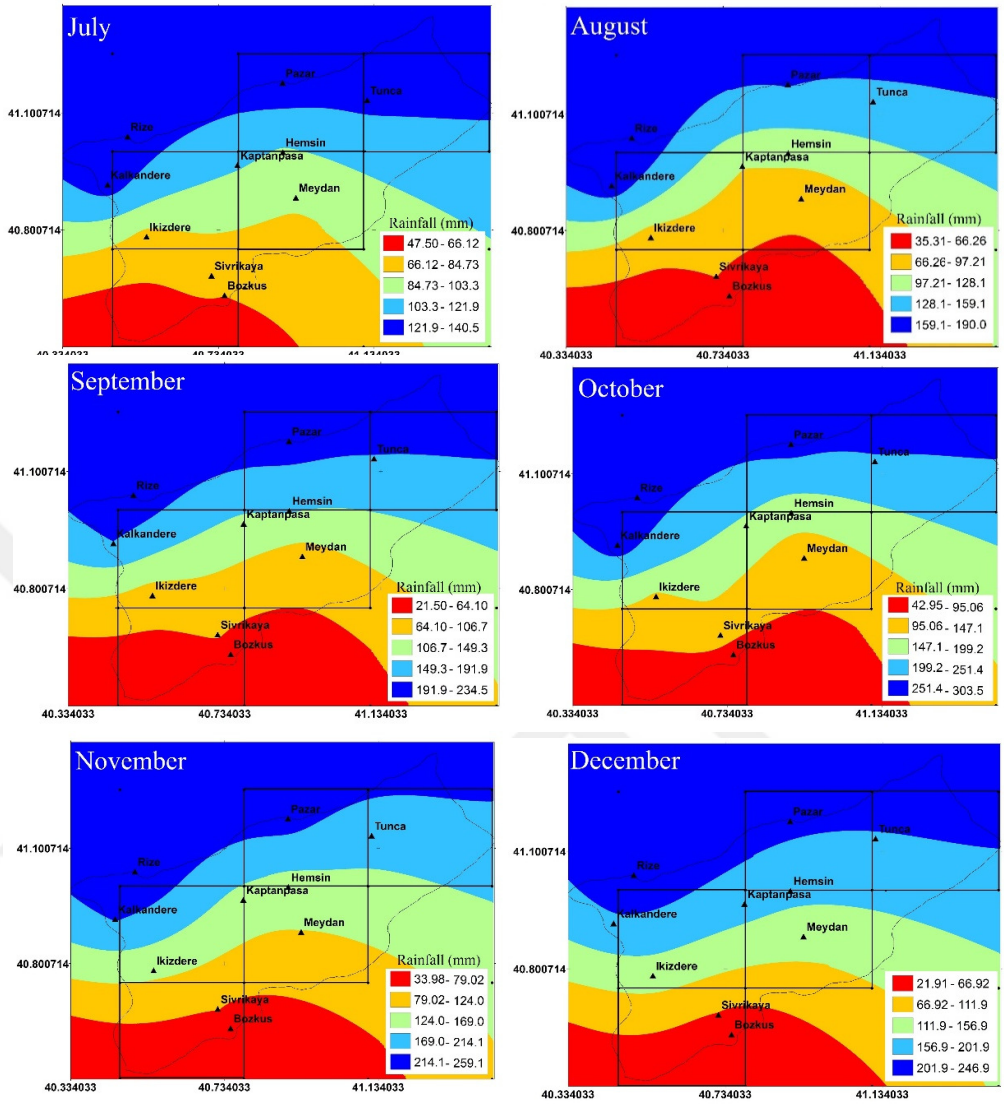

Figure 2.2 : Long-term mean of total monthly rainfall over Rize Province.

2.4 Evaluation Criteria

The relevant literature demonstrated that there is no agreement in the scientific community how to determine the best-performing climate models. It is up to the researchers which performance measure are used. In the present study, precipitation realization performance of the GCM-RCM combinations is analysed based on five measures comprising (i) scatterplots of simulation/observation, (ii) bias (simulation – observation), (iii) mean absolute bias (MAB), (iv) root mean squared error (RMSE), and (v) model performance index (MPI). The first four measures are applied only to individual climatic statistics whereas the last one (i.e. MPI) is applied to represent the overall performance of the model. The bias is the amount by which the value

reproduced by the model differs from the target observation (Liu et al. 2014). The MAB, equation (2.1), is a quantity used to measure how close model outputs are to the target values. The RMSE, equation (2.2), is a quadratic scoring rule, which measures the average magnitude of the error. The error is the amount by which the value reproduced by the model differs from the target observation.

$$\mathrm{MAB} = \frac{\sum_{i=1}^{n} \left| X_i^{sim} - X_i^{obs} \right|}{n} \tag{2.1}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (X_i^{obs} - X_i^{sim})^2}{n}} \tag{2.2}$$

where X_i^{obs} and X_i^{sim} are *i*th observed and simulated value of *X*, respectively and *n* is the total number of observations.

The MPI, firstly introduced by Reichler and Kim (2008), is an overall performance evaluation index, which measures the reliability of a model based on the scaled-normalized error variance (I^2) of a broad range of climate variables (v).

$$I_i^2 = e_{vi}^2 / \overline{e_{vi}^2} \tag{2.3}$$

$$e_{vi}^2 = \sum_{j=1}^{J} \left(w_i (\overline{M}_{vij} - \overline{O}_{vj}) / \sigma_{vj}^2 \right) \tag{2.4}$$

where e_{vi}^2 is normalized error variance of ith model; $\overline{e_{vi}^2}$ is the average error found in a reference ensemble of models; J is the number of grid cells in the catchment, M_{vij} and O_{vj} are the modelled and corresponding observed climatology; respectively; σ^2 is the inter-annual variance from the observations; w_i is proper weight needed for area and mass averaging model; subscripts i and j are the indices for each model and grid points, respectively and the overbar indicates averaging.

To rank climate models with respect to overall climatic statistics, the overall MPI is derived as the mean over the I^2 of all climate variables. This is based on the assumption that a reliable model should represent several components of the climate system of whole catchment.

$$MPI = \overline{I_i^2} \tag{2.5}$$

The *MPI* varies around one with values greater than one for underperforming models and values less than one for more accurate models. The expression $\overline{M} - \overline{O}$ in equation (2.4) indicates the error between simulation and observation. The smaller the error, the smaller the index, and the better the model. More details about MPI can be found in Reichler and Kim (2008).

2.5 Results and Discussion

In this section, ability of ECHAM5-RegCM3 and CCSM3-RegCM3 to reproduce five climatic statistics at each RCM grid as well as two different topographic locations in Rize Province are investigated. The climatic statistics include: (i) total monthly rainfall (i.e., sum of long-term mean of daily rainfall at each month; TMR), (ii) monthly maximum rainfall (i.e., the maximum value among long-term mean of daily rainfalls at each month; MMR), (iii) total annual rainfall (i.e., sum of long-term mean of daily rainfall; TAR), (iv) annual maximum rainfall (i.e., the maximum of daily rainfalls at each year; AMR), and (v) number of rainy days (NRD). The choice of each grid as performance evaluation zone is dictated by the type of observational data sets available at each grid. It is worth mentioning that simulation results at the reference period 1961-1990 composed of 10957 days (365×30 + 7 leap days) at each grid.

Table 2.1 : Monthly and annual rainfall climatology (mm) of utilized rain gauges.

Area	Rain gauge	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year
Coastland	Rize	223	170	147	101	102	125	131	189	218	274	254	236	2170
	Kalkandere	174	162	135	118	123	151	168	194	191	279	219	193	2107
	Pazar	182	142	111	81	83	152	142	164	210	270	235	211	1983
	Tunca	196	163	120	106	121	140	131	151	190	246	188	198	1950
Inland	Kaptanpasa	134	133	96	82	92	140	108	92	138	188	150	175	1528
	Hemsin	147	99	87	83	79	138	90	110	153	170	148	170	1474
	Meydan	121	79	73	81	112	99	79	70	88	117	123	132	1174
	Ikizdere	107	75	62	76	84	93	54	72	86	140	135	135	1119
	Sivrikaya	51	57	66	119	157	114	90	76	70	134	84	64	1082
	Bozkus	20	30	34	53	94	80	57	39	28	54	38	26	553

Table 2.2 : Climatology of Rize Province (total monthly precipitation, mm).

Area	Method	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year
Coastland	Isohyet	155	123	100	73	77	128	112	130	160	206	184	178	1625
	Arithmetic	181	145	118	97	99	144	130	156	191	247	204	204	1918
Inland	Isohyet	132	97	83	92	112	114	93	99	113	157	129	130	1351
	Arithmetic	68	61	60	85	120	95	68	61	62	103	87	80	950
Province	Isohyet	141	108	90	84	98	119	101	111	132	177	151	149	1461
	Arithmetic	134	110	94	92	108	123	104	117	137	187	156	153	1514

2.5.1 Monthly performance

It is conventional to evaluate simulation/prediction models first by scatterplots of simulation/observation, so that we will be able to realize the uncertainty features of simulation results. The scatterplots of the observed and modelled TMR and MMR data for the two portions and the entire province are provided in Figure 2.3. Significant differences between the results of different GCM-RCM combinations are clear in the figure. According to the Figure 2.3 (left), generally ECHAM5 tends to overestimate the TMR for the entire province while CCSM3 yields slightly dryer condition. In terms of MMR reproduction associated with each portion, it can be indicated that ECHAM5 resulted in higher performance than CCSM3, particularly for the inland portion (see Figure 2.3 (right)). To evaluate the efficiency of the GCM-RCM models more thoroughly, variations of the observed and modelled TMR and MMR values at each month for each portion are depicted in Figure 2.4. Differing largely from each other, both GCMs show same annual trends, which are more or less consistent with the observed climate trend. In contrast to ECHAM5, the CCSM3 underestimates the TMR amount in seasons other than winter and early spring. With respect to the MMR simulation results (see Figure 2.4(left)), both GCM-RCM combinations reproduce substantial biases in the coastland portion. Relatively good agreement with reference climatology is only provided by ECHAM5 in the inland portion. The coastland area in eastern Black Sea frequently receives more rainfall in the second half-year (July–November), whilst both ECHAM5 and CCSM3 simulated the opposite pattern of approximately first half-year rainfall (December–May) greater than the second half-year. Apart from February, March, and April, ECHAM5 outperforms CCSM3 at whole-year. It can be concluded that the major uncertainty in maximum monthly simulation for the province belongs to the coastland portion, which may originate from the local effects of Black Sea that is not well simulated by the climate models.

In order to provide profound discussion on simulation results, Table 2.3 presents quantitative errors of climate models for the both TMR and MMR statistics at each portion. Considering TMR, the highest bias value comes from ECHAM5 reproduction at the inland portion, whereas it interestingly reproduces the lowest bias for the coastland estimation. The spatial inconsistency of a climate model in fine

resolution reveals the fact that GSM-RCM intercomparison studies should be carried out in catchment-scale if the results are expected to be applied in hydrological design. The RMSE values of ECHAM5 and CCSM3 for the TMR estimation vary from 67mm to 188mm and 70mm to 93mm, respectively. According to the mean observed monthly rainfall for the entire province being equals to 122mm, the results indicated up to 154 and 76% uncertainty for ECHAM5 and CCSM3, respectively. Regarding the mean of observed MMR (equal to 9.0mm and 13mm at the inland and coastland portions, respectively; see Figure 2.4), the best simulation performance for reproducing MMR over the province belongs to ECHAM5 with 32% and 47% uncertainty at inland and coastland portions, respectively. Corresponding uncertainty values for CCSM3 results are 52% and 71%, respectively.

2.5.2 Annual performance

In order to identify uncertainty features of annual rainfall reproductions, three different statistics including TAR, AMR, and NRD are considered in this thesis. Owing to types of the data sets available in TSMS and DSI rain gauges (see Section 3), the annual performance is examined in grid-based level where at least one TSMS rain gauge is located in the grid of interest. In grid-based analysing, expected results are more reliable for different part of the province; however one can develop portion-based deduction using the mean of performance of grids laying at each portion. In the present study, we derived overall result based upon both arithmetic mean and MPI calculated at grid level. Areal distribution of TSMS and DSI rain gauges across the province as well as grids identification number adopted for the study have been displayed in Figure 2.5.

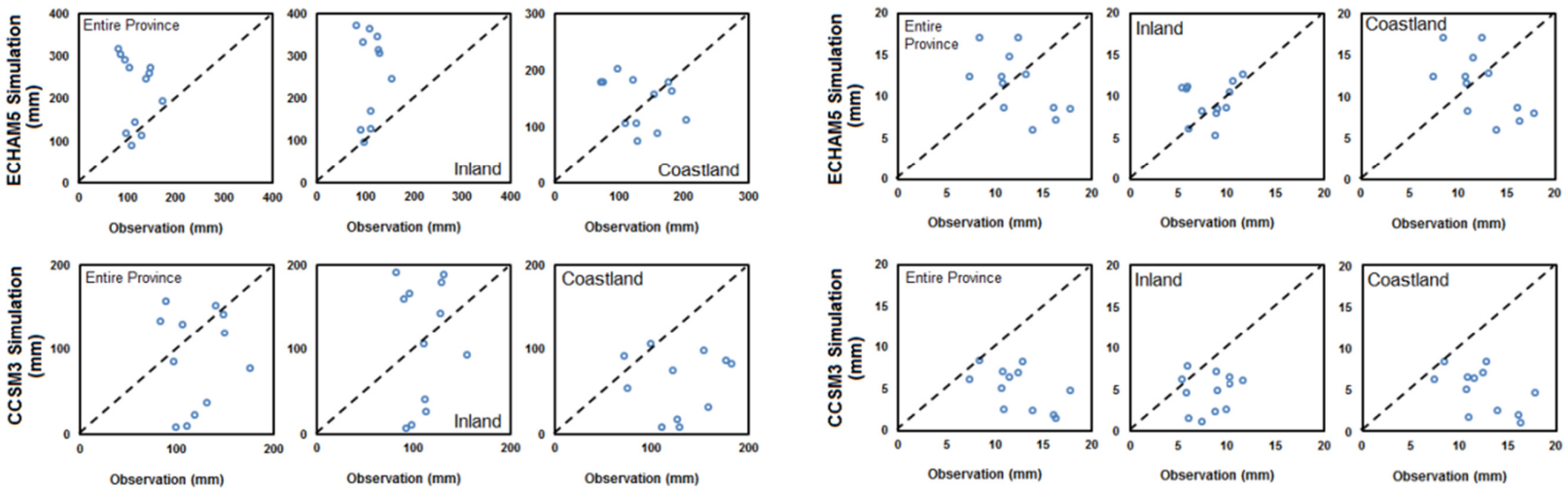

Figure 2.3 : Scatter plots of total monthly rainfall (left) and monthly maximum rainfall (right) at each portion and entire Rize Province for the period 1961-1990.

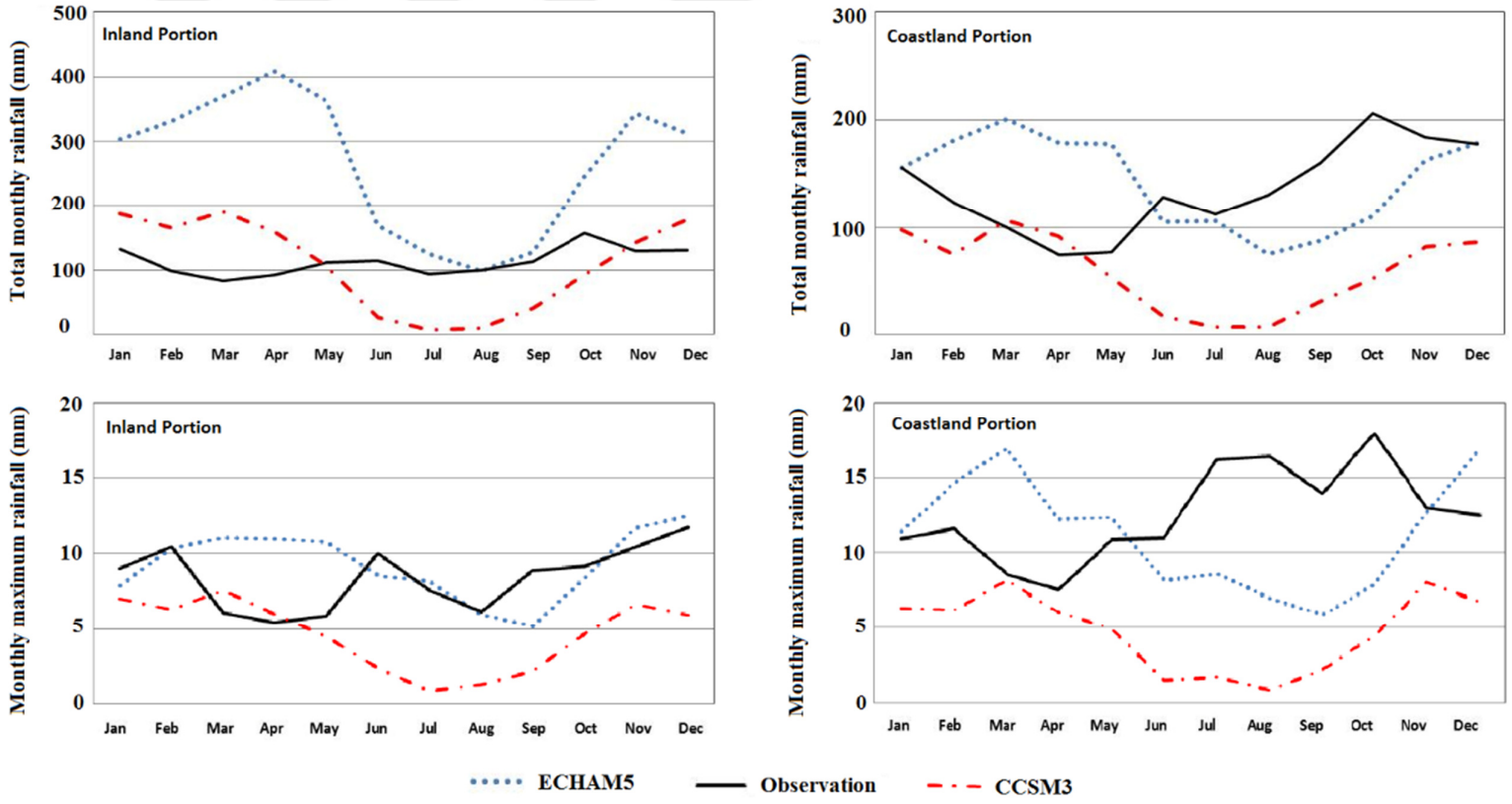

Figure 2.4 : Total monthly rainfall (top) and Monthly maximum rainfall (bottom) climatology of Rize Province for the period 1961-1990 of Rize Province.

Table 2.3 : Portion-based bias, MAB and RMSE of GCMs driven RCM simulations (mm).

Statistic	Model	Portion	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	MAB	RMSE
TMR	ECHAM5	Inland	172	234	288	317	251	55	31	-2.0	14	89	215	182	154	188
		Coastland	0.1	58	101	105	101	-23	-6.0	-55	-73	-95	-21	1.0	53	67
	CCSM3	Inland	56	68	108	-68	-6.0	-88	-87	-89	-73	-64	14	48	64	70
		Coastland	-58	-49	6.0	18	-24	-112	-105	-123	-130	-154	-102	-92	81	93
MMR	ECHAM5	Inland	-1.1	-0.1	5.1	5.6	4.9	-1.5	0.6	-0.2	-3.7	-0.7	1.2	0.8	2.1	2.9
		coastland	0.5	3.0	8.4	4.7	1.5	-2.9	-7.6	-9.4	-8.1	-10	-0.4	4.4	5.1	6.1
	CCSM3	Inland	-2.1	-4.2	1.6	0.6	-1.4	-7.6	-6.7	-4.8	-6.7	-4.5	-3.8	-5.8	4.1	4.7
		Coastland	-4.6	-5.5	-0.4	-1.5	-6.0	-9.5	-14.5	-15.6	-11.8	-13.5	-5.0	-5.8	7.8	9.2

Thirty-year scatterplots of the TAR and AMR for grids 2 and 4 as well as 23-year scatterplot for grids 3 (due to seven-year recording gap in Kaptanpasa station) have been presented in Figure 2.6. Because of the lack of daily precipitation records at grids 1 and 5, the comparison was only performed for grids 2 through 4. Highly scattered simulations of CCSM3-RegCM3 particularly in grids 2 and 4, which roughly represent the coastland portion, indicates the inability of the combination to simulate reliable TAR and AMR specifically in coastland portion. According to the figure, ECHAM5-RegCM3 provides wetter condition than CCSM3-RegCM3 in all grids. Dense scatter of ECHAM5-RegCM3 outputs around the perfect model diagonal line in all grids implies superiority of ECHAM5 to CCSM3 for both total annual and annual maximum rainfall simulation at Rize Province. Thus, projected AMR series produced by ECHAM5-RegCM3 may be more reliable to extreme rainfall analysis in the next chapter.

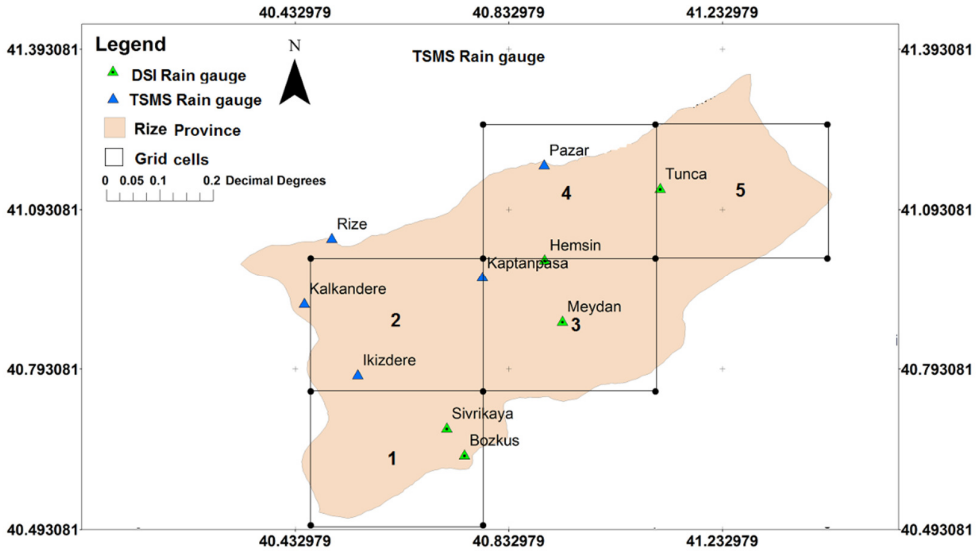

Figure 2.5 : Distribution of TSMS and DSI rain gauges in RegCM3 grids.

Owing to the importance of projected AMR series on extreme events, efficiency of the climate models is considered to supplementary evaluation using boxplot of the observed and modelled AMR series at the reference period (see Figure 2.7). The horizontal line inside the boxes represents the median and the upper and lower vertical lines, whiskers, indicate the 75% and 25% percentiles, respectively. The

figure depicts a positive spatial trend in both minimum and maximum values over the province, which indicates an increasing rainfall pattern from southwest to northeast of the study region. Both ECHAM5 and CCSM3 more or less follow this pattern in all grids, except in grid 1. The figure shows considerable differences between modelled and observational extremes. Only in grid 2, CCSM3 reproduced maximum value of AMR series (equals to 139 mm) and its event year (i.e., 1979) identical to those of observational data series interestingly. Considering the medians, the figure also implies that ECHAM5 reproduces closer values to the observed median than those of CCSM3.

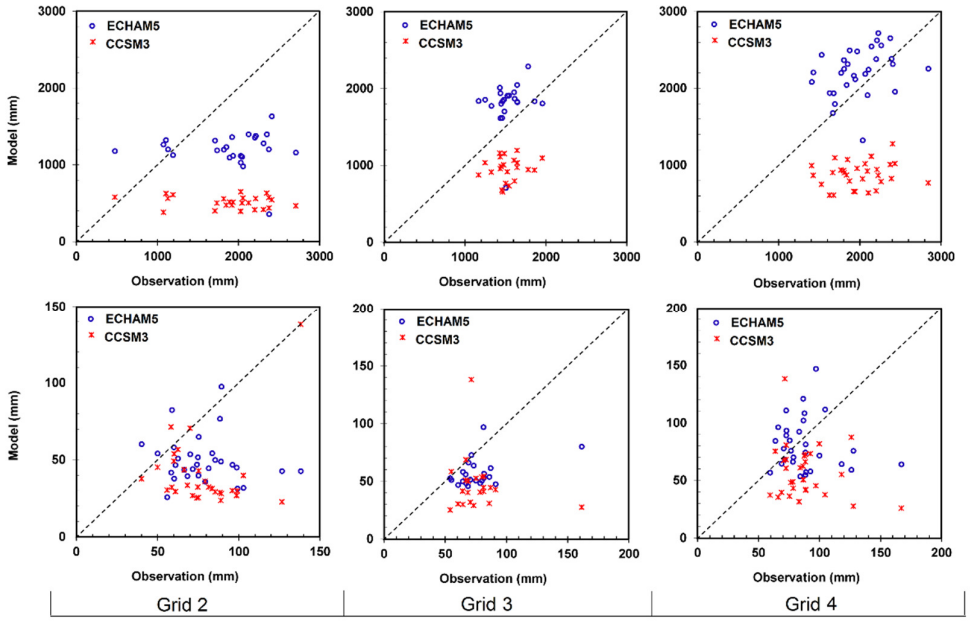

Figure 2.6 : Scatter plots of modelled and observational total annual rainfall (top) and annual maximum rainfall (bottom). Plots were given for the RegCM3 grid possessing at least one TSMS rain gauge.

