

152273

**ISTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY**

**NEW STATISTICAL DOWNSCALING METHODS  
AND APPLICATIONS FOR TURKEY**

**Ph.D. Thesis by  
Hasan TATLI, M.Sc.**

**(511972050)**

**Date of submission: 16 June 2004**

**Date of defence examination: 20 December 2004**

**Supervisor (Chairman): Prof.Dr. H. Nüzhet DALFES**

**Co-Supervisor: Ass. Prof.Dr. Ş. Sibel MENTEŞ**

**Members of the Examining Committee: Prof.Dr. Metin DEMİRALP**

**Prof.Dr. Mehmet KARACA**

**Prof.Dr. Mete TAYANÇ (MU.)**

**Assoc. Prof.Dr. Murat TÜRKEŞ (ÇOMU.)**

**Assoc. Prof.Dr. Yurdanur S. ÜNAL**

**DECEMBER 2004**

**YENİ İSTATİSTİKSEL ÖLÇEK KÜÇÜLTME  
YÖNTEMLERİ VE TÜRKİYE İÇİN UYGULAMALAR**

**DOKTORA TEZİ**

**Met. Y. Müh. Hasan TATLI**

**(511972050)**

**Tezin Enstitüye Verildiği Tarih: 16 Haziran 2004**

**Tezin Savunulduğu Tarih: 20 Aralık 2004**

**Tez Danışmanı: Prof.Dr. H. Nüzhet DALFES**

**Tez Eş-Danışmanı: Yard. Doç.Dr. Ş. Sibel MENTES**

**Diğer Jüri Üyeleri: Prof.Dr. Metin DEMİRALP**

**Prof.Dr. Mehmet KARACA**

**Prof.Dr. Mete TAYANÇ (MÜ.)**

**Doç.Dr. Murat TÜRKEŞ (ÇOMÜ.)**

**Doç.Dr. Yurdanur S. ÜNAL**

**ARALIK 2004**

## **ACKNOWLEDGEMENTS**

I would like to thank the members of my advisory committee (Professor H. Nüzhet Dalfes, Asst.Prof.Dr. Ş. Sibel Menteş, Professor Metin Demiralp, Professor Mete Tayanç, Assoc.Prof.Dr. Yurdanur S. Ünal) and, especially, chair and co-chair, Professor H. Nüzhet Dalfes and Asst.Prof.Dr. Ş. Sibel Menteş for the enthusiasm, knowledge, guidance, and encouragement that they have provided over the years.

I am also grateful to Professor H. Nüzhet Dalfes for introducing me to the field of climate-dynamics and large scale climatology.

I am grateful to many individuals who have contributed not only to this dissertation, but also my education in general. In particular, I thank Professor Yunus Borhan, Professor Mikdat Kadioğlu, Assoc.Prof.Dr. Kasım Koçak, Professor Mehmet Karaca and the staff of the Department of Meteorology Engineering who provided many hours of guidance and assistance throughout this study.

I would also like to thank Assoc. Prof.Dr. Murat Türkeş from Çanakkale 18 Mart University, who offered valuable suggestions that made it possible to finalize this research.

I would also like to thank Barış Önel for supplying some of the data sets used to illustrate the methods.

Most of all, I want to acknowledge the support and sacrifices of my wife, Handan, and our son, Mert Deniz.

**Hasan TATLI**

**December, 2004**

## CONTENTS

<b>ABBREVIATIONS</b>	<b>vi</b>
<b>LIST OF TABLES</b>	<b>vii</b>
<b>LIST OF FIGURES</b>	<b>viii</b>
<b>LIST OF SYMBOLS</b>	<b>xi</b>
<b>SUMMARY</b>	<b>xii</b>
<b>ÖZET</b>	<b>xiv</b>
<b>1. INTRODUCTION</b>	<b>1</b>
<b>2. DATA AND METHODS</b>	<b>9</b>
2.1 Data and Data Preparation	9
2.2 Methods	11
2.2.1 Singular Value Decomposition (SVD)	11
2.2.2 Principal Component Analysis (PCA), Factor Analysis (FA) and Independent Component Analysis (ICA)	12
2.2.3 Canonical Correlation Analysis (CCA)	16
2.2.4 Redundancy Analysis (RA)	17
2.2.5 Sampson Correlation Ratio ( $R_s$ )	18
2.2.6 Singular Spectrum Analysis (SSA)	18
2.2.7 Neural Networks	20
2.3 Statistical Significance Tests	24
2.3.1 Tests for PCA, FA and CCA	24
2.3.2 Tests for the Sample Correlation	25
<b>3. PROPOSED DOWNSCALING MODELS</b>	<b>27</b>
3.1 Downscaling Methods	27
3.2 Proposed Downscaling Model Based on Multivariate Techniques	29
3.3 Proposed Downscaling Model Based on Recurrent Neural Networks	31
<b>4. A GENERAL LOOK AT THE CLIMATE-CHARACTERISTICS OF PRECIPITATION AND TEMPERATURE OVER TURKEY</b>	<b>36</b>
4.1 Precipitation	36
4.1.1 Precipitation in Summer Months	36
4.1.2 Precipitation in Winter Months	37
4.2 Temperature	39
<b>5. RESULTS</b>	<b>41</b>
5.1 Downscaling Monthly Total Precipitation over Turkey	41
5.1.1 Problem Statement	42
5.1.2 Sampson Correlation Pattern Analysis of Turkish Rainfall Associated with Large Scale Climate Features	46
5.1.3 Downscaling Results	50
5.1.4 Summary	54

5.2 Surface Air Temperature Variability over Turkey and Its Connection to the Large Scale Upper Air Circulation	55
5.2.1 Problem Statement	55
5.2.2 Sampson Correlation Pattern Analysis of Turkish Surface Air Temperatures Associated with Upper Air Circulations	58
5.2.3 Downscaling Results	63
5.2.4 Summary	70
<b>6. SUMMARY AND CONCLUSIONS</b>	<b>72</b>
<b>REFERENCES</b>	<b>77</b>
<b>APPENDIX</b>	<b>88</b>
<b>CURRICULUM VITAE</b>	<b>92</b>



## **ABBREVIATIONS**

<b>AR</b>	: Autoregressive
<b>ARMA</b>	: Autoregressive and Moving Average
<b>CCA</b>	: Canonical Correlation Analysis
<b>ENSO</b>	: El Niño Southern Oscillation
<b>FA</b>	: Factor Analysis
<b>FASTICA</b>	: Fast Independent Component Analysis
<b>GCM</b>	: General Circulation Model
<b>hPa</b>	: Hecto Pascal
<b>ICA</b>	: Independent Component Analysis
<b>IPCC</b>	: Intergovernmental Panel on Climate Change
<b>NAO</b>	: North Atlantic Oscillation
<b>NCEP-NCAR</b>	: National Centers for Environmental Prediction-National Center of Atmospheric Research
<b>NWP</b>	: Numerical Weather Prediction
<b>MOS</b>	: Model Output Statistics
<b>PCA</b>	: Principal Component Analysis
<b>PP</b>	: Perfect Prognosis
<b>RA</b>	: Redundancy Analysis
<b>RNN</b>	: Recurrent Neural Network
<b>SSA</b>	: Singular Spectrum Analysis
<b>SVD</b>	: Singular Value Decomposition

**LIST OF TABLES**

	<u><b>Page No</b></u>
<b>Table 2.1</b> The look up table of SSA of the maximum near-surface air temperature series for some Turkey stations .....	20
<b>Table 3.1</b> The proposed method algorithm.....	34
<b>Table 5.1</b> The Turkey stations of monthly total precipitation in the application.....	44
<b>Table 5.2</b> The significant principal components of the large scale variables according to the criterion of Jöreskog and Sörbom, (1989).....	45
<b>Table 5.3</b> The canonical correlation variates of the large scale variables (7 <sup>th</sup> canonical correlation variate is more significant than the 6 <sup>th</sup> canonical correlation variate).....	46
<b>Table 5.4</b> Turkey stations for near-surface air temperature used in the application.....	57

## LIST OF FIGURES

	<u>Page No</u>
<b>Figure 1.1</b> Spatial correlation functions of monthly anomalies: 1 precipitation, longitudinal direction, (65°N, January); 2 the same but for July; 3 North Atlantic Sea Surface Temperature (SST), latitudinal direction; 4 North Atlantic SST, longitudinal direction (Dobrovolski, 2000).....	5
<b>Figure 1.2</b> The results of the 17 Climate Scenarios from IPCC (Houghton <i>et al.</i> , 2001) → <a href="http://ipcc-ddc.cru.uea.ac.uk/">http://ipcc-ddc.cru.uea.ac.uk/</a> .....	6
<b>Figure 2.1</b> Distribution of the 221 grid points (2.5° longitude by 2.5° latitude) for NCAR/NCEP re-analysis data field. The area of Turkey enclosed by the rectangular box is shown in additional detail in Figure 2.2.....	10
<b>Figure 2.2</b> Geographical distribution for Turkey stations: (a) The stations of the monthly total precipitation; (b) the stations of the near-surface air temperatures .....	11
<b>Figure 2.3</b> Structure of a formal neuron (a), and a multilayer network (b).....	22
<b>Figure 2.4</b> Structure of the Jordan-type recurrent neural network.....	23
<b>Figure 2.5</b> Structure of the Elman type recurrent neural network.....	23
<b>Figure 3.1</b> Formal representation the issues of scales associated with the climate features.....	29
<b>Figure 3.2</b> The proposed model components.....	30
<b>Figure 3.3</b> The flow chart of the components in the proposed downscaling approach.....	34
<b>Figure 5.1</b> The first three principal components and independent components of mean sea level pressure series.....	43
<b>Figure 5.2</b> The spectral densities of the first three CCVs. The horizontal and vertical axes represent frequency (in <i>cycle/month</i> ) and spectral density, respectively: (a) the large scale variables (on left); (b) the local scale variables (on right).....	45
<b>Figure 5.3</b> The $R_s$ pattern between the MSLP and monthly total precipitation series: (a) winter months; (b) summer months.....	47
<b>Figure 5.4</b> The $R_s$ pattern between the 500hPa geopotential heights and the monthly total precipitation series: (a) winter months; (b) summer months.....	48
<b>Figure 5.5</b> The $R_s$ pattern between the 700 hPa geopotential heights and the monthly total precipitation series: (a) winter months; (b) summer months.....	48
<b>Figure 5.6</b> The $R_s$ pattern between the 500-1000 hPa thicknesses and the monthly total precipitation series: (a) winter months; (b) summer months.....	49



<b>Figure 5.7</b>	The $R_S$ pattern between the 500 hPa vertical pressure velocities and the monthly total precipitation series: (a) winter months; (b) summer months .....	50
<b>Figure 5.8</b>	The first PC pattern of the monthly total precipitation series (the map indicates the correlation coefficients between the monthly total precipitation time series and the first PC): (a) observed; (b) downscaled.....	51
<b>Figure 5.9</b>	The second PC pattern of the monthly total precipitation series: (a) observed; (b) downscaled.....	51
<b>Figure 5.10</b>	The third PC pattern of the monthly total precipitation series: (a) observed; (b) downscaled.....	51
<b>Figure 5.11</b>	Geographical distribution of the mean square errors of the proposed model output for the validation period of the 1980-1998 interval.....	52
<b>Figure 5.12</b>	The scatter plots of the identified model outputs versus actual monthly total precipitation series. The correlation coefficients between the observed and model outputs are abbreviated as $r_1$ for the training part and $r_2$ for non-training part respectively: (a) Göztepe (İstanbul); (b) Ankara; (c) Diyarbakır; (d) İzmir; (e) Rize; (f) Adana; (g) Erzurum.....	53
<b>Figure 5.13</b>	$R_S$ pattern between the 500 hPa geopotential heights and maximum temperature series: (a) winter; (b) spring; (c) summer; and (d) autumn.....	58
<b>Figure 5.14</b>	$R_S$ pattern between the 500-1000 hPa thicknesses and maximum temperature series: (a) winter; (b) spring; (c) summer; and (d) autumn.....	59
<b>Figure 5.15</b>	$R_S$ pattern between the 500 hPa geopotential heights and minimum temperature series: (a) winter; (b) spring; (c) summer; and (d) autumn.....	60
<b>Figure 5.16</b>	$R_S$ pattern between the 500 hPa-1000 hPa thicknesses and minimum temperature series: (a) winter; (b) spring; (c) summer; and (d) autumn.....	61
<b>Figure 5.17</b>	$R_S$ pattern between the 500 hPa geopotential heights and mean temperature series: (a) winter; (b) spring; (c) summer; and (d) autumn.....	62
<b>Figure 5.18</b>	$R_S$ pattern between the 500-1000 hPa thicknesses and mean temperature series: (a) winter; (b) spring; (c) summer; and (d) autumn.....	63
<b>Figure 5.19</b>	The varimax rotated principal component patterns of the downscaling model outputs and the actual maximum temperature: a and b show the first two PC patterns of noise-free data; c and d indicate the first two PC patterns of model outputs; e, f, and g indicate the first three PC patterns of actual raw data. The third PC patterns of both noise-free and model-output data are not significant in contrast to the third PC pattern of raw data (the patterns represent correlation coefficients between individual station's temperature series and rotated principal components. These patterns are generally called loadings). The period of the extracted principal components is the period of validation period of the suggested model (18 years data)	64

<b>Figure 5.20</b>	The varimax rotated principal components of the downscaling model outputs and the actual minimum temperature: a and b show the first two PCs of noise-free data; c and d indicate the first two PCs of model outputs; e and f indicate the first two PCs of actual raw data. The third PCs of the noise-free, model-output, and actual raw data are not significant.....	65
<b>Figure 5.21</b>	The varimax rotated principal components of the downscaling model outputs and actual mean temperature: a and b show the first two PCs of noise-free data; c and d indicate the first two PCs of model outputs; e and f indicate the first two PCs of actual raw data. The third PCs of the noise-free, model-output, and actual raw data are not significant.....	66
<b>Figure 5.22</b>	Geographical distribution the mean square errors during the period of 1980-1998 (validation part): (a) Maximum temperatures; (b) mean temperatures; and (c) minimum temperatures.....	67
<b>Figure 5.23</b>	The frequency scatter plots of the minimum temperatures of the observations and model outputs during the calibration period of the proposed model. The observations are the series of which are not filtered by SSA. The ellipses indicate the 95% confidence interval, and the data in the outside of the ellipses is the unpredictable part of the data by the proposed model: (a) Göztepe (Istanbul); (b) Ankara; (c) Diyarbakır; (d) İzmir; (e) Adana; (f) Adana; (g) Erzurum.....	68
<b>Figure 5.24</b>	The frequency scatter plots of the observed minimum temperatures and model outputs during the validation period of the proposed model. The observations are the series which are not filtered by SSA. The ellipses indicate the 95% confidence interval, and the data in the outside of the ellipses is the unpredictable part of the data by the proposed model. The entire data of the validation period is 216 months (18 years). As seen in Figures 5.23 and in this figure, the outliers of the both calibration and validation part data show parallel features according to 95% significant level: (a) Göztepe (Istanbul); (b) Ankara; (c) Diyarbakır; (d) İzmir; (e) Adana; (f) Adana; (g) Erzurum.....	69

## LIST OF SYMBOLS

$\alpha$	: Statistical significance level
$\lambda$	: Eigenvalue
$\psi$	: Activation function
$\phi$	: Integration function
$\rho$	: Population correlation
$C$	: Covariance or cross-covariance matrix
$E$	: Expected value
$F$	: F-distribution statistics
$F$	: Operator
$F_X$	: Common factors of the multivariate $X$
$J$	: Negentropy (or information measure)
$N(x, w, \psi, \phi)$	: Quadruple of a formal neuron $N$
$r$	: Sample correlation
$R_S$	: Sampson correlation ratio
$S$	: Independent components of the multivariate $X$
$t$	: time index or t-distribution statistics
$t_i$	: $i$ th canonical correlation variate for predictors
$u_i$	: $i$ th canonical correlation variate for predictands
$V$	: Redundancy weights for $Y$
$W$	: Redundancy weights for $X$
$V_X$	: Principal components for $X$

## **NEW STATISTICAL DOWNSCALING METHODS AND APPLICATIONS FOR TURKEY**

### **SUMMARY**

Weather and climate have a profound influence on life on earth. Understanding the impacts of climate variability and climate change reveal the increasing need for improvements in regional and local scale climate diagnoses, which are the main goals of climate research.

The climate impact studies usually require detailed information with high resolution and accuracy for past, present, and future. The main and important tools of studying climate are the general circulation models (GCMs). Unfortunately, GCMs running on coarse resolution may not be able to detect the local scale climate variability. Correspondingly, the limitations concerning spatial resolution and accuracy of simulated data by GCMs arise to need a downscaling strategy. For instance, hydrologists ask for daily total precipitation with spatial resolutions corresponding to their catchments areas of interest. The ecologists who are studying the dynamics and responses of forests in the mountainous terrain need to know monthly mean precipitation and temperature values with a resolution of a few kilometers.

In most cases comprehensive information is required, with a spatial resolution of order of 50-100 km or less, and variables describing the surface processes. However, GCMs can not simulate such surface variables with high accuracy, and the output of those models is often implicitly considered as a continuous domain. Then, gridding is just a convenient way to store the output economically; the information determined by grid is dependable and sub-grid scale information may be recovered from the gridded data simply by spatial interpolation methods.

With such a background, it is fully acceptable to use the output of a GCM model, which runs typically on the order of  $400 \times 400 \text{ km}^2$  grid scale, and to infer the possible climate variability and climate change on Istanbul and Bolu which have different locations and different climates. Therefore, it is obvious that the resolution is one of the drawbacks associated with the GCMs to derive the local climate information.

The spatial-scale difference between climate research and climate impact studies has to be bridged by “downscaling” on the side of the climate research and “upscaling” on the side of the climate impact research.

This study is the first one to search and develop downscaling strategies for Turkey. Two new downscaling strategies for climate diagnosis are developed in this study. The proposed methods are based on artificial recurrent neural networks (RNN) and multivariate statistical techniques that derive transfer functions from the large scale free troposphere variables which are assumed to steer the regional climate over Turkey.

Any downscaling model is sensitive to the asymmetric relations between the large scale variables and the local scale variables. Therefore, before building the model the relations between the large scale and local scale variables are analyzed by Sampson Correlation ratio. Hence, the large scale influences are interpreted readily by Sampson Correlation ratio. On the other hand, extraction the relationships between local and large scale variables with the canonical correlation patterns (a common applied multivariate method based upon the symmetrical assumptions between different climate variables) observed less exposition than the Sampson correlation ratio patterns.

The first RNN model employs canonical correlation variates based on Independent Component Analysis (ICA) from significant Principal Components Analysis (PCA) of mean sea level pressure (MSLP), 500 and 700 hPa geopotential height, 500-1000 hPa thickness, and 500 hPa vertical pressure-velocity ( $\omega$ ) which are mentioned as the predictors of the monthly total precipitation for Turkey stations. The proposed RNN model originated from Jordan type RNN is able not only to reflect the influences of the large scale circulations but also impose the local scale climate features.

The performance of the proposed downscaling model decreases towards the inner parts of the Anatolian Peninsula of Turkey. However, the achievement of the model over the transition Mediterranean is relatively superior to the Black Sea region. The large errors over Black Sea region are due to the local factors such as topography that causes the rain-shadows and orographic rainfalls; thus the large scale predictors fail to capture these local climate features.

The second downscaling model is based on multivariate statistical techniques for the monthly near-surface air temperatures over Turkey. Since the periodic components in the temperature series are dominant, in the first stage of the model, Singular Spectrum Analysis (SSA) based on time lagged covariance matrix of statistical preprocessing of both the predictors and predictands is employed to filter out these periodicities.

In the second stage, principal component analysis is employed to the non-periodic components in order to reproduce the noise-free components of the near-surface air temperature series. The model is constructed via redundancy analysis (RA) due to the asymmetric relations between the large scale and local scale processes. The large scale noise-free 500 hPa geopotential heights and 500-1000 hPa thickness are selected as the predictors for the monthly near-surface air temperatures. The second model is able to explain 90% variance in the near surface air temperatures.

The results of the developed downscaling strategies demonstrate that the ability of the RNN and RA to downscale realistically the relationships between the large scale circulation fields and monthly total precipitation–near surface air temperature series. The results of this study also reveal that the statistical preprocessing by PCA, ICA, CCA, RA and SSA increases the performance of the models during simulations.



## YENİ İSTATİSTİKSEL ÖLÇEK KÜÇÜLTME YÖNTEMLERİ VE TÜRKİYE İÇİN UYGULAMALAR

### ÖZET

Hava ve iklim süreçlerinin yeryüzündeki yaşam üzerinde hayati etkileri vardır. İklim değişkenliği ve değişikliğinin etkisinin anlaşılabilmesi, bölgesel ve yerel iklim tanılarının geliştirilmesi, iklim çalışmalarının temel konularıdır.

İklim etkilerini araştıranlar, genelde iklim değişkenliğinin geçmişi, bugünü ve geleceği hakkında detaylı ve doğru bilgilere gereksinim duyarlar. İklim değişimi (değişkenliği) konusunda önemli ve temel araçlar Genel Dolaşım Modelleridir (GDM). Ne yazık ki kaba sayısal çözünürlük üzerinde çalıştırılan GDM'ler yerel iklim değişkenliği hakkında detaylı bilgi veremezler. Bu bağlamda, GDM'lerin uzaysal çözünürlükteki ve benzeşim verilerinin doğruluğundaki kısıtlamalardan dolayı “uzaysal ölçek küçültme (downscaling)” stratejilerinin geliştirilmesine gereksinim duyulur. Örneğin, hidrologistler, ele aldıkları havza ölçeğine bağlı çözünürlükte, günlük toplam yağış verilerine gereksinim duyarlar. Ekolojistler, dağlık bölgelerdeki ormanların dinamiği ve yanıtları konusunda aylık yağış ve sıcaklıkların birkaç km'lik çözünürlükteki verisine gereksinim duyarlar.

Yapılan çalışmaların büyük bir kısmında, yüzey değişkenliği (değişim) süreçlerini niteleyen yüzey değişkenlerinin, 50-100 km veya daha küçük ölçekteki benzeşimlerinin doğru ve anlaşılır bilgisine gereksinim duyulur. Ancak, GDM'ler yüzey değişkenlerin değişkenliğinin benzeşimini doğrulukla veremedikleri gibi, bu model çıktıları, çalışma ortamını, sürekli kabul ederler. Böylece, sayısal ağ, sadece ekonomik anlamda, verilerin depolanması anlamına gelir ki alt-grid (sub-grid) ölçek bilgisi gride bağımlı ve sıradan uzaysal enterpolasyona dönüşür.

Yukarıdaki bilgilere bağlı olarak; örneğin İstanbul ve Bolu gibi iklimi birbirinden farklı iki bölgenin iklim değişkenliğini veya değişimini, tipik olarak seçilen  $400 \times 400 \text{ km}^2$  sayısal-ağ ölçeğinde çalıştırılan GDM'lerden çıkarıma varmanın yanlış sonuca götürdüğü basitçe görülebilir.

İklim ve etkilerinin arasındaki bu uzaysal ölçek farkı, iklim araştırmalar için “uzaysal ölçek küçültme” ve iklim etkileri araştırmaları için de “uzaysal ölçek büyültme (upscaling)” yöntemleri ile kapatılmalıdır.

Bu çalışma, Türkiye üzerinde, uzaysal ölçek küçültme stratejilerinin geliştirilmesi konusunda ilk araştırmadır. Çalışmada, iklim tanısı için iki yeni model geliştirilmiştir. Önerilen modeller, Türkiye üzerindeki yerel iklimi idare ettiği düşünülen büyük ölçekli serbest troposferin değişkenlerinden elde edilen dönüşüm fonksiyonlarını kullanan, Yinelemeli Yapay Sinir Ağları (YYSA) ve çok değişkenli istatistik yöntemleri üzerine dayanmaktadır.

Herhangi bir ölçek küçültme modeli, büyük ölçekli değişkenler ile küçük ölçekli değişkenler arasındaki asimetric (bakışık olmayan) bağıntılara duyarlıdır; dolayısıyla

model kurulmadan önce büyük ölçekli ve küçük ölçekli değişkenler arasındaki bağıntılar, Sampson ilişki oranı ile incelemeye tabi tutulmuştur. Bu bağlamda, büyük ölçekli değişkenlerinin etkilerinin yorumlanması, Sampson ilişki oranıyla daha basite indirgenebilmektedir. Diğer taraftan, büyük ve küçük ölçekli değişkenler arasındaki bağıntıların kümelerarası ilişki desenleri ile incelenmesi (farklı iklim değişkenleri arasındaki ilişkilerin bakışık olduğunu kabul eden ve yaygın kullanılan çok değişkenli istatistiksel bir yöntem), Sampson ilişki desenleri ile incelenmesine göre daha az açıklayıcı olduğu gözlenmiştir.

Birinci uygulanan YYSa ölçek küçültme modelinde; ilk adım olarak büyük ölçekli ortalama deniz seviyesi basınçlarının (ODSB), 500 ve 700 hPa jeopotansiyel yüksekliklerinin, 500-1000 hPa jeopotansiyel kalınlıkların ve 500 hPa düşey basınç-hızlarının (omega) anlamlı Asal Bileşenleri (AB) bulunmuş, daha sonra AB'lerden Bağımsız Bileşenleri Analizi (BBA) ile bulunan Bağımsız Bileşenlerin (BB) Kümeler Arası İlişki (KAİ) değişkenleri, Türkiye'nin aylık toplam yağışlarının tahmin edicileri olarak ele alınmıştır. Önerilen Jordan türü YYSa modelinin, sadece büyük-ölçekli çevrimlerin etkisini yansıtmakla kalmayıp, yerel-ölçekli iklim özelliklerini yansıtmakta olduğu da görülmüştür.

Modelin veriminin, Türkiye'nin Anadolu Yarımadası'nın iç bölgelerine doğru azaldığı görülmüştür. Oysa Akdeniz geçiş bölgesinde, modelin verimi, Karadeniz bölgesine göre, bağıl olarak, çok daha iyi sonuç vermektedir. Karadeniz bölgesinde görülen büyük hataların ise büyük ölçekli tahmin edicilerin yakalayamadığı, örneğin topografya, ki yağmur-gölgelerine ve orografik yağışlara neden olur, gibi yerel özellikli faktörlerden kaynaklandığı sonucuna varılmıştır.

İkinci model, Türkiye yüzeye-yakın hava sıcaklıklarının ölçek küçültmesinde kullanılan, çok değişkenli istatistik yaklaşımlara dayanmaktadır. Yere-yakın hava sıcaklık serilerinde görülen baskın periyodiklerden dolayı, modelin birinci adımında, hem tahmin edicilerin hem de tahmin edilenlerin, zaman ortamında kaydırılmış özvaryans (covariance) matrisleri üzerine uygulanan tekil spektrum analizinin (TSA) istatistiksel ön-inceleme uygulanarak, bu periyodik yapılar süzgeçleşmiştir (filtering).

Modelin ikinci adımında, periyodik olmayan bileşenlere asal bileşenler analizi uygulanarak, sonuçta, gürültülerden arındırılmış bileşenlerden sıcaklık serileri tekrar elde edilmiştir. Büyük ölçekli değişkenler ile küçük ölçekli değişkenler arasındaki asimetric ilişkilerden dolayı, model, gereksizlik analizine (redundancy analysis) dayalı olarak oluşturulmuştur. Büyük ölçekli gürültüsüz 500 hPa jeopotansiyel yüksekliği ve 500-1000 hPa kalınlığı, Türkiye yere-yakın hava sıcaklıklarının tahmin edicileri olarak ele alınmıştır. İkinci modelin Türkiye yere-yakın hava sıcaklık varyansını %90 oranında yakalamayı başardığı gözlemlenmiştir.

Önerilen model sonuçlarından görüleceği üzere, modeller, başarılı bir şekilde büyük ölçekli dolaşım desenleri ile Türkiye yağışları ve yere-yakın hava sıcaklıklarının ölçek küçültmesini yapabilmektedir. Bu araştırmanın sonuçları göstermektedir ki ABA, BBA, KAİ, GA ve TSA ile istatistiksel ön-inceleme yapılması, model benzeşim verimine artırıcı etkisi olmaktadır.

## 1. INTRODUCTION

Weather and climate have a profound influence on life on earth. Understanding the causes of climate variability is one of the main goals of climate research. The increased CO<sub>2</sub> concentrations in the atmosphere due to the human activities will enhance the greenhouse effect of the atmosphere which will therefore change the global climate in the future. In order to predict how the climate will change in the future, the knowledge of the climate variability is fundamental. General Circulation Models (GCMs) with coarse-resolution are very important tools to simulate past and future climate. However, GCMs fail to simulate the climate variability on the spatial scales of which represents the local scale climate features.

Climate variability exists on all time and space scales, ranging from time scales as short as months and ending on the time scales comparable to the age of the Earth. Climate variations and change, caused by external forcing, may be partly predictable, particularly on the larger, continental and global, spatial scales. In this context the climate system is defined to consist of the entire atmosphere, the oceans, the cryosphere, and the biosphere. The sources of climate variability can therefore be attributed to external forcing and internal variability. Climate variability, which is generated within the climate subsystems, is internal variability, whereas climate variability driven by external forcing, such as solar variability, volcanoes and human activities, is considered as externally forced climate variability.

GCMs are powerful tools for analysis of large scale climate features based on physical laws for atmospheric composition through numerical solutions of partial differential equations which are highly nonlinear. However, at local scales or scales near surface, the outputs of GCMs may not be able to estimate the nature of climate.

During the first half of 20<sup>th</sup> Century, climatology was thought as a sub-discipline of meteorology and geography studying the atmosphere, oceans, cryosphere, solid earth and biosphere independently; while meteorology was concerned with the general circulation of the atmosphere, climatology was interested in the statistical properties



of climate-variables at the surface (Malone, 1951; Peixoto and Oort, 1992; Bryant, 1997; Holland *et al.*, 1999).

The Numerical Weather Prediction (NWP) models and GCMs developed as a result of more readily available computation resources. Using three-dimensional physical models, the entire climate system could be studied in an understandable framework of an integrated system consisting of the atmosphere, oceans, solid earth, cryosphere, and biosphere (Houghton *et al.*, 2001; Watson, 2002).

The principal way to validate the GCMs involves performing a simulation of the present climate and comparing the results with observations. It seems that these models can simulate the free troposphere climate variables quite well (Giorgi and Mearns, 1991; McGregor *et al.*, 1993; von Storch, 1995; McGregor, 1997). On the other hand, climate elements, such as fronts, precipitation and cloudiness are less well modeled. Consequently, all GCMs experiments have shown serious deficiencies in simulating regional or local scale surface features. Another common technical problem is that, much more data are generated than it can be stored and managed. This leads to only a selection of climate variables being archived in those experiments and even then only at a limited temporal resolution, *e.g.*, monthly temporal resolution. Due to these limitations they provide relatively little information on climate change for climate impact studies on regional or local scale.

Since this local scale information is required, a number of techniques have been suggested to infer the likely local characteristics which would accompany the changes seen in the large scale of a GCM simulations (Klein, 1982; Kim *et al.*, 1984; Wilks, 1989; Wilks, 1992; Wigley *et al.*, 1990; Karl *et al.*, 1990; Giorgi and Mearns, 1991; Cubasch *et al.*, 1992; Zorita *et al.*, 1992; McGregor *et al.*, 1993; von Storch *et al.*, 1993; Gyalistras *et al.*, 1994; von Storch, 1995, Matyasovsky *et al.*, 1994; Noguer, 1994; Bardossy, 1994; Houghton *et al.*, 1996; Hewitson and Crane, 1996; Cubasch *et al.*, 1996; McGregor, 1997; Schubert and Henderson-Sellers, 1997; Kidson and Thomson, 1998; Conway and Jones, 1998; Heimann and Sept, 2000; Smith, 1999; Sailor and Li, 1999; Solman and Nunez, 1999; Murphy, 1999; Murphy, 2000; Landman and Tennant, 2000; Wilby and Wigley, 2000; Houghton *et al.*, 2001; Stein *et al.*, 2001; Geerts, 2003; Tatl *et al.*, 2004; Tatl *et al.*, 2005).

'Large scale' is defined as anything from the grid-scale of a GCM to its skillful scale (to be defined from comparisons to observed climate) or even to the global scale (Rummukainen, 1997). One way to classify downscaling techniques is to distinguish between five types. These are:

1. Conventional methods
2. Statistical downscaling
3. Stochastic methods
4. Dynamical downscaling
5. Composite methods

Except 'statistical downscaling', the other four types are not considered in this study. The concepts of statistical downscaling are presented in Chapter 3. The aim of downscaling is to provide local scale climate data. The same goal is also addressed by purely empirical methods, which work via locating and employing analogues from the past. The conventional and composite techniques superimpose GCM-derived changes on observed local time series. In statistical downscaling, large scale free troposphere variables are used as predictors. The predictands are then the desired surface or near-surface climate variables such as surface temperature and precipitation series.

Constructing and validating statistical downscaling needs a source of local scale time series in practice, conventional, composite methods and statistical downscaling in practice work on spatial downscaling; whereas stochastic approaches include weather generators, (e.g., Richardson, 1981; Wilks, 1992) which aim at high temporal resolution and most often focus on hydrological applications.

On the other hand, dynamical downscaling involves running either a high-resolution limited area model with GCM output as boundary conditions, or performing so called 'time slice experiments'. In the latter, a high resolution AGCM (atmosphere GCM) run is performed, using the results of a coarser-resolution GCM as the initial and lateral boundary conditions. The AGCM is then integrated for a relative short period (Giorgi and Mearns, 1991; McGregor *et al.*, 1993; McGregor, 1997; Fuentes and Heimann, 2000).

