IMPROVING DATA-DRIVEN BASED STREAMFLOW FORECASTING USING WAVELET TRANSFORMATION

Sinan Jasim Hadi ALDOORI

PhD DISSERTATION

Department of Remote Sensing and Geographic Information System

Supervisor: Prof. Dr. Mustafa TOMBUL

Eskişehir

Anadolu University

Graduate School of Sciences

March 2018

This dissertation study was funded by the board of the Scientific Projects' Research Commission of Anadolu University Under Project NO. BAP- 1604F165.

FINAL APPROVAL FOR THESIS

This thesis titled "Improving Data-Driven Based Streamflow Forecasting Using Wavelet Transformation (Dalgacık Dönüşümü Kullanılarak Veriye Dayalı Akım Tahmin Modellerinin İyileştirilmesi)" has been prepared and submitted by Sinan Jasim Hadi ALDOORI in partial fulfillment of the requirements in "Anadolu University Directive on Graduate Education and Examination" for the Degree of Doctor of Philosophy (PhD) in Remote Sensing and Geographic Information System Department has been examined and approved on 20/03/2018.

Committee Members		**Signature**
Member (Supervisor)	: Prof. Dr. Mustafa TOMBUL	
Member	: Prof. Dr. Yasin Dursun SARI	
Member	: Doç. Dr. Tarkan ERDİK	
Member	: Doç. Dr. Mehmet ÖZGER	
Member	: Dr. Öğr. Üyesi Emrah PEKKAN	

Prof.Dr. Ersin YÜCEL
Director of Graduate School of Sciences

ABSTRACT

IMPROVING DATA-DRIVEN BASED STREAMFLOW FORECASTING USING WAVELET TRANSFORMATION

Sinan Jasim Hadi ALDOORI

Department of Remote Sensing and Geographic Information System

Anadolu University, Graduate School of Sciences, March 2018

Supervisor: Prof. Dr. Mustafa TOMBUL

Daily streamflow forecasting is conducted in this study using several data-driven models (DDMs): artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS) and support vector machine (SVM). Seven days ahead and one month ahead streamflow being less predictable wavelet transformation (WT) is used as preprocessing to improve the performance of the models. Two new hybrid models were proposed using multi-gene genetic programming (MGGP) and extreme gradient boosting (XGB) as a selection tool to select only the important scales obtained from the continuous wavelet transformation (CWT) to be them imposed into ANN and extreme learning machine (ELM). Hindcast and real forecast experiment are conducted to investigate the performance of the WT-based models in the real forecasting. The results show that daily forecast can be implemented successfully using DDMs and ANN has the highest performance in comparison to ANFIS and SVM. The proposed two models outperformed the models uses discrete wavelet transformation (DWT) as more information can be included in the model. Finally, the performance of the models using WT-based hybrid models for both CWT and DWT in hindcast experiment was found as increases due to the incorrect application of WT. Mostly, WT is applied onto the time series, divided into calibration and testing subsets to be then imposed into DDMs and that sends some future information into the model. In the real forecast experiment, the WT-based hybrid models have less performance than the stand-alone DDMs in which no preprocessing applied.

Keywords: Streamflow, Wavelet transformation, Artificial neural network, Extreme learning machine, extreme gradient boosting.

ÖZET

DALGACIK DÖNÜŞÜMÜ KULLANILARAK VERİYE DAYALI AKIM TAHMİN MODELLERİNİN İYİLEŞTİRİLMESİ

Sinan Jasim Hadi ALDOORI

Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Anabilim Dalı

Anadolu Üniversitesi, Fen Bilimler Enstitüsü, Mart 2018

Danışman: Prof. Dr. Mustafa TOMBUL

Bu çalışmada, yapay sinir ağı (Artificial Neural Network (ANN)), uyarlamalı nöro-bulanık çıkarım sistemi (Adaptive Neuro-Fuzzy Inference System (ANFIS)) ve destek vektör makinesi (Support Vector Machine (SVM)) gibi veriye dayalı modeller (Data-Driven Models (DDMs)) kullanılarak günlük akım tahmini yapılmıştır. Yedi gün ve bir ay ilerideki akım tahmini daha az öngörülebilir hale gelirken, modellerin performansını artırmak için ön işlem olarak dalgacık dönüşümü (Wavelet Transformation (WT)) kullanılmıştır. Sürekli dalgacık dönüşümü (Continuous Wavelet Transformation (CWT)) 'dan elde edilen önemli ölçekleri seçmek için, bir seçim aracı olan çok genli genetik programlama (Multi-Gene Genetic Programming (MGGP)) ve aşırı gradient yükseltmeyi (XGB) kullanarak iki yeni hibrid model önerilmiştir. Bu modeller yapay sinir ağları (ANN) ve Aşırı Öğrenme Makinesi (Extreme Learning Machine (ELM)) olarak uygulamaya konmuştur. Gerçek tahminlerde WT tabanlı modellerin performansını incelemek için hindcast ve gerçek tahmin denemesi yapılmıştır. Sonuçlar, günlük tahmini DDM'leri kullanarak başarılı bir şekilde uygulandığını ve ANN'nin ANFIS ve SVM'ye kıyasla en yüksek performansa sahip olduğunu göstermektedir. Modele daha fazla bilgi dahil edilebildiğinden, önerilen iki model daha iyi performans göstermiştir; o nedenle bu çalışmada ayrık dalgacık dönüşümü (Discrete Wavelet Transformation (DWT)) kullanılmıştır. Son olarak, WT'nin yanlış uygulanmasından dolayı, WT- tabanlı hibrit modellerin art arda yapılan denemelerinde, hem CWT hem de DWT için performansının arttığı bulunmuştur. Çoğunlukla, WT, zaman serilerine uygulanmakta, daha sonra DDM'lere uygulanacak kalibrasyon ve test alt kümelerine bölünmekte ve geleceğe dair bazı bilgileri modele göndermektedir. Gerçek tahmin denemesinde, WT tabanlı hibrit modellerde, ön işlem uygulanmayan bağımsız DDM'lerden daha az performans göstermiştir.

Anahtar Sözcükler: Akim, Dalgacık dönüşümü, Yapay sinir ağı, Aşırı öğrenme makinesi, Aşırı gradient yükseltmeyi.

ACKNOWLEDGMENTS

All praise to ALLAH, Most Gracious, Most Merciful, Who, Alone brings forgiveness and light and new life to those who call upon Him; and to Him is the dedication of this work.

"Read! In the Name of your Lord who has created (all that exist).
He has Created man from a clot.
Read! And your Lord as the Most Generous.
Who has taught (the writing) by the pen.
He has taught man that which he knew no."
Qur'an (Alaq) 96: 1-5

I praise ALLAH for his great loving kindness, which has brought all of us to tell and encourage each other and who has pulled us from the darkness to the light. All respect for our holy prophet (Peace be upon him), who guided us to identify our creator.

I would like to express my heartiest gratitude to my supervisor: Prof. Dr. Mustafa TOMBUL for his strong support, untiring guidance, endless encouragement, and invaluable assistance throughout the work in this project. I would like to thank Assoc. Prof. Dr. Tarkan ERDIK, Assoc. Prof. Dr. Mehmet ÖZGER, and Assoc. Prof. Dr. Uğur AVDAN for their guidance and their valuable comments and recommendations during the whole journey.

I cannot find enough words for thanking My Parents. I ask ALLAH to keep them over my head. My wife, she is the one who has to be thanked by me every day for being the biggest supporter for me in all the hard and easy times. I would like to thank my brothers and sisters also for their support, encouragement, and constant prayers.

Sinan J. H. Aldoori

27/03/2018

STATEMENT OF COMPLIANCE WITH ETHICAL PRINCIPLES AND RULES

I hereby truthfully declare that this thesis is an original work prepared by me; that I have behaved in accordance with the scientific ethical principles and rules throughout the stages of preparation, data collection, analysis and presentation of my work; that I have cited the sources of all the data and information that could be obtained within the scope of this study, and included these sources in the references section; and that this study has been scanned for plagiarism with "scientific plagiarism detection program" used by Anadolu University, and that "it does not have any plagiarism" whatsoever. I also declare that, if a case contrary to my declaration is detected in my work at any time, I hereby express my consent to all the ethical and legal consequences that are involved.

Sinan Jasim Hadi ALDOORI

TABLE OF CONTENTS

TITLE PAGE ..... **i**

FINAL APPROVAL FOR THESIS ..... **ii**

ABSTRACT ..... **iii**

ÖZET ..... **iv**

ACKNOWLEDGMENTS ..... **v**

STATEMENT OF COMPLIANCE WITH ETHICAL PRINCIPLES AND RULES ..... **vi**

TABLE OF CONTENTS ..... **vii**

TABLE OF TABLES ..... **xiii**

TABLE OF FIGURES ..... **xv**

SYMBOLS AND ABBREVIATIONS ..... **xviii**

INTRODUCTION **1**

CHAPTER ONE

1. DATA PREPARATION **5**

1.1. Comparison of Spatial Interpolation Methods of Precipitation and Temperature Using Multiple Integration Periods **5**

1.1.1. Abstract **5**

1.1.2. Introduction **5**

1.1.3. Study area and data collection **7**

1.1.4. Methods **8**

1.1.4.1. Inverse Distance Weighting (IDW) **10**

1.1.4.2. Thiessen Polygons/Nearest Neighbor (TP) **10**

1.1.4.3. Thin Plate Spline (TPS) **11**

1.1.4.4. Trend Surface Analysis/ Global Polynomial Interpolation (TSA) **12**

1.1.4.5. Local Polynomial Interpolation (LPI) **12**

1.1.4.6. Ordinary Kriging (OK) 13
1.1.4.7. Universal Kriging (UK) 14
1.1.4.8. Simple Kriging (SK) 15
1.1.5. Validation 15
1.1.6. Results and Discussion 17
1.1.7. Conclusion 22
1.2. Conversion of CRUTS 3.23 Data and Evaluation of Precipitation and Temperature Variables in A Local Scale 23
1.2.1. Abstract 23
1.2.2. Introduction 23
1.2.3. Data collection 25
1.2.4. Methodology 26
1.2.5. Results and Discussion 28
1.2.6. Conclusion 33

CHAPTER TWO

2. DATA EXPLORATION 34
2.1. Long-Term Spatiotemporal Trend Analysis of Precipitation and Temperature Over Turkey 34
2.1.1. Abstract 34
2.1.2. Introduction 34
2.1.3. Study area 37
2.1.4. Data Used 37
2.1.5. Methodology 39
2.1.5.1. Mann-Kendall Test 39
2.1.5.2. Theil-Sen's Slope Estimator 40
2.1.5.3. Pettitt-Mann-Whitney Test 41
2.1.6. Results and Discussion 41

2.1.6.1. Exploratory Data Analysis (EDA) 41
2.1.6.2. Annual trend 43
2.1.6.2.1. Precipitation 43
2.1.6.2.2. Temperature 47
2.1.6.3. Seasonal Trend 48
2.1.6.3.1. Precipitation 48
2.1.6.3.2. Temperature 51
2.1.7. Conclusion 53
2.2. Homogeneity and Trend Analysis of Rainfall and Streamflow of Seyhan Basin, Turkey 55
2.2.1. Abstract 55
2.2.2. Introduction 56
2.2.3. Study area and data collection 57
2.2.4. Methodology 58
2.2.4.1. Standard Normal Homogeneity Test (SNHT) 58
2.2.4.2. Buishand test 58
2.2.5. Results and Discussion 59
2.2.5.1. Exploratory Data Analysis (EDA) 59
2.2.5.2. Homogeneity 59
2.2.5.3. Annual trend 60
2.2.5.4. Monthly trend 63
2.2.6. Conclusion 63

CHAPTER THREE

3. DAILY STREAMFLOW MODELING USING DATA-DRIVEN MODELS 67
3.1. Forecasting of Daily Streamflow For Basins With Different Physical Characteristics Using Data-Driven Methods 67
3.1.1. Abstract 67

3.1.2. Introduction 67
3.1.3. Background of the Data-Driven Methods 69
3.1.3.1. AR model 69
3.1.3.2. Artificial Neural Network 69
3.1.3.3. Adaptive-Neuro Fuzzy Inference System (ANFIS) 70
3.1.3.4. Support Vector Machine (SVM): 72
3.1.4. Case Study 76
3.1.4.1. Study area and Data 76
3.1.4.2. Model Development and Structure 77
3.1.5. Evaluation Criteria 79
3.1.6. Results and Discussion 80
3.1.7. Conclusion 85

CHAPTER FOUR

4. IMPROVING STREAMFLOW FORECASTING USING WAVELET 87
4.1. Streamflow Forecasting Using Four Wavelet Transformation Combinations Approaches with Data-Driven Models: A Comparative Study 87
4.1.1. Abstract 87
4.1.2. Introduction 87
4.1.3. Methodology 90
4.1.3.1. Wavelet Transformation (WT) 90
4.1.3.2. ANFIS – Sub Clustering (ANFIS-SC) 92
4.1.4. Evaluation of performance 94
4.1.5. Study area and Data 94
4.1.6. Application 96
4.1.7. Results and Discussion 101
4.1.8. Conclusion 107

4.2. Improving Monthly Streamflow Forecasting Using Continuous Wavelet, and Multi-Gene Genetic Programming Combination 109

4.2.1. Abstract 109

4.2.2. Introduction 110

4.2.3. Wavelet Transformation 113

4.2.4. Wavelet Coherence Transformation (WCT) 114

4.2.5. Genetic Programming 116

4.2.6. Proposed hybrid model (WMGGPNN) 119

4.2.7. Application 119

4.2.7.1. Study area and Data 119

4.2.7.2. Methodology 122

4.2.8. Results and Discussion 126

4.2.9. Summary and Conclusion 134

CHAPTER FIVE

5. IS THE USE OF WAVELET AS PREPROCESSING IMPROVES FORECASTING? 137

5.1. Is Discrete and Continuous Wavelet Transformation Improving or Deteriorating The Performance of Data-Driven Models in Forecasting Monthly Streamflow 137

5.1.1. Abstract 137

5.1.2. Introduction 137

5.1.3. Preprocessing technique 140

5.1.3.1. Wavelet transformation (WT) 140

5.1.4. Data-Driven models 141

5.1.4.1. Extreme learning machine (ELM) 141

5.1.5. Extreme Gradient Boosting (XGB) 143

5.1.6. Proposed modeling Schema: 145

5.1.7. Experiment 147

5.1.7.1. Study area and data 147

5.1.7.2. Model development 148

5.1.7.3. Stand-alone experiment 150

5.1.7.4. Hindcast experiment 150

5.1.7.5. Forecast experiment 152

5.1.8. Evaluation Criteria 154

5.1.9. Results and Discussion 155

5.1.10. Conclusion 163

CONCLUSION 164

REFERENCES 168

CURRICULUM VITAE 190

TABLE OF TABLES

Table 1.1. *Statistics of the observed daily total precipitation (in mm) of 17 stations for the period of 1987 to 1994* 9

Table 1.2. *Statistics of the observed daily average temperature (C°) of 17 stations for the period 1987 to 1994* 9

Table 1.3. *Validation measurements: RMSE, MSRE, r2, and CE of daily, monthly, and annual total precipitation (in mm) interpolations* 19

Table 1.4. *Validation measurements: RMSE, MSRE, r2, and CE of monthly, and annual average precipitation (in mm) interpolations* 20

Table 1.5. *Validation measurements: RMSE, MSRE, r2, and CE of daily, monthly, and annual average temperature (in C°) interpolations* 21

Table 1.6. *The correlation and MSRE between CruTS and observed data for Precipitation and Temperature* 32

Table 1.7. *The P value of the mean difference, median difference and Mann-Whitney tests* 32

Table 2.1. *The results of the Pettit test and trend value (mm/year) with significance for Precipitation. Ann: Annual, DJF: December, January, and February, MAM: March, April, and May, JJA: June, July, August, SON: September, October, and November* 45

Table 2.2. *The results of the Pettit test and trend value (C°/Century) with significance for Temperature. Ann: Annual, DJF: December, January, and February, MAM: March, April, and May, JJA: June, July, August, SON: September, October, and November* 49

Table 2.3. *Statistics of the monthly total precipitation and monthly mean streamflow* 58

Table 2.4. *P values of Homogeneity tests, SNHT: Standard Normal Homogeneity Test, Buishand: Buishand Test* 60

Table 2.5. *The results of the Pettit test (change year) and trend value (mm/year) (Sen's Estimator) with significance (Mann-Kendall test) for rainfall. The underlined significant at 0.05 Bold significant at 0.1* 62

Table 2.6. *The results of the trend value (mm/year) with significance for rainfall* 66

Table 3.1. *Statistics of the daily streamflow and physical characteristics of the basins* 77

Table 3.2. *The parameters of ANN and SVM for the models* 79

Table 3.3. *The results of the 1801 basin for the different models* 81

Table 3.4. *The Peak flood values forecasting results for the three basins* 81

Table 3.5. *The results of the 1805 basin for the different models* 83

Table 3.6. *The results of the 1822 basin for the different models* 84

Table 4.1. *Characteristics of the daily temperature, rainfall, and streamflow* 96

Table 4.2. *Different combinations of models* 97

Table 4.3. *Number of inputs to the DDMs for every model and scheme* 99

Table 4.4. *The optimum number of neurons in the hidden layer of the ANN network* 100

Table 4.5. *The optimum squash factor values of the ANFIS-SC* 101

Table 4.6. *The optimum values of the SVR parameters* 101

Table 4.7. *The relative performance of DCWT with DWT in forecasting* 107
Table 4.8. *The relative performance of DCWT with DWT in forecasting peak values* 107
Table 4.9. *Statistical characteristics of the meteorological and hydrological variables* 122
Table 4.10. *Developed models' combinations* 125
Table 4.11. *The structure of the ANN approach. The structure is represented as I-H-O, I: number of input neurons, H: number of neurons in the hidden layer, O: number of output neurons* 126
Table 4.12. *Multi-Gene Genetic Programming (MGGP) parameters* 126
Table 4.13. *The performance of the ANN and MGGP approaches in the Best-Scale layout using Haar mother wavelet function* 130
Table 4.14. *The performance of the proposed WMGGPNN model* 131
Table 4.15. *The performance results of the proposed model WMGGPNN in forecasting the peak values* 133
Table 5.1. *The developed combinations* 150
Table 5.2. *The CE and RMSE of the stand-alone schema of ELM method* 156
Table 5.3. *The CE and RMSE of the stand-alone schema of XGB method* 156
Table 5.4. *The CE and RMSE of the hindcast experiment* 157
Table 5.5. *The CE and RMSE of the forecast experiment* 158
Table 5.6. *The importance matrix of the CWT scales obtained by XGB. The value in the brackets represents the scale* 162

TABLE OF FIGURES

Figure 1.1. *The distribution of the meteorological stations and the topographic map of the study area which is located in Seyhan Basin, Turkey and covers parts of Kayseri, Adana, K. Maras, and Sivas provinces* 8

Figure 1.2. *Examples of interpolated maps of monthly total precipitation (in mm) for each of the 8 interpolation methods. The month chosen as an example is June 1989* 18

Figure 1.3. *Examples of interpolated maps of monthly average temperature (in C^o) for each of the 8 interpolation methods. The month chosen as an example is June 1989* 19

Figure 1.4. *Study area with the 9 stations* 26

Figure 1.5. *Global Precipitation map of January 1973 obtained from CruTS 3.23 data* 28

Figure 1.6. *CruTS and observed Precipitation data boxplot* 29

Figure 1.7. *CruTS and observed Temperature data boxplot*.....29

Figure 1.8. *Time series plot of the monthly total precipitation obtained from CruTS and Observed data for the period 1973 - 2000*.....30

Figure 1.9. *Time series plot of the monthly average temperature obtained from CruTS and Observed data for the period 1973 - 2000*.....31

Figure 2.1. *Regions of Turkey and the distribution of the Grids. Regions; BS: Black Sea, EA: Eastern Anatolia, SEA: South Eastern Anatolia, CA: Central Anatolia, MD: Mediterranean, AE: Aegean, MA: Marmara* 38

Figure 2.2. *a) The spatial distribution of the mean of annual Precipitation. b) The spatial distribution of the mean of annual Temperature* 43

Figure 2.3. *The time series of Turkey's mean of the total annual precipitation (Top) and the average annual temperature (bottom). Solid Line: LOWESS regression, Dashed Line: Simple Linear Regression. For Precipitation(Temperature), Insignificant (Significant) trend with 0.11 mm/year (0.88 C^o/Century) Theil and Sen's Slope test value and change year 1934 (1993)*.....44

Figure 2.4. *Annual Theil and Sen's Slope test values Top) Precipitation (mm/year), Bottom) Temperature (Co/year). Circle: Not significant, Square: significant at 0.1, Triangle: Significant at 0.05*.....46

Figure 2.5. *Contour lines for the spatial distribution of the Annual Theil and Sen's Slope test values Top) Precipitation (mm/year), Bottom) Temperature (Co/decade)*.....47

Figure 2.6. *Seasonal Theil and Sen's Slope test values Top) Precipitation (mm/year), Bottom) Temperature (Co/year). Circle: Not significant, Square: significant at 0.1, Triangle: Significant at 0.05. Light to dark grey shades represents: Winter, Spring, Summer, and Autumn respectively*.....50

Figure 2.7. *Contour lines for the spatial distribution of the Theil and Sen's Slope test values of seasonal precipitation trend (mm/year) from top Winter, Spring, Summer, and Autumn. Circle: Not significant, Square: significant at 0.1, Triangle: Significant at 0.05* 52

Figure 2.8. *Contour lines for the spatial distribution of the Theil and Sen's Slope test values of seasonal temperature trend (Co/decade) from top Winter, Spring, Summer, and Autumn. Circle: Not significant, Square: significant at 0.1, Triangle: Significant at 0.05* 54
Figure 2.9. *Study area* 57
Figure 2.10. *The time series of the annual rainfall and annual streamflow. Line: Linear Regression line, Dashed Line: LOWESS regression Line* 61
Figure 2.11. *Spatial distribution of the stations with the annual trend. The size of the triangle represents the normalized magnitude of the trend. The up-pointing triangle: positive trend, the down pointing triangle: negative trend. Black: Not significant, Red: Significant at α= 0.1* 62
Figure 2.12. *Same as Figure 2.11 but for monthly trend and the Blue: Significant trend at α= 0.05*.....64
Figure 3.1. *FFNN adapted from (Nourani 2017)*.....70
Figure 3.2. *ANFIS framework* 72
Figure 3.3. *The linear SVM (Smola et al. 2004)* 73
Figure 3.4. *Study area and location of gage stations*.....77
Figure 3.5. *Time series of the Streamflow* 78
Figure 3.6. *The Autocorrelation and partial autocorrelation functions of the three basins gage stations. from top to bottom: 1801, 1805 an 1822* 80
Figure 3.7. *Daily streamflow observed vs estimated of the 1801 station. Red lines represent the best fit line*.....82
Figure 3.8. *Daily streamflow observed vs estimated of the 1805 station. Red lines represent the best fit line*.....83
Figure 3.9. *Daily streamflow observed vs estimated of the 1822 station. Red lines represent the best fit line*.....84
Figure 4.1. *The discrete wavelet decomposition framework*.....92
Figure 4.2. *location of the case study*.....95
Figure 4.3. *Time series of the temperature, rainfall, and streamflow* 95
Figure 4.4. *Schematic layouts of the applied models for forecasting streamflow* 98
Figure 4.5. *the decomposition of the daily streamflow using "dmey" function with 5 levels* 99
Figure 4.6. *The R2 and RMSE values of the three schematic layouts (ANN, CWTANN, and DWTANN)* 103
Figure 4.7. *The R2 and RMSE values of the three schematic layouts (ANFIS, CWTANFIS, and DWTANFIS)* 104
Figure 4.8. *The R2 and RMSE values of the three schematic layouts (SVR, CWTSVM, and DWTSVM)* 105
Figure 4.9. *Comparison of R2 and RMSE values between the schematic layouts of all DDMs for models M8, M9, and M10*.....105
Figure 4.10. *The time series of the observed predicted DWT, and predicted DCWT*.....108
Figure 4.11. *GP tree structure of the expression* $3.5X1 \div CosX2 + exp\,(CosX1)$.....117

Figure 4.12. *An example of the crossover and mutation operations* 117
Figure 4.13. *Multi Gene Genetic Programming (Symbolic Regression)* 118
Figure 4.14. *Schematic layout of the proposed model* 120
Figure 4.15. *Study area* 120
Figure 4.16. *Time series of the monthly variables (Right figures) and the ACF of the downstream and the CCF of the downstream and each variable (Left figures)* 122
Figure 4.17. *a) The CWT scalogram of the downstream, b, c, d, and e) The wavelet coherence scalogram of the downstream with the Upstream, Rainfall, Temperature, and Potential evapotranspiration. The surrounded areas with solid black lines are significant regions in 5% significance level. The arrows represent the phase difference between the two time series. The direction of the Arrows has the meaning; right pointing: in-phase, left pointing: Anti-phase, down pointing: x leads y by 90^0 (or y lags x by 270^0), up pointing: y leading x by 90^0 (or x lags y 270^0)* 123
Figure 4.18. *The performance of the stand-alone layout models. Black colour: Training subset, Grey Colour: Validation subset* 128
Figure 4.19. *The performance of the Best-Scale layout model of the ANN. Top figures (R2), Bottom figures (RMSE), Black colour: Training subset, Grey Colour: Validation subset* 129
Figure 4.20. *Percentage of change in R2 (left figures), and RMSE (right figures) relatively to the proposed model WMGGPNN. Black colour: Training subset, Grey Colour: Validation subset* 131
Figure 4.21. *The observed vs predicted stream flows of the best-scale, full-scale and the proposed WMGGPNN model. Redline: the best fit* 132
Figure 4.22. *The performance of the models in forecasting the peak values. Left figures: RP2, Right figures: RMSEP* 134
Figure 4.23. *The observed and the predicted peak flows of the best-scale, full-scale and the proposed WMGGPNN model* 135
Figure 5.1. *Framework of the proposed model schema* 147
Figure 5.2. *Study Area* 147
Figure 5.3. *The time series of the variables used in the development of the models* 149
Figure 5.4. *The framework of the hindcast experiment* 152
Figure 5.5. *The framework of the forecast experiment* 154
Figure 5.6. *The forecasted versus the observed downstream flow for the best model in each experiment. SA: Stand-alone, HC: Hindcasting, PM: Proposed Model, C05: combination number 5, FC: Forecasting* 159
Figure 5.7. *The reconstructed sub time series of the downstream flow using DWT with level 2* 160
Figure 5.8. *The difference between the first 200 steps of sub time series of the downstream flow using DWT with level 2 for 200 and 256 months* 161
Figure 5.9. *The coefficients of 5th scale of the downstream flow using CWT* 162
Figure 5.10. *The difference between the first 200 months of coefficients of 5th scale of the downstream flow using CWT for 200 and 256 months* 162

SYMBOLS AND ABBREVIATIONS

α	:	Significance level
ACF	:	Autocorrelation function
ANFIS	:	Adaptive neuro-fuzzy inference system
ANFIS-SC	:	ANFIS – Sub Clustering
ANN	:	Artificial neural network
AR	:	Auto regression
BADC	:	British Atmospheric Data Center
C	:	Control Coefficient
CAM	:	Climate anomaly method
CE	:	Coefficient of efficiency
cld	:	Cloud cover
Comp	:	Combination
CRUTS	:	Climate research unite time series
CWT	:	Continuous wavelet transformation
DCWT	:	Discrete continuous wavelet transformation
DDMs	:	Data-driven models
DMI	:	Devlet Meteoroloji İşleri
DS	:	down streamflow
DSI	:	Develt Su Isleri
DTRs	:	Diurnal temperature ranges
DWT	:	Discrete wavelet transformation
ELM	:	Extreme learning machine
EMD	:	Empirical mode decomposition
ε	:	Epsilon
FFNN	:	Feed-forward Back propagation network
frs	:	frost day frequency
GEP	:	Genetic Expression Programming
GIS	:	Geographic information system
GP	:	Genetic Programing
GPCC	:	Global Precipitation Climatology Center
IA	:	Index of agreement
IDW	:	Inverse distance weighting
In	:	Number of inputs
KKT	:	Karush-Kuhn-Tucker
LOWESS	:	LOcally WEighted Scatterplot Smoothing

LPI	:	Local polynomial interpolation
MCDW	:	Monthly Climatic Data for the World
MEP	:	Multi-expression Programming
MGGP	:	Multi-gene genetic programing
MK	:	Mann-Kendall test
MLPANN	:	Multilayer Perceptron Artificial Neural Network
MLR	:	Multivariate Linear Regression
MSRE	:	Mean squared relative error
NCDC	:	UK National Climatic Data Center
Ne	:	Number of hidden neurons
NMSs	:	National Meteorological Services
NOAA	:	National Oceanographic and Atmospheric Administration
OK	:	Ordinary kriging
Out	:	Number of outputs
PACF	:	Partial autocorrelation function
PET	:	Potential Evapotranspiration
pre	:	Precipitation
R	:	Rainfall
R^2	:	Coefficient of determination
RBFNN	:	Radial Basis Function Neural Network
RMSE	:	Root mean squared error
RS	:	Remote sensing
σ	:	Gaussian noise level
SK	:	Simple kriging
SNHT	:	Alexanderson's Standard Normal Homogeneity Test
SSA	:	singular spectrum analysis
SVM	:	Support vector machine
T	:	Temperature
tmn	:	Minimum temperatures
tmp	:	Mean temperature
tmx	:	Maximum temperatures
TP	:	Thiessen polygons
trimf	:	Triangle member function
TSA	:	Trend surface analysis
UK	:	Universal Kriging
US	:	Upper streamflow
vap	:	Vapor pressure
WCT	:	Wavelet Coherence Transformation

wet	:	Wet-day frequency
WMO	:	World Meteorological Organization
WT	:	Wavelet transformation
WVC	:	Second-order volterra
WWR	:	World Weather Records
XGB	:	Extreme gradient boosting
XWT	:	Cross-wavelet transformation

INTRODUCTION

The management of the water resources system including planning, designing, and operating needs short-term and long-term future events forecasting (Kişi 2008, Sanikhani et al. 2012). One of the most important part of this system is streamflow. Streamflow is an essential component of hydrologic cycle in the regional and global scale (Makkeasorn et al. 2008, Liu et al. 2014) and it is the main source of the fresh water supply and highly associated with the natural disasters such as droughts and floods. Modelling of streamflow is very important for the efficient water resources management due to its role in the mitigation of influences of the excess/scarcity of the water systems (Wood et al. 1980, Sanikhani et al. 2012). Also, accurate forecasting streamflow is essential for efficient water resources management particularly in the regions vulnerable to floods and droughts (Singh 2016). Specifically, for streamflow, Short-term forecasts (e.g. hourly or daily) are important for two main applications: forecasting of the floods and development of warning system. The long-term forecasts (e.g. Monthly or annual) are useful for several applications such as irrigation management decisions, reservoir operations, hydro-power generations, and sediments transportation.

The hydrologic modeling in general and streamflow in particular, has grown quickly in the recent years. These models can be divided into two main categories: Physically-Based models or numerical, and Data-Driven Models (DDMs) (Solomatine et al. 2008). Physically-based models are known as white box models which considers the physical interaction between the variable in modelling the hydrologic processes. Data-Driven models known as black box models that obtains the optimum solution between the inputs and outputs without taking into account the physical process; that does not mean choosing the inputs and outputs arbitrarily but selecting them in such a way that their physical contributions can be interpreted. Physical models have a disadvantage that require a big number of inputs in addition to the high cost of computation. Therefore, DDMs have become globally popular where these shortcomings are overcome.

Recently, the quick development in the computing tools and in the technology, has led to the development of several DDMs such as: Artificial Neural Network (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Support Vector Machines (SVM), Genetic Programming (GP), Extreme Learning Machine (ELM), and Extreme Gradiant Boosting (XGB) (Kişi 2009, Adamowski et al. 2010, Pramanik et al. 2011, Nourani et al. 2013, Nourani et al. 2014, Alizadeh et al. 2017, Nury et al. 2017).

The forecast of streamflow is identified as not an easy task due to the complexity of the process (Ravansalar et al. 2017). This complexity is mostly due to the non-stationarity and non-linearity in the relation between the streamflow and the basin characteristics (Nourani et al. 2011, Mehr et al. 2017). Data-Driven Models (DDMs) can be used for modeling streamflow as they have the ability to deal with non-linearity and map the relationship between the streamflow and factors driving it.

The DDMs can deal with the nonlinearity and non-stationarity in the mean and variance elements but the main shortcoming of the DDMs is that they may not be able to cope with the non-stationarity fluctuation (Nourani et al. 2011) if no preprocessing applied to the data (Cannas et al. 2006). One of the solutions is to apply the wavelet which have the ability to cope with the seasonal (cycling) nonstationary component of the time series (Daubechies 1990, Torrence et al. 1998, Anctil et al. 2004, Nourani et al. 2011). Wavelet has good ability in representing the signal locally in both time and frequency domains. The decomposition of the nonstationary time series using wavelet into several number of scales extracts the historical hidden information in both time and frequency domain (Rajaee et al. 2010, Shoaib et al. 2014). Therefore, Wavelet as preprocessing with various DDMs has been widely applied in the streamflow forecasting.

The wavelet transformation is performed by two forms namely Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). More than 80% of the hydrological studies is applying the DWT as preprocessing and not CWT (Nourani et al. 2014). This percentage is much higher in the streamflow forecasting. This is due to the nature of the hydrologic time series which gathered not in a continuous but in discrete form. Therefore, DWT which is dyadic is more appropriate for decomposing the time series which allows for reconstructing the original time series form the sub-signals. On the other hand, CWT represents the signal in many different and exact scales (periods) that in every scale there is a number of coefficients equal to the same number of time steps. Therefore, it is considered as redundant information especially when used as preprocessing to be then imposed into a DDMs as increasing the number of redundant inputs deteriorate model's ability in forecasting (Galelli et al. 2014, Taormina et al. 2016).

Several studies used CWT for forecasting streamflow by taking one scale coefficients or several scales based on the data used (Cannas et al. 2006, Adamowski 2008, Shoaib et al. 2014) but no one has tried to select only the scales that contribute in the streamflow using a hybrid model. In this study, two hybrid models are proposed in

which only the scales that important in forecasting the streamflow is chosen to be them modeled using any of the DDMs. This hypothesized based on that if the DWT which is dyadic decomposes the time series and reveals the seasonality, periodicity, and irregularity the CWT should reveal more hidden information as the decomposition is more and nowadays the advancement in the technology has overcome the shortcoming of the costly calculations.

These proposed models are implemented according to the procedure that has been well explained in the literature as many hybrid models are developed using WT and DWT in particular as preprocessing techniques with the DMMs in which the performance of these models in forecasting increased dramatically. However, most of the studies (including the proposed two models in this work) applied the hybrid models in such a way future information are sent into the model where these must not be included in the forecasting experiment. In the case of DWT, such studies implemented in such a way that all the data are decomposed and reconstructed for getting the sub time series. Then, the sub time series are divided into calibration/training and validation subsets to be after that imposed into DDMs. In the use of CWT, the same procedure followed except choosing the most contributing scale(s) to be imposed into DDM as CWT produce redundant information. Such a way of implementation sends some amount of the future information to the model and therefore this is named as hindcasting experiment in which all the data are used in calibration and forecasting.

According to that, most of the studies in the literature is applying the preprocessing WT in an incorrect way. This firmly includes the studies that has explained the procedure of the application. Therefore, a real forecast using hybrid model includes WT in both DWT and CWT is to be conducted in which only the existing data included in the model calibration and the future information used for testing the models. This forecast experiment is to be compared with the stand-alone models in which no WT-preprocessing is applied.

The objectives of this study are:

- Investigate the ability of several DDMs in forecasting daily streamflow one to seven days ahead.
- As monthly streamflow is less predictable, a new hybrid model for forecasting streamflow proposed that uses GP approach as a selection tool for the important scales of the CWT in order to be imposed after that into DDMs.

- Propose another hybrid model with the same idea of imposing the scales of CWT into XGB model which has the ability to obtain the importance of the scales in order, to then choose these important scales and use them as inputs into DDMs.
- Conduct a real forecast experiment of the hybrid models that uses WT as preprocessing and compare with the hindcast experiment.

The thesis is organized in five chapters. Chapter one consists of two studies. First one is the spatial interpolation of the temperature and rainfall as in one of the used basins no station located inside the basin and to obtain these two variables interpolation is required. Second study is evaluating the CRUTS data which is global monthly data of several meteorological variables and this study conducted due to the unavailability of evapotranspiration data which then obtained from CRUTS. Chapter two is the data exploration in which the data was explored as an essential step in data modelling. Firstly, the long-term (i.e. 1901-2015) trend and stationarity was explored for the whole Turkey including the study area. Secondly, the trend and stationarity only for the study area for the available streamflow data period are investigated. In chapter three, the forecast of one day ahead streamflow using several DDMs is conducted. Chapter four contains two studies. The first study is proposing a combination of DWT and CWT to improve seven day ahead streamflow forecasting. The second study is improving the monthly streamflow using CWT with a new hybrid model that choses only the important scales. In chapter five, another new hybrid models is proposed in which only the important scales of CWT is chosen in order to be imposed into DDMs. In this chapter, the real forecast is conducted beside the hindcast experiment in which to investigate that is including WT improving or deteriorating the models.

CHAPTER ONE

1. DATA PREPARATION

1.1. Comparison of Spatial Interpolation Methods of Precipitation and Temperature Using Multiple Integration Periods

1.1.1. Abstract

Eight spatial interpolation methods are used to interpolate precipitation and temperature over several integration periods in a local scale. The methods used are inverse distance weighting (IDW), Thiessen polygons (TP), trend surface analysis (TSA), local polynomial interpolation (LPI), thin plate spline (TPS), and three Kriging methods: ordinary, universal, and simple (OK, UK, and SK). Daily observations from 17 stations in the Seyhan Basin, Turkey, between 1987-1994 are used. A variety of parameters and models are used in each method to interpolate surfaces for several integration periods, namely, daily, monthly and annual total precipitation; monthly and annual average precipitation; and daily, monthly and annual average temperature. The performance is assessed using independent validation based on four measurements: the root mean squared error (RMSE), the mean squared relative error (MSRE), the coefficient of determination (r^2), and the coefficient of efficiency (CE). Based on these validation measurements, the method with smallest errors for most of the integration periods concerning both precipitation and temperature is IDW with a power of 3, whereas TP has the highest errors. The Gaussian model is found superior than other models with less errors in the three Kriging methods for interpolating precipitation, but no specific model is better than another for modeling temperature. UK with elevation as the external drift and SK with the mean as an additional parameter show no superiority over OK. For precipitation, annual average and monthly totals are found to be the worst and best modeled integration periods respectively, with the monthly average the best for temperature.

1.1.2. Introduction

Precipitation is the main source of soil moisture and drinking water in semiarid areas. Precipitation observations are crucial for hydrology, meteorology, geology, climatology and environmental monitoring and analysis. However, it is almost impossible to have stations (measuring precipitation, temperature, wind, etc.) that cover the whole area of interest, particularly in areas with no human activity. Meteorological stations have

limited, localized and dispersed point distributions. Spatial interpolation of weather parameters is considered to be an important aspect for most hydrological studies (Lam et al. 2015, Xu et al. 2015). Spatial interpolation is a technique that is used to estimate values for any distributed variable where none exists; develop a surface representing elevation, precipitation, temperature, pollution or other environment layers; or approximate values for any geographic location using known specific points. Meteorological stations are used as known locations, and their values are used to estimate (or predict) the values of a data surface in locations where no point data exist.

Spatial interpolation techniques can be divided into deterministic interpolation models (e.g., IDW, trend surface analysis or least square regression, RBF, and GPI) and geostatistical models (e.g., OK, SK, and UK) (Johnston et al. 2001, Wang et al. 2014). Deterministic models based on the assumption that the interpolated surface is more influenced by nearby points and less by distant points and depend on particular mathematical formulas that control the smoothness of the interpolated surface. Geostatistical models based on statistical models which include statistical relationship between the points (i.e. Autocorrelation) and rely on the assumptions that data come from stationary stochastic process. Spatial interpolation models can also be classified into three classes based on the method and scale of the interpolation application: models considered as simple interpolative, models use ancillary data, and models using complex models (Yang et al. 2015).

The desired time scale, spatial resolution (Frazier et al. 2016), the density of the station network, and the topographic complexity (Hofstra et al. 2008) are the leading factors of determining the complexity of the suitable spatial interpolation method and making the spatial pattern. Spatial variance varies especially for precipitation as the time scale varies mainly for annual means (Nalder et al. 1998, OE et al. 2008, WMO 2008, Frazier et al. 2016). The uncertainties related to the predictions increases noticeably as the time integration and station density decrease. This leads to the need of more complex models (OE et al. 2008, Frazier et al. 2016).

A number of spatial interpolation models have been created and applied in several fields of spatial interpolation (Hutchinson 1998, Goovaerts 2000, Lloyd 2005, Hofstra et al. 2008, Brunetti et al. 2014, Wang et al. 2014, Xu et al. 2015, Frazier et al. 2016). A number of studies has been conducted in using interpolation methods with a small number of stations. For example, Dirks et al. (1998) conducted a comparative study applying four

spatial interpolation methods for 13 precipitation stations, Keblouti et al. (2012) compared three methods using 10 stations, and Wang et al. (2014) compared six interpolation methods using 12 meteorological stations. These studies used several common methods among them such as IDW and some of uncommon methods such as LPI and Spline without applying several parameters and models. Dirks et al. (1998) and Keblouti et al. (2012) said IDW while Wang et al. (2014) said LPI outperformed the other methods.

The objective of this study is to compare the ability of 8 different spatial interpolation methods in modelling precipitation and temperature with different integration periods, i.e., daily, monthly, and annual total precipitation; monthly and annual average precipitation; and daily, monthly, and annual average temperature, in a local scale with a small number of stations i.e. 17 stations, using different parameters and models for each method.

The methods used are: inverse distance weighting (IDW), Thiessen polygons (TP), thin plate spline (TPS), trend surface analysis (TSA), local polynomial interpolation (LPI), ordinary Kriging (OK), universal Kriging (UK) (sometimes called Kriging with External Drift KED), and simple Kriging (SK).

1.1.3. Study area and data collection

This study aims to conduct a thorough comparison of several interpolation methods and identify the best method for generating precipitation and temperature maps. In the scope of this study, the word best refers to the model with the lowest values of errors (i.e. RMSE, and MSRE) and highest agreement (i.e. r^2 and CE) and the worst, on the contrast, refers to the model with the highest values of errors and lowest agreement.

Maps are to be generated for three sub-basins called (coded) 1822, 1801, and 1805, as seen in Figure 1.1. These three sub-basins are part of the Seyhan Basin located in southwest Turkey. To include all of the influencing stations (i.e., having spatial correlation) in the area inside the basin borders, stations were chosen inside and outside the borders. The area includes parts of four of Turkey's provinces: Adana, Kayseri, Kahramanmaras, and Sivas. During the study period, the sub-basins are: rural, agricultural, not hydrologically disturbed, and having no water storage structure. The annual average temperature of the study area is 9.1 C^0 , the maximum temperature 29.8 C^0 observed in July, and the minimum temperature -6.8 C^0 observed in January.

Observations of 17 stations were collected from DMI (Devlet Meteoroloji Isleri), the Turkish General Directory of Meteorology. The stations density is one station per 3000 Km^2. The data were collected for the period from 1/1/1987 to 31/12/1994. For every station, the collected data included total daily precipitation and average daily temperature. In some stations, some missing (non-recorded) observations were filled using the arithmetic average; they had no effect on the result due to the low amount of such instances. A simple statistical table for the collected data for all of the stations are summarized in Table 1.1 and Table 1.2.

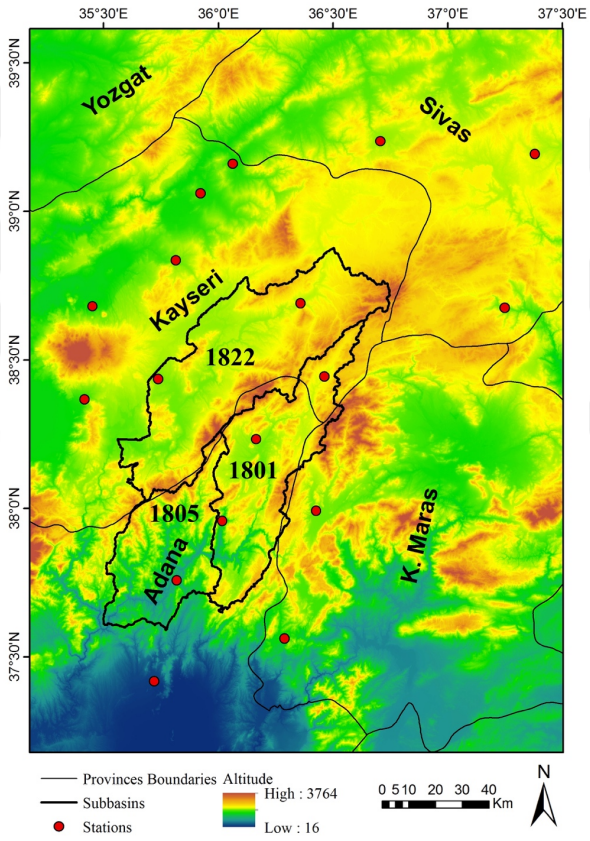


Figure 1.1. *The distribution of the meteorological stations and the topographic map of the study area which is located in Seyhan Basin, Turkey and covers parts of Kayseri, Adana, K. Maras, and Sivas provinces*

1.1.4. Methods

In this study, 8 interpolation methods were used with different parameters for deterministic methods and different models for geostatistical methods. Each method with different parameters and models was applied to every integration period. The integration

periods used are: daily, monthly, and annual total precipitation; monthly and annual average for precipitation; and daily, monthly, and annual average temperature.

The main steps of this study are (1) convert the daily (i.e., average for temperature and total for precipitation) observations for all stations to monthly and annual values; (2) apply each method considered in this study with a variety of parameters; and (3) implement cross-validation to assess the accuracy of the method using the root mean squared error (RMSE), mean squared relative error (MSRE), coefficient of determination (r^2), and coefficient of efficiency (CE). All of the steps and methods were implemented using the statistics language R with the exception of exporting several maps to ArcGIS. A brief description of each method is introduced in the next section.

Table 1.1. *Statistics of the observed daily total precipitation (in mm) of 17 stations for the period of 1987 to 1994*

Station No	Station Name	Min	Max	Range	Sum	Median	Mean	Var	Std. deviation
17162	Gemerek	0	50.6	50.6	3262.4	0	1.12	10.66	3.27
17196	Kayseri Bolge	0	51.2	51.2	3468.3	0	1.19	12.61	3.55
17762	Kangal	0	33.6	33.6	3091.5	0	1.06	8.21	2.87
17802	Pinarbasi	0	24.7	24.7	2894.4	0	0.99	7.67	2.77
17836	Develi	0	57	57	2789.6	0	0.96	9.74	3.12
17837	Tomarza	0	40.2	40.2	3154.5	0	1.08	10.75	3.28
17840	Sariz	0	48.6	48.6	4310.9	0	1.48	17.95	4.24
17841	Gurun	0	58.2	58.2	2510.5	0	0.86	8.72	2.95
17866	Goksun	0	59.1	59.1	4800.5	0	1.64	28.42	5.33
17908	Kozan	0	152.5	152.5	6614.9	0	2.26	61.43	7.84
18053	Tufanbeyli	0	79	79	4427.26	0	1.52	23.48	4.85
18056	Saimbeyli	0	126.8	126.8	7108.8	0	2.43	56.25	7.5
18156	Andirin	0	168	168	10406.24	0	3.56	140.38	11.85
18269	Feke	0	98.7	98.7	7339.89	0	2.51	65.29	8.08
18455	Bunyan	0	47	47	3945.8	0	1.35	15.53	3.94
18459	Sarioglan	0	34.5	34.5	2987.6	0	1.02	11	3.32
18466	Altinyayla (Sivas)	0	59.33	59.33	2730.72	0	0.94	9.52	3.09

Table 1.2. *Statistics of the observed daily average temperature (C^o) of 17 stations for the period 1987 to 1994*

Station No	Station Name	Min	Max	Range	Sum	Median	Mean	Var	Std. deviation
17162	Gemerek	-21	28.6	49.6	26909.6	10.4	9.21	94.53	9.72
17196	Kayseri Bolge	-16.1	32.2	48.3	29708.8	11.2	10.17	97.88	9.89
17762	Kangal	-23.2	26	49.2	17883.3	7.1	6.12	102.30	10.11
17802	Pinarbasi	-19.6	26.9	46.5	21117.2	8.5	7.23	92.73	9.63
17836	Develi	-16.8	30.7	47.5	30970.7	11.5	10.60	98.08	9.90
17837	Tomarza	-23.8	26.3	50.1	22023.9	8.7	7.54	107.48	10.37
17840	Sariz	-17	26.2	43.2	20749.4	7.7	7.10	86.57	9.30
17841	Gurun	-14.9	28.5	43.4	26338.8	9.5	9.01	95.54	9.78
17866	Goksun	-21.8	25.6	47.4	24884.7	9.4	8.52	100.07	10.00
17908	Kozan	0.7	35	34.3	55738.2	19.2	19.08	59.75	7.73
18053	Tufanbeyli	-18.6	27.5	46.1	27983.52	10	9.58	102.41	10.12
18056	Saimbeyli	-7.6	29.2	36.8	37753.7	13.1	12.92	76.04	8.72
18156	Andirin	-13.34	30.9	44.24	35844.34	12.9	12.27	68.49	8.28
18269	Feke	-6.65	32.2	38.85	43366.68	14.9	14.84	71.31	8.45
18455	Bunyan	-15.5	27.6	43.1	25039.6	9.6	8.57	83.95	9.16
18459	Sarioglan	-24.1	28.8	52.9	28397.3	11	9.72	99.91	10.00
18466	Altinyayla (Sivas)	-21.7	26.03	47.73	22066.27	8.6	7.55	90.57	9.52

1.1.4.1. Inverse Distance Weighting (IDW)

IDW is a direct deterministic interpolation method that is broadly applied in spatial interpolation applications. It was developed based on the assumption that the interpolated points are the most affected by the nearest points and the least affected by the most distant points (Wang et al. 2014, Borges et al. 2016). IDW is a local, exact, and deterministic method.

The general equation of IDW is as follows:

$$\hat{Z}(s_0) = \sum_{i=0}^{N} \lambda_i Z(s_i), \quad (1.1)$$

where $\hat{Z}(s_0)$ is the estimated values at location s_0, N is the number of points located around the point to be estimated used in the calculation, $Z(s_i)$ is the value of the known points measured at s_i, and λ_i is the weight corresponding to each known point, which is inversely proportional to the distance between the known points and the estimated point. The weights are calculated as follows:

$$\lambda_i = \frac{d_{i0}^{-p}}{\sum_{i=1}^{N} d_{i0}^{-p}}, \quad (1.2)$$

where p is the power, which controls the influence of the distance between the points on the estimation value; N is the number of points used in the estimation; and d_i is the distance between the point to be estimated s_0 and the known point s_i (Isaaks et al. 1989, Nalder et al. 1998, Borges et al. 2016). IDW was applied with three power p values: 1, 2, and 3 for all of the integration periods for precipitation and temperature.

1.1.4.2. Thiessen Polygons/Nearest Neighbor (TP)

Thiessen polygons, which are named according to its developer, Thiessen (1911), are considered to be a special case of the IDW method because only one point, i.e., the nearest point, is used in the interpolation. A Thiessen polygon consists of all of the points in a polygon that lie nearer to the point used for constructing that polygon than to any another point (Goovaerts 2000). TP is a local, exact, and deterministic method (Shope et al. 2015).

1.1.4.3. Thin Plate Spline (TPS)

Thin plate spline was described in detail by (Wahba 1990). For a bivariate thin plate spline (i.e., one type of TPS) estimations of $Z(s_i)$ for measured points $i = 1 \ldots N$, is calculated as follows:

$$Z(s_i) = f(s_i) + \epsilon(s_i), \tag{1.3}$$

whereas $\epsilon(s_i)$ represents random errors, which are assumed to be uncorrelated random errors and independent with a zero mean and variance σ^2, and $f(s_i)$ represents an unknown deterministic smooth function, which can be estimated by the minimization of the following:

$$\sum_{i=1}^{N} [Z(s_i) - f(s_i)]^2 + \lambda J_m^d(f), \tag{1.4}$$

where $f(s_i)$ is the fitted function values at the ith data point; λ is the so-called regularizer or smoothing parameter, which controls the trade-off between fitting the data as close as possible without losing the smoothness (i.e., if $\lambda = 0$ the function models the points exactly with zero noise, whereas if λ is very large, the function will be a hyperplane); and J_m^d is a measure of the smoothness of function f. The form J_m^d depends on two parameters: the number of independent variables d and the order of the partial derivatives m. For example, considering $d = 2$, which represents the two coordinates, X and Y, and $m = 2$, which represents the second partial derivative, the form is written as follows:

$$J_2^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ \left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 \right\} dxdy. \tag{1.5}$$

The minimization problem of Eq. (1.4) is well known as the thin plate smoothing spline. The term spline is given to this method because the solution of Eq. (1.4) for a univariate instance, with the second partial derivative (i.e., $m = 2$), is a natural cubic spline (Wahba 1990, Boer et al. 2001, Hancock et al. 2006).

According to (Craven et al. 1978), smoothing parameter λ, which corresponds to the spline function f, can be accurately calculated by the minimization of generalized cross validation (GCV), which has the following formula:

$$\frac{1}{N} \frac{(Z - A(\lambda)Z)^T (Z - A(\lambda)Z)}{(Tr(I - A(\lambda))/N)^2}, \tag{1.6}$$

where, according to (Wahba 1990, Hancock et al. 2006), $A(\lambda)$ is the matrix that alters the data value vector into the model predicted values vector, expressed as follows:

$$\hat{Z} = A(\lambda)Z, \tag{1.7}$$

where $\hat{Z}$ is the model predicted value vector and $A(\lambda)$ is termed the influence matrix.

1.1.4.4. Trend Surface Analysis/ Global Polynomial Interpolation (TSA)

TSA fits a smooth surface to measure points using a mathematical function. Over the area to be estimated, the surface of TSA changes gradually from one region to an adjacent region, and it has the ability to obtain a global trend in data. TSA utilizes all of the measured points to interpolate the surface in contrast to several other methods, such as IDW and LPI (i.e., briefed in the next section), which use a specific subset of measured points (Wang et al. 2014). TSA/GPI is considered to be a global, inexact, deterministic method.

The principle of TSA/GPI is that the entire study area is presented by a formula that estimates the value $\hat{\mathrm{Z}}(\mathrm{s}_i)$ at any location based on the $\mathrm{X_i}, \mathrm{Y_i}$ coordinates of that location. The general function is described as follows:

$$\hat{Z}(s_i) = f\,(X_i, Y_i). \tag{1.8}$$

The main objective in TSA/GPI is to use all of the measured points of the study area to obtain a formula that best describes that area. Many formulas exist, but the simplest one that can describe a surface is the first order bilinear surface:

$$f\,(X, Y) = b_0 + b_1 x + b_2 y + \varepsilon. \tag{1.9}$$

A second order trend surface can be represented by the following:

$$f\,(X, Y) = b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 xy + b_5 y^2 + \varepsilon. \tag{1.10}$$

In Eq. (1.10), not only the individual variables with their parameters are present, but so too is the cross product (i.e., $\mathrm{b_4xy}$). Higher order trend surfaces include more cross product terms, not only higher powers of x and y (Huisman et al. 1999). Regardless of using first, second, or even third order trend surfaces, the parameters of the used formula are obtained using a regression technique utilizing all of the measured points of the area of interest.

1.1.4.5. Local Polynomial Interpolation (LPI)

Global interpolation is based on the assumption that the entire study area can be represented by a general formula. However, in many cases, representing the study area

by the same mathematical surface misrepresents the real surface. This representation prevents variations within the natural geographic field. LPI was developed to override this shortcoming.

LPI has the same mathematical representation and fitting procedure as TSA/GPI, except LPI fits a local formula by utilizing measured points within a specified area, unlike TSA/GPI, which uses all of the measured point. The specified areas can overlap, and the estimated surface value at the center of that area is the predicted value (Huisman et al. 1999, Wang et al. 2014). LPI is a local, inexact, deterministic method.

1.1.4.6. Ordinary Kriging (OK)

Kriging is a local interpolation method that was initially developed by a mining engineer called D.G. Krige and a geostatistician called Georges Matheron. The Kriging method uses a subset of geostatistical interpolation methods with the assumption that closer points are more likely to be similar, whereas farther points are more likely to be different. To evaluate the dissimilarity between points, Kriging uses a semivariogram. The experimental semivariogram $\hat{\gamma}(\mathrm{h})$ is calculated using the following equation:

$$\hat{\gamma}(h) = \frac{1}{1N(h)} \sum_{i=1}^{N(h)} [Z(s_i) - Z(s_i + h)\,]^2\,, \tag{1.11}$$

where h is the lag, $\mathrm{Z}(\mathrm{s_i})$ is a set of data points, and N is the number of pairs of data points separated by h. A model is used to fit the experimental variogram. In this study, three models are used to compare the results: spherical, Gaussian, and linear. To save space, only the equation of the spherical model is mentioned. The spherical model is the most used model because it generally provides the best fit to the data. The spherical model is represented by the following:

$$\gamma(h) = \begin{cases} \theta_0 + \theta_s \cdot \left[\frac{3h}{2a} - \frac{1}{2}\left(\frac{h}{a}\right)^3\right] & for\ h \le a \\ \theta_s & for\ h \ge a \end{cases}, \tag{1.12}$$

where a is the range, θ_s is the partial sill, and θ_0 is the nugget (Webster R et al. 2007, Frazier et al. 2016). The lag h is considered to be the average distance between neighboring points, and the nugget θ_0 effect is attributed to the measurement errors for distances smaller than the data intervals (Borges et al. 2016).

Several types of Kriging methods have been widely used in different applications. Ordinary Kriging OK is the most used type. It is, in general, used as a base method to be compared with other methods (Goovaerts 2000, Mair et al. 2011, Juan Francisco Sanchez-Moreno et al. 2013, Frazier et al. 2016). OK assumes that the mean is constant, but its emphases on the spatial mechanisms are unknown (Isaaks et al. 1989, Cressie 1990, Borges et al. 2016). The representation of OK for a spatial process $Z(s)$ is as follows:

$$Z(s) = \mu + \delta(s), \tag{1.13}$$

where μ is an unknown expected value of a random process and δ(s) is a zero mean naturally from a stationary random process with an existing semivariogram.

The OK estimator is described exactly as Eq. (1.1) (i.e., the same equation of IDW) with the exception of weights, which are calculated based on the model fitting on the variogram differently from IDW, which is calculated according to the inverse distance. Kriging is an unbiased and optimal estimator, which means that the weights λ_i in all Kriging methods are summed to 1. The weights $\lambda_i: i = 1, \dots, N$ are calculated under the uniform unbiased condition as follows:

$$E\left(\hat{Z}(s_0)\right) = E\left(Z(s_0)\right) \qquad \sum_{i=1}^{N} \lambda_i = 1, \tag{1.14}$$

under the restriction of the minimization of the prediction error variance $\sigma(s_0)$,

$$\sigma(s_0) = E(Z(s_0) - \hat{Z}(s_0))^2. \tag{1.15}$$

OK is a local, exact, and stochastic method.

1.1.4.7. Universal Kriging (UK)

The main assumption of UK is that the data have a significant spatial trend, which makes it different from OK, which assumes no trend (Wang et al. 2014). A secondary variable is used by UK in the form of external drift or a trend (such as elevation or temperature), which gives UK the ability to account for nonstationarity in the data throughout the study area because there is a linear relationship between the interpolated variable and the used external drift, which is assessed locally (Goovaerts 2000, Frazier et al. 2016). Therefore, UK is sometimes called Kriging with external drift (KED). The trend or drift has continuous spatial variation, but is too irregular to be modelled by a simple mathematical functions. UK uses stochastic and deterministic components to incorporate the external variable into the Kriging system to overcome the irregularity (Webster R et

al. 2007). Taking into account the two constraints on the weights, UK uses $N + 2$ equations to solve for the weights. The estimator of UK or KED is described as follows:

$$\hat{Z}_{UK}(s_0) = \sum_{i=1}^{N} \lambda_i^{UK} z(s_i). \quad (1.16)$$

In this method, elevation extracted from the ASTER DEM (NASA 2016) was used as the external drift with the three models. Elevation used considering that it has high influence in the variation of the spatial pattern as it varies between 16-3764 meters (see figure 1).

1.1.4.8. Simple Kriging (SK)

Sometimes the mean of the random variable can be assumed according to the nature of the problem. This knowledge should be used to improve the estimation, and this can be done by using simple Kriging (SK). SK uses N equations to solve for N unknowns, and the estimator is described as follows:

$$\hat{Z}_{SK}(s_0) = \sum_{i=1}^{N} \lambda_i z(s_i) + \{1 - \sum_{i=1}^{N} \lambda_i\}\mu. \quad (1.17)$$

The weights λ_i are no longer constrained to the summation of 1, and unbiasedness is guaranteed by addition of the second part of Eq. (1.17). For the reason that weights are not summed to 1, covariances, C, must be used instead of semivariances, γ (Webster R et al. 2007). Comparing simple Kriging to ordinary Kriging, simple Kriging is supposed to produce smaller variance by having the mean estimated from the observed data.

1.1.5. Validation

The assessment of accuracy can be performed using a technique called cross-validation that is the most applied technique in climate studies. Cross-validation helps in making decisions about which model is best at estimating the surface using the measured points (OE et al. 2008, Borges et al. 2016). One of the cross-validation methods is leave one out cross-validation (LOOC). In this method, one point is left out of sample data, whereas the other points are used to estimate the value of the left point. This procedure continues until a value for each of the original data points is estimated (Isaaks et al. 1989).

One of the shortcomings of using cross-validation is that the interpolation model is defined using all of the sample data, which implies that the validation can be considered to be not completely independent (OE et al. 2008, Borges et al. 2016). Therefore, different

techniques were considered in this study, including separating the precipitation station data into 15 stations for interpolation and 2 stations for validation.

Four different measures used in validating interpolation methods (Robinson et al. 2006, Brunetti et al. 2014, Seo et al. 2015, Borges et al. 2016, Frazier et al. 2016) were applied in all of the methods for each integration period and used for comparing the interpolating methods:

The root mean square error (RMSE) is the square root of the average of the square of the difference between the observed (z) values and those estimated from the model ($\hat{z}$). The value of RMSE has the same unit as the variable (precipitation and temperature in this study) and should be as low as possible. It is calculated as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(\hat{z}(s_i) - z(s_i)\right)^2}{N}}. \tag{1.18}$$

The mean squared relative error (MSRE) is the average of the squared relative error, where the relative error is the difference between the observed (z) and the estimated ($\hat{z}$) values divided by the observed value. It is a ratio that indicates the overall agreement between the observed and estimated values. It is calculated as follows:

$$MSRE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{z(s_i) - \hat{z}(s_i)}{z(s_i)}\right)^2. \tag{1.19}$$

The coefficient of determination (r^2) is the square of Pearson's coefficient (Pearson 1896). It describes the proportion of the variance in the observed data that can be explained by the model. It ranges between 0 and 1, and this value represents the percentage of the variance explained and the overall agreement between the observed (z) and estimated ($\hat{z}$) values. It is calculated by the following equation:

$$r^2 = \left(\frac{\sum_{i=1}^{N}(Z(s_i) - \bar{Z})(\hat{Z}(s_i) - \tilde{Z})}{\sqrt{\sum_{i=1}^{N}(Z(s_i) - \bar{Z})^2 \sum_{i=1}^{N}(\hat{Z}(s_i) - \tilde{Z})^2}}\right)^2, \tag{1.20}$$

where $\bar{Z}$ is the mean of the observed sample points and $\tilde{Z}$ is the mean of the estimated points.

The coefficient of efficiency (CE) is also known as the Nash-Sutcliffe coefficient (Nash et al. 1970). CE has been widely applied in hydrologic models, especially precipitation. It is superior to r^2 because it takes into account the differences in the means

and variances of the observed and estimated points. It ranges between -∞ and one; one representing a perfect model—zero means the model is as accurate as the means of the observed data—and negative values indicating that the mean of the observed data is a better estimator than the model (Dawson et al. 2007). It is calculated as follows:

$$CE = 1 - \frac{\sum_{i=1}^{N}\left(Z(s_i) - \hat{Z}(s_i)\right)^2}{\sum_{i=1}^{N}(Z(s_i) - \bar{Z})^2}. \quad (1.21)$$

1.1.6. Results and Discussion

The applied 8 interpolation methods are used to provide an interpolation map for the whole study area utilizing several measured points. Therefore, a number of maps were generated. To save space, as examples, maps of every method for the monthly total precipitation and monthly average temperature for one month (i.e., June 1989) are shown in Figure 1.2 and Figure 1.3.