As another annual climatic statistic for the grid-based intercomparison, Figure 2.8 compares the observed and simulated long-term average of NRD, defined here as days of rainfall in excess of 1mm. The figure demonstrates that the annual NRD in the entire province falls in the reasonably consistent range 129-135 days. With respect to ECHAM5 and CCSM3 reproductions, the rainy days are within the wide range 158-196 and 78-115 days, respectively. The figure also indicates that ECHAM5 tends to overestimate the observational NRD for the entire province in the

range 20%-52%. By contrast, CCSM3 underestimates in the range of 15%-42%. It is obvious from the figure that the lowest percentage error (-15.4%) of rainy days estimation belongs to CCSM3 and the highest one (51.9%) belongs to ECHAM5.

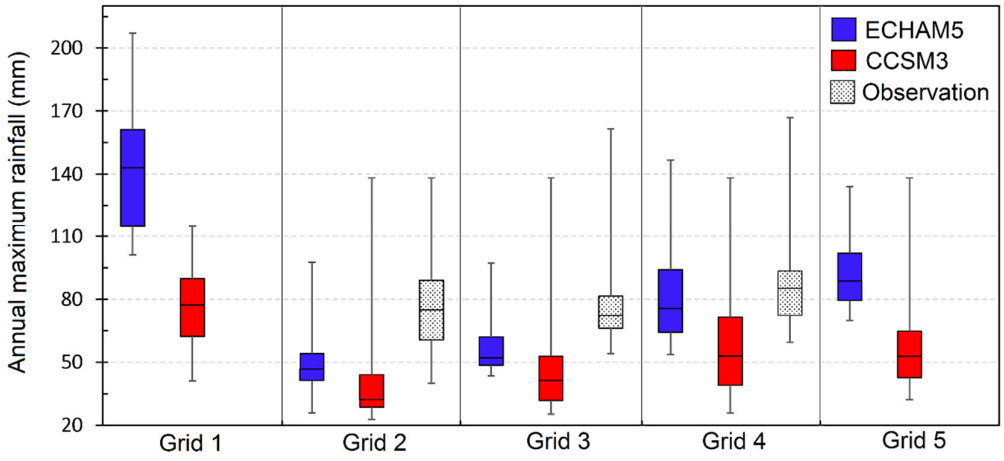

Figure 2.7 : Box and Whisker plots of modelled and observational annual maximum rainfall values at the RegCM3 grids for the period 1961-1990.

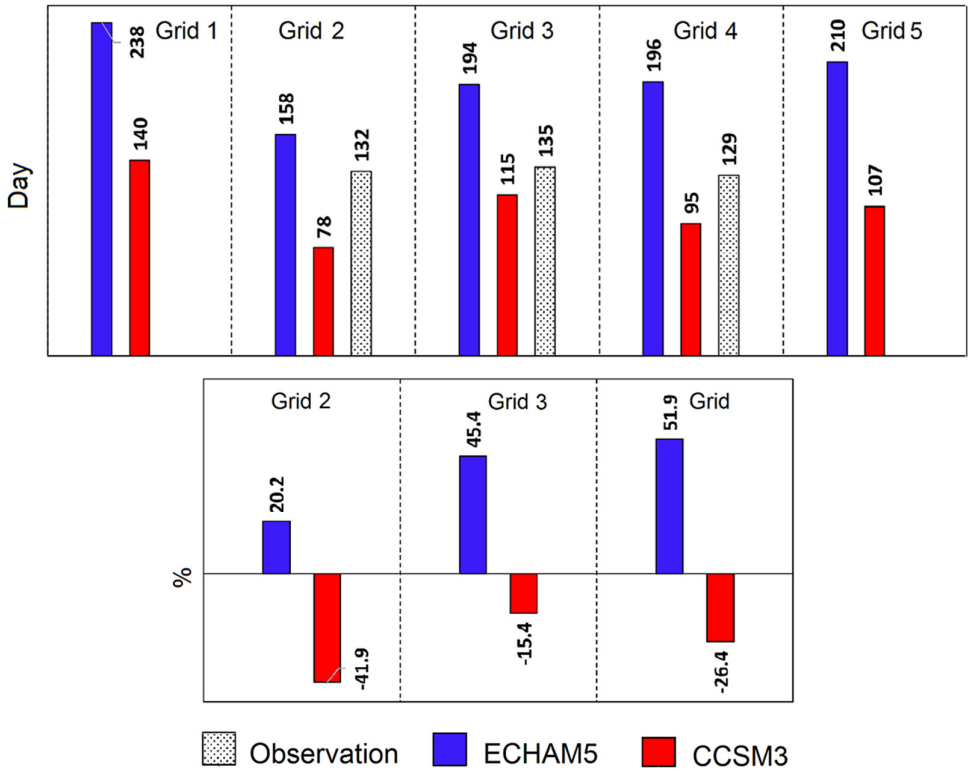

Figure 2.8 : Number of rainy days in each grid (Top, unit: days/year) and percentage of simulation error (Bottom).

The annual simulation results at each grid have been numerically compared in Table 2.4. As qualitatively discussed, the tabulated results verify the superiority of ECHAM5 to CCSM3 in terms of simulating both TAR and AMR. The averaged RMSE values of ECHAM5 and CCSM3 for TAR estimation are 578mm and 1081mm, respectively. Considering mean observed annual rainfall at grids 1-3 equals to 1800mm (not given in the table), the results indicated up to 32% and 60% uncertainty for ECHAM5 and CCSM3, respectively. In terms of simulating AMR, uncertainty value of ECHAM5 increases up to 42% whereas corresponding value at CCSM3 remains constant around 60%. With respect to the averaged error measures computed for rainy days simulation, it can be pointed out that CCSM3 provides more agreement than ECHAM5 with observed rainy days climatology.

Table 2.4 : Grid-based MAB and RMSE of GCMs driven RCM simulations.

Statistic	Model	Measure	Grid 2	Grid 3	Grid 4	Average
Total annual (mm)	ECHAM5	MAB	776	367	371	505
		RMSE	885	413	434	578
	CCSM3	MAB	1315	617	1110	1014
		RMSE	1466	616	1181	1081
Maximum annual (mm)	ECHAM5	MAB	32	30	25	29
		RMSE	42	27	34	34
	CCSM3	MAB	42	37	35	38
		RMSE	54	43	45	47
Rainy days	ECHAM5	MAB	31	61	68	53
		RMSE	34	63	71	56
	CCSM3	MAB	52	20	34	35
		RMSE	54	24	37	38

In addition to the implementation of simple arithmetic mean of grid errors, (Table 2.4, column 7), the MPI measure was also used in this study (Table 2.5) to investigate the overall performance of each model regarding the two annual statistics (i.e., TAR and AMR). Here, the author scaled the normalized error variance by the average error found in the ensemble of GCM-RCM combinations used in this study. According to the relative difference between MPI values given in the table, ECHAM5-RegCM3 is 33% more accurate than CCSM3-RegCM3 for annual rainfall reproduction. Referring to the average errors given in Table 2.4, one might claim that the superiority of the ECHAM5 implied from Table 2.5 is perhaps not so surprising. For this purpose, it should be mentioned that the averaging of grid-based errors to

judge about entire province is limited to statistics with same dimension whereas the MPI does not suffer from this drawback.

Table 2.5 : Overall comparison of the climate models.

Index	**Statistic (*v*)**	**ECHAM5**	**CCSM3**
e^2	TAR	12	13.4
	AMR	3.1	9.1
I^2	TAR	1.3	1.4
	AMR	0.3	1.0
MPI		0.8	1.2

3. EXTREME RAINFALL VARIABILITY

Introduction

In general, climate change has direct economic and social impacts on society, and thus has long been continuing to be the topic of investigations conducted in a wide spectrum of scientific communities ranging from climatology to public health (e.g., among others, Meral, et al., 2014 and 2015). Of course, many of the most direct impacts occur by means of the hydrologic cycle for which climate is a driving force. With respect to the increasing greenhouse gas concentration, prospective changes in mean and extreme weather conditions are expected (Seneviratne et al. 2012; Pinto et al. 2015). The effects of climate change on the frequency and intensity of extreme rainfall events have recently received great attention because of its important social, economic, and environmental consequences (Liuzzo and Freni 2015). Providing an accurate estimate of future changes to such extreme events is crucial for responsible decision making in flood risk management given the predictive uncertainties (Halmstad al. 2013; Notaro et al. 2015; Mirhosseini et al. 2015). A common way to detect these changes in extreme rainfall events is to use the extreme value theory (EVT) either in general GCM outputs or GCM driven RCM projections under different greenhouse gas emission/concentration scenarios. Owing to the different kind of uncertainties available in GCMs or GCM-RCM integrations (Dobler et al. 2012; Lespinas et al. 2014), a bias correction procedure is also suggested to adjust climate model outputs before applying in EVT analysis (Halmstad al. 2013; Mirhosseini et al. 2014). There is increasing interest in the hydrological community to evaluate the characteristics of projected extreme rainfall events based upon GCM-RCM integrations (e.g., Frei et al. 2006; Beniston et al. 2007; May 2008; Zhu 2013). For example, Semmler and Jacob (2004) applied the HadAM3H-REMO 5.1 combination and pointed out that in mountainous regions of Europe, the differences between simulated and observed return levels are larger than in flat regions. Possible

change in the return levels of one- to ten-day maximum rainfall amounts over Canada correspond to the special report on emissions scenarios-A2 (SRES-A2) was investigated by Mladjic et al. (2011) using Canadian Regional Climate Model. According to the results, northern Canadian climatic regions exhibit the lowest absolute but highest percentage change in 20-, 50-, and 100-yr return periods of rainfall extremes. Zhu (2013) studied climate projections of extreme rainfalls across the United States from the Hadley Regional Model version 3 (HRM3) nested in Hadley Centre Coupled Model version 3 (HadCM3). Simulation results based on the SRES-A2 scenario showed an increase in the intensity of extreme storms for most regions, with strong regional variations. Regarding the short-term extreme rainfalls over Ontario, Rudra et al. (2015) investigated change in the annual and monthly rainfall extremes and demonstrated a greater variability (increase and decrease) among stations for shorter durations (15, 30, and 60 minutes). Hosseinzadeh Talaee et al. (2015) investigated the extreme rainfall variations in Iran using the outputs of Couple Model Intercomparison Project 5 Phase. The authors showed that winter rainfall extremes are projected to decrease based on the representative concentration pathways-8.5 (RCP8.5) and to increase based on the RCP 4.5. More recently, Pinto et al. (2015) analyzed precipitation extremes across southern Africa using outputs of two RCMs forced by four GCMs. The EVT and non-parametric approaches were used by the authors to make a comparison between a possible future climate (2069–2098) under the RCPs 4.5 and 8.5 with present (1975–2005) climate conditions. They demonstrated that the total annual precipitation over southern Africa is projected to decrease while the maximum number of consecutive dry days and maximum 5-day precipitation amounts increase.

Although the abovementioned studies indicate that downscaled GCM outputs are frequently used to evaluate rainfall intensities in future climate, the application of these models has some limitations due to the uncertainty in the choice of climate models and future scenarios. In such cases, either an ensemble of climate models (scenarios) or alternative approaches such as stochastic weather generators or temporal analogues have been suggested. (Furrer and Katz 2008; Notaro et al. 2015).

In this chapter, the author investigate the spatial and temporal variability of rainfall extremes over Rize Province, Turkey, using the outputs of two GCM-RCM integrations including ECHAM5-RegCM3 and HadGEM2-RegCM4 forced by SRES-A2 and RCP8.5 scenarios, respectively. Both of these scenarios correspond to the comparatively high greenhouse gas emissions that assume a world with a continuously increasing population (Nakićenović et al. 2000; Riahi et al. 2011). The EVT analyses are used to estimate extreme rainfalls' magnitude and frequency at the both reference, 1961-1990, and three future periods 2013-2039, 2040-2070, and 2071-2100 determined using sliding-time window approach (see Waylen and Owusu, 2014). By the time of the analysis, this is the first study in the region that focuses on the comparison of projected extreme rainfall events under the both old (i.e., SRES-A2) and new (i.e., RCP8.5) greenhouse gas scenarios in catchment-scale. Additionally, the study proposes a new bias correction method that has a noteworthy importance of being applied in practice. Information about the applied models and observational data as well as the obtained results are presented in the following sections.

3.2 Models Description

As previously mentioned, the author implemented the historical and projected annual maximum rainfall (AMR) series extracted from the outputs of two GCM-RCM integrations including ECHAM5-RegCM3 and HadGEM2-RegCM4. The former was forced by SRES-A2 greenhouse gas emission scenario whereas the latter provides future climate under RCP8.5 scenario. The atmospheric simulations of ECHAM5-RegCM3 on 27km horizontal grid resolution and HadGEM2-RegCM4 on 20km resolution were performed through two different regional-scale climate change impact assessment projects entitled "Enhancing the Capacity of Turkey to Adapt to Climate Change" and "Climate Projections for in and around of Turkey", respectively. Detail information about these projects and relevant downscaling process are available in Bozkurt et al. (2012) and Demircan et al. (2014), respectively. Therefore, in this section the authors briefly describe only the models selected from these projects to use in the current study.

ECHAM5, the fifth-generation of ECHAM series, is a coupled atmosphere-ocean GCM developed at the Max Planck Institute for Meteorology, Germany. The model is evolved originally from the spectral weather prediction model of the European Centre for Medium Range Weather Forecasts. It is used in a number of configurations, which differ in the vertical extent of the atmosphere as well as the relevant processes. Compared to its predecessor, ECHAM4, a number of significant changes have been introduced in both the numeric and physics of the model. These comprise a flux-form semi-Lagrangian transport scheme for positive definite variables like water components and chemical tracers, a new longwave radiation scheme, separate prognostic equations for cloud liquid water and cloud ice, a new cloud microphysical scheme and a prognostic-statistical cloud cover parameterization. More information about ECHAM5 was given in Chapter 2. Detailed description of the model can be found in Roeckner et al. (2003).

HadGEM2, the Hadley Centre Global Environmental Model version 2, has been designed for the specific purpose of simulating and understanding the centennial scale evolution of climate including biogeochemical feedbacks (Collins et al. 2011). The HadGEM family of models encompass a range of specific model configurations incorporating different levels of complexity but with a common physical framework. They comprise a coupled atmosphere-ocean configuration, with or without a vertical extension in the atmosphere to include a well-resolved stratosphere, and an Earth-System configuration, which includes dynamic vegetation, ocean biology and atmospheric chemistry. Details on HadGEM2 can be found in Martin et al. (2011).

RegCM3 is the upgraded version of the Regional Climate Model system, originally developed at the National Centre for Atmospheric Research (NCAR). The first version of the model, RegCM1, was developed in 1989 and since then it has undergone major updates in 1993 (RegCM2), 1999 (RegCM2.5), 2006 (RegCM3) and most recently 2010 (RegCM4). The latest version of the model, RegCM4, is maintained in the Earth System Physics (ESP) section of the Abdus Salam International Centre for Theoretical Physics (ICTP) in Italy. These RCMs are limited area atmospheric models consisting of the basic equation, hydrostatic, compressible and sigma pressure levels, an ocean surface flux parameterization, a non-local

boundary layer scheme, an explicit moisture scheme, and a large-scale cloud and precipitation scheme with several options for cumulus convection scheme for the model simulations. Detailed descriptions of the RegCM family of models are available in Pal et al. (2007).

3.3 Extreme Rainfalls over the Study Region

As shown in Figure 3.1., the study region is Rize Province, Turkey, located between the Pontic Mountains and the Black Sea. According to Koppen-Geiger climate classification, the region has a borderline oceanic/humid subtropical climate with warm summers and cool winters. Annual rainfall averaging is around 2500-mm with a maximum rate commonly in late autumn (Danandeh Mehr and Kahya 2016). The study region with total area approximately equal to 3900-km^2 encompasses 12 rain gauges (see Figure 3.1). There is no evidence of annual maximum rainfall trends in historical series recorded in the rain gauges. Detail description of the gauges as well as homogeneity of annual rainfall records in the eastern Black Sea region for the period 1960-2005 has been already reported by Eris and Agiralioglu (2009).

The Rize Province, which is considered as "the wettest" corner of the country, is selected as the study region for analysis because of its topographical variability and tendency for flash floods. It characterized by rapid urbanization and increase of drained impervious area. Five RegCM3 grids with the spatial resolution of 27 km as well as 11 RegCM4 grids with the spatial resolution of 20 km covering the study area are also shown in the Figure 3.1. Increase of rainfall intensities due to the climate change may greatly increase flooding risk in the region. At the opposite, decrease in high extreme events might require additional sewer/stream maintenance.

The spatial distribution of extreme rainfall events with duration of 24 hours for the reference period 1961-1990 over the study region is shown in Figure 3.2. In this figure, the EVT (is explained later in Section 3.4) analysis using the observational annual maximum rainfall (AMR_{obs}) measures at each rain gauge is applied to map extreme rainfall events associated with 2-, 10-, 25-, 50-, 100-, and 200-yr recurrence intervals (hereafter design storms) typically used in hydrological design. It is worth mentioning that the results of EVT analysis have the characteristic of point

measurement. Hence, nonlinear co-kriging technique (Azimi-Zonooz et al. 1989) is implemented for areal rainfall estimation. It is clear from the figure that intense storms commonly occur in northeastern part of the province.

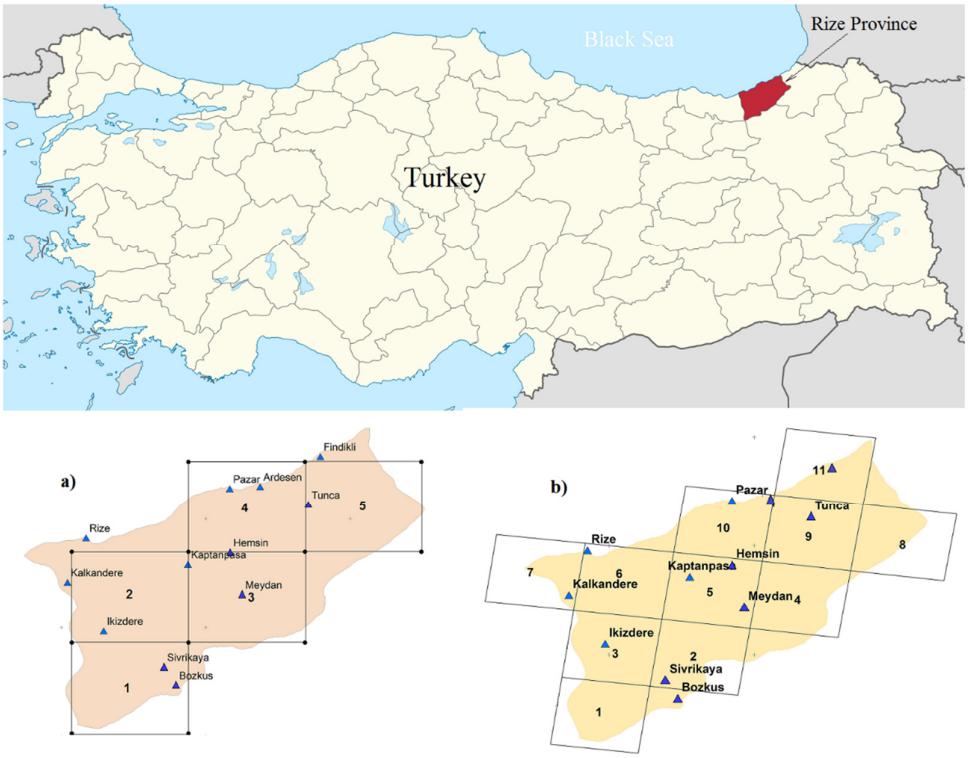

Figure 3.1 : Location of RCM grids within the Study area; a) RegCM3 and b) RegCM4.

3.4 Extreme Value Theory (EVT)

In order to study the extreme events at the both reference and future climate conditions, a parametric approach based on the Extreme Value Theory (EVT) is used in this thesis. Intensity and frequency of rare events in EVT, are generally estimated by fitting an appropriate distribution model to annual maximum series (Coles 2001; El Adlouni et al. 2008). In this study, five right-skewed distribution models, namely two-parameter lognormal (LN-2), three-parameter lognormal (LN-3), generalized extreme value (GEV), Gumbel, and log-Pearson type 3 (LP-3) are compared to describe the likelihood of extreme rainfall events at the selected duration (1 day) and

recurrence intervals (i.e., 2-, 10-, 25-, 50-, 100-, and 200-yr) for the both historical and scenario simulations.

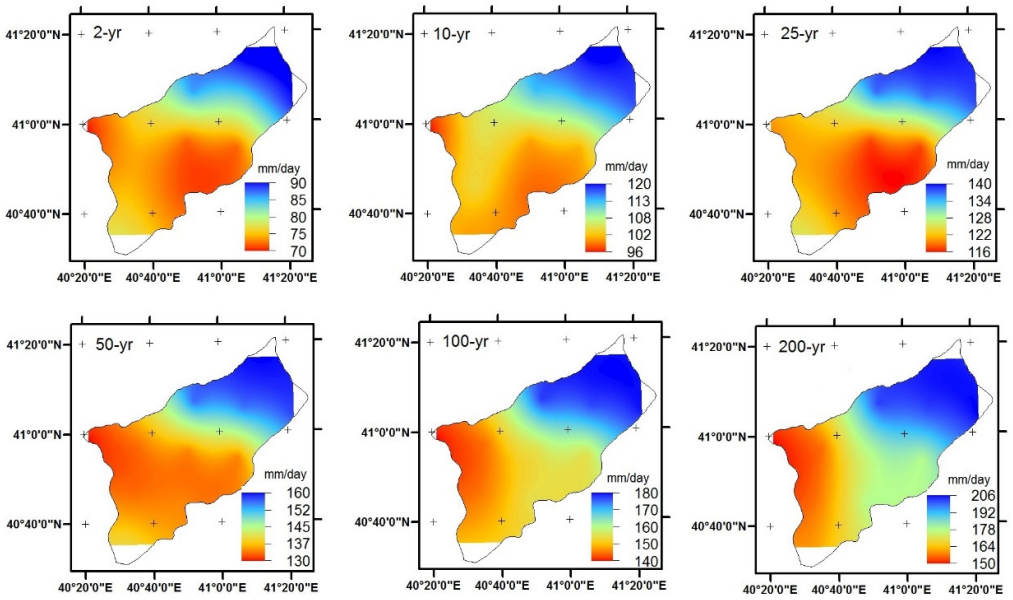

Figure 3.2 : Areal distribution of extreme rainfalls over the study area at the reference period 1961-1990.

The lognormal distribution is a probability distribution whose logarithm has a normal distribution. It is applicable when the quantity of interest must be positive, since log(x) exists only when x is positive. The probability density function of the three parameter lognormal distribution for aseries of random variable x (here AMR series) with parameters μ, σ, and γ is given by equation (3.1). The γ is continuous location parameter and $\gamma \equiv 0$ yields the two-parameter Lognormal distribution (equation (3.2)).

The location parameter describes the shift of a distribution in a given direction on the horizontal axis.

$$f(x) = \frac{\exp\left(-\frac{1}{2}\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right)^2\right)}{(x-\gamma)\sigma\sqrt{2\pi}} \qquad (3.1)$$

$$f(x) = \frac{\exp\left(-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right)}{x\sigma\sqrt{2\pi}} \tag{3.2}$$

The GEV distribution is a family of continuous probability distributions that combines the type I, II and III extreme value distributions (equation (3.3)). Among different models available in the literature the GEV and its special form, the Gumbel (EV1; equation (3.4)) are widely used in extreme rainfall modelling (El Adlouni and Ouarda 2010; Willems et al. 2012; Liuzzo and Freni 2015; Pinto et al. 2015).

$$f(x) = \begin{cases} \frac{1}{\sigma}\exp\left(-(1+kz)^{-\frac{1}{k}}\right)(1+kz)^{-1-1/k} & k \neq 0 \\ \frac{1}{\sigma}\exp\left(-z-\exp(-z)\right) & k = 0 \end{cases} \tag{3.3}$$

$$f(x) = \frac{1}{\sigma}\exp\left(-z-\exp(-z)\right) \quad k = 0 \tag{3.4}$$

Where, the k, σ, and μ are shape, scale, and location parameters, respectively. The scale parameter describes how spread out the distribution is, and defines where the bulk of the distribution lies. As the scale parameter increases, the distribution will become more spread out (Millington et al. 2011). The shape parameter is derived from skewness, as it represents where the majority of the data lies, which creates the tail(s) of the distribution. Using these parameters the parameter z can be obtained by

$$z = \frac{x - \mu}{\sigma} \tag{3.5}$$

As given in equation (3.4), when the shape parameter k equals to 0, this is the Gumbel distribution. When $k>0$, this is EV2 (frechet), and when $k<0$ is the EV3 (Weibull).