A major limitation of GCMs is that coupling between atmosphere's processes and surface conditions is limited. In particular, the coupling between biosphere (forests) and the atmosphere has been very poor. This is an important limitation as this coupling is known to be significant, *e.g.*, through precipitation, surface temperature, heat fluxes, albedo and other processes. For instance, hydrologists ask for daily precipitation data with a spatial resolution corresponding to catchments. The ecologists who are studying the dynamics and responses of forests in mountain areas need knowledge of monthly mean precipitation and temperature with a resolution of a few kilometers. Consequently, the main problem of the spatial-scale difference between climate research and climate impact studies is to be bridged by "downscaling" on the side of the climate research and "upscaling" on the side of the climate impact research (Gyalistras *et al.*, 1994; von Storch, 1995).

Analyses of climate variability show that, there is a characteristic spatial scale related to the synoptic features of the atmosphere which is about 1000 km horizontal linear scale or  $10^6 \text{ km}^2$  characteristic area scales, and evidently, a global scale of  $10^4 \text{ km}$  ( $10^8 \text{ km}^2$ ) (von Storch, 1995; Dobrovolski, 2000).

Synoptic processes impose their spatial scales not only on changes with characteristic time of several days but also on month-to-month, year-to-year, and larger climatic changes, because temporal averaging preserves the spatial structure of the processes' features. In other words, if there is a long term anomaly of some climatic parameter, this anomaly has a horizontal size comparable with the characteristic size of cyclones and anticyclones. Spatial correlation of atmospheric, oceanic, and land surface variables are indications of this feature of climatic processes. For example, spatial auto-covariance of monthly anomalies of precipitation and surface temperature are shown in Figure 1.1. This figure shows that correlation radius which is a measure of spatial structure corresponding to an auto-covariance level which equals to 0.368 for monthly anomalies is from 750 km to 1500 km and more. For many climatic parameters spatial correlations become zero at a distance of 2000 km to 2500 km (Dobrovolski, 2000).

The starting point of the study of the impact of large scale climate features on local to regional scales may be casted as Global Climate Scenarios (GCS). Regarding to IPCC (Houghton *et al.*, 2001), the mean values and the standard deviations of the 17

types of climate scenarios are demonstrated in Figure 1.2. The relations between local scale and large scale climate features based on these GCS are the key issues.

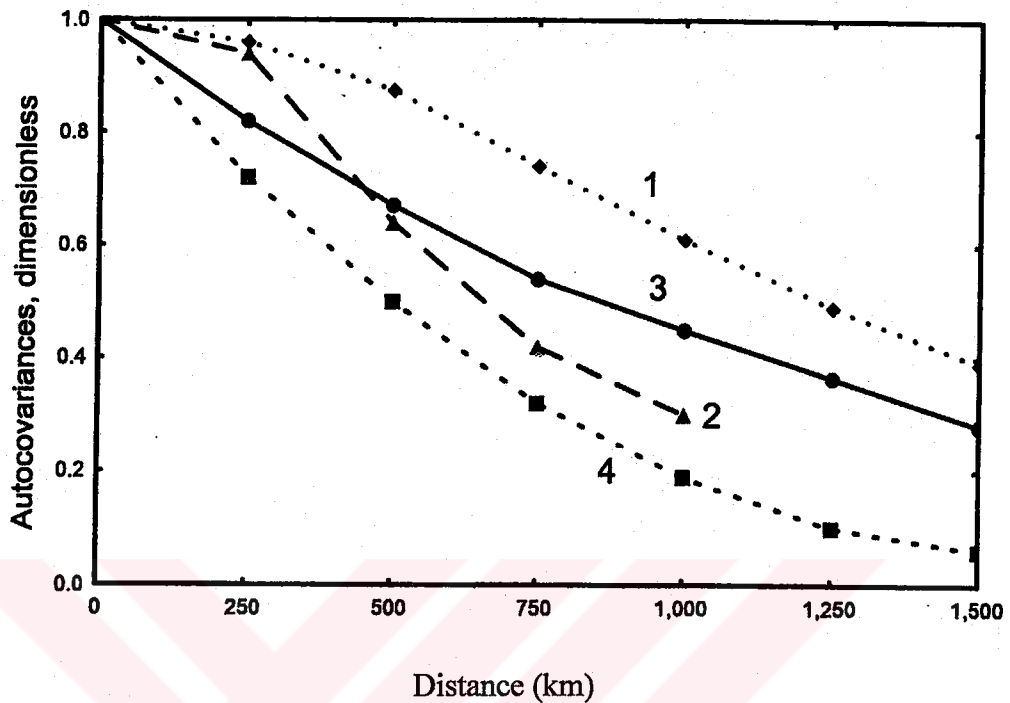


Figure 1.1. Spatial correlation functions of monthly anomalies: 1 precipitation, longitudinal direction, (65°N, January); 2 the same but for July; 3 North Atlantic Sea Surface Temperature (SST), latitudinal direction; 4 North Atlantic SST, longitudinal direction (Dobrovolski, 2000).

However, General Circulation Models (GCMs) satisfactorily describe global fields of atmospheric temperature and air pressure at global, continental, and sub-continental scales, particularly the free atmospheric variables. The major causes of these and other errors of GCMs are related to a lacking description of the specific physical, chemical, or biological mechanisms for components of the climate system. The technical problems may be due to the lack of observational data which are required to construct physical models and to validate their outputs. In the following, some examples of problems of the GCMs are investigated by Houghton *et al.* (1996):

1. Radiative effects of clouds remain an area of difficulty.
2. The large scale dynamics of current ocean models are not completely validated, in part because of a dearth of appropriate observations.
3. Fluxes at the ocean-atmosphere interface have not been yet fully examined.

4. Existing oceanic models use very crude parameterizations of sub-grid scale processes for near-surface and interior mixing and for deep convection.
5. The role of sea ice in climate change is especially uncertain because of poorly known interface feedbacks.

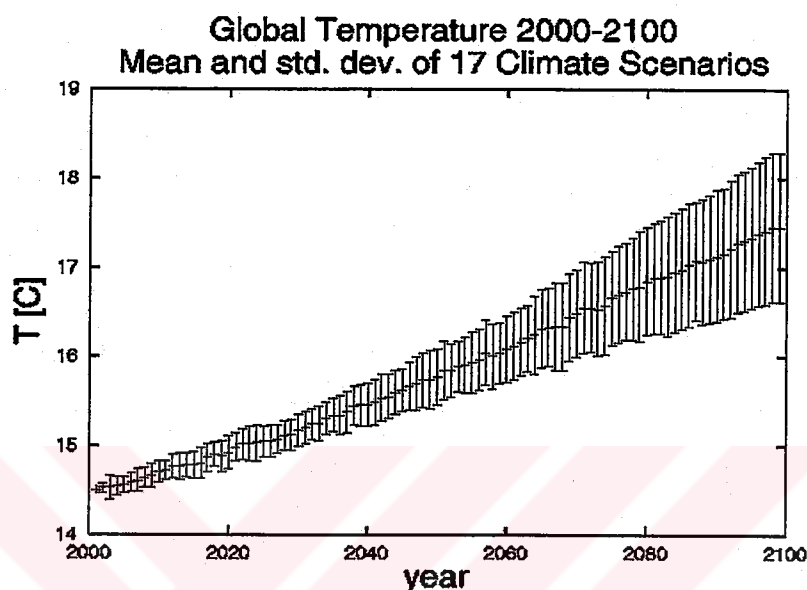


Figure 1.2. The results of the 17 Climate Scenarios from IPCC (Houghton *et al.*, 2001) → <http://ipcc-ddc.cru.uea.ac.uk/>

6. Current models do not satisfactorily simulate climate indices such as the ones related to El Niño -Southern Oscillation (ENSO) process.
7. Solar fluxes at the land surface are currently highlighted as being significantly in error compared to observations due to the inadequate treatment of clouds.
8. There are intrinsic difficulties in parametrizing slope effects to determine runoff in a climate model.
9. Biases and uncertainties in the surface energy balance and radiation and water budgets are a significant source of error in simulations of climate.
10. Models of runoff, in general show great uncertainty in global models, there are no convincing treatments of the scaling of the responsible processes over the many orders of magnitude involved, and in high latitudes of the effects of frozen soils.

11. How the integrated stomatal resistance for global vegetation might change with changing climate and CO<sub>2</sub> concentrations is largely unknown.

It was confirmed that there are considerable differences between the local scale processes and large scale processes, especially in the near-surface layer. Hence, the GCMs may be of use to understanding and describe the large scale features of climate. However, for water management, farming, energy-planning, land-use management, and for the others we need to assess the local-features such as precipitation and surface temperature in future. Since the free troposphere is more spatially and temporally homogeneous than the earth's surface, the logic behind downscaling using large scale processes from free troposphere variables is understandable.

The problem statement of the estimation of local climate variables at specific stations or within relatively small areas from large scale GCMs-generated fields is called downscaling.

There are three possible approaches to study climatic variability. The first is deterministic, *i.e.* it is postulated that the value of some variable of the climate system is explicitly determined by the value of this and other variables of the climate system, also by relevant parameters outside the climate system, at previous moment's time. From this point of view, to obtain enough detailed information to be able to give an exact forecast of the variable under consideration. In other words, if we have detailed information on the state of the atmosphere and oceanic surface for the past few weeks, we can deterministically describe and explain the present and future next few days' behavior of the atmosphere.

If we consider the climatic scale be one month or more, then the deterministic approach fails to determine the variables plainly, particular near-surface and/or in local scale. Hence, the value of the climatic variable might be considered as a random variable (Dobrovolski, 2000). So it can be determined in a probabilistic space frame. This second approach is called stochastic method.

On the other hand, GCMs might be able to determine temporally averaged free atmospheric variables (upper air circulation variables) at large scale. Bridging of the deterministic and stochastic methods leads to a new approach called as 'statistical downscaling'. In this study, this third approach named as 'grey approach' is used.



Since, the deterministic way might be seen as white-box (all the constants have physical meaning) and stochastic frame could be considered as black-box (the constants have no physical meaning), the mixture of a stochastic and deterministic approach can be called grey-box (white + black = grey).

The constraints on both the deterministic and stochastic approaches now should be valid in the third approach (statistical downscaling). Therefore, particularly in implementing the new techniques based on probability theory, such as probabilistic framework of multivariate techniques are employed.

In this study, the problem of statistical linkages between the local scale climate features over Turkey and large-scale climate features are investigated by the new proposed downscaling models based on multivariate techniques and recurrent neural networks (RNN).

The dissertation is organized as follows: In the following chapter, the data and mathematical-statistical methods are introduced and discussed. The properties of the conventional statistical downscaling types and proposed downscaling models are presented in Chapter 3, followed by a description of the near-surface air temperature and precipitation of Turkey with the view of their major characteristics in Chapter 4. The applications of the new proposed downscaling strategies are given in Chapter 5, and the study is concluded with an outlook and a summary in Chapter 6.

## 2. DATA AND METHODS

### 2.1 Data and Data Preparation

Several data sets are used in this study: gridded large scale National Centers for Environmental Prediction-National Center of Atmospheric Research (NCEP-NCAR) reanalysis data (Figure 2.1) (Kalnay *et al.*, 1996) and the data from Turkish stations.

The large scale data is used as the predictor, and the station data as the predictand. The choice of the predictands is ultimately governed by the need of the end-users, but limited by the variability of the local time series. In the study, the monthly temperature series and monthly total precipitation series are chosen, the primary climate variables, which can not be simulated as perfectly as by the GCMs. The fact that the GCMs perform best on large scales and on time-averaged basis (monthly or seasonal means), and in addition; the GCMs produce a better simulation of the free troposphere variables. Therefore the predictors are selected based upon the large scale upper air circulation and mean sea level pressure. The employed predictors are divided into the precipitation-related and temperature-related predictors. The precipitation-related large scale predictors are:

- 1) 500 hPa geopotential height
- 2) 700 hPa geopotential height
- 3) 500-1000 hPa thickness
- 4) Mean sea level pressure
- 5) 500 hPa vertical pressure-velocity

The period of the precipitation data over Turkey covers the entire months of the 1961-1998 interval. The geographical distribution of the selected 31 Turkish stations for the monthly total precipitation is given in Figure 2.2a. The reasons of why these predictors are selected are explained in the “application sections” of the study. The temperature-related predictors are:



- 1) 500 hPa geopotential height
- 2) 500-1000 hPa geopotential thickness

The record length of the both large scale temperature-related predictors and station temperature series is the 1951-1998 interval. Figure 2.2b is the geographical distribution of the 62 Turkish stations for monthly near-surface air temperature data sets (namely maximum, mean and minimum) respectively.

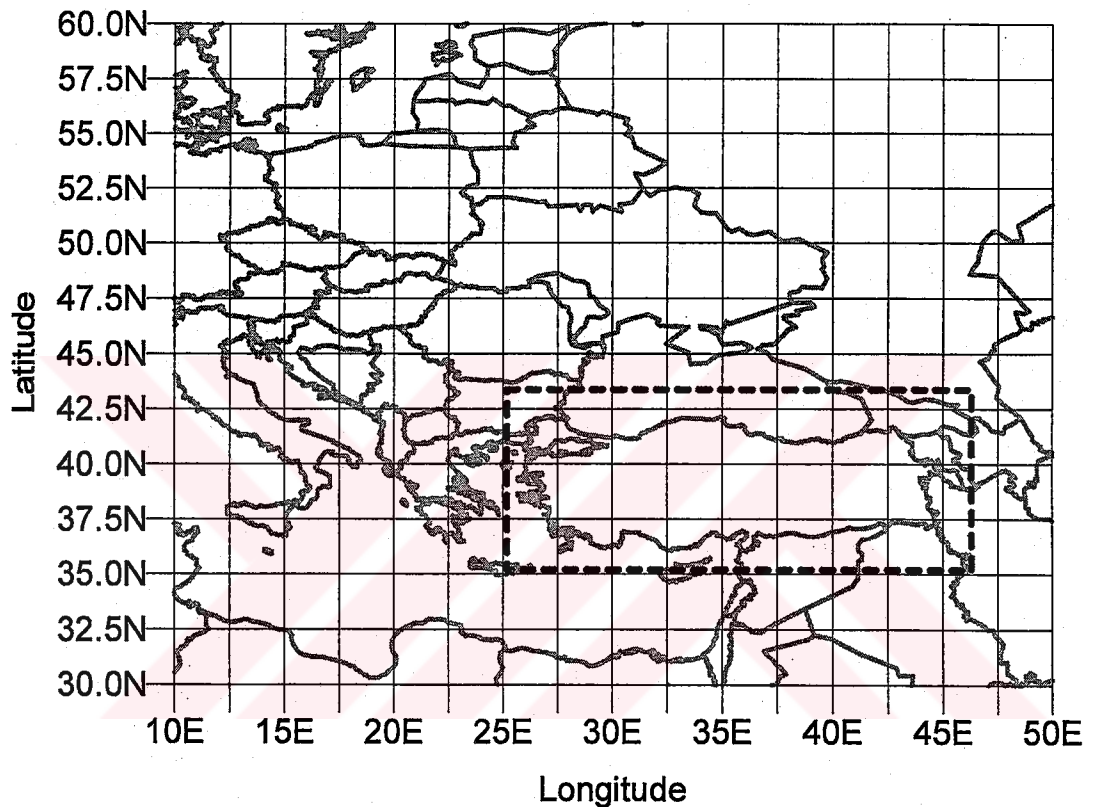
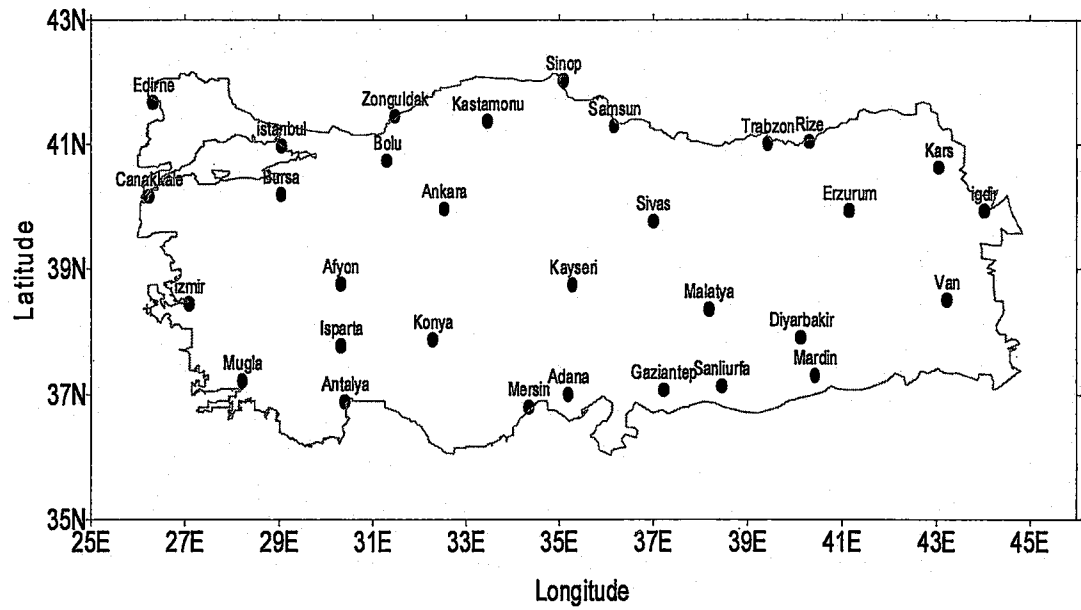
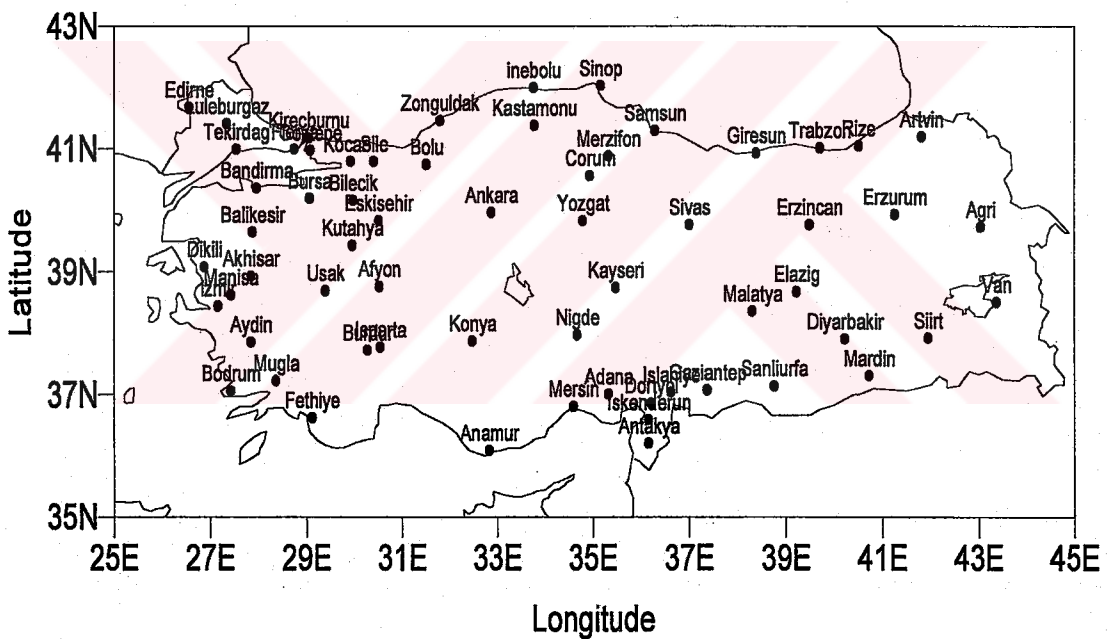


Figure 2.1 Distribution of the 221 grid points (2.5° longitude by 2.5° latitude) for NCAR/NCEP re-analysis data field. The area of Turkey enclosed by the rectangular box is shown in additional detail in Figure 2.2.

The chosen data records and stations are basis upon the quality of the data and the meaningful statistical relationships between the large scale and station scale variables in addition the performance of the typical GCMs, and also the availability of data.



(a)



(b)

Figure 2.2. Geographical distribution for Turkey stations: (a) The stations of the monthly total precipitation; (b) the stations of the near-surface air temperatures.

## 2.2 Methods

### 2.2.1 Singular Value Decomposition (SVD)

In this study, Singular Value Decomposition (SVD) of matrices which is a key method in the course of applying Principal Components Analysis (PCA), Canonical Correlation Analysis (CCA) and Redundancy Analysis (RA) which are the major

techniques are applied. Therefore, the concepts of SVD are discussed in the following, and the module of F-Language (FORTRAN) source code of SVD is also given in Appendix.

Assume  $\mathbf{X}$  is a  $p \times q$  ( $q \leq p$ ) matrix and it can be decomposed as follows (Wilkinson and Reinsch, 1971):

$$\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T \quad (2.1)$$

where  $T$  indicates the transpose of the relevant matrix. Here,  $\mathbf{U}$  is an  $p \times p$  orthogonal matrix,  $\mathbf{V}$  is an  $q \times q$  orthogonal matrix, and  $\mathbf{D} = (\lambda_1 \dots \lambda_q)$  is a diagonal matrix with elements  $\lambda_{ij}$ , where singular values are  $\lambda_{ii} = \lambda_i$  and  $\lambda_{ij} = 0$  for  $i \neq j$  respectively.

If  $\mathbf{X}$  is a symmetric matrix then the SVD is equivalent to a spectral decomposition operation. When  $\mathbf{U} = \mathbf{V}$ , this condition indicates the orthonormal (orthogonal and normal) matrix of the eigenvectors of  $\mathbf{X}$ .

### 2.2.2 Principal Component Analysis (PCA), Factor Analysis (FA) and Independent Component Analysis (ICA)

Principal component analysis and closely related to factor analysis techniques are widely applied in meteorology or climatology literature (Kidson 1975; Diaz and Fullbright, 1981; North 1984; Preisendorfer, 1998). This technique can be summarized as follows: PCA is a multivariate technique classified as a non-dependent variable method. The goal in PCA is to create a set of orthogonal variables (components) from some given data by creating linear combinations of the original data that maximizes the variance of each new variable. The goal is accomplished by rotating either the centered data or centered and scaled data, so that the axis along which the variance is maximum coincides with the axis of the first principal component (PC). The next step is to rotate the data orthogonally to the first component's axis so as to maximize the remaining variance in the second principal component. This process is repeated until a zero eigenvalue is encountered or the number of components equals the number of variables in the original data. The rotation is accomplished by using either the covariance matrix or the correlation matrix of the original data. The choice of which to use results in different solutions.

If all the variables are not measured on the same scale it is common to remove the units by using the correlation matrix to perform the rotation.

Assume the first and second order statistics of a multivariate  $\mathbf{X}$  is known or can be estimated from the sample. In PCA transform, the variables of  $\mathbf{X}$  are centered by subtracting their means, and then the covariance matrix of  $\mathbf{X}$  is defined as expected value of minor products of its variables. In the study, all matrices are organized such that rows represent simultaneous observations and columns observed variables at different sites.

$$\mathbf{S}_{\mathbf{XX}} = E(\mathbf{X}^T \mathbf{X}) \quad (2.2)$$

One has to keep in mind that PCs are not invariant under scaling. Without loss of the generality, an orthogonal decomposition of  $\mathbf{S}_{\mathbf{XX}}$  is given as follows.

$$\mathbf{S}_{\mathbf{XX}} = \mathbf{E}_{\mathbf{X}} \mathbf{D}_{\mathbf{X}} \mathbf{E}_{\mathbf{X}}^T \quad (2.3)$$

where  $\mathbf{E}_{\mathbf{X}}$  is a matrix of the orthonormal eigenvectors of  $\mathbf{S}_{\mathbf{XX}}$  and  $\mathbf{D}_{\mathbf{X}} = \text{diag}(\lambda_1, \dots, \lambda_k)$  is the diagonal matrix represents the eigenvalues of  $\mathbf{S}_{\mathbf{XX}}$  in a decreasing magnitude order, respectively. The principal components are then computed by

$$\mathbf{V}_{\mathbf{X}} = \mathbf{X} \mathbf{E}_{\mathbf{X}} \quad (2.4)$$

where  $\mathbf{V}_{\mathbf{X}}$  illustrates the PCs and the reconstruction of  $\mathbf{X}$  follows:

$$\mathbf{X} = \mathbf{V}_{\mathbf{X}} \mathbf{E}_{\mathbf{X}}^T \quad (2.5)$$

The PCA model was discussed above as a distribution free method with no underlying statistical model, but in the case of factor analysis, the multivariate  $\mathbf{X}$  is reconstructed as in the following:

$$\mathbf{X} = \mathbf{F}_{\mathbf{X}} \mathbf{A}_{\mathbf{X}} + \mathbf{G}_{\mathbf{X}} \quad (2.6)$$

where  $\mathbf{F}_{\mathbf{X}}$  and  $\mathbf{G}_{\mathbf{X}}$  are called common factors (CFs) (or latent variables) and white noise components (or specific components), respectively.  $\mathbf{A}_{\mathbf{X}}$  is called factor loadings defined as:

$$\mathbf{S}_{xx} = \mathbf{A}_x \mathbf{A}_x^T \quad (2.7)$$

There are no statistical constraints on PCs, but CFs must be Gaussian in a FA model. Furthermore, PCs are easily calculated by a unique way, whereas there is not a unique method for CFs, but a possible way according to Reymont and Jöreskog (1993) by the expression:

$$\mathbf{F}_x = \mathbf{X} \mathbf{S}_{xx}^{-1} \mathbf{A}_x^T \quad (2.8)$$

The CFs may be rotated for easier interpretation of the patterns (Richman, 1985) by a method such as varimax (Kaiser, 1959).

On the other hand, Independent Component Analysis (ICA) may be seen as a special kind of rotation of CFs (or PCs) but using high order statistics rather than the second order statistics. If probability density of concerning data set is Gaussian then FA and ICA are identical; because the high order statistics of Gaussian variables are all zero (assume the variables have zero mean and variance one). ICA has been widely used in data analysis and decomposition, particularly, in signal processing research (Jutten and Herault, 1991; Comon, 1994; Everson and Roberts, 1999; Hyvarinen *et al.*, 2001). ICA, typically aims to solve the blind source separation problem in which a set of unknown sources are mixed in some way to form the data. An ICA model assumes that the multivariate data set  $\mathbf{X}$  is a mixing of the Independent Components (ICs) as:

$$\mathbf{X} = \mathbf{S} \mathbf{A} \quad (2.9)$$

or, by the terms of PCA:

$$\mathbf{S} = \mathbf{X} \mathbf{W} = \mathbf{V}_x \mathbf{E}_x^T \mathbf{W} \quad (2.10)$$

where  $\mathbf{A}$  is called the mixing matrix and  $\mathbf{W} = \mathbf{A}^{-1}$ , and  $\mathbf{S}$  (like PCs or CFs) represents ICs, respectively. The starting point for ICA is the very simple assumption that the all ICs are statistically independent. However, the basic model of ICA does not assume that the probability distribution of data is known. After determining the matrix  $\mathbf{A}$ , its inverse  $\mathbf{W}$  is then computed to obtain the ICs by Equation (2.10) from PCs.

The restriction in a PCA (or FA) is that the entire PCs are mutually uncorrelated, but the restriction in an ICA is that the all ICs are statistically independent. Uncorrelatedness is weaker form of independence. That is, two random variables are uncorrelated if their covariance is zero. On the other hand, if the variables are independent, they are uncorrelated, which follows that uncorrelatedness does not imply independence (Comon, 1994). Two random variables  $y_1$  and  $y_2$  are said to be mutually independent, if given two arbitrary functions,  $h_1(y_1)$  and  $h_2(y_2)$  satisfy the following condition (Hyvarinen *et al.*, 2001):

$$E\{h_1(y_1)h_2(y_2)\} = E\{h_1(y_1)\}E\{h_2(y_2)\} \quad (2.11)$$

where  $E$  represents the expected value. According to the definition of statistical independency, ICA has a stronger definition than the uncorrelatedness concerning PCA or FA. Unlike PCA or FA, ICA uses high order statistics that can separate the data into the true sources. In this study, the ICs are computed via a fixed-point algorithm which is called Fast Independent Component Analysis (FASTICA) introduced by Hyvarinen and Oja (1997). The algorithm of FASTICA is based on maximizing the absolute value of kurtosis of the variable by using ‘information measure (or negentropy)’:

$$J(y) = \frac{1}{12} E\{y^3\}^2 + \frac{1}{48} kurt(y)^2 \quad (2.12)$$

where  $J(y) \in [0,1]$  indicates negentropy, and kurtosis of  $y$  is illustrated by  $kurt(y)$  defined as:

$$kurt(y) = E\{y^4\} - 3(E\{y^2\})^2 \quad (2.13)$$

Since the negentropy of Gaussian variables among equivalent variances is zero, it can be used as an indicator to measure the degree of normality in the data series.

However, there are some basic ambiguities in the ICA model:

1. The variances of ICs can not be determined, whereas the variances are explained by eigenvalues of PCs (or CFs) in a PCA (or FA) approach.
2. The order of ICs can not be determined.

In the study, the ICA process is done by the Matlab Toolbox called FASTICA which is introduced by Hyvarinen and Oja (1997).

### 2.2.3 Canonical Correlation Analysis (CCA)

The other central multivariate statistical method applied in this study, is the canonical correlation analysis (Glahn 1968; Mardia *et al.*, 1979; van de Geer, 1984; Jackson 1991; Rencher, 1992; Chen and Chang 1994; Chen *et al.*, 1994; Rencher, 1995; Wilks, 1995; TenBerge 1988; Rencher, 1998). Hotelling (1936) proposed CCA as a model to relate two sets of variables. He derived linear combinations of the  $\mathbf{X}$  variables (in this study the large scale variables) and the  $\mathbf{Y}$  variables (in this study the local scale variables) that were maximally correlated, subject to the constraint that each derived variate is uncorrelated with other variates. Denote one vector of  $\mathbf{X}$  by  $\mathbf{x}$ , and one vector of  $\mathbf{Y}$  by  $\mathbf{y}$ ; CCA can reduce linear components of  $\mathbf{x}$  and  $\mathbf{y}$ ,  $\mathbf{t}_i$  and  $\mathbf{u}_i$  as:

$$\mathbf{t}_i = \mathbf{w}_i^T \mathbf{x}, \quad \mathbf{u}_i = \mathbf{v}_i^T \mathbf{y} \quad (2.14)$$

choosing  $\mathbf{w}_i$  and  $\mathbf{v}_i$  such as to maximize the correlation between  $\mathbf{t}_i$  and  $\mathbf{u}_i$  subject to the following constraints:

$$\mathbf{w}_i^T \mathbf{C}_{XX} \mathbf{w}_j = 0, \quad \mathbf{v}_i^T \mathbf{C}_{YY} \mathbf{v}_j = 0 \quad \text{for } \forall i \neq j \quad (2.15)$$

$$\mathbf{w}_i^T \mathbf{C}_{XX} \mathbf{w}_i = 1, \quad \mathbf{v}_i^T \mathbf{C}_{YY} \mathbf{v}_i = 1 \quad (2.16)$$

where  $\mathbf{C}_{XX}$  and  $\mathbf{C}_{YY}$  denote the covariance matrices, respectively. Canonical correlation vectors (or weights)  $\mathbf{w}_i$  and  $\mathbf{v}_i$  can be obtained by orthogonal decomposition procedure of the related matrices, where  $\mathbf{w}_i$  and  $\mathbf{v}_i$  are the eigenvectors of  $\mathbf{C}_{XX}^{-1} \mathbf{C}_{XY} \mathbf{C}_{YY}^{-1} \mathbf{C}_{YX}$  and  $\mathbf{C}_{YY}^{-1} \mathbf{C}_{YX} \mathbf{C}_{XX}^{-1} \mathbf{C}_{XY}$  respectively. The square-roots of the eigenvalues of these two matrices are equal and represent canonical correlation coefficients between the pairs of  $\mathbf{t}_i$  and  $\mathbf{u}_i$ . CCA has a property of biorthogonality satisfying the diagonalizability of  $\mathbf{C}_{XY}$ :

$$\mathbf{W}^T \mathbf{C}_{XY} \mathbf{V} = \mathbf{D} \quad (2.17)$$



where  $\mathbf{D}$  is the diagonal matrix representing the squares of canonical correlation coefficients. However, a CCA approach is symmetrical and can not recognize which of the components is the predictor and predictand, that is:

$$\mathbf{V}^T \mathbf{C}_{\mathbf{YX}} \mathbf{W} = \mathbf{D} \quad (2.18)$$

Furthermore, the structure matrices for  $\mathbf{X}$  and  $\mathbf{Y}$  can be defined as the correlation matrices between variables and canonical correlation variates (CCVs):

$$\left. \begin{array}{l} \mathbf{C}_{\mathbf{XX}} \mathbf{W}, \text{ for } \mathbf{X} \\ \mathbf{C}_{\mathbf{YY}} \mathbf{V}, \text{ for } \mathbf{Y} \end{array} \right\} \quad (2.19)$$

where the maps of  $\mathbf{C}_{\mathbf{XX}} \mathbf{W}$  and  $\mathbf{C}_{\mathbf{YY}} \mathbf{V}$  denote canonical correlation patterns (structure matrices), respectively.

#### 2.2.4 Redundancy Analysis (RA)

Canonical correlation analysis finds variates that are correlated, but it is not so practical in the prediction frame since they do not explain enough covariance. A statistical prediction model must satisfy asymmetric features, such as to predict local scale processes from large scale processes (downscaling) may not be invertible to predict large scale features from local scale features (upscaling), particularly in climate processes. Van den Wollenberg (1977) devised a linear model of RA as an alternative CCA that avoids this problem. The weights for redundancy variates (RVs),  $\mathbf{w}_i$  are determined by solution of the following equation:

$$\left[ \mathbf{C}_{\mathbf{XY}} \mathbf{C}_{\mathbf{YX}} - b_i \mathbf{C}_{\mathbf{XX}} \right] \mathbf{w}_i = 0 \quad (2.20)$$

where  $b_i$  is the covariance explained by the  $i$ -th variate pair. In RA technique, the redundancy weights for  $\mathbf{X}$  and  $\mathbf{Y}$  can be obtained as (Gower, 1975; Tyler 1982):

$$\mathbf{v}_i = b_i^{-1/2} \mathbf{C}_{\mathbf{YX}} \mathbf{w}_i \quad (2.21)$$

The same redundancy weights can also be obtained by using SVD:

$$\left. \begin{array}{l} \mathbf{C}_{\mathbf{XX}}^{-1/2} \mathbf{C}_{\mathbf{XY}} = \mathbf{H} \mathbf{E} \mathbf{V}^T \\ \mathbf{W} = \mathbf{C}_{\mathbf{XX}}^{-1/2} \mathbf{H} \end{array} \right\} \quad (2.22)$$



where  $\mathbf{H}$  is the left-side eigenvector matrix for  $\mathbf{C}_{\mathbf{XX}}^{-1/2}\mathbf{C}_{\mathbf{XY}}$ ,  $\mathbf{W}$  is the matrix of variates for the  $\mathbf{X}$  variables and  $\mathbf{V}$  is the matrix of the variates for the  $\mathbf{Y}$  variables respectively, and  $\mathbf{E}$  is the diagonal matrix whose elements are square roots of  $b_i$ .