By examining these figures, the differences among the used methods and their theoretical nature can be clearly seen. The IDW interpolation map constructs circles or ellipsoids around the points used for the interpolation, and every one of these shapes represents a value. These shape values decrease, going further from the center, which highlights the theoretical background, as represented by Eq. (1.1) and (1.2). The Thiessen polygon interpolation map shows that every known point is represented by a polygon, which gives it a district nature. The LPI map is different from the TSA map in that the former has local shapes separated from the entire area and has internal discontinuities due to the local nature of the method. TSA uses one formula to model the whole area, whereas LPI uses many formulas in several parts of the area. The three maps of the Kriging methods (i.e., OK, UK, and SK) look similar, with slight differences in the values in some parts of the study area. For example, for the precipitation monthly maps shown in Figure 1.2, OK has slightly higher values in the upper middle part of the study area, whereas SK has slightly lower values in the center of the study area. In the temperature maps shown in Figure 1.3, UK has slightly lower values in the right part of the study area SK and UK.

After obtaining the results of the interpolation of each method, cross-validation measurements were obtained to make comparisons and identify the capability of these methods to spatially interpolate each weather parameter in several integration periods

with a small number of stations. The obtained measurements are summarized in Table 1.3-Table 1.5.

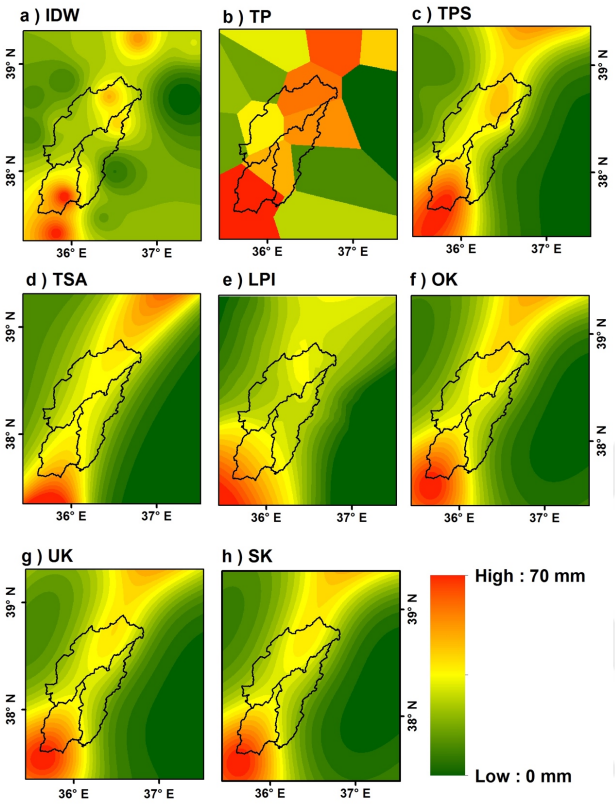


Figure 1.2. *Examples of interpolated maps of monthly total precipitation (in mm) for each of the 8 interpolation methods. The month chosen as an example is June 1989*

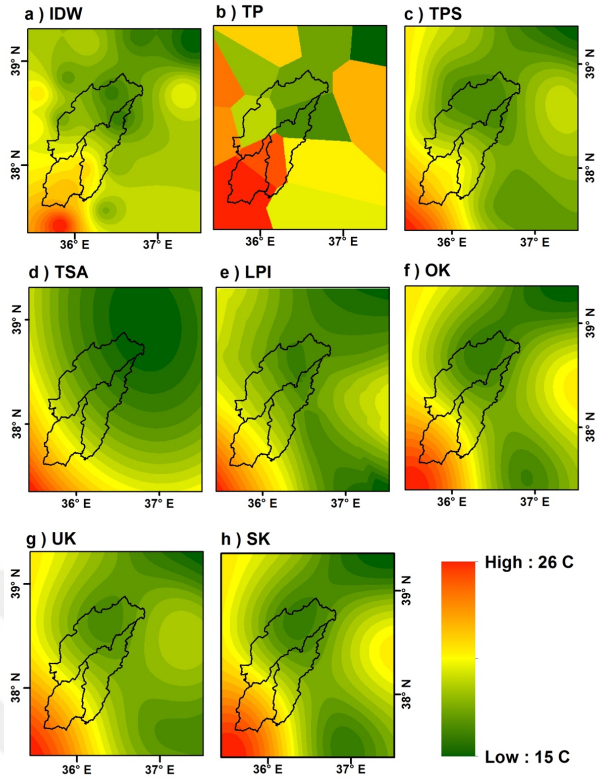


Figure 1.3. *Examples of interpolated maps of monthly average temperature (in C°) for each of the 8 interpolation methods. The month chosen as an example is June 1989*

Looking at Table 1.3 - Table 1.5, TP has the highest values of RMSE and MSRE, that this method performed the worst. Similarly, the OK method with a spherical model has a poor performance for both precipitation and temperature in all of the integration periods. Notwithstanding the high r^2 values for the average temperature in all the integration periods, it has low values of CE and high values of MSRE and RMSE.

Table 1.3. *Validation measurements: RMSE, MSRE, r2, and CE of daily, monthly, and annual total precipitation (in mm) interpolations*

Method	Power	Daily Total				Monthly Total				Annual Total			
		RMSE	MSRE	r2	CE	RMSE	MSRE	r2	CE	RMSE	MSRE	r2	CE
IDW	1	3.64	1.29	0.69	0.62	22.44	1.55	0.82	0.81	118.49	0.10	0.80	0.69
IDW	2	3.38	1.38	0.73	0.67	19.64	1.38	0.86	0.86	102.10	0.08	0.86	0.77
IDW	3	3.31	1.52	0.73	0.68	18.97	1.46	0.87	0.87	103.13	0.08	0.86	0.76
	Degree												
TSA	1	3.70	1.39	0.67	0.60	21.90	1.40	0.83	0.82	96.53	0.06	0.81	0.79
TSA	2	3.52	1.52	0.69	0.64	20.97	1.59	0.84	0.84	98.82	0.06	0.87	0.78
TSA	3	3.60	2.21	0.71	0.63	21.92	2.24	0.86	0.82	142.15	0.15	0.80	0.55
LPI	1	3.43	1.35	0.71	0.66	19.75	1.12	0.86	0.86	93.81	0.06	0.86	0.80
LPI	2	3.49	2.46	0.72	0.65	21.82	2.20	0.85	0.82	130.90	0.14	0.82	0.62
TP		5.36	5.19	0.65	0.17	44.97	4.88	0.82	0.25	358.32	0.49	0.79	-1.85
TPS		3.51	1.82	0.70	0.64	20.20	1.89	0.86	0.85	114.92	0.09	0.77	0.71
	Model												
OK	Sph	4.58	3.17	0.47	0.11	23.44	1.85	0.79	0.73	112.56	0.11	0.30	-0.73
OK	Gau	3.15	1.25	0.67	0.63	19.17	1.47	0.86	0.84	104.05	0.09	0.84	0.55
OK	Lin	3.29	1.55	0.62	0.59	20.15	1.34	0.82	0.81	147.49	0.10	0.54	0.52
UK	Sph	5.22	3.02	0.44	-0.06	28.49	2.38	0.68	0.61	285.36	0.44	0.01	-2.63
UK	Gau	3.21	1.65	0.62	0.56	19.08	1.82	0.84	0.83	126.61	0.13	0.73	0.48
UK	Lin	3.13	0.98	0.65	0.63	19.87	1.67	0.82	0.81	137.20	0.10	0.63	0.58
SK	Sph	3.21	0.14	0.46	0.38	23.60	1.48	0.75	0.72	173.28	0.20	0.44	0.27
SK	Gau	2.53	0.13	0.57	0.56	20.26	1.40	0.81	0.80	132.73	0.13	0.61	0.43
SK	Lin	2.76*	0.13	0.53	0.51	23.96	1.24	0.72	0.72	164.17	0.21	0.45	0.35

Table 1.4. *Validation measurements: RMSE, MSRE, r2, and CE of monthly, and annual average precipitation (in mm) interpolations*

Method	Power	Monthly Average				Annual Average			
		RMSE	**MSRE**	**r2**	**CE**	**RMSE**	**MSRE**	**r2**	**CE**
IDW	1	1.67	57.68	0.83	0.04	1.63	2.43	0.83	-6.91
IDW	2	1.64	54.38	0.87	0.07	1.61	2.31	0.87	-6.69
IDW	3	1.67	53.05	**0.88**	0.04	1.63	2.31	**0.87**	-6.86
	Degree								
TSA	1	1.53	54.53	0.83	0.19	1.45	**1.75**	0.78	-5.26
TSA	2	1.61	51.10	0.85	0.10	1.53	2.16	0.87	-5.99
TSA	3	1.79	53.23	0.87	-0.11	1.73	2.61	0.82	-7.91
LPI	1	1.60	51.88	0.87	0.12	1.55	2.17	0.86	-6.19
LPI	2	1.77	53.46	0.86	-0.08	1.70	2.56	0.84	-7.56
TP		2.48	59.65	0.82	-1.13	2.23	3.35	0.79	-13.82
TPS		1.68	51.89	0.86	0.03	1.59	2.24	0.78	-6.56
	Model								
OK	Sph	1.84	38.56	0.74	-0.78	1.46	1.82	0.60	-11.28
OK	Gau	1.52	43.85	0.85	0.04	1.62	2.36	0.84	-6.77
OK	Lin	1.53	43.09	0.81	0.06	1.52	2.18	0.71	-5.87
UK	Sph	2.49	42.51	0.38	-1.38	1.69	2.62	0.74	-7.46
UK	Gau	1.48	40.75	0.86	-0.01	1.67	2.53	0.80	-7.28
UK	Lin	1.45	41.60	0.82	0.13	1.45	1.96	0.76	-6.34
SK	Sph	2.31	15.89	0.35	-1.28	1.46	1.77	0.14	-5.34
SK	Gau	1.25	12.60	0.72	0.32	1.53	2.06	0.78	-5.96
SK	Lin	**1.22***	**11.53**	0.69	**0.36**	**1.39**	1.79	0.74	**-4.75**

* Bold numbers are the lowest in RMSE and MSRE and the highest in r^2 and CE.

IDW with a power of three is identified as the best at modelling precipitation with the exception of some measurements where SK with the linear model performed better regarding the average precipitation. This finding is not consistent with: (Apaydin et al. 2004), who found that in 19 studies, IDW was not recommended (i.e., mentioned in his research) for interpolating precipitation, except in one study, and (Wang et al. 2014) who found LPI outperformed IDW. On the contrast, this finding agrees with (Dirks et al. 1998, Keblouti et al. 2012). IDW was also found to be the best method for interpolating temperature (Table 1.5), with the highest (high r^2 and CE and low RMSE and MSRE) validations.

TPS can also be considered to be a well-behaving model because it has high values for all of the integration periods, except CE values for the temperature, which are slightly low. Among the Kriging methods, OK, UK, and SK all with the Gaussian model performed better by having less error and higher agreement in modeling precipitation for the various integration periods, which indicates that the Gaussian model better models the precipitation data than the two other models used. However, OK with the linear model, UK with the Gaussian model, and SK with the linear model have higher results in

modeling the temperature, which gives preference, in general, to the linear model for the temperature data.

Table 1.5. *Validation measurements: RMSE, MSRE, r2, and CE of daily, monthly, and annual average temperature (in C°) interpolations*

Method	Power	Daily Average				Monthly Average				Annual Average			
		RMSE	MSRE	r2	CE	RMSE	MSRE	r2	CE	RMSE	MSRE	r2	CE
IDW	**1**	2.91	2.26	0.97	0.96	2.19	0.67	0.98	0.97	0.79	0.01	0.68	0.65
IDW	**2**	2.67	1.87	0.97	0.96	2.00	0.52	0.99	0.98	0.90	0.01	0.68	0.54
IDW	**3**	2.48	**1.53**	**0.97**	**0.97**	1.80	0.37	**0.99**	**0.98**	0.85	**0.01**	0.75	0.59
	Degree												
TSA	**1**	3.91	6.73	0.93	0.92	3.37	1.76	0.94	0.94	2.48	0.07	0.15	-2.48
TSA	**2**	3.80	5.11	0.94	0.93	3.26	1.06	0.95	0.94	1.75	0.03	0.62	-0.74
TSA	**3**	3.73	2.71	0.95	0.93	3.22	0.92	0.96	0.94	1.02	0.01	0.73	0.41
LPI	**1**	2.55	1.71	0.97	0.97	1.82	0.34	0.99	0.98	0.75	0.01	0.83	0.68
LPI	**2**	3.14	1.67	0.96	0.95	2.56	**0.32**	0.98	0.96	1.54	0.02	0.73	-0.34
TP		4.32	8.37	0.93	0.91	3.80	1.87	0.94	0.92	3.42	0.14	0.14	-5.63
TPS		3.08	3.60	0.96	0.95	2.40	0.92	0.98	0.97	1.26	0.02	0.87	0.10
	Model												
OK	**Sph**	2.82	18.39	0.91	0.90	2.13	1.84	0.94	0.91	**0.31**	0.00	0.95	**0.93**
OK	**Gau**	2.15	3.56	0.96	0.96	1.30	0.80	0.98	0.98	1.48	0.02	0.78	-0.25
OK	**Lin**	2.09	3.25	0.96	0.96	**1.22**	2.82	0.99	0.98	1.11	0.02	**0.97**	-0.45
UK	**Sph**	2.01	4.36	0.94	0.94	1.40	2.07	0.96	0.96	1.03	0.01	0.71	-0.01
UK	**Gau**	**1.44***	2.70	0.96	0.96	0.86	0.43	0.98	0.98	0.87	0.01	0.86	0.29
UK	**Lin**	1.99	3.68	0.95	0.95	1.48	2.16	0.97	0.96	0.95	0.01	0.75	0.14
SK	**Sph**	3.62	8.12	0.94	0.94	2.80	8.43	0.96	0.96	0.65	0.01	0.83	0.76
SK	**Gau**	2.63	3.69	0.97	0.97	1.76	2.54	0.99	0.98	1.50	0.03	0.78	-0.27
SK	**Lin**	2.64	3.65	0.97	0.97	1.69	3.13	0.99	0.98	1.23	0.02	0.78	0.14

* Bold numbers are the lowest in RMSE and MSRE and the highest in r^2 and CE.

Even though, the elevation was used as an external drift with hope of improving the result, as in (Goovaerts 2000, Moges et al. 2007), UK with the three models has shown no advantage, similar to the findings of (Mair et al. 2011, Wang et al. 2014, Frazier et al. 2016), in comparison with OK and SK.

In terms of integration periods, the annual cross validation results have shown instability in the interpolation methods unlike the daily and monthly results, which changed together in most of the methods. Monthly average precipitation and temperature have the highest r^2 values, whereas monthly total precipitation and monthly average temperature have the highest CE.

Based on the CE results shown in Table 1.4, the models of annual average precipitation are not better than using the mean of the observed data because they all have negative CE values. This is due to the low density of the stations network and the high spatial variability in the annual average. It is also important to mention that the length of the time series used which is only 8 years is not enough to obtain considerable results of the annual integration time.

In precipitation, the best interpolated integration period is the monthly total, and the worst is the annual average, whereas the monthly average of temperature is the best in most of the used methods with the various parameters.

Specifically, the values of the cross-validation measurements slightly change from method to method and from one integration period to another. In one method, considering a specific measurement can be found as having the lowest error while in another measurement having the highest in comparison with other methods, so there is no clear evidence that one method is superior to another at modeling an area with small number of stations and highly changing elevation, with the exception of the TP, IDW and LPI methods. The obvious worst modeling is achieved using TP, and the best modeling is achieved using the other two methods. There is also no one integration period better than others significantly except the annual average precipitation which had an extreme poor performance.

1.1.7. Conclusion

Eight deterministic and geostatistical spatial interpolation techniques were used in this study, and 4 cross-validation measurements were obtained to compare the capability of the methods and identify the best modeling method of precipitation and temperature for several integration periods for the study area. Generally, no method is superior to the others from one integration period to another. Comparing the four validation measurements, IDW with a power of three is identified as the best modeling of precipitation and temperature. Not surprisingly, TP had the worst performance for all of the integration periods and for the two weather parameters.

All of the methods are better at interpolating the temperature in the different integration periods, due to the smoothness in the variation of the temperature through the area. In contrast, precipitation is worse due to the coarse variation throughout the area.

All the methods are capable of modeling the different integration periods for both precipitation and temperature with the exception that having no model and using the average is better in modeling annual average precipitation. This is due to the small number of stations in a vastly changing topography. The used period is limited to 8 years. Therefore, using longer time series data for both precipitation and temperature is recommended to examine the annual interpolation in more detail.

UK with elevation as external drift has no advantage in comparison with SK and OK even though the elevation which has high range is thought as an influencing factor in such scale and station network density. Therefore, the use of other factors, such as the distance to rivers, slope, and aspect, is recommended.

1.2. Conversion of CRUTS 3.23 Data and Evaluation of Precipitation and Temperature Variables in A Local Scale

1.2.1. Abstract

The precipitation and temperature data obtained from the version 3.23 of Climate Research Unit Time Series (CruTS) data are evaluated in a local scale. The data obtained for the period 1901-2014 consists of ten climate variables: Precipitation, Mean temperature, Diurnal Temperature Range, Wet-day Frequency, Vapour Pressure, Cloud Cover, Maximum temperatures, Minimum temperatures, frost day frequency, and potential evapotranspiration. These data converted to ASCII format which is usable in GIS environments. Daily precipitation and temperature data of 9 station located in Seyhan basin in Turkey for the period 1973-2000 used for the assessment after converting them to monthly total precipitation and average temperature. The assessment implemented using several statistical measurements: correlation, RMSE, mean difference, median difference, and Mann-Whitney test. The study found that CruTS has high agreement with the observed data in the meaning of the correlation and low RMSE in all stations. The result of mean, median, and Mann-Whitney tests indicate no significant difference between CruTS and the observed data in all station with the exception of two stations; Tomarza station has significant values in the temperature values in all tests; Goksun station has significance difference in the mean test in the two climate variables.

1.2.2. Introduction

Climate Research Unit Time Series (CruTS) was named based on the name of the producer of these data i.e. the Climate Research Unit, University of East Anglia, Norwich, UK. Initially these data developed by New et al. (2000) for the period 1901- 1996, updated by Mitchell et al. (2005) for the period 1901 - 2002, and finally updated one more time by Harris et al. (2014) for the period 1901-2009. The data are still being added to the collection and corrected.

The 3.23 version of CruTS datasets consists of 0.5° latitude-longitude grid cells covering the global with the exception of Antarctica region. The datasets provide monthly values of ten climate variables. Six of these variables are independent: Precipitation (pre), Mean temperature (tmp), Diurnal Temperature Range (dtr), Wet-day Frequency (wet), Vapour Pressure (vap), and Cloud Cover (cld). The remain four (i.e. Maximum temperatures (tmx), Minimum temperatures (tmn), frost day frequency (frs), and potential evapotranspiration (pet)) are calculated and estimated using the independent variables (Jones et al. 2008).

CruTS is routinely updated using several resources: CLIMAT monthly data for about 2400 station, for the last 4-5 years, produced by the collaboration between World Meteorological Organization (WMO) and the US National Oceanographic and Atmospheric Administration (NOAA), Monthly Climatic Data for the World (MCDW) produced by UK National Climatic Data Center (NCDC) for incorporating in WMO by 200 stations, World Weather Records (WWR) decadal data for about 1700 stations during 1991-2000 published as exchange between the archive center of NCDC and the National Meteorological Services (NMSs). The aforementioned sources provide a systematic incorporation in the updating process.

The importance of CruTS datasets has been increasing in Geosciences researches with a focus on the climate change studies due to two reasons: the long period that covering from 1901 – 2014 that give the ability for studying the trend and the change in the climate, and the high spatial resolution which helps in local not only regional studies. CruTS with its different versions have been widely used in many researches and in many applications such as trend analysis, climate change, desertification and many others. For example, but not limited to (Kottek et al. 2006, Hellden et al. 2008, Hoff et al. 2010, Garcia et al. 2014, Poulter et al. 2014, Taxak et al. 2014).

In order to use this data for applications, they need to be validated in different scales. (Harris et al. 2014) created the CruTS 3.10 data using Climate anomaly method (CAM) (Peterson et al. 1998) and validated them in a sub-continental scales by using: temperature dataset developed by University of Delaware (UDEL) based on GHCN-M (Peterson et al. 1998) for the period 1901-2008, and precipitation data developed by Global Precipitation Climatology Center (GPCC) for the period 1901-2009. They found a tight agreement between the temperature data of CruTS and UDEL. CruTS and GPCC precipitation data have high agreement also, but not as tight as temperature data.

The main objective of this study is to evaluate the last version of CruTS 3.23 released 1 September 2015 in a local scale. The dataset to be compared with the observed monthly total precipitation and mean temperature data of 9 stations located in Seyhan basin, Turkey for the period 1973 – 2000. Additionally, the CruTS 3.23 dataset for the period 1901 – 2014 for all the variables converted to GIS-based ASCII files to make them usable in GIS environments.

1.2.3. Data collection

CruTS data are originally available at Centre for Environmental Data Archival of the British Atmospheric Data Center (BADC) which can be reached by (http://www.badc.nerc.ac.uk). The data provided in two format: ASCII and NetCDF. In the same time, a temporal or spatial subset of the datasets can be obtained with CSV file format from CEDA OGC WEB SERVICES/ Web Processing Service reached by (http://wps-web1.ceda.ac.uk/).

The Consortium for Spatial Information, CSI, which is the CGIAR spatial scientists community (http://www.cgiar-csi.org/) processed the CruTS 3.10 data which covers the period 1901 to 2009 and converted them to ASCII in order to facilitate their use in GIS environment.

In this study, CruTS 3.23 data for the period 1901 – 2014 released 1 September 2015 evaluated by comparing them with the observed data. The data set collected from CEDA OGC WEB SERVICES/ Web Processing Service with CSV format. The collected data set converted to ASCII format to provide a full set contains the entire available variables and period.

Observations of 9 stations were collected from DMI (Devlet Meteoroloji Isleri); the Tukrish General Directory of Meteorology. These stations located at three sub-basins called (coded): 1822, 1801, and 1805 as part of Seyhan basin located in the southwest of Turkey. The area is taking part from four Turkey's provinces: Adana, Kayseri, Kahramanmaras, and Sivas, as seen in Figure 1.4.

The data were collected for the period from 1/1/1973 to 31/12/2000. For every station, the collected data were total daily precipitation and average daily temperature. These data converted to monthly total precipitation and monthly average temperature to be compared, after that, with the CruTS data.

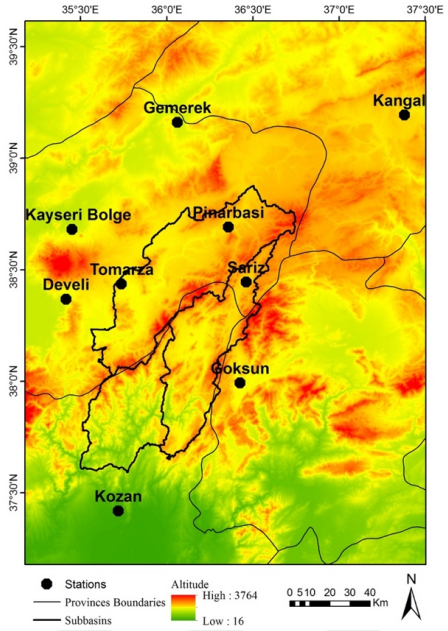


Figure 1.4. *Study area with the 9 stations*

1.2.4. Methodology

The obtained CruTS data of ten climate variables for the period 1901 -2014 have csv file format as one file for every month. MATLAB was used for reading files and converting them to ASCII after adding the required text to those files. The obtained ASCII files can be then easily used in any GIS software for further analysis.

The obtained files of the precipitation and temperature were then imported to R (i.e. Statistical Programming Language) for implementing the main objective of this study. CruTS data has a period from 1901 -2014 while the observed data 1973 -2000, therefore subsets of the CruTS data were extracted. The values of the CruTS ASCII files cells where the stations are located were extracted that 9 values obtained for each month which represented by one ASCII file.

The strength of the relationship between the values extracted from CruTS data and the observed data were evaluated by calculating the correlation coefficient which is given as:

$$r_{X,Y} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}}, \quad (1.22)$$

where x_i is the value of the CruTS data for month i, y_i is the observed value of the same month, $\bar{x}$ is the mean of the CruTS data, $\bar{y}$ is the mean of the observed data, and N the

number of data which represents months in this case. The result of the correlation for every station and for the two climate parameters; precipitation and temperature shown in Table 1.6.

In order to measure the error between the two datasets, Root Mean Squared Error (RMSE) was calculated for each station using:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(x_i - y_i)^2}{N}}. \quad (1.23)$$

The result of RMSE listed in Table 1.6.

The significance of the difference between the mean of the CruTS data μx and the mean of observed data μy was tested using a hypothesis test:

Ho: μx = μy

Ha: μx ≠ μy

The test conducted for each station using t student's test:

$$t = \frac{\bar{x} - \bar{y}}{SE}. \quad (1.24)$$

The standard error SE is calculated by:

$$SE = \sqrt{\frac{{s_x}^2}{N_x} + \frac{{s_y}^2}{N_y}}, \quad (1.25)$$

where s_x and s_y are the standard deviation of the CruTS and Observed data respectively (Helsel et al. 2002, Diez et al. 2015).

After calculating the t statistic, P – value calculated for every station in order to compare them with the significance level.

Some arguments saying that mean should not be used in case of the data are not normally distributed unless the number of observations large enough which is the case here in the precipitation data. Although the samples used here can be considered as sufficiently large, the median test was implemented.

In addition, Non-Parametric test named Mann-Whitney (known as Rank-Sum test and Wilcoxon Rank-Sum test) was used. This test is used for checking whether one group has higher or lower values than a second group. The two sided hypothesis of this test is:

Ho: Prob[x > y] = 0.5

Ha: Prob[x > y] ≠ 0.5

where x the first group data (i.e. CruTS data) and y the second group data (i.e. the observed data). In words, if the null hypothesis rejected it means that the data of two groups are different and vice versa. To save some space, the details of median and Mann-Whitney test are not mentioned here; for detailed description see (Hájek et al. 1999, Helsel et al. 2002, McCuen 2002, Hollander et al. 2014). The P-values of mean, median, and Mann-Whitney test are shown in Table 1.7.

1.2.5. Results and Discussion

The main objective of this study is to evaluate the CruTS 3.23 data by comparing them with the observed data of 9 stations located at the southwest of Turkey. To achieve this objective and to make the CruTS data (i.e. obtained with CSV file) usable in several software, especially GIS softwares, the files of the monthly 10 climate variables mentioned earlier were converted to ASCII. An example of the obtained files (i.e. the first month January of 1973 of precipitation) was added to ArcGIS and symbolized shown in Figure 1.5.

In order to ease the access, the ASCII files of the ten climate variables for the period 1901-2014 are available for download as a compressed file for every variable from (https://goo.gl/vFGNdp) or all the variables in one compressed file from (https://goo.gl/TXfxZm).

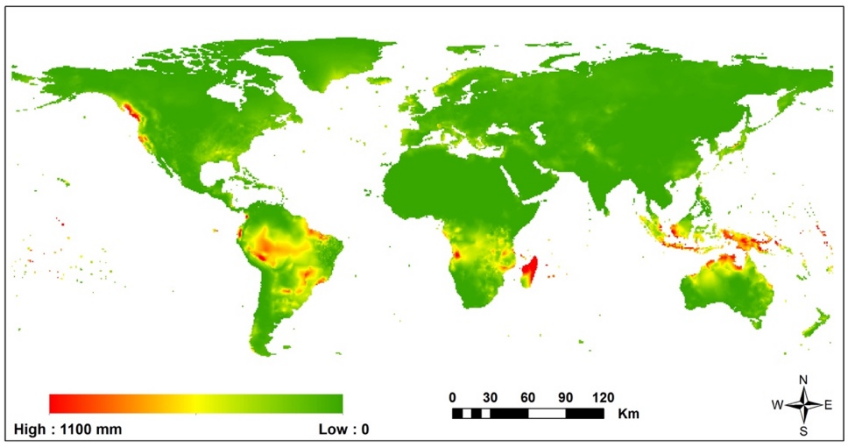


Figure 1.5. *Global Precipitation map of January 1973 obtained from CruTS 3.23 data*

Precipitation and Temperature variables for the period 1973-2000 were extracted from the CruTS data in order to have the same monthly data collected from the DMI for the 9 stations. By having subsetting the original data, consistent data are obtained that can be used for comparison.

Box plots were created for precipitation and temperature (i.e. Figure 1.6 and Figure 1.7 respectively) illustrating the data of every station from the two sources CruTS and observation side by side. Looking at Figure 1.6, Kozan station has the highest values of the total monthly precipitation while other stations are close to each other. In all stations, data coming from CruTS and observed data are close to each other in terms of the 0.25, 0.5 (i.e. median), and 0.75 quantiles with an exception of Goksun station in which the observed data has obvious higher values than CruTS data. Tomarza station shows a significant difference in the median between the two data sources. Based on the time series plot of the precipitation (i.e. Figure 1.8), there is an obvious mismatch between the two data sources; highest in Tomarza station, a bit lower in Goksun station while other stations matching better.

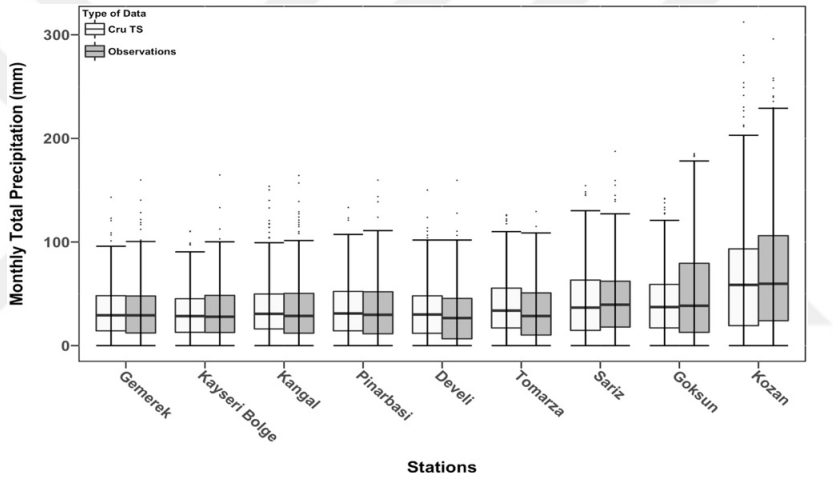


Figure 1.6. *CruTS and observed Precipitation data boxplot*

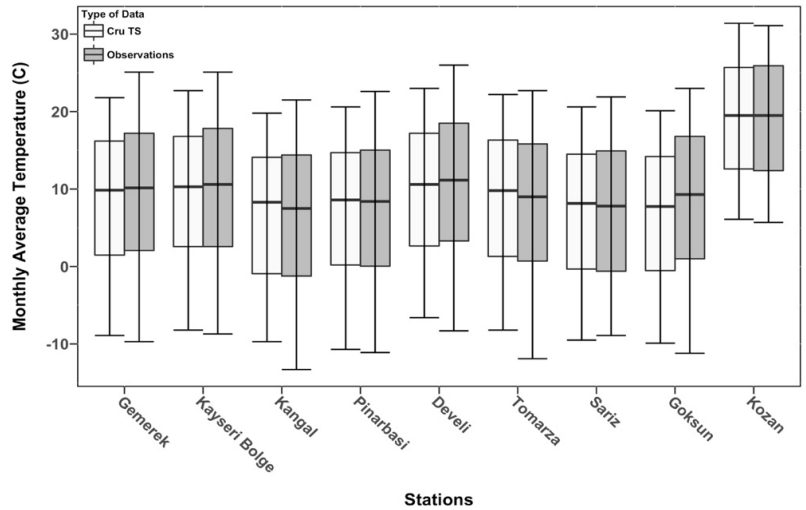


Figure 1.7. *CruTS and observed Temperature data boxplot*

Based on the temperature box plot Figure 1.7, Kozan station has, also, the highest temperature values in comparison with other stations which are not much different from each other. Similarly to the precipitation, Goksun station has the highest difference between CruTS and observed data where as other stations have slight close differences. According to the temperature time series plot Figure 1.9, the temperature in comparison with precipitation has changed, generally, smoothly. In addition, all stations show a high matching except Goksun station where there is a slight mismatch between CruTS and observed data. According to Figure 1.4 which shows the Digital Elevation Model (DEM), Kozan station is located at a low altitude in comparison with other station which are located in higher altitude and that could be the reason that this station has highest values in both precipitation and temperature.

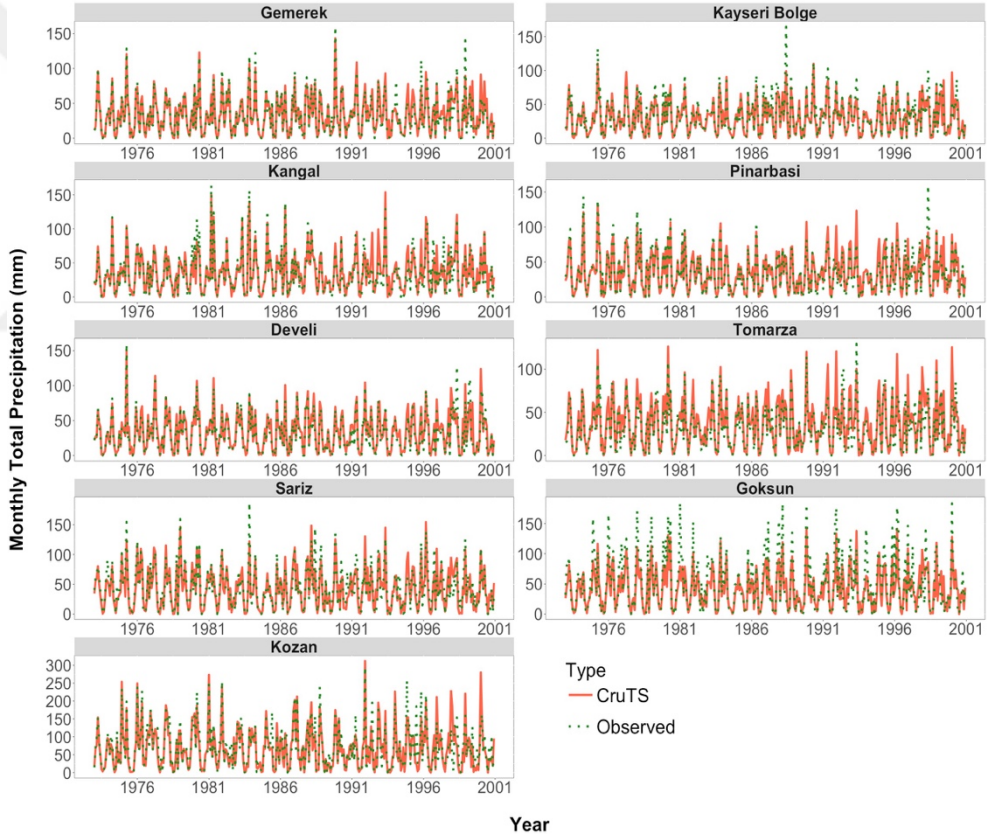


Figure 1.8. *Time series plot of the monthly total precipitation obtained from CruTS and Observed data for the period 1973 - 2000*

Observed precipitation and temperature data were used for the comparison with the CruTS 3.23 in the meaning of several statistical measurements. Correlations listed in Table 1.6 showed very high agreement for all stations in the temperature data ranges from 0.989 for Kangal station to 0.996 for Kayseri Bolge station. From the same table, RMSE

which measures the error between the two datasets, illustrates that Kozan station has the lowest while Goksun station has the highest error between the CruTS 3.23 and observed temperature data which agrees with what have been resulted from Figure 1.7.

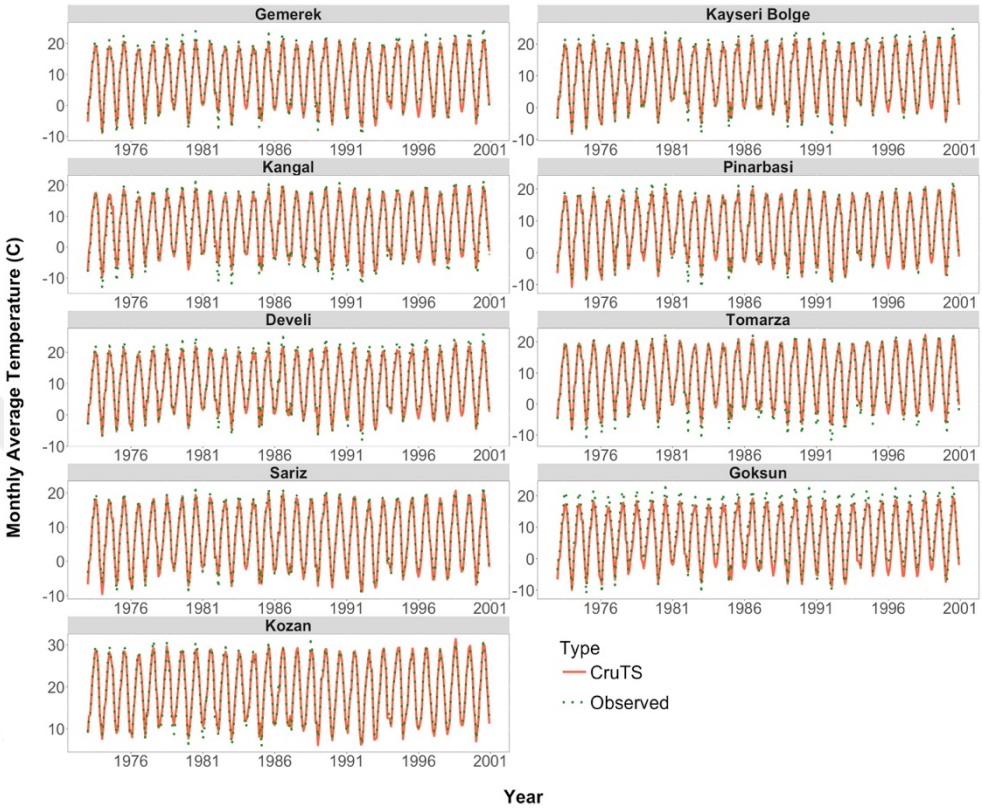


Figure 1.9. *Time series plot of the monthly average temperature obtained from CruTS and Observed data for the period 1973 - 2000*

In general, the agreement between the precipitation data was lower than temperature data. The precipitation correlations lowest value is 0.87 for Kozan station while the highest value is 0.942 for Gemerek station (Table 1.6). RMSE results for precipitation, also, showed higher error values. The highest error belongs to Kozan station while Gemerek station has the lowest. This lower agreement in the precipitation could be resulted from the nature of creating this data as these data created by interpolating a number of stations. The temperature data has more homogeneous distribution than precipitation around the interpolated area regardless the used method.

The P values of testing the difference between the means of two sources data for two variables and every station listed in Table 1.7. The result shows no significant values in most of the stations which means fail to reject the null hypothesis that the there is no difference between the means with several exceptions; Tomarza station has a significant

precipitation mean difference at the significance level $\alpha = 0.05$ and not significant at $\alpha = 0.01$ where P values is 0.018, Goksun station is significant in both precipitation and temperature at both significance levels with values 0.005 and 0.003 respectively.

Table 1.6. *The correlation and MSRE between CruTS and observed data for Precipitation and Temperature*

Station Name	Precipitation		Temperature	
	Correlation	RMSE(mm)	Correlation	RMSE (C)
Gemerek	0.947	9.007	0.992	1.383
Kayseri Bolge	0.916	10.866	0.996	1.178
Kangal	0.924	11.694	0.989	1.562
Pinarbasi	0.920	11.288	0.991	1.217
Develi	0.929	10.618	0.995	1.409
Tomarza	0.891	13.663	0.992	1.506
Sariz	0.885	16.320	0.994	0.948
Goksun	0.870	25.387	0.992	2.414
Kozan	0.870	30.326	0.992	0.930

Table 1.7. *The P value of the mean difference, median difference and Mann-Whitney tests*

Station Name	Mean		Median		Mann-Whitney test	
	Precipitation	Temperature	Precipitation	Temperature	Precipitation	Temperature
	P value					
Gemerek	0.582	0.324	0.977	0.693	0.343	0.222
Kayseri Bolge	0.397	0.452	0.862	0.762	0.793	0.338
Kangal	0.524	0.580	0.464	0.340	0.257	0.718
Pinarbasi	0.518	0.813	0.462	0.873	0.355	0.721
Develi	0.130	0.161	0.083	0.507	0.060	0.102
Tomarza	0.018*	0.265	0.029*	0.490	0.021*	0.365
Sariz	0.548	0.743	0.289	0.806	0.594	0.677
Goksun	0.005**	0.003**	0.823	0.139	0.291	0.001**
Kozan	0.614	0.949	0.911	1.000	0.618	0.980

* significant at $\alpha = 0.05$.

* *significant at $\alpha = 0.01$.

The median P value results listed in the same table indicate that two variables in all stations have no significant difference in the median with the exception of the precipitation variable in Tomarza station which is significant at the significance level $\alpha = 0.05$ and not significant at $\alpha = 0.01$ with the value 0.029. Surprisingly, temperature variable in Kozan station has a P value = 1 which indicates the median are same in the two data sources and that agrees with very high value of this station's mean result which is 0.949.

The Non – Parametric Mann-Whitney test p values are also listed in Table 1.7. These values indicate fail to reject the null hypothesis (i.e., no difference between the two

groups) as no significance difference for the two variables in all station except two stations i.e., Tomarza and Goksun stations. Precipitation variable is significant in Tomarza station at the significance levels $\alpha = 0.05$ only which agrees with the two mean and median tests. Goksun station has a significant value in the temperature variable at the two significance levels $\alpha = 0.01$ and 0.05 which matches the mean test.

In general, CruTS 3.23 data has a tight agreement with the observed data. Most of the used 9 stations has; high correlations; low errors; and no significance in the differences tests. The result agrees with the result of (Harris et al. 2014) who found close-fitting matches in a sub-continental scale during their updating of CruTS 3.10 data. As an exception, Goksun station has a statistical significance in the difference in several test which agrees with the low correlation and high error in comparison with other stations. However, although Kozan station has the lowest correlation (sharing the same value with Goksun) and highest error in the precipitation variable, it has no significance in any of the tests.

1.2.6. Conclusion

This study aims to evaluate the CruTS 3.23 data for only precipitation and temperature variables in a local scale. To have suitable format of the data, files of monthly data of ten climate variables; precipitation, mean temperature, diurnal temperature range, minimum and maximum temperature, vapour pressure, cloud cover, rain days, frost days and potential evapotranspiration for the period January 1901 to December 2014 was converted to GIS-based ASCII formats.

The local scale comparison results indicate that the two sources of data have a tight agreement in both precipitation and temperature, however, the agreement of temperature variable is higher.

The three tests (i.e., Mean difference, Median difference, and Mann-Whitney) results show no significance difference between the two data sets in all stations with an exception of two stations; Tomarza and Goksun stations. Tomarza station has significant difference only in the precipitation variable in the three tests while Goksun station has significant difference only in the mean test of the two variables. Kozan station shows the opposite of these two stations as, for example, it has a difference in the median reaches to zero. This difference among stations originated from the nature of the CruTS data that the different versions were created by interpolating a number of station around the world.

CHAPTER TWO

2. DATA EXPLORATION

2.1. Long-Term Spatiotemporal Trend Analysis of Precipitation and Temperature Over Turkey

2.1.1. Abstract

Monthly precipitation and temperature data for Turkey were obtained from the Climate Research Unit Time Series (CRU TS V3.23) dataset (1901-2014) and analyzed. The records of 81 grids, each located in a province, were converted into annual and seasonal data. The Mann-Kendal test was used to determine the trend significance. The Theil-Sen's slope estimator was used for estimating the magnitude of the trend. The Pettitt-Mann-Whitney test was used on the annual data series to identify the "most probable change year" for precipitation and temperature. The spatial pattern of the trend was analyzed using GIS environment. The results showed that the annual areal precipitation of entire Turkey is not significantly increasing with a magnitude of 0.11 mm/year, but the annual areal temperature is increasing significantly with a magnitude of 0.88 C^o/Century. The grids analysis indicated that annual precipitation is decreasing in the South-Eastern Anatolia and Mediterranean regions, and increasing in the Marmara and Black Sea regions, while other regions have no specific pattern. Though the trend is present in most of the grids, but the insignificance is dominant. Annual precipitation had no obvious most probable change year, with most frequent year 1969. Seasonal precipitation has increasing and decreasing trends in the same regions that showed such trends in annual precipitation. Annual temperature had a significant, increasing trend in all regions. 1993 was found as the most probable change year. Seasonal temperature showed an increasing trend in all regions. The highest increasing trend for temperature was found in the summer.

2.1.2. Introduction

Precipitation and temperature are the most important climatic variables in the hydrology/climate fields, and in water resource management, due to their critical effect on the temporal and spatial patterns of water availability. Precipitation is an essential factor in the rainfall-runoff relationship, mitigation measures, and flood and drought assessments, while temperature is a critical factor in water availability, transpiration, and evaporation. Changes in precipitation and temperature over a long period are vital for

water resources planners and climatologists to understand the climate change impacts(Luís et al. 2000, Taxak et al. 2014, Jones et al. 2016).

The identification, verification, and quantification of trends in precipitation and temperature are important challenges due to the significant changes in the global climate in the last ten decades. According to (IPCC 2001, Xu et al. 2005), mean annual precipitation has increased over the last century by 2% for latitudes between 0-55^{o} in the southern hemisphere, and by 7%-12% for the latitudes between 30-85^{o} in the northern hemisphere. The mean annual air temperature of both land surface and the ocean, according to (IPCC 2013), increased globally between 1880-2012 by 0.65-1.06 °C. Global and continental historical climate observations have, however, been identified as being less useful when planning at local and regional scales (Brekke et al. 2009, Raucher 2010, Taxak et al. 2014). Analysis of local and regional historical trends is therefore important.

Several studies on trend analysis of precipitation and temperature have been conducted in different regions and countries around the world. Some of these studies only used the trends in precipitation/rainfall, for example: (Buffoni et al. 1999, De Luís et al. 2000, Yue et al. 2003, Gemmer et al. 2004, Marengo 2004, Feidas et al. 2007, Pujol et al. 2007, Kampata et al. 2008, de la Casa et al. 2010, Karpouzos et al. 2010, Longobardi et al. 2010, Oguntunde et al. 2011, Tabari et al. 2011, Villarini 2012, Duhan et al. 2013, Haktanir et al. 2013, Mahsa et al. 2016) while some studied only the temperature trends, for example: (Easterling et al. 1997, Smadi 2006, Tabari et al. 2011, Saboohi et al. 2012, Robaa et al. 2013, Mamara et al. 2016, Kisi et al. 2017). Some studies have, however, investigated the trends in both precipitation and temperature, for example: (Feizi et al. 2014, Keggenhoff et al. 2014, Subash et al. 2014, Mondal et al. 2015, Chattopadhyay et al. 2016).

Most of these studies used observations collected from stations distributed around the studied area. Several studies have used interpolated data for large regions around the globe, such as CRU TS data. For example, (Taxak et al. 2014) used CRU TS 3.21, which covers 1901-2012, to study the trend and homogeneity of precipitation over the Wainganga basin in central India. (Jones et al. 2016) also used the same data set to study the trend in precipitation and temperature across the Caribbean.

Precipitation trends over Turkey have been a matter of concern for many researchers. For example, (Türkeş 1996) used total monthly rainfall recorded in 91

stations between 1930-1993 to study the temporal and spatial trends in annual rainfall data over Turkey. He identified a decreasing trend in the region-based areal rainfall for the whole of turkey, which was most apparent in the Black Sea and Mediterranean regions, but this trend was not identified as significant. Partal and Kahya (2006) studied the trend in Turkish precipitation using the observations of 96 stations around Turkey covering the period 1929-1993. They identified a decrease in the annual mean precipitation in the southern and eastern parts of Turkey, and the Black Sea coast. Türkeş et al. (2009) investigated the spatiotemporal variability of precipitation, using 97 stations' records covering the period of 1930-2002, and found an increasing trend in the spring, summer and autumn, and a decreasing trend in winter, which was very apparent in the Mediterranean regions. Yavuz and Erdoğan (2012) identified a positive trend in precipitation in March, April, and October, by studying observations from over 32 years, extending from 1975-2009, that were collected from 120 stations over the entire Turkey.

Trends in temperature have not been studied as intensively as precipitation. For example, (Türkeş et al. 1996) studied trends and changes in the mean seasonal daily maximum and minimum temperatures, and diurnal temperature, using the observations of 59 stations throughout Turkey over the 1930-1993 period. They found an increasing trend in mean minimum temperature in most of the used stations in all the seasons except winter. An insignificant decreasing trend in the mean maximum temperature in all seasons except spring was obtained. In the same study, a significant increasing trend in the spring minimum temperature over the Aegean, Black Sea, and the South-Eastern Anatolia regions was also identified. Kadioğlu (1997) used 18 stations' observations of the mean seasonal, and annual maximum and minimum temperatures, and identified a statistically insignificant increasing trend in the mean annual temperature for the period 1939-1989 and a decreasing trend for 1955-1989 over the whole Turkey. Türkeş et al. (2002) studied the trend in temperature over Turkey using observations of the mean, minimum, and maximum temperatures collected from 70 stations, covering the period of 1929-1999. They found an increasing trend in the annual, winter and spring mean temperature over the southern region of Turkey, whereas a decreasing trend over the northern and inner regions for the summer and autumn mean temperature.

Türkeş and Sümer (2004) investigated the spatial and the temporal trend of diurnal temperature ranges (DTRs) and the role of the minimum and maximum temperature in this trend for the same period and number of stations used in (Türkeş et al. 2002). They

found a decrease in the autumn and summer DTRs which is higher than the decrease in the winter and spring. They concluded that the trend in the minimum and maximum temperature is the leading factor of the decrease in the DTRs and the increase in the night temperature of spring and summer.

All the trend studies of two climate variables, precipitation and temperature, over Turkey have been conducted with periods not longer than 70 years. This study aims to investigate variations in spatial and temporal trends of these two climate variables for 114 years using the latest version of CRU TS data (3.23), covering the period of 1901-2014.

2.1.3. Study area

Turkey lies between 36-42°N and 26-45°E and is in the northern subtropical climate zone of the Earth. The climate of Turkey differs from region to region due to the existence of several mountains running along the coasts, even though it is in the large Mediterranean geographical area (Sensoy et al. 2008). The annual average precipitation of Turkey is about 574mm for the period 1981-2010. The Mediterranean and Aegean regions have rainy winters with annual precipitation varies from 580 to 1300 mm. The Black Sea region receives the greatest amount of precipitation and it is the only region which receives precipitation throughout the year. This region receives annually 2200 mm in the eastern part. Istanbul and Marmara region have a moderate climate with average temperature 4^{o}C in winter and 27^{o}C in summer. Black sea region is wet and humid with 7^{o}C in winter and 23^{o}C in summer. A long winter is witnessed in the Eastern Anatolia region as the snow remains on the ground for about 6 months extending from November to April. The average temperature is -13^{o}C in winter and 17 ^{o}C in summer. The South-Eastern Anatolia region has hot and dry summer with an average temperature over 30 ^{o}C while spring and autumn are mild in general. The Anatolia plateau has a steppe climate with average temperature -2^{o}C in winter and 23 ^{o}C in summer. The Mediterranean region has a mild climate with average temperature 9^{o}C in winter and 29 ^{o}C in summer (Sensoy et al. 2016). Figure 2.1 shows the map of Turkey with the regions divisions and the location of the grids used in this study.

2.1.4. Data Used

CRU TS is the abbreviation of the Climate Research Unit Time Series, named after the Climate Research Unit, University of East Anglia, Norwich, UK, who produce these

data sets. CRU TS 3.23 is the version released in September 2015 (Harris et al. 2014, Cru 2016), which contain monthly values of ten climate variables: precipitation (pre), mean temperature (tmp), diurnal temperature range (dtr), wet-day frequency (wet), vapor pressure (vap), cloud cover (cld), maximum temperatures (tmx), minimum temperatures (tmn), frost day frequency (frs), and potential evapotranspiration (pet) (Jones et al. 2008). The CRU TS 3.23 dataset covers the earth, with the exception of the Antarctica region, with 0.5^{o} latitude-longitude grid cells for the period of 1901-2014. Only the precipitation and mean temperature for Turkey were used in this study. The data were obtained from the Centre for Environmental Data Archival of the British Atmospheric Data Center (BADC) (http://www.badc.nerc.ac.uk), and CEDA OGC WEB SERVICES/ Web Processing Service (http://wps-web1.ceda.ac.uk/).

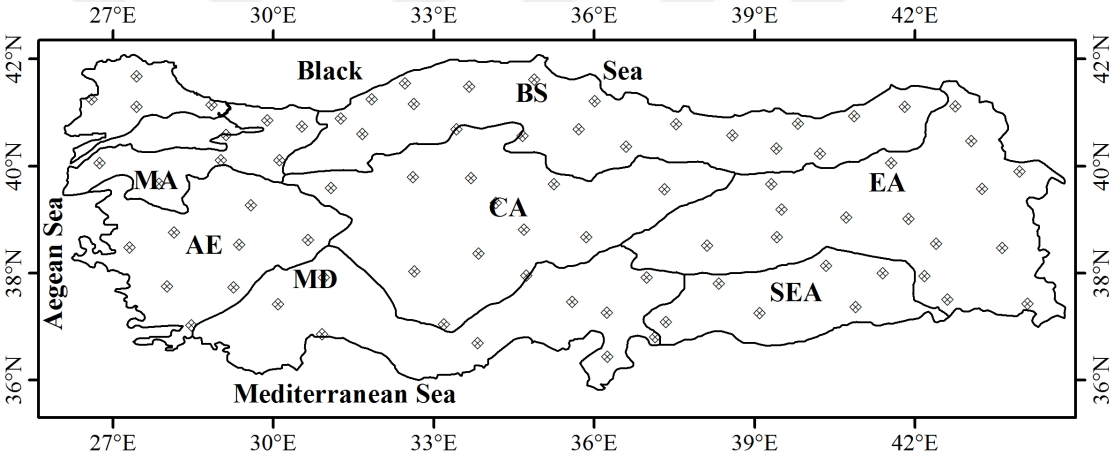


Figure 2.1. *Regions of Turkey and the distribution of the Grids. Regions; BS: Black Sea, EA: Eastern Anatolia, SEA: South Eastern Anatolia, CA: Central Anatolia, MD: Mediterranean, AE: Aegean, MA: Marmara*

For applying this data in applications, validation in different scales must be examined. Harris et al. (2014) validated the CruTS 3.10 data sets in a sub-continental scale using two different datasets: precipitation datasets for the period 1901-2009 which created by Global Precipitation Climatology Center (GPCC), and the temperature dataset for the period 1901-2008 which created by University of Delaware (UDEL). They found a very tight agreement between the UDEL and the CruTS temperature datasets. A very high agreement is present also between the GPCC and CruTS precipitation dataset but not as tight as temperature data. Locally, the evaluation and quality assessment of the CruTS 3.23 dataset were done by (Hadi et al. 2017). They used precipitation and

temperature observed data of 9 stations in Turkey and approved that these datasets are reliable for being used in applications due to the high agreement with the observed data.

2.1.5. Methodology

Monthly data for precipitation and temperature were prepared for 81 grids; each grid represents a province of Turkey. These grids were chosen in the center of the province as each grid has dimension of 0.5^{o} latitude-longitude. Those data were converted to total annual and seasonal precipitation, and average annual and seasonal temperature. Seasons are defined as follows - Winter: December, January, and February, Spring: March, April, and May, Summer: June, July, August, Autumn: September, October, and November.

Several tests were then applied to the series: Mann-Kendall test (MK) was used to detect the presence and significance of trends, Theil-Sen's slope estimator was used to estimate the magnitude of any trends, and Pettitt's test was applied on the annual data series to estimate the most probable change year. The simple linear regression and LOcally WEighted Scatterplot Smoothing (LOWESS) were used in the annual average areal precipitation and temperature time series for a graphical analysis.

Spatial variation is important for large countries, such as Turkey, as it is greatly affected by regional and local variation. A spatial pattern of the trends over Turkey was prepared using spatial interpolations through the GIS environments, using the magnitude of the trend for each grid. The regions of Turkey were not gathered for calculations, rather the calculations were implemented for each grid separately and then the results and graphs were organized according to regions, so they could be interpreted regionally.

2.1.5.1. Mann-Kendall Test

The Mann-Kendall test was first created by (Mann 1945) to be used for testing the significance of trends by testing the significance of (Kendall 1938)'s tau. It has since become the most widely used non-parametric test for several applications and specially for meteorological applications.

The test statistic S is calculated by:

$$S = \sum_{i<j} y_{ij} \,, \tag{2.1}$$

where

$$y_{ij} = sign(x_j - x_i) = sign\left(R_j - R_i\right) = \begin{Bmatrix} 1 & x_j < x_i \\ 0 & x_j = x_i \\ -1 & x_j > x_i \end{Bmatrix}, \quad (2.2)$$

where R_i and R_j are the ranks of the x_i and x_j observations of the time series.

According to (Kendall 1975), the mean and variance of S are:

$$E(S) = 0, \quad (2.3)$$

$$V_0(S) = n(n-1)(2n+5)/18\ , \quad (2.4)$$

where n is the number of observations.

The standardized Mann-Kendall variable Z_{MK} is calculated by:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{V_0(S)}} & S > 0 \\ 0 & S = 0\ . \\ \dfrac{S+1}{\sqrt{V_0(S)}} & S < 0 \end{cases} \quad (2.5)$$

The significance of the test is determined by comparing the standardized variable Z_{MK} with the standard normal distribution quantile at the desired significance level.

2.1.5.2. Theil-Sen's Slope Estimator

Theil-Sen's Slope estimator (Theil 1950, Sen 1968), is a non-parametric test for fitting a robust line to a set of points. The test identifies the median of the lines' slopes between every pair of points in the sample. The slopes are estimated by:

$$Q_i = \frac{y_j - y_k}{t_j - t_k} \qquad \text{for } i = 1, 2, 3, \ldots, \text{N}\ , \quad (2.6)$$

where y_j and y_k are observations at times t_j and t_k for all $t_j > t_k$, and, in the case of having only one data point in every time point:

$$N = n(n-1)/2\ , \quad (2.7)$$

where n is the number of time points. The N values of the slopes Q_i are ranked by $Q_1 \leq Q_2 \leq Q_3 \ldots \leq Q_N$ for the Sen's estimator $\hat{\beta}_1$, estimated using:

$$\hat{\beta}_1 = \begin{cases} Q_{\left[\frac{N+1}{2}\right]} & if\ N\ is\ odd \\ \dfrac{1}{2}\left(Q_{\left[\frac{N}{2}\right]} + Q_{\left[\frac{N+2}{2}\right]}\right) & if\ N\ is\ even\ . \end{cases} \quad (2.8)$$

2.1.5.3. Pettitt-Mann-Whitney Test

The Pettitt test (Pettitt 1979) is a non-parametric test that focuses on the case of a two-tailed test, but one-side tailed tests are also possible. The null hypothesis (H_0) of a Pettitt test is that the time series follows one or more distributions while having the same location parameters, while the alternative hypothesis (H_a) for a two-tailed test is that a change point occurs, and the alternative hypothesis (H_a) for a one-side test is that a time (t) exists from which the variables' location parameters are reduced or augmented.

Considering T, which is the length of the time series, and t, which represents the time points of the most probable change year, and a single time series divided into two samples $x_1 \dots x_t$ and $x_{t+1} \dots x_T$, using a Mann-Whitney test for two samples:

$$D_{ij} = sign\left(x_j - x_i\right), \tag{2.9}$$

where:

$$Sign\,(x) = \begin{Bmatrix} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{Bmatrix}. \tag{2.10}$$

Then,

$$U_{t,T} = \sum_{i=1}^{t} \sum_{j=t+1}^{T} D_{ij}\,. \tag{2.11}$$

The value of $U_{t,T}$ is same as the value of Mann-Whitney test for testing two groups $x_1 \dots x_t$ and $x_{t+1} \dots x_T$ that come from the same distribution. The value of $U_{t,T}$ is calculated for every t where $1 \leq t < T$. The point in time with the max value of $U_{t,T}$ is considered to be the most probable change point.

$$K_T = \max\left(U_{t,T}\right) \qquad \text{for } 1 \leq t < T. \tag{2.12}$$

The probability of the change point being the maximum value of $U_{t,T}$ is calculated by:

$$p = 1 - \exp\left(\frac{-6\,K_T^2}{T^3 + T^2}\right). \tag{2.13}$$

2.1.6. Results and Discussion

2.1.6.1. Exploratory Data Analysis (EDA)

The annual and seasonal total precipitation and average temperature for the entirety Turkey are under consideration in this study. The highest annual areal precipitation for the whole of Turkey is 771 mm in 1940, while the lowest value is 471 mm in 1932.

Analysis of each grid showed that the lowest annual precipitation is 124 mm in Karaman in 1932, and the highest is 1677 mm in Artvin in 1988. The statistics of the annual precipitation of each grid implied that the mean varies between 304 mm in Konya and 1145 mm in Artvin, while the standard deviation varies between 63 mm and 212 mm in Tokat and Muğla. The coefficient of variation is 11.6% in Giresun which is the lowest value, and 24.7% in Şanlıurfa which is the highest value. The skewness values vary greatly, from several negative values to mostly positive values. The lowest value is -0.28 in Ordu, while the highest value is 1.2 in Tunceli. The highest seasonal precipitation is 1054 mm in the winter of 1938 in Muğla, while the lowest is 0 in Kilis in the summers of 1944 and 1985, and in Şanlıurfa in the summers of 1955 and 1966.

The analysis of temperature for the whole of Turkey indicated that the highest annual areal temperature was 13.3 °C in 2010, while the lowest was 9.7 °C in 1911. Analysis of each grid showed that the lowest annual temperature is 0.3 °C in Kars in 1992, while the highest is 20 °C in Adana in 2010. A grid-based statistical analysis of the annual temperature showed that the mean varies from 2.6 °C in Kars to 18 °C in Adana. The standard deviation varies from 0.59 °C to 0.89 °C in Muğla and Iğdır, respectively. The coefficient of variation varies greatly that the lowest value is 3.6% at Adana where for most of the year the temperature is considered to be high. The highest value is 32% in Kars, where the temperature varies greatly throughout the year. The skewness was negative in some grids, but in most of them is positive and varies from -0.25 to 0.62, with an average value of 0.19. The highest seasonal temperature is 31 C^o in the summer of 2010 in Şanlıurfa, while the lowest is -14.5 C^o in the winters of 1933 and 1977 in Ağri and Kars, respectively.

Figure 2.2 shows the spatial distribution of the mean annual precipitation (a) and temperature (b). According to Figure 2.2 (a), the mean annual precipitation varies greatly between regions of Turkey. Central Anatolia and a large part of Eastern Anatolia receive the smallest amount of rainfall, the south and west parts of the Mediterranean regions, the coastal part of the Aegean region, and the Black Sea region receive the highest amount, while the Marmara region receives a modest amount of annual precipitation. In general, the coastal boundaries of Turkey receive the highest volume of precipitation. Temperature shows a different pattern of distribution. Eastern Anatolia experiences the lowest while South-Eastern Anatolia experiences the highest values. The western part of the Mediterranean, Aegean, and Marmara regions experience modest and similar

temperature values to each other. Central Anatolia experiences a modest to low temperature, sharing that with a large part of the Black Sea region.