The LP-3 distribution has been one of the most frequently used distributions for hydrologic frequency analyses since the recommendation of the Water Resources Council (1967, 1982) of the United States as to its use as the base method (Singh 1998).

$$f(x) = \frac{1}{\beta x \Gamma(\alpha)} \left(\frac{\ln x - \gamma}{\beta} \right)^{\alpha - 1} \exp\left(- \frac{\ln x - \gamma}{\beta} \right) \quad (3.5)$$

Where $\alpha>0$, $\beta>0$, and $0<\gamma<\ln x$ are the shape, scale, and location parameters, respectively. The $\Gamma(.)$ represent the Gamma function.

All the distribution parameters are estimated by maximum likelihood estimates method involving the maximization of the log-likelihood function. The best distribution among the models are selected based upon two explanatory goodness-of-fit tests including Kolmogorov–Smirnov (K-S) and Anderson-Darling (A-D). The K-S is a nonparametric hypothesis test that shows how well a theoretical distribution fits to our data. It is distribution-free with no dependency on the sample size. The null and the alternative hypotheses are: the data follow the specified distribution and the data do not follow the specified distribution, respectively. The A-D is a modification of the K-S test (Stephens 1974) and gives more weight to the tails than does the K-S test (Ahmad et al. 1988). The critical values for the A-D test are dependent on the specific distribution that is being tested. More information for the K-S and A-D tests as well as an example of their application are provided in Appendixes A and B, respectively.

3.5 Bias Correction

The term "bias correction" describes the post-processing process of re-scaling climate model outputs to reduce the effects of systematic errors in the model (Teutschbein and Seibert, 2012). Initial conditions as well as resolution of different climate models are commonly identified causes of biases. Although correction of RCM outputs significantly improves hydrological simulations (Shrestha et al. 2011; Berg et al. 2012; Shrestha et al. 2016), there is a major shortcoming; all bias correction methods rely on the assumption of stationary model errors. This implies that the correction algorithm for reference climate condition is assumed to also be valid for future climate.

Several bias correction methods have been developed to adjust inevitable bias of meteorological variables from GCM, RCM, or global reanalysis products, ranging

from the simple linear scaling approach to sophisticated distribution mapping (Teutschbein and Seibert, 2012; Fang et al. 2015; Seyyedi et al. 2015; Demirel and Moradkhani 2016; Shrestha et al. 2016). An overview of the literature shows that available methods generally use daily GCM/RCM outputs (e.g., Piani et al. 2010; Johnson and Sharma, 2011; Najafi et al., 2011; Berg et al. 2012; Halmstad et al. 2013; Shrestha et al. 2016) that are not necessarily required when the GCM/RCM outputs are used to reproduce the specific climatic statistic of historical events. In such cases, pattern oriented bias correction has been applied to reproduce the specific patterns of historical events rather than an attempt to ‘‘correct’’ day-by-day GCM/RCM outputs (e.g., Dettinger et al. 2004; Zhu 2013; Nzemosadat et al. 2016). In order to investigate extreme rainfall variability, it would be very useful for hydrologists if a bias correction method perfectly adjusts the extreme patterns. Addressing this issue, this study presents a new bias correction methodology that aims to match the annual mean of simulated AMR series with that of observed ones. To this end, a scaling factor is developed between the long-term mean of AMR_{obs} and corresponding simulated values over a reference (control run) period at each grid and then is applied for the projected (future) period (equation (3.6)) to produce future predictions of AMR series at given grid.

$$AMR_{BC}^{i} = AMR_{sim}^{i} \times \frac{\overline{AMR}_{obs}^{i}}{\overline{AMR}_{sim}^{i}} \tag{3.6}$$

where, AMR_{BC}^{i} is adjusted annual rainfall maxima at the grid number “i” and the overbar symbol indicates long-term annual mean. By definition, the scaling factor (hereafter, bias correction factor (BCF)) varies among grids with the perfect value of unity for the best control run simulation. Values more than unity indicate underestimated simulations, while values less than unity show overestimations. It is worth mentioning that the underlying concept of the proposed methodology is to some extent similar to the well-known linear scaling (Lenderink et al. 2007) approach. Using daily outputs, the linear scaling approach perfectly matches the monthly mean of model and observed data; however the new methodology use

annual maximum series to match annual mean of observed and simulated AMR series.

3.6 Efficiency Tests for the New Bias Correction Methodology

Following Fang et al. (2015), two statistical methods are used to test the efficiency of the new bias correction methodology in this study. The first one is frequency-based performance that deals with the ability of the methodology to reproduce the frequency-based statistics of observed AMR series at each grid. The second method quantifies accuracy of the proposed methodology using time series-based performance indicators. The frequency-based statistics include mean, median, standard deviation, 25th percentile, and 75th percentile whereas the time series-based metrics comprise root mean squared errors (RMSE), Nash–Sutcliffe efficiency (NSE), percentage bias (PBIAS), and ratio of RMSE to standard deviation (RSR). Such classification among goodness-of-fitness tests has been previously utilized by Fang et al. (2015) to compare the efficiency of six bias correction methods in correcting RCM-simulated rainfall and temperature data in China.

The RMSE, equation (see equation 2.2) shows the average magnitude of the error. The PBIAS given in equation (3.7), which is recommended in recent studies (e.g., Fang et al. 2015; Shrestha et al. 2016), measures the average tendency of RCM outputs to their observed counterparts. The positive and negative results represent the model underestimation and overestimation, respectively; hence, its optimum value is zero. The NSE (equation (3.8); Nash and Sutcliffe 1970) is a normalized statistic determining the relative magnitude of estimation error in comparison with the measured data variance). It indicates how well the plot of observed data versus predicted data fits the 1:1 line. It ranges between $(-\infty, 1]$ with optimum value one for the perfect model. The RSR, equation (3.9), is a normalized statistic with optimum value zero (Singh et al. 2004). The higher the RSR values, the lower the performance.

$$\text{PBIAS} = \frac{\sum_{i=1}^{n}(X_i^{obs} - X_i^{sim})}{\sum_{i=1}^{n}(X_i^{obs})} \times 100 \tag{3.7}$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n}(X_i^{obs} - X_i^{sim})^2}{\sum_{i=1}^{n}(X_i^{obs} - \text{AVE}_{obs})^2} \tag{3.8}$$

$$\text{RSR} = \frac{\text{RMSE}}{\text{STDEV}_{obs}} \tag{3.9}$$

where X_i^{obs} and X_i^{sim} are *i*th observed and simulated value of *X* (here Rainfall), respectively. AVE_{obs} and STDEV_{obs} are mean value and standard deviation of observed data sets, respectively and *n* is the total number of observations (equal to 30 in this study).

3.7 Results and Discussion

3.7.1 Control run evaluation

Table 3.1 shows the BCF values computed by the long-term mean of observed and simulated AMR series at five RegCM3 grid cells (i.e., evaluation grids). They are stronger in grid#2 and grid#3 for ECHAM5-RegCM3 simulations because of the lower magnitude of the model outputs in these grids, whereas grid #5 reveals the highest BCF for HadGEM2-RegCM4 probably due to different climate prediction mechanisms between the models. In all the grids, except in grid#1, the BCF values are more than unity. This implies that the climate models tend to underestimate extreme rainfall events over the study region. In order to adjust the historic and future simulated AMR series, the BCF given for each grid is multiplied by each value of the AMR_{sim} series for the same grid. Given that the corrections are derived solely with a constant BCF at each grid, all events are adjusted with the same correction factor.

Table 3.1 : Results from bias Bias correction factors obtained for the ECHAM5-RegCM3and HadGEM2-RegCM4 simulations over the historic period 1961–1990.

Grid No.	Observations	ECHAM5-RegCM3		HadGEM2-RegCM4	
	$\overline{AMR_{obs}}$	$\overline{AMR_{sim}}$	BCF	$\overline{AMR_{sim}}$	BCF
1	80.0	142.7	0.56	68.9	1.16
2	78.0	49.7	1.57	68.9	1.13
3	76.8	57.3	1.34	61.5	1.25
4	88.0	81.1	1.08	68.2	1.29
5	97.0	91.2	1.06	68.2	1.42

Table 3.2 lists the frequency-based statistics of observed, raw and bias corrected RCM-simulated AMR series at the evaluation grids. Compared to the observations, the raw simulations deviate significantly from the observations, with underestimation of all the statistics in all the grids, except in grid#1. The table indicates that the new bias correction methodology not only adjusts the mean but also improves the median perfectly. The method also provides good estimations in the 25th and 75th percentiles; however, there is a slight discrepancy in the standard deviation. This discrepancy is partly attributed to the RCM grids possessing a perfect estimation for standard deviation (e.g., ECHAM5-RegCM3 outputs at grid#4).

Time series-based performances are evaluated at the evaluation grids and the results are summarized in Table 3.3. The worst performance belongs to ECHAM5-RegCM3 at grid#1 with RMSE=69.1mm, PBIAS=-78.28%, NSE=-13.71, and RSR=3.84. In contrast, the best performance is generated by HadGEM2-RegCM4 at gtrid#3 with RMSE=25.73mm, PBIAS=8.43%, NSE=0.65, and RSR=1.23. As indicated by the table, the new bias correction methodology more or less improves all the statistics at all the grids. For instance, the worst and the best NSE values improves up to -0.27 and 0.79, respectively. Regardless of the worst PBIAS value obtained by the raw outputs of ECHAM5-RegCM3 at grid#1, this statistic indicates that the raw RCM simulations were underestimated with varying PBIAS in the rate 6% - 36%. The tabulated results also show that the PBIAS values were significantly improved after implementing the bias correction procedure; however it is not a perfect correction method for the RMSE and RSR metrics.

Table 3.2 : Frequency-based statistics of observed and simulated annual maxima series (mm/day) at the RegCM3 grid cells over the historic period 1961–1990.

Grid No.	Data set		Mean	Median	Standard deviation	25th percentile	75th percentile
1	Observed		80.1	80.0	18.0	62.1	90.0
	ECHAM5-RegCM3	Raw	142.7	143	28.7	114.8	162.6
		Bias Corrected	80.1	80.2	16.1	64.4	91.2
	HadGEM2-RegCM4	Raw	68.9	66.9	10.9	59.5	76.7
		Bias Corrected	80.1	77.7	12.6	69.2	89.2
2	Observed		78.0	74.9	21.0	60.7	89.1
	ECHAM5-RegCM3	Raw	49.7	46.8	14.9	41.3	53.9
		Bias Corrected	78.0	73.4	23.3	64.7	84.6
	HadGEM2-RegCM4	Raw	68.9	66.9	10.9	59.5	76.7
		Bias Corrected	78.0	75.7	12.3	67.4	86.8
3	Observed		76.8	72.3	20.9	66.2	81.5
	ECHAM5-RegCM3	Raw	57.3	52.1	13.2	48.5	62.6
		Bias Corrected	76.8	69.7	17.7	65.0	83.8
	HadGEM2-RegCM4	Raw	61.5	57.5	21.9	47.6	65.0
		Bias Corrected	76.8	71.7	27.3	59.4	81.1
4	Observed		88.0	85.2	22.1	72.4	94.6
	ECHAM5-RegCM3	Raw	81.1	75.6	22.1	64.2	92.9
		Bias Corrected	88.0	82.0	24.0	69.6	100.7
	HadGEM2-RegCM4	Raw	68.2	65.9	13.2	57.5	77.9
		Bias Corrected	88.0	84.9	17.0	74.1	100.4
5	Observed		97.0	93.6	15.7	85.1	102.6
	ECHAM5-RegCM3	Raw	91.2	88.7	15.8	79.4	98.8
		Bias Corrected	97.0	94.4	16.8	84.4	105.1
	HadGEM2-RegCM4	Raw	68.2	65.9	13.2	57.5	77.9
		Bias Corrected	97.0	93.6	18.8	81.7	110.8

Comparing to the corresponding raw values, the values of improvement percentage for all the time series-based metrics are illustrated in Figure 3.3. The figure shows that the method is very successful in terms of PBIAS with the improvement rate between 68% - 100%. It also improves the other metrics within the rate 6% - 100%. With respect to the average of the improvement rates computed over all the evaluation grids, the method decrease the overall RMSE and RSR metrics equally with the rate 25%. It is also able to enhance the overall NSE measure up to 50%.

Table 3.3 : Time series-based performance statistics of new bias correction method at the RegCM3 grid cells over the historic period 1961–1990.

Grid No.	Model		RMSE (mm)	PBIAS (%)	NSE	RSR
1	ECHAM5-RegCM3	Raw	69.10	-78.28	-13.71	3.84
		Bias Corrected	20.33	0.08	-0.27	1.13
	HadGEM2-RegCM4	Raw	23.67	13.99	-0.73	1.31
		Bias Corrected	21.79	0.00	-0.46	1.21
2	ECHAM5-RegCM3	Raw	39.23	36.26	-2.48	1.87
		Bias Corrected	33.25	0.00	-1.50	1.58
	HadGEM2-RegCM4	Raw	26.81	11.67	-0.62	1.27
		Bias Corrected	25.06	0.00	-0.53	1.19
3	ECHAM5-RegCM3	Raw	27.05	25.12	0.61	1.29
		Bias Corrected	20.03	0.00	0.79	0.96
	HadGEM2-RegCM4	Raw	31.1	26.66	0.49	1.49
		Bias Corrected	25.73	8.43	0.65	1.23
4	ECHAM5-RegCM3	Raw	33.80	7.75	-1.33	1.53
		Bias Corrected	29.53	0.00	-0.78	1.34
	HadGEM2-RegCM4	Raw	29.24	22.45	-0.74	1.32
		Bias Corrected	22.86	0.00	-0.07	1.03
5	ECHAM5-RegCM3	Raw	25.58	6.02	-0.73	1.63
		Bias Corrected	20.39	0.00	-0.69	1.30
	HadGEM2-RegCM4	Raw	31.58	29.67	-3.06	2.02
		Bias Corrected	15.52	0.00	0.02	0.99

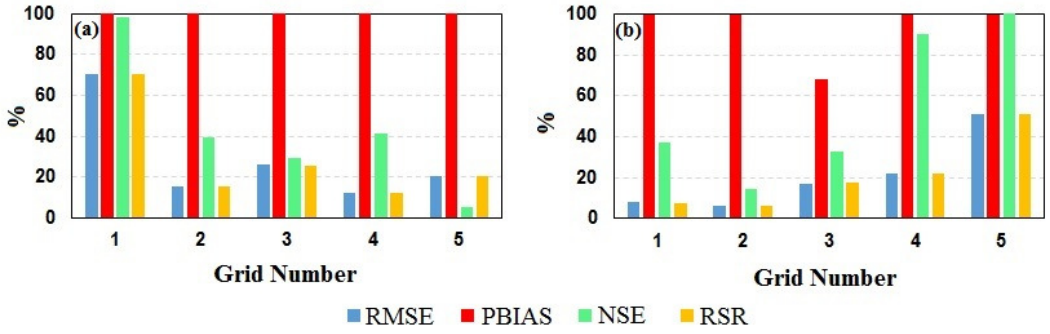

Figure 3.3: Improvement rate (%) in time series-based performance statistics as a result of bias correction applied over AMR_{sim} series at the historic period 1961–1990; (a) ECHAM5-RegCM3 (b) HadGEM2-RegCM4.

Owing to the importance of statistical distribution of historical AMR series on future extreme rainfall estimation, efficiency of the new bias correction methodology is considered to supplementary evaluation using cumulative distribution functions

(CDFs) of the observed, raw, and bias corrected RCM-simulated AMR series at the reference period. Examples of CDF curves depicted for the observed, raw and bias corrected AMR series over ECHAM5-RegCM3 grid#2 and grid#4 are presented in Figure 3.4. These two grids were given because they represent the highest, 1.57, and relatively low, 1.08, BCFs (see Table 3.1). It is clear from the figure that the raw RCM-simulated AMR sets are heavily biased (as also shown by statistics in Table 3.2) and the proposed bias correction method effectively, but in different extent, corrects the bias in the raw RCM-simulated CDFs.

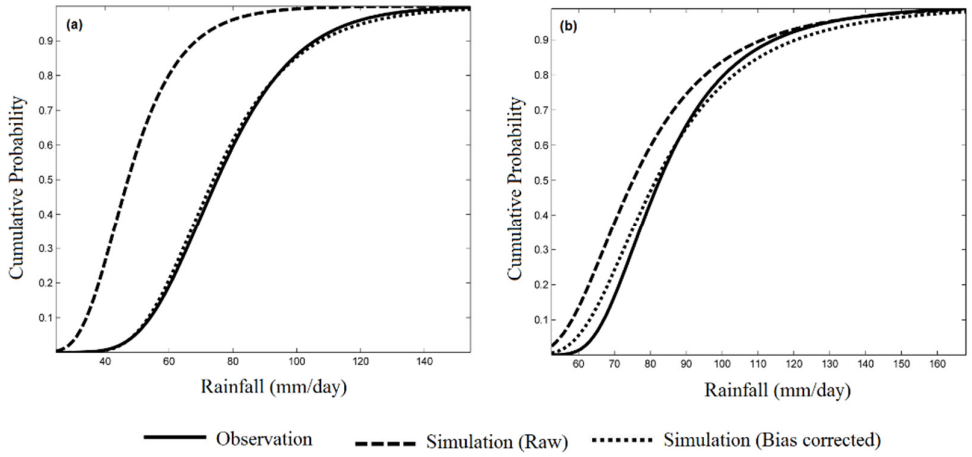

Figure 3.4: Examples of observed and simulated data CDFs over the historic period 1961–1990; (a) grid #2 (b) grid #4.

3.7.2 Projected changes in design storms

Results from the EVT analysis at the reference and three future periods are displayed in Figure 3.5. Each row in the figure exhibits the observational and scenario simulations correspond to RegCM3 grids. The projected design storm magnitudes at all the future periods expose noticeable differences between both scenario simulations as well as historic values, particularly over the regions within grids #1 and #5. Most future simulations reveal a substantial decrease in the magnitude of future design storms. It is notable that only the ECHAM5-RegCM3 outputs (see SRES-A2 in Figure 3.5) contain magnitudes more than historic values for the all recurrence intervals at grid #2. Owing to the frequent use of 25- and 100-yr storms in the design and planning of hydraulic infrastructures, spatial and temporal distribution of them over the study region are presented in Figures 3.6 and 3.7, respectively.

Based upon ECHAM5-RegCM3-SRES-A2 simulation results, the 25-year storm magnitudes vary substantially across the region, with values within the broad ranges 22-164 mm/day and 40-178 mm/day at near future (2013-2040) and far future (2071-2099), respectively. The range is narrowed to 22-141 mm/day at middle of the century (2041-2070). By contrast, the 25-yr storms do not vary dramatically during the future periods when the HadGEM2-RegCM4-RCP8.5 outputs are considered. The lower limit of 61 mm/day and upper limit of 124 mm/day are projected for the far future and middle of the century, respectively. Such a significant difference between the projections is also observed for the 100-yr storms. For example, the heaviest 100-year storm amount projected by SRES-A2 scenario (for the far future) equals to 214 mm/day, whereas the corresponding value for RCP8.5 is 125 mm/day. It should be mentioned that that such difference is not merely due to the different future climate scenario, but due to different driving GCMs and different versions of the RegCM.

The spatial distribution of these magnitudes is also informative. The extreme rainfall pattern of the study area, intense rainfalls on the north-eastern and less rainfall falling at the south-western regions, in ECHAM5-RegCM3 simulations is not discernible. However, the pattern more or less is followed by HadGEM2-RegCM4 simulations, which is probably due to the higher spatial resolution of RegCM4 grids as well as the influence of GCM driving conditions on the behavior of RCMs.

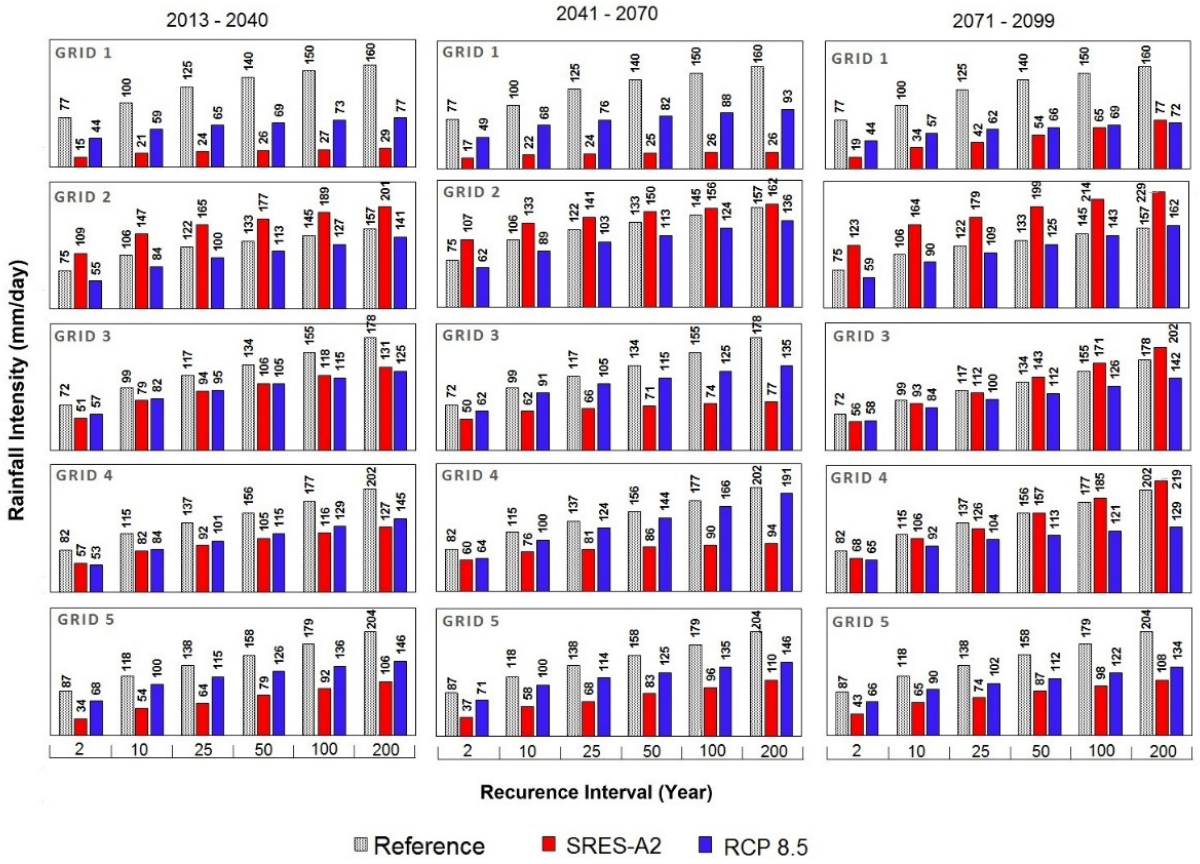

Figure 3.5 : Observational and projected design storms over Rize province, Turkey

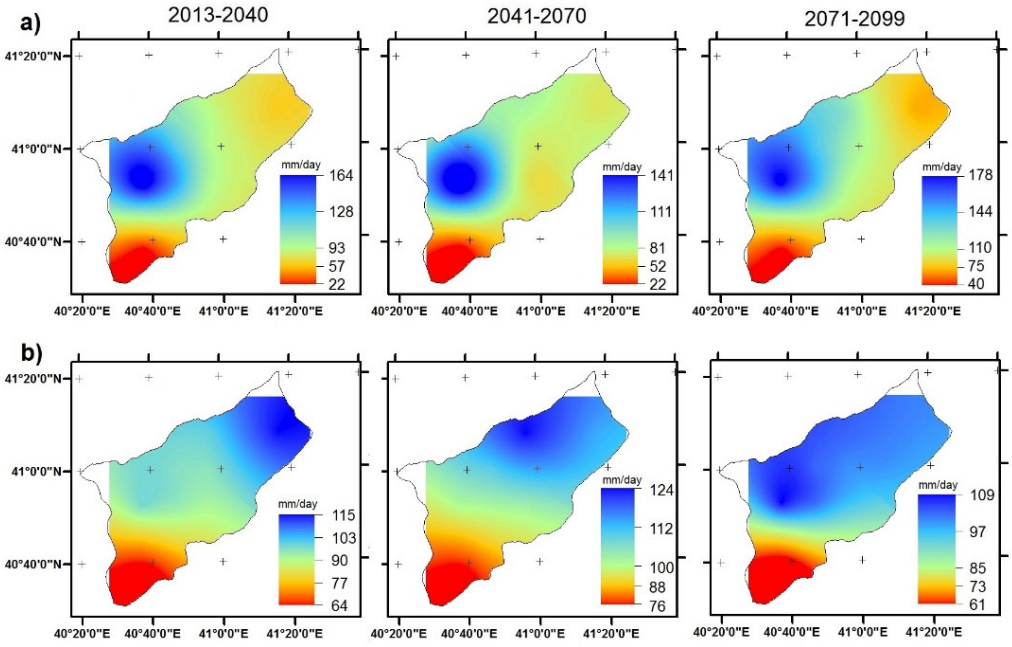

Figure 3.6 : 25-yr projected design storm over Rize Province: (a) SRES-A2 and (b) RCP8.5.