However, obtaining RVs from asymmetric matrices is not a preferred way; Tyler (1982) proposed another method for RA from symmetric matrices. First, the orthonormal eigenvectors of  $\mathbf{C}_{\mathbf{YX}}\mathbf{C}_{\mathbf{XX}}^{-1}\mathbf{C}_{\mathbf{XY}}$ ,  $\mathbf{U}$  and the corresponding matrix of  $\mathbf{R}$  roots are obtained, and then the weight matrix for RVs can be written as:

$$\mathbf{W} = \mathbf{C}_{\mathbf{XX}}^{-1}\mathbf{C}_{\mathbf{XY}}\mathbf{U}\mathbf{R}^{-1/2} \quad (2.23)$$

where  $\mathbf{Y}$  and  $\mathbf{X}$  represent the local scale predictands and large scale predictors, respectively.

### 2.2.5 Sampson correlation ratio ( $R_S$ )

The first step in downscaling is to determine the appropriate predictors. There are many statistical methods according to the scientific literature, two such possibilities are CCA and stepwise regression (Noguer 1994; von Storch and Zwiers 1999). But in this study, Sampson (1984) correlation ratio is used to recognize the predictors and predictands (*e.g.* Tatlı *et al.*, 2004, 2005). It is defined as:

$$R_S = \sqrt{\frac{\text{Tr}(\mathbf{C}_{\mathbf{YX}}\mathbf{C}_{\mathbf{XX}}^{-1}\mathbf{C}_{\mathbf{XY}})}{\text{Tr}(\mathbf{C}_{\mathbf{YY}})}} \quad (2.24)$$

where  $\mathbf{C}_{\mathbf{XY}}$ ,  $\mathbf{C}_{\mathbf{XX}}$  and  $\mathbf{C}_{\mathbf{YY}}$  indicate multivariate cross-covariance and covariance matrices and  $\text{Tr}$  denotes the trace of the related matrices, respectively. Sampson correlation ratio,  $R_S \in [0,1]$ , represents the correlation between two matrices and it is not symmetric in particular, whereas Pearson moment correlation is symmetric representing the correlation between two vectors. In this study, the contour map based on data sets between the set of time series of gridded data and the multiple time series of observations is named Sampson correlation pattern ( $R_S$  pattern).

### 2.2.6 Singular Spectrum Analysis (SSA)

Before constructing a downscaling model, a statistical preprocessing is required in order to determine the significant large scale variables in a compressed manner (such

as principal components or independent components (Tatlı *et al.*, 2004)). If we plot the time series of the monthly temperature series of station data series, and, for instance, of 500 hPa geopotential heights and 500-1000 hPa geopotential thicknesses, then the irregular noise components emerge. In order to remove of such noise in data sets, Fourier analysis (Richardson 1981) is a possible way, but Fourier analysis assumes the stationarity constraint in the process. Singular spectrum analysis is another possible way based on eigenvalue techniques without satisfying stationarity constraints. The SSA approach or temporal principal component analysis is well-known in climate studies (*e.g.* Ghil and Vautard 1991; Vautard *et al.*, 1992; Green *et al.*, 1993; Elsner and Tsonis 1994, Elsner and Tsonis 1996; Schlesinger and Ramankutty 1994; Allen and Smith 1996).

To point out the method, let  $\mathbf{x}(t)$  be a univariate time series that can be decomposed into three components as:

$$\mathbf{x}(t) = \mathbf{x}_d(t) + \mathbf{x}_s(t) + \mathbf{x}_e(t) \quad (2.25)$$

where  $\mathbf{x}_d(t)$  is the deterministic component (trend and/or periodic cycles) being predictable by mathematical models (*e.g.* GCMs);  $\mathbf{x}_s(t)$  represents the statistical component which can be estimated by statistical models (*e.g.* regression, AR, ARMA); and  $\mathbf{x}_e(t)$  indicates the noise component being not predictable by both mathematical and statistical models, respectively. However, a problem might arise if the process is constituted by the product of the deterministic and statistical components in addition to the noise component as:

$$\mathbf{x}(t) = \mathbf{x}_d(t) \times \mathbf{x}_s(t) + \mathbf{x}_e(t) \quad (2.26)$$

The constraint of our approach is based on the assumption of Equation (2.25) the analysis of Equation (2.26) is beyond the scope of this study.

SSA consists of the diagonalisation of the lagged auto-covariance (or auto-correlation) matrix, like in PCA, the eigenvectors represent patterns of temporal behavior, and PCs are significant embedding dimensions (or characteristic series). When two eigenvalues of the lagged auto-covariance (or auto-correlation) matrix are nearly equal and their corresponding eigenvectors are 90° phase-shift they represent oscillation. If there is a trend, then the first eigenvalue, the biggest, represents the

trend, and the following nearly equal eigenvalue pairs represent periodic cycles. In this study, the reconstruction of the univariate time series is mentioned as a low-pass filter. Table 2.1 represents the concept of SSA.

Table 2.1. The look up table of SSA of the maximum near-surface air temperature series for some Turkey stations

Cycles	Eigenvalue	Sites and Captured Variance (%)					
Meaning	No	Mersin	Antalya	Adana	Gaziantep	Şanlıurfa	Muğla
Periodic Cycles	1	38.67	46.883	42.701	47.131	47.634	46.644
	2	38.56	46.708	42.545	46.949	47.457	46.489
Non-Periodic Components	3	5.046	0.157	3.037	0.621	0.551	0.355
	4	5.013	0.154	3.013	0.617	0.549	0.355
	5	0.301	0.131	0.274	0.132	0.122	0.127
	6	0.284	0.129	0.274	0.116	0.114	0.125
	7	0.268	0.112	0.184	0.111	0.102	0.119
	8	0.245	0.112	0.181	0.109	0.097	0.118
	9	0.245	0.111	0.160	0.107	0.089	0.113
	10	0.226	0.110	0.158	0.095	0.069	0.112

Due to the huge volume size of the predictors, any model structure will be so complex and the parsimonies properties of the model may disappear. Furthermore, the noise may also have an effect the model parameters while constructing an economical downscaling model. In other words, the model tries to estimate the predictable components, but on the other hand it may estimate the non-predictable components due to the noise. To solve this problem, both the SSA and PCA approaches are proposed to remove these non-predictable components. After SSA, the residuals are decomposed via PCA and then the following step is the reconstruction of the data series with the significant PCs. Consequently, the resulting time series are now noise-free according to the applied methods.

### 2.2.7 Neural networks

Neural networks, which are based on biological structures, were investigated in the 1940s and in the 1950s. The well-known neural networks were modeled by

McCulloch and Pitts, and *perceptron* proposed by Rosenblatt (Haykin, 1999), on the other hand. The negative results occurred during the investigation of networks and/or perceptron caused a stagnation of neuro-computing in the sixties and seventies. Later, the rediscovery of the back-propagation algorithm (Rumelhart and McClelland, 1986), which was developed in beginning of the seventies by Hopfield (Haykin, 1999) entailed intensive research in various disciplines up to the present.

A formal neuron  $N$  is a quadruple of  $(\mathbf{x}, \mathbf{w}, \psi, \phi)$  which consists of a vector of inputs  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ , a vector of weights  $\mathbf{w} = (w_1, w_2, \dots, w_n)^T$ , an integration function  $\psi(\mathbf{x}, \mathbf{w})$ , and an activation function  $\phi(\psi)$ . An important point of this definition is the locality of the information processed by the formal neurons. Each neuron gets information only by its incoming connections. Therefore, it is possible for a formal neuron to possess local memory, *i.e.*, there is no global information interchange. This allows parallel information processing.

Neurons within a neural network are normally arranged in layers (Figure 2.3). If the input connections of a layer gain information only from the output of the proceeding layer the net is called *feed forward* network.

Several alternative functions have been used in the processing elements in a neural network, but the more popular training procedures is the back-propagation (Rumelhart and McClelland, 1986) which requires  $\phi$  being a continuously differentiable and bounded scalar function (known as basis function). In this regard, sigmoid neurons are often selected for network training.

The matching functions are continuous and characterized by the following limits.

$$\lim_{\xi \rightarrow \infty} \phi(\xi) = 1, \quad \lim_{\xi \rightarrow -\infty} \phi(\xi) = -1 \quad (2.27)$$

The functions,

$$\phi(\xi) = \tanh \xi, \quad \phi(\xi) = \frac{1}{1 + e^{-c\xi}} \quad (2.28)$$

are examples of such sigmoid functions. The back-propagation algorithm is based upon the method of steepest descent. The goal is minimizing the following error function.

$$R = \frac{1}{2}(y - \tilde{y})^2 \quad (2.29)$$

where  $y$  is the observed value, and  $\tilde{y}$  is the estimated value from the network. For example, for training the element of  $w_i$  with respect to minimizing Equation (2.29); we assume that the model is represented by  $\tilde{y} = \tilde{y}(a_j(w_i))$  then minimizing procedure follows the following steps.

$$1. \quad \frac{\partial R(w_i)}{\partial w_i} = \frac{\partial R(w_i)}{\partial \tilde{y}} \frac{\partial \tilde{y}}{\partial a} \frac{\partial a_j}{\partial w_i} \quad (2.30)$$

$$2. \quad w_i^{\text{new}} = w_i^{\text{old}} - \eta \frac{\partial R(w_i)^{\text{old}}}{\partial w_i^{\text{old}}} + \delta (w_i^{\text{old}} - w_i^{\text{old}-1}) \quad (2.31)$$

where  $\eta \in (0,1)$  and  $\delta \in [0,1]$  represent 'the learning rate' and 'the momentum factor', respectively. At the first step of the training,  $\delta$  is taken as zero. Additionally, at the initial step, the elements of  $w$  are fed randomly. The detailed information about neural networks may be found in (Hornik, 1991; Becker and Hinton, 1992; Kohonen, 1995; Haykin, 1999; Haykin, 2001).

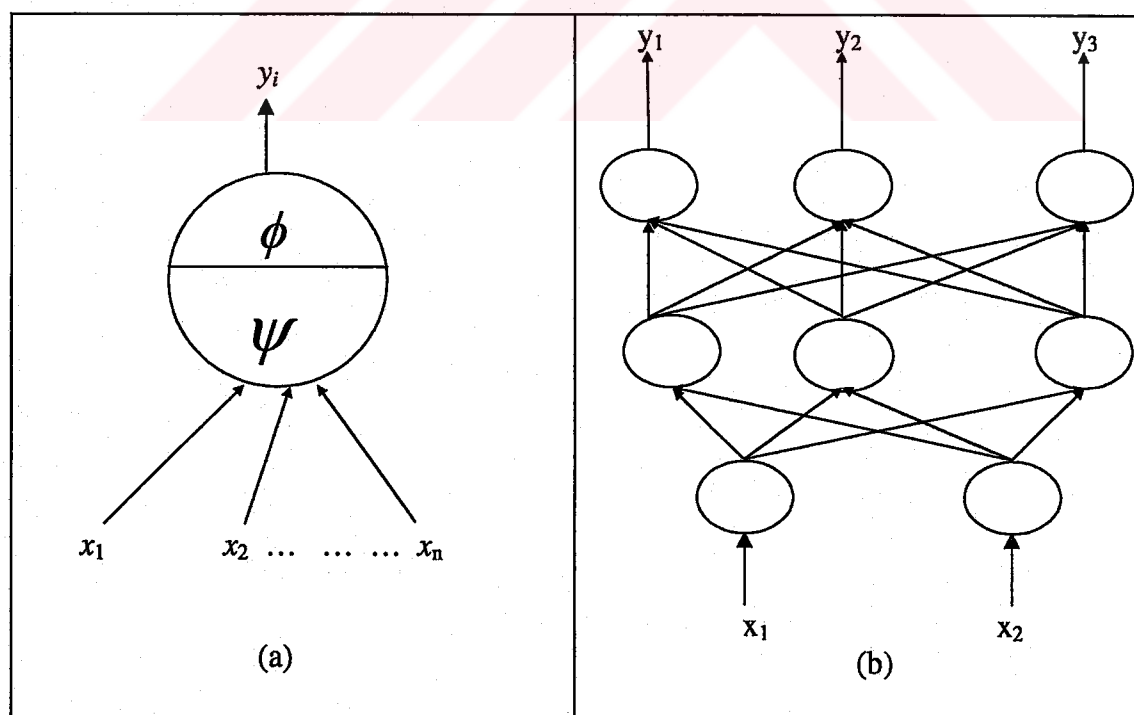


Figure 2.3. Structure of a formal neuron (a), and a multilayer network (b)

Also, there are some structurally different neural networks which may be found in (Ito, 1991; Park and Sandberg, 1991; Pati and Krishnaprasad, 1993).

Moreover, if the input connections of a network layer gain information both from the outputs of the proceeding layer and intermediated layer and the input layer then the net is so called Recurrent Neural Networks (RNN), where Figure 2.4 shows such a network type of Jordan (1986) and Figure 2.5 is the Elman (1990) RNN.

Detailed information about RNN can be found in (Jordan, 1986), Elman (1990), Robinson and Fallside (1991), Puskorius and Feldkamp (1994), Connor *et al.* (1994), Pearlmutter (1995), Haykin (1999, 2001), and Hochreiter *et al.* (2001).

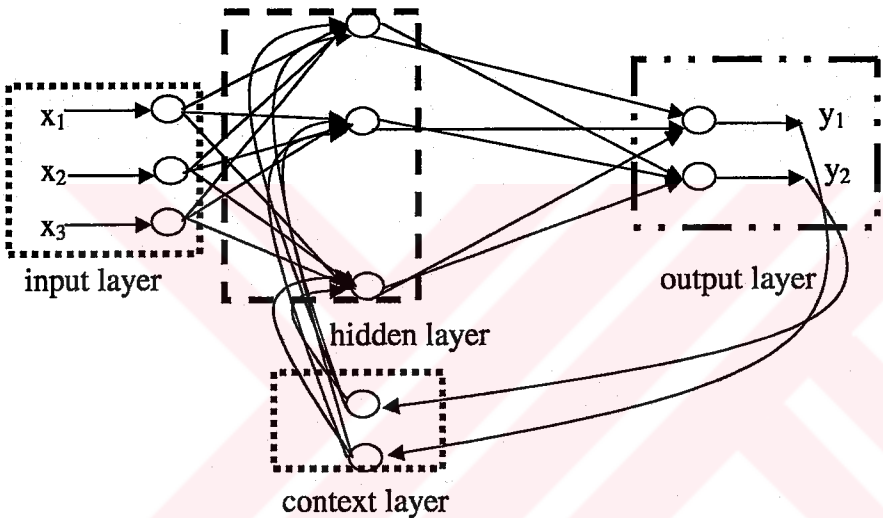


Figure 2.4. Structure of the Jordan-type recurrent neural network

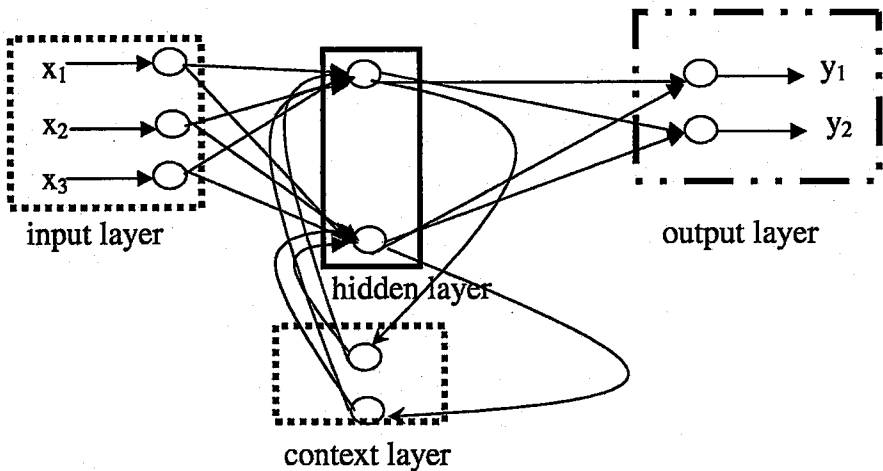


Figure 2.5. Structure of the Elman type recurrent neural network

## 2.3 Statistical Significance Tests

### 2.3.1 Tests for PCA, FA and CCA

In this work, the maximum-likelihood method (ML) of Jöreskog and Sörbom (1989) is employed in order to assess the fit of the number of PCs and CFs. Let  $\mathbf{S}$  indicate the correlation matrix of  $p$ -dimensional data and  $\mathbf{C}$  indicate the correlation matrix of  $q$ -dimensional data ( $q < p$ ). Herein  $q$  is the number of significant PCs and  $p$  is the number of entire PCs. Then ML method suggests minimizing the following equation (Jöreskog and Sörbom, 1989).

$$\log|\mathbf{C}| + \text{Tr}(\mathbf{S}\mathbf{C}^{-1}) - \log|\mathbf{S}| - p \rightarrow \min \quad (2.32)$$

If the correct number of PCs is  $q$  then Equation (2.36) is approximately distributed in large samples as  $\chi^2$  (chi-squared) with degrees of freedom expressed by:

$$\frac{1}{2}[(p-q)^2 - (p+q)] \quad (2.33)$$

A small value of  $\chi^2$ , relative to the degrees of freedom means a good fitting of the model.

Furthermore, there are other alternatives to test for PCA, CCA and FA which can be written in terms of eigenvalues, in which  $\lambda_1 > \lambda_2 > \dots > \lambda_k$ , where  $k = \min(p, q)$  indicates the rank of the corresponding data matrix. There are four most common test-statistics according to Rencher, 1992; Rencher, 1995; Rencher, 1998; Jackson, 1991; Jackson, 1993; Johnson and Wichern, 1998. These are:

$$\left. \begin{array}{ll} \text{Wilk's lambda} & \Delta = \prod_{i=1}^k \frac{1}{1 + \lambda_i} \\ \text{Roy's root} & \theta = \frac{\lambda_1}{1 + \lambda_1} \\ \text{Pillai's trace} & \Gamma = \sum_{i=1}^k \frac{\lambda_i}{1 + \lambda_i} \\ \text{Lawley - Hotelling} & H = \sum_{i=1}^k \lambda_i \end{array} \right\} \quad (2.34)$$

The squared canonical correlation for each canonical variate can be written in terms of its associated eigenvalue as:

$$r_i^2 = \frac{\lambda_i}{1 + \lambda_i} \text{ for } i = 1, 2, \dots, k \quad (2.35)$$

and is interpreted in the same way as the univariate *r-square* (Jong and Kotz, 1999).

For interpretation of the CFs, Kaiser (1959) test is employed. This test is so simple that the corresponding eigenvalues of greater than 1 are attained as the significant common factors.

### 2.3.2 Tests for the sample correlation

To test the sample correlation,  $r$ , we have employed the statistics of  $t$ -distribution is applied with the following hypothesis (Berenson *et al.*, 1993).

$$\left. \begin{array}{l} H_0 : \rho = 0 \\ H_1 : \rho \neq 0 \end{array} \right\} \quad (2.36)$$

where  $\rho$  represents the population correlation. Then we have

$$t = \frac{r - \rho}{S_r} \approx t_{n-2} \quad (2.37)$$

here  $n$  is the number of observations with  $n-2$  degrees of freedom, and the standard error of the correlation coefficient is given by

$$S_r = \sqrt{\frac{1 - r^2}{n - 2}} \quad (2.38)$$

Using an  $\alpha$ -level of significance,  $H_0$  may be rejected if  $t > t_{1-\alpha/2; n-2}$  or if  $t < t_{\alpha/2; n-2}$ .

Unless  $\rho$  equals zero, the sampling distribution of the sample correlation coefficient  $r$  is not normally distributed. In order to adjust its distribution to normal distribution, Fisher  $z$ -transformation can be used by the following procedure.

1. In the first step  $r$  is transformed to the  $z$  by:

$$z_{Fr} = \frac{1}{2} \ln[(1 + r)/(1 - r)] \quad (2.39)$$



2. Instead of  $S_r$ , the estimated standard error of the transformed correlation value  $\sigma_{Fr}$  is used.

$$\sigma_{Fr} = \frac{1}{\sqrt{n-3}} \quad (2.40)$$

3. Instead of  $t$ -distribution,  $Z_{1-\alpha/2}$  is used.
4. The  $100(1-\alpha)$  confidence interval is estimated as:

$$Z_{Fr} \pm Z_{1-\alpha/2} \sigma_{Fr} \quad (2.41)$$

5. After determining the reconverting the lower and upper limits of Equation (2.41), the confidence interval estimate of  $Z_\rho$ , back to units of  $r$  and then the  $100(1-\alpha)$  confidence interval estimate of  $\rho$  is obtained.

On other hand, in order to test the Sampson correlation, both  $t$ -distribution and  $F$ -distribution statistics are employed. The  $F$ -statistics can be obtained from the  $r$ -square statistics as:

$$F = \frac{r^2}{1-r^2} \frac{n-p-1}{p} \quad (2.42)$$

where  $n$  is the sample size and  $p$  is the number of variables, and

$$r^2 = \frac{S_{YX} S_{XX}^{-1} S_{XY}}{S_{YY}} \quad (2.43)$$

where  $S_{\cdot}$  are the sums of squares and cross product matrices, or equivalently, the corresponding covariance (or correlation) matrices.

### 3. PROPOSED DOWNSCALING MODELS

In the conventional statistical downscaling methods, the models are generally constructed based on the statistical relationships between local scale observed variables and large scale troposphere variables. The relations are intended to represent the effects of the climate-modifying local factors, such as variable topography (Rummukainen, 1997). The local sites should cover an area comparable to the GCM grid box size. The necessary constraint is that the large scale variables are selected subject to represent the variability of large scale climate features.

Based on these arguments; in this study, the area of interest to represent those climatic fluctuations is selected between 30–60°N and 10–50°E (Figure 2.1) in order to capture the large scale climatic fluctuations which affect local scale climate variability over Turkey. Some of these large scale climate characteristics are the Mediterranean high-low situations, Azores high, Siberian high, north-western part of the Monsoon low, and additionally, and the northerly effects of Sahara.

#### 3.1 Downscaling Methods

There are three very-well known methods to determine the relationships between the large scale and local scale variables concerned with downscaling in the literature. These are:

1. *Downscaling with surface variables*: This method is self-explanatory to involve the establishment of statistical relations between large scale aerial averages of surface variables and local scale variables (e.g., Kim *et al.*, 1984; Wilks, 1989; Wigley *et al.*, 1990). In the application, the same local scale surface variables are selected as predictands while the predictors for each predictand could be the relevant large scale average or a combination of large scale averages. However, downscaling with surface variables is not a good approach, since the GCMs are not capable of simulating surface variables.

2. *Perfect Prognosis (PP)*: This methodology is a familiar concept from numerical weather prediction. The constraint of this method involves the development of relations between large scale free atmospheric observations and local scale surface observations (e.g., Klein, 1982). Furthermore, the grid-scale predictors of large scale variables may be used. For example, in Jonsson *et al.* (1994) and in Johanneson *et al.* (1995), the training data for downscaling purpose are gridded 500 hPa geopotential heights and 500-1000 hPa geopotential thicknesses related to observed station data to obtain the regression relationships between the large scale and local scale variables.

The other example may be seen in Tatlı *et al.* (2004) in which NCEP-NCAR reanalysis data sets (Kalnay *et al.*, 1996) was applied as large scale predictors (namely 500 hPa heights, 700 hPa heights, 500-1000 hPa thicknesses, 500 hPa vertical pressure velocities ( $\omega$ ), and mean sea level pressure) for downscaling Turkish rainfall data.

3. *Model Output Statistics (MOS)*: This method is similar to the *PP*, but the large scale free atmospheric variables are obtained from GCMs outputs. In the case of *MOS*, the applicability of the technique is related to the performance of the model providing the free atmospheric data. Hence, this approach is sensitive to the model types, as the model of the large scale data is changed then the relationships might have to be redone.

On the other hand, *PP* based downscaling is a model-type-free method (independent of the model-type) but, of course, depends on the quality and extent of the observed data.

The methods which are traditionally applied for downscaling purposes depend on the variability of predictors according to scale-based features. Figure 3.1 illustrates such issues of the spatial-scales according to the climate variability.

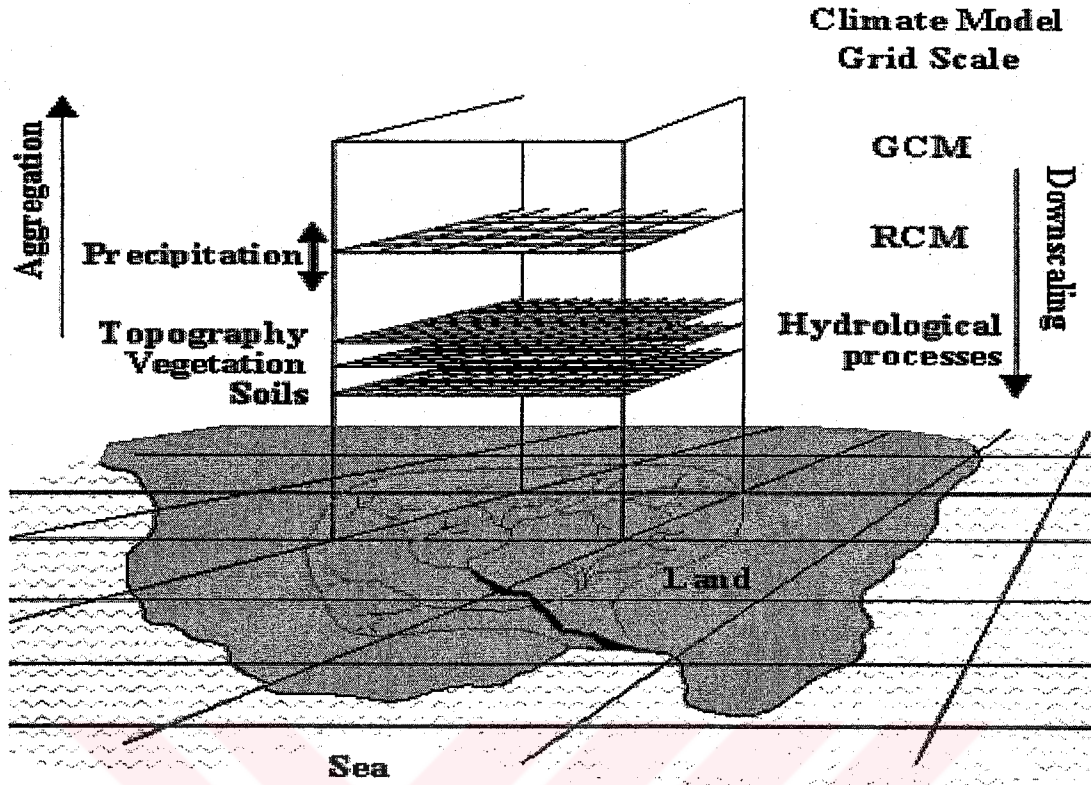


Figure 3.1 Formal representation the issues of scales associated with the climate features

In the following sub-sections, the proposed downscaling models are presented of which are the applications for Turkey in this work.

### 3.2 Proposed Downscaling Model Based on Multivariate Techniques

The details of the redundancy analysis are given in the Section 2.2.4. If it is assumed that  $X$  represents the large scale variables (predictors) and  $Y$  the local scale variables (predictands) then once the redundancy variates are obtained by Equation (2.22) or Equation (2.23) as  $W$  then the regression equation for downscaling can be written as in the following.

$$Y = XWW^T C_{XY} \quad (3.1)$$

where  $C_{XY}$  indicates the cross-covariance matrix. This proposed method was applied to downscale monthly surface temperature series in Turkey from large scale upper air circulations (Tatlı *et al.*, 2005). The proposed downscaling model components are given in Figure 3.2.

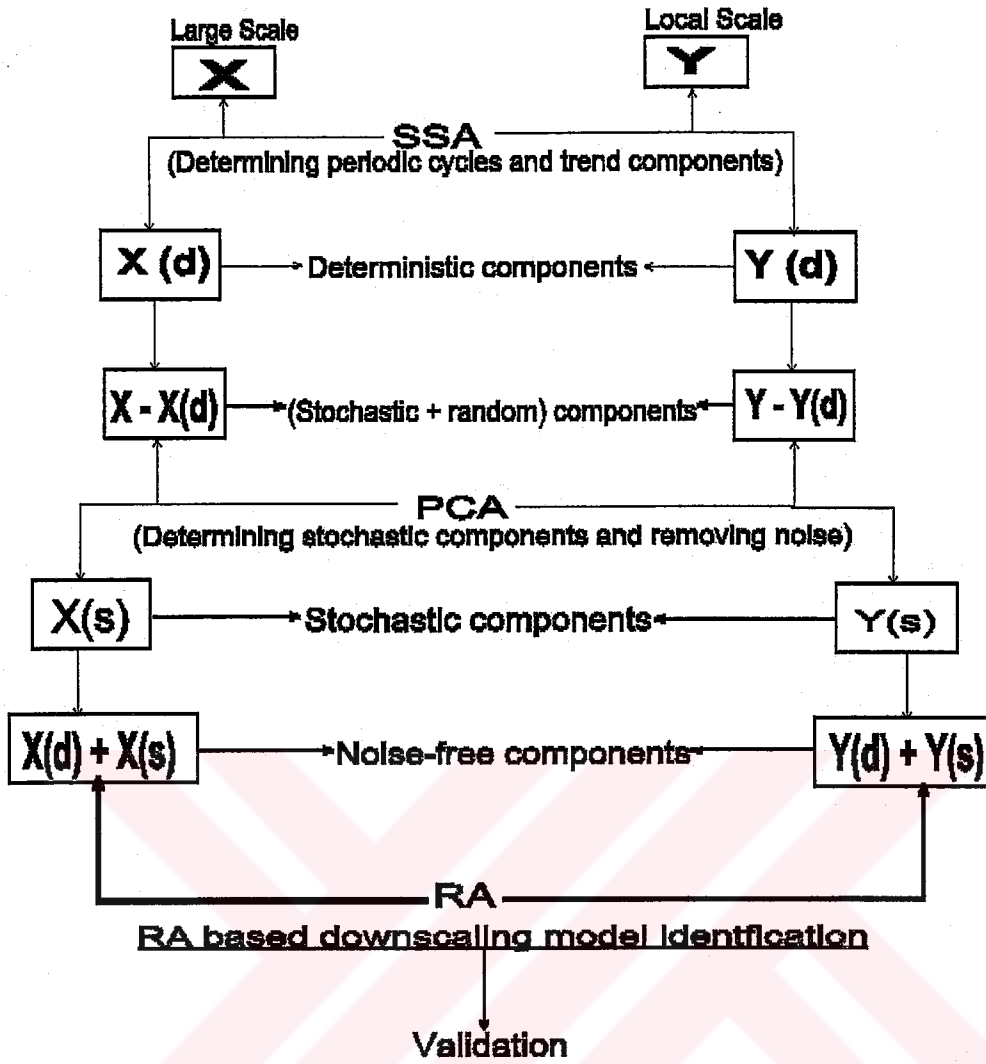


Figure 3.2. The proposed model components

The monthly surface temperature series have dominant periodic components that cause to prevent a true statistical analysis, for instance, in the trend and, both in the stochastic components and noise, on the other hand. Hence, Fourier analysis (Richardson, 1981) may not be directly applied since Fourier method needs the assumption of stationarity in the process. Thus, SSA may be performed to decompose the variables as into the deterministic, mixed statistical and random components. The main concepts of SSA are not described here, since this method has been introduced in Section 2.2.6.

After determining the periodic and/or the trend components; principal component analysis is performed on the residuals to extract the statistical and noise components. In order to obtain the noise-free spatial modes of climatic variability, the deterministic components (periodic cycles and/or trend) are subtracted from the raw

data series and then the significant spatial principal components are selected for residuals of the data sets assessing the fit of the model based on maximum likelihood method under  $\chi^2$  (chi-squared) distribution of 95% significance level of unequal variances of PCs with the degrees of freedom expressed by:  $1/2[(p-q)^2 - (p+q)]$  (Jöreskog and Sörbom, 1989). Here  $p$  and  $q$  represent the entire PCs and the number of unequal-significant PCs, respectively. In the final stage, the both predictors and predictands are assumed to be now noise-free according to the methods.

### 3.3 Proposed Downscaling Model Based on Recurrent Neural Networks

A downscaling model may be divided into a *naïve* (e.g., AR, ARMA) and causal (e.g., regression) components. One must distinguish between causal models and naive models. A naive model approach involves using only historical data. Thus it assumes that the trend of the dependent variables, which in this study are the monthly total precipitation series, remains constant over the time, and the value of the dependent variables  $Y(t+k)$  where  $t$  is the present month and  $k$  is the number of the predicted months which can be extrapolated from historical data.

On the other hand, a causal model approach assumes that there are some (causal) variables (e.g., GCMs outputs) that are responsible for changes in the dependent variables. In the causal models, previous information is available from the GCMs, which can be considered in the process of the model building for downscaling purposes (i.e. a Bayesian approach).