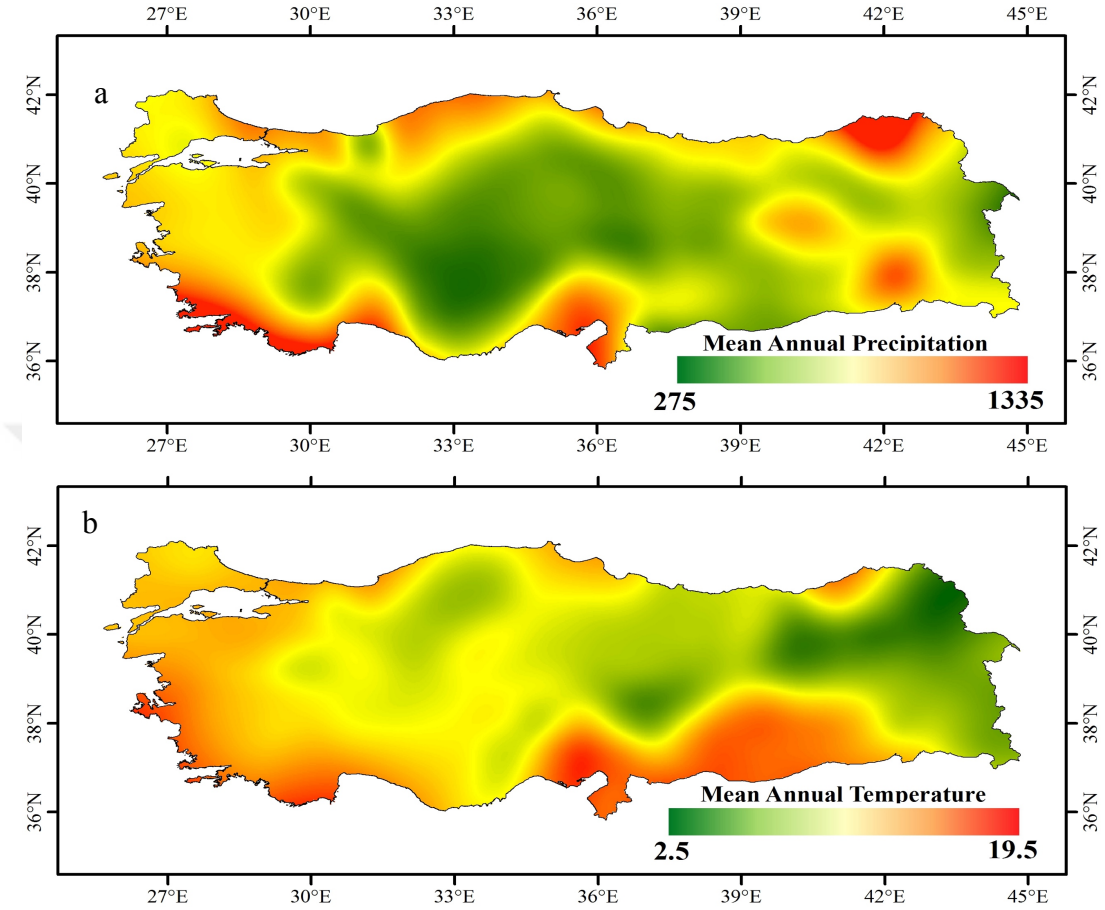


Figure 2.2. *a) The spatial distribution of the mean of annual Precipitation. b) The spatial distribution of the mean of annual Temperature*

2.1.6.2. Annual trend

2.1.6.2.1. Precipitation

The average of the total annual precipitation for the whole of Turkey is plotted as a time series (Figure 2.3), along with the simple linear and LOWESS regression lines. The linear line indicates an increasing trend, which is consistent with the Theil-Sen's Slope test value of 0.11 mm/year, however the trend was determined to be insignificant by the Man-Kendall test. This insignificance is also found by (Türkeş 1996, Unal et al. 2012), who studied the period 1930-1993 and 1961–2008 respectively. On the contrast, these two studies identified a decreasing trend for the whole of Turkey, which is not consistent with the result here. The LOWESS regression line shows an increasing trend until the end of the 1950s, where it starts to decrease until the early years of the 1990s, where it starts

increasing again. The most probable change year identified in this study, using the Pettitt test, is 1934.

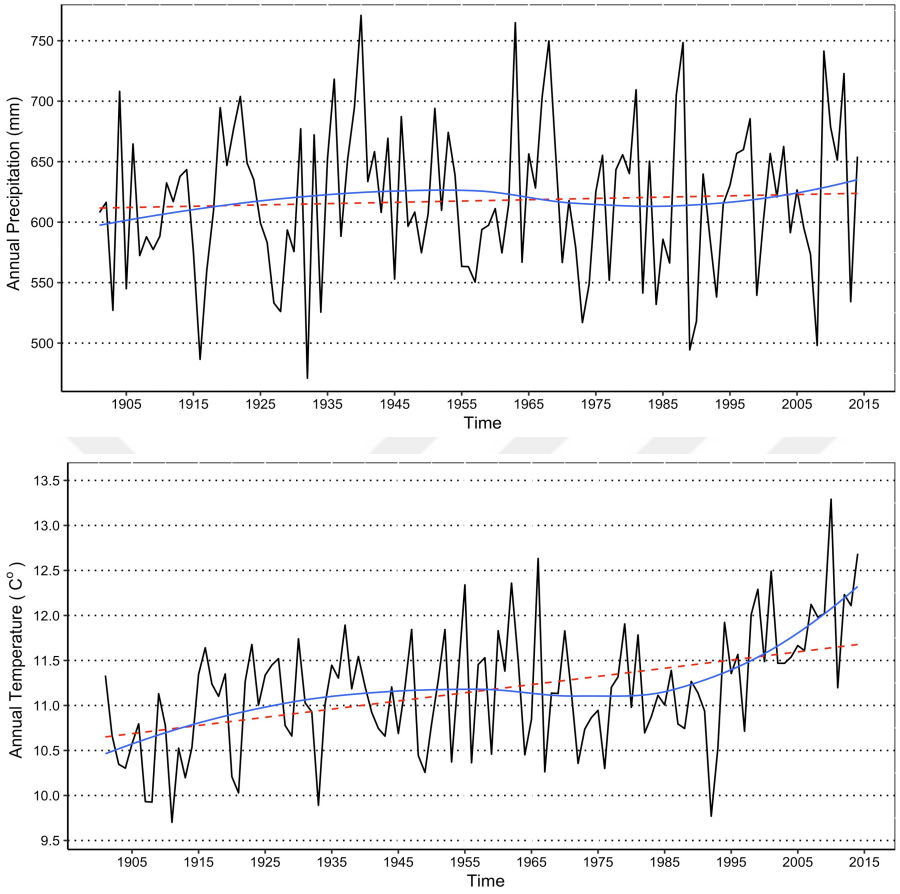


Figure 2.3. *The time series of Turkey's mean of the total annual precipitation (Top) and the average annual temperature (bottom). Solid Line: LOWESS regression, Dashed Line: Simple Linear Regression. For Precipitation(Temperature), Insignificant (Significant) trend with 0.11 mm/year (0.88 C°/Century) Theil and Sen's Slope test value and change year 1934 (1993)*

According to Table 2.1 which shows the results of grids' analysis, no year dominated the most probable change year of the grids but the most frequent year is 1969, which repeated in 11 grids. The trend in precipitation varies greatly from grid to grid, it has negative values with a lowest value of -0.77 mm/year in the Hakkâri, and positive values, with a highest value of 0.94 mm/year in Kocaeli.

Table 2.1. *The results of the Pettit test and trend value (mm/year) with significance for Precipitation. Ann: Annual, DJF: December, January, and February, MAM: March, April, and May, JJA: June, July, August, SON: September, October, and November*

Stations	Year	Ann	DJF	MAM	JJA	SON	Stations	Year	Ann	DJF	MAM	JJA	SON
Adana	1924	-0.12	0.03	0.02	0.10	0.17	K.Maraş	1942	0.04	-0.03	0.01	0.02	0.17
Adiyaman	1969	-0.35	-0.08	-0.27	0.01	0.10	Karabük	1961	**0.69**	0.13	0.20	0.22	0.08
Afyon	1937	0.00	0.11	0.04	0.02	0.00	Karaman	1954	**-0.34**	***-0.25***	-0.01	0.04	-0.04
Agri	1951	-0.08	0.09	0.14	-0.17	0.01	Kars	2000	0.11	0.11	0.17	-0.16	0.12
Aksaray	1954	-0.16	-0.08	0.01	0.00	0.09	Kastamonu	1959	***0.82***	0.20	0.25	**0.34**	0.06
Amasya	1964	0.09	-0.14	0.15	**0.17**	0.04	Kayseri	1986	0.25	0.00	***0.2***	0.09	**0.16**
Ankara	1961	0.11	-0.10	0.16	0.12	0.00	Kilis	1981	***-0.61***	-0.26	***-0.34***	-0.01	0.02
Antalya	1931	-0.59	**-0.78**	0.10	**0.08**	-0.29	Kirkkale	1925	-0.14	-0.14	0.08	0.05	0.06
Ardahan	1935	0.39	0.17	0.22	0.12	0.01	Kirklareli	1954	***0.69***	0.06	0.18	0.10	**0.34**
Artvin	1918	0.02	0.12	-0.05	-0.08	0.31	Kirşehir	1925	-0.13	-0.12	0.02	0.03	0.12
Aydin	1921	0.10	-0.32	0.18	0.05	0.13	Kocaeli	1934	***0.94***	0.25	***0.28***	0.21	0.19
Balikesir	1921	0.18	-0.18	0.18	0.03	0.25	Konya	1952	-0.23	-0.10	-0.07	0.00	0.02
Bartın	1959	***0.79****	0.27	0.17	0.29	0.10	Kütahya	1934	0.36	0.02	0.17	**0.18**	0.07
Batman	1969	**-0.71**	-0.35	-0.28	0.00	-0.21	Malatya	1962	0.17	0.00	0.04	0.05	0.17
Bayburt	1951	0.08	0.01	0.03	-0.02	0.06	Manisa	1921	0.13	-0.08	0.19	0.03	0.11
Bilecik	1937	**0.43**	0.13	0.18	0.14	0.10	Mardin	1969	**-0.64**	-0.31	-0.35	0.01	-0.06
Bingöl	1946	-0.29	-0.26	-0.07	-0.02	-0.08	Mersin	1954	***-0.74***	***-0.66***	-0.04	0.09	-0.05
Bitlis	1953	-0.22	-0.10	0.05	-0.06	-0.04	Muğla	1946	-0.35	***-0.94***	0.21	0.04	0.06
Bolu	1958	***0.64***	0.13	0.22	**0.21**	0.12	Muş	1946	-0.22	-0.01	0.05	**-0.12**	-0.02
Burdur	1969	-0.35	***-0.44***	0.04	0.10	-0.05	Nevşehir	1994	0.17	0.01	0.18	0.07	0.10
Bursa	1934	***0.62***	0.03	0.21	0.14	0.31	Niğde	1937	0.06	-0.01	0.05	0.03	0.17
Çanakkale	1921	0.36	-0.06	0.13	0.06	0.03	Ordu	1929	***0.47***	-0.05	**0.24**	0.10	***0.31***
Çankiri	1950	0.20	-0.07	0.17	0.16	0.05	Osmaniye	1942	-0.49	-0.08	-0.33	-0.03	0.15
Çorum	1925	0.01	-0.17	0.14	0.18	0.04	Rize	1925	0.14	0.09	-0.07	0.06	0.16
Denizli	1969	-0.14	-0.34	0.16	0.06	0.00	Sakarya	1948	***0.73***	0.13	0.24	0.23	0.22
Diyarbakir	1988	-0.15	-0.12	-0.21	0.03	0.04	Samsun	1932	0.17	-0.14	0.14	***0.28***	0.00
Düzce	1947	**0.41**	0.13	0.08	0.14	0.16	Sanliurfa	1969	-0.53	-0.22	**-0.32**	0.01	-0.02
Edirne	1950	0.45	-0.02	0.09	0.12	0.08	Siirt	1954	-0.37	-0.21	-0.09	-0.02	-0.06
Elazığ	1910	0.01	-0.12	-0.04	0.03	0.15	Sinop	1982	0.33	-0.09	0.14	***0.27***	0.12
Erzincan	1962	0.21	-0.02	0.21	-0.01	0.16	Şirnak	1969	**-0.73**	-0.13	-0.35	-0.04	-0.26
Erzurum	1946	-0.20	-0.05	0.04	-0.17	0.02	Sivas	1984	**0.38**	0.03	**0.26**	0.04	**0.19**
Eskişehir	1956	0.23	0.02	0.11	0.09	0.07	Tekirdağ	1951	0.47	-0.02	0.06	0.09	0.21
Gaziantep	1954	-0.51	-0.20	-0.26	-0.01	0.05	Tokat	1976	**0.47**	-0.06	***0.28***	**0.13**	0.17
Giresun	1976	**0.4**	-0.01	0.11	0.07	***0.25***	Trabzon	1984	**0.43**	0.06	0.09	0.10	0.21
Gümüşhane	1976	**0.42**	0.05	0.13	0.09	**0.20**	Tunceli	1969	0.02	-0.06	-0.02	0.01	0.21
Hakkari	1969	***-0.77***	-0.05	**-0.36**	-0.09	-0.23	Uşak	1928	0.13	-0.17	0.24	0.13	0.09
Hatay	1954	-0.76	-0.29	-0.41	-0.05	0.05	Van	1969	-0.24	-0.03	-0.08	-0.12	0.04
Iğdır	2001	0.05	0.10	0.14	-0.08	-0.02	Yalova	1934	**0.72**	0.13	0.23	0.17	0.15
Isparta	1969	-0.37	-0.20	-0.08	0.04	-0.01	Yozgat	1925	0.08	-0.11	**0.20**	0.10	0.10
Istanbul	1973	***0.92***	0.21	0.20	0.22	0.13	Zinguldak	1946	**0.73**	0.21	0.16	0.22	0.24
Izmir	1921	0.45	-0.09	0.19	-0.02	0.23							

* The Bold values are significant at significance level $\alpha = 0.1$ and the bold italic underlined values are significant at significance level $\alpha = 0.05$.

All of the grids in South-Eastern Anatolia, and the Mediterranean have a negative trend, except for one grid. In contrast, all the grids in the Black Sea and Marmara regions have a positive trend. The Central Anatolia, Eastern Anatolia, and Aegean regions have no constant trend, with several grids showing a decreasing trend and others showing an increasing trend. There was no obvious pattern present in the magnitudes of the trend between regions; however, the Marmara region has the highest increasing trends while South-Eastern Anatolia has the highest decreasing trends (Figure 2.4).

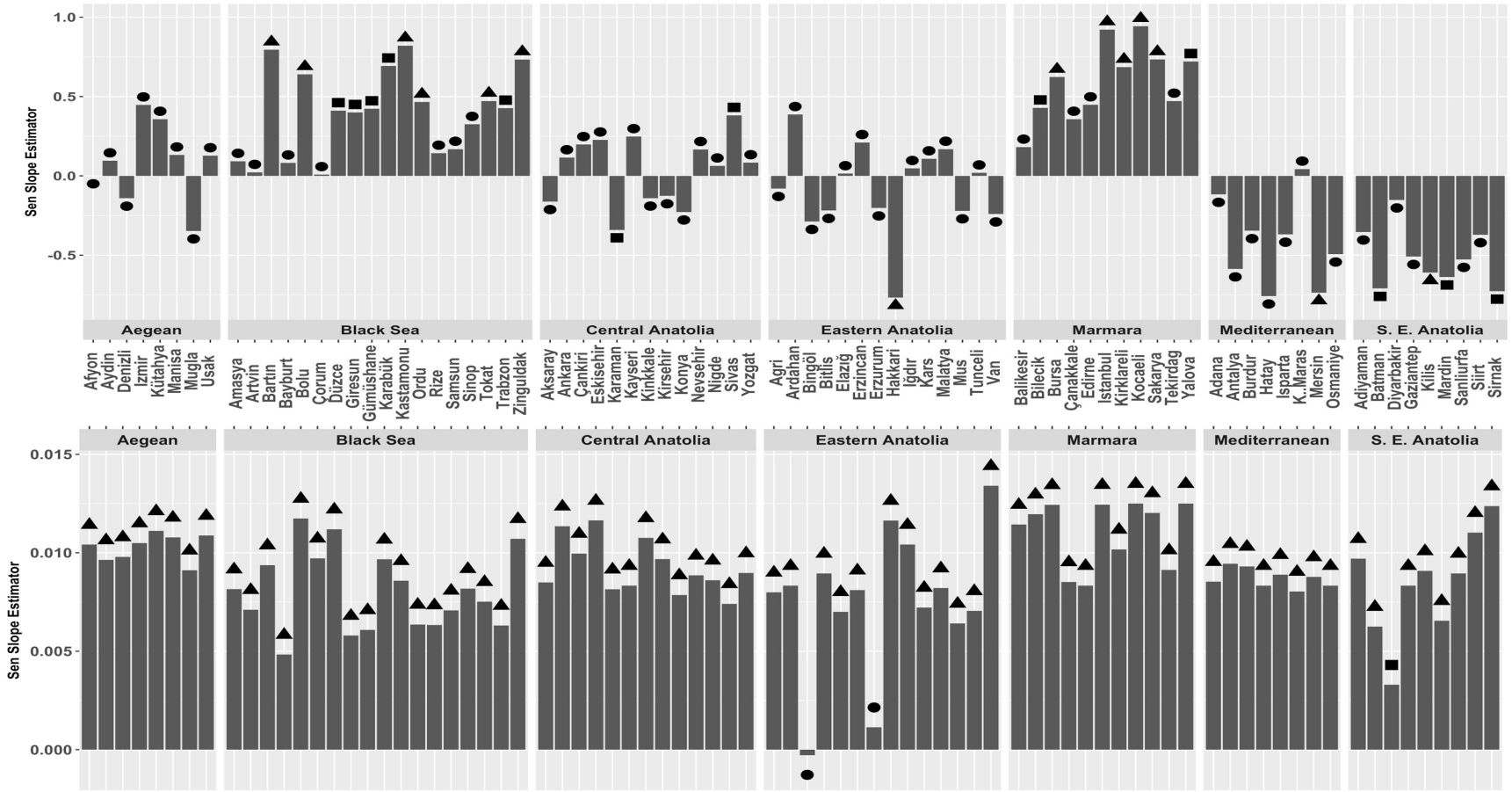


Figure 2.4. *Annual Theil and Sen's Slope test values Top) Precipitation (mm/year), Bottom) Temperature (Co/year). Circle: Not significant, Square: significant at 0.1, Triangle: Significant at 0.05*

Figure 2.5 shows the spatial pattern of trends, indicating a decreasing trend in the south of Turkey, with an area wider in the east than the west. The north has a predominantly increasing trend, with an area wider in the west than the east. The area between the north and the south, moving from the south-east and through the center towards the south-west, shows no trend. These results slightly differ from the results reported by (Unal et al. 2012), which could be due to differences in the data used, which were for different time periods. Out of 26 significant grids 11 of them had significant trends at 0.05, while 15 had trends significant at 0.1. Only the grids on the coastal borders with the Black and Mediterranean seas, and several grids in the south east of Turkey, have a statistically significant trend. Other regions have insignificant trends indicating a random change.

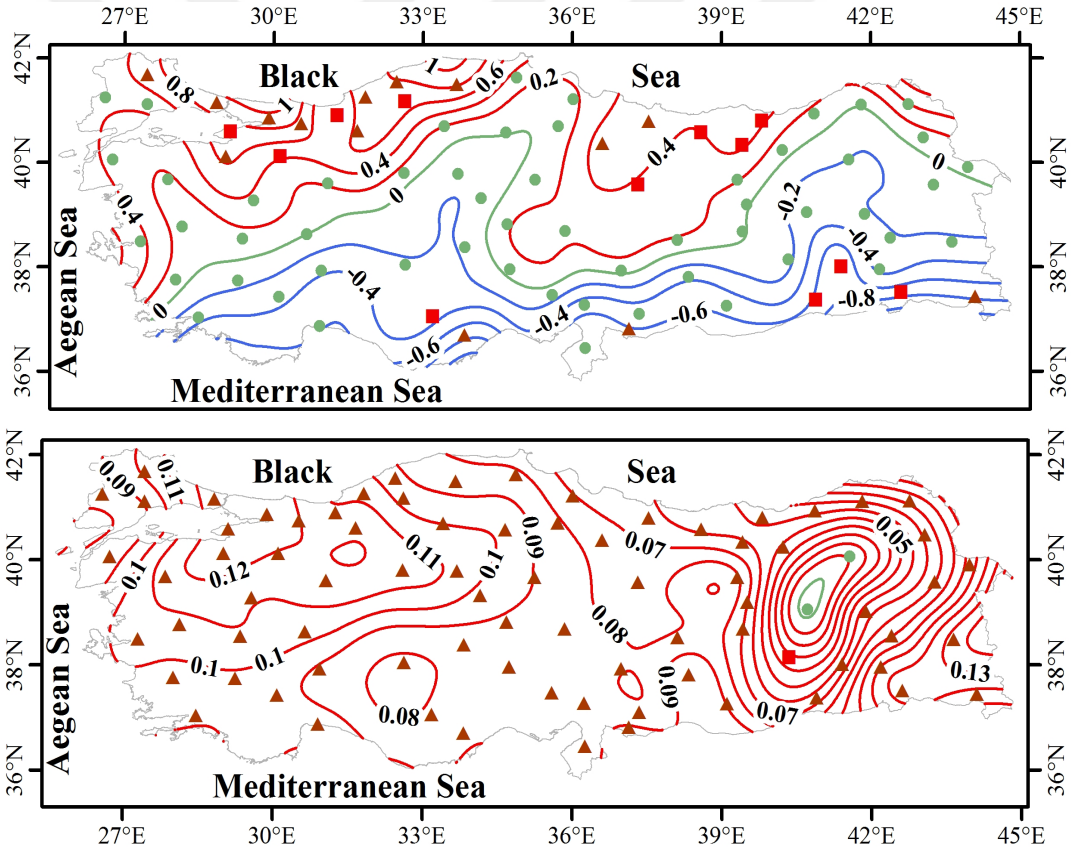


Figure 2.5. *Contour lines for the spatial distribution of the Annual Theil and Sen's Slope test values Top) Precipitation (mm/year), Bottom) Temperature (Co/decade)*

2.1.6.2.2. Temperature

Similar to precipitation, the average annual temperature of the whole of Turkey was plotted in Figure 2.3. The linear regression line indicates an apparent increasing trend for

the whole period. The Theil-Sen's Slope estimator value was 0.88 °C/century (calculated by multiplying the trend magnitude by 100), which was statistically significant. This result differs from the results found by (Kadioğlu 1997) who identified an insignificant increase. The LOWESS line shows an increase in the trend from the beginning of the series until the early 1950s, where it decreases slightly until the mid-1980s, then dramatically increases again. This was also reported by (Kadioğlu 1997), who found an increasing trend for the period of 1939-1989 and a decreasing trend for the period of 1959-1989. Using the Pettit test, the most probable change year identified as 1993.

Table 2 shows that 1993 is the most probable change year for temperature, which is repeated in 46 grids. The annual temperature trend does not vary greatly with lowest value of -0.03 °C/century in Bingöl being the only decreasing trend, and highest value of 1.34 °C/century in Van being the highest increasing trend.

Figure 2.4 shows that grids in all regions have an increasing temperature trend, except for one grid - Bingöl - located in Eastern Anatolia. Excluding this decreasing trend, the trend magnitudes are very similar, and the grids in the Marmara region have slightly higher magnitudes.

Figure 2.5 shows the spatial pattern of the decadal temperature trend (°C/decade). No obvious spatial pattern can be identified in any specific region, but generally the region about 5 degrees west of the eastern border of Turkey has the lowest trend in temperature, while the far south-eastern and north-western areas have the highest trend. In contrast with precipitation, except for the Bingöl and Erzurum grids (which are not significant) and Diyarbakir (which is significant at the 0.1 significance level), all the grids have a significant temperature trend at a level of 0.05.

2.1.6.3. Seasonal Trend

2.1.6.3.1. Precipitation

Precipitation trends differ from season to season (Table 2.1). Winter (i.e. DJF) has a decreasing trend in most of the grids, with the highest decrease in Muğla, but an increasing trend is observed in others, with Bartın having the highest value. In spring (MAM), Hatay has the most decreasing trend, while the most increasing trend is observed in Kocaeli. In the summer (JJA), however, Erzurum has the most decreasing trend and Kastamonu has the highest increasing trend. In autumn (SON), Antalya has the most decreasing trend, while Kirklareli has the most increasing trend. Although the trends have

great variation, most of the grids have no significant trend, and only 5, 9, 8, and 6 grids are significant for DJF, MAM, JJA, and SON, respectively.

Table 2.2. *The results of the Pettit test and trend value (C°/Century) with significance for Temperature. Ann: Annual, DJF: December, January, and February, MAM: March, April, and May, JJA: June, July, August, SON: September, October, and November*

Stations	Year	Ann	DJF	MAM	JJA	SON	Stations	Year	Ann	DJF	MAM	JJA	SON
Adana	1953	***0.85****	0.57	***1.21***	***1.48***	0.28	K.Maraş	1993	***0.80***	0.75	***1.09***	***1.39***	0.15
Adiyaman	1993	***0.97***	***1.00***	***1.36***	***1.43***	0.12	Karabük	1993	***0.97***	***1.08***	***0.76***	***1.45***	0.43
Afyon	1933	***1.04***	***1.05***	***1.19***	***1.69***	**0.58**	Karaman	1922	***0.81***	0.73	***0.89***	***1.38***	**0.52**
Agri	1976	***0.80***	0.83	***1.13***	***1.05***	0.30	Kars	1997	***0.72***	0.69	***0.91***	***1.05***	0.17
Aksaray	1950	***0.85***	0.75	***0.95***	***1.43***	0.45	Kastamonu	1950	***0.86***	***1.01***	***0.75***	***1.36***	0.28
Amasya	1993	***0.82***	0.71	***0.83***	***1.54***	0.26	Kayseri	1950	***0.83***	0.68	***1.03***	***1.45***	0.21
Ankara	1993	***1.13***	***1.42***	***1.06***	***1.59***	0.49	Kilis	1993	***0.91***	***0.79***	***1.28***	***1.43***	0.08
Antalya	1936	***0.94***	***0.80***	***1.01***	***1.44***	0.43	Kirkkale	1993	***1.08***	***1.21***	***1.08***	***1.57***	0.40
Ardahan	1993	***0.83***	0.85	***1.00***	***1.13***	0.38	Kirklareli	1993	***1.02***	***1.09***	***1.03***	***1.56***	0.44
Artvin	1993	***0.71***	0.39	***0.95***	***1.49***	0.13	Kirşehir	1950	***0.97***	**0.92**	***1.06***	***1.52***	0.38
Aydin	1949	***0.96***	***0.86***	***1.18***	***1.61***	**0.47**	Kocaeli	1993	***1.25***	***1.40***	***1.29***	***1.85***	***0.72***
Balikesir	1949	***1.14***	***1.06***	***1.22***	***1.76***	**0.58**	Konya	1922	***0.79***	0.74	***0.77***	***1.42***	0.54
Bartın	1993	***0.94***	***1.06***	***0.82***	***1.41***	0.37	Kütahya	1993	***1.11***	***1.13***	***1.22***	***1.79***	0.61
Batman	1951	***0.63***	0.37	***1.13***	***1.11***	-0.17	Malatya	1993	***0.82***	**0.87**	***1.11***	***1.32***	0.05
Bayburt	1993	***0.48***	0.19	***0.89***	***1.26***	-0.18	Manisa	1949	***1.08***	***0.98***	***1.23***	***1.74***	***0.58***
Bilecik	1993	***1.20***	***1.31***	***1.23***	***1.79***	***0.68***	Mardin	1953	***0.65***	0.68	***0.95***	***1.16***	-0.19
Bingöl	1986	-0.03	-0.56	**0.64**	***0.65***	***-0.74***	Mersin	1953	***0.88***	***0.81***	***1.03***	***1.41***	***0.54***
Bitlis	1976	***0.90***	0.88	***1.40***	***1.27***	0.10	Muğla	1949	***0.91***	***0.79***	***1.14***	***1.53***	0.38
Bolu	1993	***1.17***	***1.34***	***1.07***	***1.67***	***0.69***	Muş	1976	***0.64***	0.35	***1.28***	***1.11***	-0.08
Burdur	1936	***0.93***	***0.86***	***1.03***	***1.70***	0.37	Nevşehir	1950	***0.89***	0.81	***1.08***	***1.47***	0.30
Bursa	1993	***1.24***	***1.28***	***1.30***	***1.85***	***0.71***	Niğde	1953	***0.86***	0.69	***1.14***	***1.47***	0.36
Çanakkale	1993	***0.85***	**0.74**	***0.92***	***1.51***	0.40	Ordu	1993	***0.64***	0.70	***0.75***	***1.30***	0.05
Çankiri	1993	***1.00***	***1.19***	***0.87***	***1.46***	0.33	Osmaniye	1993	***0.83***	0.61	***1.21***	***1.44***	0.23
Çorum	1993	***0.97***	***1.00***	***0.95***	***1.59***	0.32	Rize	1993	***0.63***	0.13	***0.88***	***1.57***	0.07
Denizli	1936	***0.98***	***0.83***	***1.17***	***1.67***	**0.45**	Sakarya	1993	***1.20***	***1.37***	***1.15***	***1.75***	***0.69***
Diyarbakir	1914	**0.33**	0.47	**0.60**	***0.93***	**-0.49**	Samsun	1993	***0.71***	0.62	***0.70***	***1.49***	0.18
Düzce	1993	***1.12***	***1.21***	***0.95***	***1.64***	***0.67***	Sanliurfa	1993	***0.90***	***0.95***	***1.28***	***1.35***	0.00
Edirne	1993	***0.83***	**0.78**	***0.89***	***1.47***	0.37	Siirt	1976	***1.10***	***1.30***	***1.50***	***1.40***	0.27
Elazığ	1993	***0.70***	0.93	***0.87***	***1.11***	-0.07	Sinop	1993	***0.82***	**0.82**	***0.81***	***1.51***	0.18
Erzincan	1993	***0.81***	**1.00**	***1.08***	***1.33***	0.16	Şirnak	1976	***1.24***	***1.61***	***1.59***	***1.47***	0.28
Erzurum	1914	0.11	-0.46	***0.74***	***0.70***	-0.47	Sivas	1993	***0.74***	**0.81**	***0.94***	***1.27***	0.04
Eskişehir	1993	***1.16***	***1.34***	***1.16***	***1.72***	**0.65**	Tekirdağ	1993	***0.91***	***0.90***	***0.98***	***1.48***	0.39
Gaziantep	1993	***0.83***	**0.72**	***1.25***	***1.39***	0.08	Tokat	1993	***0.75***	0.74	***0.84***	***1.43***	0.10
Giresun	1993	***0.58***	0.50	***0.76***	***1.31***	0.00	Trabzon	1993	***0.63***	0.36	***0.88***	***1.47***	0.00
Gümüşhane	1993	***0.61***	0.46	***0.82***	***1.39***	0.00	Tunceli	1993	***0.71***	0.94	***0.96***	***1.18***	0.00
Hakkari	1993	***1.16***	***1.72***	***1.53***	***1.37***	0.15	Uşak	1993	***1.09***	***1.03***	***1.23***	***1.77***	***0.60***
Hatay	1993	***0.83***	**0.60**	***1.22***	***1.45***	0.18	Van	1977	***1.34***	***1.97***	***1.70***	***1.52***	0.10
Iğdır	1977	***1.04***	**1.32**	***1.21***	***1.08***	**0.56**	Yalova	1976	***1.25***	***1.33***	***1.33***	***1.88***	***0.72***
Isparta	1924	***0.89***	**0.86**	***0.93***	***1.58***	0.46	Yozgat	1950	***0.90***	**0.84**	***0.98***	***1.58***	0.29
Istanbul	1993	***1.24***	***1.38***	***1.32***	***1.84***	***0.78***	Zinguldak	1993	***1.07***	***1.17***	***0.90***	***1.58***	**0.61**
Izmir	1949	***1.05***	***0.87***	***1.20***	***1.62***	**0.55**							

***** The Bold values are significant at significance level $\alpha = 0.1$ and the bold italic underlined values are significant at significance level $\alpha = 0.05$.

Figure 2.6 shows that most of the grids in the South-Eastern Anatolian, Mediterranean, and Aegean regions have a decreasing precipitation trend in the winter. All the seasons, except winter, have an increasing trend in the Aegean region. The grids of the Marmara and Black Sea regions have an increasing trend in all seasons, with exception of several grids that have a decreasing trend in winter. Other regions have no noticeable pattern as trends in some grids increase or decrease in different seasons. The trend values for any season have no difference with any other season as they vary from grid to grid, but the most consistent values between grids can be identified in the spring.

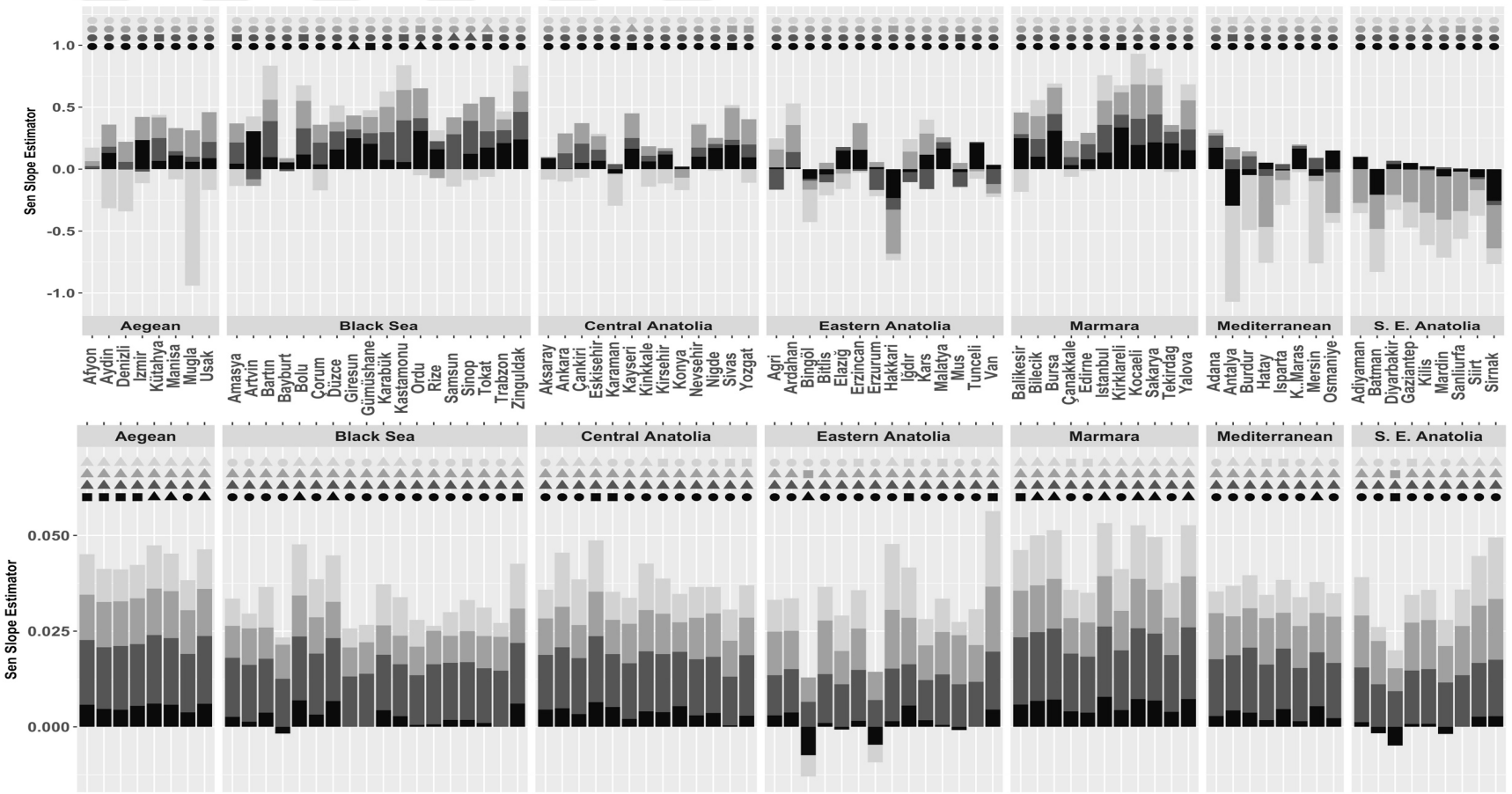


Figure 2.6. *Seasonal Theil and Sen's Slope test values Top) Precipitation (mm/year), Bottom) Temperature (Co/year). Circle: Not significant, Square: significant at 0.1, Triangle: Significant at 0.05. Light to dark grey shades represents: Winter, Spring, Summer, and Autumn respectively*

According to Figure 2.7, winter has a decreasing precipitation trend that is observed in most of Turkey, with the highest decreasing trend in the south-western area which includes the grids where this trend is significant. There is a slight increasing trend in the north-western and south-eastern parts of Turkey. This finding is consistent with (Unal et al. 2012), who studied the period of 1961-2008, but differs from the findings of (Türkeş et al. 2009), who studied the period of 1953-2002. No dominant spatial pattern for precipitation over the whole of Turkey was observed in spring. A decreasing trend in the South-Eastern Anatolia, and a small part of Central Anatolia was identified, while an increasing trend was found in the remaining areas. Only several grids were identified as significant grids at the two significance levels 0.1 and 0.05. These results agree to that reported by (Türkeş et al. 2009, Unal et al. 2012). Although significance is only identified in several grids located mostly in the Black Sea region, a general increasing trend was observed in the summer with small areas having no trend and some areas having a decreasing trend in the Eastern Anatolia region. With only four significant grids, autumn has a similar spatial pattern to spring with a decreasing trend in the South-Eastern Anatolia and the western part of the Mediterranean region. This result of summer and autumn are consistent with (Türkeş et al. 2009) but differs slightly from (Unal et al. 2012).

2.1.6.3.2. Temperature

A decreasing temperature trend in the winter is only found in two grids - Bingöl and Erzurum - while the others have an increasing trend, with highest observed in Van (Table 2.2). An increasing trend is observed in spring in all grids, with the lowest value in Diyarbakir and the highest in Van. In the summer, Bingöl was found to have the lowest increasing trend and Yalova the highest. Varying trend values are observed throughout the grids in autumn, and several of them have a decreasing trend, with the highest decrease in Bingöl, several grids have no trend, but most have an increasing trend. In contrast with precipitation, over half of the grids are found with significant trends in winter. In spring, all the grids are significant at the 0.05 significance level, with two grids - Bingöl and Diyarbakir - significant at 0.1. All grids have a significant trend in summer, whereas trends in most grids in autumn are not significant.

Figure 2.6 indicates a noticeable increasing trend in all regions and in all seasons. There is a decreasing trend in autumn in several grids in the South-Eastern Anatolia and Eastern Anatolia regions, and in one grid in the Black Sea region. In winter, only two

grids in Eastern Anatolia have a decreasing trend. The lowest trend magnitudes are in autumn and highest in summer, and winter and summer have relatively close trend magnitudes.

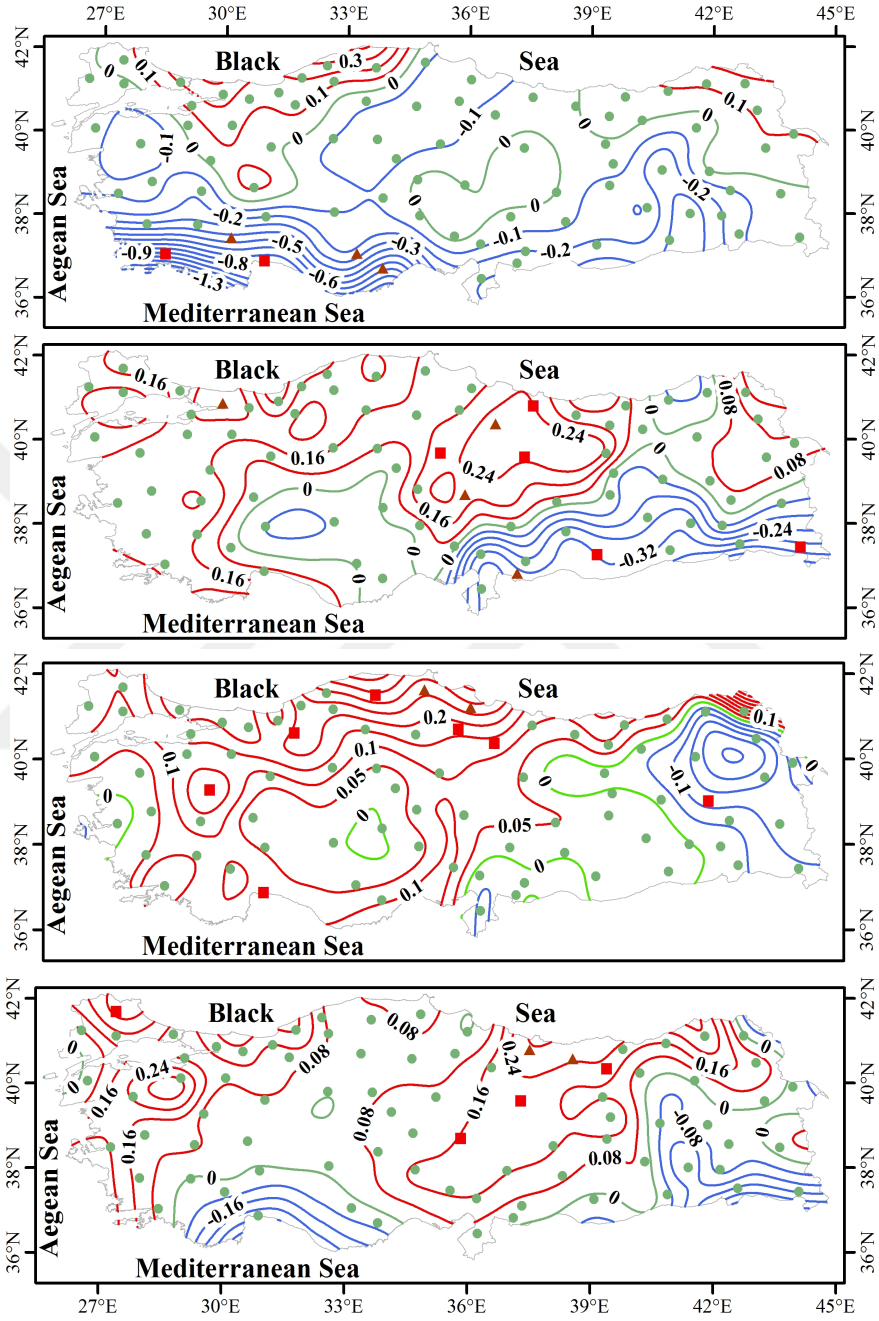


Figure 2.7. *Contour lines for the spatial distribution of the Theil and Sen's Slope test values of seasonal precipitation trend (mm/year) from top Winter, Spring, Summer, and Autumn. Circle: Not significant, Square: significant at 0.1, Triangle: Significant at 0.05*

The winter spatial pattern (Figure 2.8) indicates that an increasing trend is dominant in most of Turkey, with a small area located in the middle of Eastern Anatolia which has a decreasing trend. The significance is observed in the western and far eastern areas. Similar to the summer pattern, the spring pattern features an increasing trend throughout the whole of Turkey, with slightly higher magnitudes in the summer. For spring and summer, all grids have significant trends. In the autumn, the trend has a similar pattern to winter, with a larger decreasing trend area in the middle of Eastern Anatolia. The magnitudes of the trend are smaller than that observed in winter. Significance is also similar to winter, but it is mostly located in the west rather than the east.

These findings differ from the results reported by (Türkeş et al. 2002) as they reported a decreasing trend in the winter in a wide part of Turkey, whereas only two grids featured an increasing trend in this study. They stated that several areas had an insignificant decreasing trend in the spring, which is not matching the findings here in which a significant increasing trend in all grids was found. The autumn results also differ in that they identified the dominance of the decreasing trend, while in this study a decreasing trend is only found in eastern Turkey. This difference might be due to the length of the periods studied; the period covered by (Türkeş et al. 2002) is 1929-1999, while this study covers 1901-2014. This is the most likely reason for the disparity between results, because after the mid years of 1990s, trends increased dramatically.

2.1.7. Conclusion

The main objective of this study is to analyze the trends over Turkey of two climate variables, precipitation and temperature, for a long-term period covering 1901-2014 using the CRU TS V3.23 data. Annual and seasonal trends were analyzed by using grids, with each grid represents a province. The annual areal precipitation for the whole of Turkey had an insignificant increasing trend while annual areal temperature had a significant increasing trend. According to the regional grid-based analysis, the Aegean, Black Sea, and Marmara regions have an increasing trend in annual precipitation, while the South-Eastern Anatolia, and Mediterranean regions have decreasing trends. Other regions have variation between grids with no dominant pattern. The insignificance is identified in most of the grids. The analysis of annual temperature indicated that all regions have a significant increasing trend. The conclusion of the seasonal analysis was that winter has a decreasing trend of precipitation in general, with the exception of the Marmara and

Black Sea regions, in which an insignificant increasing trend was identified. All the other seasons have predominant increasing trends in all regions but South-Eastern Anatolia, in which a decreasing trend is found in all seasons. In contrast with precipitation trends, seasonal temperature has a significant increasing trend in all seasons and regions.

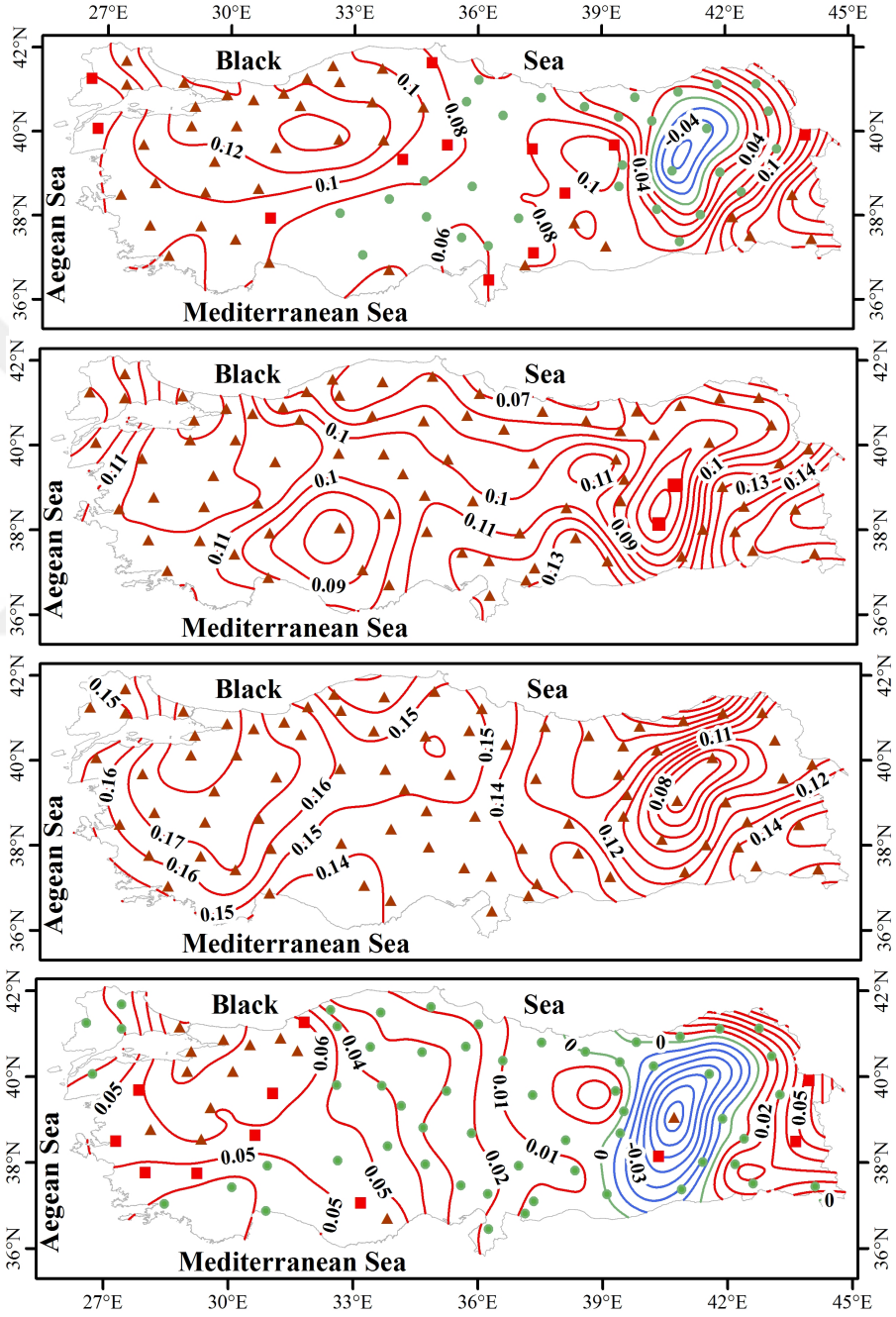


Figure 2.8. *Contour lines for the spatial distribution of the Theil and Sen's Slope test values of seasonal temperature trend (Co/decade) from top Winter, Spring, Summer, and Autumn. Circle: Not significant, Square: significant at 0.1, Triangle: Significant at 0.05*

These results are consistent with some findings, but differ with the same studies in other findings, even in aspects that should be consistent. For example, the annual areal precipitation for the whole of Turkey showed different results. The annual and seasonal temperature trends were also found to differ from the earlier studies. This inconsistency might be due to the local and regional heterogeneity of a large country like Turkey, and the usage of different data that cover different periods.

This study used the analysis of the period from 1901-2014 as one epoch. Dividing the period into two or more and analyzing the trends for each period separately is recommended. The period following 1985 is considered to be a change point in Turkey due to a dramatic increase in economical and industrial developments, which could cause significant climate change (Unal et al. 2012).

2.2. Homogeneity and Trend Analysis of Rainfall and Streamflow of Seyhan Basin, Turkey.

2.2.1. Abstract

Changes in the historical rainfall and streamflow observations are important for understanding the climate change. The homogeneity of the annual and monthly times series of observations of 9 rainfall stations and 3 streamflow gages located at Seyhan basin in Turkey was examined using standard normal homogeneity test and Buishand test. The trend was studied by the mean of Mann-Kendall test which is used for identifying the significance of the trend, and the Sen's slope estimator which provides the magnitude of the trend. Finally, the Pettit test was also applied on the annual time series only for obtaining the most probable change year. The entire time series found as homogeneous although several significant p values obtained in the 0.05 but not in the 0.01 significance level. The annual trend varied from station to another but found significant in one rainfall and one streamflow stations and significant at 0.1 significance level only. Monthly time series found to have most significant trend in January but still significant only in 0.05 significance level. Several months have no station with significant trend while others have only one station with significant trend. The Pettit test results showed no consistent change year through all station although 1981 is most obtained year.

2.2.2. Introduction

Rainfall is considered as one of the most important climatic variables in the fields of hydrology and climatology. It is also crucial in water resources management because of the critical importance in the spatial and temporal water availability. In the rainfall-streamflow relationships, drought and flood assessment and control, mitigation measures, rainfall is critical factor. Therefore, changes in historical rainfall observations are essential for water resources planners and climatologists (Luís et al. 2000, Taxak et al. 2014, Jones et al. 2016). Streamflow is also extremely important in several aspects such as keeping the rivers and lakes full of water and monitoring the land change by the effect of erosion.

Many studies examined the trends in precipitation/rainfall around the globe, for example: (Turkes 1996, Buffoni et al. 1999, De Luís et al. 2000, Yue et al. 2003, Gemmer et al. 2004, Marengo 2004, Partal et al. 2006, Pujol et al. 2007, Kampata et al. 2008, Türkeş et al. 2009, de la Casa et al. 2010, Karpouzos et al. 2010, Kumar et al. 2010, Longobardi et al. 2010, Oguntunde et al. 2011, Tabari et al. 2011, Yavuz et al. 2012, Duhan et al. 2013, Haktanir et al. 2013, Akinsanola et al. 2015, Ay et al. 2015, Yürekli 2015)

Some of these studies investigated the trend in for the entire Turkey and some in local scale with several stations. Here are some examples of these studies. (Türkeş et al. 2009) used 97 rainfall stations' observations covering the whole Turkey for the period of 1930-2002 for investigating the seasonal spatial and temporal variation. They found a deceasing trend in winter and increasing in spring, summer, autumn. The decreasing trend was cleary obvious in the Mediterranean region. The monthly trend of whole Turkey using rainfall observations of 120 stations covering the period 1975-2009 was studied by (Yavuz et al. 2012). They identified an increasing trend in March, April, and October. The trend of rainfall in six provinces of Turkey was studied by (Ay et al. 2015) for the objective of comparing the Mann-Kendall test with the Şen's test which was recently developed by (Şen 2012). They found different patterns of trend and concluded that the new method is reliable.

The objective of this study is to investigate the homogeneity and the trend of the monthly and annual rainfall and streamflow of three subbasins located at Seyhan basin, Turkey using observations extends for the period 1973-2000.

2.2.3. Study area and data collection

This study goal is to investigate the trend in the monthly and annual rainfall and streamflow data for the period of 1973-2000 in 9 rainfall and 3 streamflow stations located in three sub-basins called (coded) 1822, 1801, and 1805, as seen in Figure 2.9. These three sub-basins are part of the Seyhan Basin located in the southwest of Turkey. Several stations were chosen outside the subbasin for including all spatially correlated rainfall station which could affect the streamflow value in addition to have wide spatial distribution. The study area covers parts of the following provinces: Sivas, Adana, Kayseri, and Kahramanmaras. During the study period, the sub-basins are: rural, agricultural, not hydrologically disturbed, and having no water storage structure.

Rainfall observations of 9 stations and streamflow of the 3 subbasins outlets collected from the the Turkish General Directory of Meteorology DMI (Devlet Meteoroloji Isleri) and General Directorate of State Hydraulic Works DSI (Develt Su Isleri). The consistent available data for all stations are for the period 01/02/1973 to 30/09/2000. The collected data were daily total of rainfall and average of streamflow. Basic statistics of the monthly data are shown in Table 2.3.

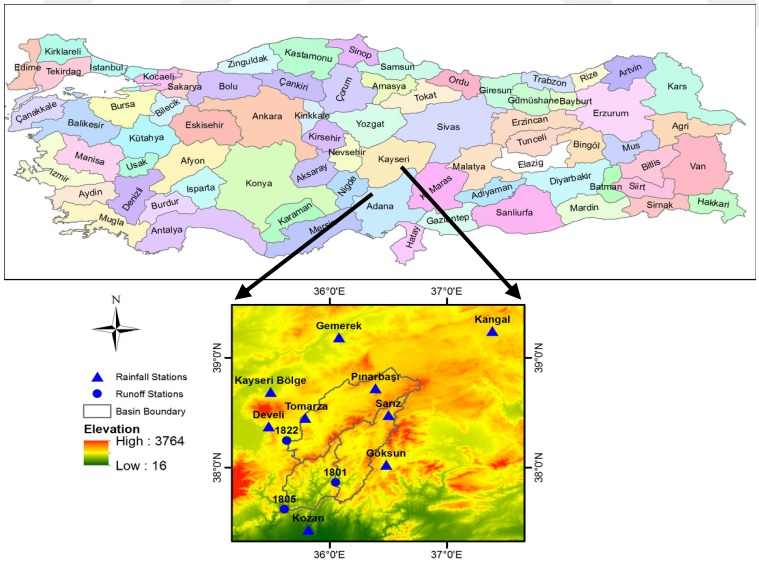


Figure 2.9. *Study area*

2.2.4. Methodology

Collected daily data were converted to monthly and annual total for rainfall and monthly and annual average for streamflow. The homogeneity of the monthly and annual time series was examined by the mean of two tests: Alexanderson's Standard Normal Homogeneity Test (SNHT), Buishand tests. Mann-Kendall test was applied for the data in order to identify the significance of the trend. Theil and Sen's slope estimator applied for obtaining the magnitude of the trend regardless its significance, Pettitt's test used for detecting the most probable change year and it was applied only for the annual series. The statistical language R was used for implementing these tests. ArcGIS was used for the presentation of the stations on the map with their trend significance and magnitude.

Table 2.3. *Statistics of the monthly total precipitation and monthly mean streamflow*

Stations	Name	No. Obs.	Min	Max	Range	Median	Mean	Std.dev	Coef.var
17162	Gemerek	332	0	159.9	159.9	29.4	33.4	27.9	0.8
17196	Kayseri	332	0	164.7	164.7	28.2	33.5	26.8	0.8
17762	Kangal	332	0	164.2	164.2	28.8	35.1	30.4	0.9
17802	Pınarbaşı	332	0	159.8	159.8	29.9	35	28.6	0.8
17836	Develi	332	0	159.6	159.6	26.6	30.5	27.1	0.9
17837	Tomarza	332	0	129.4	129.4	28.7	32.8	26.1	0.8
17840	Sarız	332	0	187.5	187.5	39.5	44.6	34.5	0.8
17866	Göksun	332	0	185.3	185.3	38.5	51.2	46	0.9
17908	Kozan	332	0	296	296	60.5	71	59.6	0.8
ST1801	Göksu–Himmetli	332	9.2	149.4	140.2	20.1	29.4	23.3	0.8
ST1822	Göksu–Gökdere	332	5	91.5	86.5	14.9	20.3	14.7	0.7
ST1805	Zamanti–Fraktin	260	14.9	267.1	252.2	40.2	60.7	50.8	0.8

2.2.4.1. Standard Normal Homogeneity Test (SNHT)

Alexanderson's Standard Normal Homogeneity Test (SNHT) (Alexandersson 1986) firstly developed for the identification of the occurrences of a change in a time series of rainfall data. The null hypothesis (H_0) of this test is that the tested variable is independent and identically distributed while the alternative hypothesis (H_a) is a change (break or step shift) is present in the mean.

2.2.4.2. Buishand test

Buishand test (Buishand 1982) is particularly studied for a normal distribution but still can be used irrespective to the distribution of the tested variable. This test has three cases: the first case is two-tailed, the second and third are left-tailed. For all cases, the null hypothesis (H_0) is that the variable has one or more distribution but with the same

mean. The alternative hypothesis (H_a) is: a change in the mean occurred in a t time point for the two-tailed, and the mean is reduced or augmented for one-tailed test.

2.2.5. Results and Discussion

2.2.5.1. Exploratory Data Analysis (EDA)

This study is to investigate and analyze the spatial and temporal homogeneity and trend in the monthly and annual rainfall and monthly and annual streamflow. According to Table 2.3 which shows the basic statistics of all stations, most of the rainfall stations have close values with the exception of Kozan station. For example, the maximum values of the monthly rainfall are close for all stations that varies between 129 and 187 mm while for Kozan station far from this range with a value 296 mm. The median and the mean of all stations are also close to each other vary from about 29 to 39 mm and 30 to 51 mm while for Kozan are 60.5 and 71 mm respectively. The coefficient of variation values are 0.8 and 0.9 and almost not different among all rainfall stations. Streamflow values varied greatly from one station to another. The minimum and maximum values are not close to each other and the same thing can be seen in the median values. The mean values vary from about 20 to 60 m^3/s. The areas of the subbasins: 2400, 1788, 3610 km2 for 1801, 1805 and 1822 respectively. For the subbasin 1801 and 1822, the values of the streamflow are rational with their areas. The subbasin 1805 has the smallest area and highest streamflow values which is due to that the subbasin 1801 outlet is an inlet to 1805 subbasin. In other words, 1805 subbasin is an accumulation of both 1801 and 1805 subbasins which lead to high streamflow.

2.2.5.2. Homogeneity

Homogeneity tests are important for identifying whether a change occurred in a time series or it is homogenous. The P values of the homogeneity tests (i.e. Standard Normal Homogeneity Test, and Buishand Test) are listed in Table 2.4. The results show that the null hypotheses of the two tests cannot be rejected for 0.01 significance level in all station for the two temporal resolutions: Annual and monthly. For the monthly time series, the null hypothesis using Buishand test of only one rainfall station 17802 and one streamflow station 1805 can be rejected and only in the 0.1 and 0.05 significance level respectively. For the annual series, null hypothesis of 17162 rainfall stations using SNHT test, and 17802 using SNHT and Buishand test can be rejected and only in the significance

level of 0.05. In general, the entire rainfall and streamflow stations are homogeneous and the no significant change occurred during the studied period.

The most probable change year which represents the results of the Pettit-Mann-Whitney test showed no consistent result through the entire stations. The most frequent year is 1981, and 1988 which appears 4 and 3 times respectively while others have no consistent change year.

Table 2.4. *P values of Homogeneity tests, SNHT: Standard Normal Homogeneity Test, Buishand: Buishand Test*

	Annual		Monthly	
	SNHT	**Buishand**	**SNHT**	**Buishand**
17162	0.027**	0.473	0.640	0.990
17196	0.576	0.334	0.605	0.368
17762	0.423	0.229	0.710	0.430
17802	0.031**	0.025**	0.187	0.075*
17836	0.828	0.904	0.661	0.891
17837	0.519	0.931	0.497	0.988
17840	0.097	0.260	0.440	0.463
17866	0.404	0.543	0.840	0.708
17908	0.627	0.344	0.977	0.706
ST1801	0.333	0.828	0.256	0.381
ST1805	0.261	0.240	0.132	0.039**
ST1822	0.185	0.753	0.133	0.488

* Significant at $\alpha = 0.1$.
** Significant at $\alpha = 0.05$.

2.2.5.3. Annual trend

The regression line fitted to the time series of all stations (Figure 2.10) shows that in all stations including streamflow the trend is negative except in 17196 rainfall station which has positive trend. The steepest trend is seen in 17908 as a rainfall station while among the streamflow stations 1805 has the steepest trend. These lines interpretations are consistent with the Sen's estimator results shown in Table 2.5. According to Table 2.5, the trend varies between the highest negative value -4.30 which is for 17908 rainfall station and the lowest negative value -0.4 which is for 17837. The only positive trend is in 17196 stations. Streamflow trend varies between -1.3 and -0.03.

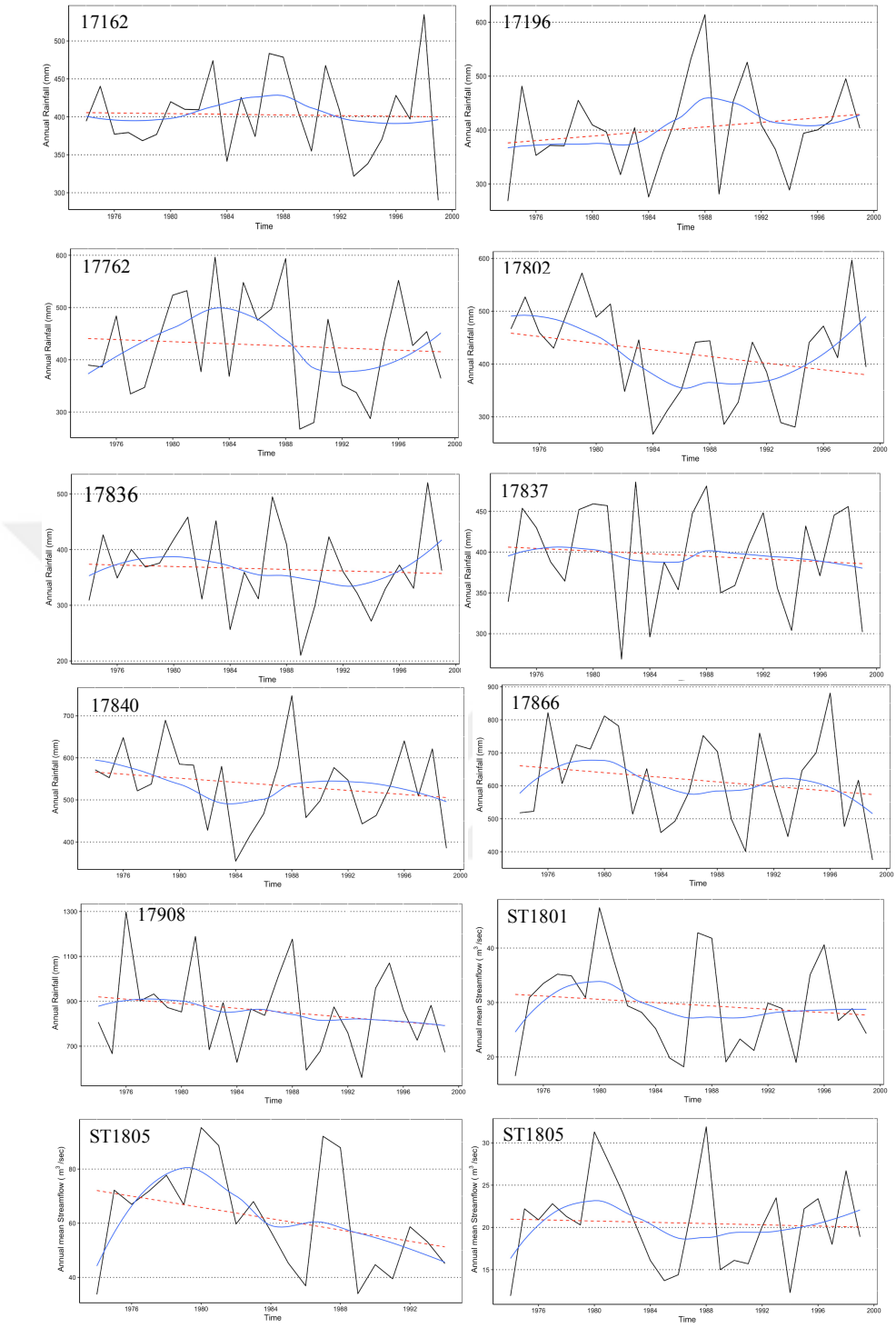


Figure 2.10. *The time series of the annual rainfall and annual streamflow. Line: Linear Regression line, Dashed Line: LOWESS regression Line*

The significance of the trend which obtained by Mann-Kendall test is shown in Table 2.5. It is also shown in Figure 2.11 that the station has significant trend colored with red color. Regardless the magnitude of the trend mentioned earlier, only the rainfall

station 17802 and streamflow station ST1805 has a significant annual trend with a 0.1 significance level.