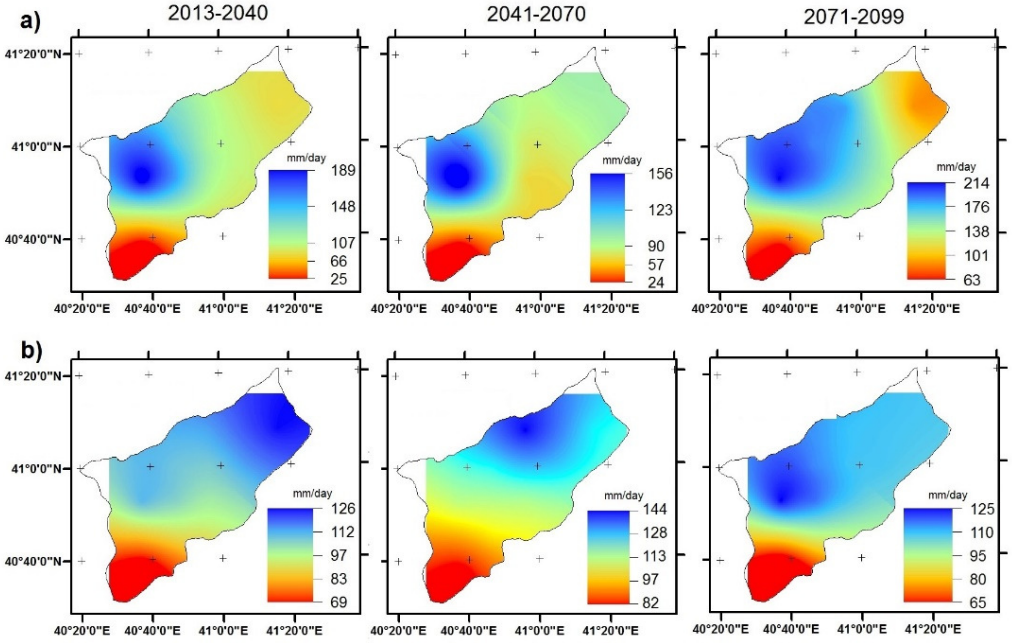

Figure 3.7 : 100-yr projected design storm over Rize Province: (a) SRES-A2 and (b) RCP8.5.

In order to analyze the variation and central tendency of projected design storms during the future periods, regional summaries of the projected changes in storm

values are presented in Figure 3.8. Each box in the figure represents the distance between the first and third quartile, interquartile range, of the data and the horizontal line inside the boxes shows the median. The upper and lower vertical lines, whiskers, display the highest and lowest change, respectively. The asterisk symbols indicate the outliers, which are any values outside of 1.5 times the box. The figure shows the models agreement on the sign of the change at median, lower and upper quartiles for the near future and the middle of the century. Due to relatively large interquartile range provided by SRES-A2, there is no agreement between scenarios on the sign of the change at median and upper quartiles of higher recurrence intervals (i.e., greater than 50-yr) for the far future period. In almost all recurrence intervals, negative change (i.e., decrease in intensity) is projected independent of the future time slices and greenhouse gas scenarios considered. The percentage of median changes of projected storms are almost the same for the near future period, however the interquartile range is generally larger in SRES-A2 compared to RCP8.5 probably due to different climate sensitivity in the models and feedback mechanisms. The box plot analysis also reveal that the higher amount of changes are found by SRES-A2 and RCP8.5 for the middle and end of the century, respectively. Overall, the interquartile ranges of the RCP8.5 simulations are dramatically smaller than those of the SRES-A2, which is a good indicator of lower variability of RCP8.5 simulations across the study region. For this scenario, plots are skewed to the lower value, which means that the concentration of values is closer to the 25th percentile. If ensemble mean of the scenarios is considered, approximately 30% decrease in the median value of extreme rainfall events is projected over the study region for the near future and middle of the century. This change dramatically decreases to 15% of its historic value at the end of the the century.

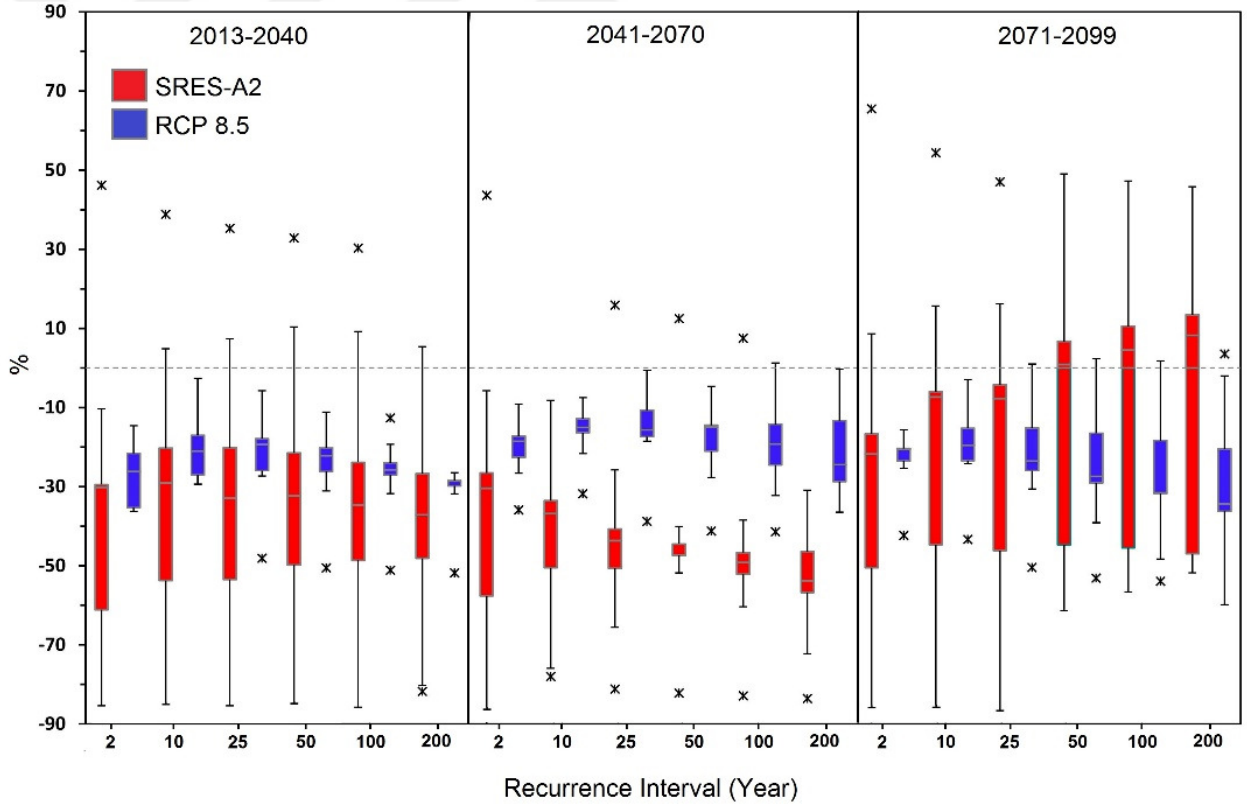

Figure 3.8 : Projected changes (%) in design storms over the future periods as differences relative to the reference period (1961–1990) for SRES-A2 and RCP8.5 scenarios.

3.7.3 Comparison among distribution models

Combining the five grids data sets produced from the ECHAM5-RegCM3 and the 11 grids data sets of HadGEM2-RegCM4 at the three different future period, a total of 48 data sets were used in the goodness-of-fit tests. The results from goodness-of-fit tests among the five distribution models used in this study are tabulated in Table 3.4. Considering the K-S and A-D tests separately, each value in the table represents the frequency of an individual model providing the best fit for a grid. Out of the 48 data sets, the K-S test showed that the GEV distribution has the best fit of the five distributions with a total of 23 times (i.e., 48% of the times). The A-D test also demonstrated that the GEV distribution is the best with a total of 31 times (65% of the times). It is worth mentioning that the critical values of the tests at 5% significance level did not reject GEV distribution in any circumstance.

Table 3.4 : Frequency of success for each distribution model as the best fit for RCM data.

		GEV	**LN-3**	**LN-2**	**LP-3**	**Gumbel**
Scenario	Test	2013-2040				
SRES-A2	K-S	4	1	-	-	-
	A-D	4	1	-	-	-
RCP8.5	K-S	7	2	2	-	-
	A-D	8	1	1	1	-
		2041-2070				
SRES-A2	K-S	2	3	-	-	-
	A-D	2	3	-	-	-
RCP8.5	K-S	4	4	-	1	2
	A-D	8	3	-	-	-
		2071-2099				
SRES-A2	K-S	1	2	-	1	1
	A-D	3	2	-	-	-
RCP8.5	K-S	5	2	1	1	2
	A-D	6	2	-	3	-

It is also evident from the table that the LN-3 is the second best distributions for characterizing the behavior of projected extreme rainfalls over the study region. The results also reveal that all of the models at least one time outperform their counterparts in terms of goodness-of-fit. In fact, the best distribution model varies among the both RCM grids and time slices. This means that if a distribution model is

superior for a grid, it is not necessarily effective model for the adjacent grids. Thus, instead of implementation of a single distribution a comparative analysis among different models would be crucial when using RCM data for EVT analysis. Comparative analysis among the storm magnitudes estimated by the best fit and the others at each grid also indicates that when the return periods associated with the design storm estimates greatly exceed the length of record available (30-year), discrepancies between the used distributions tend to increase.

4. PROJECTED RAINFALL-RUNOFF MODELLING

4.1 Intrudoction

Future changes in runoff (streamflow) can have significant effects on human life. Such hydrological changes are likely to affect virtually every aspect of human well-being, from municipal and industrial water supply to agricultural productivity and wildlife management. A number of models are available to estimate projected runoff under different climate change scenarios. At the global scale, global hydrological models (GHMs) such as WaterGAP (Alcamo et al. 1997) driven by a single or an ensemble of general circulation models (GCMs) are commonly implemented (e.g., Milly et al. 2005; Betts et al. 2007; Kling et al. 2012; Hagemann et al. 2013; Davie et al. 2013; Bozkurt et al. 2015). At the regional scale or even smaller scales (catchment-scale, hillslope), standalone catchment hydrological models such as SWAT (Soil and Water Assessment Tool, Arnold et al. 1999; Neitsch et al. 2002), HBV (Hydrologiska Byrans Vattenbalansavdelning, Bergström 1976), Hydro-BEAM (Hydrological River Basin Environmental Assessment Tool), Sacramento Soil Moisture Accounting Model, Burnash et al. 1973), and SWIM (Soil and Water Integrated Model; Krysanova et al. 1998) driven by high resolution GCM data or regional climate model (RCM) outputs are frequently applied. For instance, Betts et al. (2007) used an ensemble of experiments with a GCM that includes a vegetation component to assess the impact of doubled carbon dioxide (CO2) concentration on future changes in continental runoff. The authors demonstrated that physiological effect of doubled CO2 concentration on plant transpiration increases simulated global mean runoff by six percent compared to pre-industrial level. Using the outputs of one GHM forced by two GCMs (Parallel Climate Model and Model for Interdisciplinary Research On Climate), Barnett et al. (2008) investigated future changes in the surface hydrology variables of the western USA, including snow pack, seasonal timing of streamflow in major rivers, and average January to March

daily minimum temperature over the study region. The authors concluded that there is a detectable and attributable anthropogenic signature on the hydrology of the study region. With respect to the A1B scenario of Intergovernmental Panel on Climate Change (IPCC), Lee et al. (2014) stated that the Regional Model Program can be adequately used to detect future variability of surface runoff over Korea. More recently, Bozkurt et al. (2015) utilized the Hydrological Discharge (HD) model of Max Planck Institute for Meteorology forced by a variety of GCM-RCM datasets to investigate the future of discharge in Euphrates–Tigris Basin. They demonstrated that HD simulations forced by relatively low-resolution GCM outputs are not good enough to reproduce the seasonal cycle of discharge. By contrast, high-resolution RCM-forced simulations reproduce the annual cycle of discharge reasonably well. A review concerning existing GHMs and their structure, strengths and shortcomings has been well documented by Sood and Smakhtin (2015).

The scientific literature of the past two decades contains a large number of papers dealing with the application of different standalone hydrological models for the catchment-scale assessment of climate change effects on river flows. Most of the studies were focused on North America, Europe, and Australasia, with a small number of papers from Asia and Africa (e.g., Marshall and Randhir 2008; Minville et al. 2008; Bekele and Knapp 2010; Veijalainen et al. 2010; Hattermann et al. 2011; Islam et al. 2012; Lespinas et al., 2014; Sellami et al. 2015). For example, satisfactory implementations of the U.S. Geological Survey's Precipitation Runoff Modeling System (PRMS) for studying the potential impacts of climate and land use changes on Trent River basin in the US and Brahmani River basin in India were reported by Qi et al. (2009) and Islam et al. (2012), respectively. In another impact assessment study in the US, Bekele and Knapp (2010) used SWAT in conjunction with the Illinois State Water Survey's GCM outputs to model future water availability in Fox River basin. Gosain et al. (2006) used SWAT in combination with the outputs of the Hadley Centre Regional Climate Model to assess the effect of climate change on different catchments in India. Their results indicated an increase in the severity of drought and intensity of floods in different parts of the country. The study also revealed that the rainfall intensification due to the climate change does not result in an increase in the surface runoff as it was generally expected. The newly

developed Intermediate Space Time Resolution Hydrological Model (In-STRHyM) feed by projections of three RCMs has been used by Senatore et al. (2011) to assess the hydrological impacts of climate change in a relatively small catchment in southern Italy. With respect to the SRES A2 and A1B greenhouse gas emission scenarios, the authors reported a general reduction in water resource availability for the period 2070–2099. More recently, Sellami et al. (2015) employed SWAT model feed by projections from an ensemble of four climate models in two small catchments in French and Tunisia. The authors demonstrated that both catchments are likely to experience a decrease in precipitation and increase in temperature in the near future (2041–2071).

On the one hand, quality of catchment-scale runoff projections obtained by hydrological models is highly dependent on the accuracy and density of field measurements, particularly rainfall and runoff, which are required for model calibration/verification (Ajami et al. 2004; Kamali et al. 2013). Hence, use of these models is practically impossible for ungauged catchments and problematic for poorly gauged catchments where a regionalisation method is required to estimate parameters of interest (Parajka et al. 2005; Bardossy 2007; Bulygina et al. 2009). On the other hand, according to the IPCC's new assessment report (IPCC 2013), the Middle East has been identified as one of the most vulnerable regions of the world to climate change. The region encompasses numerous ungauged catchments, thereby limiting the ability to efficiently manage their water resources under changing climate. Consequently, there is a need for new tools/models with confined or even without calibration requirements to simulate future hydrology of ungauged catchments. It is worth mentioning that in terms of hydrological predictions much of the success has so far belonged to gauged catchments. This is the reason why International Association of Hydrological Sciences (IAHS) recommends that more studies should be conducted to improve modelling techniques and provide better understanding of different kinds of data and their possible uses in ungauged catchments (Hrachowitz et al. 2013).

It would be very helpful for hydrologists if data-driven technology incorporating the existing runoff observations and forthcoming climatologic projections can be

integrated into the catchment-scale rainfall-runoff modelling. This chapter for the first time delineates the development and implementation of two advancements of genetic programming (GP) approach including linear GP (LGP) and multigene GP (MGGP) for projected rainfall-runoff (hereafter PR-R) modelling in daily horizon at a poorly gauged catchment in Turkey. In addition, efficiency measures of these models are compared with those of obtained by the well-known feed-forward backpropagation (FFBP) artificial neural network (ANN) method used as a bench mark. To this end, projected rainfall and temperature data sets produced by a GCM-RCM integration as well as a limited set of observational runoff (streamflow) are utilized as model inputs. The idea behind predicting future runoff using projected climatic variables is based on two similar works accomplished by Tung et al. (2009) and Chen et al. (2010), respectively. In the former, The GP approach was successfully applied to project climate change impacts on the population of Formosan Landlocked Salmon in Dajia Creek, Taiwan. In the latter, the projected shallow groundwater level in North China was investigated using an ANN approach trained under different greenhouse gas emission scenarios. By the time of this study, it was only few works in the literature (e.g., Steynor et al. 2009) consider any data-driven model for projected runoff estimation. Using the outputs of three GCMs forced by SRES-A2 emission scenario, Steynor et al. (2009) demonstrated that self-organising maps could be successfully applied to project Breede River runoff in South Africa.

The remainder of this chapter is organized as follow. In sections 4.2 and 4.3, the concepts of implemented GP and ANN variants are briefly reviewed, respectively. In sections 4.4 and 4.5, study area, proposed models and combinations as well as the implemented efficiency criteria are explained. The modelling procedure and applied parameters are presented in section 4.6. In section 4.7, the results and model's performances are evaluated, discussed and compared with each other. Concluding remarks are presented in the next chapter of the thesis.

4.2 Genetic Programming (GP)

The state-of-the-art GP (Koza 1992) is a machine learning method that evolves computer programs to solve problems using Darwinian natural selection. In hydrological applications, GP is commonly used to infer the underlying structure of either a natural (e.g., Ghorbani et al. 2010; Danandeh Mehr et al. 2013, 2014a, 2014b; Sattar and Gharabaghi 2015) or an experimental processes (e.g., Khan et al. 2012; Roushangar et al 2014; Uyumaz et al. 2014) process. In such applications, GP generates some possible solutions to identify the process numerically. When the task is to build an empirical model of data acquired from a process or system, the GP is often known as symbolic regression. GP is a self-structuring technique without requiring the user to know or specify the form of the solution in advance. It differs from either traditional regression analysis or other artificial intelligence techniques, in which modeller must specify the structure of process. As shown in Figure 4.1, in GP, potential solutions are usually represented by tree structures with a root node, inner node(s) and leaves. Population of initial solutions is generated through a random processes such as the full, grow, and ramped half-and-half methods (Poli et al. 2008). Subsequent generations are commonly evolved through three genetic operators, namely reproduction, crossover, and mutation, to create better solutions. An example of crossover and mutation operators to generate two new population (offsprings) is presented in Figure 4.2. This process is iterated until the population contains individuals that solve the problem well.

The major inputs for a standard GP-based modelling are: (i) patterns for training/validation; (ii) fitness function (e.g., mean squared error) for tournament selection; (iii) functional (or nodes) and terminal (or leaves) sets for structural identification; and (iv) GP parameters for formation of a syntax tree (i.e., a function/potential solution). Depends on the degree of complexity of the process of interest, which is addressed to be formulated by GP, the functional set may contain the basic arithmetic operators (i.e., +, - , ×, ÷) or more complicated mathematical operators such as Sin, Exp, and others. In order to generate an initial program, an operator is chosen randomly from the predefined functional set to fill the root node of a GP tree. Then inner nodes are filled randomly by a member of either functional

or terminal sets. Finally, leaves are filled by only predefined members of terminal set may comprise independent variables, numerical/logical constants, or arguments for the implemented functions. Owing to the probabilistic essence of GP approach, it is possible to derive a large number of potential solutions (functional expressions) for a problem. An experienced GP modeller could choose the best solution via different ways. More details on GP can be obtained from Babovic and Keijzer (2000).

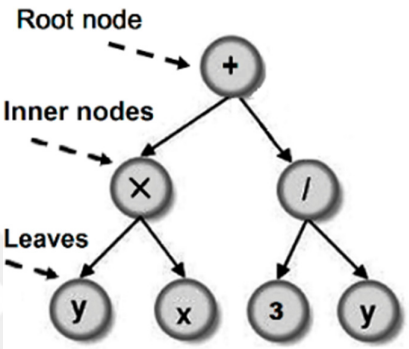

Figure 4.1 : Tree representing of function ((x .y + 3/y)).

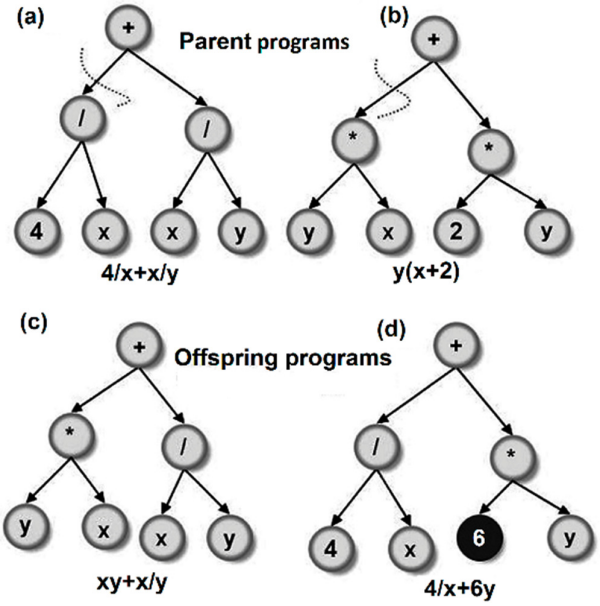

Figure 4.2 : An example of crossover and mutation operators in GP.

4.2.1 Linear genetic programming (LGP)

Besides the traditional GP (Koza 1992), there are several linear variants of GP such as linear genetic programming (LGP), multigene genetic programming (MGGP),

multiexpression programming (MEP), gene expression programming (GEP), Cartesian genetic programming (CGP) and grammatical evolution (GE). All of these variants have a clear distinction in their genotype of a chromosome (or an individual). A number of researchers have reported successful application of the linear genome in GP (e.g. Brameier and Banzhaf 2007; Guven 2009; Shiri et al. 2012; Danandeh Mehr et al. 2013, 2014a and 2014b; Shoaib et al 2015).

Comparing to Koza-style GP, where the tree-based programs correspond to expressions from a functional programming language (like LISP), LGP evolves sequences of instructions from an imperative programming language (like C/C++) or machine language (Brameier and Banzhaf 2007). Thus, evolving programs are represented as linear entities (strings) rather than in expressions of a functional programming language. The term "linear" refers to the imperative program representation. It does not mean that the method provides a linear function as a solution (Danandeh Mehr et al. 2014a). An example of LGP evolved program in C language describing the flow of water (Q) through porous media, the well-known Darcy's Law: Q=K.I.A, is illustrated as follows:

```
L0:  f [0] = 0.0;

L1:  f [0] += I;

L2:  f [0] *= K;

L3:  f [0] *= A;
```

where I = pressure gradient, K = hydraulic conductivity, A = area, and f[0] is an accumulator variable representing the final output (i.e., Q) of the evolved program. LGP employs such temporary variable to store values while performing calculations. The temporary variable equals to zero by default in the first line (i.e., L0) of the program and the output is the value remaining in the last line. It should be mentioned that in this program, evolving introns have been removed previously. In analogy with natural introns, DNA parts of genes with information that is not expressed in proteins, an intron in LGP is defined as a program portion without any effect on the calculation of the output(s) for all possible inputs. A simple example of an intron is:

```
i) L0:  f [0] += -1.00f;
```

L1: f [0] += +1.00f;

The program structure in LGP allows introns to be detected and removed much easier than in tree-based GP. Similar to traditional GP, the user prespecified functions (or instruction set) in LGP can be composed of arithmetic operations (+, - , ×, ÷), Boolean logic functions (AND, OR, NOT), conditionals (IF, THEN, ELSE), and any other mathematical functions such as ABS, Sin, Ln, EXP, Sqrt, and others. If only the mathematical functions specified, an LGP program can be turned into a functional representation by successive replacements of variables that will be more useful in hydrological studies. More details on the application of LGP in hydrological modelling can be obtained from Danandeh Mehr et al. (2013) and Uyumaz et al. (2014). The LGP modelling attributes and adopted parameters used in this thesis are explained in the Section 4.6.

4.2.2 Multigene genetic programming (MGGP)

Multigene genetic programming (Searson 2009) is another advancement of GP that linearly combines low depth GP trees in order to improve fitness of traditional GP approach. Owing to the use of smaller trees, the MGGP is expected to provide simpler models than those of traditional monolithic GP. In MGGP, predictand variables are computed by the weighted output of each genes (i.e. trees) in the multigene program plus a bias term. For example, a pseudo-linear MGGP individual shown in Figure 4.3 predicts the predictand variable using two genes comprising two input variables x and y. Mathematically, this model can be written as $d_0 + d_1 (4x+x/y) + d_2 (4/x+Sqrt(y)$ where d_0 = bias (error) term and d_1 and d_2 are the regression coefficients (i.e. gene weights). Commonly, the coefficients are determined by the ordinary least squares method for each MGGP individual (Searson et al. 2010). Therefore, MGGP employs the power of classical linear regression method to capture nonlinear behaviours without need for prespecified nonlinear structure.

In MGGP, each gene is a straightforward GP tree without any implicit or explicit reference to any of the other genes in the same chromosome. The Chromosomes containing standard low-depth GP trees create the initial population. Modelers define maximum number of each tree. Then, offspring is acquired by standard GP subtree crossover, direct reproduction and mutation transformation.

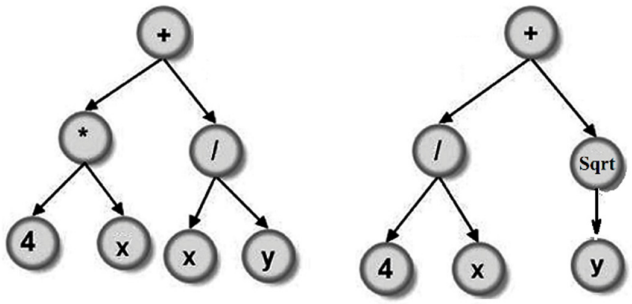

Figure 4.3 : An example of MGGP individuals.