Assuming now that  $\Delta t$  illustrates the prediction time step and  $G$  is an unknown transformation (such as AR coefficients in the AR model), then to predict the multivariate  $Y$  from its own finite countable previous states, the following relation can be used:

$$Y(t + \Delta t) = G\left[\sum Y(t - k\Delta t)\right] + e(t); k = 0, 1 \dots n \quad (3.2)$$

In order to distinguish the models, let  $Y_1(t)$  and  $Y_2(t)$  denote the outputs of the causal model and the naive model, respectively. Assume that the static transformation (causal) is linear and the dynamic prediction model (naive) is first order (Richardson 1981); then these two separated prediction models can be written as:



$$Y_1(t + \Delta t) = A[X(t + \Delta t)] \quad (\text{causal model}) \quad (3.3)$$

$$Y_2(t + \Delta t) = B[Y(t)] \quad (\text{naïve model}) \quad (3.4)$$

here  $X(t)$  represents causal variables, and for simplification, the error term  $e(t)$  is omitted. The main problem arises to select the best outputs from these models. However, it is not clear which of the model outputs is important for local scale processes, but one may propose a multivariate difference equation as in the following (Bras and Rodriguez-Iturbe, 1993).

$$Y(t) = M_1[Y(t - \Delta t)] + M_2[X(t)] + M_3[e(t)] \quad (3.5)$$

where  $M_1$ ,  $M_2$  and  $M_3$  are linear operators, and the vector  $e(t)$  consists of uncorrelated white noise. The vector covariance of  $e(t)$  is a diagonal matrix,

$$D_e = E(e(t)^T e(t)) \quad (3.6)$$

where  $E$  illustrates the expected value, but in practice it is generally taken as sample average. The definition of Equation (3.5) forms a possible way for accommodating of the naive and a causal model into a new model with linear structure. Meteorological or climate processes are generally expected to behave nonlinearly, whereas the process represented by Equation (3.5) has a linear static part and a linear dynamic part, one can claim that it represents a linearization approximation. Therefore, Equation (3.5) may not be enough to satisfy non-linearity constraint. Hence, a more sophisticated approximate approach for solving this problem is developed. If two downscaled sets from Equation (3.3) and Equation (3.4) are combined as  $Z(t) = [Y_1(t), Y_2(t)]$ , then the suggested model can be written as:

$$Y(t) = M[H[Z(t)]] + e(t) \quad (3.7)$$

where  $e(t)$  is the multivariate Gaussian noise.  $M$  is a linear operator that it predicts local variables from the  $\{H[Z(t)]\}$  basis functions, and  $H$  is a non-linear scalar function given in Equation (3.8) that can generate the basis functions, respectively. This is a new version of the downscaling method which constitutes both the knowledge from the static model of the causal large scale variables given in Equation



(3.3), and the knowledge from the first order local linear dynamic model given in Equation (3.4).

However, the regression equation in Equation (3.3) can be applied for homogenous sets of the predictors and predictands, for which the regression is evaluated. Therefore, linear regression given in Equation (3.3) can not be applied due to heterogeneity, if heterogeneity is introduced to the dependence. In this study, the 'homogeneity in dependency' is an assumption to describe the invariance of the relation (deterministic rule) between predictors and predictands. Hence, a method is proposed to solve this problem. First by canonical correlation analysis, which may be seen as a classification of predictors and predictands, the significant canonical correlation variates are determined. Second, the so-called causal model is evaluated between these CCVs. In this study, CCA is employed after independent component analysis.

Now the structure of the scalar function  $H$  can be determined. It should be non-linear in its structure to satisfy the constraint of non-linearity. So, we transform  $\mathbf{Z}(t)$  element by element to generate the basis variables of the system in Equation (3.7) with a non-linear  $H$  function (Equation (3.8)), thereafter the degrees of problem may be reduced. Two such possible non-linear functions from neural network applications are (Connor *et al.*, 1994; Haykin 1999; Haykin 2001):

$$\left. \begin{aligned} H_1(y) &= \tanh(\alpha y) \\ H_2(y) &= \frac{1}{1 + \exp(-\alpha y)} \end{aligned} \right\} \quad (3.8)$$

where  $\alpha$  is a scalar constant. The preceding proposed method algorithm stages are summarized in Table 3.1, and a flow chart diagram that describes the model components is also shown in Figure 3.3.

The proposed approach is indeed an originating form of a recurrent neural network (RNN) model (Jordan, 1986; Elman, 1990; Robinson and Fallside, 1991; Puskorius and Feldkamp, 1994; Connor *et al.*, 1994; Pearlmutter, 1995; Haykin, 1999; Haykin, 2001; Hochreiter *et al.*, 2001).

Table 3.1. The proposed method algorithm

1. Perform a statistical large scale analysis by ICA based on PCA, and employ CCA between ICs and monthly total precipitation series for reducing the proper predictors and predictands (see text).
2. Divide all the data sets into two parts. Select one part for model identification and leave the second part for validation.
3. Normalize all data sets in use to make their means zero and variances one.
4. Construct a naive and a causal model separately.
5. Transform the outputs of the linear static (causal) and dynamic (naive) models by one of the functions given in Equation (3.8).
6. Construct a multivariate linear regression model between the responses of the fifth step and the predictand variables.
7. Apply a training through-time algorithm (e.g. Puskorius and Feldkamp, 1994; Pearlmutter, 1995; Haykin, 2001; Hochreiter *et al.*, 2001) a case in the RNN models via part (causal) by part (naive)
8. Test the model to see whether it is appropriate or not (validation)

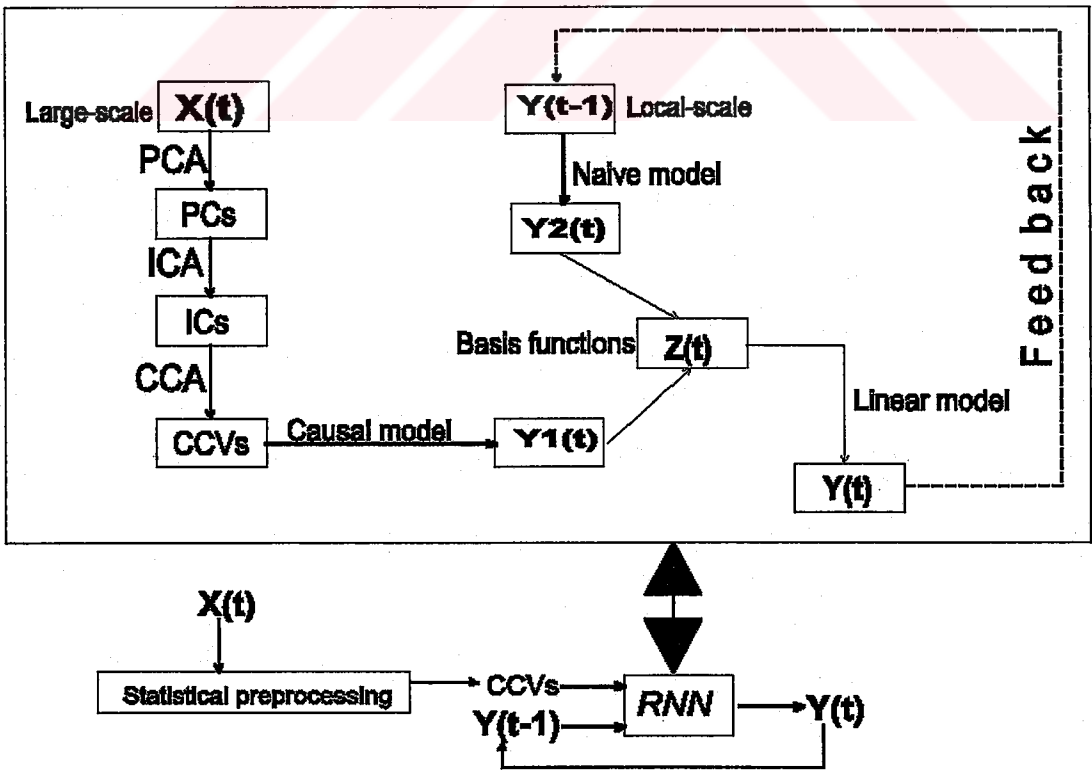


Figure 3.3. The flow chart of the components in the proposed downscaling approach

In the conventional RNN models, the training process is based on minimizing the variance of the error between estimated and observed values. Hence the structure of a RNN is sensitive to dynamics of the process. The RNN model of Jordan (1986) could be represented as:

$$\begin{aligned} \mathbf{Z}(t + \Delta t) &= H[\mathbf{M}_1 \mathbf{Y}(t) + \mathbf{M}_2 \mathbf{X}(t + \Delta t)] \\ \mathbf{Y}(t + \Delta t) &= \mathbf{M}_3 \mathbf{Z}(t + \Delta t) \end{aligned} \quad (3.9)$$

where  $H$  is a non-linear scalar function defined as in Equation (3.8) and  $\mathbf{Z}$  illustrates the states of the system (outputs of hidden neurons), and  $\mathbf{M}_1$ ,  $\mathbf{M}_2$  and  $\mathbf{M}_3$  are linear operators.

In the proposed RNN, the initial weights ( $\mathbf{M}_1$  and  $\mathbf{M}_2$  in Equation (3.9)) between inputs and hidden neurons are selected as  $\mathbf{A}$ , given in Equation (3.3), and as  $\mathbf{B}$ , given in Equation (3.4). If the initial weights are selected randomly, then the conventional RNN may capture only the dynamical features of the local scale process. To solve this problem, the training process is employed part by part. The elements of  $\mathbf{A}$  (naive) are kept constant while updating the elements of  $\mathbf{B}$  (causal); similarly the elements of  $\mathbf{A}$  are updated while keeping constant the elements of recently updated  $\mathbf{B}$ . Consequently, the proposed RNN can satisfy not only the non-linearity such as in the case of conventional RNN but also satisfies and distinguishes the naive and the causal relationships (linear or non-linear).

In this study, the well-known training process algorithms of RNN are not described, but one may find the effective algorithms, *e.g.*, in Puskorius and Feldkamp (1994), Pearlmutter (1995), Haykin (2001), Hochreiter *et al.* (2001), for example.

## **4. A GENERAL LOOK AT THE CLIMATE-CHARACTERISTICS OF PRECIPITATION AND TEMPERATURE OVER TURKEY**

### **4.1 Precipitation**

In general terms, Turkey is associated with a subtropical climatic regime that is referred to as Mediterranean. Because of its geographic location, the atmospheric systems that shape the weather regimes affecting Turkey have both polar and tropical origins, which are dominant in the winter and summer respectively.

During the autumn and winter, polar fronts have a determining effect on the mid-latitudes. Large scale atmospheric motions are translated into local weather conditions due to different land characteristics. These local conditions have both thermal and dynamical features. On the other hand, the different local conditions have sparse local climate features due to the land-sea distribution, elevation and other local physical-geographical properties that make those macro-climate features highly variable.

Turkey (despite being surrounded by the sea on three sides), because of its coastal regions characterized by high topography, includes a rather large inland area (central Anatolia) that has a continental climate character (Taha *et al.*, 1981; Erinc, 1984; Kadioğlu, 2000). The topography causes strong variations in the climate. The coastal sides of the mountains receive heavy precipitation, but there are also thermal effects due to topographic height differences coupled with other factors.

#### **4.1.1 Precipitation in summer months**

Since the maritime-polar (mp) and continental-polar (cp) air systems move towards northern regions, the tropical air systems have a dominant effect the Mediterranean basin in this season. The southern and southeastern parts of the country are under the influence of continental tropical air system due to the Azores high and the circulation-based Monsoon low which are quite dry and warm. The maritime tropical air system that travels from the Atlantic towards Turkey (a northwesterly motion)

gets heated and loses most of its relative humidity during its passage across the mainland and presents a stable structure. The Azores high-pressure centre locates over Europe during its northward and eastward displacement, and this situation affects the weather and climate, especially for the western parts of Turkey.

Both the marine and continental air systems, especially during the June–August period, have stable structures characterized by hot and dry conditions, and cloudless summers are quite typical for Turkey, except in the Black Sea region.

Exceptions do occur, as sometimes the western, and especially northwestern, parts of the country get heavy precipitation. The precipitation is the result of the frontal systems forming as the northeastern Atlantic originated air system interacts with the tropical maritime air system.

The northern part of the country, having high topography along the Black Sea coast, also experiences some heavy precipitation of the orographic origin. In the west, the mountains are not parallel to the coast; as a result, the orographic precipitation has a more local character. The mountains also divert the air flow in various ways, causing central Turkey to possess a peculiar wind regime.

#### 4.1.2 Precipitation in winter months

The Mediterranean basin, which includes all the European countries having a coast on the Mediterranean, becomes an active frontogenesis region at the beginning of autumn. The Azores high shifts to south and the Siberian high-pressure system starts to affect the northern and eastern Europe (due to thermal reasons); the Mediterranean belt becomes a convergence zone. The frontal systems that travel along various trajectories towards the east affect Turkey. These cyclonic depressions have two main paths across Turkey.

The first affects the western and southern parts of the country and then continues towards the southeastern parts. The second has a trajectory towards the northeast and is the cause of most of the precipitation in the northwestern, northern and central parts of the country.

The other trajectory of the polar front that affects Turkey is felt when the Azores high becomes strong enough to affect the Western Europe. When this comes onto the scene the depressions travel from the Thrace–Marmara and western part of the Black

Sea region towards southeastern Turkey by northerly and northwesterly flows; most of Turkey is affected by those motions. The effect is more pronounced when the mechanisms related to the Azores High create northerly and northwesterly winds over the Anatolian Plateau.

The principal effects of the Mediterranean cyclones on Turkey are southwesterly winds on the Turkish Mediterranean coast. Topography creates important local effects in the western and Black Sea coasts of Turkey: the seaward (wind) sides of the mountains are washed by heavy precipitation and the precipitation decreases towards the inner parts of the Anatolian Peninsula. There exists a very sharp negative orographic precipitation gradient from the coastal regions towards central Anatolia. In central Anatolia, there is a plateau effect rather than a sharp topography effect.

During the winter, a convergence zone also forms in the eastern part of the Black Sea. This is a consequence of the interaction between the north-easterly and easterly currents created by the Asian high (Siberian high) and the local southerly currents created by the small-scale high-pressure centers due to the thermal effects in the eastern Anatolia. Foehn winds and rain shadows are, therefore, common in this region during cold, winter periods. Dry conditions (drier than normal) associated with the Foehn winds (southerly and south-easterly upper air flows) are common during the spring months on the middle and eastern Black Sea coastal belt.

In the winter the passage of the frontal cyclones causes precipitation in the coastal regions, but central and eastern Anatolia remain largely continental in terms of climatic features and become a divergence field.

During the spring, depressions gradually decrease over Turkey, and the depression frequencies reaching a minimum during the summer (Tatlı *et al.*, 2003). The depressions that are seen during the summer mainly affect the northern part of the Marmara region, the Black Sea coastal area and the north-eastern Anatolia sub-region.

Spring and summer are seasons that are dominated mainly by the local effects. The mountains that run parallel to the Black Sea coast of Turkey prevent the northerly currents from penetrating inland. However, the situation in the west is different, as the mountains run perpendicular to the coast, so they cannot block the westerly

currents very effectively, and the mid-latitude depressions penetrate into the mid-western parts of the country through the gaps between these mountains.

## 4.2 Temperature

Turkey is in a region which is often described to have a warm and moderate climate (Erinç, 1984). The highest maximum temperature is observed in the south-east of the country, particularly in summer months. Towards the north-west and north-east the temperature decreases gradually, yet this decrease is less strong during the summers due to the continental effects of the inner regions. At low altitudes, the coastal regions are warmer than the inner regions which are separated from the former by high mountains. On average the Mediterranean coast has the highest temperature, followed by the Aegean, the eastern part of the Black Sea region and the Marmara coasts. In the eastern Anatolia, due to the continental effects and high altitudes, there is a widespread temperature decrease.

The extensions of the continental and topographic effects are important in determining the distribution of temperature variability. The most interesting thermal character of Turkey is the rise of temperature in all the regions due to continental effects and thereby a decrease in regional contrasts during summer. The resulting temperature field has a negative gradient towards the north-west. Central Anatolia becomes very hot during the summer and its temperature contrast with the coastal areas decreases. Roughly speaking, from a macro-climate point of view, Turkey represents a more homogeneous region during the summers.

During winter, continental influence, topographic-height, and of course the latitude are determining factors for the minimum temperature. The main difference between summer and winter is that during winter the regional temperature contrasts are very strong, implying that there is negative temperature gradient from the coasts towards the inner parts. In the inner parts, the continental effects are pretty much the same and there are no huge temperature differences from region to region, but still there is a gradual temperature decrease towards the east within the Anatolian Plateau.

There is a negative temperature gradient towards the west in the Thrace region. This means mostly closed isotherms over Turkey and positive temperature anomalies



towards the coasts while negative temperature anomalies are observed towards the inner parts.

The highest temperatures during winter are observed along the Mediterranean coast which is influenced by the air systems of sub-tropical origin. In the Aegean and eastern part of the Black Sea region the temperature does not fall very much due to marine effects (also due to the foehn winds in the case of the eastern part of the Black Sea region). Southerly winds are generally common in the central part and move towards the convergence field in the eastern Black Sea. The lowest temperature values are observed in the north-eastern Anatolian Plateau.

The annual temperature is least contrasting at the coasts (especially in the eastern Black Sea coast due to the additional effects of the foehn winds). The annual temperature contrast increases towards the inner parts. The transitions from the cold to the warm period are more abrupt in the inner parts than in the coastal areas.

The daily temperature has a similar character with a small contrast at the coasts, and a high contrast in the inner parts. Especially in the high plateaus of Anatolia, during winter there is a strong daily contrast when frontal activities are common, but also in the transitional seasons of spring and autumn.

In Turkey, the monthly extreme temperature averages are quite different from the monthly averages. This is mainly due to the continental effects. The above mentioned difference increases towards the inner parts of the country and becomes very large in the eastern Anatolia. During summer, the differences between the monthly mean temperature series and those of the extremes are the largest in the coastal areas. The extreme temperatures have maxima in the south-eastern region of Anatolia during summer. The average maxima are smallest in the Black Sea region. In central Anatolia it is common for the monthly mean temperature to fall below zero in January. In winter, these values are usually well below zero towards the eastern Anatolia, while moderate values are observed in the coastal areas.

## 5. RESULTS

### 5.1 Downscaling Monthly Total Precipitation over Turkey

In this section, a statistical downscaling approach of monthly total precipitation over Turkey, which is an integral part of system identification for analysis of local scale climate variables, is investigated. Based on perfect prognosis, a new computationally effective working method which was introduced in Section 3.2 by the proper predictors selected from NCEP-NCAR reanalysis data sets of which are simulated as possible as perfectly by the GCMs during the period of 1961-1998.

Sampson correlation ratio is used to determine the relationships between the monthly total precipitation series and set of large scale processes (namely 500 and 700 hPa geopotential height, mean sea level pressure, 500 hPa vertical pressure velocity and 500-1000 hPa geopotential thickness).

In the study, statistical preprocessing is implemented by independent component analysis rather than principal component analysis or principal factor analysis. The proposed downscaling method is originating from a recurrent neural network model of Jordan (1986) that it uses not only the large scale predictors but also the previous states of the relevant local scale variables. Finally, some possible improvements and suggestions for further study are mentioned.

A number of studies regarding the spatial and temporal properties of precipitation or rainfall in Turkey have been published (*e.g.*, Türkeş, 1996; Türkeş, 1998; Türkeş *et al.*, 2002a; Kutiel *et al.*, 1996; Kutiel *et al.*, 2001; Kadioğlu, 2000; Touchan *et al.*, 2003). However, this work is the first study regarding the statistical downscaling of the local variables for Turkey (Tatlı *et al.*, 2004).

In the study, the monthly total precipitation data for 31 Turkish stations has been selected as an application of the proposed approach. It is demonstrated that the proposed method can well suppress the increase of the downscaling error due to nonlinearities, improving greatly the accuracy of the downscaling compared to the conventional regression-based techniques.

### 5.1.1 Problem statement

The basic data consists of the monthly total precipitation for 31 Turkey stations Turkey (Table 5.1 and the additional detail is shown in Figure 2.2a) with a record length 38 years, during the period 1961-1998. The selected meteorology stations have no missing data in the prescribed period, and represent the major characteristics of the precipitation over Turkey.

In order to perform a successful statistical large scale analysis (Rummukainen, 1997), NCEP-NCAR reanalysis data sets are windowed between 10–50°E and 30–60°N; since this area is assumed large enough to represent the large scale climate, or synoptic features being considered to affect the monthly total precipitation series over Turkey. The following large scale data sets from NCEP-NCAR reanalysis data sets are used as predictors in this study (the details of the geographical distribution of both the large scale and local scale variables is given in Chapter 2).

- 1) The interactions of long (Rossby) and short waves are identified at 500 hPa level. At this level, the long waves generally show a barotropic character whereas the short waves are baroclinic. Therefore, the short waves at this level have determining effects on the development of surface cyclonic and anti-cyclonic systems. The cyclonic and anti-cyclonic vorticity at all high levels is a result of the interactions of short waves of high level flows and surface systems. As a consequence of these arguments; 500 hPa geopotential heights are considered as predictors.
- 2) 700 hPa geopotential heights are also considered as one of the predictors. The clouds and precipitation occur by the vertical motions in regions with short waves. Especially, the local precipitation is significantly related to the short-waves at 700 hPa and 500 hPa levels.
- 3) Due to the thermal effects of troposphere, 500-1000 hPa geopotential thicknesses are considered.
- 4) In order to represent surface weather systems features, mean sea level pressure (a prognostic variable in the GCMs) is considered. MSLP is the most important causal variable of the proposed model, since (Kutiel *et al.*, 2001) have recently studied the relationships of the MSLP patterns associated with the dry or wet monthly rainfall conditions in Turkey.

- 5) The diagnostic predictor is 500 hPa pressure vertical-velocities ( $\omega$ ), since it is considered to represent the vertical dynamics of the troposphere (Holton, 1992; Bluestein, 1993), and vertical dynamics properties are associated with occurrence of precipitation process.

In order to represent jet streams in the troposphere, one may take 200 hPa and 300 hPa wind speeds (*e.g.*, Xoplaki *et al.*, 2003), but as seen below, while preprocessing large scale variables, 500 hPa vertical pressure velocities can satisfactorily represent the vertical dynamics characteristics.

The significant canonical correlation variates (CCVs) are selected such that the each of the CCVs can explain the maximum variance of the monthly total precipitation series for Turkey stations. The significant PCs and CCVs are shown in Table 5.2 and Table 5.3, and Figure 5.1 displays the contrast between the PCs and ICs.

In actual fact, a true statistical downscaling may be done within the data sets with similar power spectra. Indeed the CCVs of the large scale processes and CCVs of the local scale processes show similar spectral spectra. As a consequence, the downscaling process is said to be in conformance with the constraint of 'homogeneity in dependence'. As a result, the low frequency large scale predictors induce instabilities in the prediction of the high frequency local scale predictands, and *vice-versa*. The first three spectral densities of the CCVs of the large scale and the local scale data sets are computed by Climlab2000 Software Package of Tourre (2000), and the results are shown in Figure 5.2.

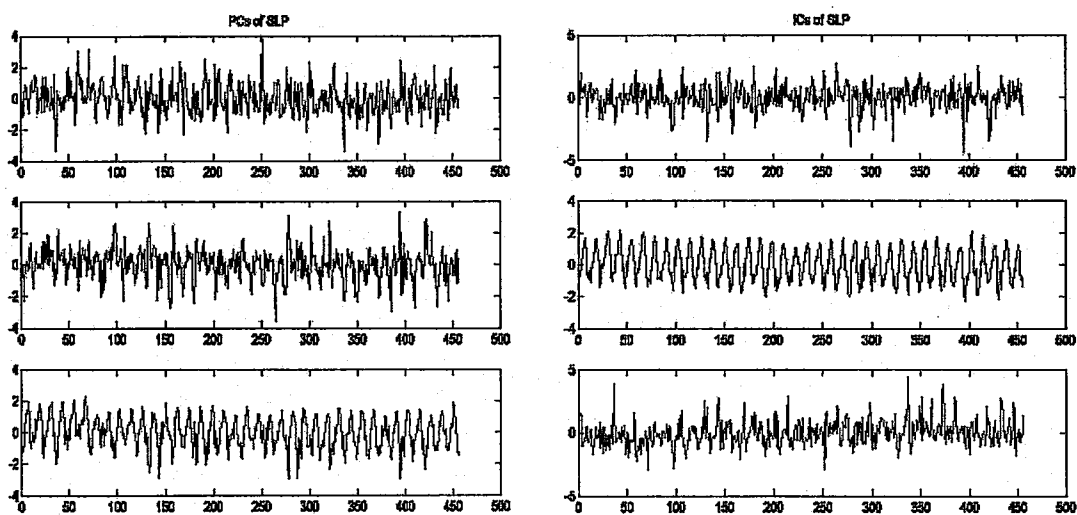


Figure 5.1. The first three principal components and independent components of mean sea level pressure series

Table 5.1. The Turkey stations of monthly total precipitation in the application

Station Name	Latitude (° North)	Longitude (° East)	International Station Number
Zonguldak	41.27	31.48	17022
Sinop	42.01	35.10	17026
Samsun	41.17	36.18	17030
Trabzon	41.00	39.43	17037
Rize	41.02	40.31	17040
Edirne	41.40	26.34	17050
Göztepe (İstanbul)	40.58	29.05	17062
Bolu	40.44	31.31	17070
Kastamonu	41.22	33.47	17074
Sivas	41.00	39.43	17090
Erzurum	37.18	40.44	17096
Kars	41.27	31.48	17098
Iğdır	42.01	35.10	17100
Çanakkale	40.09	26.25	17112
Bursa	40.11	29.04	17116
Ankara	39.55	44.03	17130
Van	41.40	26.34	17172
Afyon	38.45	30.32	17190
Kayseri	41.17	36.18	17196
Malatya	39.45	37.01	17199
İzmir	38.26	27.10	17220
Isparta	40.58	29.05	17240
Konya	41.02	40.31	17244
Gaziantep	40.11	29.04	17261
Şanlıurfa	40.37	43.06	17270
Mardin	40.44	31.31	17275
Diyarbakır	39.55	41.16	17280
Muğla	37.13	28.22	17292
Antalya	36.53	30.42	17300
Mersin	36.48	34.36	17340
Adana	37.00	35.20	17351

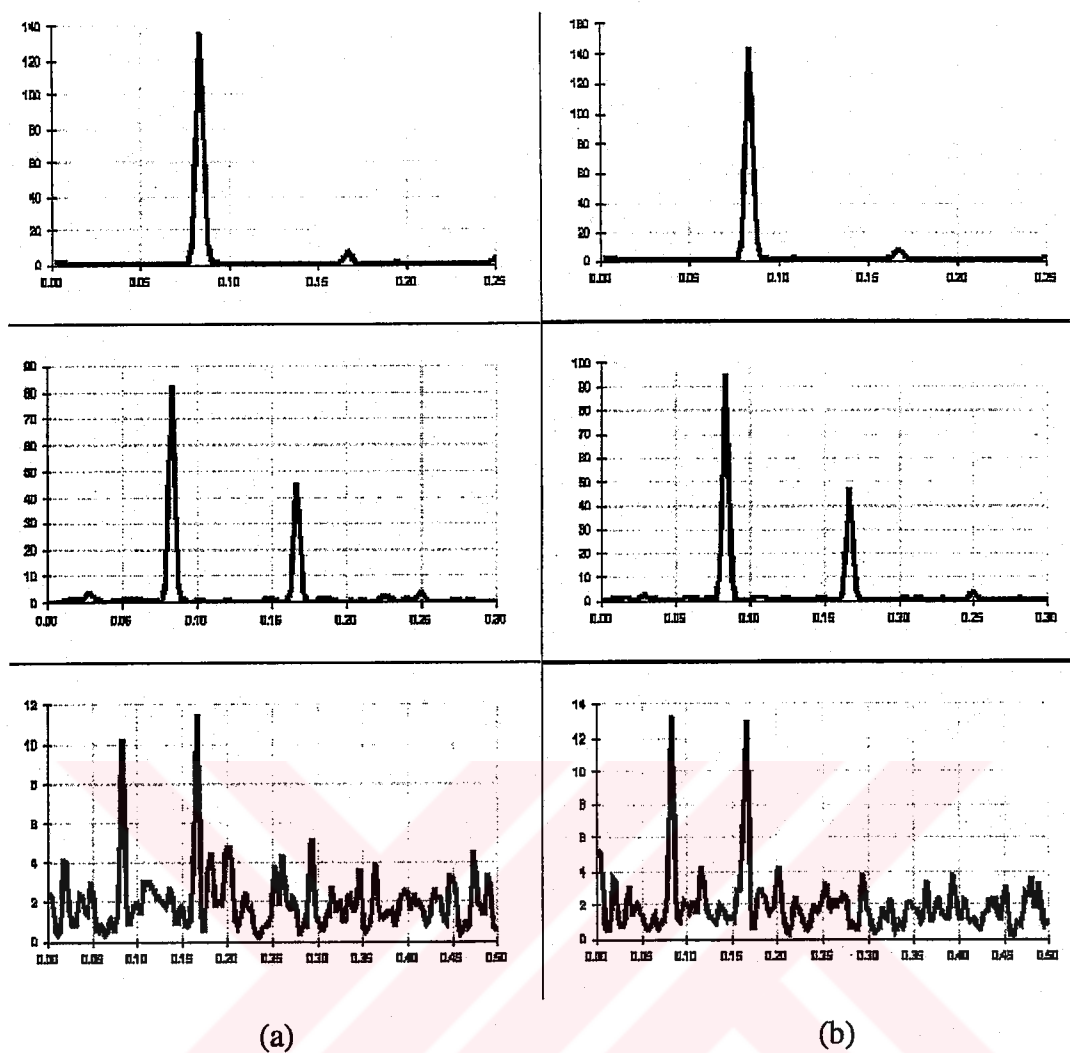


Figure 5.2. The spectral densities of the first three CCVs. The horizontal and vertical axes represent frequency (in *cycle/month*) and spectral density, respectively: (a) the large scale variables (on left); (b) the local scale variables (on right).

Table 5.2. The significant principal components of the large scale variables according to the criterion of (Jöreskog and Sörbom, 1989)

Large Scale Predictors	Sum of Explained Variance (%)	Number of PCs
500 hPa geopotential height	96.4	3
700 hPa geopotential height	94.8	3
Mean sea level pressure	90.4	3
500-1000 hPa thickness	96.6	2
500 hPa vertical pressure-velocity	69.7	6

Table 5.3. The canonical correlation variates of the large scale variables (*7th* canonical correlation variate is more significant than the *6th* canonical correlation variate)

CCVs of The Large Scale Variables	Explained Variance of the Monthly Total Precipitation Series (%)
CCV <sub>1</sub>	46.213
CCV <sub>2</sub>	18.931
CCV <sub>3</sub>	10.375
CCV <sub>4</sub>	7.748
CCV <sub>5</sub>	2.934
CCV <sub>7</sub>	1.581
Sum	87.782

### 5.1.2 Sampson correlation pattern analysis of Turkish rainfall associated with large scale climate features

In order to explore the relationships between the large scale data predictors and monthly total precipitation series; the Sampson correlation patterns are computed between the individual large scale variables (namely 500 hPa and 700 hPa geopotential heights, mean sea level pressures, 500 hPa vertical pressure velocities and 500-1000 hPa geopotential thicknesses) and precipitation series for four seasons; but in this study only the results for winter and summer are shown in Figures 5.3 to 5.7.

In these figures, the significance of  $R_s$  correlations (based on  $\alpha$ -level statistics) is shown in a gray-scale from light to dark to represent the decreasing order of significance.

Figure 5.3 shows the Sampson correlation coefficients between the seasonal sea-level atmospheric pressures (wet and dry seasons being characterized by the winter and the summer respectively) and the precipitation series of Turkey. The most important correlations are observed in the region that encompasses Italy, the entire Balkans region, Turkey and Eastern Mediterranean.



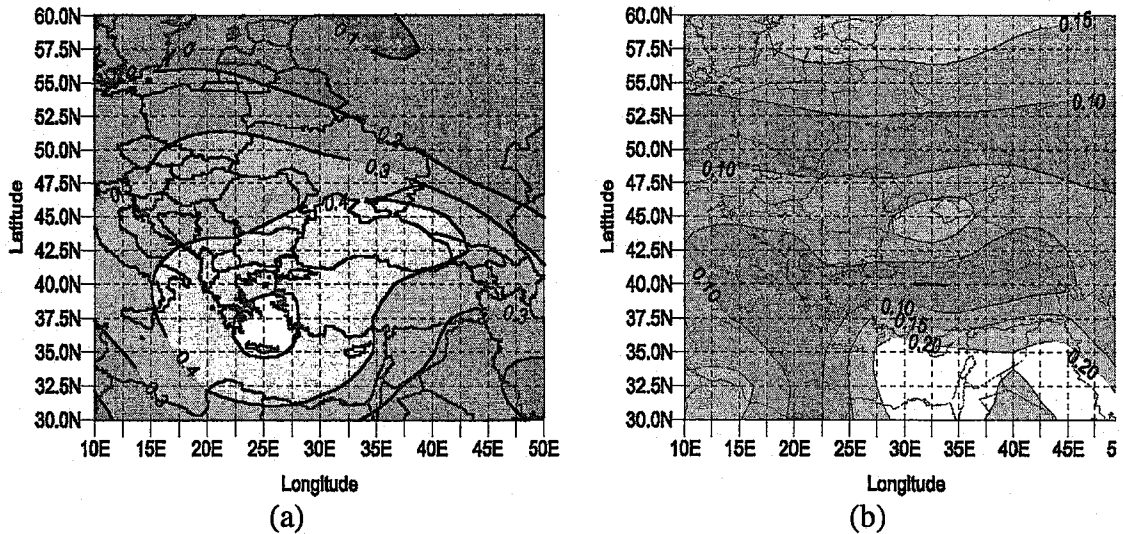


Figure 5.3. The  $R_s$  pattern between the MSLP and monthly total precipitation series: (a) winter months; (b) summer months

This explains the effects of the low pressure anomalies on the Turkish precipitation regimes during winter. These correlations are in agreement with the work of Kutiel *et al.* (2001). The strongest correlations are observed in the Central and Southern Aegean. For the summer, weaker correlations exist for the Eastern Mediterranean.