Table 2.5. *The results of the Pettit test (change year) and trend value (mm/year) (Sen's Estimator) with significance (Mann-Kendall test) for rainfall. The underlined significant at 0.05 Bold significant at 0.1*

Stations	Change Year	Sen Slope
17162	1992	-0.41
17196	1985	1.78
17762	1988	-0.82
17802	1981	-3.53*
17836	1983	-0.67
17837	1981	-0.40
17840	1981	-2.71
17866	1981	-4.27
17908	1988	-4.30
ST1801	1982	-0.21
ST1805	1988	-1.30*
ST1822	1982	-0.03

* Significant at $\alpha = 0.1$

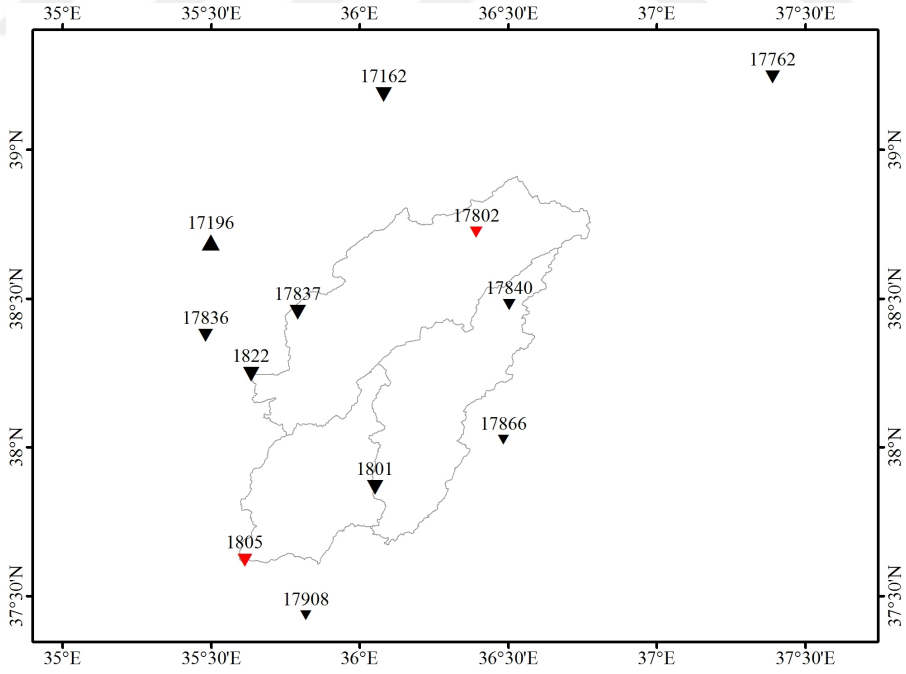


Figure 2.11. *Spatial distribution of the stations with the annual trend. The size of the triangle represents the normalized magnitude of the trend. The up-pointing triangle: positive trend, the down pointing triangle: negative trend. Black: Not significant, Red: Significant at α= 0.1*

2.2.5.4. Monthly trend

The monthly trend results are shown in Table 2.6 and Figure 2.12. Table 2.6 shows the magnitude of the trend for all stations in each month. According to the table, January has the largest values of trend in most of the rainfall stations. Interestingly, the trend in all rainfall station is negative which indicate a decrease in the amount of the received rainfall in this month. The other months trend varies from negative to positive to no trend in all station that no consistent pattern can be indicated.

The streamflow stations have a positive trend for two stations 1801, 1822 and negative for 1805 in January. The trend is negative through the months starting from February till September and from February till July in the streamflow stations 1801 and 1805 respectively. The 1805 streamflow station has a positive trend only in November.

The significance of the trend for every month is shown in Figure 2.12. According to the figure, January has the largest number of stations which have a significant trend. All the stations have a significant trend in January except one rainfall station and two streamflow stations that only the 1805 streamflow station has a significant trend. The significance in the other months is not dominant like January. In February, March, August, September, October, and November, no station including rainfall and streamflow has a significant trend. All the other months have only one station with significant trend except July which has two significant stations.

2.2.6. Conclusion

The aim of this study is to investigate the homogeneity and trend of observation 9 rainfall stations and 3 streamflow stations located at Seyhan basin, Turkey for the period 1973 - 2000. The homogeneity test conducted by the mean of two tests: Standard Normal Homogeneity Test (SNHT) and Buishand test. The entire stations were found, in general, homogenous and no change has occurred in the time series through the studied period although several significant values obtained in 0.05 significance level. The trend was examined by Mann-Kendall test which gives the significance of the trend and Sen's slope estimator which measures the magnitude of the trend. The annual trend found as not significant only in one rainfall and one streamflow station and the monthly trend found as mostly significant in January while other months have no or only one station significant. Pettit test used for the identification of the most probable change year. No consistent year found through all the stations but most obtained year is 1981.

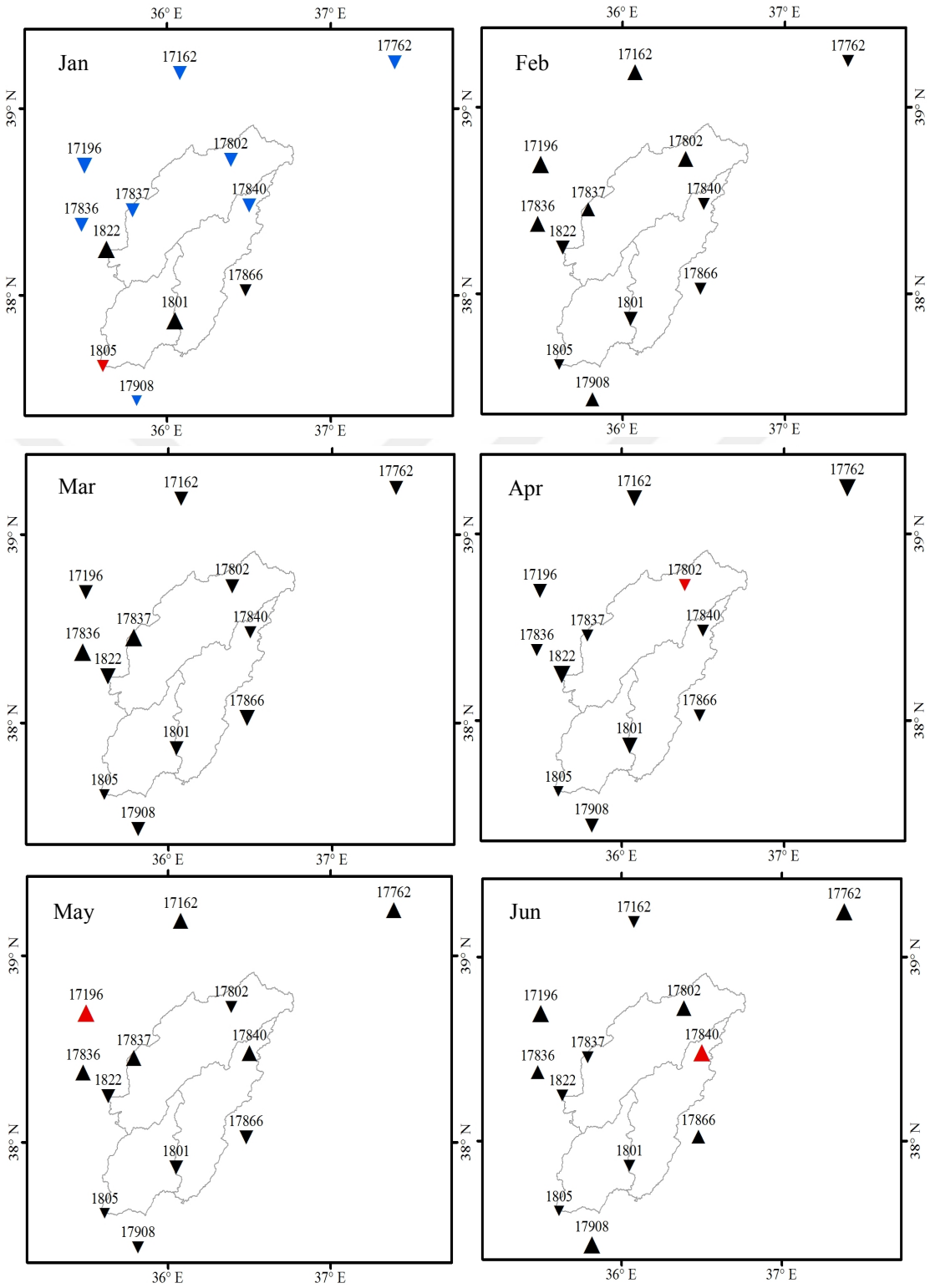


Figure 2.12. *Same as Figure 2.11 but for monthly trend and the Blue: Significant trend at α= 0.05*

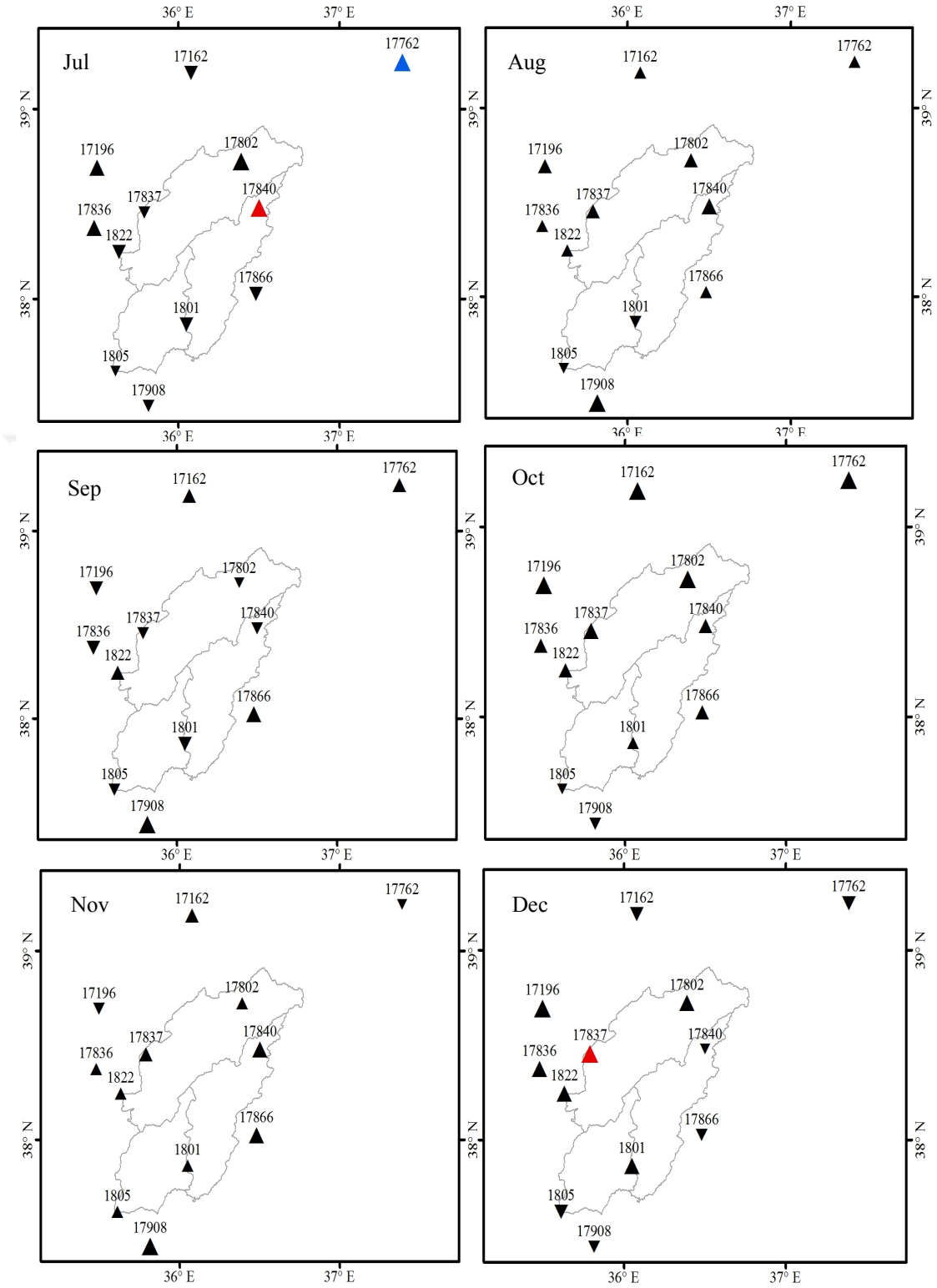


Figure 2.12. *Continued*

Table 2.6. *The results of the trend value (mm/year) with significance for rainfall*

Stations	January	February	March	April	May	June	July	August	September	October	November	December
17162	-1.26**	0.40	-0.39	-0.62	0.26	-0.15	-0.03	0.00	0.00	0.55	0.79	-0.10
17196	-0.55**	0.56	-0.35	-0.77	1.24*	0.83	0.01	0.10	-0.07	0.58	-0.04	0.87
17762	-1.47**	-0.57	-0.40	-0.13	0.06	0.88	0.25**	0.00	0.01	0.54	-0.29	-0.28
17802	-1.20**	0.41	-0.30	-1.45*	-1.12	0.41	0.17	0.08	-0.53	0.50	0.04	0.13
17836	-1.12**	0.33	0.17	-1.12	0.42	0.11	0.00	0.02	-0.02	0.15	0.09	0.42
17837	-1.03**	0.04	0.07	-1.08	0.29	-0.17	-0.15	0.13	-0.28	0.28	0.51	0.81*
17840	-1.44**	-0.68	-0.88	-1.08	0.14	1.20*	0.33*	0.21	-0.27	0.09	0.95	-1.16
17866	-2.02	-1.13	-0.03	-1.03	-0.29	0.24	-0.04	0.00	0.16	0.11	1.13	-0.61
17908	-4.63**	0.07	-0.35	-0.73	-1.38	0.94	-0.09	0.40	0.58	-0.01	2.46	-0.54
1801	0.22	-0.01	-0.58	-0.67	-0.42	-0.14	-0.04	-0.03	-0.04	0.00	0.06	0.22
1805	-1.97*	-2.25	-1.79	-3.13	-2.96	-0.24	-0.29	-0.21	-0.19	-0.29	0.25	-0.37
1822	0.10	-0.16	-0.20	-0.25	-0.09	-0.05	-0.02	0.00	0.04	0.07	0.12	0.12

CHAPTER THREE

3. DAILY STREAMFLOW MODELING USING DATA-DRIVEN MODELS

3.1. Forecasting of Daily Streamflow For Basins With Different Physical Characteristics Using Data-Driven Methods

3.1.1. Abstract

Modelling streamflow is essential for activities, such as flood control, drought mitigation, and water resources utilization and management. Artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), and support vector machines (SVM) are techniques that are frequently used in hydrology to specifically model streamflow. This study compares the accuracy of ANN, ANFIS, and SVM in forecasting the daily streamflow with the traditional approach known as autoregressive (AR) model for basins with different physical characteristics. The accuracies of the models are compared for three basins, that is, 1801, 1805, and 1822, at the Seyhan River Basin in Turkey. The comparison was performed by using coefficient of efficiency, index of agreement, and root-mean-square error. Results indicate that ANN and ANFIS are more accurate than AR and SVM for all the basins. ANN and ANFIS perform similarly, while ANN outperformed ANFIS. Among the models used, the ANN demonstrates the highest performance in forecasting the peak flood values. This study also finds that physical characteristics, such as small area, high slope, and high elevation variation, and streamflow variance deteriorate the accuracy of the methods.

3.1.2. Introduction

Hydrology is a science that studies water life cycle and its movements in terms of topographical characteristics. Streamflow, or the flow in general, is crucial for many aspects of water resource management, such as the amount of water supplied from the river for irrigation or drinking networks. Thus, modeling the flow is essential for many activities, such as flood or drought control, and proper water resources utilization (Singh 2016).

Owing to the importance of water resources planning, the branch of flow modeling, in particular, remains an active area of study. The models used in hydrology can be divided into two broad categories, namely, data-driven and physically based models. Physically based models are white-box models that consider physical interaction for modeling hydrological process. By contrast, data-driven models are black-box models

that obtain the optimum relationship between inputs and outputs but disregard the physical process. Several computing tools, such as artificial neural network (ANN), support vector machine (SVM), fuzzy logic, and a hybrid model, namely, adaptive neuro-fuzzy inference system (ANFIS), are applied in streamflow modeling.

Autoregressive (AR) model is a time-series based technique that is used in modeling hydrological process, especially the streamflow with timestamp information. The AR is popular given its simplicity and leniency. Accordingly, AR is applied in several flow modeling studies and compared with other similar models. The comparative works can be found in (Kişi 2008, Tan et al. 2015). The AR may not consistently and appropriately model the possibly complex relationship between the streamflow and previous lags because the AR is a linear model. By contrast, ANN is a nonlinear data-driven model that can overcome this shortcoming. Therefore, the ANN is increasingly adopted in hydrology. Several studies have used the ANN for modeling streamflow (Maier et al. 2000, Zadeh et al. 2010, Isik et al. 2013, Shiau et al. 2016, Abdollahi et al. 2017, Mehr et al. 2017). ANN theory and its applications in hydrology are reviewed by (Govindaraju 2000) and (Govindaraju 2000).

Moreover, ANN is used to tune the membership function parameters of a fuzzy inference system (FIS). The ANFIS (Jang et al. 1997) is a hybrid model of the ANN and fuzzy logic. This hybrid model is also used for hydrological application, particularly in streamflow and runoff prediction (Chang et al. 2001, Firat et al. 2007, Khadangi et al. 2009, Kişi et al. 2012, Sanikhani et al. 2012, Akrami et al. 2014, Tan et al. 2015).

SVM is not much adopted in the hydrological applications and daily streamflow modeling, unlike the ANN and ANFIS. Several examples of SVM that is applied in streamflow modeling include (Lin et al. 2006, Aggarwal et al. 2012, Kişi 2015, Ghorbani et al. 2016, Seyam et al. 2017). A comprehensive review of the applications of data-driven models in hydrology can be found in (Yaseen et al. 2015, Fahimi et al. 2016).

However, studies that compared the accuracy of these models are limited. For example, (Wang et al. 2009) compared the capability of several artificial intelligence (AI) models (i.e., AR moving average (ARMA), ANN, ANFIS, and genetic programming [GP]) in forecasting the monthly discharge of two rivers. The comparison among various models concludes that ANFIS is better than GP, but GP outperforms SVM. (Kişi et al. 2012) compared the accuracy of several AI models in forecasting daily intermittent streamflow. The authors noted that ANN, ANFIS, and SVM perform better than local

linear regression and dynamic local linear regression. Moreover, the ANN and ANFIS demonstrate results that are better than SVM. However, these results can be strengthened by using data from various regions. By contrast, (He et al. 2014) conducted a comparative study by using several data-driven models to illustrate streamflow in the semi-arid mountainous area. The authors concluded that SVM outperforms ANN and ANFIS and hence suggested that the accuracy of the model in forecasting high flood values requires further investigation.

However, literature that compares the performance of these methods for basins with different physical characteristics and the effect of these characteristics on the accuracy remains scarce. This study is aimed at examining the capability of data-driven methods, such as AR, ANN, ANFIS, and SVM, in modeling the daily streamflow for basins with largely different physical characteristics, such as area, elevation, and slope. The accuracy of each method in forecasting the peak flood values must also be investigated.

3.1.3. Background of the Data-Driven Methods

3.1.3.1. AR model

AR is a mostly used model in time-series modeling. The AR model predicts a value of a variable one-step ahead based on the values that emerged previously. In particular, the AR model is the regression of values based on prior occurrence. Thus, the AR model for an order p is written as AR(p) and represented by

$$X_t = \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_p X_{t-p} + \epsilon_t, \quad (3.1)$$

where ϵ_t is a white noise with $E(\epsilon_t) = 0$ and $VAR(\epsilon_t) = \sigma_e^2$. The parameters $\beta_1, \beta_2, \ldots, \beta_p$ are the AR coefficients (Ihaka 2005).

3.1.3.2. Artificial Neural Network

ANN as a black box model has been applied in numerous number of applications especially in hydrology and water resources. ANN has several types classified according to the the network structure and the learning algorithm; Feed Forward Back Propagation (or commonly known as Feed Forward Neural Network FFNN) being the most widely used type in engineering shown in Figure 3.1. The combination between inputs and outputs of the three-layered FFNNs is linear but nonlinearity is obtained by the transfer and activation functions used in the hidden and output layers. The term feed forward

refers to that the neurons of a layer are connected with neurons of the next layer but not connected with each other.

The expression of a three-layered FFNNs for obtaining an output is (Nourani 2017):

$$\hat{y}_k = f_0\left[\sum_{j=1}^{M_N} w_{kj}.f_h\left(\sum_{i=1}^{N_N} w_{ij}\,x_i + w_{j0}\right) + w_{k0}\right], \tag{3.2}$$

where i, j, and k are the neurons in the input, hidden, and output layers, w_{ij} is the weight of the connection between the neuron i^{th} (i.e. in the input layer), and the neuron j^{th} (i.e. in the hidden layer), w_{j0} is the bias of the j^{th} neuron in the hidden layer, w_{kj} is the weight of the connection between j^{th} and the k^{th} neuron in the output layer, w_{k0} is the bias of the k^{th} neuron, f_h is the transfer function of the hidden layer, f_0 is the activation function of the output layer, x_i is the i^{th} input of the input layer, $\hat{y}_k$ is the computed value, N_N and M_N are the number of neurons in the input and hidden layer. In the back-propagation algorithm, the computed value $\hat{y}_k$ is compared with observed value y (i.e. target in ANN terminology) and the difference (i.e. error) between them back propagated to the network to adjust the weights. This process is called training process.

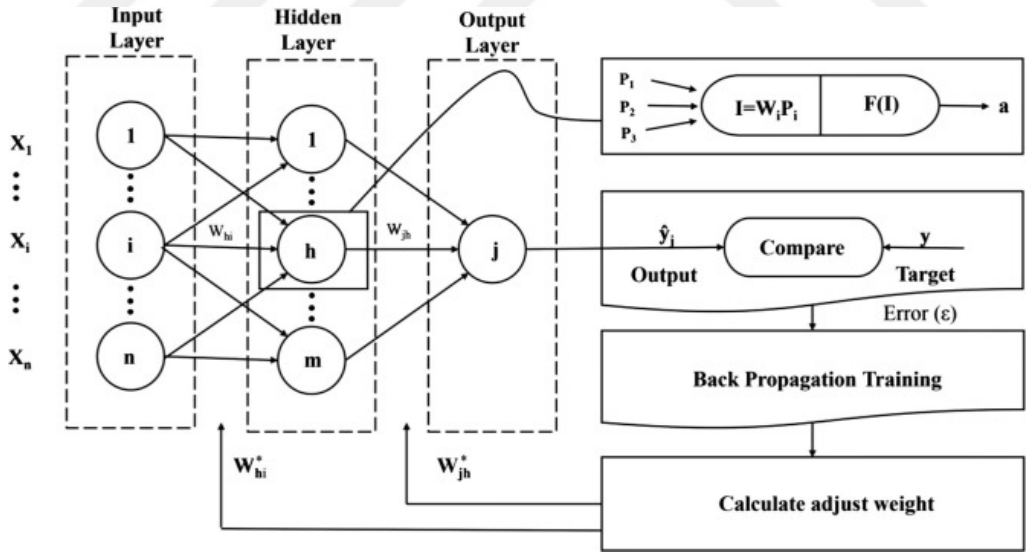


Figure 3.1. *FFNN adapted from (Nourani 2017)*

3.1.3.3. Adaptive-Neuro Fuzzy Inference System (ANFIS)

ANFIS is firstly proposed by (Jang 1993) and further defined by (Jang et al. 1997). ANFIS consists of fuzzy inference system (FIS) combined with Neural Network (NN) algorithm, where NN is used for tuning parameters of the membership functions for

getting better fitting to the training data. The fuzzy inference system can be classified into three classes based on the inference operation in the "if-then rule": Mamdani (Mamdani et al. 1975), Tsukamoto (Tsukamoto 1979), and Sugeno (Takagi et al. 1985). Although mamdani is the most used system with FISs, Sugeno is more efficient, more solid, and computationally intractable in the defuzification operation because its output is crispy and, therefore, more appropriate for adaptive techniques (Sanikhani et al. 2012).

In order to explain the ANFIS framework shown in Figure 3.2, suppose we have two inputs x1 and x2 and one output y. in the case of this study, x1 and x2 represent the flow of first and the second antecedent days Q_{t-1} and Q_{t-2} while y represents the current flow Q_t. The typical two If-Then rules of the first order of Sugeno system are:

Rule1: If x_1 is A_1 and x_2 is B_1 , then $f_1 = p_1x_1 + q_1x_2 + r_1$

Rule2: If x_1 is A_2 and x_2 is B_2 , then $f_2 = p_2x_1 + q_2x_2 + r_2$

The if part is called premise part and the "then" part is called consequent part. A and B are the linguistics terms of the membership function. p, q, and r are the parameters known as consequent parameters.

ANFIS system consists of five layers:

Layer 1: in this layer, inputs are fuzzified and the grade of the membership function is produced using the expression:

$$\begin{cases} For\ x1, O_{1i} = \varphi_{Ai}(x) \\ For\ x2, O_{1i} = \varphi_{Bi}(x) \end{cases}; i = 1,2, \tag{3.3}$$

where x_i are the inputs to the i^{th} node, O_{1i} is the output, $\varphi_{Ai}(x)$ is the membership function. Gaussian function is one of the most common used function which is adapted in this study:

$$\varphi_{\mathrm{A}i}(x) = e^{\frac{-(x-c_i)^2}{2\sigma_i^2}}, \tag{3.4}$$

where c_i and σ_i are the function parameters known as premise parameters.

Layer 2: this nodes in this layer labelled as $\prod$ as its objective is to multiply the coming signals to produce the weight or what is called rule's firing strength:

$$O_{1i} = w_i = \varphi_{Ai}(x_1).\varphi_{Bi}(x_2); i = 1,2\,. \tag{3.5}$$

Layer 3: the nodes of this layer labelled as N which comes from Normalized. In this layer the weight/firing strength ratio (sometimes called Normalized Firing strength) of the i^{th} rule is calculated.

$$O_{1i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} ; i = 1,2, \tag{3.6}$$

Layer 4: The nodes in this layer are adaptive with node functions. The objective of this layer is that the effect of the i^{th} layer on the output by the multiplication of the weight ratio by the first order polynomial function:

$$O_{4i} = \bar{w}_i f_i = \bar{w}_i (p_i x_1 + q_i x_2 + r_i); i = 1, 2 \ . \tag{3.7}$$

Layer 5: This layer contains consists of one node only labelled as Σ. In this layer, all the coming signals are summed to produce the output:

$$O_{5i} = \sum_{i=1} \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{3.8}$$

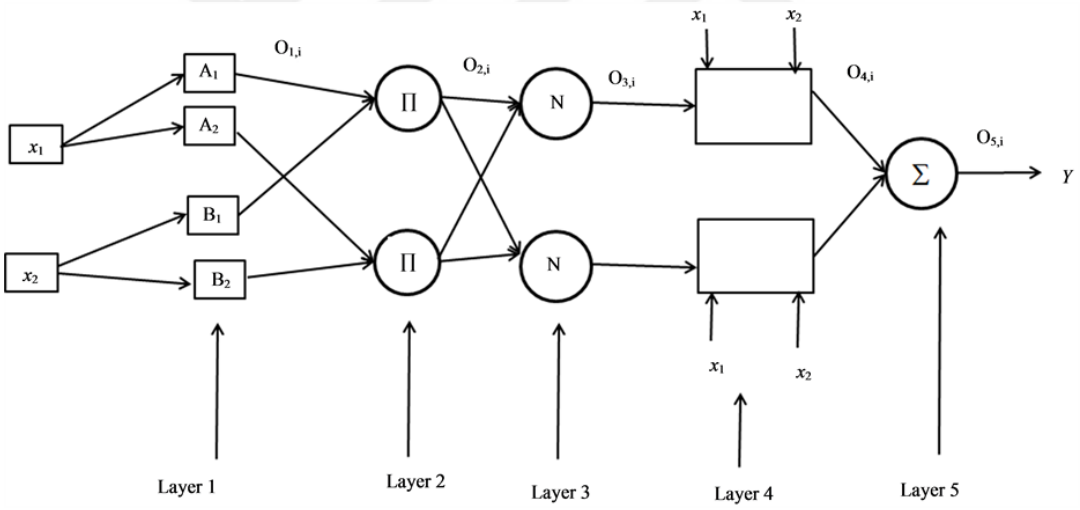


Figure 3.2. *ANFIS framework*

3.1.3.4. Support Vector Machine (SVM):

Support Vector Machine (SVM) is a data-driven machine learning method used for both classification and regression. It is firstly identified by (Vapnik 1992). Support Vector regression (SVR) is commonly used to refer to the regression of SVM. SVM relies on kernel function and therefore, it is considered as a nonparametric technique. The type of SVM that depends on a loss function which ignores the error situated inside a specific distance from the true value is called Epsilon-Intensive-SVM or $\varepsilon - SVM$ (also known

as L1 loss). The main objective of $\varepsilon - SVR$ is to find a function $f(x)$ that have a deviation from y_n not greater than ε for every sample x but as flat as possible at the same time.

Consider having a training dataset consists of a multivariate set $x_n$with an N observation and observed values y_n.

The linear function which is to be found is expressed as:

$$f(x) = x'\beta + b\,. \tag{3.9}$$

It is found with the minimal value of the norm $\beta'\beta$ to ensure that the function is as flat as possible. It is formulated as a convex optimization problem for the reason of minimizing the following:

$$J(\beta) = \frac{1}{2}\beta'\beta\,. \tag{3.10}$$

It is also subject to that all residuals have a value less than ε:

$$\forall n: |y_n - f(x_n)| \leq \varepsilon\,.$$

That is:

$$\forall n: |y_n - (x_n'\beta + b)| \leq \varepsilon\,.$$

Such as function $f(x)$ that satisfies these two constrains for all samples may not be exist. Therefore, slack variables ξ_n, and ξ_n^* are introduced to deal with the infeasible constraints. Like the concept of soft margin in SVM classification, these slack variables allow the regression error to rise up to ξ_n, and ξ_n^*, however, it still meets the conditions See Figure 3.3.

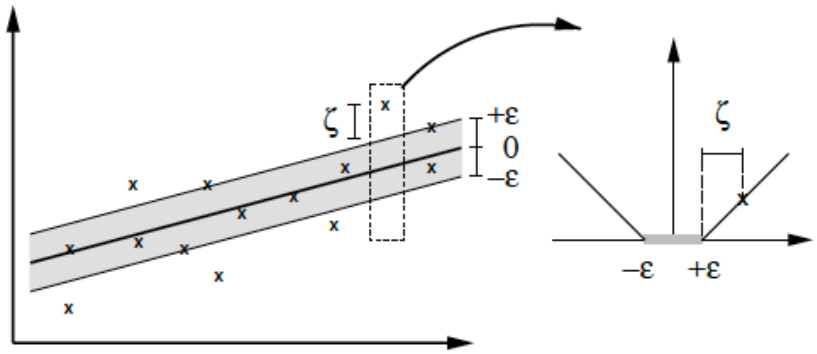


Figure 3.3. *The linear SVM (Smola et al. 2004)*

The objective function (also known as primal function) is obtained by adding the slack variables:

$$J(\beta) = \frac{1}{2}\beta'\beta + C\sum_{n=1}^{N}(\xi_n + \xi_n^*)\,. \tag{3.11}$$

Subject to the constraints:

$\forall n: y_n - (x_n'\beta + b) \leq \varepsilon + \xi_n$

$\forall n: (x_n'\beta + b) - y_n \leq \varepsilon + \xi_n^*$

$\forall n: \xi_n^* \geq 0$

$\forall n: \xi_n \geq 0\,,$

where C is a positive value known as control coefficient that used for regularization to prevent the overfitting, and for controlling the penalty imposed on the values located outside the ε margins. It controls the trade-off between the maximum tolerated error over ε and the flatness of $f(x)$.

The loss function of the linear ε-insensitive treats the errors within ε distance from the observed values as equal to zero. The distance between the observed values and ε margin is the loss in this case:

$$L_\varepsilon = \begin{cases} 0 & if\ |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon & otherwise \end{cases}. \tag{3.12}$$

The solution of the optimization problem of Eq. (3.11) is solved simpler with Langrange dual formulation as the dual problem solution obtains a lower bound to the primal (minimization) problem. The difference between the optimal values of the dual and primal problems, which need not to be equal, is called "duality gap". In the case of having problem that convex and meets the constraints conditions, the optimal solution value is obtained by the solution of the dual problem.

The dual formula is obtained by constructing Langrange function the primal formula by adding two multipliers a_i and a_i^* for every data point x_i. Thus, the dual formula to be minimized is:

$$L(a) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(a_i - a_i^*)\left(a_j - a_j^*\right)x_i'x_j + \varepsilon\sum_{i=1}^{N}(a_i - a_i^*) + \sum_{j=1}^{N} y_i(a_i - a_i^*)\,. \tag{3.13}$$

Subject to:

$\sum_{i=1}^{N}(a_i - a_i^*) = 0$, $0 \leq a_i \leq C$, and $0 \leq a_i^* \leq C$.

The β is described as linear combination of observations of the training set by:

$$\beta = \sum_{i=1}^{N}(a_i - a_i^*)\,x_i\,. \tag{3.14}$$

By this the function of predicting new values which depends on the support vectors only is:

$$f(x) = \sum_{i=1}^{N} (a_i - a_i^*)\ (x_i' x) + b\ . \tag{3.15}$$

To obtain b, The Karush-Kuhn-Tucker (KKT) complimentary conditions are required:

$\forall i: a_i\ (\varepsilon + \ \xi_i - y_i + x_i'\beta + b) = 0$

$\forall i: a_i^*\ (\varepsilon + \ \xi_i^* + y_i - x_i'\beta - b) = 0$

$\forall i:\ \xi_i\ (C - a_i) = 0$

$\forall i: \xi_i^*\ (C - a_i^*) = 0\ .$

According to these constraints, a_i and a_i^* are strictly zero for all observations inside the ε tube. If one of a_i and a_i^* is not zero, the corresponding value is called support vector.

Till this point, the SVM is a linear technique but the real applications especially the hydrologic problems are mostly nonlinear. The nonlinear SVM can be obtained by easily by replacing the dot product $x_i' x_j$ with the nonlinear kernel function. In general, Kernel function is expressed by:

$$G(x_i, x_j) = < \varphi(x_i), \varphi(x_j) >\ , \tag{3.16}$$

where $\varphi(x)$ is the transformation that maps x to high dimensional space.

The nonlinear SVM dual formula which to be minimized becomes:

$$L(a) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(a_i - a_i^*)(a_j - a_j^*)\, G(x_i, x_j) + \varepsilon\sum_{i=1}^{N}(a_i - a_i^*) + \sum_{j=1}^{N} y_i(a_i - a_i^*)\ . \tag{3.17}$$

With the same constraints mentioned earlier. The function of the new values predictions becomes:

$$f(x) = \sum_{i=1}^{N} (a_i - a_i^*)\ G(x_i, x_j) + b\ . \tag{3.18}$$

With the same KKT complementary conditions.

In this study, Gaussian kernel used which has a formula:

$$G(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \tag{3.19}$$

where σ is the Gaussian noise level of standard deviation. For the full details of SVM in both types classification and regression see (Gunn 1998, Smola et al. 2004).

3.1.4. Case Study

3.1.4.1. Study area and Data

This study is aimed at comparing the capabilities of ANN, ANFIS, and SVM in one day ahead forecasting the streamflow in comparison with modeling streamflow by using the traditional AR model. The accuracy of these models is evaluated in an area that is located at the Seyhan River Basin in Southwest Turkey. The selected basin comprises three gauging stations, which are located at three basin outlets, where each basin has different physical characteristics. These stations are labeled according to their basin codes and names. The basin codes are 1801, 1805, and 1822, and their names are Göksu–Himmetli, Göksu–Gökdere, and Zamanti Nehri–Fraktin, correspondingly. The basins will be referred by their codes throughout the manuscript. The study area and location of the stations are depicted in Figure 3.4.

Measurements of the daily streamflow of the three stations are collected from the General Directorate of Water Affairs that is part of the Ministry of Forests and Water Affairs from 1973–2000. The data for station 1805 are available only for 1973–1994. The streamflow time series of the three basins are demonstrated in Figure 3.5.

The basins considered in this study display different physical properties, especially in the area, elevation, and slope (Table 3.1). These factors are considered in the modeling because these factors affect the accuracy of the models. The areas considered are 2400, 1788, and 3610 for 1801, 1805, and 1822, respectively. Each of these basins has different variations in the elevation, although the basins have similar maximum elevation. The elevation varies from 667 to 2880 in basin 1801, 323 to 2800 in basin 1805, and from 1255 to 2800 in basin 1822. Among these basins, 1805 has the smallest area but has an extremely high streamflow variance (i.e., 4155) compared with the two other basins (i.e., 774 and 252 for basins 1801 and 1822, correspondingly). Basin 1805 has the smallest area but has the highest slope at 22°, while the lowest slope at 11° for basin 1822. Basin 1801 has the middle values with 15°. Basin 1805 not only accumulates water but also receives the water that accumulated in the adjacent basin 1801 as the outlet of basin 1801 pours into basin 1805.

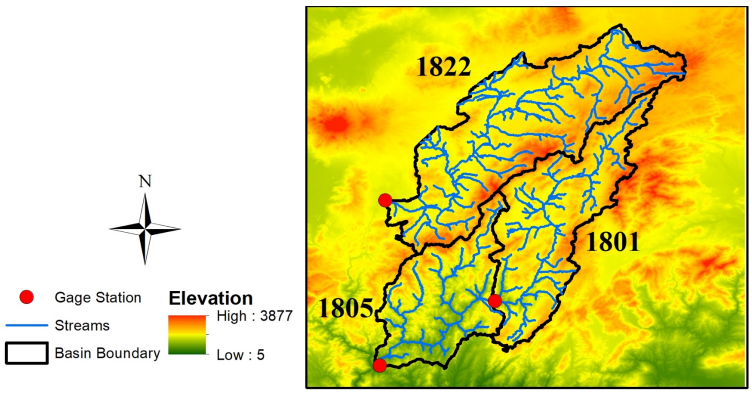


Figure 3.4. *Study area and location of gage stations*

Table 3.1. *Statistics of the daily streamflow and physical characteristics of the basins*

Basin	Period	Area (km^2)	Min. Elevation (m)	Max. Elevation (m)	Slope (Degree)	Min (m^3/s)	Max (m^3/s)	Range (m^3/s)	Median (m^3/s)	Mean (m^3/s)	Var.	Coef.var
1801	1973-2000	2400	667	2880	15	4.8	536	531.2	19	29.3	774	0.9
1805	1973-1994	1788	323	2800	22	14.3	1273	1258.7	37	60.6	4155	1.1
1822	1973-2000	3610	1255	2800	11	3.4	156	152.6	14.1	20.3	252	0.8

3.1.4.2. Model Development and Structure

First of all, normalization applied for the data used using

$$Q_n = \frac{Q - Q_{min}}{Q_{max} - Q_{min}}, \tag{3.20}$$

where Q_n is the normalized flow, Q_{min}= the minimum value of the flows, and Q_{max} is the maximum value of the flows. The data for all models are divided into three subsets: training, validating, and testing subsets. The training subset is used only to train the model, while the validation and testing subsets are used for evaluating the models.

Cox-Box transformation is applied and the transformed time series are used in the AR model to meet the normality assumptions of the time series modeling. Two stationary tests (i.e., augmented Dickey-Fuller and Phillip-Perron tests) are also applied to the three stations as an assumption of the AR model. Following the results of the two tests, all stations are found stationary.

The models considered are one-step ahead forecast, which is expressed as

$$M(p) = Q_t = \Gamma(Q_{t-1}, Q_{t-2}, \dots, Q_{t-p}). \tag{3.21}$$

In this study, $M(p)$ represents the model name and number of inputs. For example, $M(4)$ represents the 4th model of the AR (4), and ANN, ANFIS, and SVM indicates four input lags $Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4}$ and the output Q_t to be used as entries.

The autocorrelation function (ACF) and partial ACF (PACF) are used to identify the number of lag in days to be considered in the time series. All lags that have significant correlation are considered though certain lags are minimal.

In terms of the model structure, ANN has structures (In, Ne, and Out), where In represents the number of inputs, Ne is the number of neurons in the hidden layer, and Out is the number of nodes in the output layer. The optimum structure of each model's combination that is obtained by trial and error is listed in Table 3.2. The ANFIS structure used is (trimf, 2), where trimf is the triangle member function, and 2 is the number of members for all combinations. Three parameters for SVM, including the Gaussian kernel used in this study, lead to the forecasting accuracy, including the Gaussian kernel used in this study. These parameters (C, ε, and σ) correspond to control coefficient, epsilon, and Gaussian noise level. These parameters are obtained for each combination by using Bayesian optimization algorithm in which the least error is obtained.

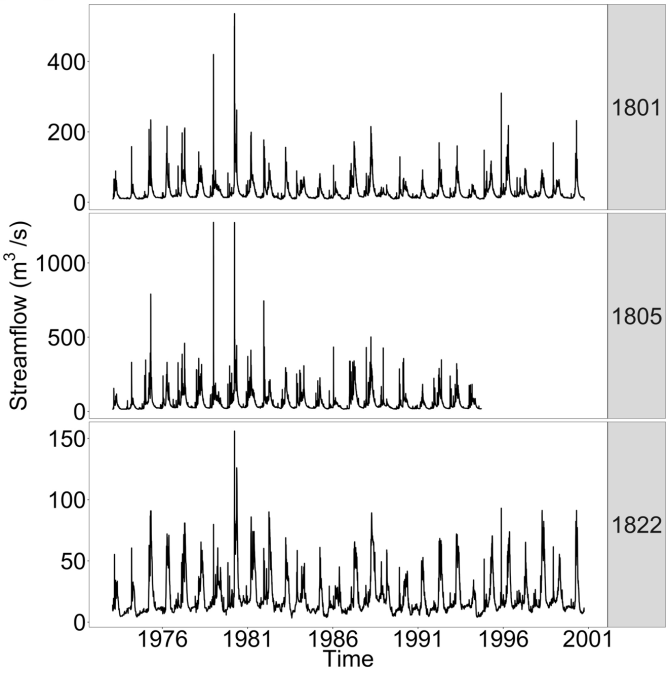


Figure 3.5. *Time series of the Streamflow*

Table 3.2. *The parameters of ANN and SVM for the models*

	1801				1805				1822			
M	**ANN**	**SVM**			**ANN**	**SVM**			**ANN**	**SVM**		
	In-Ne-Out	**C**	**ε**	**σ**	**In-Ne-Out**	**C**	**ε**	**σ**	**In-Ne-Out**	**C**	**ε**	**σ**
1	1-1-1	0.25	8.79E-04	0.14	1-6-1	0.04	4.68E-03	2.76	1-1-1	0.24	3.89E-04	0.60
2	2-4-1	0.73	1.78E-04	0.68	2-14-1	0.36	3.60E-05	1.41	2-5-1	800.95	6.20E-03	0.30
3	3-5-1	2.22	3.88E-04	1.08	3-6-1	0.48	3.22E-05	2.43	3-8-1	0.26	1.54E-04	0.40
4	4-3-1	18.44	3.77E-05	5.90	4-8-1	4.98	1.51E-03	3.32	4-2-1	0.79	1.63E-03	0.79
5	5-5-1	29.46	5.39E-05	5.98	5-8-1	972.49	6.00E-05	120.56	5-3-1	990.87	7.38E-05	72.02
6	6-5-1	154.65	5.13E-04	7.26	6-8-1	0.06	3.39E-05	0.03	6-5-1	0.37	2.04E-04	0.43
7	7-3-1	10.45	7.56E-05	8.82	7-20-1	2.37	3.24E-05	0.04				
8	8-8-1	30.55	2.44E-04	6.28	8-20-1	21.07	3.17E-05	0.41				
9					9-24-1	92.03	2.32E-04	2.87				
10					10-24-1	42.47	3.27E-05	0.43				
11					11-25-1	100.66	3.20E-05	3.41				

3.1.5. Evaluation Criteria

The evaluation and comparison of the models, that is, root-mean-square Error (RMSE), coefficient of efficiency (CE), and index of agreement (IA), are conducted with the mean of two measurements. The RMSE is the square root of the average of the square of the difference between the observed (Q) values and those estimated from the model ($\widehat{\mathrm{Q}}$). The value of RMSE comprises the same unit as the variable (Q) and should be as low as possible. The value is calculated as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\widehat{\mathrm{Q}}_i - Q_i)^2}{N}}, \tag{3.22}$$

where N is the number of the samples.

The CE is also known as the Nash–Sutcliffe coefficient (Nash et al. 1970). This coefficient ranges between $-\infty$ and one; one representing a perfect model, zero means the model is as accurate as the means of the observed data; negative values indicate that the means of the observed data is a better estimator than the model (Dawson et al. 2007). CE is calculated as follows:

$$\mathrm{CE} = 1 - \frac{\sum_{i=1}^{N}\left(\mathrm{Q}_i - \hat{Q}_i\right)^2}{\sum_{i=1}^{N}(Q_i - \bar{Q})^2}, \tag{3.23}$$

where $\overline{\mathrm{Q}}$ is the mean of the observed sample points.

IA can detect the additive and proportional differences in the predicted and observed means and variances (Legates et al. 1999). The IA varies between 0 and 1 with 1 indicating perfect model and 0 indicating no agreement (Willmott 1981). The IA is calculated as follows:

$$\mathrm{IA} = 1 - \frac{\sum_{i=1}^{N}\left(\mathrm{Q}_i - \hat{Q}_i\right)^2}{\sum_{i=1}^{N}\left(\left|\hat{Q}_i - \bar{Q}\right| + |Q_i - \bar{Q}|\right)^2}. \tag{3.24}$$

The accuracy of modeling the peak flood values is assessed using CE_p:

$$\mathrm{CE}_p = 1 - \frac{\sum_{i=1}^{N}(\mathrm{Q}_{pi} - \hat{Q}_{pi})^2}{\sum_{i=1}^{N}(Q_{pi} - \bar{Q}_p)^2}, \tag{3.25}$$

where Q_{pi} represents the peak values, $\hat{Q}_{pi}$ is the fitted values of similar dates of Q_{pi}, and $\bar{Q}_p$ is the mean of the observed peak values.

The CE, RMSE statistics, complexity of the model, and trade-off between the data subsets are used as criteria for evaluating the models rather than using only the highest CE, IA, and lowest RMSE.

3.1.6. Results and Discussion

The ACF and PACF are obtained for each station to identify the number of lags to be considered (Figure 3.6). The ACF exhibits a monotonic descending correlation while the lag increases, while PACF demonstrates a sharp decreasing trend after the first lag starts decreasing. According to (Shumway et al. 2010, Machiwal et al. 2012), the order of the AR lag is identified by using PACF; in particular, the lag is where the correlation terminates. The PACF of the stations cut off at 8, 11, and 6 lags for stations 1801, 1805, and 1822, respectively. The models are considered up to these lags, although the correlation after the first lag is low. Therefore, the number of developed models is equal to the lags considered for each station.

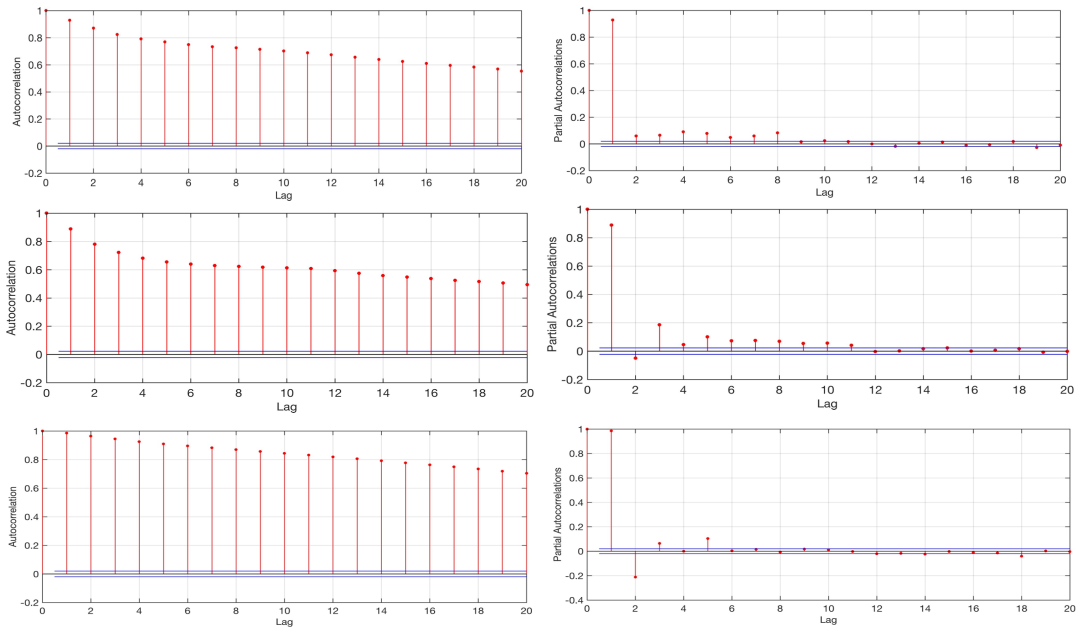


Figure 3.6. *The Autocorrelation and partial autocorrelation functions of the three basins gage stations. from top to bottom: 1801, 1805 an 1822*

Table 3.3 presents that the optimal AR model for basin 1801 is M(7), and M(8) is the least model for validation and test. M(8) is identified as the optimal model that consists of eight lag days that the AR model is AR(8) by using the evaluation criteria. For the same station, ANN displays a variation in the values of the evaluation criteria with optimum values of validation and testing in M(6), thereby leading directly to the selection of M(6) as the superior model. The ANFIS results indicate an optimal model in validation as M(2), and M(6) and M(2) are selected for testing. M(3) is identified as the superior model for validation and test by using SVM. In general, ANN, SVM, and ANFIS exhibit close results to each other and are all better than AR. The results of model accuracy in modeling the peak values listed in Table 3.4 indicate the superiority of ANN with highest CE, IA, and lowest RMSE. The AR has the lowest accuracy. Figure 3.7 illustrates the scatter plots of the observed versus the predicted values for each model. The AR model estimates the low streamflow values with high accuracy but under and overestimated the moderate and high values. ANN, ANFIS, and SVM generally underestimate the moderate and high values, except SVM, which overestimates several moderate values.

Table 3.3. *The results of the 1801 basin for the different models*

AR							ANN						
	CE		IA		RMSE (m^3/s)			CE		IA		RMSE (m^3/s)	
M	Val.	Test	Val.	Test	Val.	Test	M	Val.	Test	Val.	Test	Val.	Test
1	0.945	0.881	0.985	0.968	5.512	8.267	1	0.945	0.885	0.986	0.970	5.478	8.131
2	0.941	0.816	0.984	0.953	5.718	10.249	2	0.944	0.894	0.986	0.972	5.550	7.797
3	0.945	0.860	**0.986**	0.962	5.507	8.960	3	0.947	0.896	0.986	**0.973**	5.375	7.704
4	0.945	0.868	0.986	0.964	5.503	8.683	4	0.945	0.888	0.986	0.972	5.477	8.021
5	0.945	0.881	0.986	0.968	5.505	8.247	5	0.947	0.895	0.986	0.971	5.381	7.747
6	0.945	0.884	0.986	**0.969**	5.501	8.152	**6**	**0.948**	**0.899**	0.986	0.972	**5.364**	**7.622**
7	**0.946**	0.884	0.986	0.968	5.493	8.152	7	0.947	0.896	**0.987**	0.972	5.418	7.731
8	0.946	**0.885**	0.986	0.969	**5.490**	**8.123**	8	0.948	0.896	0.985	9.973	5.355	7.716
			ANFIS							SVM			
1	0.942	0.881	0.984	0.967	5.652	8.252	1	0.941	0.889	0.984	0.969	5.716	7.975
2	**0.943**	0.880	**0.985**	0.968	**5.631**	8.274	2	0.947	0.890	**0.986**	**0.971**	5.412	7.932
3	0.939	0.879	0.984	0.968	5.810	8.303	**3**	**0.948**	**0.894**	0.986	0.971	**5.346**	**7.771**
4	0.940	0.886	0.984	0.970	5.785	8.078	4	0.946	0.885	0.985	0.969	5.485	8.124
5	0.942	0.891	0.985	0.971	5.671	7.904	5	0.946	0.886	0.986	0.969	5.465	8.064
6	0.939	**0.893**	0.984	**0.972**	5.826	**7.827**	6	0.947	0.888	0.986	0.969	5.439	7.998
7	0.936	0.882	0.983	0.969	5.959	8.228	7	0.944	0.882	0.985	0.967	5.578	8.202
8	0.931	0.871	0.982	0.967	6.159	8.593	8	0.946	0.886	0.985	0.969	5.477	8.076

Bold values are the highest for CE and lowest for RMSE

Table 3.4. *The Peak flood values forecasting results for the three basins*

Station / Model	1801			1805			1822		
	CE	**IA**	**RMSE (m^3/s)**	**CE**	**IA**	**RMSE (m^3/s)**	**CE**	**IA**	**RMSE (m^3/s)**
AR	0.659	0.904	22.14	0.482	0.845	63.35	0.837	0.959	6.26
ANN	**0.735**	**0.920**	**19.54**	**0.583**	0.872	**56.88**	**0.889**	**0.971**	**5.17**
ANFIS	0.669	0.898	21.81	0.508	0.847	61.75	0.888	**0.972**	5.19
SVM	0.693	0.905	21.01	0.577	**0.880**	57.31	0.887	0.971	5.22

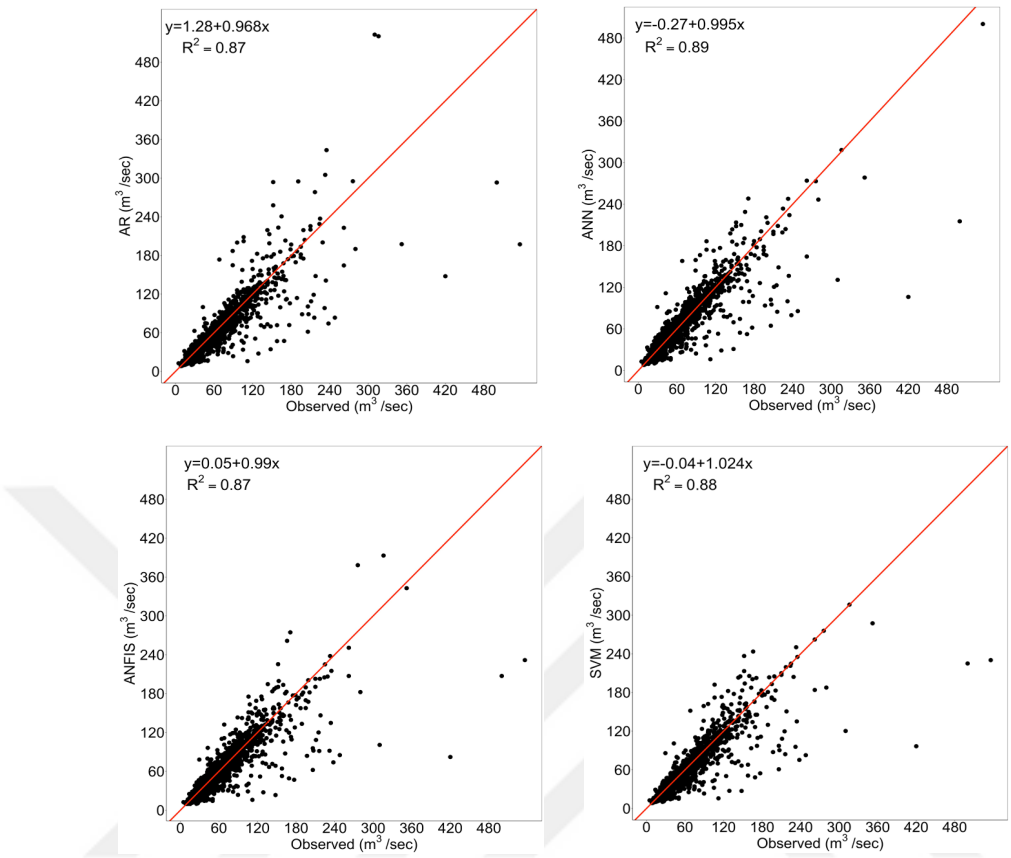


Figure 3.7. *Daily streamflow observed vs estimated of the 1801 station. Red lines represent the best fit line*

The application of the same evaluation criteria on the results of station 1805 shown in Table 3.5, M(8), M(3), M(3), and M(8) is selected as the optimal model for AR, ANN, ANFIS, and SVM. The ANN records the highest accuracy, whereas the AR indicates the lowest accuracy in forecasting the peak values (Table 3.4). The behavior of each method is similar to basin 1801 and that AR has a general underestimation. ANN, ANFIS, and SVM have general underestimation of the high moderate to high values (Figure 3.8).

The results of station 1822 indicate that M(1), M(3), M(5), and M(3) are the optimal models for AR, ANN, ANFIS, and SVM, correspondingly (Table 3.6). ANN is identified as the most accurate method with highest CE and IA and lowest RMSE, although the results from ANN and ANFIS are close in forecasting the peak flood values (Table 3.4). The AR, similar to the previous two stations, has the lowest accuracy. The accuracy of forecasting streamflow in the outlet of this basin is high. Hence, all the methods

demonstrate minimal over and underestimation throughout the whole values, regardless of being low, moderate, or high (Figure 3.9).

Table 3.5. *The results of the 1805 basin for the different models*

AR							ANN						
	CE		IA		RMSE (m^3/s)			CE		IA		RMSE (m^3/s)	
M	Val.	Test	Val.	Test	Val.	Test	M	Val.	Test	Val.	Test	Val.	Test
1	0.830	0.800	0.953	0.944	25.417	21.935	1	0.838	0.800	0.954	0.945	24.718	21.969
2	0.820	0.764	0.952	0.936	26.178	23.839	2	0.834	0.809	0.955	0.949	25.030	21.480
3	0.848	0.776	0.958	0.938	24.011	23.247	**3**	**0.848**	0.812	**0.959**	**0.950**	**23.933**	21.316
4	0.851	0.793	0.959	0.943	23.771	22.316	4	0.832	0.807	0.959	0.950	25.119	21.611
5	0.852	0.797	0.959	0.943	23.702	22.117	5	0.830	0.807	0.954	0.948	25.276	21.576
6	0.852	0.800	0.959	0.944	23.703	21.937	6	0.823	0.810	0.949	0.946	25.842	21.394
7	0.854	0.802	**0.960**	0.945	23.527	21.848	7	0.829	0.803	0.954	0.947	25.399	21.795
8	**0.856**	0.804	0.960	0.945	**23.403**	21.730	8	0.817	**0.816**	0.950	0.949	26.246	**21.070**
9	0.856	**0.805**	0.960	0.945	23.405	21.704	9	0.823	0.804	0.954	0.948	25.852	21.780
10	0.856	0.805	0.960	0.945	23.362	21.700	10	0.831	0.813	0.955	0.950	25.240	21.272
11	0.856	0.805	0.960	**0.946**	23.370	**21.679**	11	0.820	0.807	0.952	0.949	26.039	21.604
ANFIS							SVM						
1	0.829	**0.798**	0.951	0.941	25.396	**22.051**	1	0.731	0.725	0.896	0.892	31.882	25.753
2	0.829	0.791	0.954	0.944	25.376	22.448	2	0.828	0.790	0.953	0.942	25.463	22.514
3	**0.833**	0.795	**0.955**	0.944	**25.123**	22.216	3	0.831	0.802	0.952	0.944	25.224	21.865
4	0.828	0.797	0.953	**0.945**	25.506	22.137	4	0.832	0.798	0.953	0.943	25.205	22.045
5	0.790	0.780	0.944	0.941	28.144	23.020	5	0.828	0.803	0.950	0.942	25.464	21.791
6	0.781	0.777	0.941	0.940	28.759	23.203	6	0.749	0.726	0.913	0.910	30.804	25.716
7	0.761	0.779	0.936	0.940	30.010	23.105	7	0.762	0.447	0.932	0.858	29.981	36.500
8	0.783	0.786	0.943	0.943	28.594	22.732	**8**	0.801	**0.820**	0.946	**0.950**	27.421	**20.809**
9	0.791	0.784	0.944	0.943	28.101	22.806	9	**0.836**	0.795	**0.954**	0.943	**24.912**	22.252
10	0.768	0.779	0.937	0.940	29.611	23.072	10	0.779	0.817	0.939	0.947	28.918	20.997
11	0.766	0.777	0.937	0.939	29.698	23.195	11	0.833	0.793	0.954	0.943	25.080	22.334

- Bold values are the highest for CE and lowest for RMSE

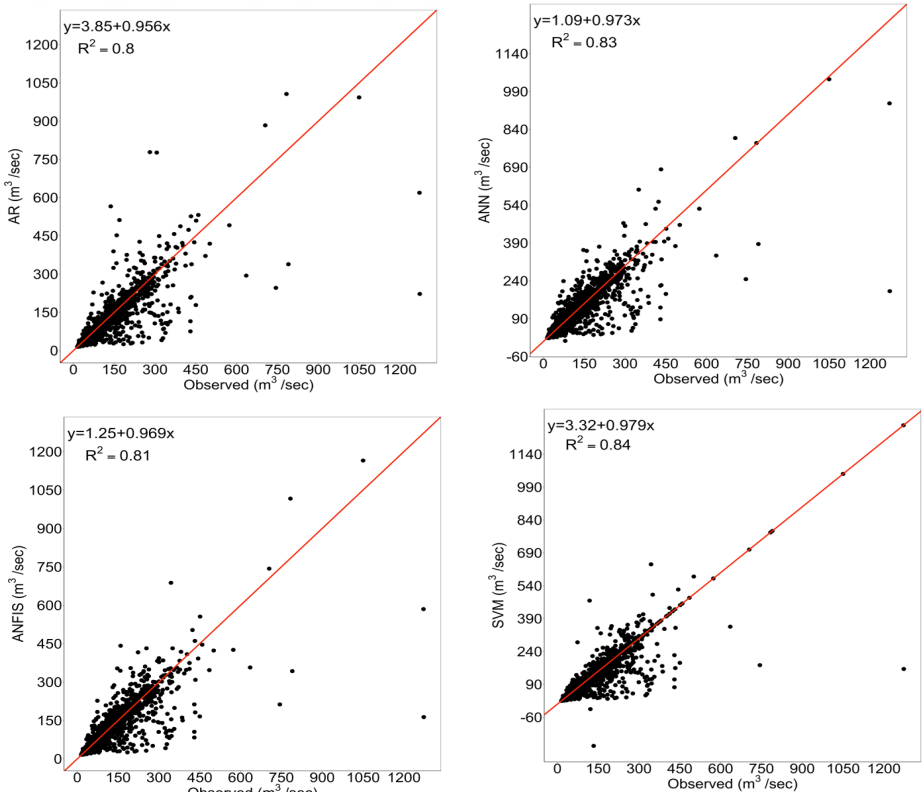


Figure 3.8. *Daily streamflow observed vs estimated of the 1805 station. Red lines represent the best fit line*

Table 3.6. *The results of the 1822 basin for the different models*

AR							ANN						
	CE		IA		RMSE (m^3/s)			CE		IA		RMSE (m^3/s)	
M	Val.	Test	Val.	Test	Val.	Test	M	Val.	Test	Val.	Test	Val.	Test
1	0.980	**0.974**	**0.995**	**0.993**	1.939	**2.509**	1	0.980	0.974	0.995	0.993	1.934	2.509
2	**0.982**	0.971	0.995	0.993	1.873	2.618	2	0.982	**0.978**	0.995	**0.994**	1.822	2.301
3	0.982	0.973	0.995	0.993	1.870	2.523	**3**	**0.983**	0.978	**0.996**	0.994	**1.804**	**2.276**
4	0.982	0.973	0.995	0.993	1.870	2.534	4	0.982	0.978	0.996	0.994	1.825	2.310
5	0.982	0.973	0.995	0.993	**1.866**	2.525	5	0.983	0.978	0.995	0.994	1.811	2.308
6	0.982	0.973	0.995	0.993	1.868	2.523	6	0.982	0.978	0.996	0.994	1.815	2.313
			ANFIS							SVM			
1	0.980	0.974	0.995	0.993	1.937	2.501	1	0.981	0.974	0.995	0.993	1.922	2.506
2	0.982	0.976	0.995	**0.994**	1.848	2.401	2	0.982	0.977	0.995	0.994	1.866	2.340
3	0.982	0.975	**0.996**	0.994	1.836	2.420	**3**	**0.983**	**0.979**	**0.996**	**0.995**	**1.803**	**2.252**
4	0.982	**0.977**	0.995	0.994	1.848	**2.358**	4	0.983	0.978	0.996	0.994	1.811	2.272
5	**0.983**	0.977	0.996	0.994	**1.823**	2.363	5	0.982	0.975	0.995	0.994	1.848	2.434
6	0.982	0.976	0.996	0.994	1.829	2.397	6	0.983	0.978	0.996	0.994	1.796	2.284

- Bold values are the highest for CE and lowest for RMSE

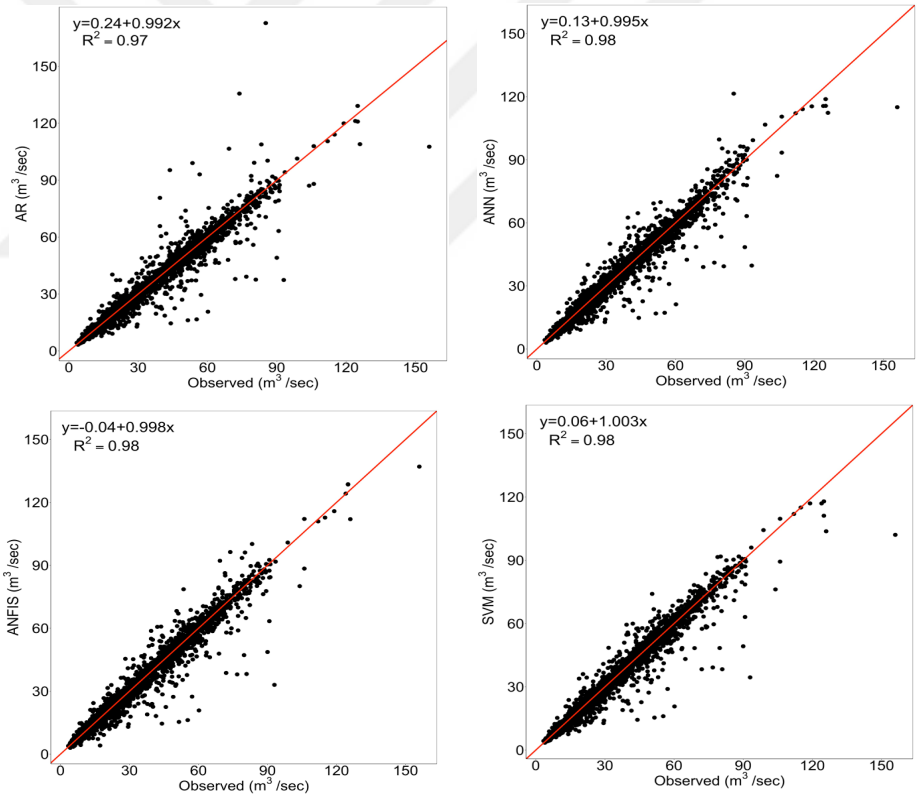


Figure 3.9. *Daily streamflow observed vs estimated of the 1822 station. Red lines represent the best fit line*

In the three basins, ANN and SVM have close accuracy results, while ANN outperforms AR and ANFIS. In terms of the forecasting of peak values, ANN and SVM exhibit the highest accuracy, the lowest for AR, and ANFIS being in the middle. The forecasting accuracy among all methods is highest for basin 1822 using SVM. Basin 1822

has the largest area, smallest variation in the elevation, smallest streamflow variance, and smallest slope with CE 0.983 and 0.979, IA 0.996 and 0.995, and RMSE 1.803 and 2.252 (m^3/s) for validation and test, respectively. The lowest accuracy is obtained in basin 1805, which has the smallest area, largest variation in the elevation, largest streamflow variance, and largest slope with CE 0.848 and 0.812, IA 0.959 and 0.950, and RMSE 23.933 and 21.316 (m^3/s) for validation and test, correspondingly. Superior results for basin 1805 are obtained using ANN. Moreover, basin 1801 has values between the largest and smallest in all the characteristics that stayed in the middle with CE 0.948 and 0.899, IA 0.986 and 0.972, and RMSE 5.364 and 7.622 (m^3/s) for validation and test obtained by using ANN.