In addition to the traditional GP recombination operator, in MGGP, a tree crossover operator called “two point high-level crossover” is also used to exchange of genes between chromosomes. The modeller also sets the relative probabilities of each of these GP operators. More details on the application of MGGP for performing symbolic regression can be obtained from Searson et al. (2010) and Gandomi and Alavi (2012). The MGGP modelling attributes used in this study are explained in Sub-section 4.6.

4.3 Multi-layer Perceptrons

The development of layered feed-forward networks began in the late 1950's by Rosenblatt's perceptron as the first model for learning with a teacher (i.e., supervised learning), which is often referred to as single layer perceptrons (Rosenblatt 1958). Basically, they consist of a single neuron with adjustable synaptic weights and bias. The weights are initialized randomly and then are adjusted as a function of estimation error by an iterative process called training. The mechanism by which the weights are adjusted is called learning rule. Single layer perceptrons can classify only linearly separable patterns (Hykin 2008). Moreover, when the training data is linearly separable, there can be an infinite number of solutions. These limitations of a single layer network has led to the development of multi-layer perceptrons (MLPs), which have one or more hidden layers. An MLP can approximate any continuous multivariate function to any degree of accuracy, provided there are sufficiently many hidden neurons (Hornik et al. 1989; Abrahart et al. 2012). MLP is one of the most popular feedforward ANNs that performs tasks such as function fitting and pattern

recognition problems. It has already been demonstrated that an MLP with supervised back propagation training algorithm (hereafter FFBP) is satisfactory for hydrological applications (e.g., Hornik et al, 1989; Rogers 1992; Rogers and Dowla 1994; Yang et al. 1997; ASCE 2000; Solaimani 2009; Nourani et al. 2011; Tahershamsi et al. 2012; Danandeh Mehr et al. 2015). An FFBPP network with one hidden layer is illustrated in Figure 4.4. It shows that a neuron connection only exist from a neuron in the input layer to other neurons in the hidden layer or from a neuron in the hidden layer to other neurons in the subsequent output layer. The letters M, N and O in the figure denote the number of neurons in input, hidden and output layers, respectively. The weights are different in the hidden and output layers, and their values are adjusted during the back propagation training process.

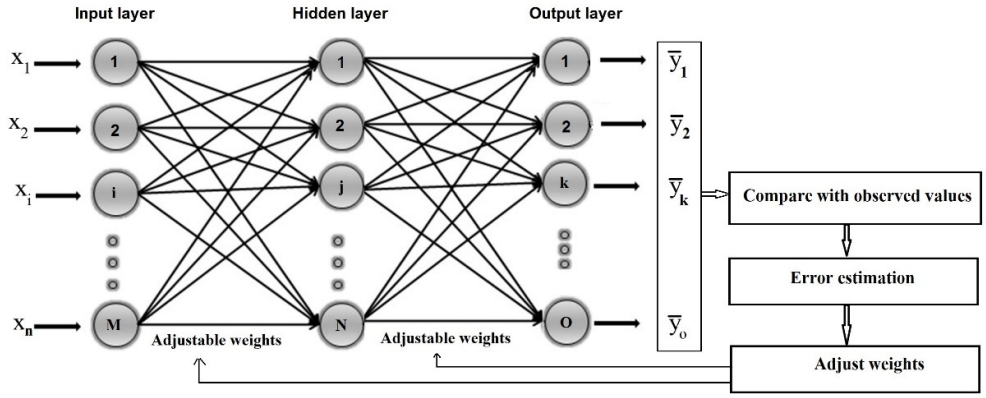

Figure 4.4 : A three-layered FFBP network.

The explicit expression for an output variable of abovementioned network is presented in Equation (4.1). Any computed output is based on a linear combination of input variables (x_i) that are transformed by linear and/or non-linear activation functions. The possibility of using nonlinear activation functions at the hidden and output layers provides the capability of capturing the complexity and nonlinearity features in the systems being modelled.

$$\bar{y}_k = f_O\left[\sum_{j=1}^{N} W_{kj}.f_h\left(\sum_{i=1}^{M} W_{ji}x_i + B_{jo}\right) + B_{ko}\right] \tag{4.1}$$

where B_{jo} is the bias for the j^{th} hidden neuron, B_{ko} is the bias for the k^{th} output neuron, W_{ji} is a weight in the hidden layer connecting the i^{th} neuron in the input layer

and the j^{th} neuron in the hidden layer, W_{kj} is a weight in the output layer connecting the j^{th} neuron in the hidden layer and the k^{th} neuron in the output layer, and f_o and f_h are the activation functions for the output and hidden neurons, respectively. More information about the FFBP structure utilized in this study is presented in Sub-section 4.6.

4.4 Study Area and Data

4.4.1 Observational runoff data

The proposed PR-R models are trained and verified for the unregulated portion of Senoz Catchment located in Rize Province, Turkey (Figure 4.5). The province, which is considered as the wettest corner of the country, is characterized by high rainfall and rough terrain covered by forest and tea gardens experiencing occasional landslides and floods. According to Koppen-Geiger climate classification, the catchment has a borderline oceanic/humid subtropical climate with warm summers and cool winters. Mean annual rainfall and temperature in the province is around 2500 mm and 14.1 °C, respectively. The entire catchment area is 379 km^2 and its longest waterway is 46 km long. The topography is steep, with an average slope of 45%. The catchment area at the point of the hydrometric station, Kaptanpasa Station, is 231 km^2 and its elevation varies between 400m to 3500m above sea level. There is no rain gauge within the catchment area and daily runoff measures are only available for the period 1990–2011. The observational runoff time series for the period 2001-2011 used in this study is presented in Figure 4.6.

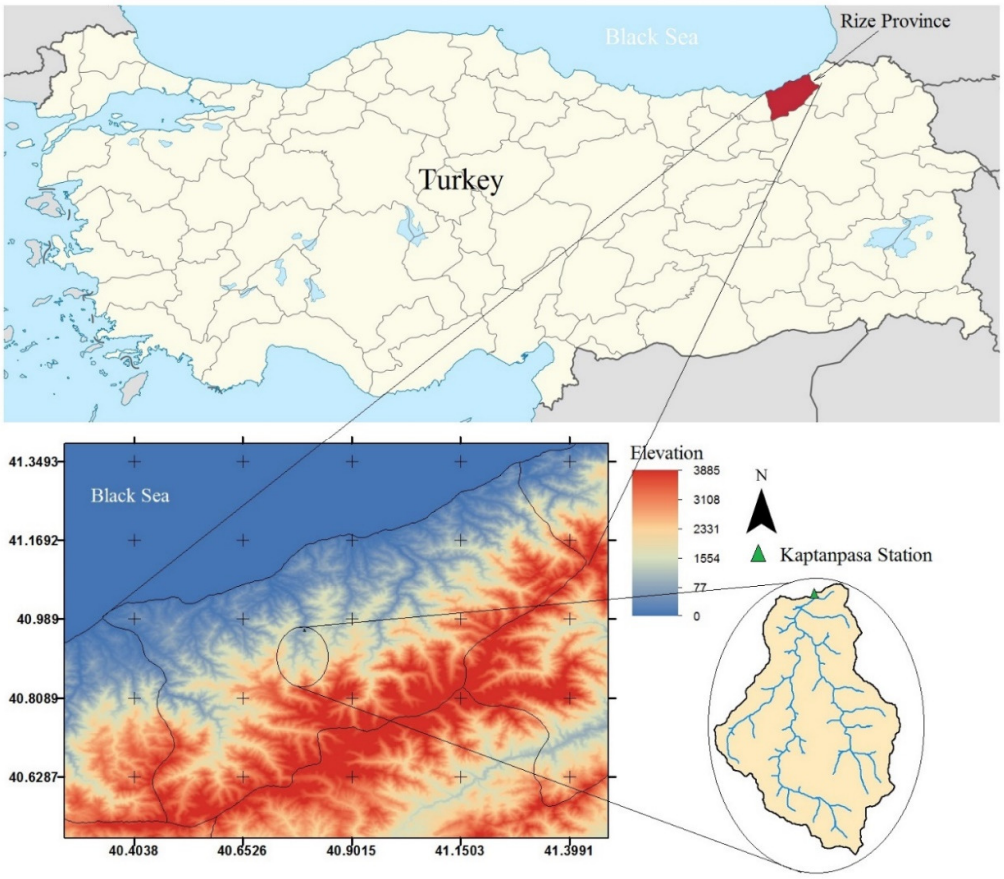

Figure 4.5 : Location of study area (Senoz Catchment).

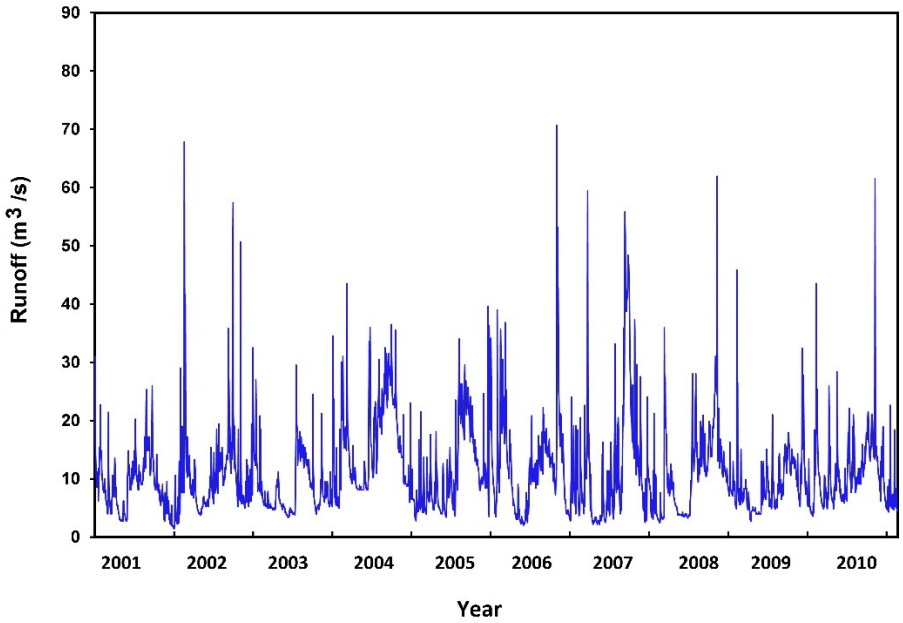

Figure 4.6 : Daily runoff observations at Senoz Stream (Kaptanpasa Station).

4.4.2 Projected rainfall and temperature data

Projected daily rainfall and temperature data (hereafter RCM rainfall and temperature) were taken from the outputs of Max Planck Institute for Meteorology ECHAM5-GCM, which were dynamically downscaled over Turkey using International Centre for Theoretical Physics RegCM3-RCM. The relevant downscaling works were accomplished at Istanbul Technical University for simulation of the future period 2000-2099 with respect to A2 greenhouse gas emission scenario. Considering the adjustment period (spin-up), we discarded the first year of the simulation outputs and took the following ten years (i.e., 2001–2011) data for PR-R modelling. Detailed descriptions of physical parameterizations of the climate models can be found at Bozkurt and Sen (2013). The RCM rainfall and temperature time series, which were calculated by areal weighted averaging over the two RCM grid cells and covers portion of the catchment is presented in Figure 4.7. It should be mentioned that RCM rainfall values less than 1mm are considered to be zero in this study.

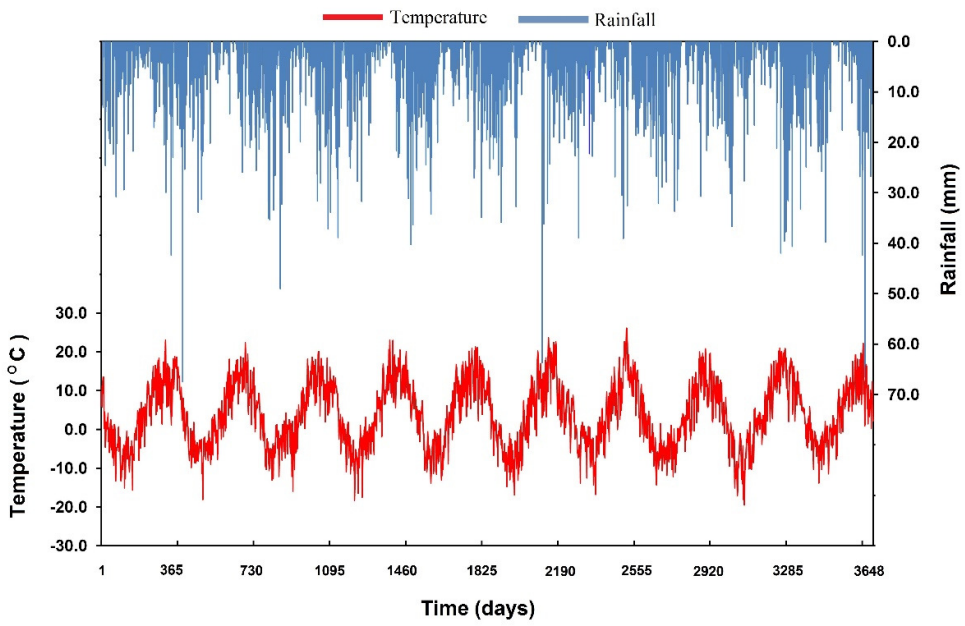

Figure 4.7 : Rainfall and temperature data generated by ECHAM5-RegCM3 integration at Senoz Catchment.

4.5 Forecasting scenarios and efficiency criteria

In this section, different one-day ahead bivariate (i.e., Comb (1) to Comb (4)) and multivariate (i.e., Comb (5) to Comb (8)) forecasting combinations are introduced to formulate the PR-R process of the study catchment.

Comb (1): $Q_t = f(I_t, Q_{t-1})$

Comb (2): $Q_t = f(I_{t-1}, I_t, Q_{t-1})$

Comb (3): $Q_t = f(I_{t-2}, I_{t-1}, I_t, Q_{t-1})$

Comb (4): $Q_t = f(I_{t-2}, I_{t-1}, I_t, Q_{t-2}, Q_{t-1})$

Comb (5): $Q_t = f(T_t, I_t, Q_{t-1})$

Comb (6): $Q_t = f(T_{t-1}, T_t, I_{t-1}, I_t, Q_{t-1})$

Comb (7): $Q_t = f(T_{t-2}, T_{t-1}, T_t, I_{t-2}, I_{t-1}, I_t, Q_{t-1})$

Comb (8): $Q_t = f(T_{t-2}, T_{t-1}, T_t, I_{t-2}, I_{t-1}, I_t, Q_{t-2}, Q_{t-1})$

where Q_t, I_t, and Tt represent observational daily runoff (streamflow), projected daily rainfall and projected daily temperature values at the present time instance, respectively. The indices t-1 and t-2 represent the one and two-day lags, respectively. It is worth mentioning that such combinations using observational sets of rainfall, runoff and temperature are frequently used in the rainfall-runoff modelling literature (e.g., Nourani et al. 2009 and 2011; Wu and Chau 2011).

Highly scattered distribution of observed runoff data with respect to the simultaneous RCM temperature and rainfall values are shown in Figure 4.8. For such complexity, data-driven approaches may be the best choice to discover underlying structure between the data sets. Prior to training a data-driven approach, a normalization procedure is commonly applied to make input/output variables dimensionless and rescales them within a certain range. In this study, the input/output variables are normalized by dividing each value for the variable by the maximum value observed/projected for that variable. Similar to other data-driven modelling techniques, a modeller must divide the data into training and validation subsets. Training data is a set of existing input values (here I, T, and Q) and corresponding output values (Q). An Initial population of GP-based/ANN models is generated using

training data. Validation data is another set of input and corresponding output values that is used after each run to evaluate the initial models. As well as training and validation data sets, an additional holdout testing data set can be specified to help mitigate overfitting problem. It also serves to measure how well the evolved models generalise to unseen data set. Our data set comprises 3652 data points. This is sequential data, so the first 60% of entire data is used as training set and the second 20% is used for validation of the proposed models. The rest 20% of the data points is held out for testing.

Table 4.1 shows the statistical parameters of each set. Since the statistical characteristics of all of the subsets are close to each other, division of data is satisfied for the both GP-based and ANN modelling.

Many studies have indicated that a hydrological model can be sufficiently evaluated by NSE (see equation (3.8)) RMSE and (see equation (2.2)) measures (e.g., Legates and McCabe 1999; Nourani et al., 2012). Given the extra challenge of this study, three other efficiency metrics, including Kling-Gupta efficiency (KGE), volumetric efficiency (VE), and persistence Index (PI), are also used for evaluation and mathematical comparison of the proposed models.

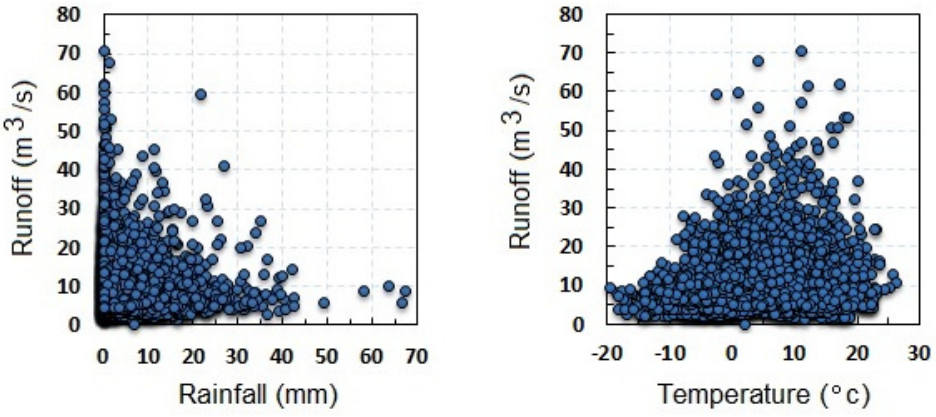

Figure 4.8 : Observational Runoff at Senoz catchment with respect to RCM rainfall and temperature data.

The PI (Kitanidis and Bras 1980) given in equation (4.2) is one-step persistence index that has been recommended to distinguish timing error in data-driven approaches (de Vos and Rientjes 2008; Wu and Chau 2010). The VE, equation (4.3),

is an improvement to NSE for streamflow model evaluation representing the fraction of water delivered at the proper time (Criss and Winston 2008). The KGE, equation (4.4), is a goodness-of-fit measure that provides a diagnostically interesting decomposition of NSE (Gupta et al. 2009). The measure facilitates the analysis of the relative importance of different components of NSE (correlation, bias and variability) in the context of hydrological modelling.

$$PI = 1 - \frac{\sum_{i=1}^{n} (X_i^{obs} - X_i^{pre})^2}{\sum_{i=1}^{n} (X_i^{obs} - X_{i-1}^{obs})^2} \tag{4.2}$$

$$VE = 1 - \frac{\sum_{i=1}^{n} \left| X_i^{pre} - X_i^{obs} \right|}{\sum_{i=1}^{n} X_i^{obs}} \tag{4.3}$$

$$KGE = 1 - ED$$
$$ED = \sqrt{(r-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \tag{4.4}$$

where X_i^{obs} = observed value of X (here runoff), X_i^{pre} = predicted value X_{mean}^{obs} = mean value of observed data, X_{i-1}^{obs} = observed value of X at time instance i-l, n = number of observed data. In equation (4.4), r = linear correlation coefficient, α = the ratio between the variance predicted and variance observed values, β = the ratio between the mean predicted and mean observed values, and ED is the Euclidian distance from the ideal point where all three measures (i.e., r, α, and β) have their ideal values of unity. Higher values for NSE, PI, VE, and KGE (all up to one) and lower values for RMSE (zero for the perfect model) stand for more efficient models.

Table 4.1 : Descriptive statistics of the modelling parameters.

parameters	Runoff (m^3/s)				Rainfall (mm)				Temperature (°C)			
	Entire data	Training	Validation	Testing	Entire data	Training	Validation	Testing	Entire data	Training	Validation	Testing
Number	3648	2191	731	726	3648	2191	731	726	3648	2191	731	726
Maximum	70.7	70.7	61.9	61.5	67.5	67.5	39.0	66.5	26.1	23.6	26.1	22.2
Minimum	1.4	1.4	2.2	2.7	0.0	0.0	0.0	0.0	-19.5	-18.3	-16.8	-19.5
Average	11.1	11.3	11.7	9.9	4.6	4.6	4.4	4.8	3.8	3.9	3.5	4.0
Standard Deviation	7.6	7.5	9.5	5.4	7.0	7.0	6.4	7.7	8.4	8.5	8.3	8.3
Skewness	2.2	2.0	2.0	2.5	2.6	2.5	2.1	2.9	0.0	0.0	0.3	-0.1

4.6 LGP, MGGP, and ANN Modelling

In GP-based modelling, the first step is creating an initial population of genes. To this end, a modeller may specify a primitive sets of terminals and functions. In this study, the predefined input variables of each combination and a set of random constants in the range [-1.0 and +1.0] are chosen as members of terminal sets for the both LGP and MGGP runs. In order to evolve a proper solution, an appropriate guess is required for selection of functions. For highly nonlinear systems, a brute-force search method may be applied as an alternative to realize the best functions (Danandeh Mehr et al. 2013). In the present study, we use three basic arithmetic functions (i.e., addition, subtraction, and multiplication) for both LGP and MGGP runs. The division operation is omitted from the function set to protect the runs from division by zero. Other kinds of functions such as Boolean and conditional functions have been left out in order to avoid increasing the complexity so that we are able to convert the evolved programs or GP trees in mathematical expressions.

After the creation of initial genes, one must define a fitness function to rank individuals. The mean squared error is utilized here as the fitness function for both LGP and MGGP runs. In order to minimise the error between observed and projected outputs, GP operators such as crossover and mutation are used. We applied Discipulus®, the LGP software package developed by Francone (2010), and GPTIPS®, the multigene symbolic regression code developed by Searson (2009) in this study. The other adopted parameters for experimental setup in this study are tabulated in Table 4.2.

Table 4.2 : Setting for LGP and MGGP runs.

Parameter	LGP	MGGP
Initial Population (programs)	100	100
Mutation Rate	% 10	% 10
Crossover Rate	% 50	% 50
Maximum Program Size	512 (bytes)	4 (genes)
Maximum Numbers of generations	300	300
Max tree depth	-----	6
Max nodes per tree	-----	infinitive

In order to find the best FFBP-ANN structures for the combinations (i.e., Comb (1) to (8)), several three layer FFBP networks have been trained and tested by the

Levenberg-Marquardt (LM) algorithm. All of the networks possess a single hidden layer with a sigmoid transfer function and a single output layer with a linear transfer function. Satisfactory application of the sigmoid and linear transfer functions as well as LM algorithm in FFBP networks were widely reported (e.g. de Vos and Rientjes 2005; Chang et al. 2007; Kisi 2010, Londhe and Charhate 2010; Rezaeianzadeh 2013; Danandeh Mehr et al. 2015). Determination of an optimum number of neurons in the hidden layer is one of the most important aspects of ANN modelling. In this study, a brute-force search method among one to 10 neurons is adopted to avoid overtraining as well as to select the optimum number of neurons in the hidden layer. The mean squared error of the validation step at each trial was used as a criterion for selecting the optimal number of neurons in hidden layer. When the number of hidden neurons was increased above the threshold (i.e., 10 neurons in hidden layer), no significant improvement in model performance was found..

4.7 Results and Discussion

4.7.1 Bivariate PR-R models

The efficiency results of the best MGGP, LGP and ANN models calculated for the bivariate combinations (i.e., Combs (1) to (4)) are presented in Table 3. Since a GP system derives different formulas with varying complexity and almost equal accuracy, a pareto-optimal model among the 30 best solutions has been used to select as the optimal model for each combination in this study. The best model in the GP community often is the model providing the best goodness of fit in the validation period. However, the pareto-optimal model is the model that is not outperformed by any other model in terms of both fitness (accuracy) and complexity. To quantify the complexity of a GP solution, the total number of nodes (bytes) of the corresponding MGGP tree (LGP program) has been used in this study.

In order to optimize the complexity and accuracy of the proposed ANN-based models, we initially optimized (fixed) the number of hidden neuron (complexity) for each combinations by brute-force search method. Then, keeping the number of hidden neurons fixed, we performed 30 different runs for each combination to figure out the most accurate model with respect to the validation set.

Considering the NSE, KGE, and VE measures tabulated in Table 4.3, the both LGP and MGGP approaches are satisfactory for bivariate PR-R modelling, whereas the ANN is not. The given RMSE values show that the mean prediction error of GP-based models is less than the minimum observed flow of corresponding subsets (i.e., training, validation, and testing sets, see Table 4.1). By contrast, ANNs resulted in mean prediction error more than twice the minimum observed flow, which warrants superiority of the both LGP and MGGP to the ANN. According to the PI values given for different models/combinations, there is a salient difference between GP-based and ANN models. The reasons for this will become clearer later in this section, but it is useful to keep in mind that the PI focuses more on peak flow accuracy than low flow. Eventually, by comparing the performance of LGP and MGGP models, it can be concluded that LGP outperforms MGGP. Therefore, the LGP approach is evaluated as the best bivariate data-driven method for PR-R modelling in Senoz Catchment. Among different combinations, the LGP model developed for the Comb (2) exhibits the best overall behaviour for the study region, closely followed by MGGP. The dimensionless LGP and MGGP formulas are presented by equations (4.5) and (4.6), respectively. Though we examined two time-steps lag as the memory of the GP and ANN models, Equations (4.7) and (4.8) show that the dominant lag of the catchment is one step.

$$Q_t = Q_{t-1} + Q^3_{t-1}(I_t \times (I_{t-1} - 2.048) - 0.132) \tag{4.5}$$

$$Q_t = 1.075Q_{t-1} - 0.2206I_{t-1} \times Q_{t-1} - 0.2206Q_{t-1}(1.739Q_{t-1} - I_t \times I_{t-1}) + 0.0039 \tag{4.6}$$

The projected runoff hydrographs depicted by equations (4.5) and (4.6) for validation and testing (unseen) periods as well as those of the best ANN model for the Comb (2) are shown in Figure 4.9 and compared with corresponding observational hydrographs in daily time scale. In these hydrographs, the portions of data belonging to validation and testing periods have been magnified. The figure illustrates that the proposed LGP, MGGP, and ANN models are able to pick up the strong periodicity of observational hydrograph. The both LGP and MGGP estimate the global maximum and local maxima satisfactorily, but the ANN outputs shows lagged prediction that is the reason behind low PI values of ANN models tabulated in Table 3. In conclusion,

Lagged predictions can be considered as the main drawback of the ANN method for PR-R modelling.