As seen in Figure 5.4, there is an important correlation between the values of the 500 hPa geopotential heights and winter precipitation series. The correlation pattern in the area that includes the western part of Turkey, Adriatic Sea, the Balkans and eastern part of Italy is at appreciable level. In winter, the geopotential low center at 500 hPa affects Turkey from the north-west. Such conditions are observed at high levels (500 hPa) when the low pressure centers and related frontal systems are observed at surface.

Significant correlations for the summer occur in the Eastern Mediterranean (namely Southern Turkey and Cyprus). During the summers, Turkey is under the influence of warm air masses (systems) of tropical origin, which is visible in terms of positive anomalies for the 500 hPa geopotential heights. Spring and autumn form transitions between the above-mentioned regimes.

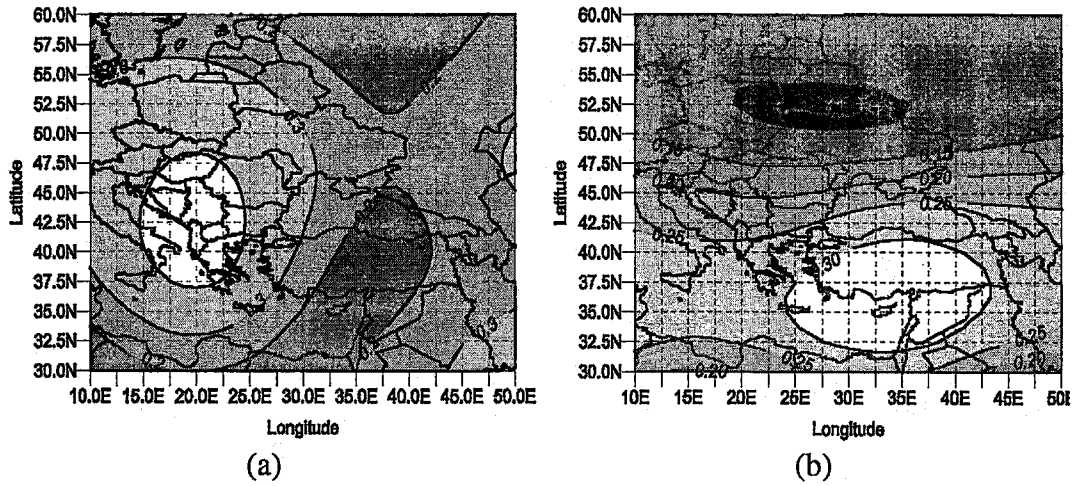


Figure 5.4. The  $R_s$  pattern between the 500hPa geopotential heights and the monthly total precipitation series: (a) winter months; (b) summer months

In Figure 5.5, 700 hPa geopotential heights data set gives a similar result to that of 500 hPa in term of its correlations with the precipitation regimes during the winter. In this figure, the significant correlations are seen over western-part. However, no such a significant correlation pattern exists for the summer (according to 95 % of statistical alpha level).

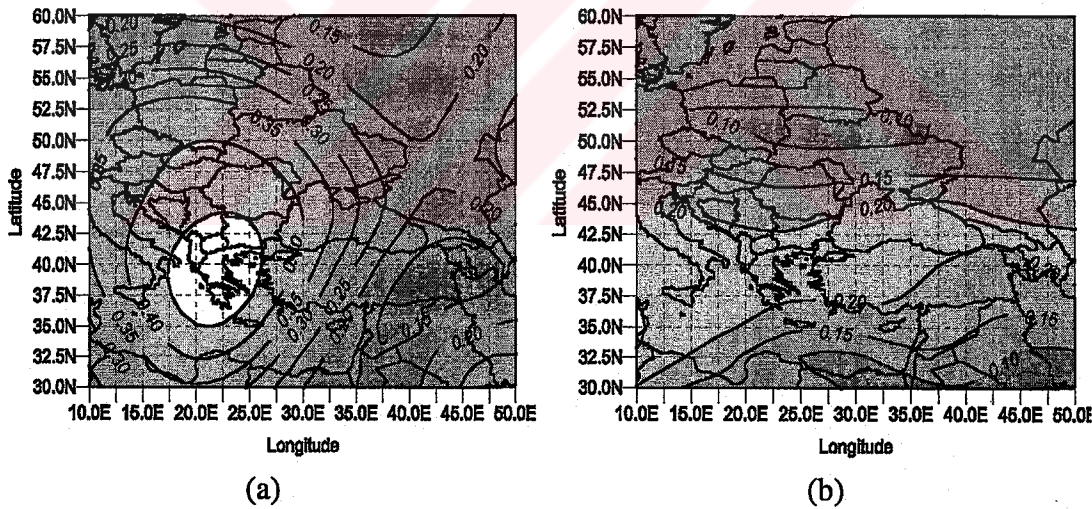


Figure 5.5. The  $R_s$  pattern between the 700 hPa geopotential heights and the monthly total precipitation series: (a) winter months; (b) summer months

The anomalies related to the geopotential thicknesses of 500-1000 hPa (in a similar way to the 500 hPa geopotential heights themselves) in Central Europe have a correlation with the Turkish precipitation regimes as it can be seen in Figure 5.6. It can be postulated that, especially during the winter, the negative thickness anomalies are related to the cold advection, because the thickness itself is an indicator of thermal advection.



As it can be seen from the contours, a similar correlation is valid for the Caucasus and the Caspian region as well (this correlation is related to the positive thickness anomalies that indicate the lack of precipitation during winter, *i.e.* winter with dry conditions). As discussed for Figure 5.3, Turkey is under the influence of Iceland and Mediterranean Lows in winter, therefore those regions are characterized as convergent fields. That is, the upward motions are effective. The pronouncement of such cold advection at the high levels (from 850 hPa to 500 hPa) has a determining effect on the occurrence of the precipitation.

In winter, on the eastern part of Turkey (especially the eastern Black Sea part of Turkey) there is an important convergent field (see Section 4) while exists of cold advection over of these upward motions has important effects on the formation of clouds and precipitation occurrence conditions. In summer, the effect of the air system which is originated from Sahara (dry and warm) is clearly seen in Figure 5.6b. This period indicates the dry period for Turkey.

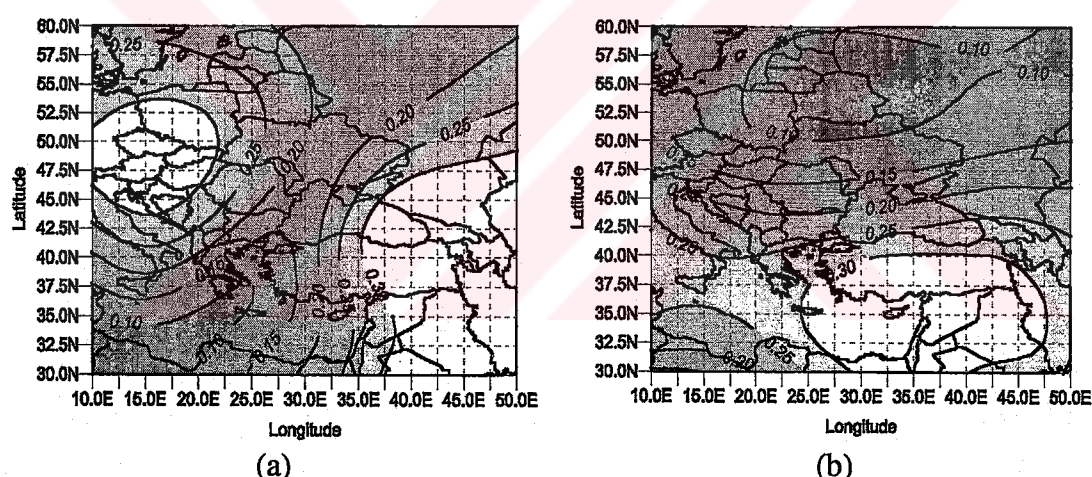


Figure 5.6. The  $R_s$  pattern between the 500-1000 hPa thicknesses and the monthly total precipitation series: (a) winter months; (b) summer months

The correlation patterns between the 500 hPa vertical pressure velocity ( $\omega$ ) series and the precipitation series can be observed on Figure 5.7. For both seasons, there is a strong correlation pattern for the area that includes the east of Greece, entire Black Sea and entire Eastern Mediterranean. There is also a strong correlation pattern for the southwestern corner of Turkey and Cyprus.

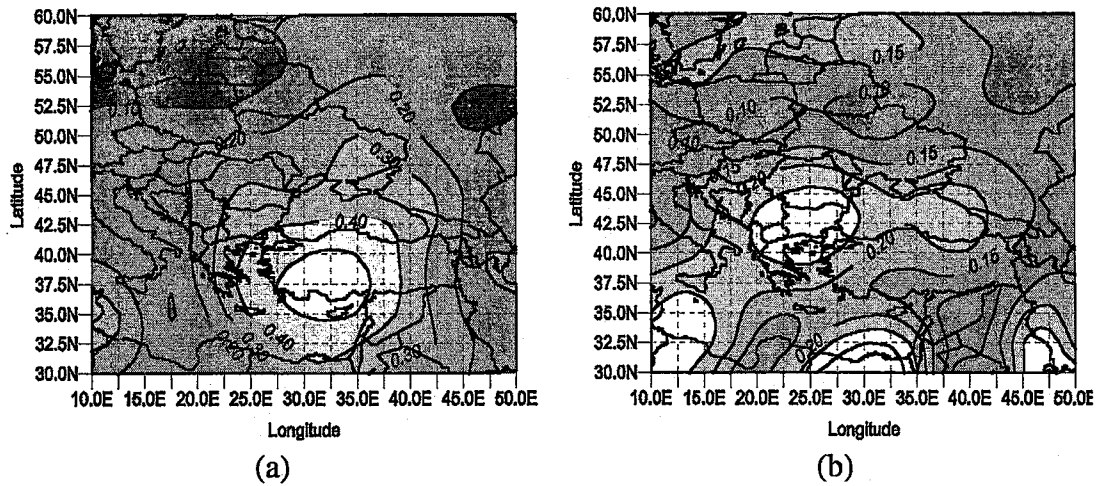


Figure 5.7. The  $R_s$  pattern between the 500 hPa vertical pressure velocities and monthly total precipitation series: (a) winter months; (b) summer months

During the winter, the area of omega-anomalies is on the trajectory of Mediterranean low pressure systems (Alpert *et al.*, 1990). For the summer the strong correlation patterns have a more 'patchy' nature, having maxima in the Balkans, Northern Africa (Sahara), Central Mediterranean, Arabic Peninsula and Persian Gulf. These centers have a potential to be related to the droughts as they are mainly characterized by the Azores high connected to sinking motions, and northwestern part of the monsoon low (the Arabic low) which has a dry characteristic during summers.

### 5.1.3 Downscaling results

In order to show the proposed model spatial performance, the pseudo principal component patterns of both the observed and downscaled precipitation series are shown in Figures 5.8 to 5.10 (since the amount of data is insufficient to extract the true principal component patterns; these are termed as pseudo-patterns). The model performance is also shown in Figure 5.11 to underline how the mean square errors are distributed over Turkey for the validation period of the proposed model output (the last half part of the data, *i.e.*, 228 months).

The scatter plots of the downscaled versus observed monthly total precipitation time series for Göztepe (Istanbul), Ankara, Diyarbakır, İzmir, Rize, Adana and Erzurum are selected from the rainfall regions according to Türkeş (1996) and Türkeş (1998) are shown in Figure 5.12. In order to distinguish the performance for validation, the correlation coefficients ( $r_1$ ,  $r_2$ ) between the observed and downscaled data sets are computed. In this figure,  $r_1$  and  $r_2$  represent the correlation coefficients for training

part of the data (the data part for identifying the model) and non-training part of the data (validation part) respectively.

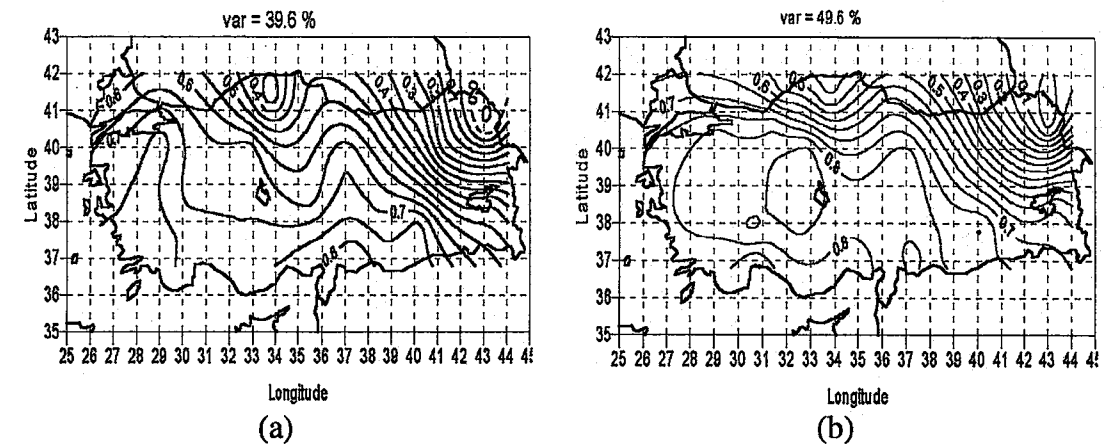


Figure 5.8. The first PC pattern of the monthly total precipitation series (the map indicates the correlation coefficients between the monthly total precipitation time series and the first PC): (a) observed; (b) downscaled

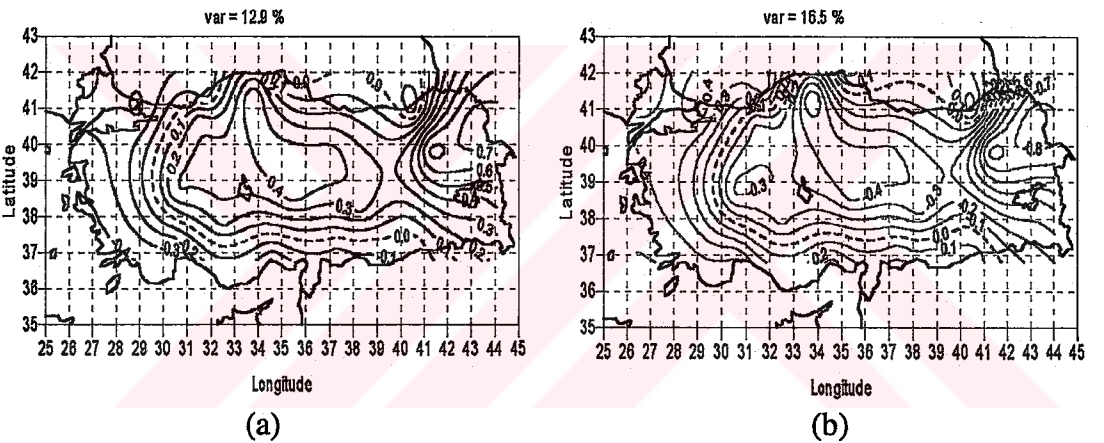


Figure 5.9. The second PC pattern of the monthly total precipitation series: (a) observed; (b) downscaled

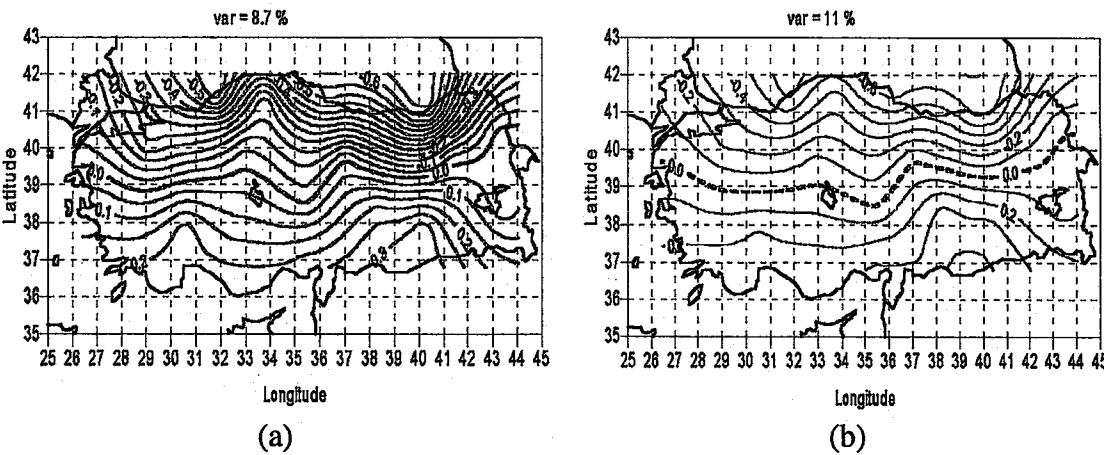


Figure 5.10. The third PC pattern of the monthly total precipitation series: (a) observed; (b) downscaled

The performance of the proposed model increases from the continental regions of Turkey to the coastal regions of Turkey except in the Black Sea and in southern Anatolia regions. This result is expected, because the precipitation conditions over the western and southern coastal regions (including mainly the Marmara Transition and Mediterranean rainfall regions of Turkey) are dominantly controlled by the large scale systems, whereas in the continental parts the added precipitation is due to the local conditions.

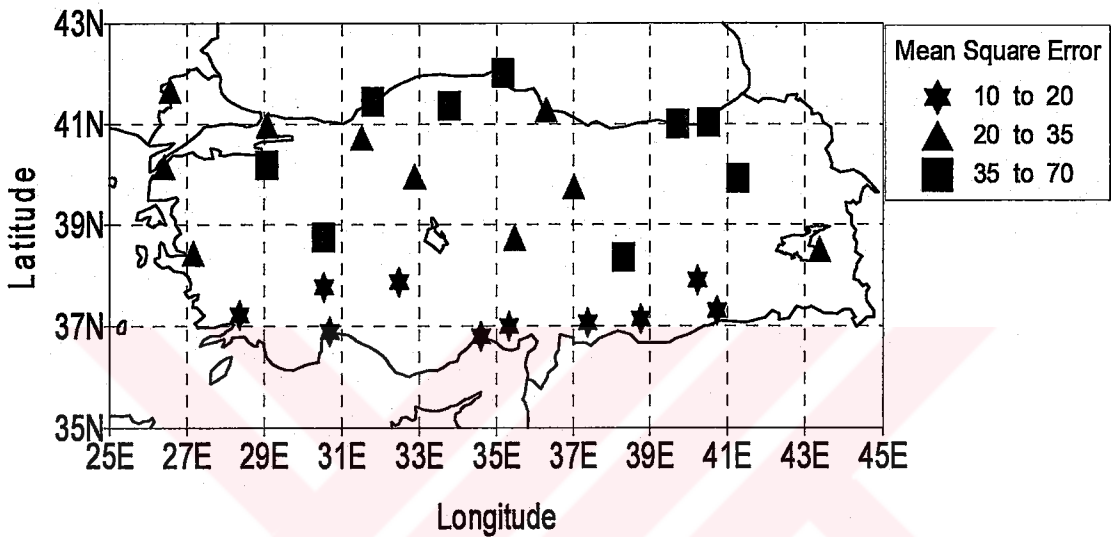


Figure 5.11. Geographical distribution of the mean square errors of the proposed model output for the validation period of the 1980-1998 interval.

In the Black Sea region, the additional precipitation is due to the topographical characteristics of a region includes such as the rain-shadows and/or exposure that the large scale processes may not capture.

The other unexpected result is for Diyarbakır. The performance of downscaling at this station is superior to the other non-coastal stations. Diyarbakır’s precipitation is affected by both the Mediterranean frontal low-pressure systems and large scale dry climate conditions (especially in summers). Therefore, Diyarbakır is not a typical continental station; this result is in agreement with the works of Türkeş (1996) and Türkeş (1998), where he named this region as “Continental Mediterranean”.



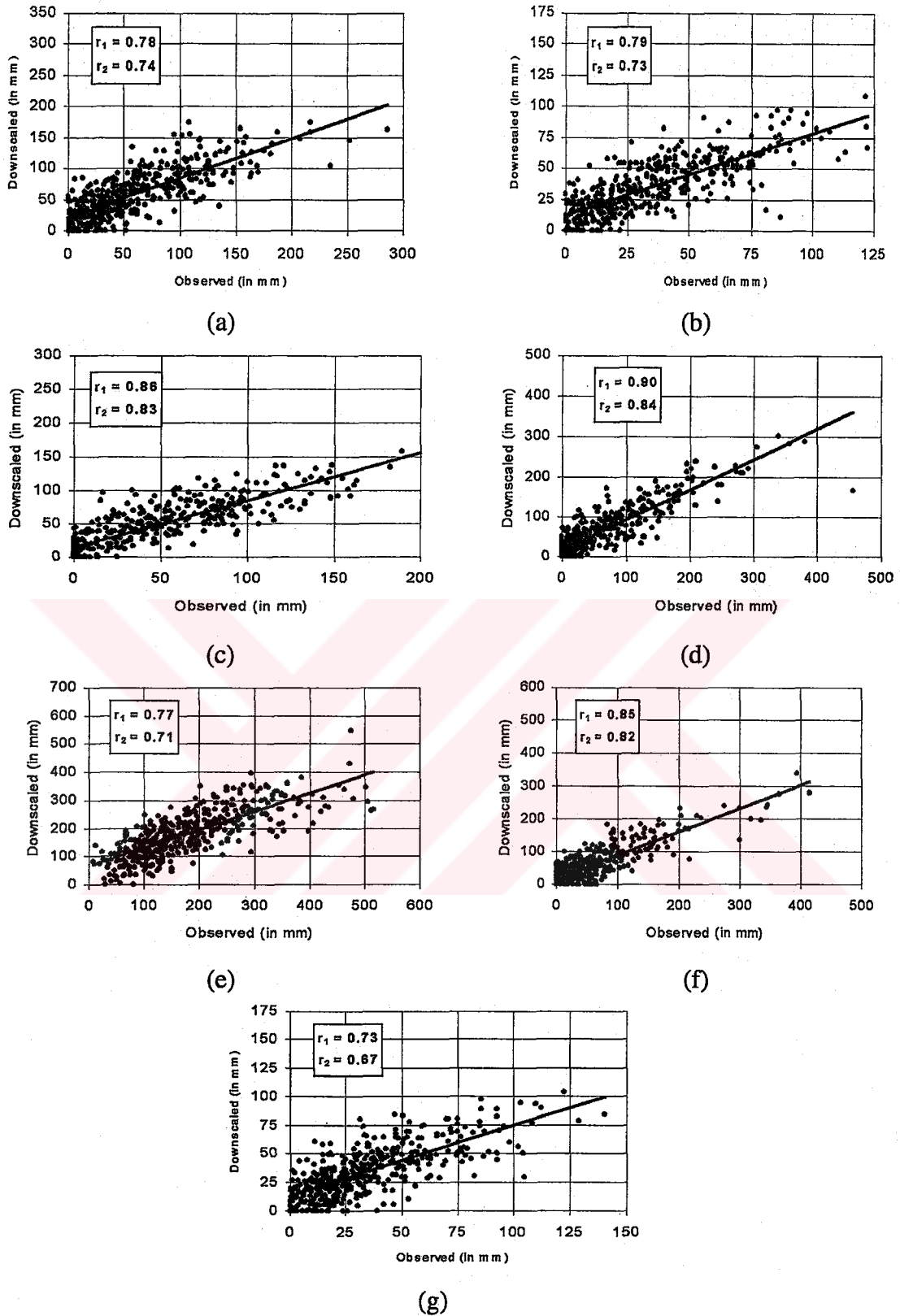


Figure 5.12. The scatter plots of the identified model outputs versus actual monthly total precipitation series. The correlation coefficients between the observed and model outputs are abbreviated as  $r_1$  for the training part and  $r_2$  for non-training part respectively: (a) Göztepe (İstanbul); (b) Ankara; (c) Diyarbakır; (d) İzmir; (e) Rize; (f) Adana; (g) Erzurum



#### 5.1.4 Summary

In this section, a new proposed method which is introduced in the Section 3.2 for downscaling regional climate processes is presented. The model results show that the precipitation regime (both wet and dry periods) of the coastal regions for Turkey (the Mediterranean, Aegean and Marmara), and region of southern Anatolia are under the influence of large scale pressure system and upper-air circulation.

On the other hand, especially in the Black Sea region, in addition to the large scale processes, the local features (namely topographical conditions) determine the likelihood and intensity of precipitation. For inland regions, the local processes are more effective than the large scale processes. The southeastern part of the country is affected by both the Mediterranean and Monsoon lows. Therefore, this region can be called a Transition Mediterranean precipitation regime, which is in agreement with the works of Türkeş (1996, 1998).

Before model building, a large scale analysis is needed because of the relationships between the large scale processes and local processes based on the meaningful statistical linkages. However, according to many studies (*e.g.*, Nicholls, 1987; Özeren *et al.*, 2003; Türkeş and Erlat, 2003) the other climate indices, such as North Atlantic oscillation, may be considered to relate remote effects on the local precipitation series.

The model is constructed following a three-stage process. In the first stage, the potential predictors are preprocessed with PCA based on the maximum likelihood criterion of Jöreskog and Sörbom (1989). In this way, 1105 time series (five type predictors on 221 grid points) are reduced to 17 time series of PCs (Table 5.2).

If the data in use do not have a Gaussian distribution, as is the case of the precipitation series, then CCA after PCA may not able to produce true correlated components; therefore, the significant PCs are transformed into ICs by the ICA procedure to satisfy the probability distribution constraint. Since a prediction technique's performance is sensitive to the independence of the predictors; a better performance of CCA with ICs is not a surprising result according to this study. After employing CCA, six significant CCVs of ICs of the predictor time series are produced based on their explanatory performance of the maximum variances in the precipitation series for Turkey stations (Table 5.3).

In the second stage, a causal model is identified between the monthly total precipitation series and CCVs of ICs of the large scale data sets. Thereafter, a naive model is identified as a first order model (Richardson, 1981). In this step, the causal and naive models are run in parallel (or online), and the outputs of the models are transformed into the basis functions by the hyperbolic tangent function (see Section 3.2) to satisfy nonlinearity requirements.

In the third stage, a classical multivariate regression model is built between the extracted basis functions in the second stage and the set of the monthly total precipitation series. Finally, for the proposed RNN, the training through time process is applied to the naive and causal components of the model separately. The method is equivalent to the Jordan-type recurrent neural network in terms of its structure. The final model is so simple in its structure that it can easily be updated when new data sets are available.

Generally, we have two types of knowledge about large scale processes: those coming from the models and those inferred from the observations. Knowledge from the observations may be divided into three parts: physical based linkages, such as weather types (Lamb, 1972; Conway and Jones, 1998); 'summarized' knowledge, such as climate indices; and time series of large scale observations or GCM-generated fields. In this study, we only developed statistical linkages based on the large scale NCEP–NCAR reanalysis data sets.

Finally, one has to emphasize that the performance of the resulting prediction scheme (grey = white + black) depends not only on the performance of the statistical (black-box) downscaling techniques, but also on the GCM's (white-box) performance in simulating large scale fields.

## **5.2 Surface Air Temperature Variability over Turkey and Its Connection to the Large Scale Upper Air Circulation**

### **5.2.1 Problem statement**

In this Section, the problem of statistical linkages between the large scale and local scale processes is investigated through noise-reduction by singular spectrum analysis and spatial principal component analysis in order to construct an appropriate statistical model to downscale the local scale climate variables from the large scale

climate processes for Turkey stations. This study presents an approach for downscaling monthly near-surface air temperature series over Turkey by the upper air circulations from NCEP-NCAR reanalysis data sets (namely 500 hPa geopotential heights and 500-1000 hPa thicknesses) which are windowed the range of 10-50°E and 30-60°N where additional details were previously shown in Figure 2.1 and Figure 2.2b in Section 2.

The proposed approach procedure consists of three stages. First, the available data sets are separated into deterministic and statistical components), and random components by SSA.

Secondly, the deterministic components are saved and the random components are eliminated by spatial PCA. Later, the statistical components and deterministic components are combined of which the new data set is called noise-free. Secondly, so called Sampson correlation patterns are determined between the noise-free large scale and local scale variables for interpreting the large scale processes impacts on local scale features for Turkey stations.

Thirdly, the significant redundancy variates based on CCA are extracted in order to identify the statistical-downscaling model for the temperature series of 62 Turkey stations (Table 5.4). The results show that the interpretation of the local scale processes with the noise-free data sets is more significant than the raw data sets.

A number of applications regarding to the temperature series of Turkey associated with the statistical analysis have been published, *e.g.*, (Jones *et al.*, 1986; Jones, 1995; Kadioğlu, 1997; Tayanç *et al.*, 1997; Kömüştü, 1998; Türkeş *et al.*, 1995; Türkeş *et al.*, 2002b).

The predictands are monthly maximum, mean and minimum temperature series of 62 stations in Turkey, during the period of 1951 to 1998. Since data sets of the temperature time series have dominantly periodic cycles these components can easily be estimated by any linear prediction model. On the other hand, these periodic components may hide the non-seasonal components which are called non-periodic or/and trend components (linear or non-linear) according to the stochastic literature (Richardson 1981; Bras and Rodriguez-Iturbe 1993; Box *et al.*, 1994; Hipel and Mcleod 1994; von Storch and Zwiers 1999), respectively.

Table 5.4. Turkish stations for near-surface air temperature used in the application

Station Name	Longitude (°E)	Latitude (°N)	Station Name	Longitude (°E)	Latitude (°N)
Zonguldak	31.48	41.27	Afyon	30.32	38.45
Inebolu	33.46	41.59	Kayseri	35.29	38.44
Sinop	35.10	42.01	Malatya	38.19	38.21
Samsun	36.18	41.17	Elazig	39.14	38.40
Giresun	38.24	40.55	Siirt	41.57	37.55
Trabzon	39.43	41.00	Izmir	27.10	38.26
Rize	40.31	41.02	Aydin	27.51	37.51
Artvin	41.49	41.11	Burdur	30.17	37.43
Edirne	26.34	41.40	Isparta	30.33	37.46
Tekirdag	27.33	40.59	Konya	32.29	37.52
Kirecburnu	29.03	41.10	Nigde	34.41	37.58
Goztepe	29.05	40.58	Gaziantep	37.23	37.04
Kocaeli	29.56	40.47	Sanliurfa	38.46	37.08
Bolu	31.31	40.44	Mardin	40.44	37.18
Kastamonu	33.47	41.22	Diyarbakir	40.14	37.54
Merzifon	35.20	40.52	Bodrum	27.26	37.03
Corum	34.57	40.33	Mugla	28.22	37.13
Sivas	37.01	39.45	Fethiye	29.07	36.37
Erzincan	39.30	39.45	Anamur	32.50	36.05
Erzurum	41.16	39.55	Mersin	34.36	36.48
Agri	43.03	39.43	Adana	35.20	37.00
Bandirma	27.58	40.21	Iskenderun	36.10	36.35
Bursa	29.04	40.11	Luleburgaz	27.21	41.24
Bilecik	29.59	40.09	Sile	30.25	40.47
Yozgat	34.48	39.49	Florya	28.45	40.59
Balikesir	27.53	39.38	Kutahya	29.58	39.25
Van	43.23	38.30	Dortyol	36.13	36.51
Dikili	26.53	39.04	Islahiye	36.38	37.02
Akhisar	27.51	38.55	Antakya	36.10	36.12
Manisa	27.26	38.37	Eskisehir	30.31	39.49
Usak	29.24	38.41	Ankara	32.53	39.57

The variability of climate may usually be in the non-seasonal components which may show non-stationary characteristics. Hence, SSA is applied to both the large scale predictors and local scale predictands before preprocessing the data sets by spatial PCA in order to extract the predictable components. In SSA procedure, if the length of a univariate time series is  $N$  then the lag-window is generally selected between  $N/3$  and  $N/4$  interval according to Elsner and Tsonis (1996). Accordingly, in this study, the selected lag-window is  $N/4$ .

Both the NCAR/NCEP reanalysis data sets and observations are divided into four seasons in order to study the relationships between the temperature series and large scale processes. To test the performance of the proposed model, the first 360 months



(30 years) are used during the model calibration and the other 216 months (18 years) are used for model validation.

### 5.2.2 Sampson correlation pattern analysis of Turkish surface air temperatures associated with upper air circulation

The relationships between monthly near-surface air temperature series in Turkey and upper air circulations (500 hPa geopotential heights and 500-1000 hPa geopotential thicknesses) are shown in Figures 5.13-5.18 as Sampson correlation patterns in the seasonal-scale months for four seasons.

As depicted in these figures, based on the 99% ( $R_s > 3$ ) significance level of correlations between maximum temperatures and 500 hPa geopotential heights, the area that includes the southern and western parts of Turkey is at appreciable levels in summer and spring.

