

Climate Change Modeling, Mitigation, and Adaptation



ASCE

EDITED BY

Ken E. Salway
Yueqi Zhang, C. D. Topp
D. D. George, R. D. Topp
and C. M. Rice



Impact of Climate Change on Hydrometeorological Variables in a River Basin in India for IPCC SRES Scenarios

Aavudai Anandhi, V. V. Srinivas, and D. Nagesh Kumar

12.1 Introduction

In the recent years, the new paradigm of abrupt climate change has been well established, and a major global concern is to assess implications of climate change on hydrology of river basins and availability of water, which is considered to be a vulnerable resource. The future climate is unknown and uncertain. Hence to evaluate plausible impacts of climate change on the hydrology of a river basin, it is necessary to develop plausible future projections of hydrometeorological processes in the river basin for various climate scenarios. For this purpose, a variety of methods are available. The classical methods use climate variables simulated by General Circulation Models (GCMs) for projected changes in GCM boundary conditions based on emissions scenarios. Among the scenarios available in literature, those that were published in the Special Report on Emissions Scenarios (SRES) (Nakicenovic et al. 2000) are widely used, and are known as SRES scenarios.

The GCMs are among the most advanced tools, which use transient climate simulations to simulate the climatic conditions on earth, several decades into the future. For quantitative climate impact studies in hydrological processes, the various projections of variables output from the GCM simulations are studied. Since GCMs are run at coarse resolutions, the output climate variables from these models cannot be used directly for impact assessment on a local scale. Hence in the past two decades, several downscaling methodologies have been developed to transfer information from the GCM simulated climate variables to local scale.

The remainder of this chapter is structured as follows. First, a brief description of SRES scenarios, downscaling methods and Support Vector Machine (SVM) is provided in section 12.2. The description of Malaprabha river basin in India, which is considered for case study, is provided in section 12.3. Subsequently, the SVM based methodology suggested for downscaling precipitation and temperature in the river basin is presented in section 12.4. Following this, results of downscaling models are

presented in section 12.5, and possible consequences on hydrology of the river basin are discussed. Finally, summary and conclusions drawn based on the study and some of the conceptual and philosophical issues concerning the use of downscaling models are provided in section 12.6.

12.2 Background

In this section a general description on the various SRES scenarios, downscaling methods and SVM is provided.

12.2.1 SRES

The SRES scenarios are constructed based on the major driving forces or factors (e.g., human development including economic, demographic, social and technological changes) that are suited for climate impact assessment. These factors play significant role in energy consumption, land use changes and emissions, and represent a diverse range of different development pathways of the world for impact assessment. Hence they are useful for research on sustainable development and impact assessment, serving as inputs for evaluating climatic and environmental consequences of future greenhouse gas emissions and for assessing alternative mitigation strategies. These SRES scenarios were constructed with different ranges for each projection called “storyline.” There are four storylines (A1, A2, B1 and B2), describing the way the world population, land use changes, new technologies, energy resources, economies and political structure may evolve over the next few decades. Thus different world futures are represented in two dimensions, with one dimension representing economic or environmental concerns, and the other representing global or regional development patterns. For each storyline several emission scenarios were constructed, producing four “scenario families.” Ultimately, six SRES marker scenarios were defined: A1 has three marker scenarios (A1B, A1FI and A1T) and the others have one each.

A1 Story-line. This scenario represents very rapid economic growth with increasing globalization, an increase in general wealth, with convergence between regions and reduced differences in regional per capita incomes. Materialist–consumerist values will be predominant, with rapid technological change and low population growth when compared to A2 scenario. Three variants within this family make different assumptions about sources of energy for this rapid growth: fossil intensive (A1FI), non-fossil fuels (A1T), and a balance across all sources (A1B).

A2 Story-line. A2 scenario is represented as a heterogeneous, market-led world, with rapid population growth but less rapid economic growth than A1. The underlying theme is self reliance and preservation of local identities. Economic growth is regionally oriented, and hence both income growth and technological change are regionally diverse. Fertility patterns across regions converge slowly, resulting in high population growth.

B1 Story-line. This scenario represents low population growth as A1, but development takes a much more environmentally sustainable pathway with global-scale cooperation and regulation. Clean and efficient technologies are introduced. The emphasis is on global solutions to achieving economic, social and environmental sustainability.

B2 Story-line. B2 scenario represents population increase at a lower rate than A2, but at a higher rate than A1, with development following environmentally, economically and socially sustainable local oriented pathways.