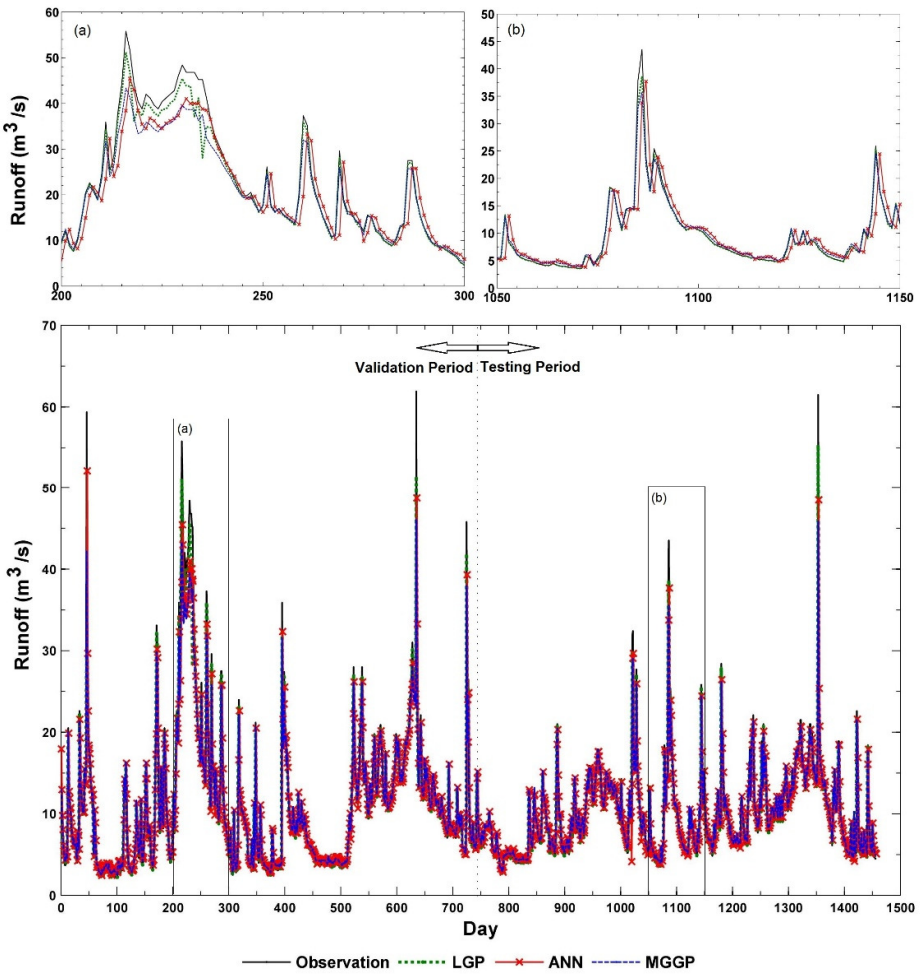

Figure 4.9 : Best bivariate models outputs compared to the observational runoff.

The issue of lagged predictions in the ANN models for naive forecasting has been discussed by some researchers. For example, de Vos and Rientjes (2005) and Wu and Chau (2011) described that this drawback of ANNs originates from relatively high autocorrelation between the target variable and its lags, which are used as model inputs, so that the ANN tends to give the most weight to the latest target variable for calculating the output. To reduce such error (which is out of the scope of our study), application of different data pre-processing methods such as moving averages over time of modelling inputs, singular spectrum analysis, and continues/discrete wavelet transforms integrated with data-driven models have been suggested (Jain and

Srinivasulu 2004; de Vos and Rientjes 2005 and 2008; Muttil and Chau 2006; Wu and Chau 2011).

Table 4.3 : Goodness of fitness results of bivariate PR-R models for Senoz Catchment.

Comb	Model	NSE			PI			KGE			VE			RMSE (m^3/s)		
		Train	Val	Test	Train	Val	Test	Train	Val	Test	Train	Val	Test	Train	Val	Test
	MGGP	0.973	0.958	0.978	0.900	0.803	0.944	0.870	0.838	0.897	0.952	0.936	0.961	1.221	1.957	0.800
1	LGP	0.988	0.972	0.980	0.957	0.870	0.973	0.920	0.890	0.938	0.979	0.967	0.987	0.785	1.590	0.547
	ANN (2-8-1)*	0.759	0.774	0.661	0.101	0.020	0.144	0.830	0.847	0.789	0.844	0.816	0.843	3.662	4.532	3.118
	MGGP	0.973	0.965	0.976	0.900	0.836	0.940	0.871	0.843	0.894	0.949	0.934	0.956	1.223	1.786	0.825
2	LGP	**0.989**	**0.972**	**0.985**	**0.959**	**0.870**	**0.974**	**0.925**	**0.890**	**0.938**	**0.980**	**0.968**	**0.987**	**0.785**	**1.588**	**0.547**
	ANN (3-5-1)	0.759	0.774	0.661	0.101	0.020	0.144	0.831	0.847	0.789	0.844	0.816	0.843	3.661	4.532	3.118
	MGGP	0.973	0.976	0.949	0.899	0.888	0.871	0.871	0.865	0.885	0.948	0.937	0.951	1.226	1.475	1.212
3	LGP	0.983	0.977	0.984	0.938	0.893	0.960	0.904	0.877	0.916	0.967	0.953	0.974	0.961	1.442	0.673
	ANN (4-5-1)	0.760	0.793	0.665	0.105	0.044	0.152	0.812	0.791	0.780	0.842	0.808	0.841	3.652	4.335	3.102
	MGGP	0.967	0.903	0.970	0.879	0.546	0.923	0.874	0.82	0.897	0.947	0.921	0.954	1.344	2.971	0.935
4	LGP	0.974	0.977	0.976	0.902	0.893	0.938	0.899	0.885	0.907	0.972	0.962	0.981	1.208	1.442	0.873
	ANN (5-7-1)	0.769	0.793	0.673	0.159	0.044	0.174	0.82	0.784	0.761	0.848	0.816	0.847	3.583	4.334	3.061

* The ANN architecture is: 2-8-1 (2 input units, 8 hidden neurons, 1 output neuron)

4.7.2 Multivariate PR-R models

As we mentioned previously, in order to investigate the capability of GP and ANN techniques to model the PR-R process using multivariate inputs, four supplementary combinations (i.e., Combs (5) to (8)) are examined in the present study. For this problem, our modelling attempts revealed that inclusion of the RCM temperature in model inputs does not improve the prediction accuracy noticeably, which means temperature is an ineffective input for PR-R modelling. This is consistent with Senoz Catchment hydrology, in which surface runoff is groundwater dominated. Therefore, the potential solutions (formulas) evolved by LGP and MGGP are mostly bivariate programs. Table 4 displays the efficiency results of the optimal multivariate models generated by the LGP, MGGP, and ANN. Among the 30 best potential solutions, the LGP did not generated any multivariate model possessing RCM rainfall and temperature inputs for the Combs (7) and (8), whereas the MGGP did. For this reason we didn't give the LGP's efficiency results in the table. The reason behind such a difference between LGP and MGGP may lie in the robustness of self-tuning feature of Discipulus®, which is higher than GPTIPS®, so that the probability of appearance of ineffective variables (e.g., RCM temperature with two time-steps lag) in LGP-based potential solutions is smaller than in MGGP. This evolutionary optimisation feature of GP systems help modellers to identify more dominant inputs even without a sensitivity analysis among input variables. By contrast, the addition of one or more input nodes in ANN systems increases the model complexity, and therefore the risk of overfitting. As a result, ANN modellers need to optimize the number of input nodes via a brute force method (as we have done in this study) or an external optimization/determination technique (for more information refer to Bowden et al. 2005; Fernandeo et al. 2009).

Similar to the results derived from Table 3, the efficiency measures given in Table 4 indicate the promising role of both LGP and MGGP for multivariate PR-R modelling. The ANN results are still contaminated by timing error and the LGP outperforms both the MGGP and ANN. The dimensionless mathematical expression of the best multivariate LGP program and MGGP tree (i.e., Comb (6)) are presented in Eqs. (4.9) and (4.10), respectively.

$$Q_t = Q_{t-1} - 0.3639\left[0.2878\left(\left(T_t - I_{t-1}\right)I_t\right)^2 - 0.4282\right) + Q_{t-1}\Big]^2 \tag{4.7}$$

$$Q_t = 0.00174 - 0.05647Q_{t-1}\left(I_{t-1}Q_{t-1} + 4.181\right) - 0.4026Q_{t-1}\left(Q_{t-1} + I_tQ_{t-1} + T_tI_{t-1}Q_{t-1}^2 - 3.285\right) \tag{4.8}$$

The projected runoff hydrographs resulted from equations (4.7) and (4.8) as well as those of ANN method at the validation and testing period with respect to the corresponding observational hydrographs are depicted in Figure 4.10. In these hydrographs, the portions of data belonging to validation and testing period have been magnified. Akin to the projections illustrated in Figure 4.9, the low and medium runoff values are mostly well predicted by both LGP and MGGP. This is due to the frequent occurrences of the low flow regime in which allow the trained models to have better generalization of the study catchment. All the previously presented discussions for Figure 4.9 are also valid for Figure 4.10.

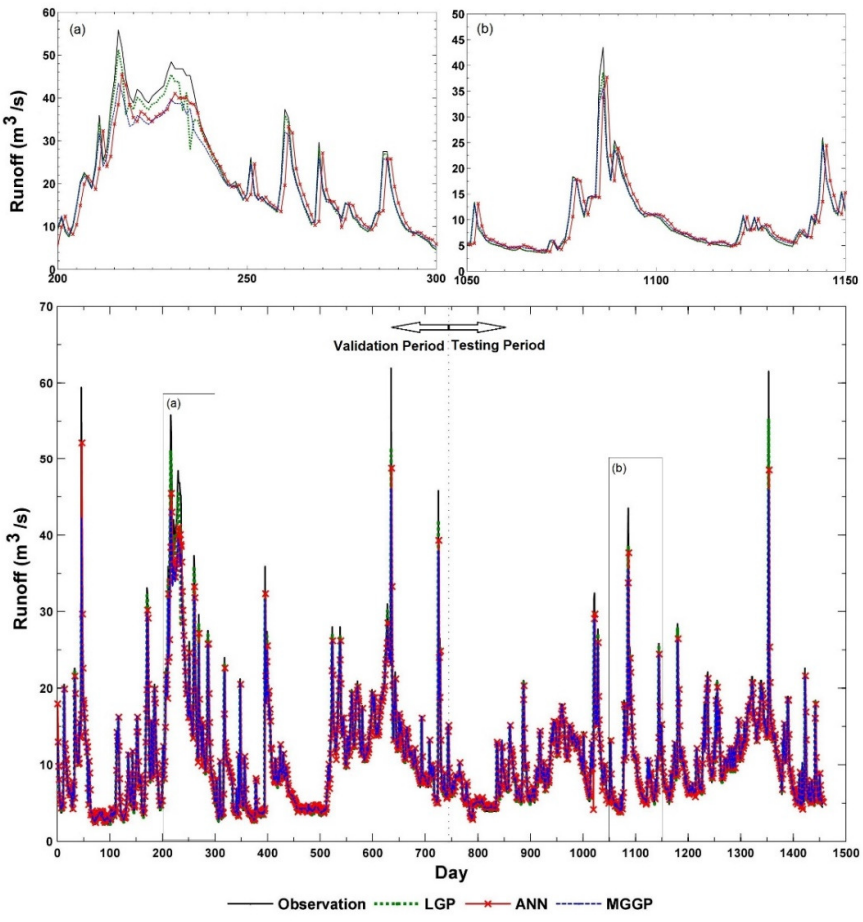

Figure 4.10 : Best multivariate models outputs compared to the observational runoff.

Table 4.4 : Goodness of fitness results of multivariate PR-R models for Senoz Catchment.

Comb	Model	NSE			PI			KGE			VE			RMSE (m^3/s)		
		Train	Val	Test	Train	Val	Test	Train	Val	Test	Train	Val	Test	Train	Val	Test
5	MGGP	0.972	0.951	0.975	0.894	0.773	0.935	0.87	0.828	0.895	0.952	0.932	0.961	1.257	2.101	0.852
	LGP	0.975	0.968	0.980	0.906	0.850	0.950	0.904	0.874	0.922	0.977	0.963	0.987	1.183	1.71	0.756
	ANN (3-6-1)*	0.760	0.799	0.661	0.104	0.069	0.144	0.825	0.800	0.790	0.843	0.812	0.843	3.654	4.278	3.117
6	MGGP	0.972	0.958	0.975	0.897	0.804	0.938	0.871	0.835	0.892	0.951	0.934	0.958	1.236	1.954	0.838
	LGP	**0.979**	**0.974**	**0.983**	**0.925**	**0.879**	**0.961**	**0.897**	**0.870**	**0.922**	**0.972**	**0.956**	**0.984**	**1.059**	**1.535**	**0.665**
	ANN (5-9-1)	0.759	0.796	0.670	0.102	0.056	0.166	0.811	0.789	0.787	0.837	0.808	0.834	3.658	4.308	3.077
7	MGGP	0.973	0.964	0.975	0.898	0.833	0.938	0.872	0.843	0.896	0.949	0.934	0.956	1.234	1.803	0.839
	ANN (4-5-1)	0.757	0.791	0.670	0.094	0.033	0.167	0.81	0.793	0.790	0.841	0.811	0.842	3.674	4.36	3.075
8	MGGP	0.975	0.966	0.982	0.908	0.841	0.954	0.878	0.849	0.901	0.953	0.913	0.961	1.174	1.756	0.726
	ANN (8-1-1)	0.758	0.791	0.672	0.097	0.043	0.171	0.829	0.804	0.807	0.844	0.815	0.846	3.668	4.343	3.068

* The ANN architecture is: 3-6-1 (3 input units, 6 hidden neurons, 1 output neuron)

In the last phase of the chapter, a visual evaluation is presented in Fig. 11 to compare the variation of efficiency measures among all combinations. It can be seen that the LGP model evolved for Comb (2) provides the best overall behaviour for the study region, closely followed by MGGP. In addition, the figure demonstrates that the PI provides the most meaningful difference on the prediction accuracy of LGP, MGGP, and ANN. Although KGE or VE are introduced in the literature as improvements for NSE, they do not eliminate the need for other useful metrics such as PI for naive forecast models.

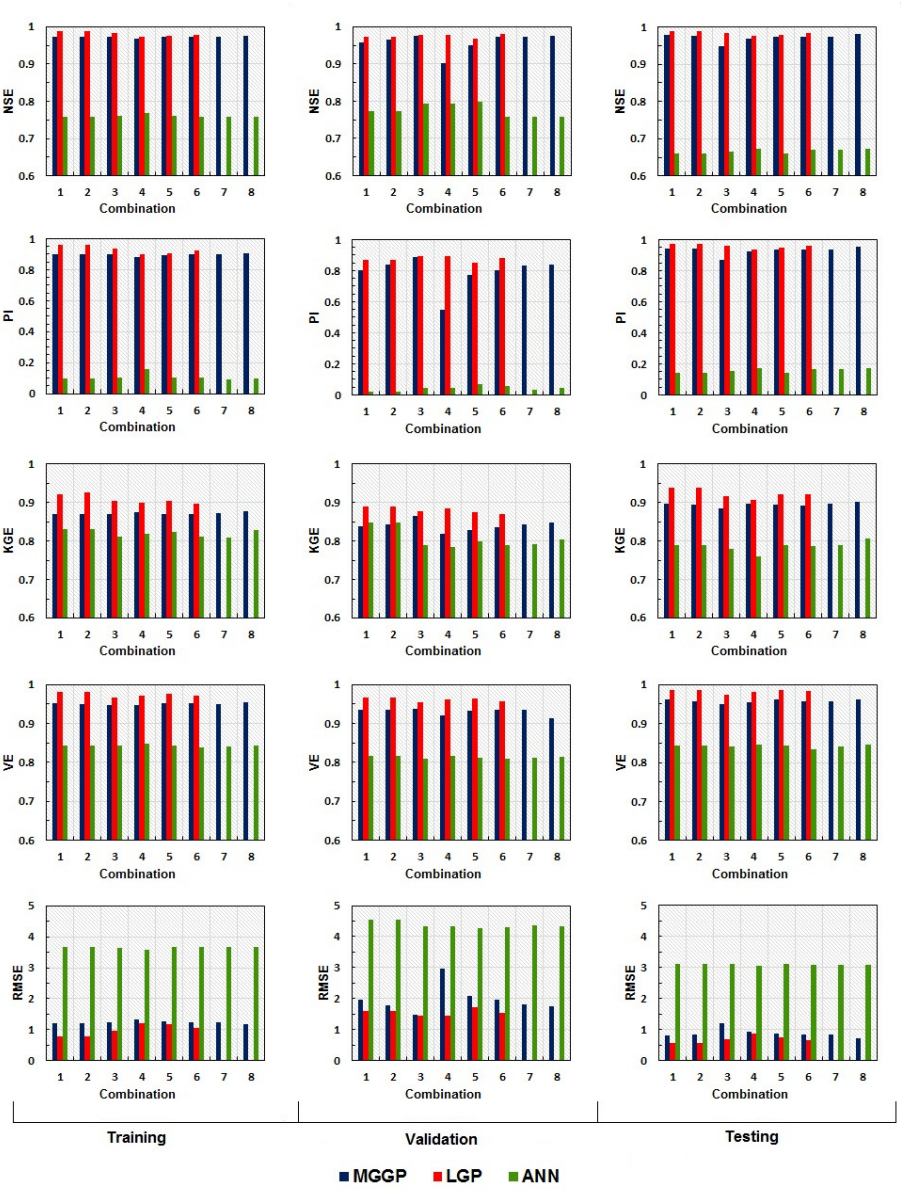

Figure 4.11 : Variation in efficiency measures of the best MGGP, LGP, and ANN models for Senoz catchment.

5. CONCLUSIONS AND RECOMMENDATIONS

Catchment-scale projected extreme rainfall estimation is an important task in water resources engineering, which this study has just begun to address for Rize Province, Turkey. The task has been discussed and accomplished in two parts; the first part, which was presented in chapter 2, is motivated by the idea of grid-based performance evaluation and intercomparison of two GCM-RCM combinations for rainfall reproduction in the province. The second part, presented in chapter 3, was completed by the grid-based extreme value analysis of projected AMR series under SRES-A2 and RCP8.5 greenhouse gas scenarios. In addition, a GP-based PR-R model, for the firs time, was presented in Chapter 4 that is motivated by the idea of improving the hydrological modelling techniques and better understanding of different kinds of data and their possible uses in poorly gauged catchments. Based upon the results discussed in pervious chapters, this section presents the concluding remarks and recommendations for future studies relevant to each of these chapters.

Chapter 2 demonstrates that catchment-scale performance evaluation of GCM-RCM combinations is required for selecting a proper set of GCM-RCM combinations to run hydrological models in small catchments. This important issue was addressed in this study to investigate precipitation realization performance of two GCM-RCM combinations, ECHAM5-RegCM3 and CCSM3-RegCM3, over 10 small to midsize catchments situated in Rize Province, North-eastern part of Turkey. To this end, different daily, monthly and annual precipitation statistics as well as two performance evaluation frameworks, grid- and portion-based, are implemented, which are of interest to a wide audience of hydrologists and practitioners. The results showed that both ECHAM5-RegCM3 and CCSM3-RegCM3 reproduce the monthly rainfall cycle rather weak. In general, the former tends to overestimate total monthly rainfall while the latter yields slightly dryer condition in the entire province. Consequently, CCSM3-RegCM3 underestimates observational total annual rainfall. For monthly and annual maximum rainfall statistics, ECHAM5-RegCM3 resulted in higher performance than CCSM3-RegCM3, particularly in the inland portion.

Relating to the number of rainy days in different parts of the study area, ECHAM5-RegCM3 tends to overestimate the relatively constant rainy days of entire province. By contrast, CCSM3-RegCM3 underestimates the statistic in the range 20-50 days. Since the major uncertainty of both models appeared in the coastland portion, implementation of model projections in this portion evidently requires special cares such as a bias correction procedure.

The evaluation performed in Chapter 2 also showed that both of the GCM-RCM combinations reproduces unsatisfactory results for the entire province. When a combination performs well for a portion/grid of study region, it is not necessarily a well combination for the other portions/grids. Similarly, the evaluation for different variables showed that no combination was best/worst in all variables even though all of the variables are from a same family (i.e., rainfall-born statistics). In other words, if a combination performs well for reproducing a specific statistic, it is not necessarily a well combination to reproduce the other statistics. These conclusions are consistent with those of reported by Kjellström et al. (2010) at regional-scale evaluation of GCM-RCM combinations. Therefore, it is suggested that when the interest is impact assessment at even relatively small areas, as it is the case of FRA-PFC Rize project, validation of GCM-RCM outputs should be conducted at a single-grid level, rather than at regional scales. In this case, it is also necessary to check models' performance in reproducing different prescribed observations in order to make a fair judgment among models available for the catchment of interests.

With respect to SRES-A2 and RCP8.5 scenarios, **Chapter 3** assessed the postulate climate change impacts on extreme rainfall variability over the Rize Province, Turkey. To this end, outputs of two GCM-RCM integrations (i.e., ECHAM5-RegCM3 and HadGEM2-RegCM4) at three future time periods 2013-2040, 2041-2070, and 2071-2100 were campared to the extreme events estimated by observational records at the reference period 1961-1990. Using five different probability distribution models, the EVT analysis were applied to the data to recognize extreme rainfall pattern over the study area. In order to find the best distribution, the K-S and A-D goodness of fit tests are used. In addition, a new bias correction method was developed to adjust the climate models outputs. The Results of the control run evaluation, bias correction procedure, and information from the future extreme rainfall events were described, including the spatial and temporal

distribution of 2-, 10- , 25-, 50-, 100-, and 200-yr storms as well as the appropriateness of representative distribution models at each RCM grids. This chapter presented two key results regarding the use of dynamically downscaled climate data sets over Rize Province. First, applying a bias correction scheme to the dynamically downscaled data sets is a necessary and important step yielding more accurate results. For the both ECHAM5-RegCM3 and HadGEM-RegCM4 data sets, the proposed bias correction procedure not only matched the mean of the AMR series, as expected, but also significantly improved other statistics such as median, 25th percentile and 75th percentile. In addition, the method effectively reduced PBIAS as well as the difference between PDFs of observed and simulated rainfall over the study region. From a practical standpoint, the method is rapid and has a relatively simple structure, which are of interest to a wide audience of practitioners. This encourages the application of this method in extreme rainfall studies. Second, implementing fundamentals of EVT to projected data sets provides estimates of changes to variable values, such as rainfall, because of climate change. Using the different distribution models and emission scenarios, this study obtained estimates of changes to extreme rainfall intensities over the Rize Province. The ensemble mean results indicated that the extreme rainfall intensities will decrease in the all future periods compared with simulations over the reference period 1961–1990. In addition, checking the appropriateness of the distributions using two goodness-of-fit tests demontrated that the GEV is the best model to characterize the behavior of the projected extreme rainfalls over the study region. This conclusion is consistent with the results of a similar study accomplished by Solaiman (2011), which reports the GEV as the most appropriate distributions model for the Upper Thames River Watershed under different climate models.

While almost all previous studies applies knowledge-driven models, such as physically based or lumped hydrological models, to understand the hydrological response of a river basin to climatic change, **chapter 4,** for the first time, investigated the data-driven techniques of LGP, MGGP, and ANN to model projected runoff in a poorly gauged catchment. the results indicated a promising role of both LGP and MGGP for PR-R modelling in general, and showed higher performance level of LGP over MGGP and FFBP neural networks particularly in Senoz catchment. The best bivariate LGP model showed a NSE fitness of 0.972 and 0.985 for validation and

unseen data sets, respectively. From a point of timing accuracy, the results also demonstrated the outstanding superiority of GP-based methods over the ANN. According to the PI values, the poor performance of ANN models was mainly due to the existence of the lagged prediction. Since, most data-driven hydrological models have not addressed this issue, the thesis strongly recommend the PI measure be used for comparing observed and data-driven-modelled forecasts particularly when antecedent values of predictand are used as predictor. The chapter was limited to (i) one-day ahead forecasting combinations and (ii) row outputs of a single GCM-RCM integration forced by one greenhouse gas emission scenario. Since the reliability of runoff projections depends on the accuracy of RCM rainfall and temperature feed into the PR-R model, implementation of an ensemble of bias corrected GCM-RCM outputs is suggested for real world applications. To more efficiently use the runoff projections for many water resource applications, one way to future work, following the presented methodology, could be testing the capability of GP-based techniques for PR-R modelling on a long-term basis. From the standpoint of improving modelling results, performing multiple GP runs at different mutation and crossover rates may also be interesting for future studies.