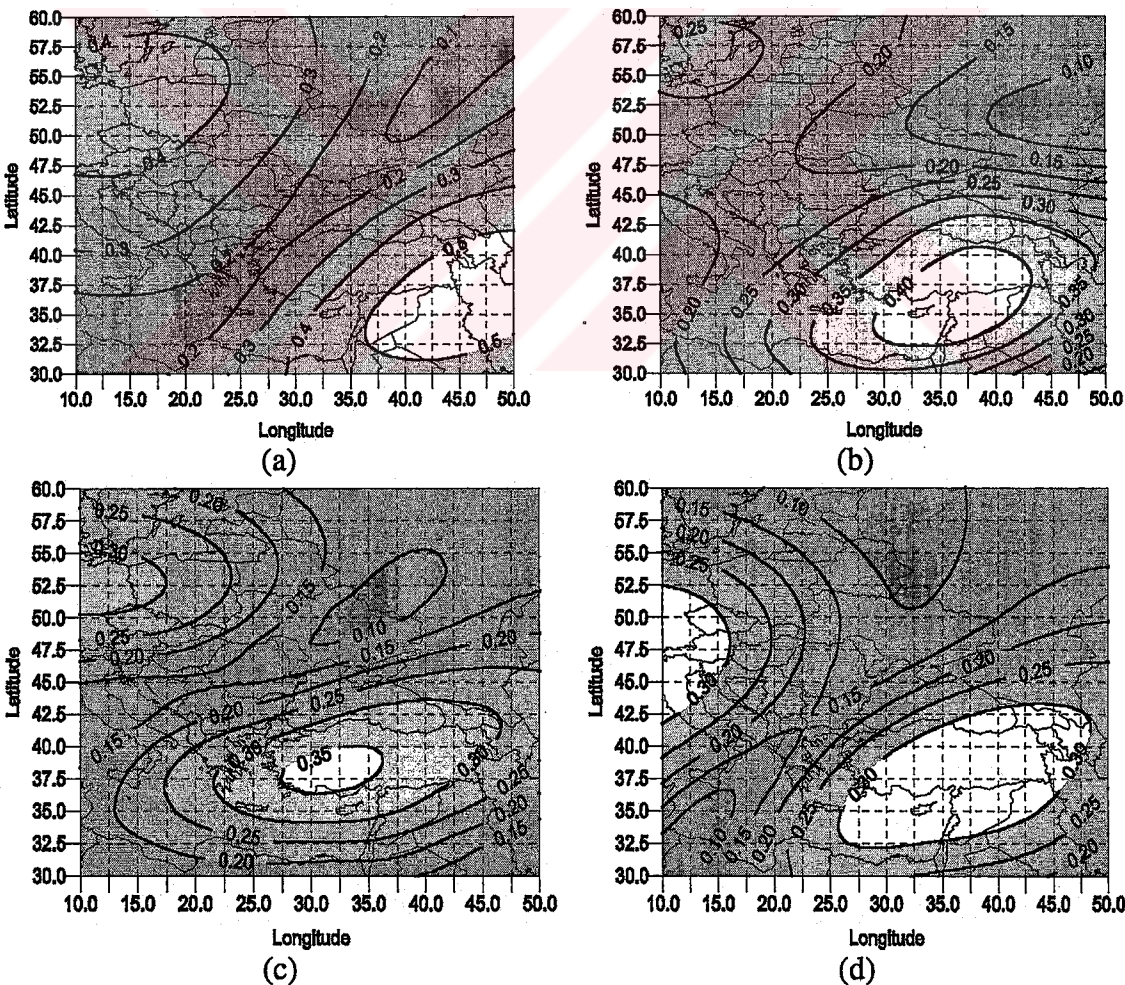


Figure 5.13.  $R_s$  pattern between the 500 hPa geopotential heights and maximum temperature series: (a) winter; (b) spring, (c) summer; and (d) autumn months



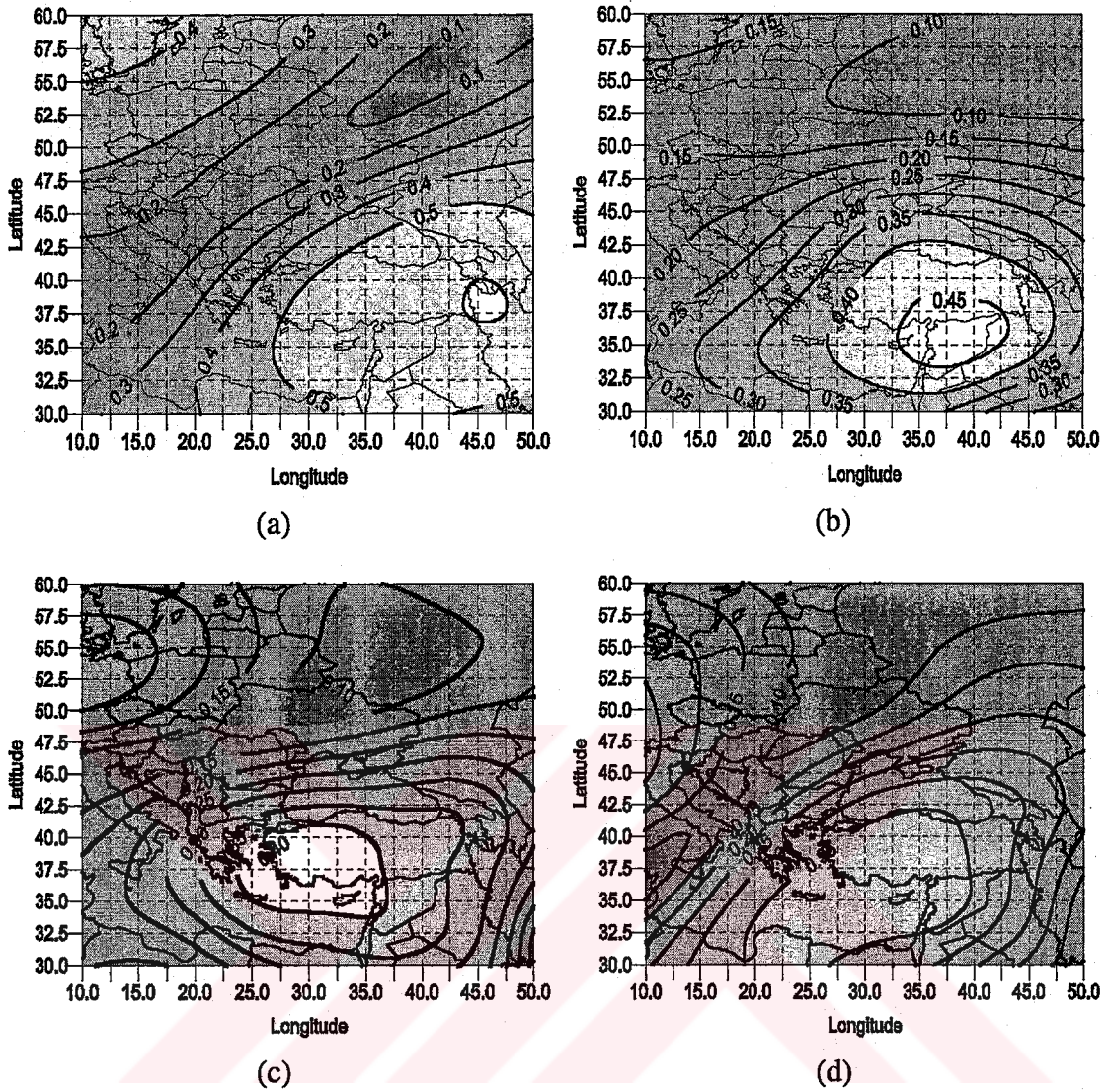


Figure 5.14.  $R_s$  pattern between the 500-1000 hPa thicknesses and maximum temperature series: (a) winter; (b) spring; (c) summer; and (d) autumn

Since in the beginning of spring, Turkey gets under the influence of warm sub-tropical air systems, 500 hPa geopotential heights are transformed into ridges over the corresponding locations. This leads to an increase of maximum temperatures. On the other hand, while the 500 hPa high center extends its influence over Europe, further effects on maximum temperatures arise. Displacing of the Azores high center over Europe leads northerly further east currents in summer. As a result of this mechanism, the maximum temperature decreases especially in the regions of western and northern parts of Turkey.

During winter, sub-tropical air systems just affect the southern parts of the country. The second effect is seen from central and northern Europe. In this season, the Iceland low pressure and Siberian high pressure systems start to affect northern and

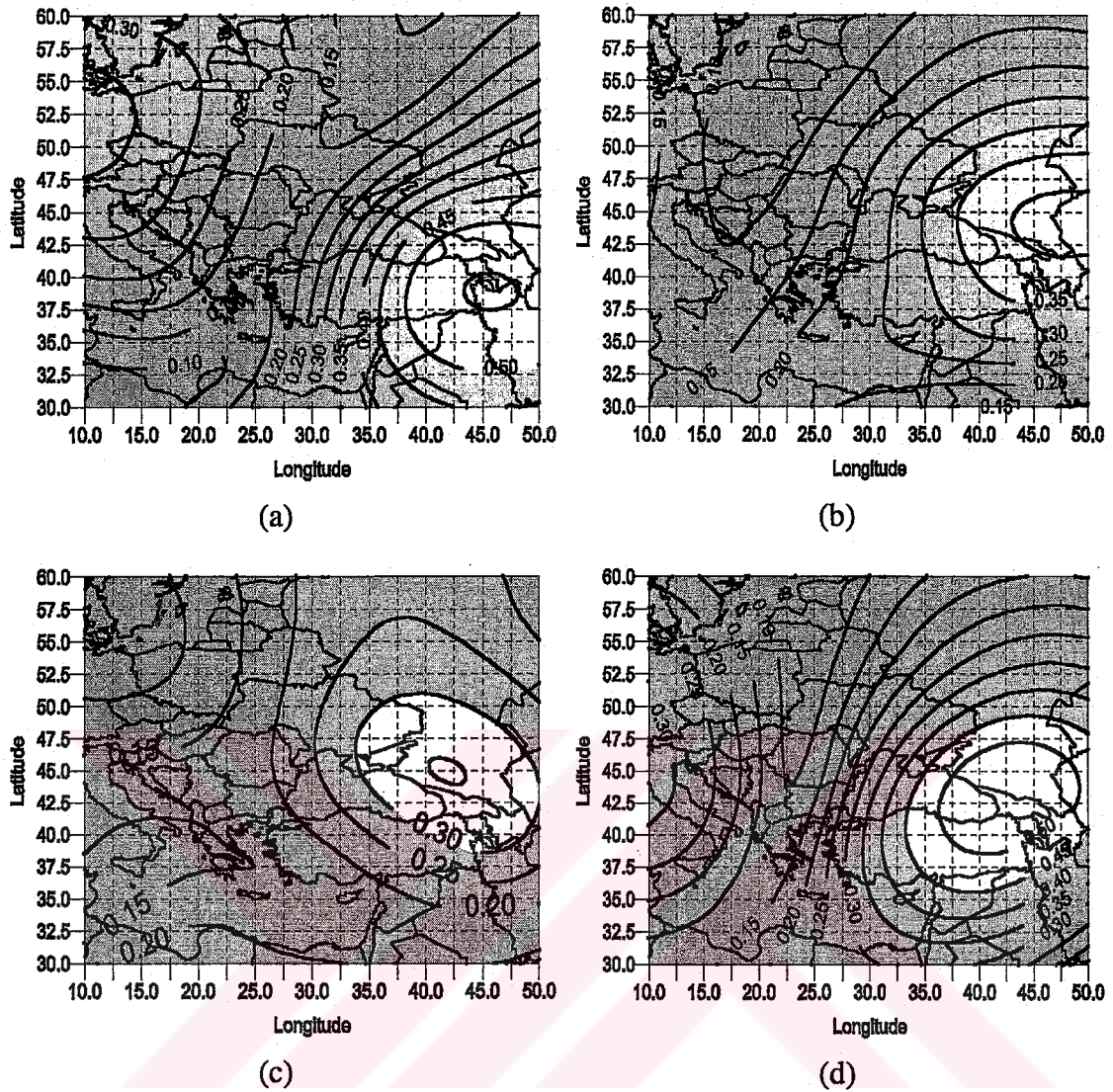


Figure 5.15.  $R_s$  pattern between the 500 hPa geopotential heights and minimum temperature series: (a) winter; (b) spring; (c) summer; and (d) autumn

eastern Europe (due to dynamical and thermal reasons) since both systems have cold characteristics leading a decrease of 500 hPa geopotential heights. The northerly currents have effects on decreasing maximum temperatures in Turkey (especially in northern parts). The effect of the Icelandic low is mentioned through its connection with the North Atlantic Oscillation (NAO) during winters (Türkeş and Erilat, 2003).

The effects of large scale upper air circulations on maximum, minimum and mean near-surface air temperatures show similar patterns during the winter season. On the other hand, those effects mostly resemble to the  $R_s$  correlation patterns for maximum and mean temperatures, whereas the minimum temperature series show very different patterns in other seasons (summer, autumn and spring).



In autumn, the influenced areas by sub-tropical air systems which are effective during summer are shifted towards southern and south-eastern part of Turkey. According to 500 hPa geopotential heights, another area of influence is linked to the Azores high affects the Western Europe encompassing Italy and France.

There are strong correlations between 500-1000 hPa thicknesses and maximum-mean temperatures. The thicknesses have positive effects to increase the Turkish maximum-mean temperatures, starting in spring and continuing in summer. In this season, the movement of the tropical (Monsoon) low to higher latitudes over the eastern and south-eastern part of Turkey (including the eastern and central Black Sea part of Turkey) has important effects for those regions. The warm features of this air system lead to increased thicknesses and ridges at the 500 hPa level.

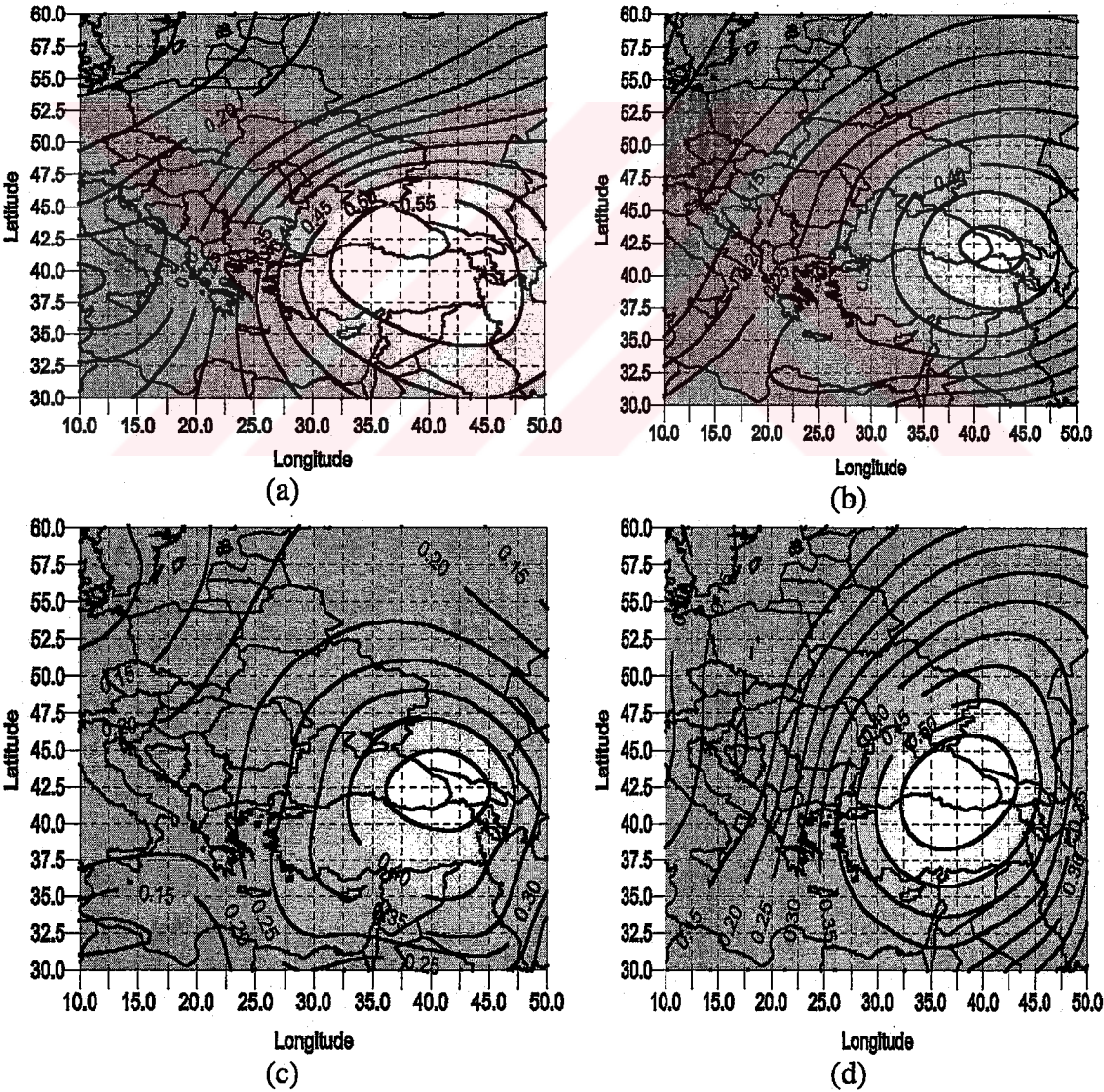


Figure 5.16.  $R_s$  pattern between the 500 hPa-1000 hPa thicknesses and minimum temperature series: (a) winter; (b) spring); (c) summer; and (d) autumn



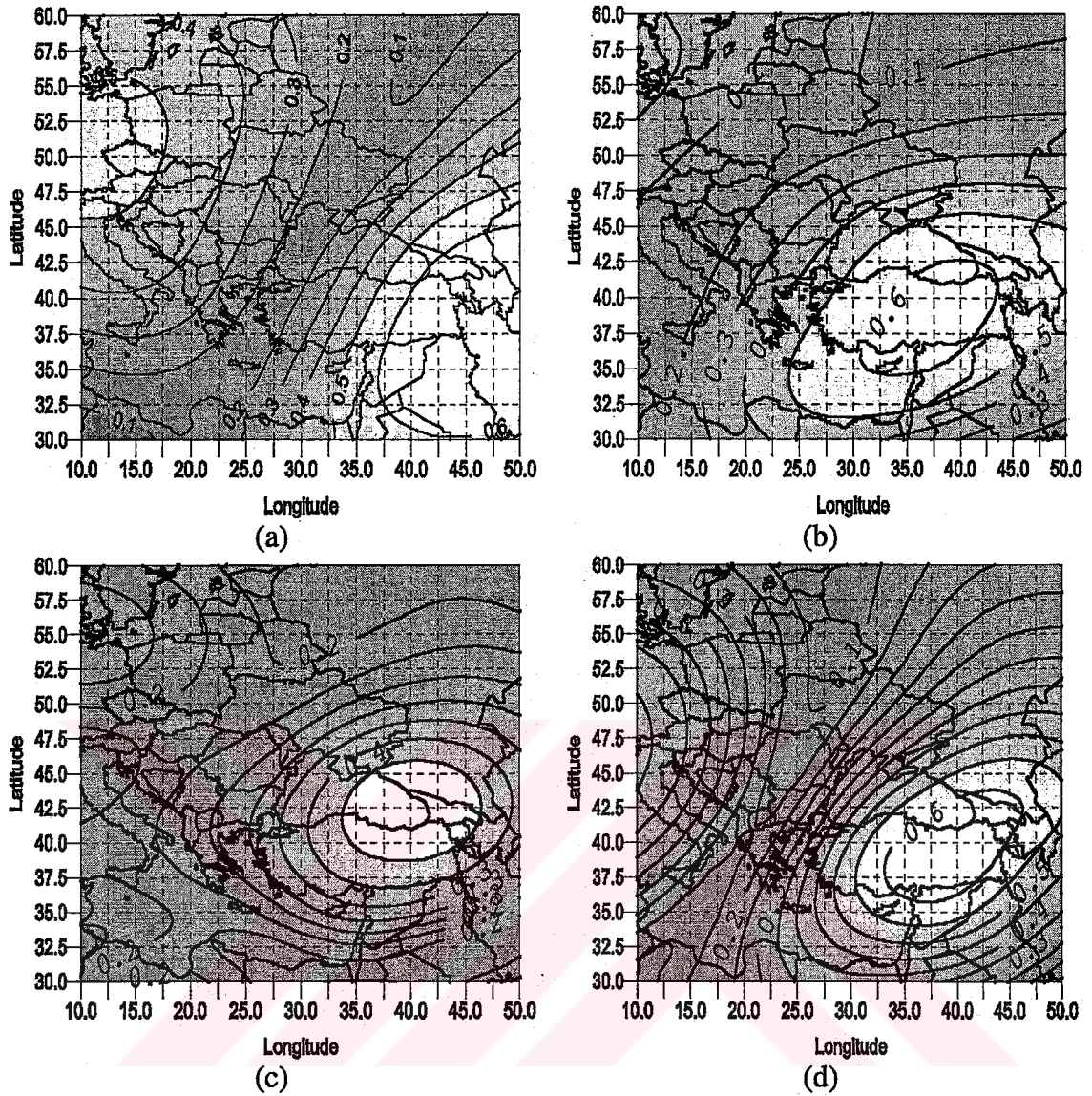


Figure 5.17  $R_s$  pattern between the 500 hPa geopotential heights and mean temperature series: (a) winter; (b) spring; (c) summer; (d) autumn

In autumn, tropical air systems shift to the south as a result of cooling land areas. In summer, the significant correlations are observed over north-western parts of Europe. Those areas and the influence of Azores high center's area are overlapping.

As may be seen in these figures, the discussion of  $R_s$  patterns for the maximum and mean temperatures is also valid for the minimum temperature series during winter.

However, in summer and the transition seasons, the locations of significant correlations are different since the large scale circulation features of winter are more effective than these of three seasons. In other words, Turkey is under the influence of large scale systems in winter, hence similar correlation patterns are expected for all three temperature types. Minimum temperature patterns, in addition to large scale

features, are also affected by so-called local topographical conditions. In spring, summer and autumn, the significant correlations are observed in eastern and north-eastern parts of the country.

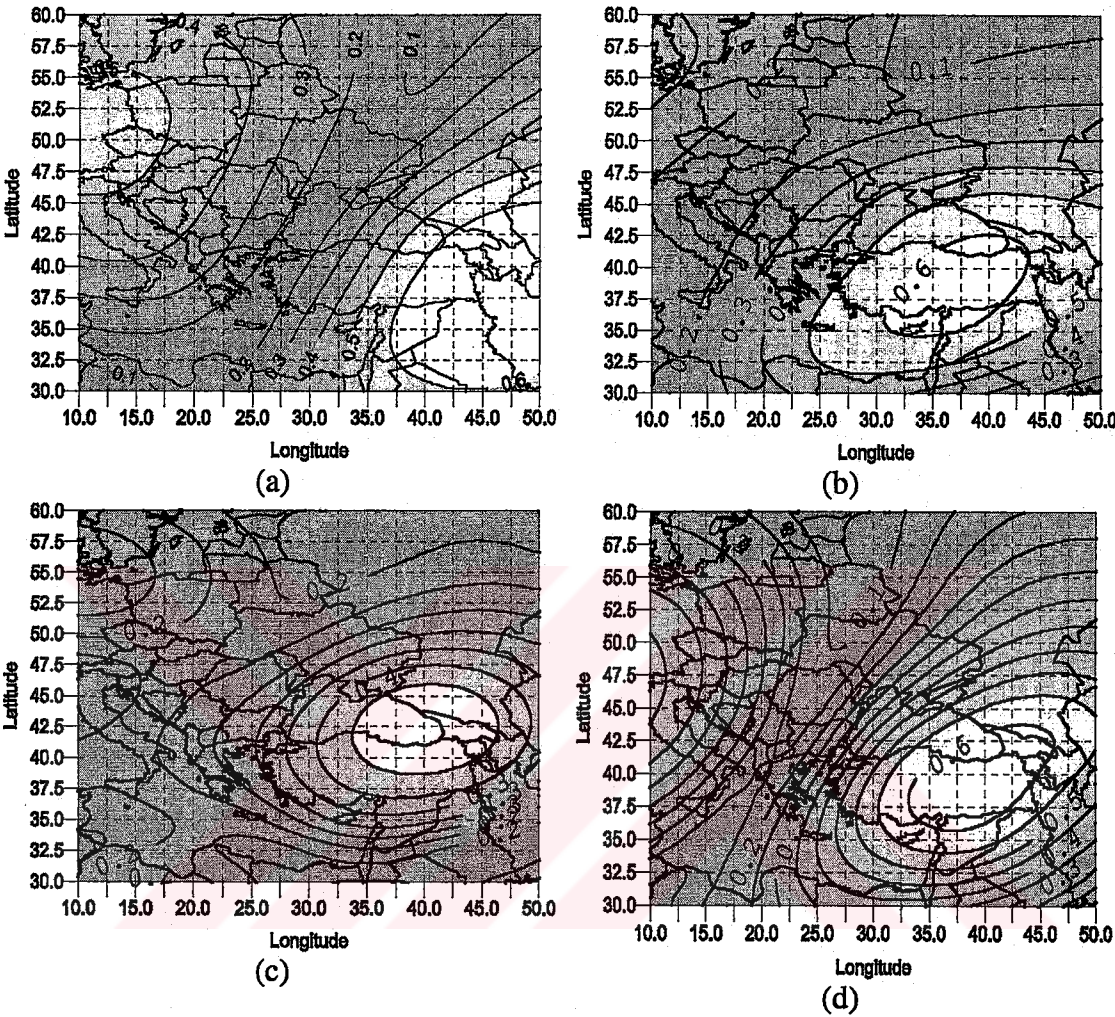


Figure 5.18.  $R_s$  pattern between the 500-1000 hPa thicknesses and mean temperature series: (a) winter; (b) spring; (c) summer; and (d) autumn

### 5.2.3 Downscaling results

To show spatial performance of the model, the significant varimax rotated (Kaiser 1959) principal components of both the downscaled and the observed monthly temperature series are shown in Figures 5.19-5.21; additionally Figure 5.22 demonstrates which parts of the country's temperature fields can be estimated from the large scale upper-air circulations by the proposed downscaling approach by the means of the mean square error geographical distribution.



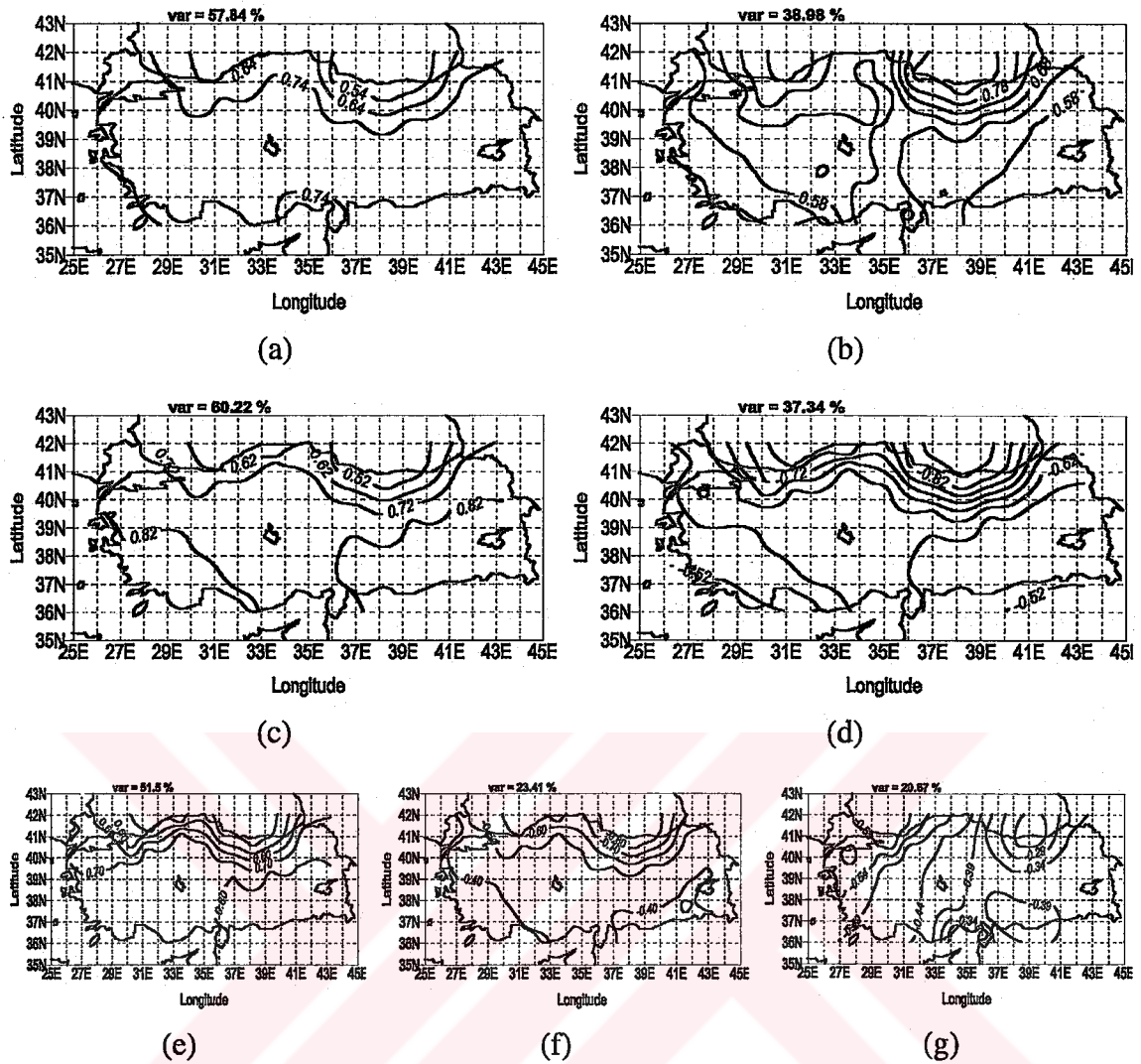


Figure 5.19. The varimax rotated principal component patterns of the downscaling model outputs and the actual maximum temperature: a and b show the first two PC patterns of noise-free data; c and d indicate the first two PC patterns of model outputs; e, f, and g indicate the first three PC patterns of actual raw data. The third PC patterns of both noise-free and model-output data are not significant in contrast to the third PC pattern of raw data (the patterns represent correlation coefficients between individual station's temperature series and rotated principal components. These patterns are generally called loadings). The period of the extracted principal components is the period of validation period of the suggested model (18 years data).

As may be seen in these figures, without noise-reduction procedure the number of the significant PCs is increased, only for maximum temperature series. However, the third PC of maximum temperature series can not be connected with the large scale processes; therefore noise-free data sets may be seen more appropriate for downscaling purposes. As seen in these figures, temperature variability over Turkey is represented by two PC patterns. One indicates seaside regions and the other

represents inland regions. The similar patterns are also extracted from proposed downscaling model.

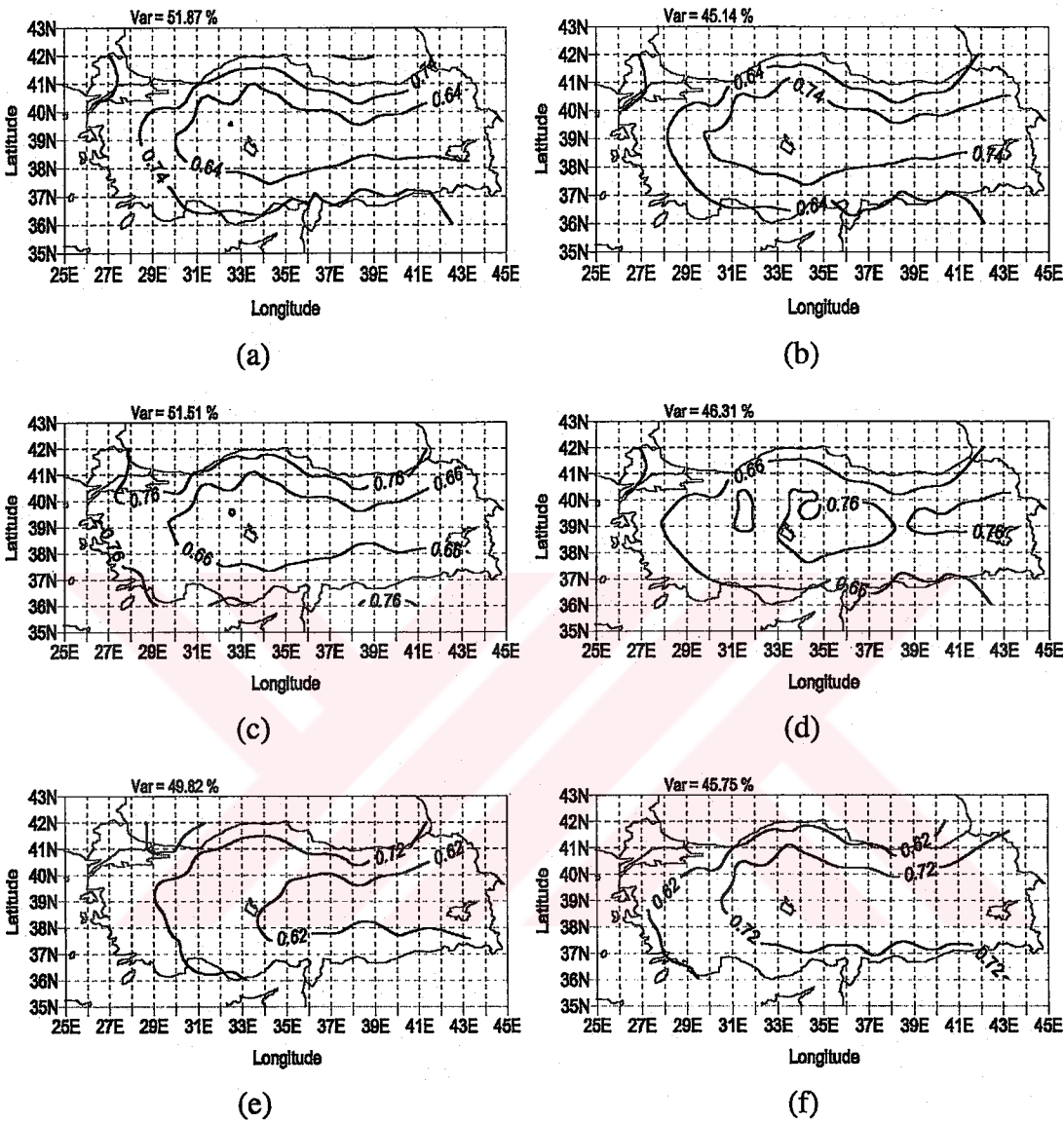


Figure 5.20. The varimax rotated principal components of the downscaling model outputs and the actual minimum temperature: a and b show the first two PCs of noise-free data; c and d indicate the first two PCs of model outputs; e and f indicate the first two PCs of actual raw data. The third PCs of the noise-free, model-output, and actual raw data are not significant.