The accuracy of peak value forecasting is similar to all streamflow values. ANN still obtained the highest accuracy in basin 1822 with 0.889, 0.971, and 5.17 m^3/s for CE, IA, and RMSE, respectively; ANN obtained the lowest basin 1805 by 0.583 CE, 0.872 IA, and 56.88 m^3/s RMSE, respectively. Moreover, basin 1801 falls in the middle with 0.735 CE, 0.920 IA, and 19.54 (m^3/s) RMSE. The behavior of all methods in terms of the estimation of low or high values in basin 1801 and 1805 is similar, but basin 1822 is reduced regardless of the value. These results and differences show a robust relationship between the models and physical characteristics of the basins. The variations in the physical characteristics of the basins, such as the elevation, slope, and area, significantly affect the performance of all methods. The increase in various elevations and the slope deteriorate performance and decrease the area, thereby causing the deterioration. This result indicates that the physical attributes of the basins should also be considered in the streamflow modeling.

3.1.7. Conclusion

The daily streamflow data from the gauge stations of three basins are used in this study to compare the performance of modeling three data-driven methods, namely, ANN, ANFIS, and SVM versus the linear AR model. The comparison is conducted to investigate the effect of the basin's physical characteristics on the model performance. The ACF and PACF are used to identify the required number of lags. The evaluation of the models is achieved using several criteria, that is, the values of CE, IA, and RMSE, trade-off in the performance of a model in the two data subsets (i.e., validation and test subsets), and the complexity of the model represented by the number of inputs to the model.

The ANN indicates the highest performance among the four methods. However, in some cases, the ANN and SVM are indicated to obtain close results. Simultaneously, the SVM performed better than AR and ANFIS when the latter indicates the least performance among all stations. The ANN outperformed all methods in forecasting peak flood values with the close result to SVM, which is far from the least performance obtained by AR, which is less than ANFIS.

The physical characteristics, such as small area, high variation in elevation, and high slope, significantly deteriorate the model performance, although the methods are data-driven models. These physical characteristics can support the evidence of the deterioration or improvement in the forecasting accuracy regardless of the applied data-driven model. Hybrid data-driven models are recommended, especially for forecasting peak flood values, to increase the accuracy of forecasting. The increase in the number of basins that are used to conduct a statistical comparison between the characteristics and accuracy of the methods are also recommended.

CHAPTER FOUR

4. IMPROVING STREAMFLOW FORECASTING USING WAVELET

4.1. Streamflow Forecasting Using Four Wavelet Transformation Combinations Approaches with Data-Driven Models: A Comparative Study

4.1.1. Abstract

The importance of short-term and long-term forecast of the streamflow in the efficient management of the water resources is important. Streamflow forecasting plays a crucial role in the mitigation of the excess or lack influence on the water system. Therefore, the proposed study herein investigated the use of wavelet transformation (WT) as preprocessing tool in data-driven models (DDMs) for forecasting streamflow seven days ahead. Continuous wavelet transformation (CWT), discrete wavelet transformation (DWT), and the proposed discrete continuous wavelet transformation (DCWT) in addition to the single DDMs were used to create four different schematic layouts. The DDMs applied were artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), and support vector machines (SVM). The lagged rainfall, temperature, and streamflow were incorporated as inputs into the models. Data collected over the seventeen-year period from February, 1973 and September, 2000. The results indicated that the use of WT improved the performance of the single DDMs. In addition, it was found that CWT improved the forecasting accuracy of models which only included the rainfall and temperature but not the streamflow. Moreover, DWT improved the models dramatically particularly the models with streamflow. Notably, DWT layout outperformed CWT layout in general despite the fact that CWT layouts resulted in higher improvement to the models with rainfall and temperature only. The proposed DCWT in which CWT applied on the rainfall and temperature variables and DWT applied on the streamflow improved the forecasting ability in several models when ANN was applied. Nevertheless, improvement in the forecasting accuracy was deteriorated in those with SVM while no improvement was observed with ANFIS. ANN outperformed both ANFIS and SVM while ANFIS performed better than SVM.

4.1.2. Introduction

The hydrologic modeling in general and streamflow in particular, has grown rapidly in recent years. Management of the water resources systems including planning, designing, and operating requires short-term and long-term streamflow forecasting (Kişi

2008, Sanikhani et al. 2012). Accurate streamflow forecasting allows for efficient water resources management due to its role in the mitigation of influences of the excess or scarcity of the water systems (Wood et al. 1980, Sanikhani et al. 2012). Models for streamflow can be divided into three main categories including physical, mathematical, and empirical which also known as data-driven models (DDMs). Physical and mathematical models which also known as white box models involve the physical interactions between the variables in modelling the hydrologic processes. In contrast, DDMs or black box models obtain the optimum solution between the inputs and outputs without involving the physical process by selectively choosing the inputs and outcomes to interpret their physical contributions (Solomatine et al. 2008). DDMs are widely used as they are inexpensive and requires fewer inputs compared to the physical models.

Data-driven approaches such as artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), support vector machines (SVM), and genetic programming (GP) have been used widely for modelling of hydrological systems Nonetheless, utilization of the non-stationary data is the main disadvantage of the data-driven models (Nourani et al. 2009, Kisi et al. 2012, Abdollahi et al. 2017). Thus, the wavelet transformation (WT) is one of the solutions to overcome the non-stationarity issue. The transformation is performed by two forms namely continuous wavelet transform (CWT) and discrete wavelet transform (DWT). Wavelet is a mathematical technique that cut up the time series into various frequencies (Lee et al. 1994) to uncover the hidden information of the data and produce larger data than the original time series (Shoaib et al. 2014).

WT has been applied for modelling the daily streamflow to preprocess the data to be then entered to the DDMs. This includes wavelet-based ANN, (Kişi 2009, Adamowski et al. 2010, Pramanik et al. 2011, Nourani et al. 2013, Nourani et al. 2014, Alizadeh et al. 2017, Nury et al. 2017), wavelet-based ANFIS (Shiri et al. 2010, Nourani et al. 2011, Badrzadeh et al. 2013), and wavelet-based SVM (Adamowski et al. 2012, Komasi et al. 2016).

There are very limited number of studies comparing more than two approaches as most of the studies in the literature compares the approach itself with the approach conjugated with WT. For example, (Kişi 2009) compared one approach that ANN with WT-ANN. Another example of comparing two approaches (Abdollahi et al. 2017) who compared ANN and Genetic Expression Programming (GEP) with WT-ANN and WT-

GEP. Also, there are very limited number in the entire hydrology forecasting (Sanikhani et al. 2012) not only in the daily streamflow forecasting studies using ANFIS combined with subclustering. This trigger the need of a comparison of more approaches in one place that eases the exploration of the approaches behavior under several circumstances.

Several researchers used CWT such as (Nakken 1999, Cannas et al. 2006, Ren et al. 2013), however, according to (Nourani et al. 2014), till the date of the reviewer paper publishing, not more than 20% of the studies in all the hydrological applications used Continuous Wavelet Transformation (CWT), and the majority used Discrete Wavelet Transformation (DWT). This percentage is much less in the forecasting of the streamflow. This is due to the nature of the hydrologic time series which gathered not in a continuous but in Discrete form. Therefore, DWT which is dyadic is more appropriate for decomposing the time series which allows for reconstructing the original time series from the sub-signals. Also, CWT, in the time domain, represents the time-frequency of the signal at many number of time and exact periods that at every scale there is a coefficient for every time step in the time series. This provides a redundant information which is costly in terms of the calculations. Nowadays with the advancement in the technology, the speed of these calculations is not a problem anymore. The use of CWT in hydrological applications was suggested by (Nourani et al. 2014). (Shoaib et al. 2014) examined the effect of the choice of the mother function and the decomposition level (i.e. in DWT) of the wavelet by applying DWT and CWT with Multilayer Perceptron Artificial Neural Network (MLPANN) Radial Basis Function Neural Network (RBFNN) in rainfall-runoff modelling. They concluded that applying the WT is increasing the performance of the models. However, this study applies CWT on the rainfall time series only for forecasting runoff. Accordingly, a question can be raised here what if CWT applied on the rainfall, DWT applied on the streamflow and gathered in one model? Will the model's performance increase?

A study concluded that wavelet-based model is more accurate for shorter lead-time (one and three days) but less efficient for longer lead-time (seven days) for both river flow and flood forecasting (Adamowski 2008). Additionally, with WT, daily streamflow forecasting accuracy was improved for shorter lead-time (one and two days) compared to the longer lead-time (six days). In contrast, coupling DWT with a second-order volterra (WVC) yielded better performance in the streamflow forecasting for longer lead-time (Maheswaran et al. 2013). Hence, as the longer lead-time (E.g 7 days) affects the accuracy

of the streamflow forecasting, it is crucial to choose an appropriate lead time at the daily timescale (Nourani et al. 2014).

In view of the above, the primary aim of this study was to investigate the robust modelling approaches of WT-DDMs combinations, especially for the long-time streamflow forecasting, through assessing both modelling accuracy and performance. In this study, WT included were DWT, CWT and DCWT, while DDMs used were ANN, ANFIS and SVM. In addition, combination of the data-driven modelling and wavelet transformation approaches were used and compared using data collected from the basin in Goksu-Himmeti, Turkey. This study used seven days lead-time.

4.1.3. Methodology

4.1.3.1. Wavelet Transformation (WT)

A wavelet is defined as a limited duration small wave with a zero-average value. Wavelet is a mathematical function or tool that decomposes the time series into different frequency components in order to capture the characteristics of the time series that has the advantage of overcoming the issue of the non-stationarity faced by the DDMs (Shoaib et al. 2014, Alizadeh et al. 2017). A family of wavelet has two orthogonal function: father or scaling function $\emptyset(\mathrm{t})$ and the mother or wavelet function $\psi(t)$.

Time series data transformation is done either by translation or by scaling or dilating. Translation is moving the wavelet along the axis of the time at different positions while scaling or dilation is squeezing or stretching the wavelet (Percival et al. 2006). Generally, WT divided into two forms: continuous wavelet transform (CWT) and discrete wavelet transform (DWT). The CWT is used for revealing the characteristics of the series in multi-temporal scales. The continuous wavelet coefficient $W_f(a,b)$ is defined as:

$$W_f(a,b) = \int_{-\infty}^{+\infty} f(t)\, \psi^*_{a,b}(t)dt\,, \tag{4.1}$$

where:

$$\psi_{a,b}(t) = |a|^{-1/2}\psi\left(\frac{t-b}{a}\right) \quad a,b \in R, a \neq 0, \tag{4.2}$$

where * represents the complex conjugate of the function, a and b are the parameters known as the scale and translation parameters respectively. The coefficients of the CWT are calculated in each a (scale) and b (translation) which generate a large amount of data.

DWT is the solution for this issue because it uses scaling function and wavelet function known as low pass filter and high pass filter respectively. The DWT is defined by:

$$W_f(j,k) = \int_{-\infty}^{+\infty} f(t)\, \psi_{j,k}^{*}(t)dt, \tag{4.3}$$

where:

$$\psi_{j,k}(t) = a_0^{-j/2}\psi\left(a_0^{-j}t - b_0k\right), \tag{4.4}$$

where a_0 and b_0 are the constants, j is the decomposition level, and k is the factor of time translation.

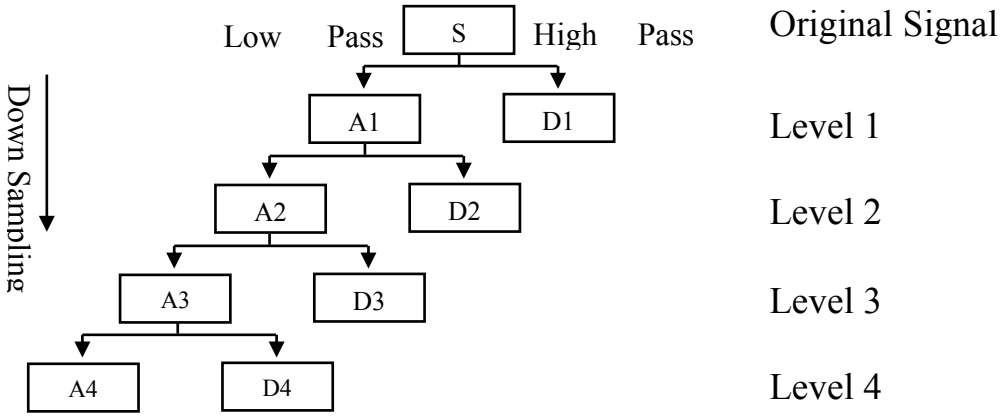
Frequently used DWT is the dyadic DWT by assigning $a_0 = 2$ and $b_0 = 1$ and that leads to the function of Eq. (4.5) (Daubechies 1992):

$$\psi_{j,k}(t) = 2^{-j/2}\psi\left(2^{-j}t - k\right). \tag{4.5}$$

In the first stage, using dyadic DWT, the signal S passed through the high and low pass filters resulting in two types of coefficients sets: approximation produced by low pass filter, and detail produced by high pass filter. In the subsequent levels, only the approximation coefficients obtained from the previous level were analyzed (Figure 4.1). This process is called downsampling as only one data point out of two kept in both filters. The length of the signal or time series controls the maximum possible number of the levels but the desired level can be identified. This is one of the most important aspects of using DWT with DDMs (Sang 2013, Nourani et al. 2014). In addition to the decomposition level, the wavelet has another key issue in Eq. (4.5) that it must meet the regularity condition in Eq. (4.6) with N order of regularity:

$$\int_{-\infty}^{+\infty} t^k\, \psi(t)dt = 0, k = 1, \dots, N-1, \tag{4.6}$$

There are six wavelets families meet the condition namely ReverseBior (rbioM.N), Symlets (symN), Daubechies (dbN), DMeyer (dmey), BiorSplines (biorM.N), and Coiflets (coifN). Furthermore, there are several wavelet families that mainly used only for CWT as they failed to meet the regularity condition namely Meyer (meyr), Mexican_hat (mexh), Morlet (morl), and Gaussian (gaus) (Sang 2013). For more details about DWT, readers can refer to the work of (Mallat 1989, Daubechies 1992, Addison 2002).



A: The approximation coefficients, D: The detail coefficients

Figure 4.1. *The discrete wavelet decomposition framework*

4.1.3.2. ANFIS – Sub Clustering (ANFIS-SC)

Grid partitioning model is a commonly used ANFIS model. In this model, each input variable is divided into a number of membership functions and each of the membership function has a rule created with all membership function of the same and other variables. Therefore, the number of all rules increases exponentially with the number of variables. By having n number of variables and m number of membership functions for each input variable, the number of rules is equal to m^n. For example, in one of the models' combinations used in this study, there were 32 input variables with two membership functions, equals 2^{32} which produce a large number of rules. Therefore, the use of ANFIS with the subclustering model which is superior to the ANFIS-GP (Sanikhani et al. 2012) was used in this study.

SC method is an extension of mountain clustering method (Yager et al. 1994). Combination of SC and the fuzzy methods provides a fast and robust algorithm in any numerical data. Two main advantages of this method are the elimination requirement to the grid resolution where the trade-off between the accuracy and the complexity is a must and the equality of the effective grid points to be evaluated to the data points independently from the problematic dimension.

It is assumed that a collection of n data points $\{x_1, x_2, x_3 \dots, x_n\}$ exist in an m dimensional space. Then, the data points are assumed as normalized in each dimension, without losing the generality, the ranges of their coordinates are equal in each dimension which denotes that the data points are bounded by a hypercube. Starting by considering

each data point as a potential cluster center and the potentiality of any data point x_i is measured by:

$$P_i = \sum_{j=1}^{n} e^{-a\|x_i - x_j\|^2}, \tag{4.7}$$

where

$$a = \frac{4}{r_a^2}, \tag{4.8}$$

where $r_a$is a positive constant and $\|x_i - x_j\|$ is the Euclidean distance.

After calculation of the potential of each data point, the point data with the highest potential is selected as the first cluster center. Assigning x_1^* to the location of the first cluster center and P_1^* to its potential value, the potential of every data point x_i is then revised by:

$$P_i \Leftarrow P_i - P_1^* \, e^{-\beta\|x_i - x_1^*\|^2}, \tag{4.9}$$

where

$$\beta = \frac{4}{r_b^2}, \tag{4.10}$$

where r_b is a positive constant known as squash value or squash factor (Ping et al. 2006). According to Eq.(4.9), the amount of potential of the cluster center is subtracted from the potential of each data as the function of its distance from the first cluster center. The potentiality of the data points near the cluster center are reduced which lower their probability to be chosen as the next cluster center. Squash factor r_b , is a radius which identifies the extent in which the reduction in the potentials is measurable (Chiu 1994). The study suggested using value of r_b greater than r_a and recommended a relation between them r_b = 1.5 r_a.

After revision of all data points' potentials, the point with the highest remaining potential is selected as the second cluster center. Then, the potential of every data point is reduced according to their distances from the second cluster center. Generally, after obtaining k^{th} cluster centers, the potentials are revised using:

$$P_i \Leftarrow P_i - P_k^* \, e^{-\beta\|x_i - x_k^*\|^2}, \tag{4.11}$$

where x_k^* and P_k^* are location and the potential value of the k^{th} cluster center respectively. For more details of the method, the readers can refer to (Chiu 1994).

4.1.4. Evaluation of performance

Performance of the models evaluated by the use of two measurements known as root mean square error (RMSE) and coefficient of efficiency (CE).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(\widehat{\mathrm{Q}}_i - Q_i\right)^2}{N}}, \tag{4.12}$$

where N is the number of the samples, Q is the observed values and $\widehat{\mathrm{Q}}$ is the estimated values from the model. RMSE has the same unit of (Q) and should be as low as possible.

The coefficient of efficiency (CE) is also known as Nash-Sutcliffe coefficient (Nash et al. 1970). CE) is calculated as follows:

$$\mathrm{CE} = 1 - \frac{\sum_{i=1}^{N}\left(\mathrm{Q}_i - \hat{Q}_i\right)^2}{\sum_{i=1}^{N}(Q_i - \bar{Q})^2}, \tag{4.13}$$

where $\overline{\mathrm{Q}}$ is the mean of the observed sample points. The values of CE varies between -∞ and one. The value of one denotes a perfect model while the value of zero denotes that the model is not giving different value from the mean of the observed data, and negative values indicate that the mean of the observed data is a better estimator than the model (Dawson et al. 2007).

4.1.5. Study area and Data

A basin named 1801 or Goksu-Himmeti, situated in a large basin namely Seyhan basin was used as the case study area as shown in Figure 4.2. The basin located between 36°1'33"E - 36°40'48"E and 37°40'9"N-38°40'30"N which situated in southwest of Turkey. The basin has area covers a surface of about 2400 km^2 and an elevation varying between 667 – 2880m. The daily temperature and rainfall data were measured by meteorological station for the period of Feb-1973 to Sep-2000. The data was obtained from the general directory of meteorology which is part of the Ministry of forests and water affairs.

The daily streamflow measured at Goksu-Himmeti station located on the outlet of the basin for the same period were obtained from the general directory of Water affairs which is also part of the Ministry of forests and water affairs. The time series of the temperature, rainfall, and streamflow are shown in Figure 4.3.

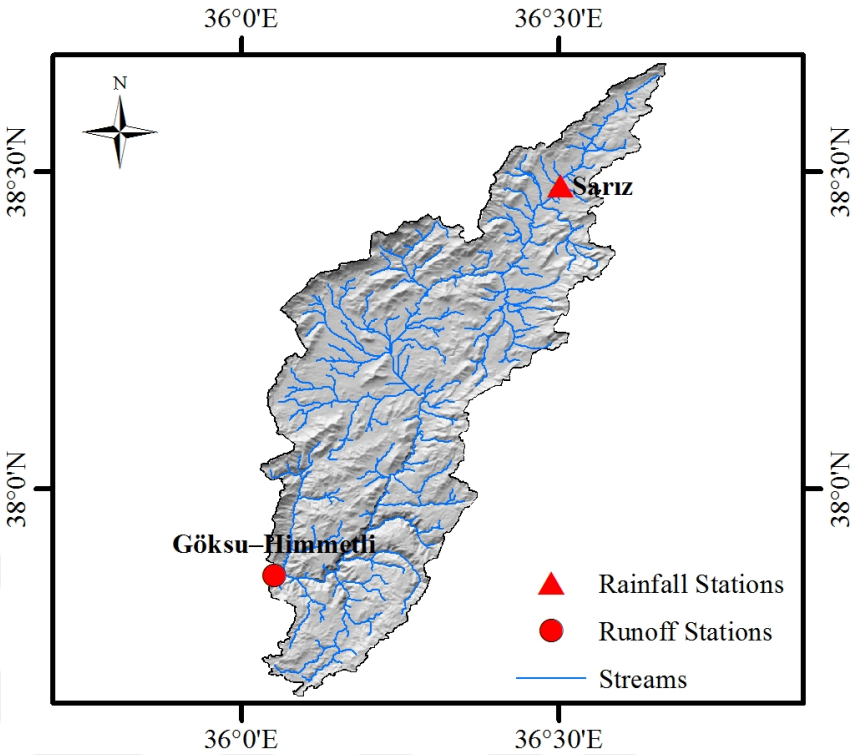


Figure 4.2. *location of the case study*

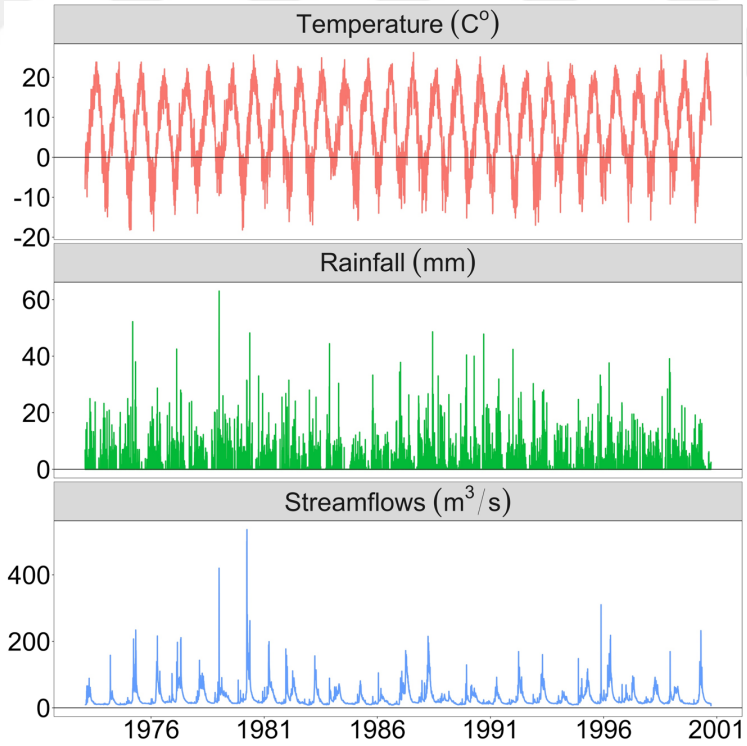


Figure 4.3. *Time series of the temperature, rainfall, and streamflow*

In DDMs, to maintain generality, the dataset divided into two subsets, each dedicated for training and validating the network where the latter not to be used for the training process. The dataset can also be divided into three subsets, each dedicated for training, validating, and testing the network, where the last subset is used after training and validation. In this study, the dataset was divided into two subsets in which 50% used for training from the period of February 1973 to November 1986 and 50% used for validation from the period of December 1986 to September 2000. Characteristics of the datasets used in the study are summarized in Table 4.1.

Table 4.1. *Characteristics of the daily temperature, rainfall, and streamflow*

Climate Var.	Temperature (C^o)		Rainfall (mm)		Streamflow (m^3/s)	
Subset	Training	Validation	Training	Validation	Training	Validation
Min	-18.4	-17	0	0	8.3	8.6
Max	25.6	26.2	63	48.6	536	310
Range	44	43.2	63	48.6	527.7	301.4
Median	8.1	7.8	0	0	18.3	19.9
Mean	7.6	7.3	1.5	1.5	29.2	29.5
Variance	79.3	83	16.3	16.6	912.3	635.8
CV	1.2	1.2	2.8	2.8	1	0.9

4.1.6. Application

In order to have datasets with the same range, increased accuracy, and increased pace of the calculations, normalization was applied to all datasets by using the following formula:

$$y = \frac{x - x_{min}}{x_{max} - x_{min}}, \tag{4.14}$$

where y is the normalized value of x, x_{min} and x_{max} are the minimum and the maximum values of the x variable respectively.

Several combinations of the lagged rainfall, temperature, and streamflow were created to compare the ability of several single DDMs with same models conjugated with WT in forecasting streamflow for seven days ahead. The output of each model is the current day streamflow, Q_t.

The models' combinations and models' numbers are shown in Table 4.2. For the readability of this manuscript, the model numbers are mentioned throughout the manuscript.

Table 4.2. *Different combinations of models*

Model No.	Combination	Notes
1	R_{t-7}	R: Rainfall t-7: lagged for seven days
2	$R_{t-7} + R_{t-8}$	
3	T_{t-7}	T: Temperature
4	$T_{t-7} + T_{t-8}$	
5	Q_{t-7}	Q: Streamflow
6	$Q_{t-7} + Q_{t-8}$	
7	$Q_{t-7} + Q_{t-8} + Q_{t-9}$	
8	$R_{t-7} + Q_{t-7} + Q_{t-8} + Q_{t-9}$	
9	$T_{t-7} + Q_{t-7} + Q_{t-8} + Q_{t-9}$	
10	$R_{t-7} + T_{t-7} + Q_{t-7} + Q_{t-8} + Q_{t-9}$	

The four schematic layouts with their preprocessing (i.e. WT) which were adapted in this study are shown in Figure 4.4. The first schematic was for the single DDMs in which the combination of model 1 to 10 (Table 4.2) were entered as inputs to the specific approach without applying any WT. Secondly, the CWT was applied to the normalized time series using scales from 0-400. The scale in which the coefficients of the CWT had the highest correlation with the streamflow identified as the best scale in which the specific coefficients were used as inputs of the DDMs. The third schematic was the DWT applied to the normalized time series and then DWT components of every model were entered to the DDMs. In the fourth schematic, DWT and CWT were gathered to produce a DCWT model. In this model, DWT was applied only on the streamflow and CWT only on rainfall and temperature.

In WT analysis, the wavelet function and the scale or decomposition level are two main issues that should be addressed. In CWT analysis, the “db2” which is one of the Daubechies wavelet family functions (Daubechies 1988) was found as the best function similar to (Shoaib et al. 2014). The scale where the coefficients had the highest correlation with the streamflow was identified and its coefficients were used as input.

In DWT analysis, several wavelet functions were examined and ‘dmey’ (i.e. Discrete Meyer function) (Meyer 1990) was found as the function with the highest accuracy for all variables This was in agreement with (Alizadeh et al. 2017) but not with (Cannas et al. 2006, Singh 2012, Shoaib et al. 2014) in which the study suggested db2, db4, and db8 functions. Additionally, best decomposition level was found as five for streamflow and temperature and seven for the rainfall.

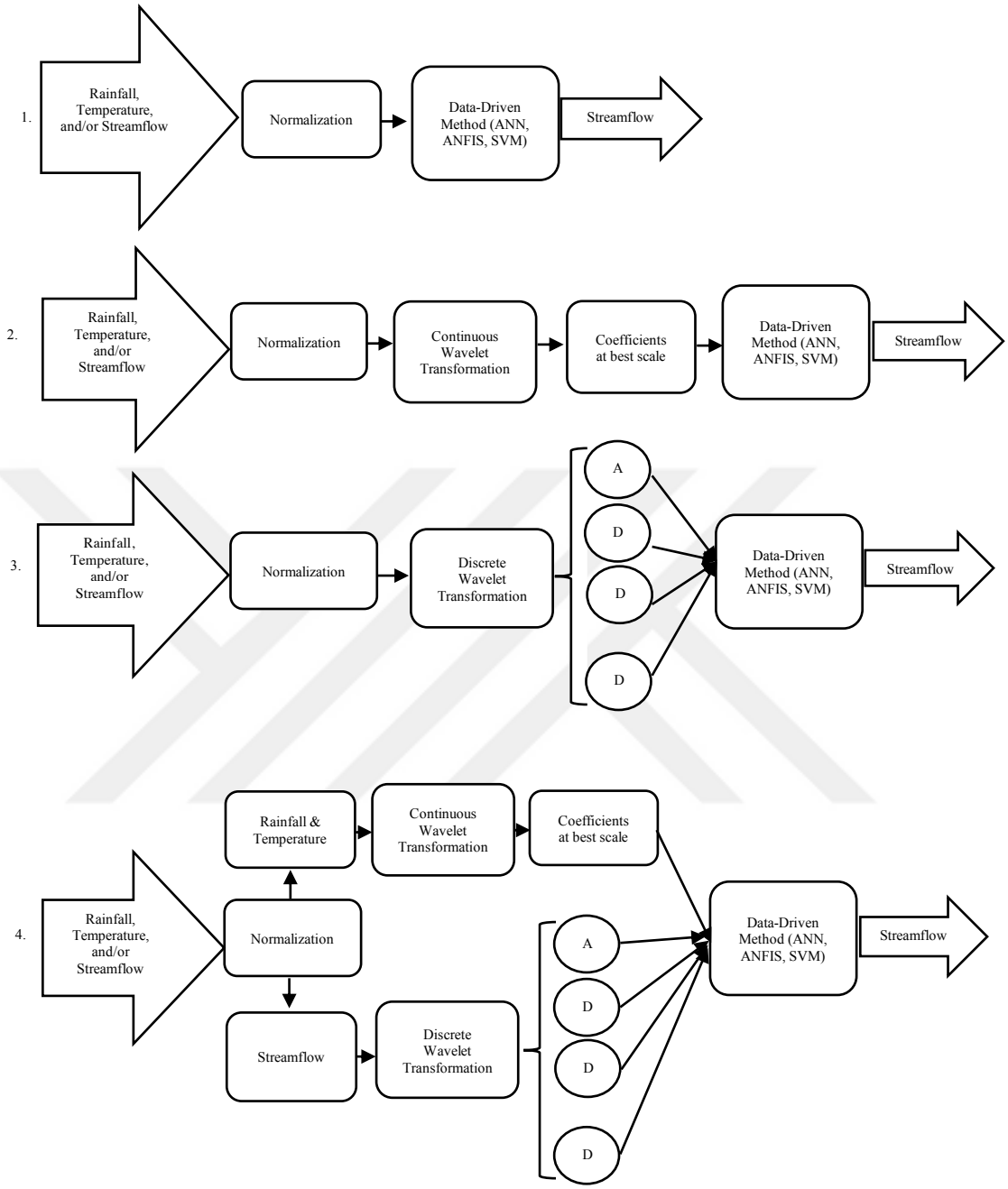


Figure 4.4. *Schematic layouts of the applied models for forecasting streamflow*

The numbers of inputs to the DDMs for every schematic are shown in Table 4.3. The coefficients with the best scale were chosen as the input for every variable in CWT schematic and, therefore, each variable had one transformed variable. For DWT schematic, the inputs relied on the decomposition level. In one rainfall variable (lag) there is 8 variables are to be used as inputs; 7 for the details of every level and one is the approximation coefficients of the eighth level. For the streamflow and temperature, there

are 6 series of coefficients 5 for details of the five levels and one for the approximation of the last level. For the DCWT schematic, the inputs for each streamflow variable (lag) were 6 for 5 decomposition level of the DWT decomposition, and 1 input for each temperature and rainfall obtained from the CWT. An example of the approximation and detail coefficients of the 7-days lagged streamflow Q_{t-7} is shown in Figure 4.5.

Table 4.3. *Number of inputs to the DDMs for every model and scheme*

Model	Direct	CWT	DWT	DCWT
1	1	1	8	
2	2	2	16	
3	1	1	6	
4	2	2	12	
5	1	1	6	
6	2	2	12	
7	3	3	18	
8	4	4	26	19
9	4	4	24	19
10	5	5	32	20

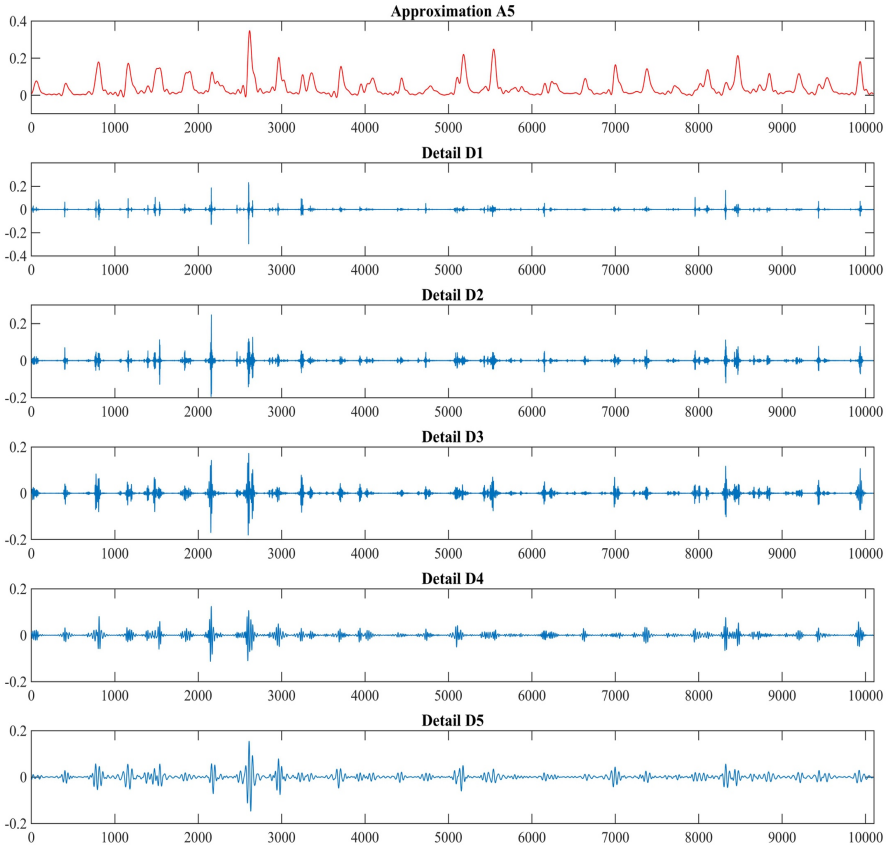


Figure 4.5. *the decomposition of the daily streamflow using "dmey" function with 5 levels*

Feed-forward Neural Network (FFNN) or MLP (FFNN will be used throughout the manuscript) is one of the several ANN structures that has been approved as an appropriate approach for the majority of hydrological applications (Hsu et al. 1995, Govindaraju 2000, Maier et al. 2000). Additionally, the three-layer network containing the input layer, hidden layer, and output layer has been useful in the hydrological applications (Hornik et al. 1989, Kişi 2007). Therefore, a three-layered FFNN was implemented in this study. The number of neurons in the input layer for the model combination and the schematic layout are shown in Table 4.3. Table 4.4 demonstrates the optimum number of neurons in the hidden layer obtained by trial and error process. It should be noted that the output layer in each model contains one neuron as the models' output is only the current streamflow. The current study applied the Levenberg-Marquardt algorithm as the training method (Kişi 2007). Tangent hyperbolic used as a transfer function in the hidden layer and the linear transfer function were used in the output layer.

Table 4.4. *The optimum number of neurons in the hidden layer of the ANN network*

Model	Single ANN	CWT-ANN	DWT-ANN	DCWT-ANN
1	1	2	2	
2	1	11	5	
3	6	4	7	
4	8	13	4	
5	3	8	13	
6	6	11	15	
7	14	7	5	
8	18	13	5	8
9	14	15	9	9
10	9	10	6	9

ANFIS with subtractive clustering (ANFIS-SC) method was used in this study. In order to generate the fuzzy rule of the ANFIS-SC, there are two important factors should be taken into account. Firstly, the cluster radius which defines the area that data points outside have reduced effect on the cluster center. Value of the radius varies between 0 and 1, where small values lead to more cluster and vice versa. It was reported previously that good radius values varied between 0.35 and 0.45 (Ping et al. 2006), and 0.2 and 0.5 (Samhouri et al. 2009). In the current study, 0.4 was found as an optimum radius value for all the model combinations by performing repeated trial and error approach. Hence, this value was used as constant throughout all models' combinations. Secondly, the squash factor which represents the radius where the reduction in the potential of any point

being a cluster center is measurable. By using trial and error in every model the optimum squash factor was identified and they are tabulated in Table 4.5.

Table 4.5. *The optimum squash factor values of the ANFIS-SC*

Models	Single ANFIS	CWT-ANFIS	DWT-ANFIS	DCWT-ANFIS
1	0.4	0.9	1.2	
2	0.4	0.9	1.1	
3	0.7	0.8	1.2	
4	1.1	1.1	1.5	
5	0.2	0.4	0.8	
6	0.3	0.3	0.7	
7	0.4	0.3	0.7	
8	0.4	0.9	0.7	1.1
9	0.6	0.5	1.1	1.2
10	0.6	1.2	1.6	0.9

The accuracy of the SVM relies on several parameters such as the control coefficient C that controls the trade-off between the maximum tolerated errors over ε and the flatness of the function, Epsilon ε which represents the distance in which the error is ignored, and Sigma σ which is a parameter of Gaussian kernel that controls the scale of the kernel. The optimum values of the above parameters were obtained using hyper-parameter tuning function. The optimum values of the SVR parameters obtained are listed in Table 4.6.

Table 4.6. *The optimum values of the SVR parameters*

	Single ANFIS			CWT-ANFIS			DWT-ANFIS			DCWT-ANFIS		
Model	C	ε	σ	C	ε	σ	C	ε	σ	C	ε	σ
1	0.001	3.1E-05	0.001	878.8	4.0E-05	0.102	0.009	1.6E-03	0.002			
2	0.001	3.1E-05	0.013	0.823	1.5E-02	0.046	102.8	3.6E-04	0.965			
3	0.013	2.8E-02	0.774	914.4	3.9E-02	45.53	99.6	1.8E-04	0.142			
4	0.005	3.1E-02	0.081	0.561	2.9E-02	0.956	99.6	1.8E-04	0.379			
5	252.8	9.2E-05	0.002	882.7	1.1E-02	17.55	949.5	3.1E-03	5.751			
6	998.5	3.2E-02	0.020	136.3	1.6E-02	0.106	17.04	3.6E-05	0.986			
7	0.315	3.4E-05	0.023	2.590	4.9E-05	0.150	686.7	2.9E-05	1.467			
8	0.048	2.9E-05	0.044	2.583	1.4E-02	0.114	2.223	2.9E-05	0.470	997.4	2.1E-03	4.77
9	0.062	3.2E-05	0.110	1.599	5.9E-04	1.485	696.1	5.3E-04	1.781	962.1	1.1E-02	3.80
10	962.1	3.9E-03	737.0	0.470	3.0E-05	202.5	380.2	2.9E-03	1.660	904.9	2.4E-03	3.36

4.1.7. Results and Discussion

Streamflow forecasting was performed using four schematic layouts which included Single, CWT, DWT and DCWT. The fourth schematic layout proposed in this study included both DWT and CWT where the first applied on the streamflow variables and the latter on the rainfall and temperature. Therefore, to have homogeneous comparison, the evaluation of the models' performance will be conducted in the following

manner: Firstly, only the first three schematic layouts (i.e. Single, CWT, and DWT) will be compared in order to compare the fourth layout using only the last three models' combinations. The model performance (in terms of model accuracy) of the models' combinations were evaluated in this study using *RMSE* and R^2.

Figure 4.6 shows the R^2 and RMSE of the ten models using the single ANN, CWT-ANN, and DWT-ANN schematic layouts. In the single ANN layout, the first two models M1 and M2 which included the 7 and 8 days delayed rainfall showed lowest R^2 and highest RMSE which indicated as bad models. The models with 7 and 8 days delayed temperature showed better result but stay giving bad result in general. Using the delayed streamflow as inputs increased the model performance dramatically. Models M5-M7 which included the 7, 8, and 9 days delayed streamflow respectively showed reliable results compared to the first models. In the last three models, namely M8, M9 andM10, the rainfall and temperature were added to the inputs of M7. Nevertheless, this addition resulted in a modest increase in the performance of the M7 model. In general, performance of the models using only ANN is not giving a sufficient accuracy. By using CWT as pre-processing, performance of the first two models contained only the rainfall improved dramatically compared to the single ANN layout. Nevertheless, the model accuracy did not improve by including the 8-days delayed rainfall. Similarly, the performance of the models that have temperature only increased however, M4 with 7 and 8 days delayed temperature outperformed M3 with only 7 days delayed temperature. In contrast, results from the trade-off between the training and validation subsets demonstrated that the accuracy of all other models M5-M10 neither increased nor decreased. In the layout of DWTANN, the first model, M1 was better than single ANN but worse than CWTANN, but adding the 8 days delayed rainfall improved the M2 performance was better than CWTANN. In the temperature models namely M3 and M4, M3 performed better than single ANN, but not CWTANN while M4 performed better than single ANN but relatively same as the CWTANN. Comparing DWTANN with both single ANN and CWTANN, the performance of M5-M10 were increased dramatically. Adding only the rainfall or only the temperature to the model M7 in M8 and M9 did not improve the performance but adding both of them in M10 added a modest increase to the accuracy. In general, using DWT with ANN demonstrated the superior performance among the three schematic layouts.

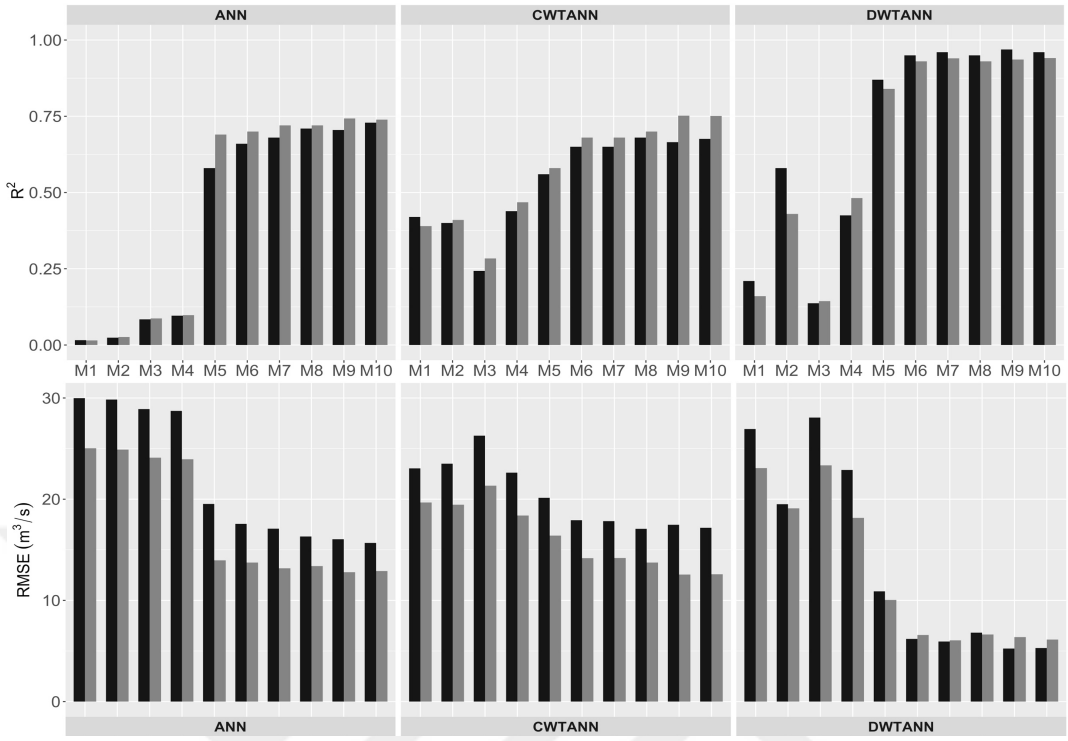


Figure 4.6. *The R2 and RMSE values of the three schematic layouts (ANN, CWTANN, and DWTANN)*

The three schematic layouts with ANFIS-SC showed similar trend to those with ANN (Figure 4.7). The first four models M1-M4 in which only included the rainfall and temperature had better accuracy in CWTANFIS layout compared to the single ANFIS. The models in DWTANFIS layout had higher accuracy than single ANFIS but lower accuracy than CWTANFIS. The models M5-M10 had the best result in DWTANFIS; single ANFIS and CWTANFIS had similar results.

Figure 4.8 shows that the performance of layouts single SVM and CWTSVM were different from the same layouts with the ANN and ANFIS, while the layout of DWT was similar with the three DDMs. Similar to the ANN and ANFIS results, the first four models, in the single SVM, has the lowest performance with M1, and M2 have negative R^2 values. The performance of the models that includes the delayed streamflow as inputs increased dramatically. Unlike ANN, and ANFIS, the accuracy deteriorated in the models M8-M10 with the rainfall and temperature especially in M10. In CWTSVM, the first four models' performance improved in comparison to single SVM. Similar to the CWT layout of ANN and ANFIS, the remaining models had similar results with the single SVM. In contrast, the performance of M5 and M10 decreased. In DWTSVM, the models

performed analogously to those of the same layout in ANN and ANFIS in which models with one delayed variable in rainfall and temperature performed better than single SVM but not in CWT. The models with two delayed variables performed better than single SVM but relatively similar to CWTSVM. Moreover, rest of the models demonstrated similar performance in ANN and ANFIS. On the other hand, in single SVM and CWTSVM, adding the rainfall and temperature neither improved nor deteriorated the performance of models.

According to Figure 4.8, it is clear that M8-M10 models had the highest performance compared to other models' combinations. Moreover, the proposed DCWT layout can only be applied to M8-M10 models. Therefore, Figure 4.9 shows the comparison of R^2 and RMSE between the schematic layouts of all DDMs for M8, M9 and M10 models. Generally, the single models had better performance than the CWT with nominal differences in SVM. Despite of having small differences between each other, the performance of DWT and DCWT in all DDMs is higher than the other two layouts. The performance of the three models' combinations M8-M10 was consistent. Nevertheless, single SVM and CWT of SVM demonstrated fluctuation in the results in contrast to DWT and DCWT of the same approach which behaved consistently.

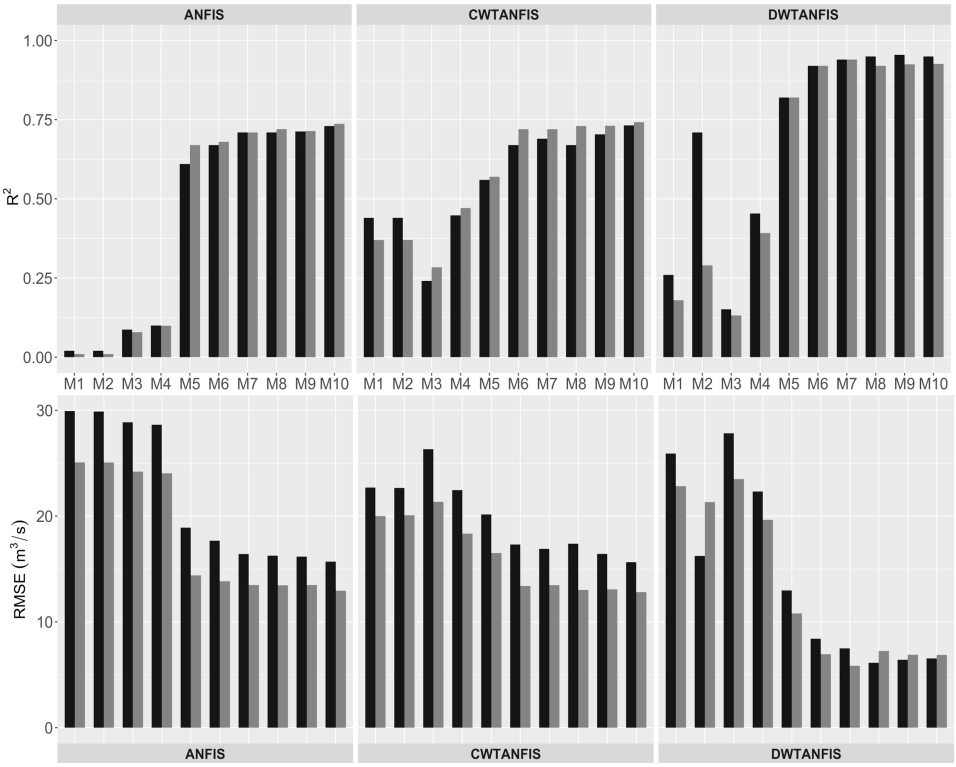


Figure 4.7. *The R2 and RMSE values of the three schematic layouts (ANFIS, CWTANFIS, and DWTANFIS)*

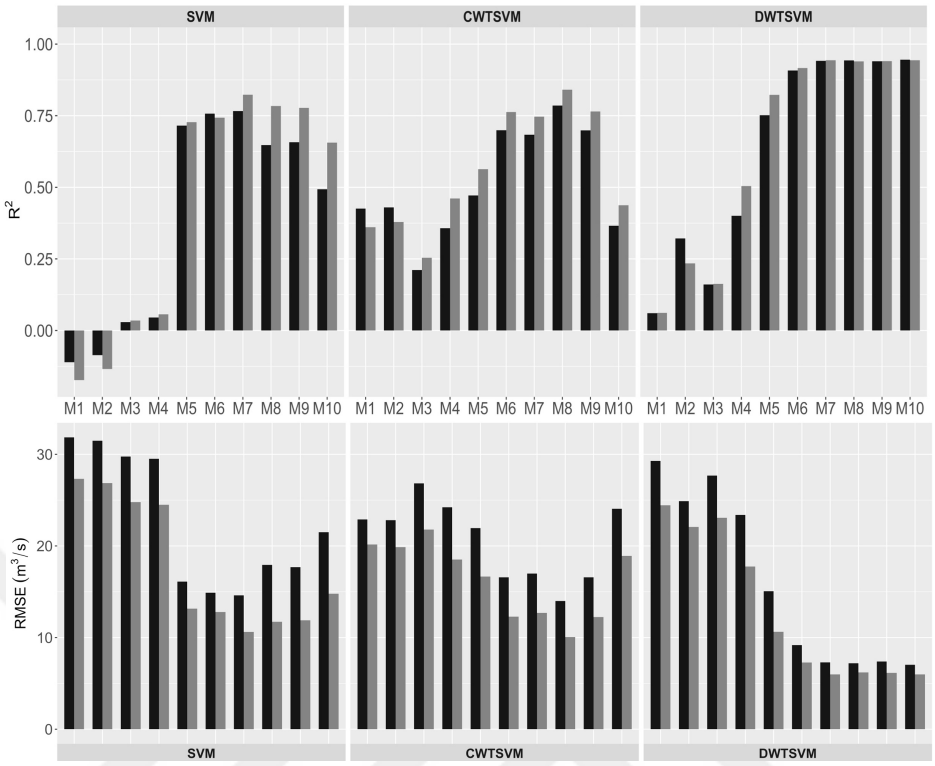


Figure 4.8. *The R2 and RMSE values of the three schematic layouts (SVR, CWTSVM, and DWTSVM)*

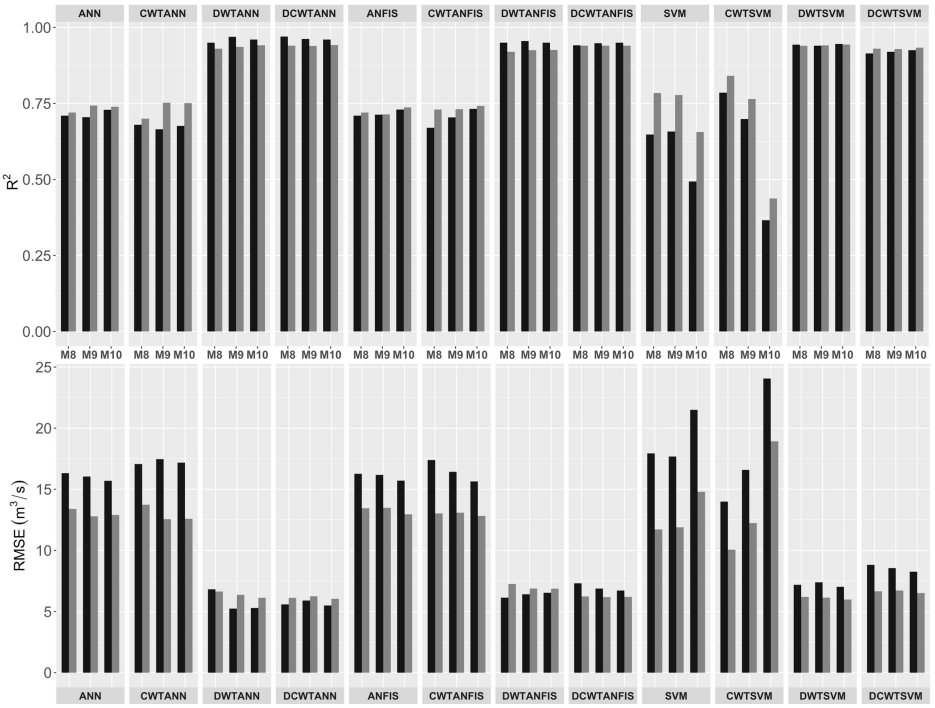


Figure 4.9. *Comparison of R2 and RMSE values between the schematic layouts of all DDMs for models M8, M9, and M10*

DCWT is the application of CWT on the rainfall and temperature and DWT on the streamflow variables. Hence, this combination was investigated in this study to determine whether it improves the models' accuracy. An increased accuracy of the models with only rainfall and temperature variables was observed by applying CWT compared to DWT. In contrast, an increased performance observed in the models with streamflow variables by applying DWT compared to CWT. The results of the relative performance of DCWT with the DWT indicated by the change percentage in R^2 and RMSE by applying DCWT with respect to DWT are listed in Table 4.7. In ANN, particularly for M8, R^2 increased by 2.11 and 1.08 % and RMSE decreased by 18.04 and 7.83 respectively for the training and validation. In M9, the performance of the training was deteriorated but improved in the validation applying trade-off between the two subsets the deterioration in the performance is identified. In M10, accuracy was unchanged. That leads to the conclusion that, CWT applied on the rainfall in M8 improved the accuracy of forecasting in the overall model. However, with the temperature input in M9, performance deteriorated slightly while M10 had no change. In ANFIS, contradictory changes observed in M8 and M9 between the training and validation subsets in which trade-off resulted in no change. Slight improvement in the performance was observed by including both rainfall and temperature in M10. By applying this layout using SVM, decreased in R^2 and increased in RMSE values were observed in the three models' combinations. Hence, the performance of all the three models' combinations deteriorated although CWT was identified previously to perform better than DWT in models contained only rainfall and temperature. The ability in forecasting the peak values also increased by applying DCWT with ANN in model M8 but not in M9 and M10 (Table 4.8). In ANFIS, the performance of the model M10 increased while in SVM the performance deteriorated for all models.

Figure 4.10 shows the time series of the observed, predicted DCWT, and predicted DWT. Using ANN, DCWT had a higher agreement with the observed data than DWT in M8 but DWT performed better in M9 and M10. Nonetheless, there was no superiority observed between DCWT and DWT in all three models of ANFIS while the higher agreement of the DWT with the observed data was seen using SVM. Although the forecasting ability was good, there were underestimation of the high values with good forecasting of the highest two flood values. Additionally, ANN had the highest R^2 and lowest RMSE values in all models combinations and in both layouts of DWT and DCWT as shown in Table 9. ANFIS and SVM had relatively similar performance in the DWT

layout while ANFIS outperformed SVM in the DCWT layout. Notably, the results indicated that ANN in all three models had the highest agreement, and ANFIS had a better agreement than SVM.

Table 4.7. *The relative performance of DCWT with DWT in forecasting*

			Change (%)		
			M8	M9	M10
ANN	R2	Training	2.11	-0.72	0.00
		Validation	1.08	0.32	0.11
	RMSE	Training	-18.04	12.57	3.77
		Validation	-7.83	-2.04	-1.31
ANFIS	R2	Training	-0.95	-0.73	0.00
		Validation	2.17	1.62	1.51
	RMSE	Training	19.06	7.32	2.75
		Validation	-13.93	-10.30	-9.88
SVM	R2	Training	-3.03	-2.17	-2.19
		Validation	-0.99	-1.24	-1.11
	RMSE	Training	22.66	15.77	17.54
		Validation	7.42	9.38	8.89

Table 4.8. *The relative performance of DCWT with DWT in forecasting peak values*

		Change (%)		
		M8	M9	M10
ANN	R^2	2.86	-0.85	-0.92
	RMSE	-11.43	4.02	4.64
ANFIS	R^2	-1.30	-0.31	1.40
	RMSE	4.64	1.17	-5.07
SVM	R^2	-4.74	-3.23	-3.96
	RMSE	14.23	9.84	12.65

4.1.8. Conclusion

Using a thirteen-year data series from a basin located in Goksu-Himmet, the ability of ANN, ANFIS and SVM combination models to forecast seven days ahead streamflow was evaluated and compared by utilizing four different schematic layouts. The results indicated that utilization of WT improved the ability of the DDMs in forecasting the 7-days ahead streamflow. Moreover, it was found that both ANN and ANFIS methods provided accurate results, with the best ANN model slightly outperformed the ANFIS model.

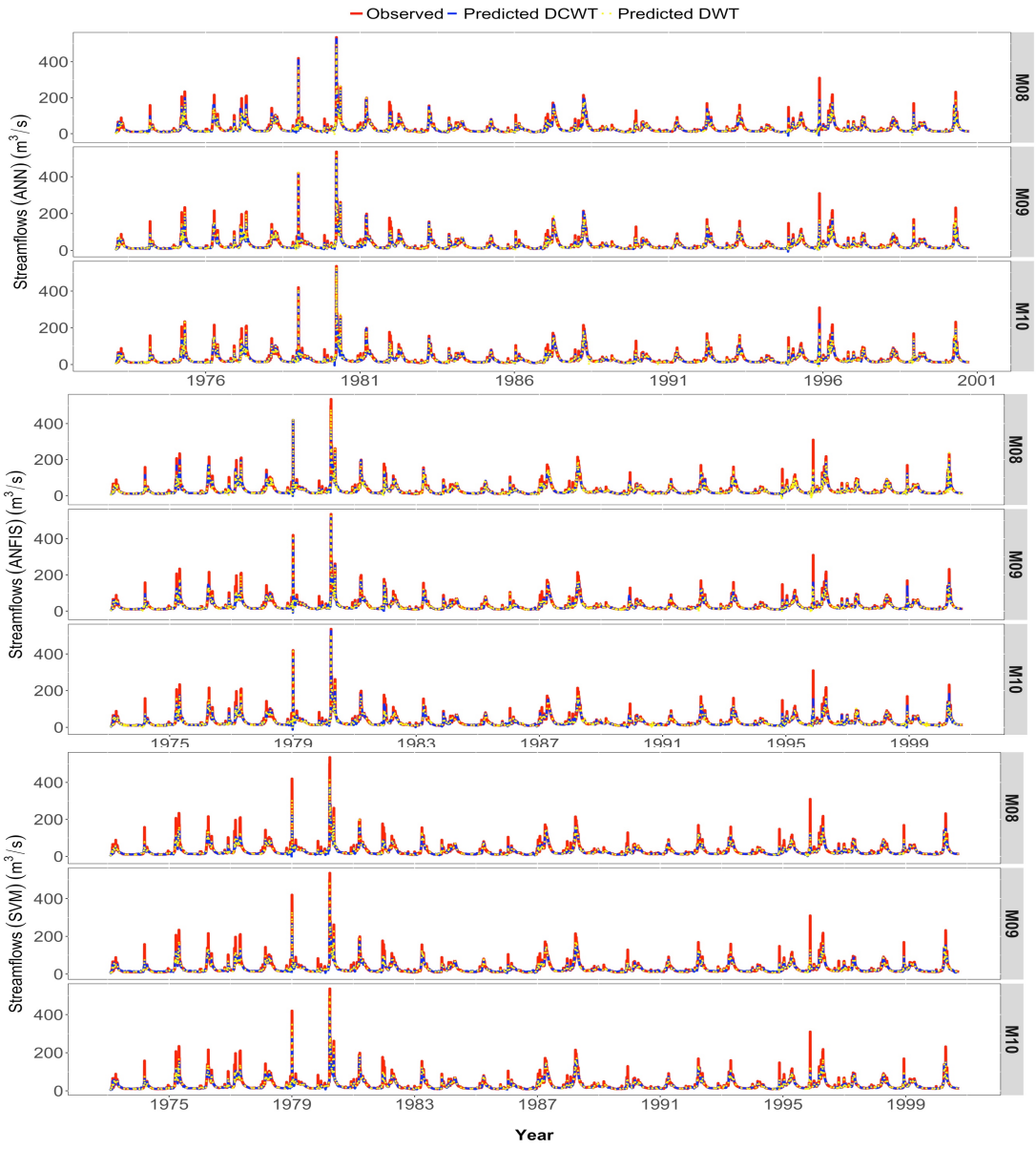


Figure 4.10. *The time series of the observed predicted DWT, and predicted DCWT*

In specific several conclusions are remarked:

- CWT improves only the models that contains only the rainfall and temperature in which the single models have very low accuracy, but not those which includes the delayed streamflow in which the accuracy higher.
- CWT increase the accuracy more than the DWT in the models that contains only rainfall and temperature while in the other models in which the streamflow is included the use of DWT is greatly giving higher accuracy.

- DWT improves the models' accuracy in comparison with the single model dramatically and that greatly depends on the choice of the appropriate wavelet function and the decomposing level.
- The combination of the CWT which applied on the rainfall and temperature only and DWT which applied on the streamflow only namely Discrete-Continuous-Wavelet transformation (DCWT) improved several models in comparison with DWT; especially those models include rainfall and used with ANN, but, the models imposed to the SVM had a deteriorated performance and that due to the unexpected behavior of SVM with the use of CWT in general.
- The use of DWT and DCWT are the schematic layout with the highest accuracy in comparison with single and CWT layouts.
- The use of DWT and DCWT but not single or CWT as preprocessing with DDMs is sufficient in forecasting the 7-days ahead streamflow when several climatological variables are used as inputs.
- ANN outperformed ANFIS and SVM and ANFIS has higher accuracy than SVM.