REFERENCES

Abrahart, R. J., Anctil, F., Coulibaly, P., et al. (2012). Two decades of anarchy? Emerging themes and outstanding challenges for neural network modelling of surface hydrology. *Prog. Phys. Geog., 36* (4)*,* 480-513.

Afiesimama, E. A., Pal, J. S., Abiodun, B. J., Gutowski, W. J., Adedoyin A. (2006). Simulation of West African monsoon using the RegCM3. Part I: model validation and inter annual variability. *Theor. Appl. Climatol., 86*, 23–37.

Ahmad, M.I. Sinclair C.D., Spurr, B.D. (1988). Assessment of Flood Frequency Models Us-ing Empirical Distribution Function Statistics, *Water Resour. Res., 24* (8)*,* 1323-1328.

Alcamo, J., Döll, P., Kaspar, F., Siebert, S. (1997). *Global change and global scenarios of water use and availability: an application of WaterGAP 1.0.* Center for Environmental Systems Research (CESR), University of Kassel, Germany, 1720.

Ajami, N. K., Gupta H., Wagener T., Sorooshian S. (2004). Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system, *J. Hydrol., 298* (1-4), 112-135.

Anandhi, A., Nanjundiah R. S. (2015) Performance evaluation of AR4 climate models in simulating daily precipitation over the Indian region using skill scores. *Theor. Appl. Climatol.*, *119*, 551-566.

Arnold, J. G., Williams, J. R., Srinivasan, R., King K. W. (1999). *SWAT: soil and water assessment tool*. Temple, USDA, Agricultural Research Service.

ASCE Task committee. (2000). Artificial neural networks in hydrology: hydrologic applications. *J. Hydrol. Eng., 5* (2)*,* 124–137.

Babovic, V., Keijzer, M. (2000). Genetic programming as a model induction engine. *J. Hydroinform.*, *2*, 35–60.

Bardossy, A. (2007). Calibration of hydrological model parameters for ungauged catchments, *Hydrol. Earth Syst. Sci., **11***, 703–710.

Barfus, K., Bernhofer, C. (2014). Assessment of GCM performances for the Arabian Peninsula, Brazil, and Ukraine and indications of regional climate change. *Environ. Earth Sci.*, *72* (12)*,* 4689-4703.

Barnett, T. P., Pierce, D. W., Hidalgo, H. G., et al. (2008). Human-induced changes in the hydrology of the western United States. *Science*, *319* (5866), 1080-1083.

Bastola, S., Murphy, C., Sweeney, J. (2011). The role of hydrological modelling uncertainties in climate change impact assessments of Irish river catchments. *Adv. Water Resour., 34* (5), 562-576.

Bekele, E. G., Knapp, H. V. (2010). Watershed modeling to assessing impacts of potential climate change on water supply availability. *Water Resour. Manag*., *24* (13), 3299-3320.

Bergström, S. (1976). *Development and application of a conceptual runoff model for Scandinavian catchments.* Department of Water Resources Engineering, Lund Institute of Technology, Bulletin Series A-52, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden.

Beniston, M., Stephenson D. B., Christensen, O. B., et al. (2007). Future extreme events in European climate: An exploration of regional climate model projections. *Clim. Change, 81* (1), 71–95.

Berg, P., Feldmann, H., Panitz, H. J. (2012). Bias correction of high resolution regional climate model data. *J. Hydrol., 448-449*, 80-92.

Betts, R.A., Boucher O., Collins M., et al. (2007). Projected increase in continental runoff due to plant responses to increasing carbon dioxide. *Nature*, *448*, 1037-1041.

Bozkurt, D. (2013). Climate change impacts in the Euphrates–Tigris Basin. *PhD Thesis,* ITU, Istanbul, Turkey.

Bozkurt, D., Sen O. L. (2013). Climate change impacts in the Euphrates–Tigris Basin based on different model and scenario simulations. *J. Hydrol., 480*, 149-161.

Bozkurt, D., Sen, O. L., Hagemann, S. (2015). Projected river discharge in the Euphrates-Tigris Basin from a hydrological discharge model forced with RCM and GCM outputs. *Climate Res., 62*, 131-147.

Bozkurt, D., Turuncoglu, U., Sen, O. L. et al. (2012) Downscaled simulations of the ECHAM5, CCSM3 and HadCM3 global models for the eastern Mediterranean-Black Sea region: Evaluation of the reference period. *Clim. Dynam., 39*, 207-225.

Boville, B. A., Rasch, P. J., Hack, J. J., McCaa, J. R. (2006). Representation of clouds and precipitation processes in the Community Atmosphere Model (CAM3). *J. Climate, 19,* 2184-2198.

Bowden, G. J., Dandy, G. C., Maier, H. R. (2005) Input determination for neural network models in water resources applications. Part 1—background and methodology. *J. Hydrol., 301* (1), 75-92.

Brameier, M., Banzhaf, W. (2007). Linear genetic programming. Springer Science + Business Media, LLC, New York.

Bulygina, N., McIntyre, N., Wheater, H. 2009. Conditioning rainfall-runoff model parameters for ungauged catchments and land management impacts analysis. *Hydrol. Earth Syst. Sci., 13* (6), 893-904.

Burnash, R.J.C., Ferral, R.L. McGuire, R.A. (1973). *A generalized streamflow simulation system, conceptual modeling for digital computer*. U.S. Department of Commerce, National Weather Service and State of California, Department of Water Resources, Sacramento, CA.

Chang, F.J., Chiang, Y.M., Chang, L.C. (2007). Multi-step-ahead neural networks for flood forecasting. *Hydrolog. Sci. J., 52* (1),114–130.

Chen, B., Chao, W.C., Liu, X. (2003) Enhanced climatic warming in the Tibetan Plateau due to doubling CO2: a model study. *Clim. Dynam., 20*, 401–413.

Chen, H., Wang, S., Gao, Z., Hu, Y. (2010). Artificial neural network approach for quantifying climate change and human activities impacts on shallow groundwater level: A case study of Wuqiao in North China Plain. In Geoinformatics, *18th International Conference on* (pp. 1-6). IEEE.

Christensen, J. H., Carter, T. R., Giorgi, F. (2002) PRUDENCE employs new methods to assess European climate change. *EOS*. 83:13.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values. Springer Series in Statistics*. Springer Verlag. London.

Collins, W.D., Bitz, C.M., Blackmon, M.L. et al. (2006) The Community Climate System Model Version 3 (CCSM3). *J. Climate, 19*, 2122–2143.

Criss, R.E., Winston, W.E. (2008). Do Nash values have value? Discussion and alternate proposals. *Hydrol. Process., 22* (14), 2723.

Danandeh Mehr, A., Kahya, E. (2016). Catchment-scale performance evaluation of two GCM-RCM combinations for rainfall reproduction. *Theor. Appl. Climatol.* DOI 10.1007/s00704-016-1758-1. (in press).

Danandeh Mehr, A., Kahya, E. and Özger M. (2014a). A gene-wavelet model for long lead-time drought forecasting. *J. Hydrol., 517*, 691-699.

Danandeh Mehr, A., Kahya, E. Yerdelen, C. (2014b). Linear genetic programming application for successive-station monthly streamflow prediction. *Comput. Geosci., 70*, 63-72.

Danandeh Mehr, A., Kahya E., Şahin, A. Nazemosadat M.J. (2015). Successive-station monthly streamflow prediction using different ANN algorithms. *Int. J. Environ. Sci. Technol., 12* (7), 2191-2200.

Danandeh Mehr, A., Kahya E. and Olyaie E. (2013). Streamflow prediction using linear genetic programming in comparison with a neuro-wavelet technique. *J. Hydrol., 505*, 240–249.

Dash, S. K., Shekhar, M. S., Singh, G. P. (2006). Simulation of Indian summer monsoon circulation and rainfall using RegCM3. *Theor. Appl. Climatol., 86*,161−172.

Davie, J.C.S., Falloon, P.D., Kahana, R. et al. (2013). Comparing projections of future changes in runoff from hydrological and biome models in ISI-MIP. *Earth Syst. Dynam., 4*, 359-374.

Deidda, R., Marrocu, M., Caroletti, G. et al. (2013) Regional climate models' performance in representing precipitation and temperature over selected Mediterranean areas, *Hydrol. Earth. Syst. Sci., 17*, 5041-5059.

Demircan, M., Demir, Ö., Atay, H., Et Al. (2014). Climate change projections for turkey with new scenarios. *The climate change and climate dynamics conference (CCCD)*, Istanbul, Turkey, September 8-10.

Demirel, M. C., Booij, M. J., Hoekstra, A. Y. (2013) Impacts of climate change on the seasonality of low flows in 134 catchments in the River Rhine basin using an ensemble of bias-corrected regional climate simulations. *Hydrol. Earth. Syst. Sci., 17*, 4241–4257.

Demirel, M. C., Moradkhani, H. (2016) Assessing the impact of CMIP5 climate **multi**-modeling on estimating the precipitation seasonality and timing. *Clim. Change,* 135(2), 357-372.

Dettinger, M. D., Cayan, D. R., Meyer, M. K., and Anne, E. J. (2004). Simulated hydrological responses to climate variations and change in the Merced, Carson, and American river basin, Sierra Nevada, California, 1900–2099. *Clim. Change,* 62, 283–317.

de Vos, N.J., Rientjes, T.H.M. (2005). Constraints of artificial neural networks for rainfall–runoff modeling: trade-offs in hydrological state representation and model evaluation. *Hydrol. Earth Syst. Sci., 9*, 111–126.

de Vos, N.J., Rientjes, T.H.M. (2008). Correction of timing errors of artificial neural network rainfall-runoff models. in: Abrahart, R.J., See, L.M., Solomatine, D.P. (Eds.), Practical Hydroinformatics, Computational Intelligence and Technological Developments in Water Applications Springer Berlin Heidelberg. pp. 101-112.

Di Luca, A., de Elía, R., Laprise, R. (2011) Potential for added value in precipitation simulated by high-resolution nested regional climate models and observations. *Clim. Dynam.,* 38, 1229–1247.

Dobler, C., et al. (2012). Quantifying different sources of uncertainty in hydrological projections in an Alpine watershed. *Hydrol. Earth. Syst. Sc., 16* (11), 4343-4360.

Ekström, M., Fowler, H. J., Kilsby, C. G., Jones P. D. (2005) New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 2. Future estimates and use in impact studies. *J. Hydrol., 300*, 234–251.

El Adlouni, S., Ouarda, T. B. M. J. (2010). Frequency Analysis of Extreme Rainfall Events. Rainfall: State of the Science. *Geophys. Monogr., 191*, 171-188.

El Adlouni, S., Bobée, B., Ouarda, T.B.M.J. (2008). On the tails of extreme event distributions, *J. Hydrol., 355*, 16–33.

Eris, E., Agiralioglu, N. (2009) Effect of Coastline Configuration on Precipitation Distribution in Coastal Zones. *Hydrol. Process., 23* (25), 3610-3618.

Errasti, I., Ezcurra, A., Sáenz, J., Ibarra-Berastegi, G. (2011). Validation of IPCC AR4 models over the Iberian Peninsula. *Theor. Appl. Climatol., 103* (1-2), 61-79.

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

Foley, A. M. (2010). Uncertainty in regional climate modelling: A review. *Prog. Phys. Geog., 34* (5), 647-670.

Fowler, H. J., Ekström, M., Kilsby, C. G., Jones, P. D. (2005). New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1: Assessment of control climate. *J. Hydrol., 300*, 212–233.

Francone, F. D. (1998). *Discipulus™ software owner's manual, version 3.0 DRAFT.* Machine Learning Technologies Inc., Littleton, CO, USA.

Francone, F. D. (2010). *Discipulus™ with Notitia and Solution Analytics Owner's Manual*. Register Machine Learning Technologies, Inc., Littleton, CO, USA.

Frei, C., Schöoll, R., Fukutome, S. et al. (2006). Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models. *J. Geophys. Res., 111*, D06105.

Fu, G., Liu, Z., Charles, S. P., Xu, Z., Yao, Z. (2013). A score-based method for assessing the performance of GCMs: A case study of south-eastern Australia*. J. Geophys. Res-Atmos. 118*, 4154–4167.

Furrer, E. M., Katz, R. W. (2008). Improving the simulation of extreme precipitation events by stochastic weather generators. *Water Resourc. Res., 44*, W12439.

Gao, XJ, Giorgi F. (2008). Increased aridity in the Mediterranean region under greenhouse gas forcing estimated from high resolution regional climate model projections. *Global Planet. Change, 62*, 195–209.

Gandomi, A.H., Alavi, A.H. (2012). A new multi-gene genetic programming approach to nonlinear system modeling. Part I: materials and structural engineering problems. *Neural. Comput. Appl., 21*, 171–187.

Ghorbani, M., Khatibi, R., Aytek, A., Makarynskyy, O. Shiri, J. (2010). Sea water level forecasting using genetic programming and artificial neural networks. *Comput. Geosci., 36* (5), 620–627.

Giorgi, F. (2006). Climate change hot-spots. *Geophys. Res. Lett., 33*, L08707.

Gosain, A.K., Rao, S., Basuray, D. (2006). Climate change impact assessment on hydrology of Indian River basins. *Curr. Sci., 90* (3), 346–353.

Graham, L. P., Andréasson, J., Carlsson, B. (2007a) Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods – a case study on the Lule River basin. *Clim. Change*, *81*, 293–307.

Graham, L. P., Hagemann, S., Jaun, S., Beniston., M. (2007b) On interpreting hydrological change from regional climate models. *Clim. Change*, 81, 97–122.

Grossi, G., Caronna, P., Ranzi, R . (2013). Hydrologic vulnerability to climate change of the Mandrone glacier (Adamello-Presanella group, Italian Alps) *Adv. Water Resour., 55*, 190-203.

Gupta, H. V., Kling, H., Yilmaz, K. K., Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol., 377* (1), 80-91.

Guven, A. (2009). Linear genetic programming for time-series modelling of daily flow rate. *J. Earth Syst. Sci., 118* (2), 137–146.

Hagemann, S., Chen, C., Clark, D. B. et al. (2013). Climate change impact on available water resources obtained using multiple global climate and hydrology models. *Earth Syst. Dynam., 4*, 129–144.

Halmstad, A., Najafi, M. R., Moradkhani, H. (2013). Analysis of precipitation extremes with the assessment of regional climate models over the Willamette River Basin, USA. *Hydrol. Process., 27*, 2579–2590.

Harvey, L. D. D, Wigley T. M. L. (2003). Charactering and comparing control-run variability of eight coupled AOGCMs and of observations part 1: Temperature, *Climate Dynam., 21*, 619–646.

Hattermann, F. F., Weiland, M., Huang, S., et al. (2011). Model-supported impact assessment for the water sector in Central Germany under climate change—a case study. *Water Resour. Manag. 25* (13), 3113-3134.

Hofstadter, R., Bidegain, M. (1997). Performance of general circulation models in south-eastern South America. *Climate Res. 9*, 101-110.

Hornik, K., Stinchcombe, M., White, M., (1989). Multi-layer feed forward networks are universal approximators. *Neural Networks, 2*, 359–366.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., et al. (2013). A decade of Predictions in Ungauged Basins (PUB)-a review. *Hydrolog. Sci. J. 58* (6), 1198-1255.

Hykin, S. (2008) Neural Networks and Learning Machines (3rd Edition). McMaster University Hamilton, Ontario, Canada.

IPCC (2007). Intergovernmental Panel on Climate Change fourth assessment report on scientific aspects of climate change for researchers, students, and policymakers.

IPCC (Intergovernmental Panel on Climate Change). (2013). Climate change 2013: the physical science basis. Contribution of Working Group I to the 5th Assessment Report to the IPCC. Cambridge University Press, Cambridge.

Islam, A., Sikka, A. K., Saha, B., Singh, A. 2012. Streamflow response to climate change in the Brahmani River Basin, India. *Water Resour. Manag. 26* (6), 1409-1424.

Jacob D, Bärring, L, Christensen, O. B., et al. (2007). An inter-comparison of regional climate models for Europe: model performance in present-day climate. *Climatic Change*, *81*, 31–52.

Jain, A., Srinivasulu, S. (2004). Development of effective and efficient rainfall–runoff models using integration of deterministic, real-coded genetic algorithms and artificial neural network techniques. *Water Resour. Res. 40*, W04302.

Johnson, F, Sharma A. (2011). Accounting for interannual variability: A comparison of options for water resources climate change impact assessments. *Water Resour. Res. 47*, W04508.

Jones, P.D, Reid P.A. (2001). Assessing future changes in extreme precipitation over Britain using regional climate model integrations. *Int. J. Climatol. 21*, 1337–1356.

Kamali, B., Mousavi, S.J., Abbaspour, K.C. 2013. Automatic calibration of HEC-HMS using single-objective and multi-objective PSO algorithms. *Hydrol. Process. 27* (26), 4028-4042.

Khan, M., Azamathulla, H.M.d., Tufail, M. (2012). Gene expression programming to predict pier scour depth using laboratory data. *J. Hydroinform. 14* (3), 628–645.

Kisi O. (2010). Wavelet regression model for short-term streamflow forecasting. *J. Hydrol. 389*, 344–353

Kitanidis, P.K. (1997). *Introduction to geostatistics: applications to hydrogeology. Cambridge University Press,* 249pp.

Kitadinis, P.K. Bras, R.L. (1980). Real-Time forecasting with a conceptual hydrologic model, applications and results, *Water Resour. Res.16*, 1034-1044.

Kjellström, E., Boberg, F., Castro, M., Christensen, H.J., Nikulin G., Sánchez E. (2010) Daily and monthly temperature and precipitation statistics as performance indicators for regional climate models. *Climate Res. 44*,135-150.

Kling, H., Fuchs, M., Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydrol.* ***424***, 264-277.

Krysanova, V., Müller-Wohlfeil, D. I., Becker, A. (1998). Development and test of a spatially distributed hydrological/water quality model for mesoscale watersheds. *Ecol. Model. 106* (2), 261-289.

Konare, A., Zakey, A., Solmon, F., Giorgi, F., Rauscher, S., Ibrah, S. S., Bi, X. (2008) A regional climate modeling study of the effect of desert dust on the West African monsoon. *J Geophys Res 113*, D12206.

Krichak, S. O., Alpert, P., Bassat, K., Kunin, P. (2007). The surface climatology of the eastern Mediterranean region obtained in a three-member ensemble climate change simulation experiment. *Adv. Geosci. 12*, 67–80.

Koza, J. R. (1992). *Genetic Programming: On the Programming of Computers by means of Natural Selection*. MIT Press. Cambridge, MA.

Kundzewicz, Z. W., Mata L. J., Arnell N. W. et al. (2008). The implications of projected climate change for freshwater resources and their management, *Hydrolog. Sci. J. 53* (1), 3-10.

Kundzewicz, Z. W., Kanae, S., Seneviratne S. I., et al. (2013). Flood risk and climate change: global and regional perspectives. *Hydrolog. Sci. J. 59* (1), 1–28.

Lespinas, F., Ludwig, W., Heussner, S. (2014). Hydrological and climatic uncertainties associated with modeling the impact of climate change on water resources of small Mediterranean coastal rivers. *J. Hydrol. 511*, 403–422.

Lin, S. J., Rood, R. B. (1996). Mutlidimensional flux form semi-Lagrangian transport. *Mon. Wea. Rev*., *124*, 2046-2068.

Lionello, P., Boldrin, U., Giorgi, F. (2008). Future changes in cyclone climatology over Europe as inferred from a regional climate simulation. *Clim. Dynam., 30*, 657–671.

Londhe, S., Charhate, S. (2010). Comparison of data-driven modelling techniques for river flow forecasting. *Hydrolog. Sci. J., 55* (7), 1163-1174.

Legates, D. R., McCabe Jr., G. J. (1999). Evaluating the use of goodness-of-fit measures in hydrologic and hydroclimatic model validation. *Water Resour. Res., 35* (1), 233–241.

Lee, J. W., Ham, S., Hong, S. Y., Yoshimura, K., Joh, M. (2014). Future changes in surface runoff over Korea projected by a regional climate model under A1B scenario. *Adv. Meteoro.,*753790.

Liu Z, Mehran A, Phillips TJ, AghaKouchak A. (2014). Seasonal and regional biases in CMIP5 precipitation simulations. *Climate Res., 60*, 35–50.

Marshall, E., Randhir, T. (2008). Effects of climate change on watershed system: a regional analysis. *Clim. Change, 89*, 263–280.

Massey, F. J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit." *J. Amer. Statist. Assoc., 46* (253), 68–78.

Maxino, C. C., Mc Avaney B. J., Pitman A. J., Perkins S. E. (2008). Ranking the AR4 climate models over the Murray-Darling Basin using simulated maximum temperature, minimum temperature and precipitation. *Int. J. Climatol., 28*, 1097–1112.

May, W. (2008). Potential future changes in the characteristics of daily precipitation in Europe simulated by the HIRHAM regional climate model. *Clim. Dynam*., *30*, 581–603.

Meral, G., et al. (2014). The Comparison of Menarcheal Age with Familial Pattern and Body Mass Index. *Science Journal of Public Health,* 2 (6), 584-588.

Meral, G., Uslu, A., Yozgatli, A. U. and Akçay, F. (2015) "Association of Body Mass Index and Lipid Pro- files in Children." *O. J. Ped.,* 5, 141-146.

Mirhosseini, G., Srivastava, P., and Fang, X. (2014). Developing Rainfall Intensity-Duration-Frequency Curves for Alabama under Future Climate Scenarios Using Artificial Neural Networks. *J. Hydrol. Eng.,* 10.1061/(ASCE)HE.1943-5584.0000962, 04014022.

Mirhosseini, G., Srivastava, P., and Sharifi, A. (2015). Developing Probability-Based IDF Curves Using Kernel Density Estimator. *J. Hydrol. Eng.,* 10.1061/(ASCE)HE.1943-5584.0001160, 04015002

Meylan P., Favre A. C., Musy, A. (2012). *Predictive Hydrology A Frequency Analysis Approach*. Science Publishers.

Millington, N., Das, S., Siminovic, S. P. (2011). *The Comparison of GEV, Log-Pearson Type 3 and Gumbel Distributions in the Upper Thames River Watershed under Global Climate Models*. Report No. 077, University of Western Ontario, Canada.

Milly, P. C. D., Dunne, K. A. Vecchia A. V. (2005). Global pattern of trends in streamflow and water availability in a changing climate. *Nature*, *438* (7066), 347–350.

Minville, M., Brissette, F., Leconte, R. (2008). Uncertainty of the impact of climate change on the hydrology of a Nordic watershed*. J. Hydrol., 358* (1), 70-83.

Mladjic, B., Sushama, L., Khaliq, M. N. et al. (2011). Canadian RCM projected changes to extreme precipitation characteristics over Canada. *J. Climate, 24,* 2565–2584.

Mondal, A., Mujumdar P. P. (2012). On the basin-scale detection and attribution of human-induced climate change in monsoon precipitation and streamflow, *Water Resour. Res., 48*, W10520.

Moradkhani, H., Baird, R. G., and Wherry, S. (2010). "Impact of climate change on floodplain mapping and hydrologic ecotones." *J.* Hydrol., *(Amsterdam), 395*, 264–278.

Muttil, N., Chau, K. W., (2006). Neural network and genetic programming for modelling coastal algal blooms. *Int. J. Environ. Pollut., 28* (3–4), 223–238.

Najafi, M. R., Moradkhani, H., Piechota, T. C. (2011a). Ensemble streamflow prediction: climate signal weighting vs. climate forecast system reanalysis. *J. Hydrol., 442-443*, 105–116.

Najafi M. R., Moradkhani, H., Piechota T. C. (2011b) Statistical downscaling of precipitation using machine learning with optimal predictor selection. *J. Hydrol. Eng., 16* (8), 650-664.

Nakićenović, N. et al. (2000) *IPCC special report on emissions scenarios,* Cambridge University Press, Cambridge, UK, p 599.

Nash, J.E., Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I – a discussion of principles. *J. Hydrol., 10* (3), 282–290.

Nazemosadat, M. J., Ravan, V., Kahya, E., Ghaedamini, H. (2016). "Projection of Temperature and Precipitation in Southern Iran Using ECHAM5 Simulations." *Iran. J. Sci. Technol. Trans. Sci.,* in press.

Neitsch, S.L., et al., (2002). Soil and Water Assessment Tool Theoretical Documentation, Version 2000. Grassland, Soil and Water Research

Laboratory, Temple, Texas and Blackland Research Center, Temple, Texas. Texas Water Resources Institute, College Station, Texas.

Notaro, V., Liuzzo, L., Freni, G., La Loggia, G. (2015). "Uncertainty Analysis in the Evaluation of Extreme Rainfall Trends and Its Implications on Urban Drainage System Design." *Water, 7* (12), 6931-6945.

Nourani, V., Komasi, M., Mano, A. (2009). A multivariate ANN wavelet approach for rainfall-runoff modelling. *Water Resour. Manage. 23*, 2877–2894.