The other temperature series (maximum and mean) generally follow general circulation patterns but minimum temperature series can also show local scale characteristics. In order to show proposed model performance, seven stations' minimum temperature series are selected according to Tatlı *et al.* (2004), and results are given in Figure 5.23. This figure is a 95% confidence of frequency scatter plot of the proposed model outputs versus observed series during the calibration of the

model; and Figure 5.24 is prepared to show the performance of the model for the validation part of the data.

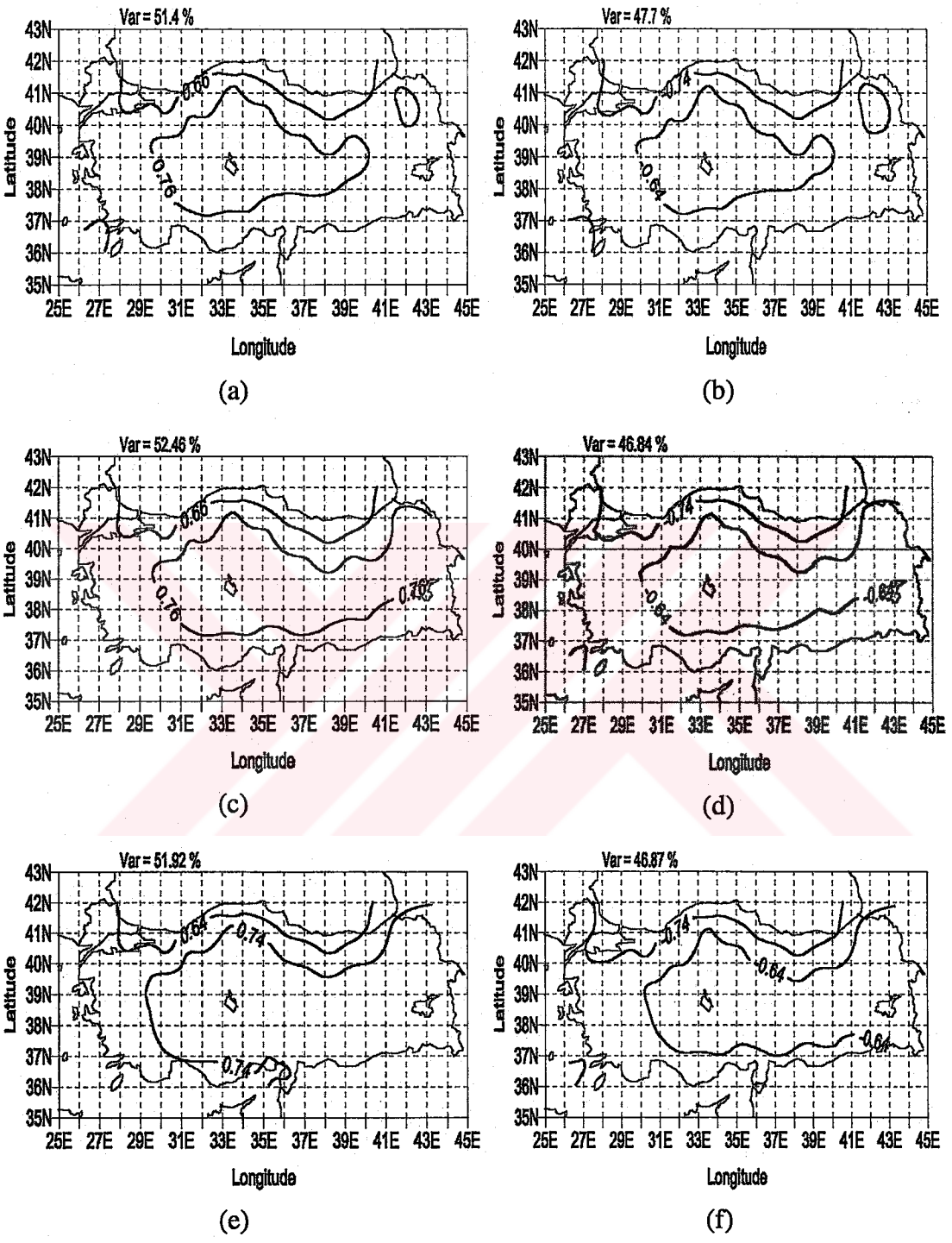


Figure 5.21. The varimax rotated principal components of the downscaling model outputs and actual mean temperature: a and b show the first two PCs of noise-free data; c and d indicate the first two PCs of model outputs; e and f indicate the first two PCs of actual raw data. The third PCs of the noise-free, model-output, and actual raw data are not significant.

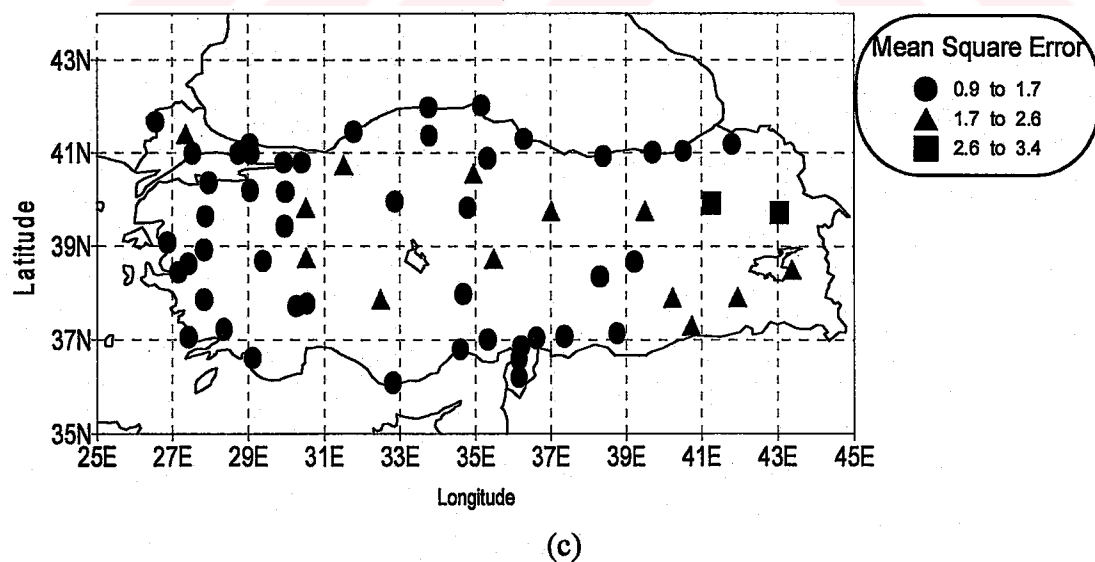
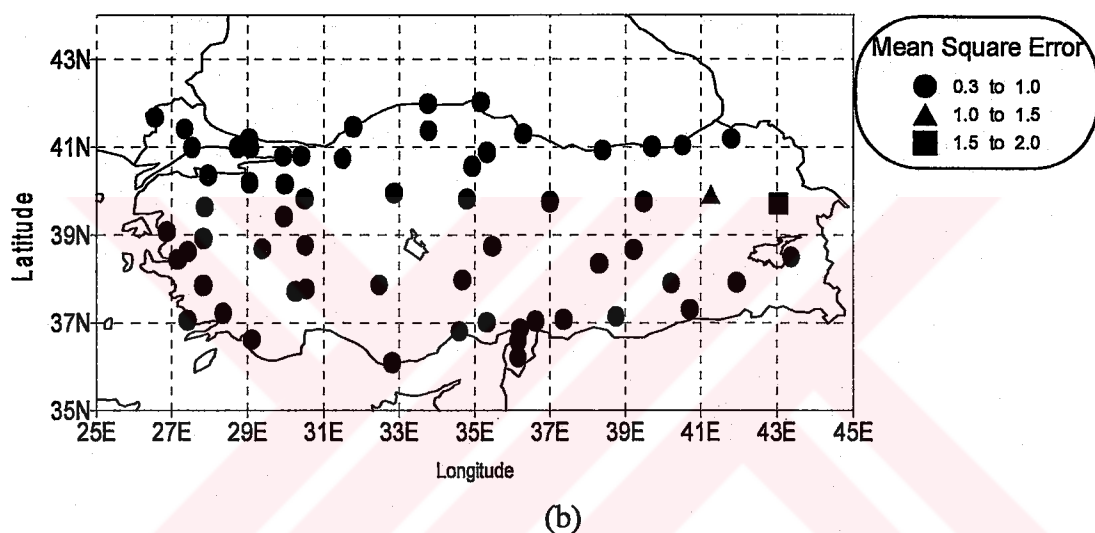
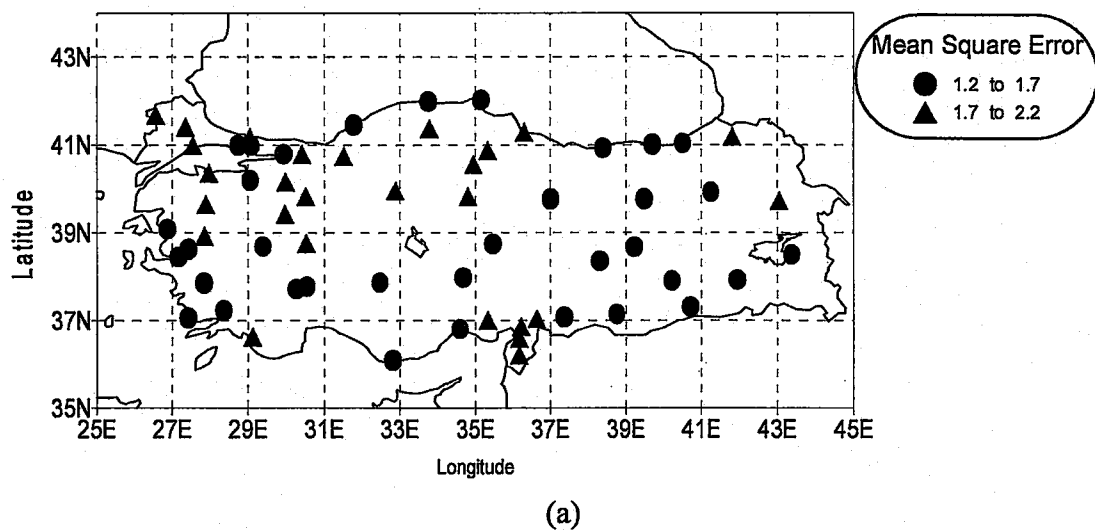


Figure 5.22. Geographical distribution of the mean square errors during the period of 1980-1998 (validation part): (a) Maximum temperatures; (b) mean temperatures; and (c) minimum temperatures.



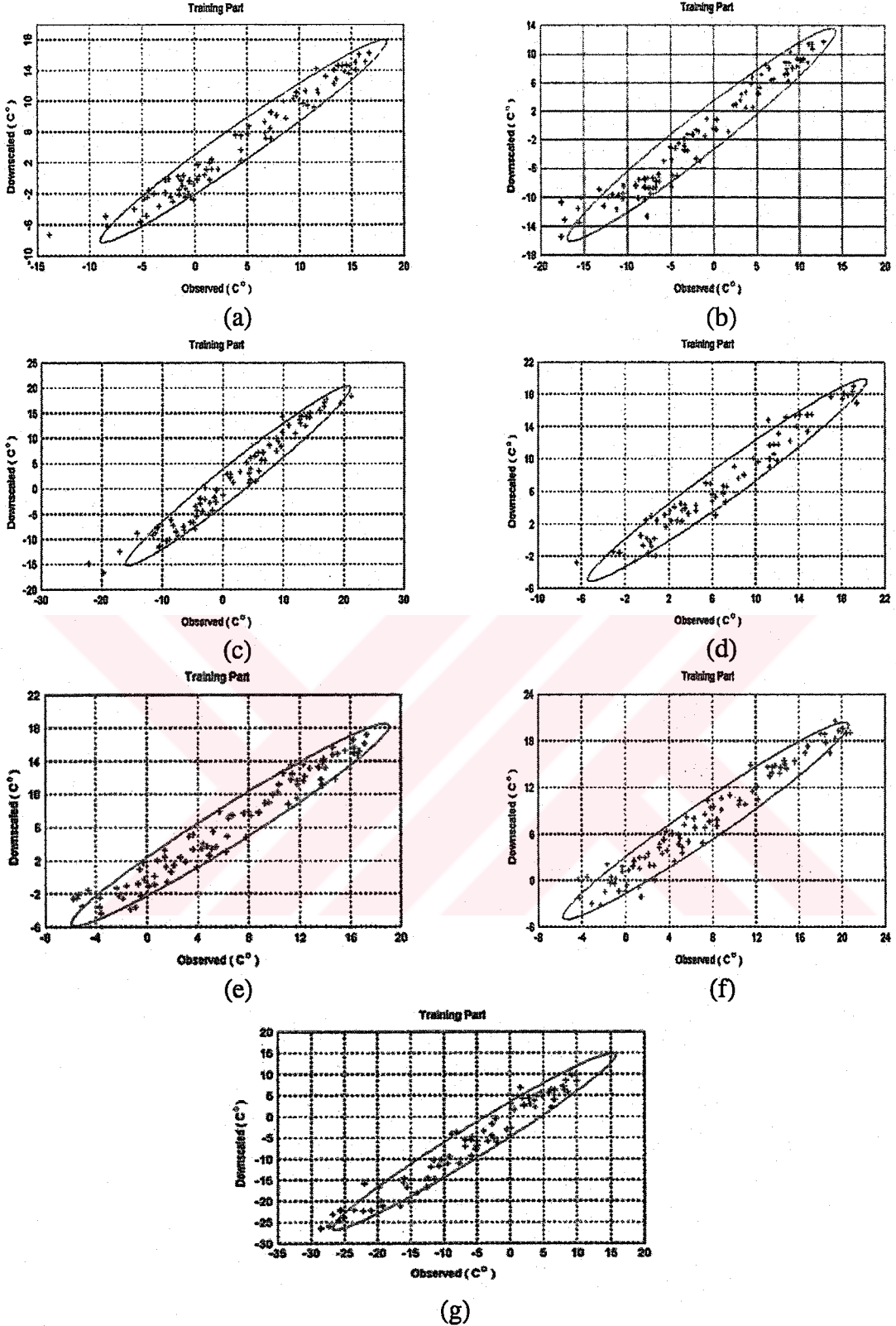


Figure 5.23. The frequency scatter plots of the minimum temperatures of the observations and model outputs during the calibration period of the proposed model. The observations are the series of which are not filtered by SSA. The ellipses indicate the 95% confidence interval, and the data in the outside of the ellipses is the unpredictable part of the data by the proposed model: (a) Göztepe (Istanbul); (b) Ankara; (c) Diyarbakır; (d) İzmir; (e) Adana; (f) Adana; (g) Erzurum.

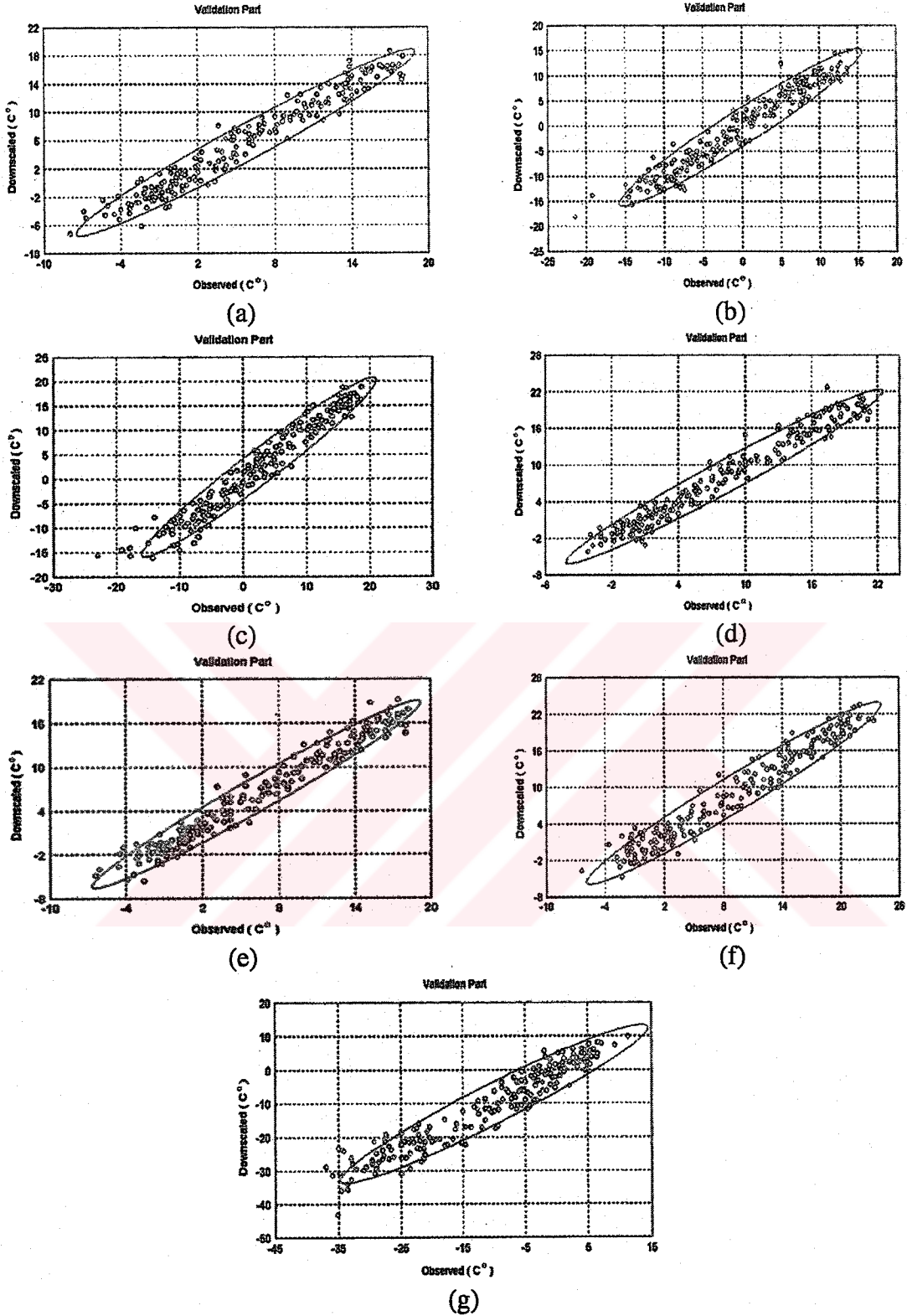


Figure 5.24. The frequency scatter plots of the observed minimum temperatures and model outputs during the validation period of the proposed model. The observations are the series which are not filtered by SSA. The ellipses indicate the 95% confidence interval, and the data in the outside of the ellipses is the unpredictable part of the data by the proposed model. The entire data of the validation period is 216 months (18 years). As seen in Figures 5.23 and in this figure, the outliers of the both calibration and validation part data show parallel features according to 95% significant level: (a) Göztepe (Istanbul); (b) Ankara; (c) Diyarbakır; (d) İzmir; (e) Adana; (f) Adana; (g) Erzurum.

#### 5.2.4 Summary

In this section, the problem of statistical linkages between the monthly near-surface air temperature series over Turkey and large scale upper air circulations from NCEP-NCAR reanalysis data sets are investigated by a particular downscaling approach based on multivariate redundancy and Sampson correlation analysis during the period of 1951-1998. The noise-reduction techniques of singular spectrum and principal component analysis are employed in order to filter data mention as the deterministic (trend and/or cycles) and the statistical components.

The suggested approach shows that the effects of the large scale upper air circulations on monthly maximum, minimum and mean temperature series show similar patterns in winter. On the other hand, these effects resemble in the Sampson correlation patterns for monthly maximum and mean temperatures whereas the minimum temperature series show very different patterns in other seasons (summer, autumn and spring). In the autumn, the influence areas of sub-tropical air systems which are effective in summer form a belt towards southern and south-eastern regions of Turkey.

The Sampson correlation patterns indicate that the correlations between maximum temperatures and 500 hPa geopotential heights in the area covering the southern and western parts of Turkey are at significant level in summer and spring. Since Turkey in under the influence of warm sub-tropical air systems in early spring, the geopotential heights having ridge characteristics lead to an increase in temperatures. On the other hand, while 500 hPa high centers extending towards Europe, and therefore, its signature may be seen as the variability of the temperature fields. In summer, displacement of Azores high center over Europe generates northerly current flows lead to decrease temperature especially in the regions of western and northern parts of Turkey.

The Sampson correlation pattern between 500-1000 hPa thickness and temperature series shows that the thickness increases the Turkish temperature starting in the spring and continuing in summer. In this season, the moving of the sub-tropical air system (Monsoon low) to high latitudes over eastern and south-eastern parts of Turkey (includes eastern and central Black Sea region of Turkey) has important

effects. The warm features of this air system lead to increase thicknesses and create ridges on 500 hPa levels.

Additional effects of thickness anomalies are observed for Caucasus and Caspian where the cold advection invades the topographic depression between Balkans highlands and the Caucasus, and even across the Anatolian upland and the Balkans, this cold air being derived mainly from Russian and Balkan sources.

The methods applied in this study are selected according to the large scale and local scale climate variability and asymmetric relationships. Data assimilation or upscaling is indeed a processing of convolution between large scale and local scale variables. The problem of reconstructing the local scale variables is, on the other hand, an inverse problem. In more general terms, the deconvolution problem is to reconstruct the inputs (local scale) of a climate system (represented by a GCM) from its outputs (large scale). However, the interactions of large scale and local scale climate features make the processes separation highly nonlinear to a great extent. In the majority of downscaling studies it is assumed that it is possible to identify the local scale variables by means of suitable analysis of free troposphere variables, for instance 500 hPa geopotential heights and 500-1000 hPa geopotential thicknesses which are also used in this study. These variables can be simulated as perfectly as by GCMs might be considered.

For further case studies over Turkey, the design of proposed model in this study may be bridged by the nested models in order to extract climate impacts with physical-based model approaches. Additionally, during the parameterization of nested models, the results of the proposed analyses methods may be supportive to explain the linkages between the local scale and large scale climate variability and the atmospheric disturbances. The proposed model ability to downscale the local scale climate features from a nested model for Turkey is still one of the open questions in this work.

## 6. SUMMARY AND CONCLUSIONS

In this study, from the presentations and discussions of the proposed downscaling strategies, it is clear that, there is no valid universal downscaling method for all variables and for all regions. Instead, statistical downscaling requires the design of statistical models on a case by case basis.

The strategy for downscaling Turkish rainfall, the proposed method is more complex than the method for downscaling surface air temperature series over Turkey. For instance, local scale climate characteristics have more influences on precipitation variations. The first model results show that the precipitation regime (both wet and dry periods) of the coastal regions of Turkey (Mediterranean, Aegean, Marmara, western Black Sea) is under the influence of large scale pressure system and upper-air circulation. However, especially in the Black Sea region, in addition to the large scale processes, the local features (namely topography and rain-shadows) determine the likelihood and intensity of precipitation.

For inland regions, the local processes are more effective than the large scale processes. The southeastern part of the country (particularly Diyarbakır) is affected by both the Mediterranean and Monsoon lows. Therefore, this region could be called a Transition Mediterranean precipitation regime.

Before identifying the first proposed downscaling model, the relationships between large scale variables (namely 500 hPa and 700 hPa geopotential heights, 500-1000 hPa geopotential thicknesses, mean sea level pressures and 500 hPa vertical pressure-velocities) and local scale monthly total precipitation over Turkey are explored by Sampson correlation patterns; where the large scale variables are extracted from National Centers for Environmental Prediction-National Center for Atmospheric Research reanalysis data sets (Kalnay *et al.*, 1996) windowed between 10-50°E and 30-60°N range. Sampson correlation patterns show that there are meaningful correlations between precipitation series and both 500 and 700 hPa geopotential heights, especially, in winter months in the area that includes the western parts of



Turkey. However, in summer months, significant but weaker correlations exist in the Eastern Mediterranean (namely southern Turkey and Cyprus).

The significant correlations associated with thickness anomalies reflect the thermal relations located in Central Europe (in a similar way to the 500 and 700 hPa geopotential heights). Additional effects of thickness anomalies are observed for Caucasus and Caspian where the cold advection invades the topographic depression between the Balkans Highlands and the Caucasus, and even across the Anatolian upland and the Balkans, this cold air being derived mainly from Russian and Balkan sources.

In summer months, the effect of air flow which is originated from Sahara (dry and warm) is clearly seen in the correlation patterns of thicknesses and precipitation. The very important correlations are observed between precipitation and both the mean sea level pressure and 500 hPa vertical pressure-velocities from reflecting Sampson correlation patterns. Those results are expected because precipitation occurrence situations are dominantly related such upward and surface weather systems, and those features are characterized being the sources of frontal systems.

In the study, the results of proposed downscaling outputs are compared with the actual precipitation series by extracting the pseudo spatial principal components and mean square errors distribution. The eigenvector patterns show that the proposed model outputs are consistent with the actual precipitation series. The first model performance is superior to the conventional techniques (a list of conventional techniques may be seen in Wilby and Wigley (2000)) because the proposed statistical model is constructed after three stages. In the first stage, the potential predictors are preprocessed with principal component analysis based on maximum likelihood criterion of Jöreskog and Sörbom (1989). The data in use do not follow Gaussian distribution, then canonical correlation analysis after PCA may not be able to reproduce the true correlated components. Due to this constraint, the data series are transformed into the independent components by independent component analysis in order to satisfy probability assumptions.

In the second stage, the naive and causal skeletons of the climate features of the local scale precipitation series can be reproduced by the proposed RNN-based downscaling model. The final stage is a simple regression model between the

independent components and precipitation series. The model training process is the training trough-time algorithm of which leads to capture the dynamical effects during downscaling procedure. This point has big advantages while comparing to classical static-relations methods; in the model, if the predictors are changed then the model can able to modify its structure without regenerating the statistical model. This is the main reason of why the model is named as a grey prediction scheme in this study.

In the second model, the problem of statistical linkages between the large scale upper-air circulations (namely 500-1000 hPa thicknesses and 500 hPa geopotential heights from NCEP-NCAR reanalysis data sets) and near-surface air temperature series (monthly mean, maximum and minimum) during the period of 1951-1998 are investigated by using the noise-reduction techniques of SSA and PCA. The noise filtered data consist of the deterministic (trend and/or seasonal cycles) and statistical components.

Before model building, Sampson correlation patterns are reproduced; similar to the first model, between the upper-air circulations and near-surface air temperatures. Both the NCEP-NCAR reanalysis data sets and observations are divided into four seasons in order to extract the seasonal climate feature statistical linkages.

The effects of the large scale upper air circulation on monthly maximum, minimum and mean temperature series show similar patterns in winter months. On the other hand, those effects resemble in the Sampson correlation patterns for the monthly maximum and mean temperatures whereas the minimum temperature series show very different patterns in the other seasons (summer, autumn and spring). In the autumn months, the influence areas of the tropical air systems which are effective in summer months form a belt towards southern and south-eastern parts of Turkey.

The Sampson correlation patterns show that the correlations between maximum temperatures and 500 hPa geopotential heights in the area covering the southern and western part of Turkey are at significant levels in summer and spring. Since Turkey is under influence of warm tropical air systems in early spring, the geopotential heights having ridge characteristics lead to an increase in the temperatures. On the other hand, 500 hPa geopotential pattern high centers extend towards Europe and its signature can be seen in the variability of the temperature fields.



In summer months, over this area, there is a ridge effect on 500 hPa level due to Azores high which shows a warm-high geopotential height center extending its influence to the Europe. Displacement of Azores High center over Europe generates northerly current flows. As a result of this mechanism, the temperature decreases especially in the regions of western and northern parts of Turkey.

The Sampson correlation ratio patterns show that there are strong positive correlations between 500-1000 hPa thicknesses and mean and maximum temperatures. The thicknesses increase Turkish mean and maximum temperatures starting in the spring and continuing in summer. In this season, the tropical air system (Monsoon low) moving to high latitudes over the eastern and south-eastern part of Turkey (includes the eastern and central Black Sea Region of Turkey) has important effects over of these regions. The warm features of this air system lead to increase the thicknesses and create ridges on 500 hPa levels.

The second model design is not more complex than the first model. Because the temperature series over Turkey show similar patterns of general circulation characteristics, evidently, the downscaling strategy for the temperature series is constructed by well-known technique of redundancy analysis. The CCA may be a way to extract the symmetrical relationships; on the other hand, RA is able to extract the asymmetric relationships. This is the major reason why RA is preferred to CCA in this study.

The downscaling model results reveal that without noise-reduction procedure, the number of the significant principal components is increased, particularly, for maximum temperature series. However, the third PC of maximum temperature series can not be connected with the large scale processes therefore noise-free data sets are more appropriate for downscaling purposes.

For further studies, the designs of proposed model schemes may be bridged by nested models in order to extract climate impacts based on physical model approaches. Additionally, during the parameterization of nested models, the results of the proposed model outputs may be very supportive to explain the large scale climate variability and atmospheric disturbances. Its ability to downscale local scale climate features from a nested model is still one of the open questions in this work.

The applications are restricted with the region over Turkey; however the proposed model designs may also be modified to apply to any region over the world. Finally, in most cases, a statistical downscaling model should be developed and tested by reproducing the observed low-frequency variability of the regional climate when GCMs simulate correctly the large scale climate variables.



## REFERENCES

- Allen, M.R. and Smith, L.A., 1996. Monte Carlo SSA: Detecting irregular oscillations in the presence of coloured noise, *Journal of Climate*, **9**, 3373-3404.
- Alpert, P., Neeman, B.U. and Shah-El, Y., 1990. Intermonthly variability of cyclone tracks in the Mediterranean, *Journal of Climate*, **3**, 1474-1478.
- Bardossy, A., 1994. Downscaling from GCMs to local climate through stochastic linkages, pp. 33-46, in *Climate Change, Uncertainty and Decision Making*, Ed. Paoli, G., Institute for Risk Research, Waterloo.
- Becker, S. and Hinton, G.E., 1992. A self-organization neural network that discovers surfaces in random-dot stereogram, *Nature (London)*, **355**, 161-163.
- Berenson, M.L., Levine, D.M. and Goldstein, M., 1993. Intermediate Statistical Methods and Applications, Prentice-Hall, New Jersey.
- Bluestein, H.B., 1993. Synoptic-Dynamic Meteorology in Midlatitudes, Vol. II, Oxford University Press, Oxford.
- Box, G.E.P., Jenkins, G.M. and Reinsel, G.C., 1994. Time Series Analysis: Forecasting and Control, 3rd edn, Prentice Hall, New Jersey.
- Bras, R.L. and Rodriguez-Iturbe, I., 1993. Random Functions and Hydrology, Dover Publications, New York.
- Bryant, E., 1997. Climate Process and Change, Cambridge University Press, New York.
- Chen, J.M., Chang, C.P. and Harr, P., 1994. A technique for analyzing optimal relationships among multiple sets of data fields. Part I: The method, *Monthly Weather Review*, **122**, 2482-2493.
- Chen, J.M. and Chang, C.P., 1994. A technique for analyzing optimal relationships among multiple sets of data fields. Part II: Reliability case study, *Monthly Weather Review*, **122**, 2494-2505.
- Comon, P., 1994. Independent component analysis-a new concept? *Signal Processing*, **36**, 287-314.

- Connor, J., Martin, R. and Atlas, L., 1994. Recurrent neural networks and robust time series prediction, *IEEE Transactions on Neural Networks*, **5**, 240–253.
- Conway, D. and Jones, P.D., 1998. The use of weather types and air flow indices for GCM downscaling, *Journal of Hydrology*, **212–213**, 348–361.
- Cubasch, U., Hasselmann, K., Höck, H., Maier-Reimer, E., Mikolajewicz, U., Santer, B.D. and Sausen, R., 1992. Time-dependent greenhouse warming computations with coupled ocean-atmosphere model, *Climate Dynamics*, **8**, 55–69.
- Cubasch, U., von Storch, H., Waszkewitz, J. and Zorita, E., 1996. Estimates of climate change in southern Europe using different downscaling techniques, **report no. 183**, Max Planck Institute für Meteorologie, Hamburg, Germany.
- Diaz, H.F. and Fulbright D.C., 1981. Eigenvector analysis of seasonal temperature, precipitation and synoptic-scale system frequency over the Contiguous United States. Part I: Winter, *Monthly Weather Review*, **109**, 1267–1284.
- Dobrovolski, S.G., 2000. Stochastic Climate Theory: Models and Applications, Springer, Berlin.
- Elman, J., 1990. Finding structure in time, *Cognitive Science*, **14**, 179–221.
- Elsner, J.B. and Tsonis, A.A., 1994. Low-frequency oscillation, *Nature*, **372**, 507–508.
- Elsner, J.B. and Tsonis, A.A., 1996. Singular Spectrum Analysis: A New Tool in Time Series Analysis, Plenum Press, New York, London.
- Erinç, S., 1984. Climatology and Its Methods, Istanbul University Press, Marine Science, Institute of Geography, Istanbul (in Turkish).
- Everson, R. and Roberts, S., 1999. Independent component analysis: a flexible non-linearity and decorrelating manifold approach, *Neural Computing*, **11**, 1957–1984.
- Fuentes, U. and Heimann, D., 2000. An improved statistical-dynamical downscaling scheme and its application to the Alpine precipitation climatology, *Theoretical and Applied Climatology*, **65**, 119–135.
- Geerts, B., 2003. Empirical estimation of the monthly-mean daily temperature range, *Theoretical and Applied Climatology*, **74**, 145–165.
- Ghil, M. and Vautard, R., 1991. Interdecadal oscillations and the warming trend in global temperature time series, *Nature*, **350**, 324–327.