It is construed that the use of DCWT could be a promising combination of the WT as a preprocessing tool that could improve the forecasting ability and useful in other hydrological applications.

4.2. Improving Monthly Streamflow Forecasting Using Continuous Wavelet, and Multi-Gene Genetic Programming Combination

4.2.1. Abstract

Streamflow is an essential component of hydrologic cycle in the regional and global scale and it is the main source of the fresh water supply and highly associated with the natural disasters such as droughts and floods. Therefore, accurate streamflow forecasting is essential. Forecasting streamflow in general and monthly streamflow in particular is a complex process cannot be handled by data-driven models (DDMs) only and requires pre-process. Wavelet transformation (WT) is one of the pre-processing techniques but applying Continuous Wavelet Transformation (CWT) producing a large number of scales which deteriorates the performance of any DDM due to the high number of redundant variables. In this study, Multigene Genetic Programming (MGGP) is proposed as a

selection tool that after the CWT analysis it selects the important scales in order to impose these scales, after that to the Artificial Neural Network (ANN). To prove the forecasting ability of the proposed model, a basin located in the southeast of Turkey is selected as case study. One month ahead downstream flow used as output, downstream flow, upstream, rainfall, temperature, and potential evapotranspiration with their associated lags used as inputs. Before the modeling, the Wavelet Coherence Transformation (WCT) analysis conducted to analyze the relation between variables in time-frequency domain. Several combinations developed to investigate the effect of the variables on the streamflow forecasting. The results showed that there is high localized correlation between the streamflow and the other variables and especially the upstream. In the models of stand-alone layout in which the data entered to ANN and MGGP without CWT the performance is found poor. In the best-scale layout in which the best scale of the CWT identified as the highest correlated scale is chosen and entered to ANN and MGGP, the performance slightly increased. Using the proposed model, due to the inclusion of several scales in which seasonality and irregularity can be captured, the performance was dramatically increased and particularly in forecasting the peak values. It was also found that using hydrological and meteorological variables improves the ability in forecasting the streamflow.

4.2.2. Introduction

Streamflow is an essential component of hydrologic cycle in the regional and global scale (Makkeasorn et al. 2008, Liu et al. 2014). It is the main source of the fresh water supply and highly associated with the natural disasters such as droughts and floods. Therefore, accurate short-term and long-term forecasting streamflow is essential for efficient water resources management particularly in the regions vulnerable to floods and droughts (Kisi et al. 2011).

The forecast of streamflow is identified as not an easy task (Ravansalar et al. 2017) due to the complexity of the process (Yaseen et al. 2015, Mehr et al. 2017). This complexity is mostly due to the non-stationarity and non-linearity in the relation between the streamflow and the basin characteristics (Nourani et al. 2011, Mehr et al. 2017). Data-Driven models (DDMs) can be used for modeling streamflow as they have the ability to deal with non-linearity and map the relationship between the streamflow and factors driving it. Another advantage of DDMs is that it obtains the optimum solution between

the inputs and outputs without involving the physical process (Solomatine et al. 2008). There are a numerous number of applications of forecasting streamflow using DDMs including Artificial Neural Network (ANN), Support Vector Machines (SVM), Adaptive Neuro-Fuzzy Inference System (ANFIS), and Genetic Programming (GP) (Tokar et al. 1999, Kişi 2007, Kisi et al. 2012, Nourani et al. 2013, Mehr et al. 2015, Abdollahi et al. 2017, Mehr et al. 2017, Mehr et al. 2017).

The DDMs can deal with the nonlinearity and non-stationarity in the mean and variance elements but the main shortcoming of the DDMs is that they may not be able to cope with the non-stationarity fluctuation (Nourani et al. 2011) if no preprocessing applied to the data (Cannas et al. 2006). One of the solutions to apply the wavelet which have the ability to cope with the seasonal (cycling) nonstationary component of the time series (Daubechies 1990, Torrence et al. 1998, Anctil et al. 2004, Nourani et al. 2011). Wavelet has good ability in representing the signal locally in both time and frequency domains. The decomposition of the nonstationary time series using wavelet into several number of scales (Lee et al. 1994) extracts the historical hiden information in both time and frequency domain (Rajaee et al. 2010, Shoaib et al. 2014).

Wavelet with various DDMs types has been widely applied in the streamflow forecasting (Shiri et al. 2010, Kisi et al. 2011, Adamowski et al. 2012, Badrzadeh et al. 2013, Komasi et al. 2016, Ravansalar et al. 2017) and especially with ANN approach (Kişi et al. 2007, Carcano et al. 2008, Kişi 2009, Adamowski et al. 2010, Araghinejad et al. 2011, Pramanik et al. 2011, Adamowski et al. 2012, Isik et al. 2013, Nourani et al. 2013, Nourani et al. 2014, Alizadeh et al. 2017).

Although ANN is a robust modelling method it is an implicit method that maps the inputs and outputs using weights matrices which leads to a model that not understandable by humans (Savic et al. 1999). Also, ANN needs an intensive trial and error processes in order to determine the optimum numbers of the input and hidden neurons (Nourani et al. 2011). Therefore, the need of GP method was initiated which is recognized as a robust explicit method that can be used for modelling (Dorado et al. 2003, Mehr et al. 2013, Mehr et al. 2014). (De Falco et al. 2005) reported several advantages of the GP over ANN: 1) Generates an explicit model understandable by humans, 2) Automatically discovers the structure of the model by utilizing the given data, 3) has an adaptive evolutionary search ability that allows for not being trapped in suboptimal unsatisfactory local solutions, 4) absence of specific knowledge. As a consequence, GP has been applied

widely in the hydrological modelling problems (Rabunal et al. 2007, Aytek et al. 2008, Ghorbani et al. 2010, Kisi et al. 2010, Mehr et al. 2013) and especially with wavelet for forecasting streamflow forecasting (Mehr et al. 2017, Mehr et al. 2017, Ravansalar et al. 2017). Regardless the inputs preprocessed or not, most of the studies applied GP as final modeling method by giving the input and output and train the model. In other words, GP is used as standalone modelling method. In addition to that GP can be used as a sensitivity analysis tools as it has the ability in formulating and structuring the equation used for forecasting. This helps in selecting the inputs in order to be after that imposed to another DDM like the model proposed by (Nourani et al. 2011). The authors proposed a model namely WGPNN which stands for Wavelet-Genetic Programming-ANN in which after the discrete wavelet transformation applied the GP applied as sensitivity analysis to be finally imposed to ANN.

The wavelet transformation is performed by two forms namely continuous wavelet transform (CWT) and discrete wavelet transform (DWT). According to (Nourani et al. 2014), more than 80% of the hydrological studies is applying the DWT as preprocessing and not CWT. This percentage is much higher in the streamflow forecasting. This is due to the nature of the hydrologic time series which gathered not in a continuous but in Discrete form. Therefore, DWT which is dyadic is more appropriate for decomposing the time series which allows for reconstructing the original time series form the sub-signals. On the other hand, CWT represents the signal in many different and exact scales (periods) that in every scale there is a number of coefficients equal to the same number of time steps. Therefore, it is considered as redundant information especially when used as preprocessing to be then imposed into a DDMs as increasing the number of redundant inputs deteriorate model's ability in forecasting (Galelli et al. 2014, Taormina et al. 2016). DWT is dyadic and therefore the coefficients are produced in a scale of 2^m modes in which to capture the annual cycling in the daily mode about 2^8 –days mode is used where the decomposition level is 8 and that produced 9 sub time series one is the approximation and 8 details. For monthly data, 2^4 mode approximately represents the annual time scale which produce 5 sub time series (Nourani et al. 2011). In CWT, the daily data produces 365 and the monthly data produces 12 sub time series to capture the annual scale which can be a huge number if several variables are transformed using CWT.

Several studies used CWT for forecasting streamflow by taking one scale coefficients or several scales based on the data used. For example, (Adamowski 2008)

used a standalone CWT method for river flow forecasting by applying numerical method on coefficients of several scales chosen based on the cross-wavelet analysis and compared the result with Multivariate Linear Regression (MLR), ARIMA, and ANN. He reported good accuracy in 1, and 2 days leading time models but ANN outperformed the proposed method in 6 days leading time. (Cannas et al. 2006) studied the use of CWT, DWT, and Data-partitioning as three different preprocessing techniques for predicting the river flow. Although they found that the first scale coefficients in the application of CWT give the best accuracy but concluded that it is not giving a significant increase in comparison to the model using the raw data. (Shoaib et al. 2014) investigated the effect of the selection of the mother wavelet function on the ability of forecasting the runoff using only the rainfall as input. They used the best scale coefficients which identified by the highest correlation with the runoff. Although they reported an increase in the accuracy but the linear correlation coefficients in not enough to detect the rainfall-runoff relation which is identified as nonlinear. Therefore, there is a need of a method that have the ability to identify the most important scales that contributes in the streamflow. The most probable candidate is Genetic programming (GP)

In this study, GP is proposed as selection tool of the most important scales of the transformed time series using CWT, to be imposed after that to ANN. The schema of this study is as follow: the seasonality (Cycling), non-stationarity, and irregularity to be extracted from the time series by applying CWT, the selection tool is then implemented by applying multigene-GP in which the most contributing scales in the streamflow forecasting are selected, and finally the nonlinearity is to be handled by imposing the selected scales coefficients to ANN.

4.2.3. Wavelet Transformation

A wavelet is defined as a small wave (i.e. has a finite length) with a zero-average value. Wavelet is a mathematical function or tool that decomposes the time series into different frequency components in order to capture the characteristics of the time series. In other words, it breaks the time series into its wavelets which are scaled and translated (shifted) version of the mother wavelet (Nason et al. 1999). It has the ability of providing time-frequency representation of the signal as it provides time and frequency information simultaneously (Cannas et al. 2006). Also, it has the advantage of overcoming the issue

of the non-stationarity faced by the DDMs (Daubechies 1990, Anctil et al. 2004, Shoaib et al. 2014, Alizadeh et al. 2017).

Time series data transformation is done either by translation or by scaling or dilating of the mother wavelet function. Translation is moving the wavelet along the axis of the time at different positions while scaling or dilation is squeezing or stretching the wavelet (Percival et al. 2006). Generally, WT divided into two forms: continuous wavelet transform (CWT) and discrete wavelet transform (DWT). Only CWT is applied in this study. The CWT is used for revealing the characteristics of the series in multi-temporal scales. The continuous wavelet coefficient $W_x(a,b)$ is defined as:

$$W_x(a,b) = \int_{-\infty}^{+\infty} f(t)\, \psi^*_{a,b}(t)dt\,, \tag{4.15}$$

where:

$$\psi_{a,b}(t) = |a|^{-1/2}\psi\left(\frac{t-b}{a}\right) \quad a,b \in R, a \neq 0\,, \tag{4.16}$$

where * represents the complex conjugate of the function, a and b are the parameters known as the scale and translation parameters respectively. $\psi\left(\frac{t-b}{a}\right)$ is the transformation function known as the mother wavelet. The term mother refers to that the transformation is implemented by scaling and translating using one main function, mother wavelet function. Wavelet means a small wave in which small refers to finite length of the function and wave is its oscillatory condition (Cannas et al. 2006).

The coefficients of the CWT ($W_x(a,b)$) are calculated in each a (scale) and b (translation) which generate a large amount of data. These coefficients are large when the time series $f(t)$ and the wavelet $\psi^*_{a,b}(t)$ are similar (Torrence et al. 1998) which, after the decomposition, provides a signal frequency at different scales that can be used for analysis.

4.2.4. Wavelet Coherence Transformation (WCT)

CWT has the ability to analyze the localized frequencies in time series but this is done only for one time series. It is also desirable to examine the relation between two time series in the time-frequency domain. Particularly, it is important to examine the regions with the high common power in the time-frequency domain if they have a consistent phase relationship which lead to casualty conclusion. Therefore, Cross-wavelet

transformation (XWT) can be useful in this situation which is constructed using two CWTs (Grinsted et al. 2004).

The XWT of two time series x and y is defined by (Jury et al. 2002) as:

$$W_{xy}(a,b) = W_x(a,b)W_y^*(a,b)\,, \tag{4.17}$$

where * represents the complex conjugate of the function. And the cross wavelet power is further defined by (Grinsted et al. 2004) as $|W_{xy}(a,b)|$. Cross wavelet spectrum also provides a local phase difference estimation $\Delta\phi(a,b)$ between the two time series for each point (a,b) in the time-frequency domain. Taking the advantage of the phase difference being independent from the amplitude, the estimation of the instantaneous phase difference between two time series is possible. The phase difference is defined by (Jury et al. 2002) as:

$$\phi(b) = tan^{-1}\frac{\int_{a1}^{a2} Im[W_{xy}(a,b)]da}{\int_{a1}^{a2} Re[W_{xy}(a,b)]da}, \tag{4.18}$$

where b corresponds to time, $a1 < a2$ represents the lower and upper scale limits respectively, Im and Re represents the imaginary and the real part of $W_{xy}(a,b)$.

XWT identifies only the regions with a high common power in the two time series. WCT has the advantage of measure how coherent the XWT in time-frequency domain that it measures how the two time series co-vary together and not necessarily have high power (Grinsted et al. 2004). The WCT is defined by (Torrence et al. 1999) as the squared absolute value of the smoothed cross wavelet spectrum normalized by the smoothed wavelet power spectra of the two time series:

$$R_b^2(a) = \frac{\left|S(a^{-1}W_{xy}(a,b))\right|^2}{S(a^{-1}|W_x(a,b)|^2)\,.\,S\left(a^{-1}\left|W_y(a,b)\right|^2\right)}, \tag{4.19}$$

where S is smoothing operator. Looking at Eq.(4.19), it can be seen that this equation looks like the traditional correlation expression and thinking WCT as a localized correlation in the time-frequency domain is useful (Grinsted et al. 2004). The smoothing parameter is written as:

$$S(W) = S_{scale}(S_{time}(W(a,b))), \tag{4.20}$$

where S_{scale} refers to smoothing along the wavelet scale axis, and S_{time} to smoothing along the wavelet time axis. For details of XWT and WCT see (Torrence et al. 1998, Torrence et al. 1999, Jury et al. 2002, Grinsted et al. 2004).

4.2.5. Genetic Programming

Genetic programming (GP) which is firstly proposed by (Koza 1992) is a type of evolutionary algorithms as one of the machine learning categories inspired by the biological evolution operations. The operations include reproduction, cross over, mutation, architecture-changing operations patterned after gene duplication, and gene deletion in nature. The algorithm of GP is evaluating a user-defined (known as fitness function) function by using these operations and randomly generates a population of computer programs (represented by tree structure). Then, the best preformed trees are bred together to generate a new population. This process which mimics the Darwinian evolution, is repeated until the population that solve the fitness function well is obtained (Searson 2015). GP can solve problems of extracting functional relationships between the features which is called symbolic regression, and grouping data into categories which is called classification. GP has the advantage over the traditional regression and other data-driven models in that it does not need for the identification of the structure as it is self-structuring method (Mehr et al. 2017). Usually, possible program is represented by a tree structure that contains root node, inner node(s), and leaves. Figure 4.11 shows an example of the representation of the tree structure of GP and the function of that example is written as:

$$F(X_1, X_2) = (3.5X_1 \div Cos(X_2) + \exp\big(Cos(X_1)\big). \quad (4.21)$$

The GP algorithm process starts with creating a random population of individuals, evaluating the fitness function of individuals, and after that choosing parents are out of these individuals. The chosen parents are then bred to yield offspring by the reproduction, mutation, and crossover operations (Sivanandam et al. 2007, Nourani et al. 2011). The offspring creation process is iterated until the creation of a specified number of offspring in a generation and a specified number of generations. The results of the entire process is a computer program or an equation which is the solution of the problem (Nourani et al. 2011). The operations used in the transformation of one population of individual to another (i.e. reproduction, mutation, and crossover) are defined in the following lines. Reproduction operation is the selection of an individual from the current population and copying it into a new population without altering. Mutation operation is applied either to a terminal node or a function node and it replaces the terminal node in the tree by another terminal and the function node by new function. Crossover operation is the switch

between two selected parents with their trees. Figure 4.12 shows an example of the operations.

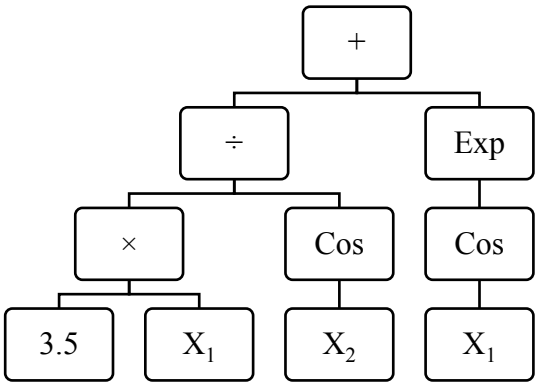


Figure 4.11. *GP tree structure of the expression* $\mathbf{3.5}\boldsymbol{X_1} \div \boldsymbol{Cos}(\boldsymbol{X_2}) + \boldsymbol{exp}\ (\boldsymbol{Cos}(\boldsymbol{X_1}))$

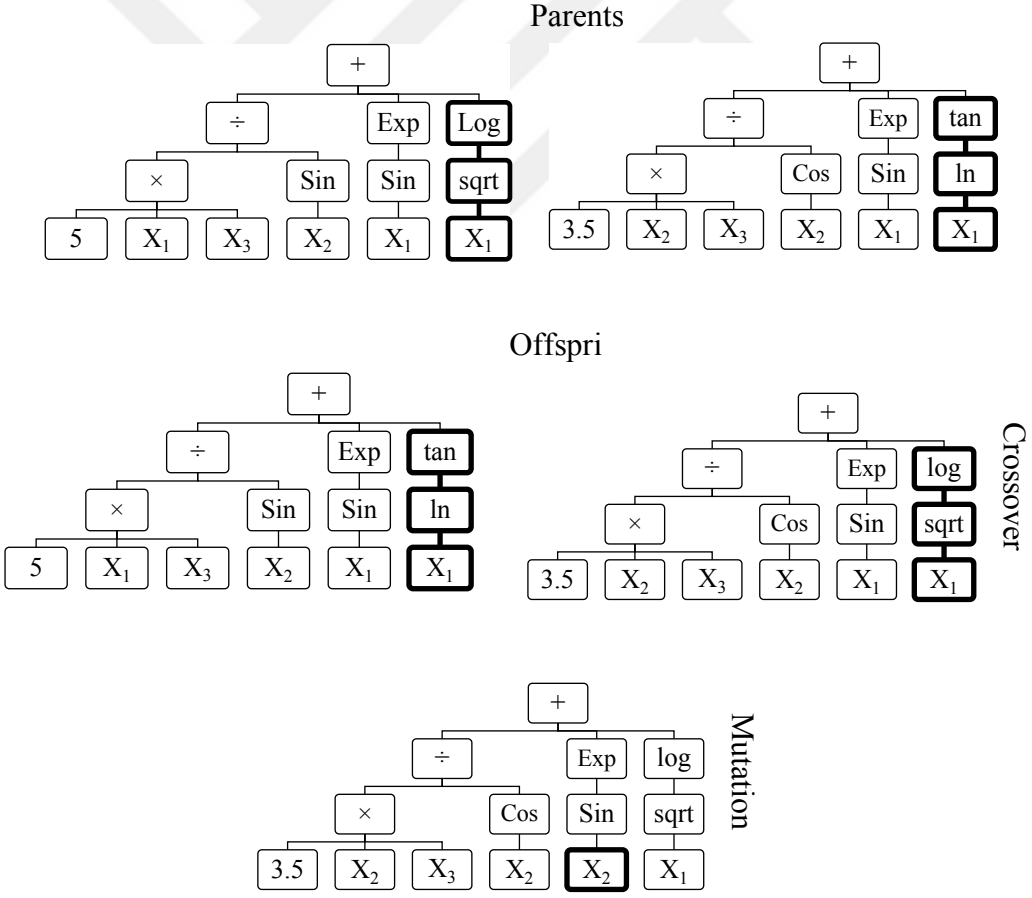


Figure 4.12. *An example of the crossover and mutation operations*

GP method has major inputs: 1) pattern of training/validation, 2) fitness function such as Root Mean Square Error (RMSE), 3) functional and terminal nodes (or inner nodes and leaves) for the structure identification, and 4) GP parameters for syntax tree formation (Mehr et al. 2017). The functional set are chosen according to the complexity of the problem to be chosen. The basic arithmetic parameters such as $+, -, \div, \times$ can be used for a simple problem but for more complex problems complicated operators such as Sin, Cos, Tan, Exp and others (Gandomi et al. 2015).

Recently, several advancements have been suggested for the classical GP such as Linear GP (LGP), Multi-expression Programming (MEP), Gene-Expression Programming (GEP), and Multi-Gene GP (MGGP) (Mehr et al. 2017). MGGP combines low depth GP trees linearly for improving the classical GP fitness. Therefore, MGGP is expected to provide models simpler than those of the classical GP as it uses smaller trees. Suppose we have a variable y to be predicted using several variables $x_1, x_2, \ldots, x_n$. In MGGP, y is predicted by the summation of the weighted outputs of each tree (gene) and the bias as the additive model is considered (Searson 2015). Referring to Figure 4.13, the mathematical expression can be written as:

$$\hat{y} = b_0 + b_1 T_1 + b_2 T_2 + \cdots + b_G T_G , \tag{4.22}$$

where $\hat{y}$ is the predicted value, T_i is the vector of the output of the ith gene/tree, b_0 is the bias, $b_1, b_2, \ldots, b_G$ are the regression coefficients and considered as the weights of the gene/tree also.

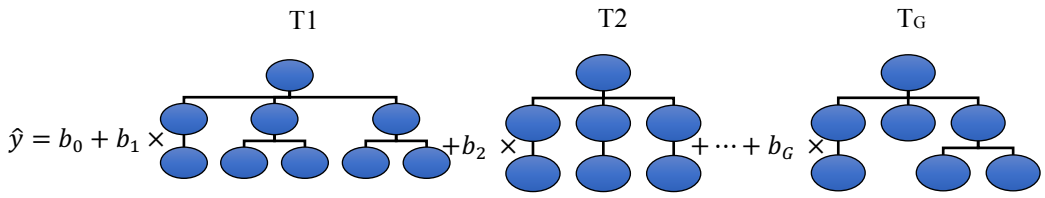


Figure 4.13. *Multi Gene Genetic Programming (Symbolic Regression)*

The coefficients of the regression are determined by the ordinary least square method for every individual MGGP. The advantage of depth restricted trees and the use of several other strategies (i.e. Expressional Complexity and Pareto tournament) in MGGP provides the possibility of having compact models which can be linearly separable and, hence they are automated post-run simplified using symbolic math

software. For the full details of MGGP and the full steps description see (Searson et al. 2010, Searson 2015).

4.2.6. Proposed hybrid model (WMGGPNN)

CWT has the advantage in the time-frequency domain in that it transforms the time series in all scales and translations unlike the DWT which does this in a dyadic form that scales are in 2^m mode. However, this produces a redundant information. For example, in this study, one of the developed models' combinations (See Methodology section) has 7 variables, every variable is transformed with CWT using 128 scales which leads to 896 scales that every scale to be used as input vector. In the case of this number of input vectors imposed into a DDM, the models performance deteriorated as most of the inputs are redundant inputs (Galelli et al. 2014, Taormina et al. 2016). The MGGP has the ability to produce an expression that contains only the input has contribution in the forecasting problem (Mehr et al. 2017). Therefore, it can be utilized as selection tool that selects the best scales that contribute in the forecasting although it is a standalone forecasting method. ANN is a robust modelling method with strong ability of capturing the non-linearity. According to that, the proposed method is as follow: The seasonality, cycling, and irregularity is captured by CWT which is applied to all variables with the highest possible scale, in this study 128; the best scales are selected by the use of MGPP by obtaining the expression from the models results; and finally, the non-linearity is cope with by imposing the best scales into ANN models to forecast the one month ahead streamflow. The schema of the proposed method is shown in Figure 4.14.

4.2.7. Application

4.2.7.1. Study area and Data

A basin named 1805 or Goksu-Gokdere, located in a large basin namely Seyhan basin was used as the study area (Figure 4.15). The basin location is between 35°34'44"E - 36°06'45"E and 37°36'07"N-38°17'20"N in the southwest of Turkey. The basin has area covers a surface of about 1790 km^2, highly varying elevation 319-2967 m, and the longest water path 192 km. The topography considered as steep as the average slope is 23%.

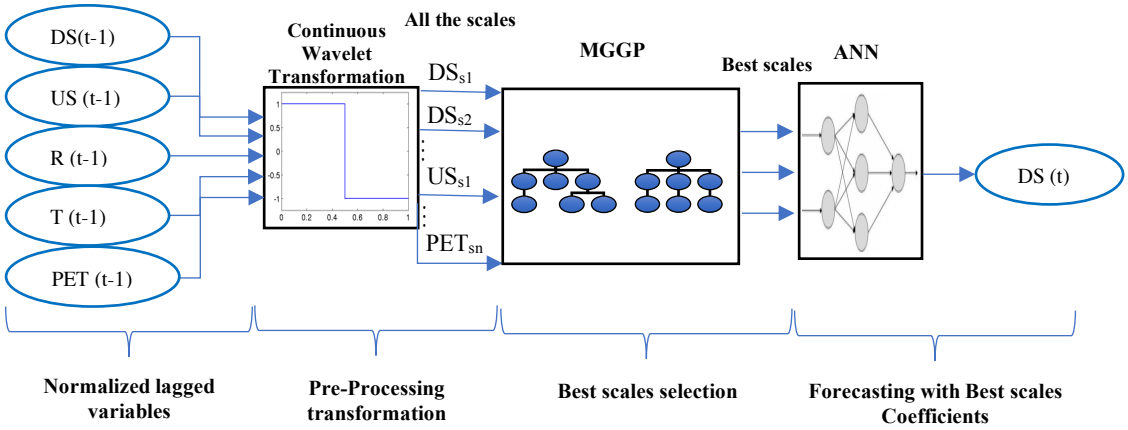


Figure 4.14. *Schematic layout of the proposed model*

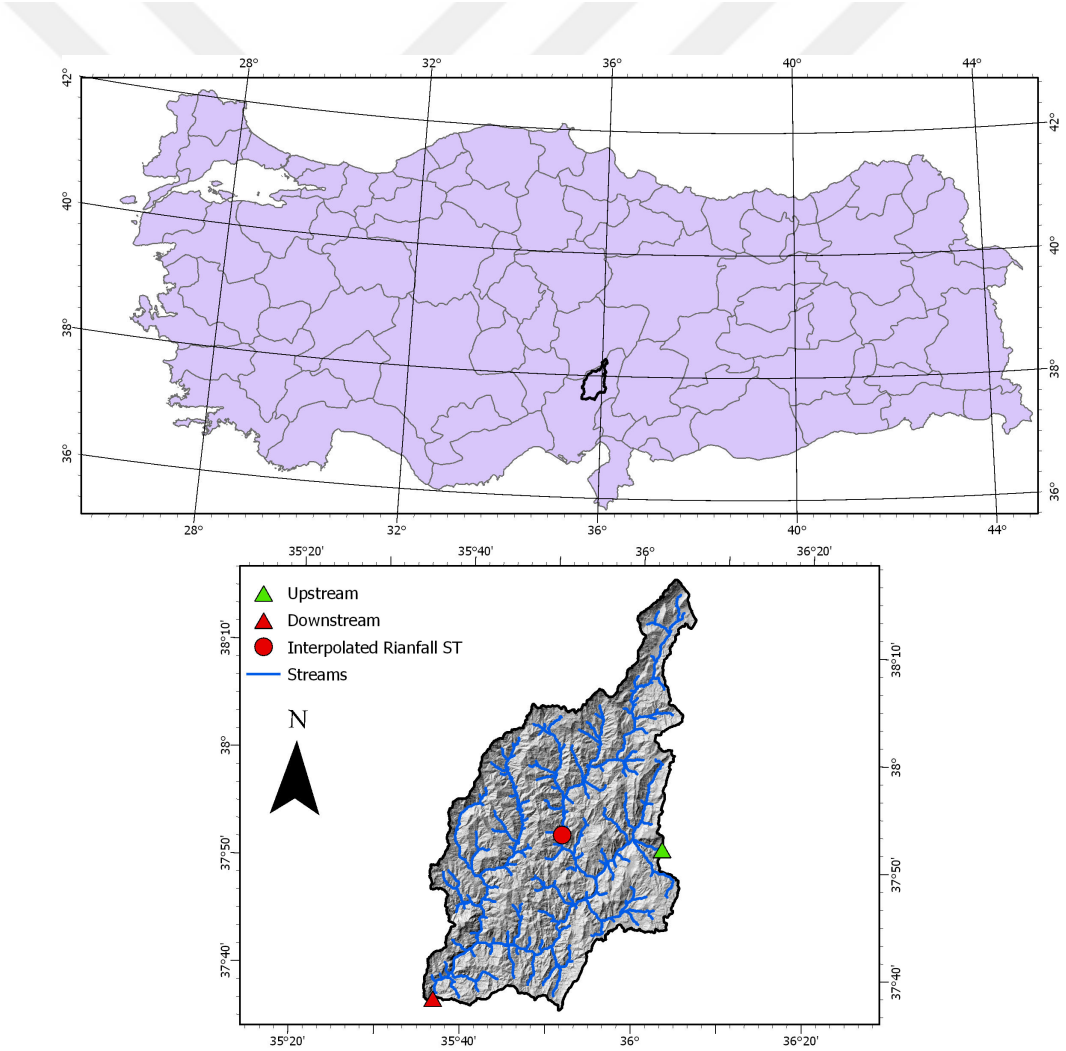


Figure 4.15. *Study area*

The monthly streamflow measured at Goksu-Gokdere station situated on the outlet of the basin for the period Feb-1973 to Sep-1994 were obtained from the general directory of Water affairs which is part of the Ministry of forests and water affairs.

In order to examine the effect of the contribution of several meteorological variables on the streamflow forecasting, Upstream, temperature, rainfall, and potential evapotranspiration were also used for the same period. Upstream data collected from the same department measured in the stations named Goksu-Himmetli. The monthly rainfall and temperature data of 17 stations was collected from the general directory of meteorology for the same period. As no meteorological station located inside the basin under the study, monthly temperature and rainfall was interpolated over a point in the middle of the basin and over the mean stream (Figure 6). The interpolation implemented using Inverse Distance Weighted (IDW) with a power of 3 which was chosen after an intensive study of the best interpolation method using the 17 stations data around the interpolated point. The examined methods are Inverse Distance Weighting (IDW), Thiessen Polygons (TP), Trend Surface Analysis (TSA), Local Polynomial Interpolation (LPI), Thin Plate Spline (TPS), and three Kriging methods: Ordinary, Universal, and Simple (OK, UK, and SK) each with several parameters and models. The evapotranspiration data collected from the same department above but with missing data more than the recorded data, therefore, it was not used in this study; but rather the data obtained from the global data CruTS3.23 (Jones et al. 2013) which is locally evaluated by (Hadi et al.) were used in this study for the same period. Time series of the used variables are shown in Figure 4.16.

In DDMs, to maintain generality, the dataset divided into two subsets, each dedicated for training and validating the network where the latter not to be used for the training process. In this study, the dataset was divided into two subsets in which 75% used for training from the and 25% used for validation. Characteristics of the datasets used in the study are summarized in Table 4.9.

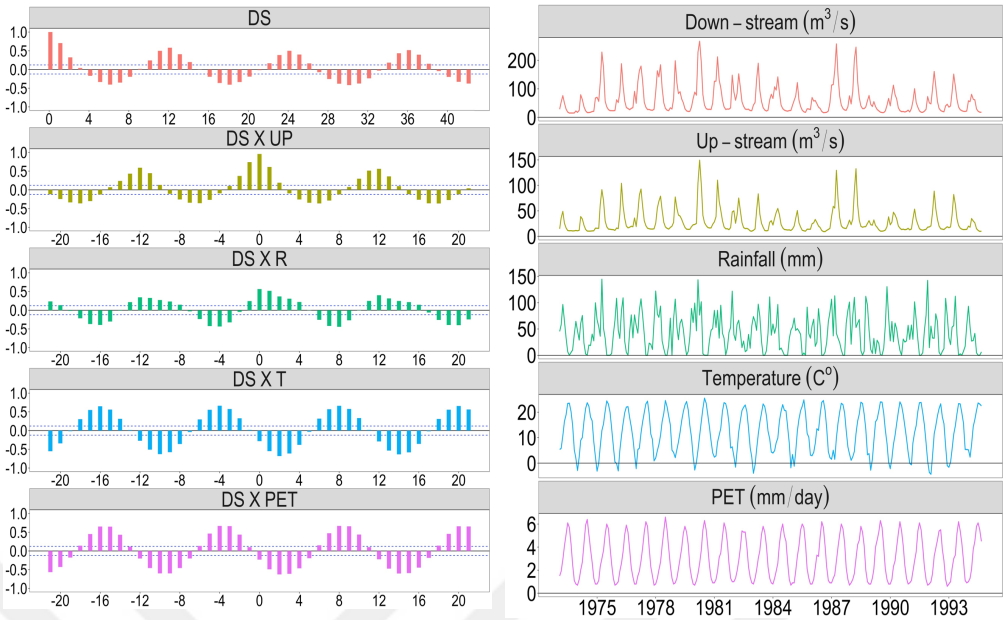


Figure 4.16. *Time series of the monthly variables (Right figures) and the ACF of the downstream and the CCF of the downstream and each variable (Left figures)*

Table 4.9. *Statistical characteristics of the meteorological and hydrological variables*

Variable	Subset	Min	Max	Range	Sum	Median	Mean	SD	C.V
Downstream	Training	14.9	267.1	252.2	12781.4	43	65.5	54.6	0.8
	Validation	16.4	161	144.6	3010.9	35.2	46.3	32.9	0.7
Upstream	Training	9.2	149.4	140.2	5984.9	20.3	30.7	25.4	0.8
	Validation	9.6	88.6	79.1	1556.6	17.8	23.9	16.4	0.7
Rainfall	Training	0	144.7	144.7	8574	40.9	44	34.9	0.8
	Validation	0	142.3	142.3	2452.1	29.7	37.7	34.2	0.9
Temperature	Training	-4	25.5	29.5	2366.2	12.6	12.1	8.2	0.7
	Validation	-4.3	24	28.4	815.3	14.6	12.5	8.6	0.7
Potential Evapotranspiration	Training	0.7	6.6	5.9	583.8	2.6	3	1.8	0.6
	Validation	0.6	6.3	5.7	206.7	3	3.2	1.9	0.6

Min: Minimum, Max: Maximum, Sum: Summation, SD: Standard Deviation, C.V: Coefficient of varation.

4.2.7.2. Methodology

First step in this study was to investigate the localized frequencies of the downstream in the time-frequency domain and that is implemented by the using the

scalogram obtained by CWT. The local correlation between the downstream and the meteorological and hydrological variables in the time-frequency domain was also examined by WCT scalogram. In addition, the phase difference was also investigated for having an indication of the lag for every variable. The scalograms of CWT and WCT are shown in Figure 4.17.

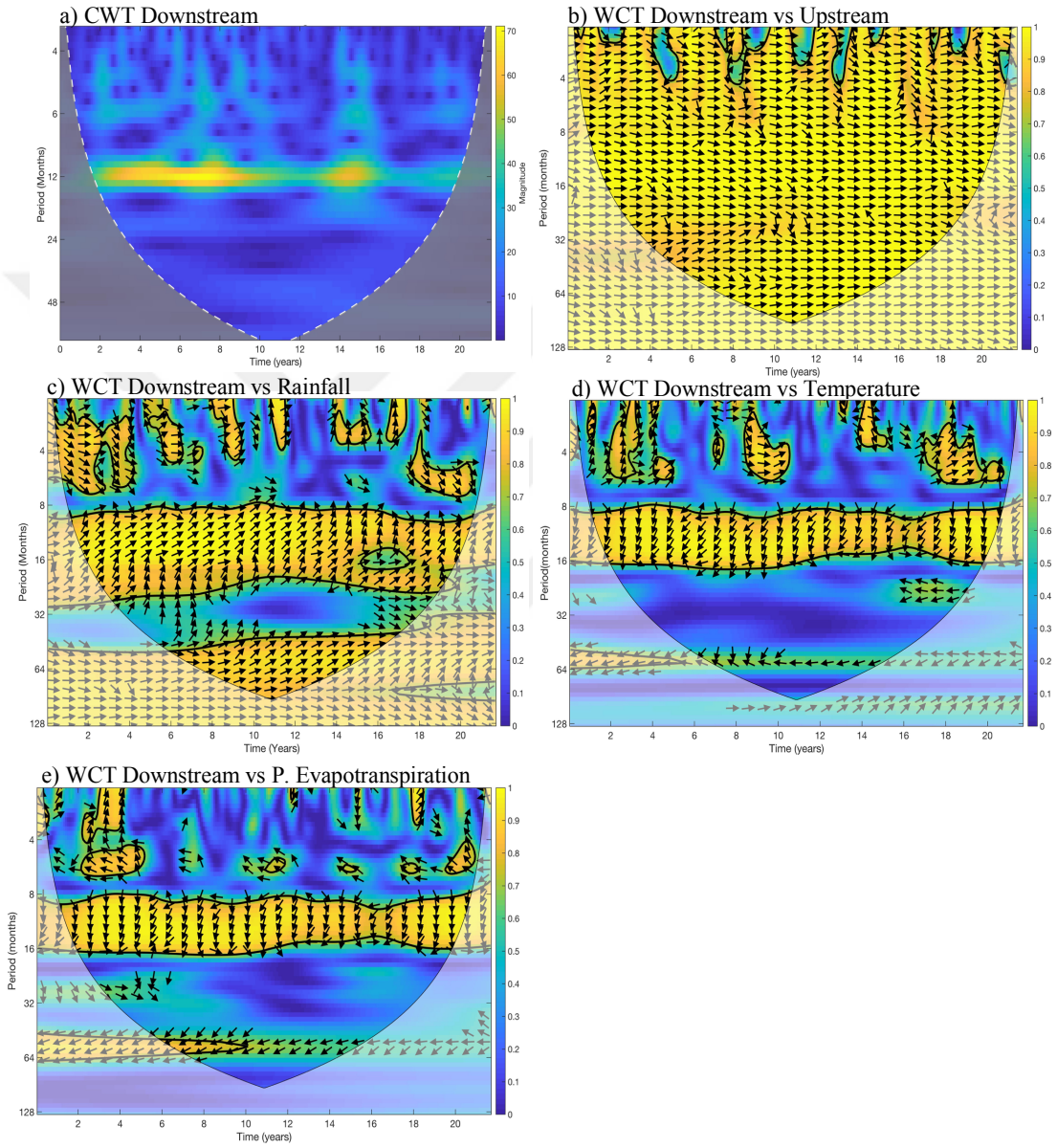


Figure 4.17. *a) The CWT scalogram of the downstream, b, c, d, and e) The wavelet coherence scalogram of the downstream with the Upstream, Rainfall, Temperature, and Potential evapotranspiration. The surrounded areas with solid black lines are significant regions in 5% significance level. The arrows represent the phase difference between the two time series. The direction of the Arrows has the meaning; right pointing: in-phase, left pointing: Anti-phase, down pointing: x leads y by 90^0 (or y lags x by 270^0), up pointing: y leading x by 90^0 (or x lags y 270^0)*

In order to obtain the optimum lags of every variable to be used as input vector, Autocorrelation Function (ACF), and Cross-Correlation Function (CCF) can be applied as they have been widely applied (Sudheer et al. 2002, Shoaib et al. 2015), but it is argued that these functions are linear and do not capture the full relationship which is mostly, in streamflow, non-linear (Nayak et al. 2004, Senthil Kumar et al. 2005, Mehr et al. 2017). Another approach used for this purpose is sequential addition of the lags that start the model with one lag, and one lag added in the second iteration and so on up to a specific number of lags and choose the number of lags according to the best performed model (Furundzic 1998, Moosavi et al. 2013). In this study, both approaches applied. Firstly, according to the Autocorrelation function of the down streamflow and the cross-correlation function between the down streamflow and the used variable shown in Figure 4.16, it can be seen that 2, 2, and 1 lag for down streamflow (DS), up streamflow (US), rainfall (R), respectively, have the correlation with DS. The temperature (T) and the Potential Evapotranspiration (PET) has a significant correlation in the no lag time but in the first lag the correlation is very small and not significant. Therefore, the sequential addition method was also used that every variable was used alone for forecasting one month ahead of DS. For down streamflow, up streamflow, rainfall the same lags of the ACF and CCF are found to be adding significant increase in the models forecasting. For temperature, and potential evapotranspiration the first lag only has a contribution to the model.

After obtaining the optimum lag of every variable, several models combinations were developed (Table 2). Starting with a model contains only the downstream, all the combinations include the lagged downstream as this component provides the auto-regressivity and the models performance is dramatically worsened without this component. The variables with their optimum lags were added to lagged DS one by one to investigate their individual effect. To examine the combined effect of these variable, models contains different combinations of the variables are also developed.

Before the data used in the different schemas, normalization was applied to all datasets to have datasets with the same range, increased accuracy, and increased pace of the calculations (Cannas et al. 2006, Nourani et al. 2009). The general formula of normalization is written as:

$$y = (b - a)\frac{x - x_{min}}{x_{max} - x_{min}} + a\,, \tag{4.23}$$

where y is the normalized value of x, x_{min} and x_{max} are the minimum and the maximum values of the x variable respectively, a and b are the lowest and highest values to which the data to be normalized. In this study, the normalization used between 0.1 and 1 and these are the values of a and b respectivley.

Table 4.10. *Developed models' combinations*

Combination No.	Abbreviation	Model Components
1	DS	DS only
2	DSUS	DS + US
3	DSR	DS + R
4	DST	DS + T
5	DSPET	DS + PET
6	DSUSPET	DS + US + PET
7	DSUST	DS + US + T
8	DSUSR	DS + US + R
9	DSUSRTPET	DS + US + R + T + PET

DS: downstream, US: upstream, R: rainfall, T: Temperature, PET: Potential evapotranspiration.

In order to evaluate the performance of the proposed model WTMGGPANN, the stand-alone and best scale schemas were also applied. In the stand-alone schema, the variables of the developed model combinations were imposed directly to ANN and MGGP. The main advantage of this schema is to assess the performance of the individual methods in which the data was imposed without any pre-processing except the normalization. In the second best-scale schema, the data was transformed using CWT for 128 scales and the scale which has the highest correlation with the one month ahead downstream flow is identified and imposed as inputs to ANN and MGGP. In the proposed model, the data are transformed using CWT for 128 scales, then all the scales entered as inputs to MGGP for selecting the best scales, and finally the selected scales are imposed into ANN for the forecasting the one month ahead downstream flow. In both ANN and MGGP, there are some parameters should be considered for getting models with higher performance. For example, the number of the hidden layers and the number of neurons in the hidden layer is important in ANN. In this study, three layered network; Input layer, hidden layer, and output layer is used as it was approved to be sufficient for such problems (Hornik et al. 1989, Kişi 2007). The number of the neurons in the hidden layer was

obtained by trial and error for every layout. The structure of the ANN in every layout is listed in Table 4.11. In MGGP several parameters are also important such as: Number of population, number of maximum generation, Mutate rate and so on (Mehr et al. 2017, Mehr et al. 2017). The parameters of the MGGP used in this study are tabulated in Table 4.12.

Table 4.11. *The structure of the ANN approach. The structure is represented as I-H-O, I: number of input neurons, H: number of neurons in the hidden layer, O: number of output neurons*

Stand-alone	Best-Scale (Haar)	Full-Scales
2-3-1	1-10-1	18-7-1
4-4-1	1-8-1	16-23-1
3-6-1	1-12-1	9-6-1
3-7-1	1-10-1	16-14-1
3-6-1	1-7-1	13-12-1
5-9-1	1-3-1	12-15-1
5-10-1	1-8-1	12-5-1
5-8-1	1-14-1	16-12-1
7-23-1	1-13-1	10-8-1

Table 4.12. *Multi-Gene Genetic Programming (MGGP) parameters*

Parameter	Value
Population	500
Generation	300
Mutation rate	0.5
Crossover rate	0.4
Reproduction rate	0.1
Maximum genes (trees)	5
Maximum depth	4

4.2.8. Results and Discussion

The main advantage of the WT on the time series in general and in hydrological data in particular is to overcome the issue of the non-stationarity confronted by DDMs. Another important advantage is representation the time series in time-frequency domain that provides a robust ability of analyzing the localized periodicities. Figure 4.17-a shows the CWT scalogram of the downstream flow. It is obvious that the highest power is around the scale 12 which is the most powerful waveband. As the data used is monthly data, this is the annual periodicity. This line shaped around the 12-period indicates that the annual signal is stationary. To examine the coherence (i.e. the co-variance) in the time-frequency

domain between the downstream flow and each of the used variables the WCT scalograms are shown in Figure 4.17 (b-e). Downstream and upstream flow have a very high localized correlation (coherence) in all the considered scales (Figure 4.17-b) and this agrees with the high correlation obtained from the cross-correlation (Figure 4.16). The phase difference generally shows in-phase that no phase difference between the two variables (Figure 4.17-b). This is expected as most of the downstream is a result of the upstream particularly the base flow which mostly flowing from the upstream. Although the area of the basin is large and other affecting variables could have more contribution, it is important to note that upstream gauge location is not in the longest water path but in the middle. Regarding the phase similarity, the monthly data used should have no difference as the difference expected between the upstream and downstream in smallest time scales; daily or smaller. In the WCT of the downstream and rainfall (Figure 4.17-c), there are scattered significant localized correlation in the scales less that 7 months, but for the scales between 11 months to 30 (i.e. approximately one year to 2.5 years) there is a harmonic coherence. The highest coherence is seen in the 12 scales which is the annual scale. After the scale of 60 months (5 years) the significant coherence is also present although most of it under the cone of influence. The phase difference shows that there is a lag about 45^0 of the downstream from the rainfall and that is due to the contribution of the rainfall. This lag is small due to the monthly time scale in which the rainfall contribution should be seen higher in the smaller time-scales. The temperature shows similar coherence in the scales less than 7 months (Figure 4.17-d). A harmonic coherence is detected between 9-15 months scales with highest being in the 12 months (i.e. Annual). The phase difference indicates that mostly there is about $220\text{-}270^0$ lag between the downstream and the temperature which may be due to that the effect of the temperature needs more time to be reflected on the streamflow. The potential evapotranspiration shows a pattern similar to the temperature and this due the fact that potential evapotranspiration is a function of the temperature and calculated using it. In general, most of the scales are important in analyzing and modelling the relation between the downstream and the other variables. Therefore, including all contributing scales would capture all the seasonality, cycling, and irregularity.

Before assessing the proposed model which proposed including all the scales, stand-alone layout was implemented. In this layout, the variables used for the forecasting of the one month downstream is entered to ANN and MGGP directly without wavelet

transformation to evaluate the performance of the DDMs on the bare variables. Although MGGP has the ability to capture the contributing inputs in the output that by including all the variables only the variables contributes in the output will be in the final expression (Nourani et al. 2011), all the models combinations were entered one by one to address the influence of every variable on the forecasting performance. In both ANN and MGGP, the result indicates that the model which contains only the downstream has poor performance in comparison with the other combinations (Figure 4.18). On the contrast of adding the rainfall which not improves the model, adding the temperature and the potential evapotranspiration increased the model performance dramatically but still not higher than the performance of the model contains all the variables which is the best performed model. It can be due to the fact that lagging the rainfall one month is too much as the effect of the rainfall does not need all this time to be reflected on the streamflow. The effect of the temperature and potential evapotranspiration needs more time to see its influence on the streamflow which is reflected in the monthly observations. As mentioned earlier, ANN and MGGP can deal with the non-linearity of the relation between the inputs and the outputs, and they are both autoregressive approaches. Therefore, the similar results of both approaches for all the models are obvious (Figure 4.18).

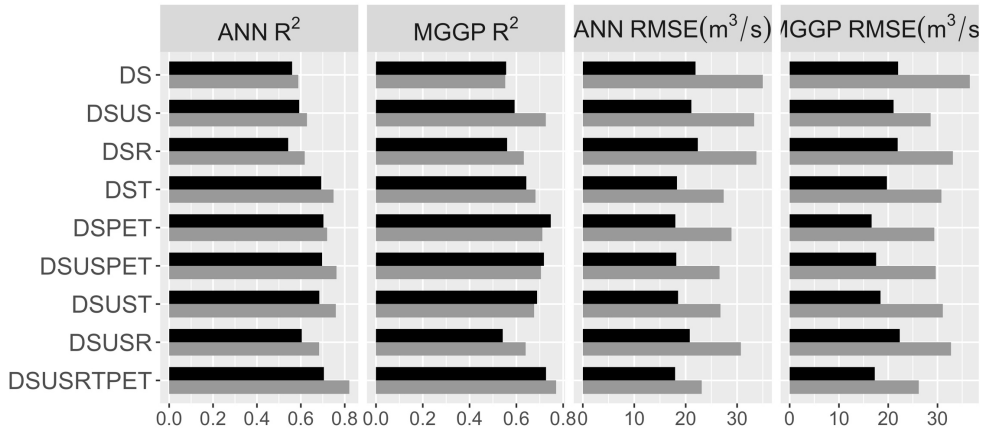


Figure 4.18. *The performance of the stand-alone layout models. Black colour: Training subset, Grey Colour: Validation subset*

ANN and MGGP stand-alone may not able to capture the seasonality of the data if no preprocessing applied to the data (Cannas et al. 2006). Therefore, WT as a solution to overcome the stationarity problem is applied as pre-processing approach. CWT can be applied in the streamflow forecasting problems by choosing the first scale (Cannas et al. 2006) or the best scale which is chosen as the scales has the highest correlation (Shoaib

et al. 2014). As the best-scale method approved as increases the performance of the models (Shoaib et al. 2014), it is adopted in this study for the comparison reason. In WT analysis, the wavelet function and the scale are two important issues that must be considered. In the Best-Scale layout, the best scale is to be considered as input to ANN and MGGP. To choose the best wavelet function, best scale was imposed to ANN after the CWT transformation by using several wavelet mother functions. The considered functions are: Haar, dmey, db2, db3, db4, db5, db6, db7, db8, Sym2, Sym3, Sym4, Sym5, Sym6, Sym7, Sym8, coif1, coif2, coif3, and coif4. The best performed wavelet function is found to be Haar (Figure 4.19) and that agrees with (Nourani et al. 2009, Nourani et al. 2011).

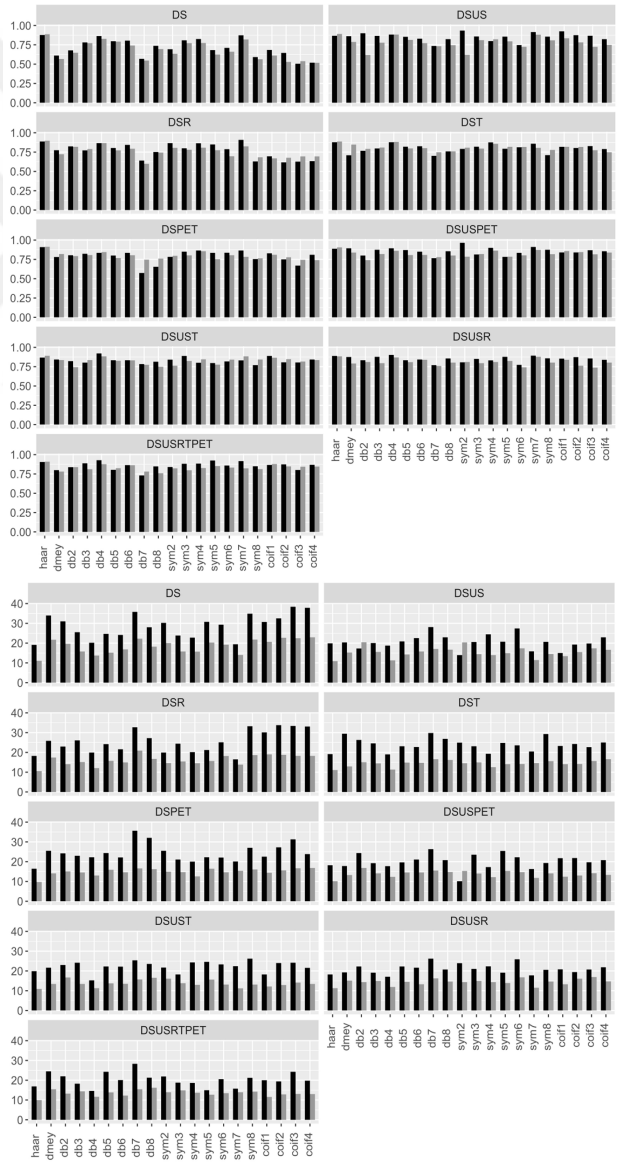


Figure 4.19. *The performance of the Best-Scale layout model of the ANN. Top figures (R2), Bottom figures (RMSE), Black colour: Training subset, Grey Colour: Validation subset*

According to Table 4.13, the results of the two approaches (i.e. ANN and MGGP) are close to each other with slight outperformance to the favor of ANN. There is no large difference between the performance of the models combinations in comparison with the first combination which contains only downstream lags with the exception that adding the potential evapotranspiration to the models improve the performance noticeably. In comparison to the stand-alone layout, the best-scale layout increased the performance of almost all the models combinations. The fact of having the highest correlated scale could improve the performance but with no much seasonality information extracted as one scale is not enough to obtain the hidden information in the time series. As seen in the WCT analysis (Figure 4.17 b-e), not only one scale has a correlation with the streamflow but several. It is also worth noting that the correlation used here is linear correlation in which the non-linear relation between the streamflow and the other variables cannot be captured.

Table 4.13. *The performance of the ANN and MGGP approaches in the Best-Scale layout using Haar mother wavelet function*

	WANNBS				WMGGPBS			
	R^2		RMSE (m^3/s)		R^2		RMSE (m^3/s)	
Comb.	Training	Validation	Training	Validation	Training	Validation	Training	Validation
DS	0.877	0.888	19.124	11.071	0.871	0.868	19.593	12.019
DSUS	0.867	0.891	19.878	10.895	0.861	0.845	20.360	13.007
DSR	0.888	0.898	18.256	10.566	0.866	0.880	20.004	11.425
DST	0.877	0.888	19.124	11.071	0.860	0.862	20.425	12.273
DSPET	0.909	0.912	16.472	9.783	0.868	0.872	19.844	11.799
DSUSPET	0.889	0.906	18.208	10.136	0.890	0.866	18.119	12.115
DSUST	0.867	0.891	19.878	10.895	0.871	0.868	19.638	12.026
DSUSR	0.888	0.884	18.267	11.268	0.879	0.871	18.971	11.852
DSUSRTPET	0.904	0.909	16.906	9.984	0.888	0.868	18.284	12.009

The use of only one scale from the CWT analysis is not preferable because of the aforementioned reasons, however, the use of all scales leads to includes redundant information and hence, the cost of calculation dramatically increased and the performance of the model decreased (Galelli et al. 2014, Taormina et al. 2016). Therefore, the WMGGPNN model is proposed. In this model, all the scales are entered to the MGGP models for the reason of selecting the contributing scales and then imposed only these scales to ANN. The number of scales selected for every model combination is listed in Table 4.11 as the ANN structure. Table 4.14 which shows the results of the proposed

model indicates high performance with high R^2 and low RMSE. In order to compare the performance of the proposed model, the percentage of change in R^2 and RMSE of the proposed model from the best scale (MGGP and ANN) and Full-scale MGGP are shown in Figure 4.20. According to the figure, there relative change percentage shows increases in the R^2 and decreases in the RMSE in all the models.

Table 4.14. *The performance of the proposed WMGGPNN model*

	WMGGPNN			
	R2		RMSE (m3/s)	
Comb.	Training	Validation	Training	Validation
DS	0.988	0.901	5.921	10.385
DSUS	0.996	0.901	3.292	10.384
DSR	0.995	0.955	3.919	6.997
DST	0.998	0.947	2.139	7.607
DSPET	0.992	0.948	4.861	7.530
DSUSPET	0.997	0.960	3.097	6.602
DSUST	0.987	0.927	6.147	8.935
DSUSR	0.930	0.924	14.465	9.108
DSUSRTPET	0.992	0.930	5.015	8.755

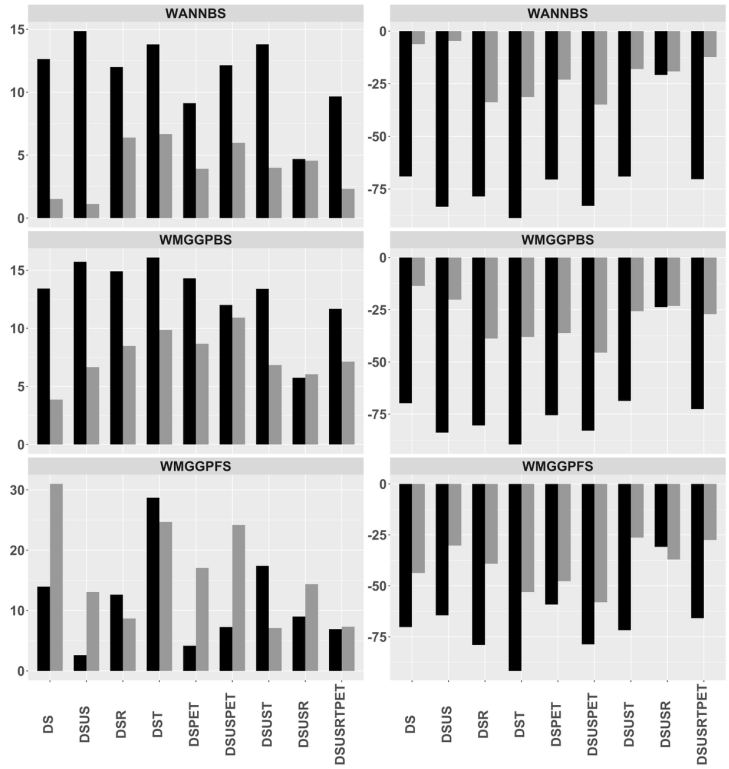


Figure 4.20. *Percentage of change in R2 (left figures), and RMSE (right figures) relatively to the proposed model WMGGPNN. Black colour: Training subset, Grey Colour: Validation subset*

The fitted versus the observed streamflow values plot of the best performed model combination (Figure 4.21) shows some of the scattered values around the best fit line in all models except the proposed model in which most of the points lay over the line. Based on that the proposed model improve the performance of the streamflow forecasting. The proposed model outperformed the best-scale of both MGGP and ANN due the fact that not only one scale but several scales are included as inputs to ANN. For example, in the first model combination in the best-scale layout the number of scales are only 2 (i.e. one for every variable), but the number of scales obtained by MGGP after entering 265 scales (i.e. all the scales for two variables) is 18 (Table 4.11). The increase in the performance can be attributable to the increase in the number of scales included in which more seasonality, cycling and irregularity information are extracted.

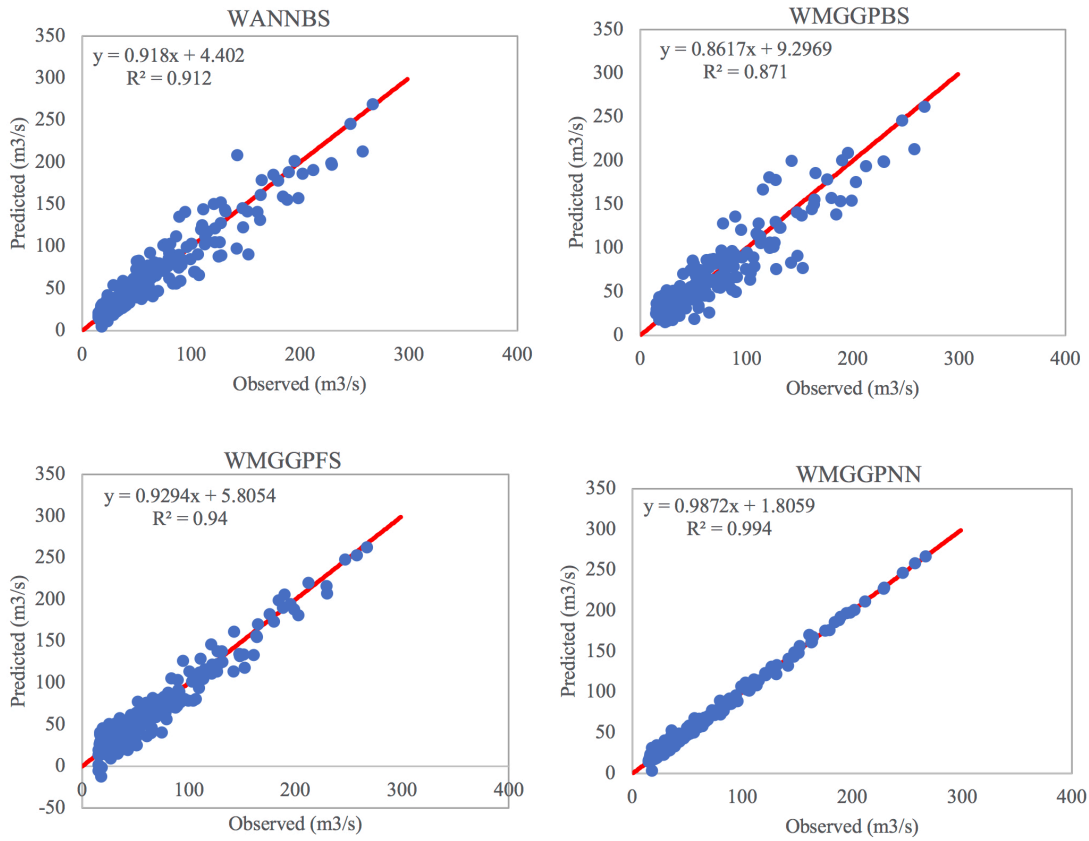


Figure 4.21. *The observed vs predicted stream flows of the best-scale, full-scale and the proposed WMGGPNN model. Redline: the best fit*

One of the most important aspects in streamflow forecasting is ability of the model in capturing the extreme/peak values. Therefore, the peak flows forecasting ability

comparison conducted using 2.5% of highest flows. As the stand-alone layout models outperformed by best-scale layout models, it was excluded from the comparison in the peak flow forecasting. The proposed model demonstrates very high accuracy in the forecasting of the peak flows being highest in the model combination includes downstream and potential evapotranspiration with R^2 equals 0.998 and RMSE 1.231 (m^3/s) (Table 4.15). In comparison with best-scale layout and full-scale MGGP model, the proposed model shows the highest performance (Figure 4.22). The highest peak flow forecasting performance is obtained in two models; the model includes downstream flow and temperature, and the model includes downstream, upstream and potential evapotranspiration. In the time of dominating underestimation of the peak values of the streamflow in general using WANNBS, WMGGPBS, and WMGGPFS, the proposed model WMGGPNN is almost perfectly overlays the observed data (Figure 4.23). Therefore, including several scales which helps in extracting the seasonality information, and using ANN which has a robust ability in dealing with the non-linearity and auto-regression, improved the models to a level that can capture most of fluctuations well.

Table 4.15. *The performance results of the proposed model WMGGPNN in forecasting the peak values*

	R2	RMSE
DS	0.952	6.941
DSUS	0.985	3.917
DSR	0.990	3.251
DST	0.998	1.231
DSPET	0.976	4.967
DSUSPET	0.997	1.767
DSUST	0.929	8.443
DSUSR	0.818	13.559
DSUSRTPET	0.986	3.797

Generally, CWT can be applied as pre-processing tool for forecasting the streamflow but this could be not better that using the DDM alone (Cannas et al. 2006, Adamowski 2008), it can also improves the runoff modelling in case of using rainfall only as inputs by choosing the best scale (Shoaib et al. 2014). In this study, the best – scale layout improved the performance of the stand-alone layout but not significantly and this is expected as the increase due to the high linear correlation is normal but in the same time still much of the information not extracted and the advantage of applying CWT not

well utilized. Including more scales by applying MGGP to the output of the CWT as selection tool leads to select the most contributing scales and throw the redundant scales. Using these scales which helps in extracting most of the hidden information with ANN dramatically improves the model performance especially in the forecasting of the peak values. Also, it was reported by (Baratti et al. 2003, Cannas et al. 2006) that monthly rainfall and temperature has at the gauge stations are not correlated with the behavior of the monthly runoff. In this study, it can be seen through all the layouts including meteorological and hydrological variables recorded in the catchment have correlation with the streamflow as they improve the performance of the models.