Nourani, V., Kisi, O., Komasi, M. (2011). Two hybrid Artificial Intelligence approaches for modeling rainfall–runoff process. *J. Hydrol., 402*, 41–59.

Nourani, V., Komasi, M., Alami, M.T. (2012). Hybrid Wavelet–Genetic Programming Approach to Optimize ANN Modelling of Rainfall–Runoff Process. *J. Hydrol. Eng., 17* (6), 724-741.

Ozturk, T., Altinsoy, H., Türkeş, M., Kurnaz, L. (2012). Simulation of temperature and precipitation climatology for the central Asia CORDEX domain using RegCM 4.0. *Climate Res., 52*, 63–76.

Ozkul, S. (2009). Assessment of climate change effects in Aegean river basins: the case of Gediz and Buyuk Menderes Basins. Clim. Change, *97*, 253–283**.**

Önol, B, Bozkurt, D, Turuncoglu, U. U. et al. (2014). Evaluation of the twenty-first century RCM simulations driven by multiple GCMs over the Eastern Mediterranean–Black Sea region. *Clim. Dynam., 42,* 1949-1965.

Önol, B., Semazzi, F. H. M. (2009). Regionalization of climate change simulations over Eastern Mediterranean. *J. Climate*, *22* (8), 1944–1961.

Özdoğan, M. (2011) Climate change impacts on snow water availability in the Euphrates–Tigris basin. *Hydrol. Earth. Syst. Sci. 15*, 2789–2803.

Pal, J.S., Giorgi, F., Bi, X. et al. (2007). Regional climate modeling for the developing world: the ICTP RegCM3 and RegCNET. *Bull. Amer. Meteor. Soc., 88* (9), 1395–1409.

Parajka, J., Merz, R., Blöschl, G. (2005). A comparison of regionalisation methods for catchment model parameters. *Hydrol. Earth. Syst. Sc. Discuss.*, *9* (3), 157-171.

Perkins, S.bE., Pitman, A. J., Holbrook, N. J., McAneney J. (2007) Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *J.Climate, 20*, 4356–4376.

Piani, C., Weedon, G.P., Best, M et al. (2010). Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *J. Hydrol. 395,* 199–215.

Piani, C., Haerter, J., Coppola, E. (2010). Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor. Appl. Climatol., 99*, 187–192, 2010.

Poli, R., Langdon, W.B., McPhee, N.F. (2008). *A Field guide to genetic programming*. Published via http://lulu.com and freely available at http://www.gp-field-guide.org.uk, (With contributions by J. R. Koza).

Qi, S., Sun, G., Wang, Y., McNulty, S. G., Myers, J. M. (2009). Streamflow response to climate and landuse changes in a coastal watershed in North Carolina. *Trans ASABE*, *52* (3), 739-749.

Rezaeianzadeh, M., Tabari, H., Abghari, H. (2013). Prediction of monthly discharge volume by different artificial neural network algorithms in semi-arid regions. *Arab. J. Geosci., 6*, 2529–2537.

Reichler, T., Kim J. (2008). How well do coupled models simulate today's climate. *Bull. Amer. Meteor. Soc.*, *89*, 303–311.

Riahi, K., Krey, V., Rao, S., Chirkov, V., Fischer, G., … & Rafai, P. (2011) RCP-8.5: exploring the consequence of high emission trajectories. *Clim Change*, *109* (1-2), 33-57.

Roeckner, E., Bäuml, G., Bonaventura L. et al. (2003). The atmospheric general circulation model ECHAM5. Part I: model description. Max Planck Institute for Meteorology Rep. *349*, p 127.

Rogers, L. L. (1992). Optimal groundwater remediation using artificial neural network and the genetic algorithm. PhD dissertation, Stanford University, Stanford, Calif.

Rogers, L. L., and Dowla, F. U. (1994). Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. *Water Resour. Res.*, *30* (2), 457–481.

Roosmalen, L. V., Christensen J. H., Butts M. B. et al. (2010) An intercomparison of regional climate model data for hydrological impact studies in Denmark. *J. Hydrol., 380,* 406-419.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev., 65* (6), 386–408.

Roushangar, K., Mouaze, D., Shiri J. (2014). Evaluation of genetic programming-based models for simulating friction factor in alluvial channels. *J. Hydrol., 517*, 1154–1161.

Sattar, A. M. Gharabaghi, B. (2015). Gene expression models for prediction of longitudinal dispersion coefficient in streams. *J. Hydrol.,* 524, 587-596.

Searson, D. (2009). GPTIPS: Genetic Programming & Symbolic Regression for MATLAB, http://gptips.sourceforge.net.

Searson, D. P., Leahy, D. E., Willis, M. J. (2010). GPTIPS: an open source genetic programming toolbox for multigene symbolic regression. *In Proceedings of the International multi conference of engineers and computer scientists* (Vol. 1, pp. 77-80).

Sellami, H., et al. (2015). Quantifying hydrological responses of small Mediterranean catchments under climate change projections, *Sci. Total Environ,* http://dx.doi.org/10.1016/j.scitotenv.2015.07.006.

Semmler, T., Jacob, D. (2004). Modeling extreme precipitation events–A climate change simulation for Europe. *Global. Planet. Change, 44,* 119–127.

Senatore, A., Mendicino, G., Smiatek, G., Kunstmann, H. (2011). Regional climate change projections and hydrological impact analysis for a Mediterranean basin in Southern Italy. *J. Hydrol., 399*, 70–92.

Seneviratne, S.I., N. Nicholls, D. Easterling, C.M. Goodess, S. Kanae, J. Kossin, Y. Luo, J. Marengo, K. McInnes, M. Rahimi, M. Reichstein, A. Sorteberg, C. Vera, Zhang, X. (2012). Changes in climate extremes and their impacts on the natural physical environment. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp. 109-230.

Simmons, A. J. et al. (1989). The ECMWF medium-range prediction model: Development of the numerical formulations and the impact of increased resolution. *Meteorol. Atmos. Phys., 40*, 28-60.

Singh, V. (1998). *Entropy-based parameter estimation in hydrology* (Vol. 30). Springer Science & Business Media.

Singh, J., Knapp, H.V., Demissie, M. (2004). Hydrologic Modeling of the Iroquois River Watershed using HSPF and SWAT. http://www.isws.illinois.edu/.

Shrestha, R. R., Dibike, Y. B., Prowse, T. D. (2011). Modelling of climate-induced hydrologic changes in Lake Winnipeg Watershed. *J. Great. Lakes. Res., 38* (3), 83–94.

Şen, Z. (2008). *Wady Hydrology*. CRC Press.

Şen, Z. (2012). A statistical downscaling model with applications in northwestern Turkey. *Int. J. Global Warming 4* (1), 31-51.

Şen, Ö. L. (2013). A holistic view of climate change and its impacts in Turkey. Report, Istanbul Policy Centre, Sabanci University, Istanbul, Turkey.

Shen, Y, Oki, T., Kanae, S., et al. (2014) Projection of future world water resources under SRES scenarios: an integrated assessment. *Hydrolog.Sci. J., 59* (10), 1775-1793.

Shiri, J., Kişi ö, Landera G., et al. (2012). Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country (Northern Spain). *J. Hydrol., 414-415*, 302-316.

Shoaib, M., Shamseldin, A. Y., Melville, B. W., Khan, M. M. (2015). Runoff forecasting using hybrid Wavelet Gene Expression Programming (WGEP) approach. *J. Hydrol., 527*, 326-344.

Smith, I., Chandler, E. (2010). Refining rainfall projections for the Murray Darling Basin of south-east Australia-The effect of sampling model results based on performance, *Clim. Change, 102*, 377–393.

Solaiman, T. A. (2011). Uncertainty estimation of extreme precipitations under climate change: A non-parametric approach, Ph.D. dissertation, Univ. of Western Ontario, ON, Canada.

Solaimani, K. (2009). Rainfall-runoff Prediction based on artificial neural network-A case study: Jarahi watershed. American-Eurasian *J. Agric. & Environ. Sci., 5* (6), 856-865.

Solmon, F, Mallet M, Elguindi N, Giorgi F, Zakey A, Konare A (2008) Dust aerosol impact on regional precipitation over western Africa, mechanisms and sensitivity to absorption properties. *Geophys. Res. Lett., 35,* L24705.

Sood, A., Smakhtin, V. (2015). Global hydrological models: a review, *Hydrolog. Sci. J., 60* (4), 549-565,

Stephens, M. A. (1974). EDF Statistics for Goodness of Fit and Some Comparisons, *J. American Statistical Association*, *69*, 730-737.

Steynor, A. C., Hewitson, B. C., Tadross, M. A. (2009). Projected future runoff of the Breede River under climate change. *Water SA*, *35* (4), 433-440.

Sylla, M. B., Coppola, E., Mariotti, L., Giorgi, F., Ruti, P. M., Dell' Aquila, A., Bi, X. (2010) Multiyear simulation of the African climate using a regional climate model (Reg CM3) with the high resolution ERA-interim reanalysis. *Clim. Dynam., 35*, 231–247.

Tahershamsi, A., Majdzade Tabatabai M. R., Shirkhani, R. (2012). An evaluation model of artificial neural network to predict stable width in gravel bed rivers. *Int. J. Environ. Sci. Technol.*, *9*, 333–342.

Tung, C. P., Lee, T. Y., Yang, Y. C. E., Chen, Y. J. (2009). Application of genetic programming to project climate change impacts on the population of

Formosan Landlocked Salmon. *Environ. Model. Soft., 24* (9), 1062-1072.

Teutschbein, C., Seibert, J. (2010). Regional Climate Models for Hydrological Impact Studies at the Catchment Scale: A Review of Recent Modeling Strategies. *Geography Compass, 4* (7), 834–860.

Teutschbein, C., Seibert, J. (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, *J. Hydrol., 456–457*, 12–29.

Tryhorn, L, DeGaetano, A. (2011). A comparison of techniques for downscaling extreme precipitation over the Northeastern United States. *Int. J. Climatol. 31*, 1975–1989.

Uyumaz, A., Danandeh Mehr A., Kahya E., Erdem H. (2014). Rectangular side weirs discharge coefficient estimation in circular channels using linear genetic programming approach. *J. Hydroinform., 16* (6), 1318-1330.

Van Der Linden, P., Mitchell J.F.B. (Eds). (2009). ENSEMBLES: Climate Change and Its Impacts: Summary of Research and Results From the ENSEMBLES Project. Met Office Hadley Centre, Fitz Roy Road, Exeter EX1 3PB, UK, 160.

Van Vliet, M. T. H., Blenkinsop, S., Burton, A .et al. (2012). A multi-model ensemble of downscaled spatial climate change scenarios for the Dommel catchment, Western Europe. *Clim. Change, 111*, 249–277.

Varis, O., Kajander T., Lemmelä, R. (2004) Climate and water: from climate models to water resources management and vice versa. *Clim. Change*, *66*, 321–344.

Vaze, J., Teng J., Chiew, F.H.S. (2011). Assessment of GCM simulations of annual and seasonal rainfall and daily rainfall distribution across south-east Australia. *Hydrol Process., 25* (9), 1486-1497.

Veijalainen, N., Dubrovin, T., Marttunen, M., Vehviläinen, B. (2010). Climate change impacts on water resources and lake regulation in the Vuoksi watershed in Finland. *Water Resour. Manage. 24* (13), 3437-3459.

Waylen, P, Owusu, K. (2014). Changes in expectations and extremes in the rainfall climatology of Accra, Ghana, 1895-2005. *Appl. Geogr., 52*, 99-109.

Willems, P., Olsson, J., Arnbjerg-Nielsen, K., Beecham, S., Pathirana, A., Gregersen, I. B., Madsen, H. (2012). Impacts of climate change on rainfall extremes and urban drainage systems. IWA Publishing, London, UK.

Wu, C. L., Chau, K. W. (2010). Data-driven models for monthly streamflow time series prediction. *Eng. Appl. Artif. Intel., 23* (8), 1350-1367.

Wu, C. L., Chau, K. W. (2011). Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis. *J. Hydrol., 399* (3), 394-409.

Xuejie, G., Zongci, Z., Giorgi, F. (2002). Changes of extreme events in regional climate simulations over East Asia. *Adv. Atmos. Sci.* ***19(5),*** 927-942.

Yang, C. C., Prasher, S. O., Lacroix, R., Sreekanth, S., Patni, N. K., and Masse, L., (1997). Artificial neural network model for subsurface drained farmland. *J. Irrig. Drain. E.-ASCE, 123* (4), 285–292.

Zanis, P., Kapsomenakis, I., Philandras, C., Douvis, K., et al. (2009). Analysis of an ensemble of present-day and future regional climate simulations for Greece. *Int. J. Clim., 29*, 1614–1633.

Zhu, J. (2013). "Impact of Climate Change on Extreme Rainfall across the United States." *J. Hydrol. Eng.,* 10.1061/(ASCE)HE.1943-5584.0000725, 1301-1309.

APPENDICES

APPENDIX A: Kolmogorov-Smirnov (K-S) test

APPENDIX B: Anderson-Darling (A-D) test

Appendix A: Kolmogorov-Smirnov (K-S) test

The choice of theoretical distribution models to be used in extreme value analysis has been a topic of interest for a long time. Different goodness-of-fit tests are available in the relevant literature to be used for choosing an appropriate distribution. These tests measure the compatibility of an experimental (random) sample with a theoretical probability distribution function (PDF). In other words, these tests show how well the distribution selected by a modeler fits to their data.

The K-S test is a test for goodness-of-fit to the shape of a PDF (Massey 1951). It has the same role as the parametric Chi-Square, χ^2, goodness-of-test. But, this is a non-parametric test. Its advantage over the χ^2 test is that the K-S test does not need arbitrarily divide a experimental data into a selection of pre-designated bins. The discriminating ability of the K-S test is dependent on sample size; the larger the sample size, the more reliable the results. When large sample sizes are available, the χ^2 test for goodness-of-fit is more powerful but requires additional manipulation of the data. Where sample sizes are small, the K-S test provides limited information but is a better choice than the χ^2 alternative.

The K-S test is used to test equality of PDF of two sample populations or whether a sample population has the same PDF as some specified or theoretical PDFs. The null and the alternative hypotheses are:

- H_0: the data follow the specified distribution; (PDF1=PDF2; i.e., PDF of population one (or empirical distribution function (ECDF)) is equal to PDF of population two (or theoretical PDF))
- H_A: the data do not follow the specified distribution; (PDF1≠PDF2).

The hypothesis regarding the distributional form is rejected at the chosen significance level (α) if the test statistic, D (Equation (A.1)), is greater than the critical value obtained from Table A.1. The fixed values of α (0.01, 0.05 etc.) are generally used to evaluate the null hypothesis (H_0) at various significance levels. A value of 0.05 is typically used for most applications, however, in some critical cases, a lower α value may be applied.

$$D = \max \left| PDF_1 - PDF_2 \right| \tag{A.1}$$

Given N ordered data points (samples: x1, x2, …, xN), the PDF1 (i.e., ECDF) is defined as

$$PDF_i = \frac{n(i)}{N} \tag{A.2}$$

where n(i) is the number of points less than xi and the xi are ordered from smallest to largest value. This is a step function that increases by 1/N at the value of each ordered data point.

Table A.1 : Critical values for K-S statistic.

Sample size, (N)	Level of Significance, α				
	0.2	0.15	0.1	0.05	0.01
3	0.565	0.597	0.642	0.708	0.828
4	0.494	0.525	0.564	0.624	0.733
5	0.446	0.474	0.474	0.565	0.669
10	0.322	0.342	0.368	0.410	0.490
15	0.266	0.283	0.304	0.338	0.404
20	0.231	0.246	0.264	0.294	0.356
25	0.21	0.22	0.24	0.27	0.32
30	0.19	0.20	0.22	0.24	0.29
35	0.18	0.19	0.21	0.23	0.27
40	0.17	0.18	0.19	0.21	0.25
45	0.16	0.17	0.18	0.20	0.24
50	0.15	0.16	0.17	0.19	0.23
Over 50	$1.07/\sqrt{N}$	$1.14/\sqrt{N}$	$1.22/\sqrt{N}$	$1.36/\sqrt{N}$	$1.63/\sqrt{N}$

For instance, the graph below is a plot of the ECDF and GEV model for bias corrected historical AMR series reproduced by ECHAM5-RegCM3 at the evaluation grid 4. The K-S test statistic based on the maximum distance between these two curves equals to 0.079. From Table A.1, the critical distance (D) for a sample size including 30 samples at significance level 0.05 equals 0.24. Therefore, the GEV distribution is the correct fit for the reprofuced AMR series at the evaluation grid 4.

The Kolmogorov-Smirnov test for goodness-of-fit is distribution free in the sense that the critical values do not depend on the specific distribution being tested and can therefore be used regardless of any distribution that the data are assumed to follow. Despite these advantages, it tends to be more sensitive near the centre of the distribution than at the tails.

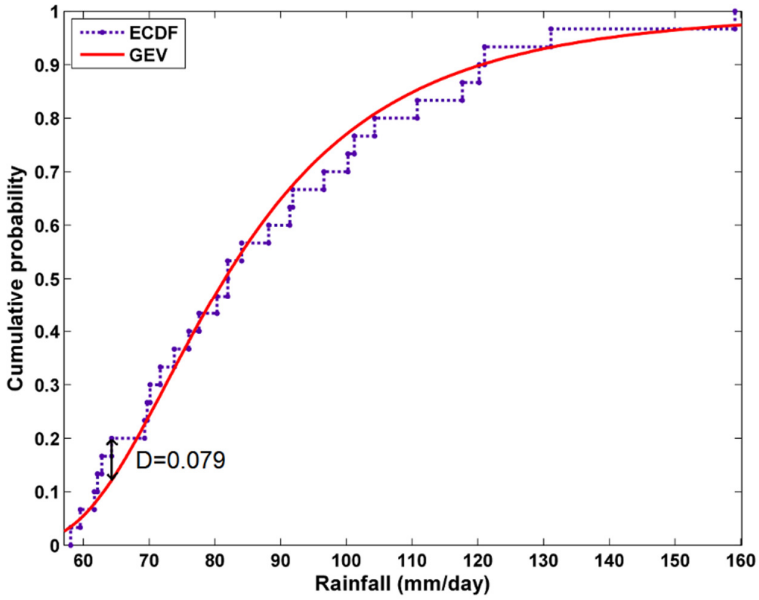

Figure A.1 : Example of goddness-of-fit test for bias corrected AMR series reprodused by ECHAM5-RegCM3 over the historic period 1961–1990 at grid #4.

Appendix B: Anderson-Darling (A-D) test

Like K-S test, the Anderson-Darling test (Stephens, 1974) is used to test if a sample of data came from a population with a specific distribution. It is a modification of the Kolmogorov-Smirnov (K-S) test and gives more weight to the tails than does the K-S test (Ahmed et al. 1988). The null and the alternative hypotheses in A-D test are the same with that of K-S test. However, The Anderson-Darling test statistic is defined as

$$A^2 = -(N + S) \tag{B.1}$$

Where

$$S = \sum_{i=1}^{N} \frac{2i-1}{N} \left[\mathrm{Ln}(\mathrm{PDF}(x_i)) + \mathrm{Ln}(1 - \mathrm{PDF}(x_{N+1-i})) \right] \tag{B.2}$$

Unlike the K-S test, the Anderson-Darling test makes use of the specific distribution in calculating critical values. This has the advantage of allowing a more sensitive test and the disadvantage that critical values must be calculated for each distribution. Currently, tables of critical values are available for different distributions (e.g., Stephens, 1974; Ahmad et al., 1988; Meylan et al. 2012). To test the goodness-of-fit of GEV distribution with three parameters, the critical values of the A^2 statistic can be found in Ahmad et al. (1988) and are reproduced in Table A.2. The test is a one-sided test and the hypothesis that the distribution is of a specific form is rejected if the test statistic, A^2, is greater than the critical value at a given significance level.

As an example, the A-D test statistic based on the GEV function and bias corrected historical AMR series (30 samples) reproduced by ECHAM5-RegCM3 at the evaluation grid 4 equals to 0.2367. From Table B.1, the corresponding critical value at significance level 0.05 equals 0.671, which is greater than A^2. Therefore, the GEV distribution is the correct fit for the simulated AMR series based upon A-D test.

Table B.1 : Values of A-D test statistic for a 3-parameter GEV distribution, taken from Ahmad et al. (1988).

Sample size,	Level of Significance, α			
(N)	0.15	0.1	0.05	0.01
10	0.469	0.525	0.617	0.808
15	0.476	0.536	0.637	0.870
20	0.481	0.543	0.651	0.925
25	0.484	0.549	0.662	0.978
30	0.487	0.553	0.671	1.04
35	0.489	0.556	0.678	1.12
40	0.491	0.559	0.683	1.22
50	0.493	0.563	0.693	1.20
100	0.500	0.575	0.721	1.15

CURRICULUM VITAE

Name Surname **:** Ali DANANDEH MEHR

Place and Date of Birth **:** Tabriz, Iran, 1979

E-Mail **:** danandeh@itu.edu.tr danandehmehr@yahoo.com

EDUCATION:

- **B.Sc.** **:** 2001, Islamic Azad University–Tabriz Branch, Faculty of Engineering, Civil Engineering Department.
- **M.Sc.** **:** 2005, Power and Water University of Technology, Faculty of water Engineering.

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2014 – 2016 Research Assistant, Istanbul Technical University, TRNC Education and Research Campuses. Famagusta, Turkis Republic of North Cyprus.
- 2009 – 2012 Senior Expert of the Bureau of Rivers & Coastal Engineering, Iran Water Resources Management Company. Tehran, Iran.
- 2007 – 2009 River Engineering Expert, Sangab Zagros Consulting Eng. Co., Shiraz, Iran.
- 2004 – 2007 River Engineering Expert, Ab-Pouyeshgaran-Novin Consulting Eng. Co., Tehran, Iran.
- 2002 – 2004 Hydrology Expert, Tarh-E-Now-Andishan Consulting Eng. Co., Tehran, Iran.

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Danandeh Mehr, A.,** and Kahya, E. (2016). A linear genetic programming approach for projected Rainfall-Runoff Modelling with limited data availability. (Submitted).
- **Danandeh Mehr, A.,** and Kahya, E. (2016). Climate change impacts on catchment-scale extreme rainfall variability. (Submitted).
- **Danandeh Mehr, A.,** and Kahya, E. (2016). Grid-based performance evaluation of GCM-RCM combinations for rainfall reproduction. *Theoretical and Applied Climatology* (in press).
- Kahya, E., **Danandeh Mehr, A.,** Şen, O., and Özger M. (2015). Climate change impacts on extreme rainfalls over Rize Province, Turkey. *10th International Congress on Civil Engineering*, 5-7 May 2015 University of Tabriz, Tabriz, Iran.
- Kahya, E., **Danandeh Mehr, A.,** Bozkurt, D., Özger M., and Şeker, D. Z. (2014). Evaluation of the performance of GCMs driven RCM outputs for rainfall prediction over Rize Catchment. *11th International Congress on Advances in Civil Engineering* (ACE 2014), Istanbul, Turkey.

OTHER PUBLICATIONS, PRESENTATIONS AND PATENTS

(Accomplished during Ph.D. career):

- **Danandeh Mehr, A.,** Kahya E., Şahin, A. and Nazemosadat M.J. (2015). Successive-station monthly streamflow prediction using different ANN algorithms. *International Journal of Environmental Science and Technolog*y, 12 (7): 2191-2200.
- **Danandeh Mehr, A.,** Kahya, E. and Özger M. (2014). A gene-wavelet model for long lead-time drought forecasting. *Journal of Hydrology* 517, 691-699.
- **Danandeh Mehr, A.,** Kahya, E. and Yerdelen, C. (2014). Linear genetic programming application for successive-station monthly streamflow prediction. *Computers & Geosciences* 70, 63-72.
- Uyumaz, A., **Danandeh Mehr, A.,** Kahya E. and Erdem H. (2014) Rectangular side weirs discharge coefficient estimation in circular channels using linear genetic programming approach. *Journal of Hydroinformatics* 16(6), 1318-1330.
- **Danandeh Mehr, A.,** Kahya, E., Bagheri, F. and Deliktas E. (2014) Successive-station monthly streamflow prediction using neuro-wavelet technique. *Earth Science Informatics* 7, 217-229.
- **Danandeh Mehr, A.,** Kahya E. and Olyaie E. (2013) Streamflow prediction using linear genetic programming in comparison with a neuro-wavelet technique. *Journal of Hydrology*, 505:240–249.
- Özger, M., **Danandeh Mehr, A.,** and Kahya, E. (2013). A fuzzy logic approach for flood risk management: An application – Mehranrood, Iran. Proceedings of *Flood Symposium*, pp., 325-336. Trabzon, Turkey (in Turkish).
- Uyumaz, A., **Danandeh Mehr**, **A.,** and Şen, O. (2013). River morphology and

fluvial processes. Proceedings of Bursa *3rd International Congress on Water*. pp., 1094-1099. Bursa, Turkey.

- **Danandeh Mehr, A.**, and Imanshoar, F. (2013). RESCDAM Method Application for Estimating Loss of Life Caused by Dam Break (Case study: Shahid Madani Dam, Iran). Proceedings of *fifth Iran water resources management Conference*, Power and Water University of Technology, Tehran, Iran (in Farsi).