- Giorgi, F. and Mearns, L.O.**, 1991. Approaches to the simulation of regional climate change: a review, *Review of Geophysics*, **29**, 191-216.
- Glahn, H.R.**, 1968. Canonical correlation and its relationship to discriminant analysis and multiple regression, *Journal of Atmospheric Science*, **25**, 23-31.
- Gower, J.C.**, 1975. Generalized procrustes analysis, *Psychometrika*, **40**, 33-51.
- Green M.C., Floccini, R.G. and Myrup, L.O.**, 1993. Use of temporal principal components analysis to determine seasonal periods, *Journal of Applied Meteorology*, **32**, 986-995.
- Gyalistras, D., von Storch, H., Fischlin, A. and Beniston, M.**, 1994. Linking GCM simulated climatic changes to ecosystem models: Case studies of statistical downscaling in the Alps, *Climate Research*, **4**, 67-189.
- Haykin, S.**, 1999. Neural Networks: A Comprehensive Foundation, 2<sup>nd</sup> edn, Prentice-Hall, New Jersey.
- Haykin, S.**, 2001. Kalman Filtering and Neural Networks, Wiley, New York.
- Heimann, D. and Sept, V.**, 2000. Climatic change of summer temperature and precipitation in the Alpine region, *Theoretical and Applied Climatology*, **66**, 1-12.
- Hewitson, B.C. and Crane, R.G.**, 1996. Climate downscaling: Techniques and application, *Climate Research*, **7**, 85-95.
- Hipel, K.W. and Mcleod, A.I.**, 1994. Developments in Water Sciences: Time Series Modeling of Water Resources and Environmental Systems, Elsevier, Amsterdam, New York.
- Hochreiter, S., Younger, A.S. and Conwell, P.R.**, 2001. Learning to learn using gradient descent, in *Proceeding of the International Conference on Artificial Neural Networks*, Vienna, Austria, 21-25 August, p. 87-94.
- Holland, W.R., Joussaume, S. and David, F.**, 1999. Modeling The Earth's Climate And Its Variability, Elsevier, Amsterdam.
- Holton, J.R.**, 1992. An Introduction to Dynamic Meteorology, 3<sup>rd</sup> edn, Academic Press, San Diego.
- Hornik, K.**, 1991. Approximation capabilities of multilayer feed forward network, *Neural Networks*, **4**, 251-257.
- Hotelling, H.**, 1936. Relations between two sets of variates, *Biometrika*, **28**, 321-377.

- Houghton, J.T., Meira Filho, L.G., Callender, B.A., Harris, N., Kattenbarg, A. and Maskell, K. (Eds.), 1996. *Climate Change 1995: The Science of Climate Change*, Cambridge University Press, Cambridge.
- Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Dai, X., Maskell, K. and Johnston, C.A. (Eds.), 2001. *Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge.
- Hyvarinen, A. and Oja, E., 1997. A fast fixed-point algorithm for independent component analysis, *Neural Computing*, **9**, 1483-1492.
- Hyvarinen, A., Karhunen, J. and Oja, E., 2001. *Independent Component Analysis*, John Wiley, New York.
- Ito, Y., 1991. Representation of functions by superposition of a step or sigmoid functions and their applications to neural network theory, *Neural Networks*, **4**, 385-394.
- Jackson, J.E., 1991. *A Users Guide to Principal Components*, John Wiley and Sons, New York.
- Jackson, D., 1993. Stopping rules in principal components analysis: A comparison of heuristic and statistical approaches, *Ecology*, **46**, 2204—2214.
- Johannesson, T., Jonsson, T., Kallen, E. and Kaas E., 1995. Climate change scenarios for the Nordic countries, *Climate Research*, **5**, 181-195.
- Johnson, R.A. and Wichern, D.W., 1998. *Applied Multivariate Statistical Analysis*, 4<sup>th</sup> edn, Prentice Hall, New York.
- Jones, P.D., Raper, S.C.B., Bradley, R.S., Diaz, H.F., Kelly, P.M. and Wigley, T.M.L., 1986. Northern Hemisphere surface air temperature variations, 1851-1984, *Journal of Climate and Applied Meteorology*, **25**, 161-179.
- Jones, P.D., 1995. Maximum and minimum temperature trends in Ireland, Italy, Thailand, and Turkey, *Atmospheric Research*, **37**, 67-78.
- Jong, J. and Kotz, S., 1999. On a relation between principal components and regression analysis, *The American Statistician*, **53**, 349-351.
- Jonsson, T., Johannesson, T. and Kallen, E., 1994. Climate change scenarios for the Nordic countries, *Icelandic Meteorological Office and National Energy Authority Report*, OS-94030/VOD-04b, 23 pp.
- Jordan, M.I., 1986. Attractor dynamics and parallelism in a connectionist sequential machine, in *Proceedings of 1986 Cognitive Science Conference*, Amherst, MA, August, p.531-546.



- Jöreskog, K.G. and Sörbom, D., 1989. LISREL-7: A Guide to the Program and Applications, 2<sup>nd</sup> edn, SPSS Publications, Chicago.
- Jutten, C. and Herault, J., 1991. Blind separation of sources. Part I: An adaptive algorithm based on neuromimetic architecture, *Signal Processing*, **24**, 1-10.
- Kadioğlu, M., 1997. Trends in surface air temperature data over Turkey, *International Journal of Climatology*, **17**, 511-520.
- Kadioğlu, M., 2000. Regional variability of seasonal precipitation over Turkey, *International Journal of Climatology*, **20**, 1743-1760.
- Kaiser, H.F., 1959. Computer program for varimax rotation in factor analysis, *Educational and Psychological Measurement*, **19**, 413-420.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woolen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D., 1996. The NCEP/NCAR Reanalysis 40-year Project, *Bulletin of American Meteorology Society*, **77**, 437-471.
- Karl, T.R., Wang, W.C., Schlesinger, M.E., Knight, R.W. and Portman, D., 1990. A method of relating general circulation model simulated climate to the observed local climate. Part I: Seasonal statistics, *Journal of Climate*, **3**, 1053-1079.
- Kidson, J.W. and Thompson, C.S., 1998. Comparison of statistical and model-based downscaling techniques for estimating local climate variations, *Journal of Climate*, **11**, 735-753.
- Kidson, J.W., 1975. Eigenvector analysis of monthly mean surface data, *Monthly Weather Review*, **103**, 177-186.
- Kim, J.W., Chang, J.T., Baker, N.L., Wilks, D.S. and Gates, W.L., 1984. The statistical problem of climate inversion. Determination of the relationship between local and large scale climate, *Monthly Weather Review*, **112**, 2069-2077.
- Klein, W.H., 1982. Statistical weather forecasting on different time scales, *Bulletin of American Meteorology Society*, **63**, 170-177.
- Kohonen, T., 1995. Self-Organizing Maps, Springer-Verlag, Berlin.
- Kömüşçü, A.U., 1998. An analysis of the fluctuations in the long-term annual mean air temperature data of Turkey, *International Journal of Climatology*, **18**, 99-213.

- Kutiel, H., Hirsch-Eshkol, T.R. and Türkeş, M., 2001.** Sea level pressure patterns associated with dry or wet monthly rainfall conditions in Turkey, *Theoretical and Applied Climatology*, **69**, 39-67.
- Kutiel, H., Maheras, P. and Guika, P., 1996.** Circulation and extreme rainfall conditions in the eastern Mediterranean during the last century, *International Journal of Climatology*, **16**, 73-92.
- Lamb, H.H., 1972.** British Isles Weather Types And A Register Of Daily Sequence Of Circulation Patterns, 1861-1971, **Geophysical Memoir 116**, HMSO, London.
- Landman, W.A. and Tennant, W.J., 2000.** Statistical downscaling of monthly forecast, *International Journal of Climatology*, **20**, 1521-1532.
- Malone, T.F., (ed.), 1951.** Compendium of Meteorology, American Meteorological Society, Boston.
- Mardia, K.V., Kent, J.T. and Bibby, J.M., 1979.** Multivariate Analysis, Academic Press, New York.
- Matyasovsky, I., Bogardi, A. and Duckstein, L., 1994.** Local temperature estimation under climate change, *Theoretical and Applied Climatology*, **50**, 1-13.
- McGregor, J.L., 1997.** Regional climate modeling, *Meteorology and Atmospheric Physics*, **63**, 105-117.
- McGregor, J.L., Walsh, K.J. and Katzfey, J.J., 1993.** Nested modelling for regional climate studies, in *Modelling Change in Environmental Systems*, pp.367-386, Eds. Jakeman, A.J., Beck, M.B. and McAleer, M.J., John Wiley and Sons, New York.
- Murphy, J.M., 1999.** An evaluation of statistical and dynamical techniques for downscaling local climate, *Journal of Climate*, **12**, 2256-2284.
- Murphy, J.M., 2000.** Predictions of climate change over Europe using statistical and dynamical downscaling techniques, *International Journal of Climatology*, **20**, 489-501.
- Nicholls, N., 1987.** The use of canonical correlations to study teleconnections, *Monthly Weather Review*, **115**, 393-399.
- Noguer, M., 1994.** Using statistical techniques to deduce local climate distributions: An application for model validation, *Meteorology Applications*, **1**, 277-287.
- North, G., 1984.** Empirical orthogonal functions and normal modes, *Journal of Atmospheric Science*, **41**, 879-887.

- Özeren, M.S., Tath, H. and Dalfes, H.N., 2003. Independent component analysis (ICA) of the influences of North Atlantic Oscillation (NAO) on local climate. *11th Congress of Signal Processing and Communication Applications (SIU 2003)*, Koç University, Istanbul, Turkey, 18-20 June, pp.19-22 (in Turkish).
- Park, J. and Sandberg, I.W., 1991. Universal approximation using radial-basis function networks, *Neural Computation*, **3**, 246-257.
- Pati, Y.C. and Krishnaprasad, P.S., 1993. Analysis of and synthesis of feed forward neural networks using discrete affine wavelet transformations, *IEEE Transactions on Neural Networks*, **4**, 96-108.
- Pearlmutter, B.A., 1995. Gradient calculations for dynamic recurrent neural networks: a survey, *IEEE Transactions on Neural Networks*, **6**, 1212-1228.
- Peixoto, J.P. and Oort, A.H., 1992. *Physics of Climate*, American Institute of Physics, New York.
- Preisendorfer, R.W., 1988. *Principal Component Analysis in Meteorology and Oceanography*, Elsevier, New York.
- Puskorius, G.V. and Feldkamp, L.A., 1994. Neurocontrol of nonlinear dynamical systems with Kalman filter trained recurrent networks, *IEEE Transactions on Neural Networks*, **5**, 279-297.
- Rencher, A.C., 1992. Interpretation of canonical discriminant functions, canonical variates, and principal components, *American Statistician*, **46**, 217-25.
- Rencher, A.C., 1995. *Methods of Multivariate Analysis*, John Wiley and Sons, New York.
- Rencher, A.C., 1998. *Multivariate Statistical Inference and Applications*, John Wiley and Sons, New York.
- Reyment, R. and Jöreskog, K.G., 1993. *Applied Factor Analysis in the Natural Sciences*, 2<sup>nd</sup> edn, Cambridge University Press, Cambridge.
- Richardson, C.W., 1981. Stochastic simulation of daily precipitation, temperature and solar radiation, *Water Resources Research*, **17**, 182-190.
- Richman, M.B., 1985. Rotation of principal components, *Journal of Climatology*, **6**, 293-335.
- Robinson, A.J. and Fallside, F., 1991. A recurrent error propagation speech recognition system, *Computer Speech and Language*, **5**, 259-274.

- Rumelhart, D.E. and McClelland, J.L. (eds.), 1986.** Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1. MIT Press, Cambridge.
- Rummukainen, M., 1997.** Methods for Statistical Downscaling of GCM Simulation, *Reports Meteorology and Climatology, report no. 80*, SWECLIM (Swedish Regional Climate Modelling Programme),
- Sailor, D.J. and Li, X.A., 1999.** A semi-empirical downscaling approach for predicting regional temperature impacts associated with climate change, *Journal of Climate*, **12**, 103-114.
- Sampson, A.R., 1984.** A multivariate correlation ratio, *Statistics and Probability Letters*, **2**, 77-81.
- Schlesinger, M.E. and Ramankutty, N., 1994.** An oscillation in the global climate systems of period 65-70 years, *Nature*, **367**, 723-726.
- Schubert, A. and Henderson-Sellers, A., 1997.** A statistical model to downscale local daily temperature extremes from synoptic scale atmospheric circulation patterns in the Australian region, *Climate Dynamics*, **13**, 223-234.
- Smith, J., 1999.** Models and scale: up and downscaling, p. 81-98, in *Data and Models in Action*, Eds. Stein, A. and Penning de Vries, F.W.T., Kluwer Academic, Dordrecht.
- Solman, S.A. and Nunez, M.N., 1999.** Local estimates of global change: A statistical downscaling approach, *International Journal of Climatology*, **19**, 835-861.
- Stein, A., Riley, J. and Halberg, N., 2001.** Issues of scale for environmental indicators, *Agriculture, Ecosystem and Environment*, **87**, 215-232.
- Taha, M.F., Harb, S.A., Nagib, M.K. and Tantawy, A.H., 1981.** Climate of Southern and Western Asia. *The Climate of the Near East*, Eds. Takahashi, K. and Arakawa, H., World Survey of Climatology Series, Vol. 9. Elsevier. Amsterdam.
- Tatlı, H., Dalfes, H.N. and Menteş, Ş.S., 2004.** A statistical downscaling method for monthly total precipitation over Turkey, *International Journal of Climatology*, **24**, 161-180.
- Tatlı, H., Dalfes, H.N. and Menteş, Ş.S., 2005.** Surface air temperature variability over Turkey and its connection to the large scale upper air circulation via multivariate techniques, *International Journal of Climatology* (in press).
- Tatlı, H., Dalfes, H.N. and Menteş, Ş.S., 2003.** Principal factor and canonical correlation analysis of sea level pressures associated with precipitation

series in Turkey. *III. Atmospheric Science Symposium*, Istanbul Technical University Istanbul, Turkey, 19-21 March, pp. 150-160, (in Turkish).

- Tayanç, M., Karaca, M. and Yenigun, O., 1997.** Annual and seasonal air temperature trend patterns of climate change and urbanization effects in relation to air pollutants in Turkey, *Journal of Geophysical Research-Atmosphere*, **102**, 1909-1919.
- TenBerge, J.M.F., 1988.** Generalized approaches to the Maxbet problem and Maxdiff problem, with applications to canonical correlations, *Psychometrika*, **53**, 487-494.
- Touchan, R., Garfin, G.M., Meko, D.M., Funkhouser, G., Erkan, N., Hughes, M.K. and Wallin, B., 2003.** Preliminary reconstructions of spring precipitation in southwestern Turkey from tree-ring width, *International Journal of Climatology*, **23**, 157-171.
- Tourre, Y.M., 2000.** *Climlab2000: A Statistical Software Package for Climate Applications*, IRI: International Research Institute for Climate Prediction, <http://iri.columbia.edu/outreach/training/climlab2000>, 2003, version 1.1.0.
- Türkeş, M. and Erlat, E., 2003.** Precipitation changes and variability in Turkey linked to the North Atlantic Oscillation during the period 1930-2000, *International Journal of Climatology*, **23**, 1771-1796.
- Türkeş, M., Sümer, U.M. and Kılıç, G., 1995.** Variations and trends in annual mean air temperatures in Turkey with respect to climatic variability, *International Journal of Climatology*, **15**, 557-569.
- Türkeş, M., 1996.** Spatial and temporal analysis of annual rainfall variations in Turkey, *International Journal of Climatology*, **16**, 1057-1076.
- Türkeş, M., 1998.** Influence of geopotential heights, cyclone frequency and southern oscillation on rainfall variations in Turkey, *International Journal of Climatology*, **18**, 649-680.
- Türkeş, M., Sümer, U.M. and Kılıç, G., 2002a.** Persistence and periodicity in the precipitation series of Turkey and associations with 500 hPa geopotential heights, *Climate Research*, **21**, 59-81.
- Türkeş, M., Sümer U.M. and Demir, I., 2002b.** Re-evaluation of trends and changes in mean, maximum and minimum temperatures of Turkey for the period 1929-1999, *International Journal of Climatology*, **22**, 947-977.
- Tyler, D.E., 1982.** On the optimality of the simultaneous redundancy transformations, *Psychometrika*, **47**, 77-86.



- van de Geer, J.P., 1984. Linear relations among k sets of variables, *Psychometrika*, **49**, 79-94.
- van den Wollenberg, A.L., 1977. Redundancy analysis, *Psychometrika*, **42**, 207-219.
- Vautard, R., Yiou, P. and Ghil, M., 1992. Singular spectrum analysis: a toolkit for short, noisy and chaotic series, *Physica (D)*, **58**, 95-126.
- von Storch, H., 1995. Inconsistencies at the interface of climate impact studies and global climate research, *Meteorologie Zeitschrift*, **4**, 72-80.
- von Storch, H. and Zwiers, F.W., 1999. Statistical Analysis in Climate Research. Springer/Cambridge University Press. Cambridge.
- von Storch, H., Zorita, E. and Cubash, U., 1993. Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime, *Journal of Climate*, **6**, 1161-1171.
- Watson, R.T. (ed.), 2002. Climate Change 2001: Synthesis Report, *Third Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge.
- Wigley, T.M.L., Jones, P.D., Briffa, K.R. and Smith G., 1990. Obtaining sub-grid-scale information from coarse-resolution general circulation model output, *Journal of Geophysical Research*, **95**, 1943-1953.
- Wilby, R.L. and Wigley, T.M.L., 2000. Precipitation predictors for downscaling: observed and General Circulation Model relationships, *International Journal of Climatology*, **20**, 641-661.
- Wilkinson, J.H. and Reinsch, C., 1971. Handbook for Automatic Computation, vol. 2, Linear Algebra. Springer-Verlag, Berlin.
- Wilks, D.S., 1989. Statistical specification of local surface weather elements from large scale information, *Theoretical and Applied Climatology*, **40**, 119-134.
- Wilks, D.S., 1992. Adapting stochastic weather generation algorithms for climate change studies, *Climate Change*, **22**, 67-84.
- Wilks, D.S., 1995. Statistical Methods in the Atmospheric Sciences: An introduction, Academic Press, London.
- Xoplaki, E., Gonzales-Rouco, J.F., Luterbacher, J. and Wanner, H., 2003. Mediterranean summer air temperature variability and its connection to the large scale atmospheric circulation and SSTs, *Climate Dynamics*, **20**, 723-739.

**Zorita, E., Kharin, V. and von Storch, H., 1992.** The atmospheric circulation and sea surface temperature in the North Atlantic area in winter: Their interaction and relevance for Iberian precipitation, *Journal of Climate*, **5**, 1097-1108.



## APPENDIX

### F-Language (FORTRAN) Source Code of SVD

The following subroutines compute the Singular Value Decomposition (SVD) of a general rectangular matrix **X**; but the declaration of the integer variable should be changed as "dr = 8" on a unix machine. If the left and right eigenvectors are represented by **u** and **v**, and the eigenvalues by **d** respectively, then the calling subroutine of 'svd' requires only the following two statements in the main program.

```
use svd_computation
call svd (x, u, d, v)
```

```
module svd_computation
implicit none
public::drotg,dswap1,drot1,svd
integer,parameter,public:: dr=2
```

```
!Based upon routines from the
!nswc (naval surface warfare
!center) which were based
!upon lapack routines
!Code converted from mixed
!F77 and F_90 to F_language by
!Hasan TATLI
!Date: 09/ 02/2002
```

contains

```
subroutine drotg (da,db,dc, ds)
```

```
!construct Givens transformation
!normally, the subprogram drot
!(n,dx,incx,dy,incy,dc,ds) will
!next be called to apply the
!transformation to a 2 by n matrix.
```

```
real(kind=dr),intent(inout) :: da
real (kind=dr), intent(in out) :: db
real (kind = dr), intent(out) :: dc
real (kind = dr), intent(out) :: ds
real (kind = dr) :: u, v, r
```

```
if (abs(da) <= abs(db)) then
  go to 10
endif
u = da + da
v = db / u
r = sqrt(0.25_dr + v**2) * u
dc = da / r
ds = v * (dc + dc)
db = ds
```

```
da = r
return
10 continue
if (db == 0.0_dr) then
  go to 20
endif
u = db + db
v = da / u
da = sqrt(0.25_dr + v**2) * u
ds = db / da
dc = v * (ds + ds)
if (dc == 0.0_dr) then
  go to 15
endif
db = 1.0_dr / dc
return
15 continue
db = 1.0_dr
return
20 continue
dc = 1.0_dr
ds = 0.0_dr
return
end subroutine drotg
!-----
subroutine dswap1 (n, dx, dy)
!Jack Dongarra, linpack,
!3/11/78.
integer, intent (in) :: n
real(kind=dr),dimension(:),
intent(inout):: dx
real(kind=dr),dimension(:),
intent(inout):: dy
real (kind = dr) :: dtemp
integer :: i, m, mp1

if (n <= 0) then
  return
endif
```

```
m = modulo (n, 3)
m=n-m
if ( m == 0 ) then
  go to 40
endif
do i = 1,m
  dtemp = dx(i)
  dx(i) = dy(i)
  dy(i) = dtemp
end do
if(n < 3 ) then
  return
endif
40 continue
mp1 = m + 1
do i = mp1,n,3
  dtemp = dx(i)
  dx(i) = dy(i)
  dy(i) = dtemp
  dtemp = dx(i + 1)
  dx(i + 1) = dy(i + 1)
  dy(i + 1) = dtemp
  dtemp = dx(i + 2)
  dx(i + 2) = dy(i + 2)
  dy(i + 2) = dtemp
end do
return
end subroutine dswap1
!-----
subroutine
drot1 (n,dx,dy,c,s)

!Jack Dongarra, Linpack,
!13/11/78.

integer, intent(in):: n
```

```

real(kind=dr),dimension(:), &
intent(inout):: dx
real(kind=dr),dimension(:), &
intent(inout):: dy
real(kind=dr), intent(in) :: c
real(kind=dr), intent(in) :: s
real(kind=dr) :: dtemp
integer :: i

```

```

if(n <= 0) then
  return
endif

```

```

do i = 1,n
  dtemp = c*dx(i) + s*dy(i)
  dy(i) = c*dy(i) - s*dx(i)
  dx(i) = dtemp
end do

```

```

return
end subroutine drot1

```

```

!-----
subroutine svd (x,u,diag,v)
integer,parameter:: job=1
real(kind=dr),dimension(:,,:), &
intent(inout):: x
real(kind=dr),dimension(:,,:), &
intent(out):: diag
real(kind=dr), &
dimension(:,,:), intent(out) :: u
real(kind=dr), &
dimension(:,,:), intent(out) :: v
real(kind=dr), &
dimension(size(x,1)):: e
real(kind=dr), &
dimension(size(x,2)):: s

```

**! Linpack, this version dated**  
**! 03/19/79.**  
**! G.W. Stewart, University of**  
**! Maryland**

```

real(kind=dr), &
dimension (size(x,1)) :: work
integer :: iter, j,jobu, k, kase, &
kk, l, ll, lls, lm1, lp1, ls, lu, &
m, maxit, mm, mm1, mp1, &
nct, nctp1, ncui, nrt, nrtp1, &
info, n, p
real (kind = dr) :: t,b,c, cs, &
el, emm1, f, g, scale, shift, &
sl, sm, sn, smm1, t1, test, &
ztest
logical :: wantu, wantv

```

```

  n=size(x,1)
  p=size(x,2)
! Maximum number of iterations
  maxit = 30
  wantu = .false.
  wantv = .false.
  jobu=2

```

```

  ncui = n
  if (jobu > 1) then
    ncui = min(n,p)
  endif
  if (jobu /= 0) then
    wantu = .true.
  endif
  if (job /= 0) then
    wantv = .true.
  endif
  info = 0
  nct = min(n-1, p)
  s(1:nct+1) = 0.0_dr
  nrt = max(0, min(p-2,n)) &
  lu = max(nct,nrt)
  if (lu < 1) then
    go to 170
  endif
  do l = 1, lu
    lp1 = l + 1
    if (l > nct) then
      go to 20
    endif
    s(l)=sqrt(sum(x(l:n,l)**2 )) &
    if (s(l) == 0.0_dr) then
      go to 10
    endif
    if (x(l,l) /= 0.0_dr) then
      s(l) = sign(s(l), x(l,l))
    endif
    x(l:n,l) = x(l:n,l) / s(l)
    x(l,l) = 1.0_dr + x(l,l)
    10 continue
    s(l) = -s(l)
    20 continue
    if (p < lp1) then
      go to 50
    endif
    do j = lp1, p
      if (l > nct) then
        go to 30
      endif
      if (s(l) == 0.0_dr) then
        go to 30
      endif
      t = -dot_product(x(l:n,l), &
        x(l:n,j)) / x(l,l)
      x(l:n,j)=x(l:n,j)+t*x(l:n,l) &
      30 continue
      e(j) = x(l,j)
    end do
    50 continue
    if (.not. wantu .or. l > nct) then
      go to 70
    endif
    u(l:n,l) = x(l:n,l)
    70 continue
    if (l > nrt) then
      cycle
    endif

```

```

  do l = 1, lu
    lp1 = l + 1
    if (l > nct) then
      go to 20
    endif
    s(l)=sqrt(sum(x(l:n,l)**2 )) &
    if (s(l) == 0.0_dr) then
      go to 10
    endif
    if (x(l,l) /= 0.0_dr) then
      s(l) = sign(s(l), x(l,l))
    endif
    x(l:n,l) = x(l:n,l) / s(l)
    x(l,l) = 1.0_dr + x(l,l)
    10 continue
    s(l) = -s(l)
    20 continue
    if (p < lp1) then
      go to 50
    endif
    do j = lp1, p
      if (l > nct) then
        go to 30
      endif
      if (s(l) == 0.0_dr) then
        go to 30
      endif
      t = -dot_product(x(l:n,l), &
        x(l:n,j)) / x(l,l)
      x(l:n,j)=x(l:n,j)+t*x(l:n,l) &
      30 continue
      e(j) = x(l,j)
    end do
    50 continue
    if (.not. wantu .or. l > nct) then
      go to 70
    endif
    u(l:n,l) = x(l:n,l)
    70 continue
    if (l > nrt) then
      cycle
    endif

```

```

    e(l)=sqrt(sum(e(lp1:p)**2 )
  ) &
  if (e(l) == 0.0_dr) then
    go to 80
  endif
  if (e(lp1) /= 0.0_dr) then
    e(l) = sign(e(l), e(lp1))
  endif
  e(lp1:lp1+p-l-1)=&
  e(lp1:p)/ e(l)
  e(lp1) = 1.0_dr + e(lp1)
  80 continue
  e(l) = -e(l)
  if (lp1 > n .or. e(l) == &
    0.0_dr) then
    go to 120
  endif
  work(lp1:n) = 0.0_dr
  do j = lp1, p
    work(lp1:lp1+n-l-1) = &
    work(lp1:lp1+n-l-1)+e(j)* &
    x(lp1:lp1+n-l-1,j)
  end do
  do j = lp1, p
    x(lp1:lp1+n-l-1,j) = &
    x(lp1:lp1+ &
    n-l-1,j)-e(j)/e(lp1)) &
    *work(lp1:lp1+ n-l-1)
  end do
  120 continue
  if (.not. wantv) then
    cycle
  endif
  v(lp1:p,l) = e(lp1:p)
end do
170 continue
  m = min(p,n+1)
  nctp1 = nct + 1
  nrtp1 = nrt + 1
  if (nct < p) then
    s(nctp1) = x(nctp1,nctp1)
  endif
  if (n < m) then
    s(m) = 0.0_dr
  endif
  if (nrtp1 < m) then
    e(nrtp1) = x(nrtp1,m)
  endif
  e(m) = 0.0_dr
  if (.not. wantu) then
    go to 300
  endif
  if (ncui < nctp1) then
    go to 200
  endif
  do j = nctp1, ncui
    u(1:n,j) = 0.0_dr
    u(j,j) = 1.0_dr
  end do
  200 continue

```

```

do ll = 1, nct
l = nct - ll + 1
if (s(l) == 0.0_dr) then
go to 250
end if
lp1 = l + 1
if (ncu < lp1) then
go to 220
end if
do j = lp1, ncu
t = -dot_product(u(l:n,l), &
u(l:n,j)) / u(l,l)
u(l:n,j) = u(l:n,j) + t * u(l:n,l)
end do
220 continue
u(l:n,l) = -u(l:n,l)
u(l,l) = 1.0_dr + u(l,l)
lm1 = l - 1
if (lm1 < 1) then
cycle
end if
u(1:lm1,l) = 0.0_dr
cycle
250 continue
u(1:n,l) = 0.0_dr
u(l,l) = 1.0_dr
end do
300 continue
if (.not. wantv) then
go to 350
end if
do ll = 1, p
l = p - ll + 1
lp1 = l + 1
if (l > nrt) then
go to 320
endif
if (e(l) == 0.0_dr) then
go to 320
endif
do j = lp1, p
t = -dot_product(v(lp1:lp1+ &
p-l-1,l), v(lp1:lp1+p-l-1,j)) &
/ v(lp1,l)
v(lp1:lp1+p-l-1,j) = v(lp1:lp1+ &
p-l-1,j) + t * v(lp1:lp1+p-l-1,l)
end do
320 continue
v(1:p,l) = 0.0_dr
v(l,l) = 1.0_dr
end do
350 continue
mm = m
iter = 0
do
if (m == 0) then
go to 620
endif
if (iter < maxit) then
go to 370

```

```

endif
info = m
go to 620
370 continue
do ll = 1, m
l = m - ll
if (l == 0) then
exit
end if
test = abs(s(l)) + abs(s(l+1))
ztest = test + abs(e(l))
if (ztest /= test) then
cycle
end if
e(l) = 0.0_dr
exit
end do
if (l /= m - 1) then
go to 410
end if
kase = 4
go to 480
410 continue
lp1 = l + 1
mp1 = m + 1
do lls = lp1, mp1
ls = m - lls + lp1
if (ls == l) then
exit
endif
test = 0.0_dr
if (ls /= m) then
test = test + abs(e(ls))
endif
if (ls /= l + 1) then
test = test + abs(e(ls-1))
endif
ztest = test + abs(s(ls))
if (ztest /= test) then
cycle
endif
s(ls) = 0.0_dr
exit
end do
if (ls /= l) then
go to 450
endif
kase = 3
go to 480
450 continue
if (ls /= m) then
go to 460
endif
kase = 1
go to 480
460 continue
kase = 2
l = ls
480 continue
l = l + 1

```

```

select case (kase)
case (1)
go to 490
case (2)
go to 520
case (3)
go to 540
case (4)
go to 570
end select
490 continue
mm1 = m - 1
f = e(m-1)
e(m-1) = 0.0_dr
do kk = 1, mm1
k = mm1 - kk + 1
t1 = s(k)
call drotg(t1, f, cs, sn)
s(k) = t1
if (k == l) then
go to 500
end if
f = -sn*e(k-1)
e(k-1) = cs*e(k-1)
500 continue
if (wantv) then
call drot1(p,v(1:,k), &
v(1:,m), cs, sn)
end if
end do
go to 610
520 continue
f = e(l-1)
e(l-1) = 0.0_dr
do k = l, m
t1 = s(k)
call drotg(t1, f, cs, sn)
s(k) = t1
f = -sn*e(k)
e(k) = cs*e(k)
if (wantu) then
call drot1(n,u(1:,k), &
u(1:,l-1),cs, sn)
end if
end do
go to 610
540 continue
scale = max(abs(s(m)), &
abs(s(m-1)),
abs(e(m-1)), abs(s(l)), &
abs(e(l)))
sm = s(m)/scale
smm1 = s(m-1)/scale
emm1 = e(m-1)/scale
sl = s(l)/scale
el = e(l)/scale
b = ((smm1+sm)*(smm1- &
sm)+ emm1**2)/2.0_dr
c = (sm*emm1)**2
shift = 0.0_dr

```



```

if (b == 0.0_dr .and. &
    c == 0.0_dr) then
    go to 550
endif
shift = sqrt(b**2+c)
if (b < 0.0_dr) then
    shift = -shift
endif
shift = c/(b + shift)
550 continue
f = (sl + sm)*(sl - sm) - shift
g = sl*el
mm1 = m - 1
do k = 1, mm1
    call drotg (f, g, cs, sn)
    if (k /= 1) then
        e(k-1) = f
    endif
    f = cs*s(k) + sn*e(k)
    e(k) = cs*e(k) - sn*s(k)
    g = sn*s(k+1)
    s(k+1) = cs*s(k+1)
    if (wantv) then
        call drotl(p, v(1:,k), v(1:,k+1),
            cs, sn)
    endif
    call drotg(f, g, cs, sn)
    s(k) = f
    f = cs*e(k) + sn*s(k+1)
    s(k+1) = -sn*e(k) +
        cs*s(k+1)
    g = sn*e(k+1)
    e(k+1) = cs*e(k+1)
    if (wantu .and. k < n) then
        call
        drotl(n,u(1:,k),u(1:,k+1),cs,sn)
    endif
end do
e(m-1) = f
iter = iter + 1
go to 610
570 continue
if (s(l) >= 0.0_dr) then
    go to 590
endif
s(l) = -s(l)
if (wantv) then
    v(1:p,l) = -v(1:p,l)
endif
590 continue
do
    if (l == mm) then
        go to 600
    endif
    if (s(l) >= s(l+1)) then
        go to 600
    endif
    t = s(l)
    s(l) = s(l+1)
    s(l+1) = t

    if (wantv .and. l < p) then
        call dswap1(p, v(1:,l),
            v(1:,l+1))
    endif
    if (wantu .and. l < n) then
        call dswap1(n,u(1:,l),u(1:,l+1))
    endif
    l = l + 1
end do
600 continue
iter = 0
m = m - 1
610 continue
end do
620 continue
diag=0
do j=1,p
    diag(j,j)=s(j)
enddo
return
end subroutine svd
!-----
end module svd_computation

```

## **CURRICULUM VITAE**

Hasan TATLI was born in Arguvan/Malatya on October 1st, 1967. He graduated from Diyarbakır Anatolian High School in 1986, attended Faculty of Medicine at Dicle University. However, he had to change the University due to the financial problems, and he attended the Department of Meteorology Engineering of Istanbul Technical University (ITU.) in 1991, graduating in 1995. He obtained the MSc degree from the Institute of Science and Technology of ITU in 1997. The same year, he was appointed as Teaching Assistant in the Department of Meteorology Engineering at ITU and began his doctoral studies at ITU, the Institute of Science and Technology with major in Meteorology Engineering.