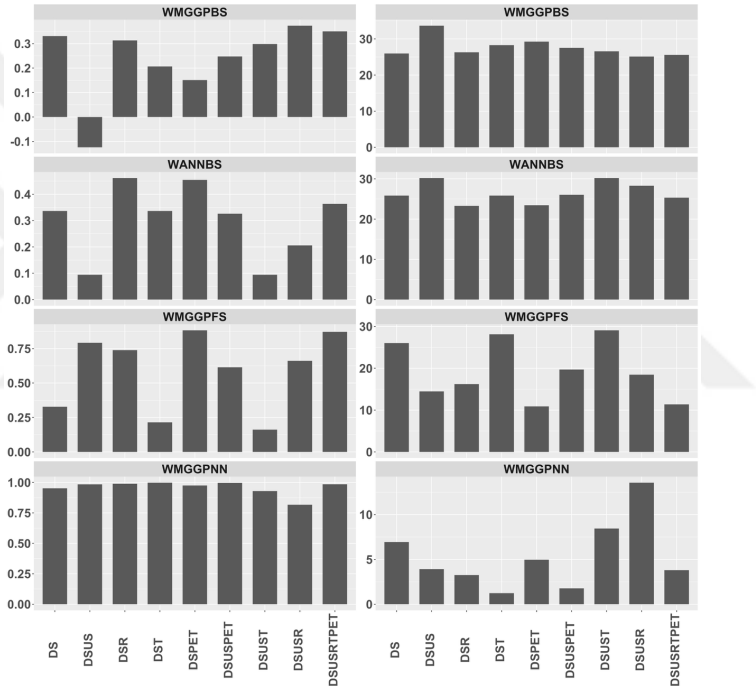


Figure 4.22. *The performance of the models in forecasting the peak values. Left figures:* $\boldsymbol{R_P^2}$*, Right figures:* $\boldsymbol{RMSE_P}$

4.2.9. Summary and Conclusion

Hydrological process such as streamflow are complex characterized by high non-linearity, seasonality, and non-stationarity. Data-driven models such as ANN and MGGP has the ability to cope with the non-linearity but not stationarity and seasonality. WT are pre-processing techniques applied to deal with non-stationarity and seasonality by transforming the data into time-frequency domain. Applying CWT is different from DWT as a large number of scales coefficients are produced in which most of them are redundant

and deteriorate the performance of any DDM. Therefore, in this study, MGGP is used as selection tool that selects only the contributing scales in order to be imposed to ANN.

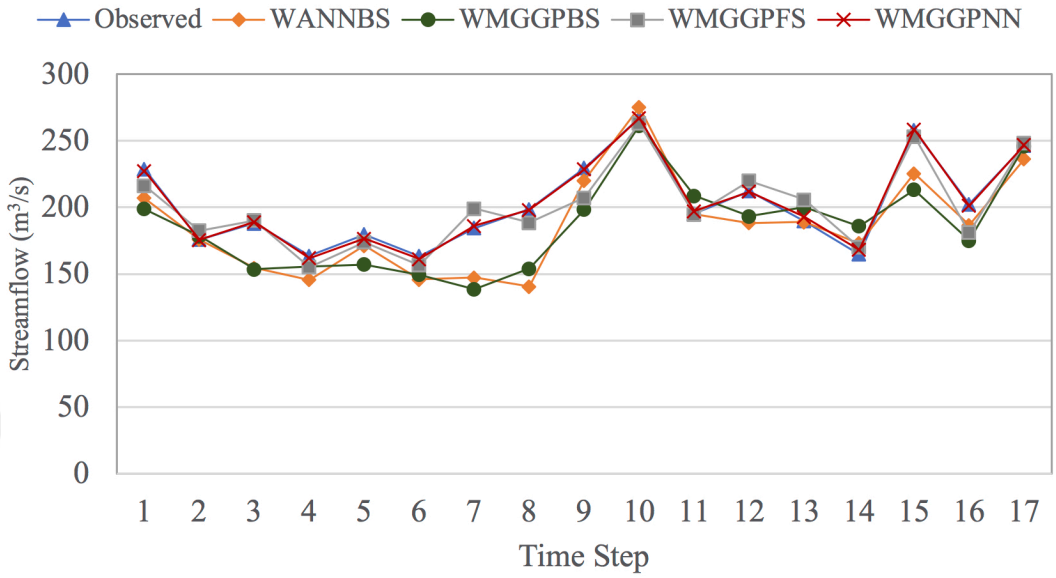


Figure 4.23. *The observed and the predicted peak flows of the best-scale, full-scale and the proposed WMGGPNN model*

To improve the ability of the proposed model in increasing the performance of the forecasting of monthly streamflow, an application is implemented by choosing a basin located in the southeast of Turkey. Monthly streamflow measured in the outlet of the basin are used. Several variables: Upstream flow, rainfall, temperature, and potential evapotranspiration are also added to the models to investigate their contribution in the downstream flow. Before the development of the model, the wavelet coherence transformation (WCT) was conducted between the downstream flow and the other variables for investigating the relation in the time-frequency domain. Bearing in the mind the output of the forecasting is one month ahead downstream flow, three layouts considered in this study: Stand-alone; the data imposed to ANN and MGGP without pre-processing, Best-Scale; the CWT scale which has the highest correlation with the downstream flow used as input to the ANN and MGGP, and Full-scale; MGGP used as selection tool to select the contributing scales and these scales were imposed to ANN. The results showed that Upstream is highly correlated with downstream in almost all the scales while other variables mostly locally correlated in the12 months scales (i.e. annual). The temperature correlation pattern is similar to one of the potential evapotranspiration as the latter is derived from the former. The stand-alone forecasting layout showed a poor

performance for all the models combinations, the best-scale layout improved the performance of the models slightly, the proposed model WMGGPNN dramatically improved the forecasting ability and lead to a model has the ability to capture the peak values excellently.

According to the WCT analysis, the streamflow is correlated with the variables: Upstream, Rainfall, temperature, and potential evapotranspiration in several scales although it shows no high linear correlation. The stand-alone layout performed poorly due to the fact that the relation between the inputs and output is not only characterized by non-linearity which can be handled by DDMs but also has the seasonality, irregularity, and non-stationarity. From this point, the need to the WT is seen. Using the best-scale layout improved the performance due to the high linear correlation which is higher than the original variable but using linear correlation is not sufficient in the existence of non-linearity. The main advantage of applying WT is to uncover the seasonality, non-stationarity and cycling information hidden in the original variable but one scale is not enough for carrying this information. Therefore, the proposed model WMGGP which after the CWT analysis selects several scales that are contributing in the output, improved the performance dramatically. The most important point is that the fluctuation in the streamflow represented by peak values are well forecasted due to the fact that the chosen scales include the sufficient seasonality and irregularity information.

According to this study, the restriction of using the CWT as pre-processing tool due to the high number of scales and redundant information for predicting the monthly streamflow is avoidable. This study is applied only on the monthly streamflow which is less predictable than daily, it is recommended to the apply the same methodology to the daily streamflow also.

CHAPTER FIVE

5. IS THE USE OF WAVELET AS PREPROCESSING IMPROVES FORECASTING?

5.1. Is Discrete and Continuous Wavelet Transformation Improving or Deteriorating The Performance of Data-Driven Models in Forecasting Monthly Streamflow

5.1.1. Abstract

Hybrid models that combining the wavelet (WT) as preprocessing tool and data-driven models (DDMs) as modelling approaches have been widely applied in the last recent years for forecasting monthly streamflow. Mostly WT is applied on the original time series before the division into calibration and testing subsets to be then entered into DDMs. This way of implementation sending some of the future information into the model which lead to the dramatic increase in the model performance. In this study, three experiments are implemented; stand-alone in which no WT applied, handcast in which data decomposed by WT, divided, and entered into Extreme learning machine (ELM) and extreme gradiant boosting (XGB), and forecast in which real forecasting is implemented supposing no future information is existing. WT is applied in the two forms discrete and continuous (DWT and CWT). A new hybrid model is proposed in this study in which XGB is used as a selection tool by utilizing the importance matrix produced by XGB to choose only the most important scales of CWT to be entered into ELM. Monthly stream flow, upstream, rainfall, temperature, and potential evapotranspiration of a basing namely 1805 located at south east of Turkey used in this study. The results showed that applying WT improving the performance in hindcast experiment especially the proposed model which showed superiority over the DWT-based models. On the contrast, WT is deteriorating the performance of the forecasting and the stand-alone models have better performance. WT increases the performance of the hindcast experiment due to the inclusion of the future information caused by the convolution of the time series while deteriorates the forecast experiment due to the border effect at the end of the time series. WT found as not a useful preprocessing technique in forecasting the streamflow.

5.1.2. Introduction

Streamflow is a crucial element of the hydrologic cycle in both global and regional scale (Makkeasorn et al. 2008, Liu et al. 2014) and it is the main source of the freshwater.

It is highly related to the droughts and floods disasters therefore, for the effective water resources management in the areas susceptible to these disasters in particular the accurate short and long-term streamflow forecasting is crucial (Kisi et al. 2011). Short-term forecasts (e.g. hourly or daily) are important for two main applications: forecasting of the floods and development of warning system (Guven 2009, Yaseen et al. 2016, Yaseen et al. 2016). The long-term forecasts (e.g. Monthly or annual) are useful for several applications such as irrigation management decisions, reservoir operations, hydro-power generations, and sediments transportation (Solomatine et al. 2009, Mehr et al. 2015, Yaseen et al. 2016).

The hydrologic models in general are divided into two main categories: data-driven models (DDMs) and physical-based model. Physical-based models are known as white box models that involves the physical process in the hydrologic modelling which leads to the need to high amount of data which is not always available. DDMs known as black box models that map the relation between the inputs and output through statistical model without involving the physical process in which not much data required. Therefore, DDMs such as artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), support vector machine (SVM), and Decision Tree (DT) have been widely used in forecasting streamflow by many researcher (Maier et al. 2000, Chang et al. 2001, Lin et al. 2006, Firat et al. 2007, Khadangi et al. 2009, Aggarwal et al. 2012, Kisi et al. 2012, Sanikhani et al. 2012, Isik et al. 2013, Akrami et al. 2014, Kişi 2015, Ghorbani et al. 2016, Shiau et al. 2016, Abdollahi et al. 2017, Mehr et al. 2017, Seyam et al. 2017). A comprehensive review of DDMs application in hydrology in general can be found in (Yaseen et al. 2015, Fahimi et al. 2016).

The streamflow forecasting is not an easy task due to the complexity in the process derives the streamflow in which this complexity is a consequence of the non-linearity and non-stationarity existed in the streamflow time series (Nourani et al. 2011, Yaseen et al. 2015, Mehr et al. 2017). DDMs has the ability to deal with nonlinearity and the non-stationarity in the mean and the variance but the main disadvantage is the failure in handling the non-stationary fluctuations (Nourani et al. 2011). Therefore, a preprocessing was found as important for improving the performance (Cannas et al. 2006). Wavelet transformation (WT, Discrete wavelet transformation (DWT) in particular), singular spectrum analysis (SSA), and empirical mode decomposition (EMD) are such

preprocessing techniques that decomposes the non-stationary and non-linear time series into several components that are easier to be modelled (Zhang et al. 2015).

In the recent years, hybrid models are developed using these preprocessing techniques with the DMMs in which the performance of these models in forecasting increased dramatically. Many researchers have investigated the application of WT (particularly DWT)-DDMs hybrid models such as (Kişi 2009, Adamowski et al. 2010, Shiri et al. 2010, Badrzadeh et al. 2013, Nourani et al. 2013, Komasi et al. 2016, Nury et al. 2017). Continuous wavelet transformation was also applied with DDMs by several researchers such as (Cannas et al. 2006, Adamowski 2008, Shoaib et al. 2014). As an example, (Badrzadeh et al. 2018) studied the hybrid model of DWT- ANFIS and found the use of DWT as preprocessing increase the models performance significantly. (Kisi et al. 2011) conducted study of investigating the improvement in the SVM model by conjugating it with DWT and reported that the conjugated model DWT-SVM has higher performance in model in forecasting monthly streamflow.

However, ,most of these studies applied the hybrid models in such a way future information are sent into the model where these must not be included in the forecasting experiment (Kim et al. 2003, Zhang et al. 2015). In the case of DWT, such studies implemented in such away that all the data are decomposed and reconstructed for getting the sub time series. Then, the sub time series are divided into calibration/training and validation subsets to be after that imposed into DDMs. In the use of CWT, the same procedure followed except choosing the most contributing scale(s) to be imposed into DDM as CWT produce redundant information. Such a way of implementation sends some amount of the future information to the model and therefore this is named as hindcasting experiment and not forecasting experiment (Karthikeyan et al. 2013, Zhang et al. 2015). One of the recent studies (Zhang et al. 2015), conducted hindcast and forecast experiment using DWT, SSA, and EMD as preprocessing and they concluded that the hybrid models performed worse than the original models. (Du et al. 2017) also investigated the hybrid models using DWT and SSA with ANN and SVM and found that the hybrid models are including some amount of future information and therefore cause incorrect increase in the performance of the original models.

The objective of this study is to conduct three experiments: stand-alone, hindcast and forecast for predicting 1 month ahead. In the stand-alone experiment no WT is applied which in both hindcast and forecast experiments the DWT and CWT applied as

preprocessing techniques with the extreme learning machine (ELM) which to the knowledge of the authors has not been applied with WT in the literature. CWT produce redundant information therefore has not been applied as much as DWT (Nourani et al. 2014) but in this study, a recent approach namely extreme gradient boosting (XGB) is applied for choosing the contributing scales only from the CWT scales to be imposed to the ELM although XGB itself is a stand-alone model.

5.1.3. Preprocessing technique

5.1.3.1. Wavelet transformation (WT)

Wavelet transformation is divided into two categories Continuous (CWT) and Discrete (DWT). CWT can be defined as the summation through the time of the signal multiplied by shifted and scaled versions of wavelet known as mother wavelet $\psi_{a,b}(t)$:

$$W_x(a,b) = |a|^{-1/2} \int_{-\infty}^{+\infty} f(t)\, \psi_{a,b}^{*}\left(\frac{t-b}{a}\right) dt, \tag{5.1}$$

where $W_x(a,b)$ are the CWT coefficients, * represents the complex conjugate of the function, a and b are the parameters known as the scale and shifting parameters respectively. According to Eq.(5.1), for every scale there are a number of coefficients equal to the length of the original time series.

DWT is thought as the dyadic sampling form of CWT with $a = 2^j$ and $b = k2^j$ where k is the location index and j is the decomposition level. The discrete wavelet is written as:

$$\psi_{j,k}(t) = 2^{-j/2}\psi\left(2^{-j}t - k\right), \tag{5.2}$$

and the DWT is written as:

$$W_\psi(j,k) = 2^{-j/2} \int_{-\infty}^{+\infty} f(t)\, \psi_{j,k}^{*}\left(2^{-j}t - k\right) dt\,. \tag{5.3}$$

For a discrete time series x_n, the DWT is:

$$W_\psi(j,k) = 2^{-j/2} \sum_{n=0}^{N} \psi\left(2^{-j}t - k\right) x_n\,, \tag{5.4}$$

where $W_\psi(j,k)$ are the DWT coefficients and to reconstruct the original time series the inverse DWT is implemented given as:

$$x_n = A + \sum_{j=1}^{J} \sum_{k=0}^{2^{J-j}-1} W_\psi(j,k) 2^{-j/2}\left(2^{-j}t - k\right), \tag{5.5}$$

and can be written as:

$$x_n = A(t) + \sum_{j=1}^{J} W_j(t), \tag{5.6}$$

where $A(t)$ is the approximation at level J and $W_j(t)$ are the details coefficients at levels $j = 1, 2, \dots, J$.

5.1.4. Data-Driven models

5.1.4.1. Extreme learning machine (ELM)

ELM is an emerging DDM of single hidden layer feed forward networks (SLFNs) of ANN firstly proposed by (Huang et al. 2006). ELM overcomes the disadvantage of over fitting, local minima, and slow learning of the traditional Back-propagation ANN. Since the development of the algorithm, it has been applied in several applications in hydrological modeling by several researchers (Abdullah et al. 2015, Mohammadi et al. 2015) and for streamflow forecasting in particular such as (Li et al. 2014, Deo et al. 2016, Lima et al. 2016, Yadav et al. 2016). One of the recent studies is conducted by (Yaseen et al. 2016) who compared the performance of ELM with SVM and generalized regression neural network and concluded the superiority of the ELM over other approaches in terms of the speed and the accuracy.

An ELM model is developed using a set of training samples as $\{(x_1, y_1), \dots, (x_t, y_t)\}$ where x_t is the independent variable and y_t is the dependent/target variable. In this study, the dependent variable which is the input vector denoted as $x_1, x_2, \dots, x_t$ defined as the lagged streamflow, or any another variable used in this study; variable used as inputs explained in the used data section. The targets or the output vector $y_1, y_2, \dots, y_t$ represent the observed one step ahead streamflow. For set of N training samples (i.e. $t = 1, 2, \dots, N$) in which $x_t \in \mathbb{R}^d$ and $y_t \in \mathbb{R}$, a SLFN with H hidden nodes is mathematically expressed as (Huang et al. 2006):

$$\sum_{i=1}^{H} B_i \, g_i(\alpha_i . x_t + \beta_i) = z_t, \tag{5.7}$$

where $B \in \mathbb{R}^H$ is the weights that to be estimated in the output layer between the hidden layer nodes and the model output (i.e. $Z(z_t \in \mathbb{R})$), $G(\alpha, \beta, x)$ is the activation function of the hidden layer, α_i represents the weights of the randomized layers, β_i represents the

biases of these randomized layers, i is the index of the specific node in the hidden layer, and d is the number of inputs.

In this study, the sigmoid function used as activation function:

$$G(x) = \frac{1}{1 + \exp(-x)}, \tag{5.8}$$

and a linear transfer function used in the output layer. According to (Huang et al. 2006), an ELM model with suitable number of hidden neurons, randomized input layer weights α, and randomized hidden layer biases β, can lead to a zero error (Eq.(5.9)) which, hence, yield to realization that the weights of the output layer B can be obtained analytically for any training sample.

$$\sum_{t=1}^{N} \|z_t - y_t\| = 0. \tag{5.9}$$

The B values of an N input-output training samples can be estimated using a system of linear equations:

$$Y = GB, \tag{5.10}$$

in which

$$\begin{aligned} G(\alpha, \beta, x) &= \begin{bmatrix} g(x_1) \\ \vdots \\ g(x_N) \end{bmatrix} \\ &= \begin{bmatrix} g_1(\alpha_1 . x_1 + \beta_1) & \cdots & g_L(w_H . x_1 + \beta_H) \\ \vdots & \cdots & \vdots \\ g_1(\alpha_N . x_N + \beta_1) & \cdots & g_L(w_H . x_N + \beta_F) \end{bmatrix}_{N \times H}, \end{aligned} \tag{5.11}$$

and

$$B = \begin{bmatrix} B_1^T \\ \vdots \\ B_H^T \end{bmatrix}_{H \times 1}, \tag{5.12}$$

and

$$Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{N \times 1}, \tag{5.13}$$

where G here known as the hidden layer output, T is the transpose of the matrix. The output weights $\hat{B}$ can be estimated by inverting the matrix of the hidden layer using Moore-penrose generalized inverse function (+):

$$\hat{B} = G^{+}Y. \tag{5.14}$$

Eventually, the estimated values $\hat{y}$ (i.e. represents the one month ahead streamflow in this study) can be determined by:

$$\hat{y} = \sum_{i=1}^{H} \hat{B}_i \, g_i(\alpha_i . x_t + \beta_i) \, . \tag{5.15}$$

5.1.5. Extreme Gradient Boosting (XGB)

XGB was first developed by (Chen et al. 2016), as an efficient, fast, and scalable implementation of the Gradient tree boosting algorithm developed by (Friedman 2001). XGB is classified as supervised learning technique uses an ensemble of decision trees (Ayumi 2016). It can be used for both classification and regression problems. Being a new algorithm, to the knowledge of the authors it has not been applied in streamflow forecasting in general and as an input selection tool with wavelet in particular. In this study, this algorithm is applied as a model itself and as a selection tool for the important scales of the CWT by taking the advantage of the feature importance characteristic of this algorithm.

XBG algorithm is a set of classification and regression trees (CART) as CART is different from DT in that each of the leaves has a real score that helps in richer interpretation that beyond the classification which is the case in DT. In CART, a single tree is used which is not strong enough in practice, therefore, an ensemble of multiple trees is proposed. For a training data (with multiple feature) x_i (i.e. lagged downstream, upstream, rainfall, temperature, and potential evapotranspiration in this study) to predict a target variable y_i (i.e. one month ahead downstream), the mathematically expression of the ensemble of K number of trees can be written as:

$$\hat{y} = \sum_{k=1}^{K} f_k \, (x_i), f_k \in F \, , \tag{5.16}$$

where f_k is a function in a set of all possible function in CARTs, F the set of these functions. The main objective function is the optimization of:

$$obj(\theta) = \sum_{i=1}^{n} l(\, y_i, \hat{y}_i) + \sum_{i=1}^{t} \Omega(\, f_i) \, , \tag{5.17}$$

where l part is the training loss function and Ω is the regularization part which is used to control the complexity and prevent the overfitting.

Since training the functions f_i of all trees in one time is not easy, an additive training is considered in which training one and fix the learned tree and add one new tree and learn again and so on. Having the prediction value at step t as $\hat{y}_i^t$ lead to:

$$\hat{y}_i^t = \sum_{k=1}^{t} f_k\,(x_i) = \hat{y}_i^{t-1} + f_t\,(x_i)\,. \tag{5.18}$$

Adding one tree in every iteration, the selected tree is the tree that minimizes the objective function in Eq. (5.17) after the substitution of the new predicted values from Eq.(5.18). Considering the mean squared errors (MSE) as the loss function, the objective function becomes:

$$obj^t = \sum_{i=1}^{n} (\,y_i - (\hat{y}_i^{t-1} + f_t\,(x_i)))^2 + \sum_{i=1}^{t} \Omega(\,f_i)\,, \tag{5.19}$$

and that can be written as:

$$obj^t = \sum_{i=1}^{n} [\,2(\,\hat{y}_i^{t-1} - y_i) f_t\,(x_i)\, + f_t\,(x_i)^2 + \Omega(f_t) + constant\,. \tag{5.20}$$

MSE in both first order or quadratic form is friendly but functions such as logistic is more complex, therefore, Taylor expansion up to the second order is applied:

$$obj^t = \sum_{i=1}^{n} [\,l\big(y_i, \hat{y}_i^{t-1}\big) + g_i\, f_t\,(x_i)\, + \frac{1}{2} h_i\, f_t\,(x_i)^2 + \Omega(f_t) + constant, \tag{5.21}$$

where $g_i = \partial_{\hat{y}_i^{t-1}} l(y_i - \hat{y}_i^{t-1})$, and $h_i = \partial^2_{\hat{y}_i^{t-1}} l(y_i - \hat{y}_i^{t-1})$. All the steps, have been done in the training loss function part; considering the regularization part and bearing in mind the tree is defined as $f_t(x) = w_{q(x)}$, the regularization is written as:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2\,, \tag{5.22}$$

where w represents the vector of scores on leaves, q represents the assigning function of each point in the data to the corresponding leaf, T is the number of leaves. After the inclusion of the tree function and the regularization part, and remove the constants from Eq.(5.21) the objective value of the t^{th} tree can be written as:

$$\begin{aligned} obj^t &\approx \sum_{i=1}^{n} [g_i\, w_{q(x_i)} + \frac{1}{2} h_i w_{q(x_i)}^2] + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \\ &= \textstyle\sum_{j=1}^{T} [\,\big(\sum_{i\in I_i} g_i\big) w_j + \frac{1}{2}\big(\sum_{i\in I_i} h_i + \lambda\big) w_j^2 + \gamma T\,, \end{aligned} \tag{5.23}$$

where $I_i = \{i|\, q(x_i) = j\}$ represents the indices of data points assigned to the j^{th} leaf. Defining $G_j = \sum_{i \in I_i} g_i$ and $H_j = \sum_{i \in I_i} h_i$, the objective function is written simply as:

$$obj^t = \sum_{j=1}^{T} [\, G_j w_j + \frac{1}{2}\big(H_j + \lambda\big) w_j^2 + \gamma T \,. \tag{5.24}$$

For a given structure of tree $q(x)$, the best w_j and objective reduction which are used for measuring how good that structure is can be obtained by:

$$w_j^* = -\frac{G_j}{H_j + \lambda}, \tag{5.25}$$

$$obj^* = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j}{H_j + \lambda} + \gamma T \;, \tag{5.26}$$

Enumerate all the possible trees and pick the best one is not an intractable way of solving the problem, therefore, Eq.(5.27) is used for optimizing one level of the tree a time and in particular it is used to obtain the score gain in splitting a leaf into two leaves. Eq. (5.27) consists of the score on the new left leaf (L), the score of the new tight leaf (R), the score in the original leaf, and the regularization term γ. In case of the gain is less than γ it would be better not to add the branch.

$$Gain = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right] - \gamma \,. \tag{5.27}$$

Using this gain information in the inclusion of the variable that measuring the gain of including a certain variable over the whole trees network, the feature importance is obtained from this algorithm. This characteristic is utilized in this study for proposing new selection tool of the important scales in the CWT for forecasting streamflow.

5.1.6. Proposed modeling Schema:

In this study, two modeling schemas are used. The first schema is that the inputs are transformed using DWT (CWT) and all the reconstructed sub time series (scales) are imposed to the XGB and ELM (only XGB for CWT). The Second schema is that the inputs are transformed using the CWT using the highest possible scales and all the scales are entered into XGB while only the important scales obtained from XGB are imposed into ELM according to their order of importance. These two schemas are applied in both hindcast and forecast experiments (see the experiment section). In the following, the

proposed schema is explained in detail. The framework of the proposed schema is shown in Figure 5.1.

CWT transforms the time series in the time-frequency domain into several number of scales for the same length of the original time series. using big number of scales produces too many scales that in case of imposing them into any DDMs produces bad models due to the inclusion of redundant information that deteriorates the models performance (Galelli et al. 2014, Taormina et al. 2016). XGB as mentioned earlier, has the ability of producing the ordered feature importance (i.e. scales in this specific study) which shows the importance of a specific feature in modeling the dependent variables starting with most important feature and ending with the less important. Features not adding any gain to the model are not even shown in the importance matrix. In this study, XGB was proposed as selection tool beside being a modelling approach itself. In both hindcast and forecast experiment, the proposed schema is applied as in the following manner:

- The lags of the used meteorological and hydrological variables $V = v_1, v_2, \dots, v_n$ are transformed using CWT with the highest possible scale 128 scales.
- For every variable, 128 time series $\{S_{v_i} = s_{vi1}, s_{vi2}, \dots, s_{vi128}, i = 1, 2, \dots, n\}$ are obtained. All the scales (i.e $TS = [S_{v_1}\ S_{v_2}\ \dots\ S_{v_n}]$) are then entered into the XGB model as inputs and the one month ahead downstream flow Q_{t+1} as output.
- The evaluation criteria are obtained for the XGB model for evaluating its performance.
- The importance of the scales is obtained from XGB in order that the most important scale is first and least important scale is last. $M = \left(s_{vij}^{m}\right)$, where i represent the variable, j represent the scale of that variable, and m represents the order of the importance of the variable and that takes values according to the number of the features/scales chosen in the importance matrix M.
- Only the important scales are then chosen and imposed to ELM in a sequential manner that the first model includes only the most important scale $Q_{t+1} = f(s_{vij}^{1})$, the second model includes the first and the second

important scales $Q_{t+1} = f(s^1_{vij}, s^2_{vij})$and this procedure continuous till the last model which contains all the scales in the importance matrix $Q_{t+1} = f(s^1_{vij}, s^2_{vij}, \dots, s^m_{vij})$.

- The highest performed model is chosen with their scales to be compared with the two others schemas; stand alone and DWT preprocessed models.

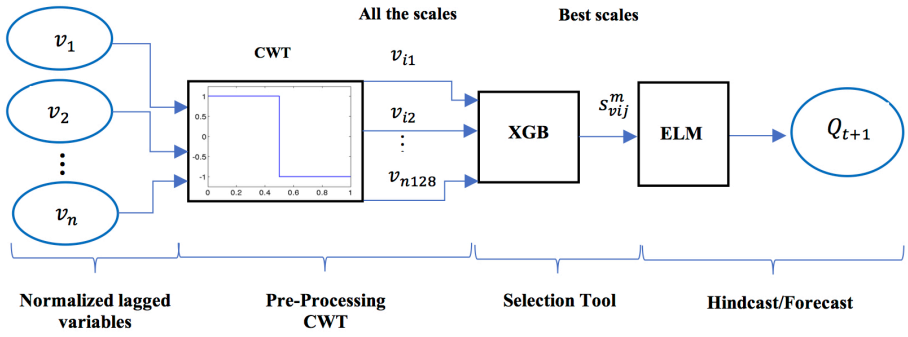


Figure 5.1. *Framework of the proposed model schema*

5.1.7. Experiment

5.1.7.1. Study area and data

In order to examine the three experiments conducted in this study, a basin located in the south east of Turkey, namly 1805 or Goksu-Gokdere was chosen (Figure 5.2). The basin located between 37°36'07"N-38°17'20"N and 35°34'44"E - 36°06'45"E. The area of the basins covers about 1790 km^2, with steep average slope 23%, highly varying elevation 319-2967m, and log water path 192 km.

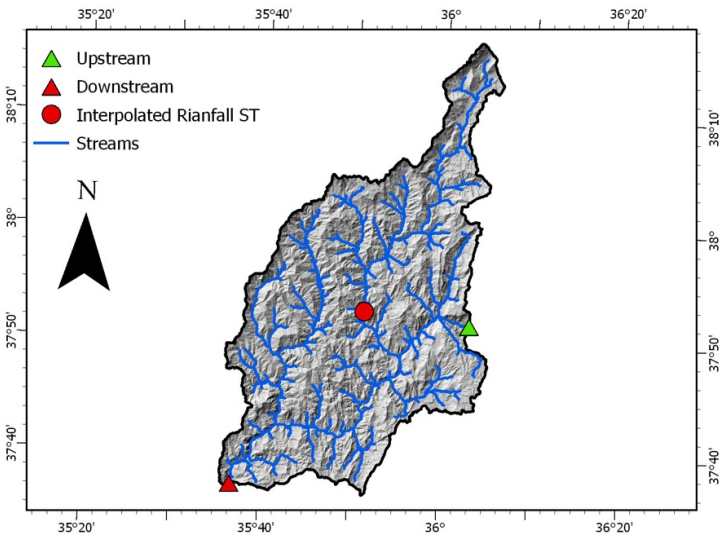


Figure 5.2. *Study Area*

The streamflow measured at Goksu-Gokdere and upstream measured at Goksu-Himmetli stations collected from the general directory of Water affairs which is part of the Ministry of forests and water affairs for the period February-1973 to September-1994. The rainfall and temperature observations were interpolated by the Inverse Distance Weight method using 17 stations around the basin due to the non-existence of any meteorological station inside the basin. The potential evapotranspiration was obtained from the CruTS3.23 (i.e. locally assessed by (Hadi et al. 2017)) as the collected observations from the general directory of Water affairs has missing values more than the available values. The time series of the used variables are shown in Figure 5.3.

5.1.7.2. Model development

One of the main issues in time series modeling is to identify the number of lags to be used in the model which increasing the model performance. One of the most common method is to use Autocorrelation Function (ACF), and Cross-Correlation Function (CCF) (Sudheer et al. 2002, Shoaib et al. 2015). The method criticized by several researchers such as (Nayak et al. 2004, Senthil Kumar et al. 2005, Mehr et al. 2017) in that the relation between the variables or lags of one variable could be non-linear and that cannot be captured by ACF and CCF which are linear. Another method is to conduct a sequential approach in which one lag is added every iteration till the model performance not increasing or start decreasing then this lag identified as the optimum lag (Furundzic 1998). In this study, both of these methods were applied, and the optimum lags found as 2, 2, 1, 1, and 1 for downstream flow, upstream flow, rainfall, temperature, and evapotranspiration respectively.

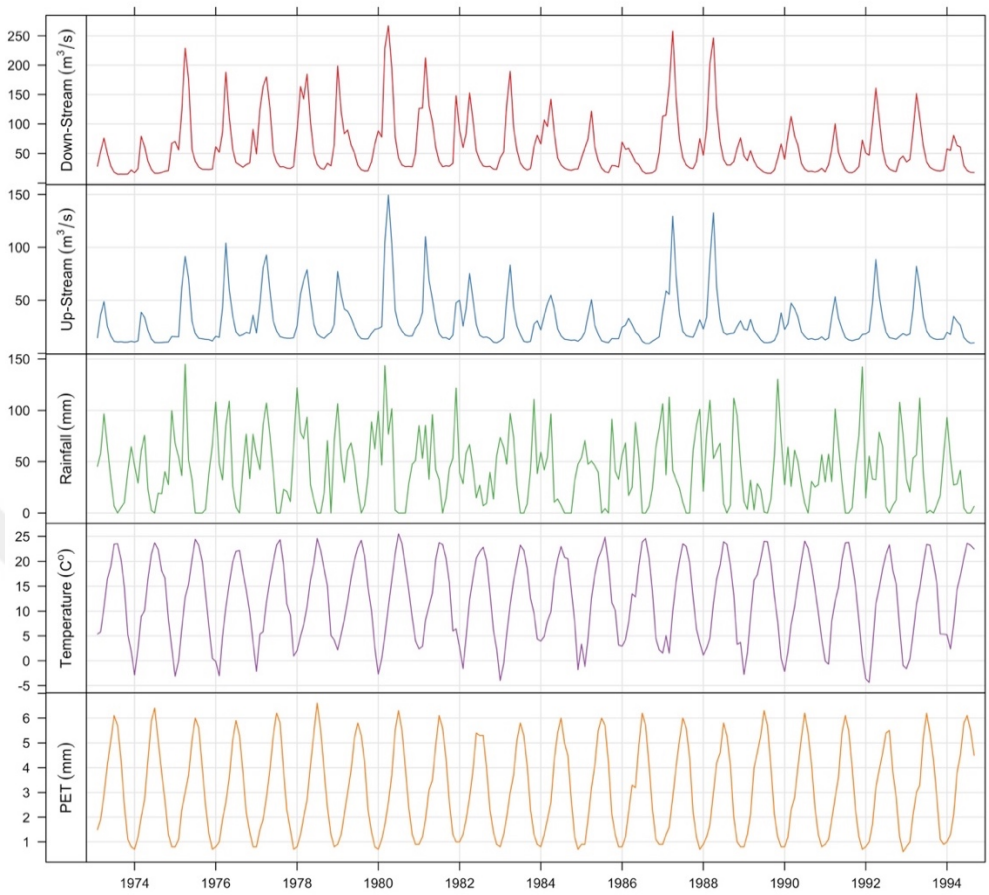


Figure 5.3. *The time series of the variables used in the development of the models*

After the determination of the optimum lags, several models combination developed starting with the model contains only the downstream variables and the other variables were added in the following models one by one to investigate their effect on the modeling performance and finally a model contains all the variables was developed (Table 5.1). It is worth mentioned that the variable in the developed model refers not to the variable itself but to its optimum lag(s).

The normalization was applied to all the data set for every model combination using:

$$y = (b - a)\frac{x - x_{min}}{x_{max} - x_{min}} + a\,, \tag{5.28}$$

where x, x_{min} and x_{max} represents the minimum and the maximum values of the variable x respectively, a and b represents the lowest and highest values of the normalized data. the normalization was implemented between 0.1 and 1 in this study.

The DWT and CWT uses a mother wavelet function to transform the time series into the time-frequency domain. There are a number of functions can be used and the identification of the function that gives the highest performance is another task in WT-based hybrid models. In this study, db7 and Haar (i.e. db1) are determined as the best functions for DWT and CWT respectively. Another important issue in DWT is the determination of the optimum decomposition level. The 2 levels was found as best in this study. For CWT, the 128 scales is used as highest possible scales based on the number of data points utilized.

Two schemas were used in this study: I schema; DWT-ELM, DWT-XGB, and CWT-XGB, II schema; CWT-XGB-ELM.

Table 5.1. *The developed combinations*

Combination No.	Abbreviation	Model Components
1	DS	DS only
2	DSUS	DS + US
3	DSR	DS + R
4	DST	DS + T
5	DSPET	DS + PET
6	DSUSRTPET	DS + US + R + T + PET

DS: downstream, US: upstream, R: rainfall, T: Temperature, PET: Potential evapotranspiration.

5.1.7.3. Stand-alone experiment

The first experiment is to forecast the one month ahead downstream flow using the bare models without WT. In this experiment the lagged variables are imposed to the XGB, and ELM directly as inputs and one month ahead flow as output. In terms of dividing the time series, the normalized data divided into 75% for training and 25% for testing and that is obviously forecasting because of the inclusion of the Q_{t+1} as output and taking the first model combination as example the inputs are the current Q_t or one month earlier Q_{t-1} downstream flow.

5.1.7.4. Hindcast experiment

A large number of studies which apply a WT-based hybrid models used a hindcast experiment and incorrectly name it as forecasting (Zhang et al. 2015, Du et al. 2017). In hindcast experiment, all the data set are transformed with WT and then divided into

training and testing subsets to be imposed after that to the DDMs. The framework of the hindcast experiment is shown in Figure 5.4 and the steps are:

1. After obtaining the optimum lags and normalizing them, a set of variables are obtained for every model combination $V = v_1, v_2, \dots, v_n$.
2. Decompose the normalized lags of the variables using a certain mother wavelet, a specific decomposition level for DWT, and the highest possible scale for CWT. For DWT, the decomposed coefficients are reconstructed to obtain the sub time series namely A_j, and $D_1, D_2, \dots, D_j$ a set of sub time series are obtained for each variable $S_{vi}^{dwt} = \{A_j, D_1, D_2, \dots, D_j\}$, j is the decomposition level. For CWT, $\{S_{vi}^{cwt} = s_{vi1}, s_{vi2}, \dots, s_{vij}, i = 1, 2, \dots, n\}$, where j is the highest possible scale and s_{vi1} is a vector of the coefficients of first scale of the i variable.
3. The sub time series/scales for all variables are gathered $TS^{dwt} = [S_{v1}^{dwt}\ S_{v2}^{dwt}\ \dots\ S_{vn}^{dwt}])$, $TS^{cwt} = [S_{v1}^{cwt}\ S_{v2}^{cwt}\ \dots\ S_{vn}^{cwt}])$ for DWT and CWT respectively and divided into two subsets 75% for calibration TS_{calib}^{dwt} (TS_{calib}^{cwt}) and 25% for testing TS_{test}^{dwt} (TS_{test}^{cwt}) for DWT (CWT).
4. For the second schema, the TS_{calib}^{dwt} (TS_{calib}^{cwt}) are imposed as inputs into ELM and XGB for training and $Q_{t+1,calib}$ as output, and then the models with the best parameters are chosen. The TS_{test}^{dwt} (TS_{test}^{cwt}) and $Q_{t+1,test}$ is imposed into the best models for testing and the evaluation criteria are obtained. In the third schema, the proposed model (the details are in Proposed modelling schema section) is applied in which the feature importance obtained from XGB is utilized that shows the importance of the scales in order $s_{vij}^1, s_{vij}^2, \dots, s_{vij}^m$ where 1 is the highest important scale while m is the least important scale. The training of ELM uses the matrix consists of only the important scale of the calibration subset $TS_{calib}^{cwt,best} = [S_{vij}^{cwt,1}\ S_{vij}^{cwt,2}\ \dots\ S_{vij}^{cwt,m}]$ and the testing is implemented using $TS_{test}^{cwt,best}$.

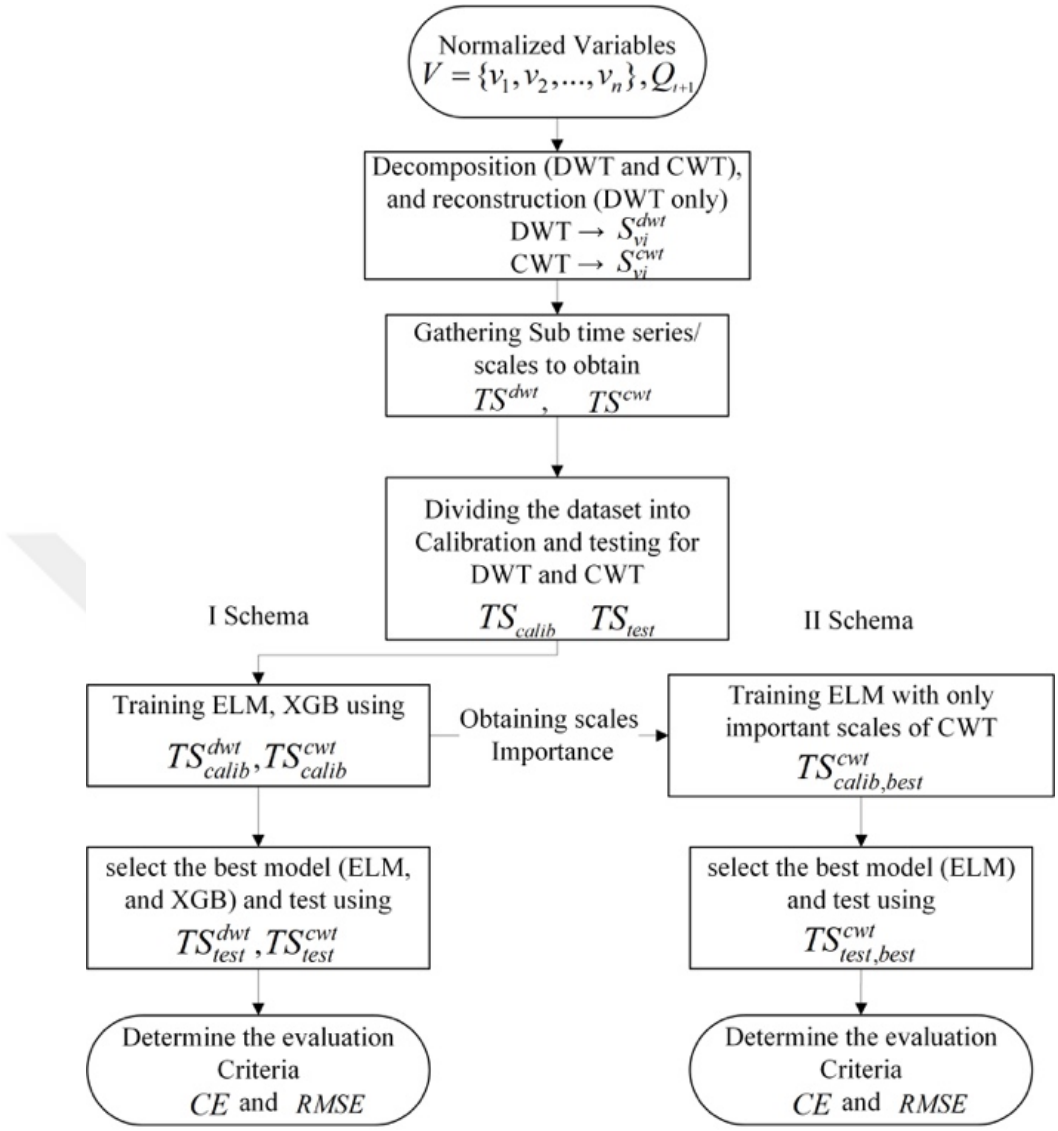


Figure 5.4. *The framework of the hindcast experiment*

5.1.7.5. Forecast experiment

In this experiment, a real forecast experiment is conducted by not including any of the future information in the model. The framework of the forecast experiment is shown in figure 5 and the steps are:

1. After obtaining the optimum lags and normalizing them, a set of variables are obtained for every model combination $V = v_1, v_2, \dots, v_n$.
2. Divide the data into two subsets 75% for calibration V_{CALIB} and 25% for testing V_{TEST}. The targets Q_{t+1} are also divided into 75% for calibration $Q_{t+1,CALIB}$ and 25% to get $Q_{t+1,TEST}$.

3. Let V_{CALIB} and $Q_{t+1,CALIB}$, series denoted by $1, 2, \dots, k$.
4. Decompose the V_{CALIB} using a certain mother wavelet, a specific decomposition level for DWT, and the highest possible scale for CWT. For DWT, the decomposed coefficients are reconstructed to obtain the sub time series namely A_j, and $D_1, D_2, \dots, D_j$. A set of sub time series are obtained for each variable $S_{vi}^{dwt} = \{A_j, D_1, D_2, \dots, D_j\}$, j is the decomposition level. For CWT, $\{S_{vi}^{cwt} = s_{vi1}, s_{vi2}, \dots, s_{vij}, i = 1, 2, \dots, n\}$, where j is the highest possible scale and s_{vi1} is a vector of the coefficients of first scale of the i^{th} variable.
5. The sub time series/scales for all variables are gathered $TS^{dwt} = [S_{v1}^{dwt}\ S_{v2}^{dwt}\ \dots\ S_{vn}^{dwt}])$, $TS^{cwt} = [S_{v1}^{cwt}\ S_{v2}^{cwt}\ \dots\ S_{vn}^{cwt}])$ for DWT and CWT respectively and divided into two parts 75% for calibration TS_{calib}^{dwt} (TS_{calib}^{cwt}) and 25% for testing TS_{test}^{dwt} (TS_{test}^{cwt}) for DWT(CWT). The target $Q_{t+1,CALIB}$ also divided into $Q_{t+1,calib}$ and $Q_{t+1,test}$ with the same percentages.
6. For the second schema, the TS_{calib}^{dwt} (TS_{calib}^{cwt}) are imposed as inputs into ELM and XGB for training and $Q_{t+1,calib}$ as output, and then the models with the best parameters are chosen and tested using TS_{test}^{dwt} (TS_{test}^{cwt}) and $Q_{t+1,test}$.
7. In the third schema, the proposed model is applied in which the feature importance obtained from XGB implemented in step 6 is utilized. The training of ELM uses the matrix consists of only the important scale of the calibration subset $TS_{calib}^{cwt,best} = [S_{vij}^{cwt,1}\ S_{vij}^{cwt,2}\ \dots\ S_{vij}^{cwt,m}]$ and the testing is implemented using $TS_{test}^{cwt,best}$.
8. In both schemas the first value in the V_{TEST} is appended into V_{CALIB} to obtain a series V_{calib} have a length of $1 - k + 1$.
9. The V_{calib} is decomposed using the same mother wavelet, decomposition level, and highest possible scale. The last value of the decomposed series $TS_{CALIB,k+1}^{dwt}$ ($TS_{CALIB,k+1}^{cwt}$) is imposed in the models obtained from 6 and 7 to predict the $Q_{t+1,CALIB,k+1}$value and save the predicted value.
10. Append the first value from V_{TEST} into V_{CALIB} and from $Q_{t+1,TEST}$ into $Q_{t+1,CALIB}$ to obtain series with lengths $1, 2, \dots, k + 1$.

11. The steps 1 – 9 is repeated till all the V_{TEST} is appended into V_{CALIB}.
12. The evaluation criteria are determined.

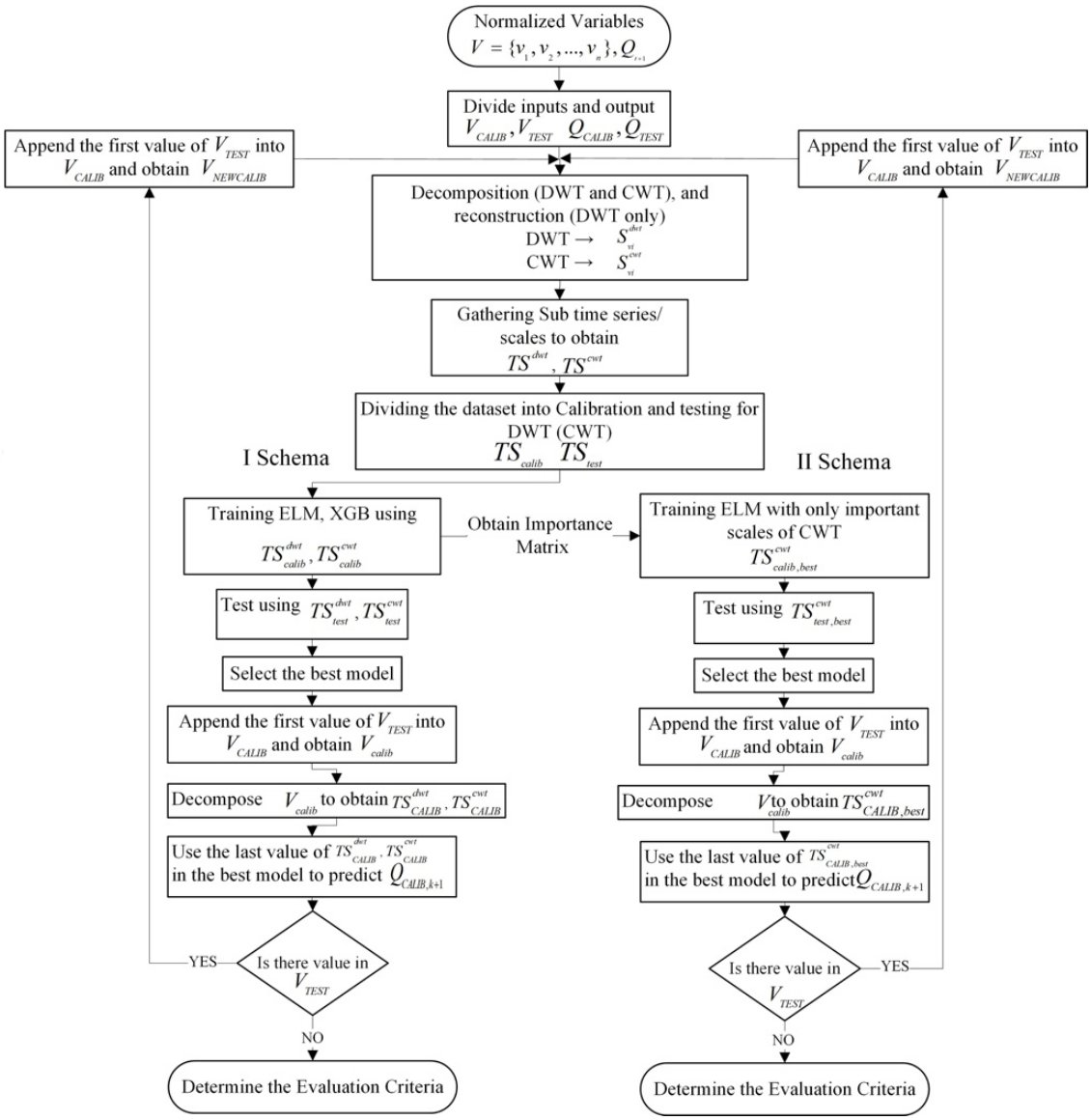


Figure 5.5. *The framework of the forecast experiment*

5.1.8. Evaluation Criteria

Performance of the models evaluated by the use of two measurements known as root mean square error (RMSE) and coefficient of determination (R^2).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(\widehat{Q}_i - Q_i\right)^2}{N}}, \tag{5.29}$$

where N is the number of the samples, Q is the observed values and $\widehat{\mathrm{Q}}$ is the estimated values from the model. RMSE has the same unit of (Q) and should be as low as possible.

The coefficient of efficiency (CE) is also known Nash-Sutcliffe coefficient (Nash et al. 1970). CE) is calculated as follows:

$$\mathrm{CE} = 1 - \frac{\sum_{i=1}^{N}\left(\mathrm{Q}_i - \hat{Q}_i\right)^2}{\sum_{i=1}^{N}(Q_i - \bar{Q})^2}, \tag{5.30}$$

where $\overline{\mathrm{Q}}$ is the mean of the observed sample points. The values of CE varies between -∞ and one. The value of one denotes a perfect model while the value of zero denotes that the model is not giving different value from the mean of the observed data, and negative values indicate that the mean of the observed data is a better estimator than the model (Dawson et al. 2007).

5.1.9. Results and Discussion

The stand-alone experiment consists of imposing the normalized lags of the different combination into the model as inputs and the one month ahead downstream flow Q_{t+1} as output after the division of the data set into calibration and test subsets. In this type of modelling, the future information is not included and considered as forecasting as no decomposition applied on the original time series. The results of the two methods ELM and XGB are listed in Table 5.2 and Table 5.3. The ELM results indicate that including any of the lags of the variables beside the lags of the downstream flow (i.e. DS combination) increases the performance of the model in general. Although including the lags of all variable (i.e. RTPETUSDS combination) increasing the performance in comparison to the DS combination, the model contains the PET beside DS (i.e. PETDS combination) has the highest performance combination. Using XGB also shows that including other variables with the DS increases the performance of the models but the highest performance is for the combination that consists of all the variables (i.e. RTPETUSDS combination) and not the PETDS combination as in ELM. XGB outperforms the ELM mostly in the calibration subset indicating overtraining and that is normal in tree-based models which need careful parameter tuning.

Table 5.2. *The CE and RMSE of the stand-alone schema of ELM method*

	Calibration		Testing	
Comb	CE	RMSE (m3/s)	CE	RMSE (m3/s)
DS	0.539	37.135	0.504	23.195
USDS	0.648	32.427	0.603	20.740
RDS	0.626	33.436	0.579	21.365
TDS	0.710	29.429	0.662	19.131
PETDS	0.789	25.141	0.680	18.635
RTPETUSDS	0.751	27.282	0.676	18.739

Table 5.3. *The CE and RMSE of the stand-alone schema of XGB method*

	Calibration		Testing	
Comb	CE	RMSE (m3/s)	CE	RMSE (m3/s)
DS	0.661	31.844	0.369	26.152
USDS	0.768	26.309	0.535	22.454
RDS	0.711	29.393	0.414	25.198
TDS	0.772	26.082	0.633	19.946
PETDS	0.831	22.493	0.594	20.986
RTPETUSDS	0.861	20.394	0.630	20.021

In both hindcast and forecast experiments two schemas are implemented; I schema in which the data decomposed using DWT (CWT). In this schema CWT-ELM is not applied as the number of inputs is huge and that deteriorates the performance dramatically. The II schema is proposed in this study in which the only the important scales of the CWT obtained by XGB are included in the model and the hybrid model produced is CWT-XGB-ELM. The hindcast experiment is implemented in such a way that the inputs are decomposed and divided into calibration and testing subsets to be then imposed into ELM or XGB. The results of the four hybrid models of this experiment is shown in Table 5.4. The use of DWT as preprocessing increased the performance of the models dramatically in comparison with the stand-alone models especially with ELM who performed better than XGB. The use of CWT-XGB hybrid method behaved better than the use of DWT with XGB in DWT-XGB and better than the stand-alone models. The proposed hybrid model CWT-XGB-ELM has the highest performance in all models with highest accuracy in the PETDS combination 0.987 and 0.973 CE and 6.182 m^3/s and 5.413 m^3/s for calibration and test respectively.

Table 5.4. *The CE and RMSE of the hindcast experiment*

	DWT-ELM				DWT-XGB				CWT-XGB				CWT-XGB-ELM			
	Calibration		Test		Calibration		Test		Calibration		Test		Calibration		Test	
Comb	CE	RMSE (m3/s)	CE	RMSE (m3/s)	CE	RMSE (m3/s)	CE	RMSE (m3/s)	CE	RMSE (m3/s)	CE	RMSE (m3/s)	CE	RMSE (m3/s)	CE	RMSE (m3/s)
DS	0.965	10.176	0.944	7.783	0.931	14.410	0.841	13.113	0.927	14.753	0.831	13.548	0.986	6.480	0.978	4.910
USDS	0.962	10.704	0.948	7.444	0.929	14.594	0.818	14.067	0.929	14.521	0.825	13.781	0.974	8.820	0.955	7.017
RDS	0.963	10.489	0.925	8.965	0.931	14.370	0.797	14.842	0.936	13.867	0.862	12.234	0.981	7.540	0.974	5.330
TDS	0.967	9.969	0.946	7.624	0.931	14.370	0.797	14.842	0.939	13.552	0.875	11.652	0.981	7.483	0.966	6.078
PETDS	0.959	11.101	0.942	7.859	0.929	14.520	0.797	14.841	0.936	13.824	0.881	11.352	0.987	6.182	0.973	5.413
RTPETUSDS	0.970	9.552	0.943	7.795	0.933	14.121	0.830	13.569	0.947	12.560	0.827	13.694	0.963	10.544	0.950	7.337

The reason of this dramatic increase in the hybrid models is that the decomposition and reconstruction of the times series imposes some amount of the future information into the model. This information is included in the decomposed time series due to the convolution process of the original time series with the filters.

In the forecast experiment, the modelling is done in such a way that to forecast the coming month all the inputs till the previous month is used for calibrating a model after the decomposition and these inputs till the current month is decomposed and imposed to the calibrated model for forecasting as in reality the data only up to the current month is existing. This procedure depends greatly on the ends of the decomposed time series which is really affected by the border effect. The results of this experiment are shown in Table 5.5. According to these results, the performance of all the combinations deteriorated in comparison to the stand-alone experiment. The DWT-ELM has shown the highest performance among all others in this forecast experiment. The proposed model CWT-XGB-ELM has the worst performance while in the hindcast experiment it showed a dramatic increase in the performance. This deterioration in the proposed model caused by the high distortion at the ends of the CWT-based transformed coefficients resulted from the boundary effect.

Table 5.5. *The CE and RMSE of the forecast experiment*

Comb	DWT-ELM		DWT-XGB		CWT-XGB		CWT-XGB-ELM	
	CE	RMSE (m3/s)	CE	RMSE (m3/s)	CE	RMSE (m3/s)	CE	RMSE (m3/s)
DS	0.40	26.11	0.36	27.05	-0.09	35.21	-0.24	37.62
USDS	0.31	27.93	0.34	27.50	-0.15	36.14	-1.24	50.57
RDS	0.41	25.91	-0.35	39.21	-0.06	34.70	-0.33	38.91
TDS	0.46	24.82	0.52	23.28	-0.15	36.18	-2.39	62.17
PETDS	0.49	24.21	0.47	24.49	-0.20	37.05	-2.11	59.57
RTPETUSDS	0.44	25.27	-0.11	35.49	-0.92	46.80	-0.92	46.80

The fitted values of the best performed models of the three experiments are plotted versus the observed values in Figure 5.6. The figure shows almost perfect agreement of the points with the best fit line for the hindcast experiment but that has been proved as incorrect application of the WT-based hybrid models. For forecasting, only 56 points are shown as only these values were forecasted while precedent used for calibration. The agreement of these points with the best fit line is much worse than the stand-alone

experiment. In brief, the use of WT in both DWT and CWT deteriorates the models performance and stand-alone models outperform hybrid models.

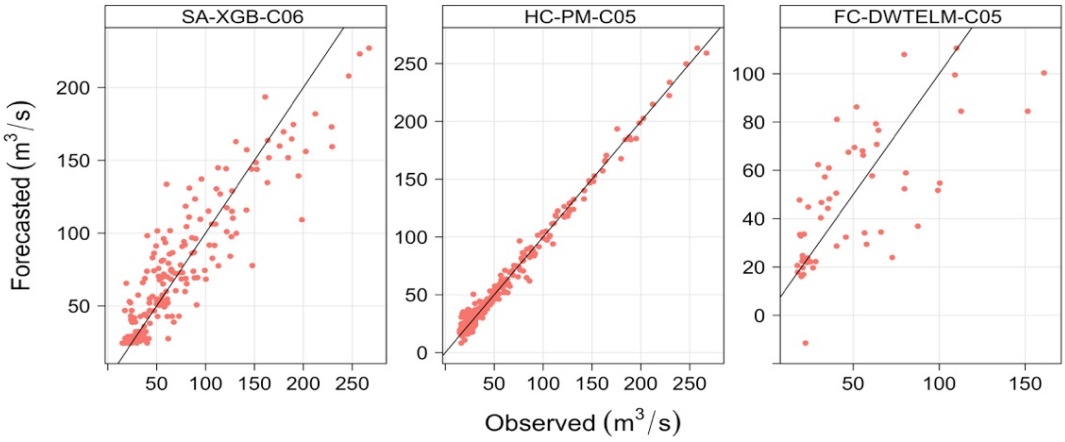


Figure 5.6. *The forecasted versus the observed downstream flow for the best model in each experiment. SA: Stand-alone, HC: Hindcasting, PM: Proposed Model, C05: combination number 5, FC: Forecasting*

The most important issue in WT-based models is the issue of the border effect on the finite time series which is the case of the hydrological time series and not only streamflow. Therefore, the decomposition of the time series before the division is an inaccurate practice of forecasting as some of the future information is included in the modeling due to the convolution. In a real example, if the streamflow of the current month to be forecasted no future information is existing as the data is recorded up to that month. Therefore, we have to decompose the time series and include all the previous records in this decomposition in order to then impose them into any DDM for forecasting. In this case, the ends are treated differently according to the method of the border effect used. These distorted ends are not included in the hindcast experiment. In fact, they are included only at the end of the whole original time series which is not the real case. Therefore, a real forecast experiment is conducted here.

For showing the effect of the border effect which leads to the dramatic decrease in the performance of the real forecasting models, a difference analysis is conducted. For DWT, level 2 is used which produces 3 reconstructed subseries: 1 approximation A and two details D1, D1. The sub time series of the 200 months and the 265 months is plotted in Figure 5.7. The difference between 200 sub time series and the first 200 steps of the 256 sub time series of the downstream flow shown in Figure 5.8 indicates that there are

differences at the end of the series due to the border effect. These differences varies based on the border effect method used which arises in dealing finite time series (Zhang et al. 2015). As the best wavelet function found is db7 for this study, the filter length is 14. According to that, the number of the different values in D1 is 12 and in D2 and A2 is 36. The calculation of the number of different values can be found in detail in (Du et al. 2017).

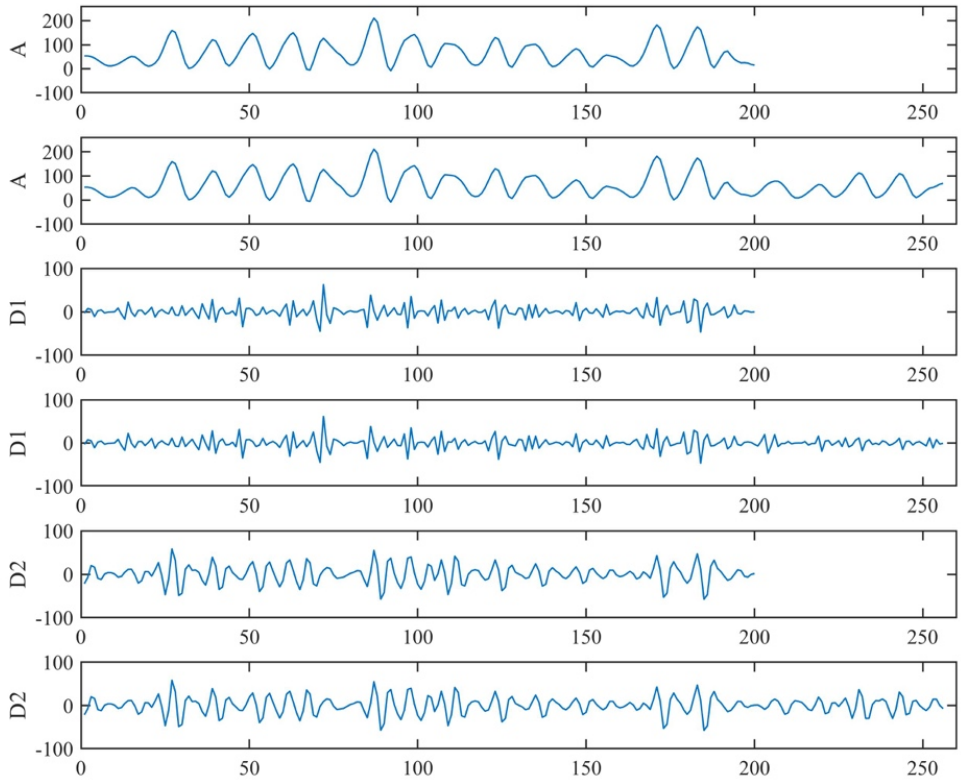


Figure 5.7. *The reconstructed sub time series of the downstream flow using DWT with level 2*

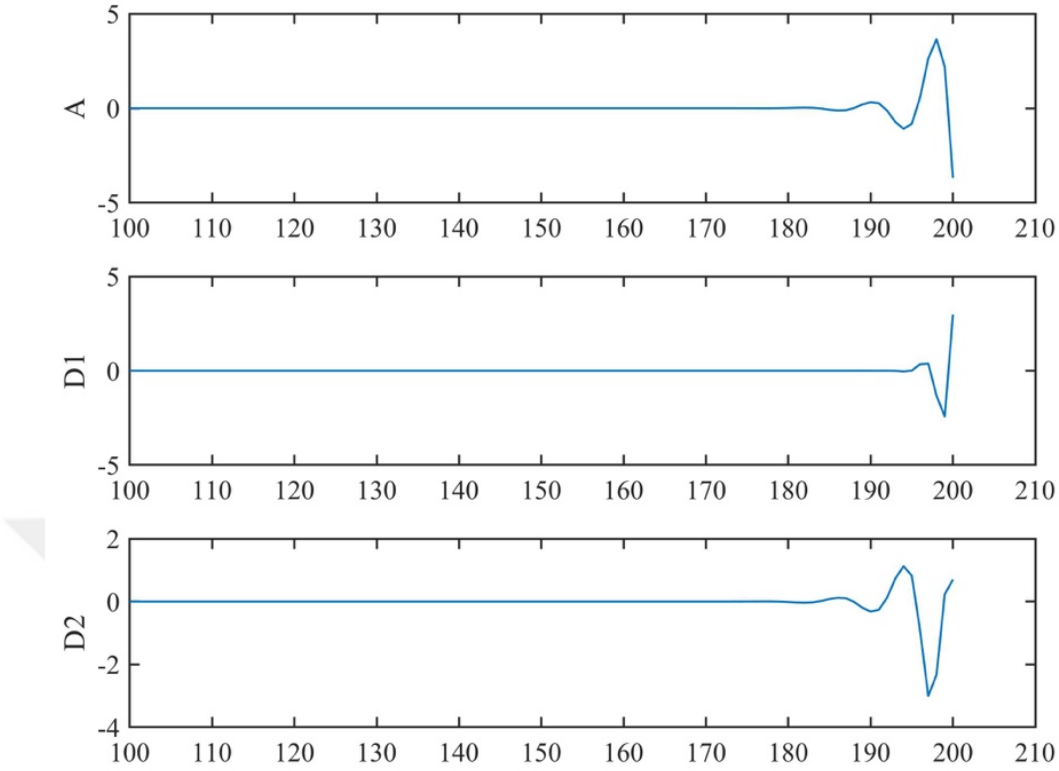


Figure 5.8. *The difference between the first 200 steps of sub time series of the downstream flow using DWT with level 2 for 200 and 256 months*

For CWT, the time series were transformed using 128 scales for every variable. Using the proposed method, the importance matrix of the most important scales involved in the modeling process is obtained using XGB. According to the results of this importance matrix, the 5th scales of the fist lag of downstream flow Q_t was found as the most important scale in all models combinations (Table 5.6). Therefore, the comparison of the 200 and 265 time series is conducted here on this scale only. Haar (i.e. db1) which has a filter length of 2 found as the best function of the CWT analysis. The transformation of the original time series (scale 5 is shown in Figure 5.9) leads to same issue of the different values at the end of the decomposed finite time series. As the filter length is only 2, there are only three different values between the 200 and 256 decomposed coefficients and the differences have very high magnitudes . The differences in this case is much higher the differences in the case of DWT, and that leads to a dramatic decrease in the performance of the CWT models in comparison with DWT models.

Table 5.6. *The importance matrix of the CWT scales obtained by XGB. The value in the brackets represents the scale*

Im. Rank	DS	USDS	RDS	TDS	PETDS	RTPETUSDS
1	$DS_{t-1}(5)$	$DS_{t-1}(5)$	$DS_{t-1}(5)$	$DS_{t-1}(5)$	$DS_{t-1}(5)$	$US_{t-1}(4)$
2	$DS_{t-1}(8)$	$US_{t-1}(4)$	$DS_{t-2}(3)$	$DS_{t-1}(8)$	$Ds_{t-1}(7)$	$DS_{t-1}(5)$
3	$DS_{t-2}(3)$	$US_{t-2}(10)$	$DS_{t-1}(8)$	$Ds_{t-1}(6)$	$DS_{t-1}(8)$	$DS_{t-2}(3)$
4	DS_{t-1} (7)	DS_{t-2} (1)	$DS_{t-1}(6)$	$DS_{t-1}(7)$	$DS_{t-2}(1)$	$US_{t-2}(6)$
5	DS_{t-1} (6)	US_{t-2} (7)	DS_{t-2} (1)	DS_{t-2} (1)	$DS_{t-2}(3)$	$DS_{t-2}(1)$
6	$DS_{t-2}(1)$	DS_{t-2} (3)	DS_{t-2} (11)	DS_{t-2} (4)	$DS_{t-1}(6)$	$US_{t-2}(7)$
7	DS_{t-2} (4)	US_{t-2} (8)		DS_{t-2} (3)	$PET_{t-1}(11)$	$US_{t-2}(12)$
8	DS_{t-2} (9)	DS_{t-1} (8)				$US_{t-2}(10)$
9		US_{t-2} (5)				$US_{t-2}(5)$
10		DS_{t-1} (6)				DS_{t-1} (2)
11		Us_{t-1} (1)				

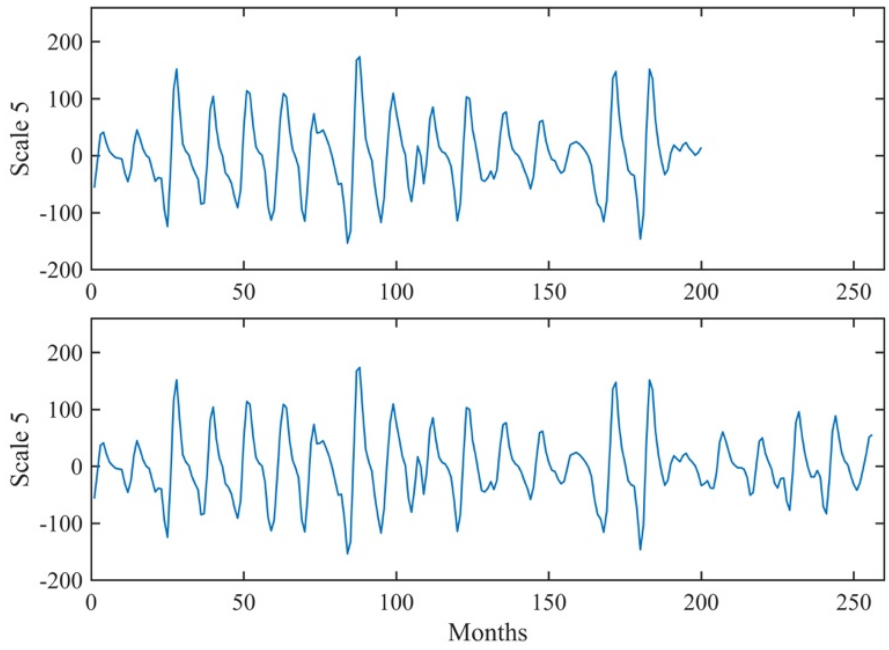


Figure 5.9. *The coefficients of 5th scale of the downstream flow using CWT*

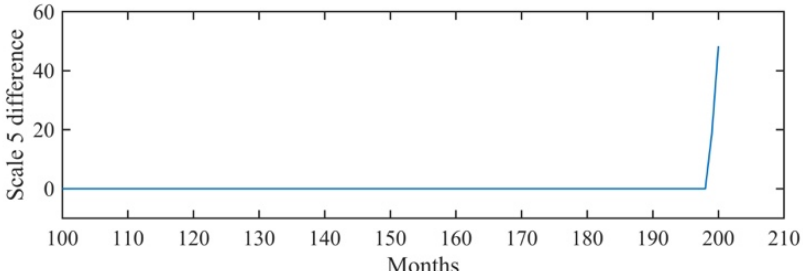


Figure 5.10. *The difference between the first 200 months of coefficients of 5th scale of the downstream flow using CWT for 200 and 256 months*

5.1.10. Conclusion

In this study, three experiments are applied stand-alone in which no WT is applied, hindcast experiment where DWT (CWT) is applied to the time series to be then divided into calibration and testing before entering them into ELM or XGB, forecast experiment in which the data is divided and then after the model is trained with the decomposed first subset a value is added from the second subset to the first subset to be decomposed and entered to the ELM or XGB to implement real forecast. In both hindcast and forecast experiment, a new proposed hybrid model is proposed in which the importance matrix that show the importance of the features (i.e. scales in this study) is utilized to choose only the important scales to be imposed into the ELM. According to the results obtained, several points are remarked:

- The use of WT-based hybrid models increases the performance of the models in hindcast experiment due to including some amount of the future information resulted as consequence of time series convolution.
- WT-based hybrid models in the forecast experiment deteriorates the performance of the models and the stand-alone models performs better and that mostly due to the border effect which distorts the ends of the decomposed time series.
- In hindcast experiment, the distortion at the ends of the decomposed finite time series caused by the border effect is involved only at the end of original time series while in the forecast experiment which is considered as real forecasting is involved in forecasting every value in the testing subset.
- The proposed hybrid model that using XGB as selection tool in addition to being a modelling approach itself has dramatically improved the performance of the hindcast experiment but not the forecast experiment. So, it has the potential to be applied on other applications of reducing the number of features imposed into the modelling approach.
- The use of metrological and hydrological variables such as temperature, and potential evapotranspiration beside the lagged streamflow, which is essential for auto regression, improves the performance of the models.

CONCLUSION

The main objective of this study is to propose a new hybrid models that improve the streamflow forecasting and especially the monthly streamflow which is less predictable than the daily. Due to the unavailability of the required data several steps had to be done before the main objective. One of the most important steps is the spatial interpolation which is done due to that one of the used basins in the study has no station located inside the basin therefore, using several stations around the around it the rainfall and temperature can be interpolated. The evaluation of eight interpolation methods is conducted and IDW with a power of three is identified as the best modeling of rainfall and temperature which then used for interpolating value in the missing station.

The evapotranspiration data collected from the official department has missing values more than the recorded values. Therefore, the use of this data is impossible. A global data set extends from 1901-2015 namely CRUTS is utilized for obtaining the potential evapotranspiration to be used in the models. But evaluating these data in a local scale is an essential step in this case. Therefore, the evaluation is implemented, and the results indicates that these dataset is reliable and has a very high agreement with the observed data.

Another important issue in modelling is to explore the trend and the stationarity of the data. This exploration is done in two scales: national scale and local scale. The national scale is conducted for the whole Turkey for each province using CRUTS data covering the period 1901-2014. The results indicate that in general rainfall has an insignificant increasing trend while temperature is increasing significantly over the whole Turkey. The local scale study conducted for the period 1973 -2000 which the period of the available streamflow data. The rainfall and streamflow data were found as homogeneous using several homogeneity tests and most of the stations has either increasing or decreasing trend but mostly not statistically significant.

Having the preliminarily steps accomplished the main objective was started with daily streamflow modeling using several DDMs including AR, ANN, ANFIS, SVM. In this study, only the lagged streamflow used. As three basins used in this study, the effect of the physical characteristics of these basins on the model's performance was also investigated. The ANN indicates the highest performance among the four methods. However, in some cases, the ANN and SVM are indicated to obtain close results. Simultaneously, the SVM performed better than AR and ANFIS when the latter indicates

the least performance among all stations. The ANN outperformed all methods in forecasting peak flood values with the close result to SVM, which is far from the least performance obtained by AR, which is less than ANFIS. The physical characteristics, such as small area, high variation in elevation, and high slope, significantly deteriorate the model performance, although the methods are data-driven models. These physical characteristics can support the evidence of the deterioration or improvement in the forecasting accuracy regardless of the applied data-driven model.

The one day ahead streamflow is giving high accuracies in comparison to the monthly and even seven days ahead. Therefore, a hybrid model is found important to improve the forecasting ability. As most of the studies uses DWT as preprocessing, in this study a combination of CWT and DWT is proposed for improving the seven day ahead. Only the scale with the highest correlation with the streamflow is used from CWT while all the sub time series of DWT is used. CWT improves only the models that contains only the rainfall and temperature in which the single models have very low accuracy, but not those which includes the delayed streamflow in which the accuracy higher while in the other models in which the streamflow is included the use of DWT is greatly giving higher accuracy. The combination of the CWT which applied on the rainfall and temperature only and DWT which applied on the streamflow only namely Discrete-Continuous-Wavelet transformation (DCWT) improved several models in comparison with DWT; especially those models include rainfall and used with ANN, but, the models imposed to the SVM had a deteriorated performance and that due to the unexpected behavior of SVM with the use of CWT in general. ANN again outperformed ANFIS and SVM and ANFIS has higher accuracy than SVM.

Monthly streamflow is less predictable than daily which brings the need of another approach to improve the forecast. In the forecasting of seven days ahead, the best-scale layout is used and it improved the performance due to the high linear correlation which is higher than the original variable but using linear correlation is not sufficient in the existence of non-linearity. Therefore, a new hybrid model is proposed that Multi-Gene Genetic Programing (MGGP) which has the ability to produce a formula for the prediction is used to obtain the important scales of CWT that appears in the that formula. Then, only the chosen scales are imposed into ANN for one month ahead forecasting. A hybrid model CWT-MGGP-ANN is developed. Therefore, the proposed model WMGGP improved the performance dramatically. The most important point is that the fluctuation

in the streamflow represented by peak values are well forecasted due to the fact that the chosen scales include the sufficient seasonality and irregularity information. According to this study, the restriction of using the CWT as pre-processing tool due to the high number of scales and redundant information for predicting the monthly streamflow is avoidable.

Another hybrid model was proposed in which XGB. XGB being a tree-based model has the ability of producing feature importance matrix that shows the importance of the scales in order starting with the most important and ending with the least important scales. Therefore, XGB is utilized as a selection tool of the most important scales of the preprocessed time series with CWT and these scales are imposed into ELM as modelling approach. In the literature including the proposed model, the application of WT in both DWT and CWT is considered as hindcast modelling and not forecasting. In other words, the WT is applied incorrectly in the applications of forecasting. Based on that, two experiments were conducted: Hindcast and Forecast experiment. In hindcast experiment the time series is decomposed, divided into calibration and test, and then imposed into the model (XGB or ELM). In the forecast experiment, the data till the current date is decomposed and used as calibration, and for forecasting the next month value is added to the time series and decomposed to use it for the prediction. The hindcast experiment is sending some amount of the future information into the models which is an incorrect application as the real forecast is done with not future information. According to the obtained results, the conclusion is derived as follow:

- The use of WT-based hybrid models increases the performance of the models in hindcast experiment due to including some amount of the future information resulted as consequence of time series convolution.
- WT-based hybrid models in the forecast experiment deteriorates the performance of the models and the stand-alone models performs better and that mostly due to the border effect which distorts the ends of the decomposed time series.
- In hindcast experiment, the distortion at the ends of the decomposed finite time series caused by the border effect is involved only at the end of original time series while in the forecast experiment which is considered as real forecasting is involved in forecasting every value in the testing subset.
- The proposed hybrid model that using XGB as selection tool in addition to being a modelling approach itself has dramatically improved the performance of the

hindcast experiment but not the forecast experiment. So, it has the potential to be applied on other applications of reducing the number of features imposed into the modelling approach.

- The use of metrological and hydrological variables such as temperature, and potential evapotranspiration beside the lagged streamflow, which is essential for auto regression, improves the performance of the models.

REFERENCES

Abdollahi, S., J. Raeisi, M. Khalilianpour, F. Ahmadi and O. Kisi (2017). "Daily Mean Streamflow Prediction in Perennial and Non-Perennial Rivers Using Four Data Driven Techniques." Water Resources Management.

Abdollahi, S., J. Raeisi, M. Khalilianpour, F. Ahmadi and O. Kisi (2017). "Daily Mean Streamflow Prediction in Perennial and Non-Perennial Rivers Using Four Data Driven Techniques." Water Resources Management 31(15): 4855-4874.

Abdullah, S. S., M. A. Malek, N. S. Abdullah, O. Kisi and K. S. Yap (2015). "Extreme learning machines: a new approach for prediction of reference evapotranspiration." Journal of Hydrology 527: 184-195.

Adamowski, J., H. F. Chan, S. O. Prasher and V. N. Sharda (2012). "Comparison of multivariate adaptive regression splines with coupled wavelet transform artificial neural networks for runoff forecasting in Himalayan micro-watersheds with limited data." Journal of Hydroinformatics 14(3): 731-744.

Adamowski, J. and S. O. Prasher (2012). "Comparison of machine learning methods for runoff forecasting in mountainous watersheds with limited data." Journal of Water and Land Development 17(1): 89-97.

Adamowski, J. and K. Sun (2010). "Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds." Journal of Hydrology 390(1): 85-91.

Adamowski, J. F. (2008). "Development of a short-term river flood forecasting method for snowmelt driven floods based on wavelet and cross-wavelet analysis." Journal of Hydrology 353(3): 247-266.

Adamowski, J. F. (2008). "River flow forecasting using wavelet and cross-wavelet transform models." Hydrological processes 22(25): 4877-4891.

Addison, P. S. (2002). The illustrated wavelet transform handbook, CRC Press.

Aggarwal, S. K., A. Goel and V. P. Singh (2012). "Stage and discharge forecasting by SVM and ANN techniques." Water resources management 26(13): 3705-3724.

Akinsanola, A. A. and K. O. Ogunjobi (2015). "Recent homogeneity analysis and long-term spatio-temporal rainfall trends in Nigeria." Theoretical and Applied Climatology: 1-15.

Akrami, S. A., V. Nourani and S. J. S. Hakim (2014). "Development of nonlinear model based on wavelet-ANFIS for rainfall forecasting at Klang Gates Dam." Water resources management 28(10): 2999-3018.

Alexandersson, H. (1986). "A homogeneity test applied to precipitation data." Journal of Climatology 6(6): 661-675.

Alizadeh, M. J., M. R. Kavianpour, O. Kisi and V. Nourani (2017). "A new approach for simulating and forecasting the rainfall-runoff process within the next two months." Journal of Hydrology 548(Supplement C): 588-597.

Alizadeh, M. J., V. Nourani, M. Mousavimehr and M. R. Kavianpour (2017). "Wavelet-IANN model for predicting flow discharge up to several days and months ahead." Journal of Hydroinformatics.

Anctil, F. and D. G. Tape (2004). "An exploration of artificial neural network rainfall-runoff forecasting combined with wavelet decomposition." Journal of Environmental Engineering and Science 3(S1): S121-S128.

Apaydin, H., F. K. Sonmez and Y. E. Yildirim (2004). "Spatial interpolation techniques for climate data in the GAP region in Turkey." Climate Research 28(1): 31-40.

Araghinejad, S., M. Azmi and M. Kholghi (2011). "Application of artificial neural network ensembles in probabilistic hydrological forecasting." Journal of Hydrology 407(1): 94-104.

Ay, M. and O. Kisi (2015). "Investigation of trend analysis of monthly total precipitation by an innovative method." Theoretical and Applied Climatology 120(3): 617-629.

Aytek, A. and Ö. Kişi (2008). "A genetic programming approach to suspended sediment modelling." Journal of Hydrology 351(3): 288-298.

Ayumi, V. (2016). Pose-based human action recognition with Extreme Gradient Boosting. Student Conference on Research and Development (SCOReD), Malaysia, IEEE.

Badrzadeh, H., R. Sarukkalige and A. W. Jayawardena (2013). "Impact of multi-resolution analysis of artificial intelligence models inputs on multi-step ahead river flow forecasting." Journal of Hydrology 507(Supplement C): 75-85.

Badrzadeh, H., R. Sarukkalige and A. W. Jayawardena (2018). "Intermittent stream flow forecasting and modelling with hybrid wavelet neuro-fuzzy model." Hydrology Research 49(1): 27-40.

Baratti, R., B. Cannas, A. Fanni, M. Pintus, G. M. Sechi and N. Toreno (2003). "River flow forecast for reservoir management through neural networks." Neurocomputing 55(3): 421-437.

Boer, E. P. J., K. M. d. Beurs and A. D. Hartkamp (2001). "Kriging and thin plate splines for mapping climate variables." International Journal of Applied Earth Observation and Geoinformation 3(2): 146-154.

Borges, P. D., J. Franke, Y. M. T. da Anuncicao, H. Weiss and C. Bernhofer (2016). "Comparison of spatial interpolation methods for the estimation of precipitation distribution in Distrito Federal, Brazil." Theoretical and Applied Climatology 123(1-2): 335-348.

Brekke, L. D., J. E. Kiang, J. R. Olsen, R. S. Pulwarty, D. A. Raff, D. P. Turnipseed, R. S. Webb and K. D. White (2009). Climate change and water resources management: A federal perspective, US Geological Survey Circular: 65.

Brunetti, M., M. Maugeri, T. Nanni, C. Simolo and J. Spinoni (2014). "High- resolution temperature climatology for Italy: interpolation method intercomparison." International Journal of Climatology 34(4): 1278-1296.

Buffoni, L., M. Maugeri and T. Nanni (1999). "Precipitation in Italy from 1833 to 1996." Theoretical and Applied Climatology 63(1): 33-40.

Buishand, T. A. (1982). "Some methods for testing the homogeneity of rainfall records." Journal of Hydrology 58(1): 11-27.

Cannas, B., A. Fanni, L. See and G. Sias (2006). "Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning." Physics and Chemistry of the Earth, Parts A/B/C 31(18): 1164-1171.

Carcano, E. C., P. Bartolini, M. Muselli and L. Piroddi (2008). "Jordan recurrent neural network versus IHACRES in modelling daily streamflows." Journal of Hydrology 362(3–4): 291-307.

Chang, F.-J. and Y.-C. Chen (2001). "A counterpropagation fuzzy-neural network modeling approach to real time streamflow prediction." Journal of hydrology 245(1): 153-164.

Chattopadhyay, S. and R. D. Edwards (2016). "Long-Term Trend Analysis of Precipitation and Air Temperature for Kentucky, United States." Climate 4(1).

Chen, T. and C. Guestrin (2016). Xgboost: A scalable tree boosting system. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, California, USA, ACM.

Chiu, S. L. (1994). "Fuzzy model identification based on cluster estimation." Journal of Intelligent & fuzzy systems 2(3): 267-278.

Craven, P. and G. Wahba (1978). "Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation." Numerische Mathematik 31(4).

Cressie, N. (1990). "The Origins of Kriging." Mathematical Geology 22(3): 239-252.

Cru. (2016). "High-resolution gridded datasets (and derived products)." Retrieved 12-9-2016, from http://www.cru.uea.ac.uk/.

Daubechies, I. (1988). "Orthonormal bases of compactly supported wavelets." Communications on pure and applied mathematics 41(7): 909-996.

Daubechies, I. (1990). "The wavelet transform, time-frequency localization and signal analysis." IEEE transactions on information theory 36(5): 961-1005.

Daubechies, I. (1992). Ten lectures on wavelets, SIAM.

Dawson, C. W., R. J. Abrahart and L. M. See (2007). "HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts." Environmental Modelling & Software 22(7): 1034-1052.

De Falco, I., A. Della Cioppa and E. Tarantino (2005). "A genetic programming system for time series prediction and its application to El Nino forecast." Soft Computing: Methodologies and Applications: 151-162.

de la Casa, A. and O. Nasello (2010). "Breakpoints in annual rainfall trends in Córdoba, Argentina." Atmospheric Research 95(4): 419-427.

De Luís, M., J. Raventós, J. C. González-Hidalgo, J. R. Sánchez and J. Cortina (2000). "Spatial analysis of rainfall trends in the region of Valencia (east Spain)." International Journal of Climatology 20(12): 1451-1469.

Deo, R. C. and M. Şahin (2016). "An extreme learning machine model for the simulation of monthly mean streamflow water level in eastern Queensland." Environmental monitoring and assessment 188(2): 90.

Diez, D. M., C. D. Barr and M. C. etinkaya-Rundel (2015). OpenIntro Statistics, Duke University.

Dirks, K. N., J. E. Hay, C. D. Stow and D. Harris (1998). "High-resolution studies of rainfall on Norfolk Island Part II: Interpolation of rainfall data." Journal of Hydrology 208(3-4): 187-193.

Dorado, J., J. R. RabuñAl, A. Pazos, D. Rivero, A. Santos and J. Puertas (2003). "Prediction and modeling of the rainfall-runoff transformation of a typical urban basin using ANN and GP." Applied Artificial Intelligence 17(4): 329-343.

Du, K., Y. Zhao and J. Lei (2017). "The incorrect usage of singular spectral analysis and discrete wavelet transform in hybrid models to predict hydrological time series." Journal of Hydrology 552: 44-51.

Duhan, D. and A. Pandey (2013). "Statistical analysis of long term spatial and temporal trends of precipitation during 1901–2002 at Madhya Pradesh, India." Atmospheric Research 122: 136-149.

Easterling, D. R., B. Horton, P. D. Jones, T. C. Peterson, T. R. Karl, D. E. Parker, M. J. Salinger, V. Razuvayev, N. Plummer, P. Jamason and C. K. Folland (1997). "Maximum and minimum temperature trends for the Globe." Science 277(5324): 364.

Fahimi, F., Z. M. Yaseen and A. El-shafie (2016). "Application of soft computing based hybrid models in hydrological variables modeling: a comprehensive review." Theoretical and Applied Climatology: 1-29.

Feidas, H., C. Noulopoulou, T. Makrogiannis and E. Bora-Senta (2007). "Trend analysis of precipitation time series in Greece and their relationship with circulation using surface and satellite data: 1955–2001." Theoretical and Applied Climatology 87(1): 155-177.

Feizi, V., M. Mollashahi, M. Farajzadeh and G. Azizi (2014). "Spatial and temporal trend analysis of temperature and precipitation in Iran." ECOPERSIA 2(4): 727-742.

Firat, M. and M. Güngör (2007). "River flow estimation using adaptive neuro fuzzy inference system." Mathematics and Computers in Simulation 75(3): 87-96.

Frazier, A. G., T. W. Giambelluca, H. F. Diaz and H. L. Needham (2016). "Comparison of geostatistical approaches to spatially interpolate month-year rainfall for the Hawaiian Islands." International Journal of Climatology 36(3): 1459-1470.

Friedman, J. H. (2001). "Greedy function approximation: a gradient boosting machine." Annals of statistics: 1189-1232.

Furundzic, D. (1998). "Application example of neural networks for time series analysis:: Rainfall–runoff modeling." Signal Processing 64(3): 383-396.

Galelli, S., G. B. Humphrey, H. R. Maier, A. Castelletti, G. C. Dandy and M. S. Gibbs (2014). "An evaluation framework for input variable selection algorithms for environmental data-driven models." Environmental modelling & software 62: 33-51.

Gandomi, A. H., A. H. Alavi and C. Ryan (2015). Handbook of genetic programming applications, Springer.

Garcia, R. A., M. Cabeza, C. Rahbek and M. B. Araújo (2014). "Multiple dimensions of climate change and their implications for biodiversity." Science 344(6183).

Gemmer, M., S. Becker and T. Jiang (2004). "Observed monthly precipitation trends in China 1951–2002." Theoretical and Applied Climatology 77(1): 39-45.

Ghorbani, M. A., R. Khatibi, A. Aytek, O. Makarynskyy and J. Shiri (2010). "Sea water level forecasting using genetic programming and comparing the performance with artificial neural networks." Computers & Geosciences 36(5): 620-627.

Ghorbani, M. A., H. A. Zadeh, M. Isazadeh and O. Terzi (2016). "A comparative study of artificial neural network (MLP, RBF) and support vector machine models for river flow prediction." Environmental Earth Sciences 75(6): 476.

Goovaerts, P. (2000). "Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall." Journal of hydrology 228(1): 113-129.

Govindaraju, R. S. (2000). "Artificial neural networks in hydrology. I: Preliminary concepts." Journal of Hydrologic Engineering 5(2): 115-123.

Govindaraju, R. S. (2000). "Artificial neural networks in hydrology. II: hydrologic applications." Journal of Hydrologic Engineering 5(2): 124-137.

Grinsted, A., J. C. Moore and S. Jevrejeva (2004). "Application of the cross wavelet transform and wavelet coherence to geophysical time series." Nonlinear processes in geophysics 11(5/6): 561-566.

Gunn, S. R. (1998). "Support vector machines for classification and regression." ISIS technical report 14: 85-86.

Guven, A. (2009). "Linear genetic programming for time-series modelling of daily flow rate." Journal of earth system science 118(2): 137-146.

Hadi, S. J. and M. Tombul Conversion of CruTS 3.23 data and evaluation of precipitation and temperature variables in a local scale, EDP Sciences.

Hadi, S. J. and M. TOMBUL (2017). Conversion of CruTS 3.23 data and evaluation of precipitation and temperature variables in a local scale. Proceedings of the International Conference on Advances in Sustainable Construction Materials & Civil Engineering Systems (ASCMCES-17), Sharjah, UAE.

Hadi, S. J. and M. TOMBUL (2017). Conversion of CruTS 3.23 data and evaluation of precipitation and temperature variables in a local scale. Proceedings of the International Conference on Advances in Sustainable Construction

Materials & Civil Engineering Systems (ASCMCES-17), Sharjah, UAE.

Hájek, J., Z. Šidák and P. K. Sen (1999). Theory of Rank Tests, Academic Press.

Haktanir, T., S. Bajabaa and M. Masoud (2013). "Stochastic analyses of maximum daily rainfall series recorded at two stations across the Mediterranean Sea." Arabian Journal of Geosciences 6(10): 3943-3958.

Hancock, P. A. and M. F. Hutchinson (2006). "Spatial interpolation of large climate data sets using bivariate thin plate smoothing splines." Environmental Modelling & Software 21(12): 1684-1694.

Harris, I., P. D. Jones, T. J. Osborn and D. H. Lister (2014). "Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset." International Journal of Climatology 34(3): 623-642.

He, Z., X. Wen, H. Liu and J. Du (2014). "A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region." Journal of Hydrology 509: 379-386.

Hellden, U. and C. Tottrup (2008). "Regional desertification: A global synthesis." Global and Planetary Change 64(3-4): 169-176.

Helsel, D. R. and R. M. Hirsch (2002). Statistical Methods in Water Resources. Techniques of Water Resources Investigations, U.S. Geological Survey. Book 4: 522.

Hoff, H., M. Falkenmark, D. Gerten, L. Gordon, L. Karlberg and J. Rockstrom (2010). "Greening the global water system." Journal of Hydrology 384(3-4): 177-186.

Hofstra, N., M. Haylock, M. New, P. Jones and C. Frei (2008). "Comparison of six methods for the interpolation of daily, European climate data." Journal of Geophysical Research-Atmospheres 113(D21).

Hollander, M., D. A. Wolfe and E. Chicken (2014). Nonparametric Statistical Methods, Wiley Online.

Hornik, K., M. Stinchcombe and H. White (1989). "Multilayer feedforward networks are universal approximators." Neural Networks 2(5): 359-366.

Hsu, K. l., H. V. Gupta and S. Sorooshian (1995). "Artificial neural network modeling of the rainfall-runoff process." Water resources research 31(10): 2517-2530.

Huang, G.-B., Q.-Y. Zhu and C.-K. Siew (2006). "Extreme learning machine: theory and applications." Neurocomputing 70(1-3): 489-501.

Huisman, O. and R. A. d. By (1999). Principles of Geographical Information Systems: An introductory Textbook. Enschede, Netherlands, The international Institute for Geo-Information Science and Earth Observation (ITC).

Hutchinson, M. F. (1998). "Interpolation of rainfall data with thin plate smoothing splines. Part I: Two dimensional smoothing of data with short range correlation." Journal of Geographic Information and Decision Analysis 2(2): 139-151.

Ihaka, R. (2005). "Time Series Analysis." Epidemiology and.

IPCC (2001). Climate change 2001: the scientific basis. contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press.

IPCC (2013). Summary for Policymakers. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. T. F. Stocker, D. Qin, G.-K. Plattner et al. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press: 1–30.

Isaaks, E. H. and R. M. Srivastava (1989). Applied geostatistics. New York ; Oxford, Oxford University Press.

Isik, S., L. Kalin, J. E. Schoonover, P. Srivastava and B. G. Lockaby (2013). "Modeling effects of changing land use/cover on daily streamflow: an artificial neural network and curve number based hybrid approach." Journal of Hydrology 485: 103-112.

Jang, J.-S. R., C.-T. Sun and E. Mizutani (1997). "Neuro-fuzzy and soft computing, a computational approach to learning and machine intelligence."

Jang, J. S. (1993). "ANFIS: adaptive-network-based fuzzy inference system." IEEE transactions on systems, man, and cybernetics 23(3): 665-685.

Jang, J. S., C.-T. Sun and E. Mizutani (1997). "Neuro-fuzzy and soft computing, a computational approach to learning and machine intelligence."

Johnston, K., J. M. Ver Hoef, K. Krivoruchko and N. Lucas (2001). Using ArcGIS geostatistical analyst, Esri Redlands.

Jones, P. and I. Harris (2013). "CRU TS3. 21: Climatic Research Unit (CRU) Time-Series (TS) version 3.21 of high resolution gridded data of month-by-month variation in climate (Jan. 1901–Dec. 2012)." NCAS British Atmospheric Data Centre.

Jones, P. D., C. Harpham, I. Harris, C. M. Goodess, A. Burton, A. Centella-Artola, M. A. Taylor, A. Bezanilla-Morlot, J. D. Campbell, T. S. Stephenson, O. Joslyn, K. Nicholls and T. Baur (2016). "Long-term trends in precipitation and temperature across the Caribbean." International Journal of Climatology 36(9): 3314-3333.

Jones, P. D. and I. Harris. (2008). "Climatic Research Unit (CRU) time-series datasets of variations in climate with variations in other phenomena." Retrieved 10/08/2016, from http://catalogue.ceda.ac.uk/uuid/3f8944800cc48e1cbc29a5ee12d8542d.

Juan Francisco Sanchez-Moreno, Chris M. Mannaerts and V. Jetten (2013). "Influence of topography on rainfall variability in Santiago Island, Cape Verde." International Journal of Climatology 34(4): 1081–1097.

Jury, M. R., D. B. Enfield and J. L. Mélice (2002). "Tropical monsoons around Africa: Stability of El Niño–Southern Oscillation associations and links with continental climate." Journal of Geophysical Research: Oceans 107(C10).

Kadioğlu, M. (1997). "Trends in surface air temperature data over Turkey." International Journal of Climatology 17(5): 511-520.

Kampata, J. M., B. P. Parida and D. B. Moalafhi (2008). "Trend analysis of rainfall in the headstreams of the Zambezi River Basin in Zambia." Physics and Chemistry of the Earth, Parts A/B/C 33(8–13): 621-625.

Karpouzos, D. K., S. Kavalieratou and C. Babajimopoulos (2010). "Trend analysis of precipitation data in Pieria Region (Greece)." European Water 30: 31-40.

Karthikeyan, L. and D. N. Kumar (2013). "Predictability of nonstationary time series using wavelet and EMD based ARMA models." Journal of hydrology 502: 103-119.

Keblouti, M., L. Ouerdachi and H. Boutaghane (2012). "Spatial Interpolation of Annual Precipitation in Annaba-Algeria - Comparison and Evaluation of Methods."

Terragreen 2012: Clean Energy Solutions for Sustainable Environment (Cesse) 18: 468-475.

Keggenhoff, I., M. Elizbarashvili, A. Amiri-Farahani and L. King (2014). "Trends in daily temperature and precipitation extremes over Georgia, 1971–2010." Weather and Climate Extremes 4: 75-85.

Kendall, M. G. (1938). "A new measure of rank correlation." Biometrika 30(1/2): 81-93.

Kendall, M. G. (1975). Rank Correlation Methods. Griffin, London.

Khadangi, E., H. R. Madvar and M. M. Ebadzadeh (2009). Comparison of ANFIS and RBF models in daily stream flow forecasting, IEEE.

Kim, T.-W. and J. B. Valdés (2003). "Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks." Journal of Hydrologic Engineering 8(6): 319-328.

Kişi, Ö. (2007). "Streamflow forecasting using different artificial neural network algorithms." Journal of Hydrologic Engineering 12(5): 532-539.

Kişi, Ö. (2008). "River flow forecasting and estimation using different artificial neural network techniques." Hydrology Research 39(1): 27.

Kişi, Ö. (2008). "Stream flow forecasting using neuro-wavelet technique." Hydrological Processes 22(20): 4142-4152.

Kişi, Ö. (2009). "Neural Networks and Wavelet Conjunction Model for Intermittent Streamflow Forecasting." Journal of Hydrologic Engineering 14(8): 773-782.

Kişi, Ö. (2015). "Streamflow Forecasting and Estimation Using Least Square Support Vector Regression and Adaptive Neuro-Fuzzy Embedded Fuzzy c-means Clustering." Water Resources Management 29(14): 5109-5127.

Kisi, O. and M. Cimen (2011). "A wavelet-support vector machine conjunction model for monthly streamflow forecasting." Journal of Hydrology 399(1): 132-140.

Kisi, O. and A. Guven (2010). "Evapotranspiration modeling using linear genetic programming technique." Journal of Irrigation and Drainage Engineering 136(10): 715-723.

Kişi, Ö. and H. Kerem Cigizoglu (2007). "Comparison of different ANN techniques in river flow prediction." Civil Engineering and Environmental Systems 24(3): 211-231.

Kisi, O., A. M. Nia, M. G. Gosheh, M. R. J. Tajabadi and A. Ahmadi (2012). "Intermittent streamflow forecasting by using several data driven techniques." Water resources management 26(2): 457-474.

Kişi, Ö., A. M. Nia, M. G. Gosheh, M. R. J. Tajabadi and A. Ahmadi (2012). "Intermittent streamflow forecasting by using several data driven techniques." Water resources management 26(2): 457-474.

Kisi, O., E. Okan and M. Cihan (2017). "Trend analysis of long-term temperatures in Tbilisi, Georgia." Journal of Technical Science and Technologies 5(2): 59-63.

Komasi, M. and S. Sharghi (2016). "Hybrid wavelet-support vector machine approach for modelling rainfall–runoff process." Water Science and Technology 73(8): 1937-1953.

Kottek, M., J. Grieser, C. Beck, B. Rudolf and F. Rubel (2006). "World map of the Köppen-Geiger climate classification updated." Meteorologische Zeitschrift 15(3).

Koza, J. R. (1992). Genetic programming: on the programming of computers by means of natural selection, MIT press.

Kumar, V., S. K. Jain and Y. Singh (2010). "Analysis of long-term rainfall trends in India." Hydrological Sciences Journal 55(4): 484-496.

Lam, K. C., R. G. Bryant and J. Wainright (2015). "Application of Spatial Interpolation Method for Estimating the Spatial Variability of Rainfall in Semiarid New Mexico, USA." Mediterranean Journal of Social Sciences 6(4).

Lee, D. T. L. and A. Yamamoto (1994). "Wavelet analysis: theory and applications." Hewlett Packard journal 45: 44-44.

Legates, D. R. and G. J. McCabe (1999). "Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation." Water resources research 35(1): 233-241.

Li, B. and C. Cheng (2014). "Monthly discharge forecasting using wavelet neural networks with extreme learning machine." Science China Technological Sciences 57(12): 2441-2452.

Lima, A. R., A. J. Cannon and W. W. Hsieh (2016). "Forecasting daily streamflow using online sequential extreme learning machines." Journal of hydrology 537: 431-443.

Lin, J.-Y., C.-T. Cheng and K.-W. Chau (2006). "Using support vector machines for long-term discharge prediction." Hydrological Sciences Journal 51(4): 599-612.

Liu, Z., P. Zhou, G. Chen and L. Guo (2014). "Evaluating a coupled discrete wavelet transform and support vector regression for daily and monthly streamflow forecasting." Journal of hydrology 519: 2822-2831.

Lloyd, C. D. (2005). "Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain." Journal of Hydrology 308(1-4): 128-150.

Longobardi, A. and P. Villani (2010). "Trend analysis of annual and seasonal rainfall time series in the Mediterranean area." International Journal of Climatology 30(10): 1538-1546.

Luís, M. D., J. Raventós, J. C. González-Hidalgo, J. R. Sánchez and J. Cortina (2000). "Spatial analysis of rainfall trends in the region of Valencia (east Spain)." International Journal of Climatology 20(12): 1451-1469.

Machiwal, D. and M. K. Jha (2012). Hydrologic time series analysis: theory and practice, Springer Science & Business Media.

Maheswaran, R. and R. Khosa (2013). "Wavelets-based non-linear model for real-time daily flow forecasting in Krishna River." Journal of Hydroinformatics 15(3): 1022.

Mahsa, F., K. Majid and C. Seyyed Ali (2016). "Rainfall trend analysis of hydrological subbasins in western Iran." Journal of Irrigation and Drainage Engineering 142(10).

Maier, H. R. and G. C. Dandy (2000). "Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications." Environmental modelling & software 15(1): 101-124.

Mair, A. and A. Fares (2011). "Comparison of Rainfall Interpolation Methods in a Mountainous Region of a Tropical Island." Journal of Hydrologic Engineering 16(4): 371-383.

Makkeasorn, A., N.-B. Chang and X. Zhou (2008). "Short-term streamflow forecasting with global climate change implications–A comparative study between genetic programming and neural network models." Journal of hydrology 352(3): 336-354.

Mallat, S. G. (1989). "Multiresolution Approximations and Wavelet Orthonormal Bases of L²(R)." Transactions of the American Mathematical Society 315(1): 69-87.

Mamara, A., A. A. Argiriou and M. Anadranistakis (2016). "Recent trend analysis of mean air temperature in Greece based on homogenized data." Theoretical and Applied Climatology 126(3): 543-573.

Mamdani, E. H. and S. Assilian (1975). "An experiment in linguistic synthesis with a fuzzy logic controller." International Journal of Man-Machine Studies 7(1): 1-13.

Mann, H. B. (1945). "Nonparametric Tests Against Trend." Econometrica 13(3): 245-259.

Marengo, J. A. (2004). "Interdecadal variability and trends of rainfall across the Amazon basin." Theoretical and Applied Climatology 78(1): 79-96.

McCuen, R. H. (2002). Modeling Hydrologic Change: Statistical Methods, CRC Press.

Mehr, A. D. and E. Kahya (2017). "A Pareto-optimal moving average multigene genetic programming model for daily streamflow prediction." Journal of Hydrology 549: 603-615.

Mehr, A. D., E. Kahya and E. Olyaie (2013). "Streamflow prediction using linear genetic programming in comparison with a neuro-wavelet technique." Journal of Hydrology 505: 240-249.

Mehr, A. D., E. Kahya and M. Özger (2014). "A gene–wavelet model for long lead time drought forecasting." Journal of Hydrology 517: 691-699.

Mehr, A. D., E. Kahya, A. Şahin and M. J. Nazemosadat (2015). "Successive-station monthly streamflow prediction using different artificial neural network algorithms." International Journal of Environmental Science and Technology 12(7): 2191-2200.

Mehr, A. D. and V. Nourani (2017). "A Pareto-optimal moving average-multigene genetic programming model for rainfall-runoff modelling." Environmental Modelling & Software 92: 239-251.

Meyer, Y. (1990). "Ondelettes ét opétateurs, I: Ondelettes." Hermann, Paris.

Mitchell, T. D. and P. D. Jones (2005). "An improved method of constructing a database of monthly climate observations and associated high-resolution grids." International Journal of Climatology 25(6): 693-712.

Moges, S. A., B. F. Alemaw, T. R. Chaoka and R. K. Kachroo (2007). "Rainfall interpolation using a remote sensing CCD data in a tropical basin - A GIS and geostatistical application." Physics and Chemistry of the Earth 32(15-18): 976-983.

Mohammadi, K., S. Shamshirband, S. Motamedi, D. Petković, R. Hashim and M. Gocic (2015). "Extreme learning machine based prediction of daily dew point temperature." Computers and Electronics in Agriculture 117: 214-225.

Mondal, A., D. Khare and S. Kundu (2015). "Spatial and temporal analysis of rainfall and temperature trend of India." Theoretical and Applied Climatology 122(1): 143-158.

Moosavi, V., M. Vafakhah, B. Shirmohammadi and N. Behnia (2013). "A wavelet-ANFIS hybrid model for groundwater level forecasting for different prediction periods." Water resources management 27(5): 1301-1321.

Nakken, M. (1999). "Wavelet analysis of rainfall–runoff variability isolating climatic from anthropogenic patterns." Environmental Modelling & Software 14(4): 283-295.

Nalder, I. A. and R. W. Wein (1998). "Spatial interpolation of climatic Normals: test of a new method in the Canadian boreal forest." Agricultural and Forest Meteorology 92(4): 211-225.

NASA. (2016). "ASTER Global Digital Elevation Map." Retrieved 26/06/2016, from https://asterweb.jpl.nasa.gov/gdem.asp.

Nash, J. E. and J. V. Sutcliffe (1970). "River flow forecasting through conceptual models part I—A discussion of principles." Journal of hydrology 10(3): 282-290.

Nason, G. P. and R. Von Sachs (1999). "Wavelets in time-series analysis." Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 357(1760): 2511-2526.

Nayak, P. C., K. P. Sudheer, D. M. Rangan and K. S. Ramasastri (2004). "A neuro-fuzzy computing technique for modeling hydrological time series." Journal of Hydrology 291(1): 52-66.

New, M., M. Hulme and P. Jones (2000). "Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate." Journal of Climate 13(13): 2217-2238.

Nourani, V. (2017). "An Emotional ANN (EANN) approach to modeling rainfall-runoff process." Journal of Hydrology 544(Supplement C): 267-277.

Nourani, V., M. T. Alami and M. H. Aminfar (2009). "A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation." Engineering Applications of Artificial Intelligence 22(3): 466-472.

Nourani, V., A. H. Baghanam, A. Y. Rahimi and F. H. Nejad (2014). Evaluation of Wavelet-Based De-noising Approach in Hydrological Models Linked to Artificial Neural Networks. Computational Intelligence Techniques in Earth and

Environmental Sciences. T. Islam, P. K. Srivastava, M. Gupta, X. Zhu and S. Mukherjee. Dordrecht, Springer Netherlands: 209-241.

Nourani, V., A. Hosseini Baghanam, J. Adamowski and O. Kisi (2014). "Applications of hybrid wavelet–Artificial Intelligence models in hydrology: A review." Journal of Hydrology 514(Supplement C): 358-377.

Nourani, V., Ö. Kisi and M. Komasi (2011). "Two hybrid artificial intelligence approaches for modeling rainfall–runoff process." Journal of Hydrology 402(1): 41-59.

Nourani, V., M. Komasi and M. T. Alami (2011). "Hybrid wavelet–genetic programming approach to optimize ANN modeling of rainfall–runoff Process." Journal of Hydrologic Engineering 17(6): 724-741.

Nourani, V., M. Komasi and M. T. Alami (2013). "Geomorphology-based genetic programming approach for rainfall–runoff modeling." Journal of Hydroinformatics 15(2): 427-445.

Nourani, V., M. Komasi and A. Mano (2009). "A multivariate ANN-wavelet approach for rainfall–runoff modeling." Water resources management 23(14): 2877-2894.

Nourani, V. and M. Parhizkar (2013). "Conjunction of SOM-based feature extraction method and hybrid wavelet-ANN approach for rainfall–runoff modeling." Journal of Hydroinformatics 15(3): 829.

Nury, A. H., K. Hasan and M. J. B. Alam (2017). "Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh." Journal of King Saud University - Science 29(1): 47-61.

OE, T., W. M, v. d. W. F and D. H (2008). The use of geographic information systems in climatology and meteorology: COST Action 719, EUR-OP, Luxembourg.

Oguntunde, P. G., B. J. Abiodun and G. Lischeid (2011). "Rainfall trends in Nigeria, 1901–2000." Journal of Hydrology 411(3–4): 207-218.

Partal, T. and E. Kahya (2006). "Trend analysis in Turkish precipitation data." Hydrological processes 20(9): 2011-2026.

Pearson, K. (1896). "Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity, and Panmixia." Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 187: 253-318.

Percival, D. B. and A. T. Walden (2006). Wavelet methods for time series analysis, Cambridge university press.

Peterson, T. C., T. R. Karl, P. F. Jamason, R. Knight and D. R. Easterling (1998). "First difference method: Maximizing station density for the calculation of long-term global temperature change." Journal of Geophysical Research: Atmospheres 103(D20): 25967-25974.

Peterson, T. C., R. Vose, R. Schmoyer and V. Razuvaëv (1998). "Global historical climatology network (GHCN) quality control of monthly temperature data." International Journal of Climatology 18(11): 1169-1179.

Pettitt, A. N. (1979). "A Non-Parametric Approach to the Change-Point Problem." Journal of the Royal Statistical Society. Series C (Applied Statistics) 28(2): 126-135.

Ping, K. S. and N. Salim (2006). Optimized subtractive clustering for cluster-based compound selection.

Poulter, B., D. Frank, P. Ciais, R. B. Myneni, N. Andela, J. Bi, G. Broquet, J. G. Canadell, F. Chevallier and Y. Y. Liu (2014). "Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle." Nature 509(7502).

Pramanik, N., R. K. Panda and A. Singh (2011). "Daily river flow forecasting using wavelet ANN hybrid models." Journal of Hydroinformatics 13(1): 49.

Pujol, N., L. U. C. Neppel and R. Sabatier (2007). "Regional tests for trend detection in maximum precipitation series in the French Mediterranean region." Hydrological Sciences Journal 52(5): 956-973.

Rabunal, J. R., J. Puertas, J. Suarez and D. Rivero (2007). "Determination of the unit hydrograph of a typical urban basin using genetic programming and artificial neural networks." Hydrological processes 21(4): 476-485.

Rajaee, T., S. A. Mirbagheri, V. Nourani and A. Alikhani (2010). "Prediction of daily suspended sediment load using wavelet and neurofuzzy combined model." International Journal of Environmental Science & Technology 7(1): 93-110.

Raucher, R. S. (2010). The future of research on climate change impacts on water: a noaa/usepa/nasa/waterrf/werf workshop focusing on adaptation strategies and information needs, Water Research Foundation.

Ravansalar, M., T. Rajaee and O. Kisi (2017). "Wavelet-linear genetic programming: A new approach for modeling monthly streamflow." Journal of Hydrology 549: 461-475.

Ren, L., X.-Y. Xiang and J.-J. Ni (2013). "Forecast Modeling of Monthly Runoff with Adaptive Neural Fuzzy Inference System and Wavelet Analysis." Journal of Hydrologic Engineering 18(9): 1133-1139.

Robaa, S. M. and Z. J. AL-Barazanji (2013). "Trends of annual mean surface air temperature over Iraq." Nature & Science 11(12): 138.

Robinson, T. P. and G. Metternicht (2006). "Testing the performance of spatial interpolation techniques for mapping soil properties." Computers and Electronics in Agriculture 50(2): 97-108.

Saboohi, R., S. Soltani and M. Khodagholi (2012). "Trend analysis of temperature parameters in Iran." Theoretical and Applied Climatology 109(3): 529-547.

Samhouri, M., M. Abu-Ghoush, E. Yaseen and T. Herald (2009). "Fuzzy clustering-based modeling of surface interactions and emulsions of selected whey protein concentrate combined to ι-carrageenan and gum arabic solutions." Journal of Food Engineering 91(1): 10-17.

Sang, Y.-F. (2013). "A review on the applications of wavelet transform in hydrology time series analysis." Atmospheric Research 122(Supplement C): 8-15.

Sanikhani, H. and O. Kisi (2012). "River Flow Estimation and Forecasting by Using Two Different Adaptive Neuro-Fuzzy Approaches." Water Resources Management 26(6): 1715-1729.

Savic, D. A., G. A. Walters and J. W. Davidson (1999). "A genetic programming approach to rainfall-runoff modelling." Water Resources Management 13(3): 219-231.

Searson, D. P. (2015). GPTIPS 2: An Open-Source Software Platform for Symbolic Data Mining. Handbook of Genetic Programming Applications. A. H. Gandomi, A. H. Alavi and C. Ryan. Cham, Springer International Publishing: 551-573.

Searson, D. P., D. E. Leahy and M. J. Willis (2010). GPTIPS: an open source genetic programming toolbox for multigene symbolic regression, Citeseer.

Sen, P. K. (1968). "Estimates of the regression coefficient based on Kendall's tau." Journal of the American Statistical Association 63(324): 1379-1389.

Şen, Z. (2012). "Innovative Trend Analysis Methodology."

Sensoy, S., M. Demircan and Y. Ulupinar (2016). Climate of Turkey. Ankara, Turkey, Turkish State Meteorological Service: 13.

Sensoy, S., M. Demircan, Y. Ulupinar and İ. Balta (2008). "Climate of Turkey." Climate of Turkey. 2007. Devlet Meteoroloji İşleri Genel Müdürlüğü, 13 Feb. 2009< http://www. dmi. gov. tr/index. aspx.

Senthil Kumar, A. R., K. P. Sudheer, S. K. Jain and P. K. Agarwal (2005). "Rainfall-runoff modelling using artificial neural networks: comparison of network types." Hydrological Processes 19(6): 1277-1291.

Seo, Y., S. Kim and V. P. Singh (2015). "Estimating Spatial Precipitation Using Regression Kriging and Artificial Neural Network Residual Kriging (RKNNRK) Hybrid Approach." Water Resources Management 29(7): 2189-2204.

Seyam, M., F. Othman and A. El-Shafie (2017). "Prediction of Stream Flow in Humid Tropical Rivers by Support Vector Machines." MATEC Web Conf. 111.

Shiau, J.-T. and H.-T. Hsu (2016). "Suitability of ANN-Based Daily Streamflow Extension Models: a Case Study of Gaoping River Basin, Taiwan." Water Resources Management 30(4): 1499-1513.

Shiri, J. and O. Kisi (2010). "Short-term and long-term streamflow forecasting using a wavelet and neuro-fuzzy conjunction model." Journal of Hydrology 394(3): 486-493.

Shoaib, M., A. Y. Shamseldin and B. W. Melville (2014). "Comparative study of different wavelet based neural network models for rainfall–runoff modeling." Journal of hydrology 515: 47-58.

Shoaib, M., A. Y. Shamseldin, B. W. Melville and M. M. Khan (2015). "Runoff forecasting using hybrid wavelet gene expression programming (WGEP) approach." Journal of Hydrology 527: 326-344.

Shope, C. L. and G. R. Maharjan (2015). "Modeling Spatiotemporal Precipitation: Effects of Density, Interpolation, and Land Use Distribution." Advances in Meteorology.

Shumway, R. H. and D. S. Stoffer (2010). Time series analysis and its applications: with R examples, Springer Science & Business Media.

Singh, R. M. (2012). Wavelet-ANN Model for Flood Events. Proceedings of the International Conference on Soft Computing for Problem Solving (SocProS 2011) December 20-22, 2011: Volume 2. K. Deep, A. Nagar, M. Pant and J. C. Bansal. India, Springer India: 165-175.

Singh, S. K. (2016). "Long-term Streamflow Forecasting Based on Ensemble Streamflow Prediction Technique: A Case Study in New Zealand." Water Resources Management 30(7): 2295-2309.

Sivanandam, S. N. and S. N. Deepa (2007). Introduction to genetic algorithms, Springer Science & Business Media.

Smadi, M. M. (2006). "Observed abrupt changes in minimum and maximum temperatures in Jordan in the 20th century." Am. J. Environ. Sci 2(3): 114-120.

Smola, A. J. and B. Schölkopf (2004). "A tutorial on support vector regression." Statistics and computing 14(3): 199-222.

Solomatine, D., L. M. See and R. J. Abrahart (2008). Data-Driven Modelling: Concepts, Approaches and Experiences. Practical Hydroinformatics: Computational Intelligence and Technological Developments in Water Applications. R. J. Abrahart, L. M. See and D. P. Solomatine. Berlin, Heidelberg, Springer Berlin Heidelberg: 17-30.

Solomatine, D. P. and D. L. Shrestha (2009). "A novel method to estimate model uncertainty using machine learning techniques." Water Resources Research 45(12).

Subash, N. and A. K. Sikka (2014). "Trend analysis of rainfall and temperature and its relationship over India." Theoretical and Applied Climatology 117(3): 449-462.

Sudheer, K. P., A. K. Gosain and K. S. Ramasastri (2002). "A data-driven algorithm for constructing artificial neural network rainfall-runoff models." Hydrological processes 16(6): 1325-1330.

Tabari, H. and P. H. Talaee (2011). "Analysis of trends in temperature data in arid and semi-arid regions of Iran." Global and Planetary Change 79(1–2): 1-10.

Tabari, H. and P. H. Talaee (2011). "Temporal variability of precipitation over Iran: 1966–2005." Journal of Hydrology 396(3–4): 313-320.

Takagi, T. and M. Sugeno (1985). "Fuzzy identification of systems and its applications to modeling and control." IEEE transactions on systems, man, and cybernetics(1): 116-132.

Tan, Q., X. Wang, S. Cai and X. Lei (2015). Daily runoff time-series prediction based on the adaptive neural fuzzy inference system. 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), IEEE.

Taormina, R., S. Galelli, G. Karakaya and S. D. Ahipasaoglu (2016). "An information theoretic approach to select alternate subsets of predictors for data-driven hydrological models." Journal of Hydrology 542: 18-34.

Taxak, A. K., A. R. Murumkar and D. S. Arya (2014). "Long term spatial and temporal rainfall trends and homogeneity analysis in Wainganga basin, Central India." Weather and Climate Extremes 4: 50-61.

Theil, H. (1950). "A rank-invariant method of linear and polynomial regression analysis III." Nederl. Akad. Wetensch., Proc. 53: 386–392.

Thiessen, A. H. (1911). "Precipitation Averages for Large Areas." 39(7): 1082-1084.

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

Torrence, C. and G. P. Compo (1998). "A practical guide to wavelet analysis." Bulletin of the American Meteorological society 79(1): 61-78.

Torrence, C. and P. J. Webster (1999). "Interdecadal changes in the ENSO–monsoon system." Journal of Climate 12(8): 2679-2690.

Tsukamoto, Y. (1979). "An approach to fuzzy reasoning method." Advances in fuzzy set theory and applications.

Turkes, M. (1996). "Spatial and temporal analysis of annual rainfall variations in Turkey." International Journal of Climatology 16(9): 1057-1076.

Türkeş, M. (1996). "Spatial and temporal analysis of annual rainfall variations in Turkey." International Journal of Climatology 16(9): 1057-1076.

Türkeş, M., T. Koç and F. Sariş (2009). "Spatiotemporal variability of precipitation total series over Turkey." International Journal of Climatology 29(8): 1056-1074.

Türkeş, M. and U. M. Sümer (2004). "Spatial and temporal patterns of trends and variability in diurnal temperature ranges of Turkey." Theoretical and Applied Climatology 77(3): 195-227.

Türkeş, M., U. M. Sümer and İ. l. Demi˙r (2002). "Re-evaluation of trends and changes in mean, maximum and minimum temperatures of Turkey for the period 1929–1999." International Journal of Climatology 22(8): 947-977.

Türkeş, M., U. M. Sümer and G. KiliÇ (1996). "Observed changes in maximum and minimum temperatures in Turkey." International Journal of Climatology 16(4): 463-477.

Unal, Y. S., A. Deniz, H. Toros and S. Incecik (2012). "Temporal and spatial patterns of precipitation variability for annual, wet, and dry seasons in Turkey." International Journal of Climatology 32(3): 392-405.

Vapnik, V. (1992). The nature of statistical learning theory, Springer science & business media.

Villarini, G. (2012). "Analyses of annual and seasonal maximum daily rainfall accumulations for Ukraine, Moldova, and Romania." International Journal of Climatology 32(14): 2213-2226.

Wahba, G. (1990). Spline models for observational data. CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, PA,.

Wang, S., G. H. Huang, Q. G. Lin, Z. Li, H. Zhang and Y. R. Fan (2014). "Comparison of interpolation methods for estimating spatial distribution of precipitation in Ontario, Canada." International Journal of Climatology 34(14): 3745-3751.

Wang, W.-C., K.-W. Chau, C.-T. Cheng and L. Qiu (2009). "A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series." Journal of Hydrology 374(3): 294-306.

Webster R and O. MA (2007). Geostatistics for Environmental Scientists. West Sussex, UK, John Wiley & Sons, Ltd.

Willmott, C. J. (1981). "On the validation of models." Physical geography 2(2): 184-194.

WMO (2008). Guide to hydrological practices. Hydrology—from measurement to hydrological information. Geneva.

Wood, E. F. and A. Szollosi-Nagy (1980). Real-Time Forecasting/Control of Water Resource Systems; Selected Papers from an IIASA Workshop, October 18-21, 1976, Pergamon Press.

Xu, W. B., Y. J. Zou, G. P. Zhang and M. Linderman (2015). "A comparison among spatial interpolation techniques for daily rainfall data in Sichuan Province, China." International Journal of Climatology 35(10): 2898-2907.

Xu, Z. X., K. Takeuchi, H. Ishidaira and J. Y. Li (2005). "Long-term trend analysis for precipitation in Asian Pacific FRIEND river basins." Hydrological Processes 19(18): 3517-3532.

Yadav, B., S. Ch, S. Mathur and J. Adamowski (2016). "Discharge forecasting using an online sequential extreme learning machine (OS-ELM) model: a case study in Neckar River, Germany." Measurement 92: 433-445.

Yager, R. R. and D. P. Filev (1994). "Approximate clustering via the mountain method." IEEE Transactions on Systems, Man, and Cybernetics 24(8): 1279-1284.

Yang, X. H., X. J. Xie, D. L. Liu, F. Ji and L. Wang (2015). "Spatial Interpolation of Daily Rainfall Data for Local Climate Impact Assessment over Greater Sydney Region." Advances in Meteorology.

Yaseen, Z. M., A. El-Shafie, H. A. Afan, M. Hameed, W. H. M. W. Mohtar and A. Hussain (2016). "RBFNN versus FFNN for daily river flow forecasting at Johor River, Malaysia." Neural Computing and Applications 27(6): 1533-1542.

Yaseen, Z. M., A. El-Shafie, O. Jaafar, H. A. Afan and K. N. Sayl (2015). "Artificial intelligence based models for stream-flow forecasting: 2000–2015." Journal of Hydrology 530: 829-844.

Yaseen, Z. M., O. Jaafar, R. C. Deo, O. Kisi, J. Adamowski, J. Quilty and A. El-Shafie (2016). "Stream-flow forecasting using extreme learning machines: a case study in a semi-arid region in Iraq." Journal of Hydrology 542: 603-614.

Yavuz, H. and S. Erdoğan (2012). "Spatial analysis of monthly and annual precipitation trends in Turkey." Water Resources Management 26(3): 609-621.

Yue, S. and M. Hashino (2003). "Long term trends of annual and monthly precipitation in Japan." JAWRA Journal of the American Water Resources Association 39(3): 587-596.

Yue, S. and M. Hashino (2003). "LONG TERM TRENDS OF ANNUAL AND MONTHLY PRECIPITATION IN JAPAN1." JAWRA Journal of the American Water Resources Association 39(3): 587-596.

Yürekli, K. (2015). "Impact of climate variability on precipitation in the Upper Euphrates–Tigris Rivers Basin of Southeast Turkey." Atmospheric Research 154: 25-38.

Zadeh, M. R., S. Amin, D. Khalili and V. P. Singh (2010). "Daily outflow prediction by multi layer perceptron with logistic sigmoid and tangent sigmoid activation functions." Water resources management 24(11): 2673-2688.

Zhang, X., Y. Peng, C. Zhang and B. Wang (2015). "Are hybrid models integrated with data preprocessing techniques suitable for monthly streamflow forecasting? Some experiment evidences." Journal of Hydrology 530: 137-152.

CURRICULUM VITAE

Name-Surname: Sinan Jasim Hadi ALDOORI

Foreign Languages: Turkish, English

Year and Place of birth: Iraq- Sallahuddin /1982

E-Mail : sinan.jasim@yahoo.com

Education and Profession:

B. sc. : University of Tikrit, Faculty of Engineering, Civil Engineering Department, 2001-2006.

M. sc. : University Putra Malaysia, Faculty of Engineering, Civil Engineering Department, 2010-2012.

PhD : Anadolu University, Earth and Space Science Institution, Department of Remote Sensing and GIS, 2013-2018.

Scientific Publications/Activities:

Hadi Sinan, J. and M. Tombul (2018). "Long-term spatiotemporal trend analysis of precipitation and temperature over Turkey." Meteorological Applications. (Article published in Early View on 6 March, 2018).

Abba, S. I., S. J. Hadi and J. Abdullahi (2017). "River water modelling prediction using multi-linear regression, artificial neural network, and adaptive neuro-fuzzy inference system techniques." Procedia Computer Science **120**: 75-82.

Hadi, S. J., H. Z. M. Shafri and M. D. Mahir (2014). "Factors affecting the eco-environment identification through change detection analysis by using remote sensing and GIS: a case study of Tikrit, Iraq." Arabian Journal for Science and Engineering **39**(1): 395-405.

Hadi, S. J., A. R. M. Shariff and M. D. Mahir (2012). "A review of NSDI aspects (Components, Challenges, and Advantages)." International Journal of Engineering Research and Applications (IJERA) **2**(3).

Hadi, S. J. and M. Tombul (2017). "Analyzing the historical rainfall observations and examining the flood event of Hopa, Turkey." International Journal of Environmental Science & Technology **1(1).**

Hadi, S. J. and M. Tombul (2017). Analyzing trend and developing an empirical formula for estimating rainfall intensity: a case study Eskisehir city, Turkey.

International Conference on Advances in Sustainable Construction Materials & Civil Engineering Systems (ASCMCES-17), UAE, EDP Sciences.

Hadi, S. J. and M. Tombul (2017). Conversion of CruTS 3.23 data and evaluation of precipitation and temperature variables in a local scale. International Conference on Advances in Sustainable Construction Materials & Civil Engineering Systems (ASCMCES-17), UAE, EDP Sciences.

Al-Sharif, A. A. A., B. Pradhan, S. J. Hadi and N. Mola (2013). Revisiting methods and potentials of sar change detection. the World Congress on Engineering Vol III, London UK.

Hadi, S. J., O. F. Althwaynee, H. Z. M. Shafri and B. Pradhan (2011). The use of GPS in mitigating and managing landslides. 10th International Symposium & Exhibition on Geoinformation (ISG 2011), Malaysia.

Hadi, S. J., H. Z. M. Shafri and M. D. Mahir (2014). Modelling LULC for the period 2010-2030 using GIS and Remote sensing: a case study of Tikrit, Iraq. 7th IGRSM International Remote Sensing & GIS Conference and Exhibition, Kuala Lumpur, Malaysia, IOP Publishing