In SRES, none of the presented scenarios explicitly assumes implementation of the United Nations Framework Convention on Climate Change or the emissions target of the Kyoto Protocol. They exclude even the outlying “surprise” or “disaster” scenarios. It is preferable to consider a range of scenarios for climate impact studies as such an approach better reflects the uncertainties of the possible future climate change. For the case study presented in this chapter, A1B, A2, B1 and COMMIT scenarios were considered. In the COMMIT scenario, the atmospheric carbon-dioxide concentrations are maintained (‘Committed’) at the same level as in the year 2000.

12.2.2 Methods of Downscaling

The various downscaling methods available in literature can be broadly classified as dynamic downscaling and statistical downscaling (Figure 12.1).

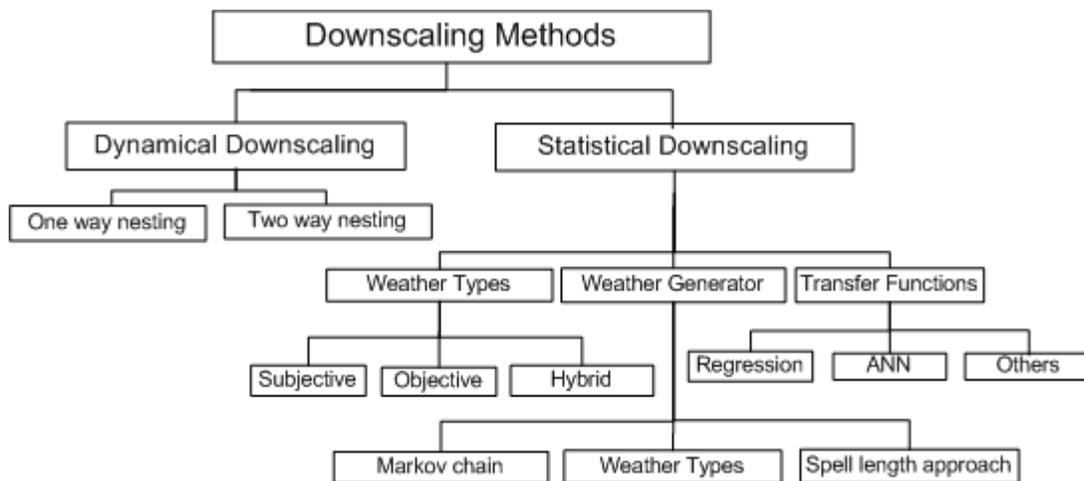


Figure 12.1. Methods of downscaling

In the dynamic downscaling method, a Regional Climate Model (RCM) is embedded into GCM. There are two types of dynamic downscaling based on the types of nesting: one-way nesting and two-way nesting (Wang et al. 2004). One-way nesting consists of driving a limited-area high-resolution RCM with low-resolution data obtained previously by a GCM or by analyses of atmospheric observations. The

one-way nesting technique does not allow feedback from the RCM to the driving data. In two-way nesting, the RCM is run simultaneously with the host GCM, and it regularly updates the host GCM in the RCM region. Models of this type are typically developed using different numeric and physical parameterizations. They are not presently in use as they are cumbersome. Benefits similar to “two-way nesting” can be derived from the use of a variable-resolution GCM.

Statistical downscaling involves developing quantitative relationships between large-scale atmospheric variables (predictors) and local surface variables (predictands). There are three types of statistical downscaling namely- weather types, weather generators and transfer functions.

Weather types or weather classification methods group the days into a finite number of discrete weather types or ‘states’ according to their synoptic similarity. These methods in turn are classified as subjective, objective, or hybrid.

In subjective classification methods, the classifications were carried out manually using empirical rules. Some of the most widely known subjective classifications are Grosswetterlagen (Hess and Brezowsky 1969) and British Isles Weather Types (Lamb 1972).

In objective classification methods, a variety of automated techniques developed using computers are used to group the weather types. The most popular objective classification methods are based on correlation based algorithms (Brinkmann 1999), clustering techniques (Huth et al. 1993; Kidson 1994) and Fuzzy rules based approaches (Wetterhall et al. 2005).

The hybrid techniques combine elements of empirical/manual and automated procedures for grouping weather types, thereby avoiding time delay and enabling the production of easily reproducible and interpretable results (Frakes and Yarnal 1997; Anandhi 2010). Some of the hybrid techniques are screening discriminant analysis (Enke and Spekat 1997) and Classification and Regression Trees (CART) (Breiman et al. 1984).

Weather generators are statistical models of observed sequences of weather variables. They can also be regarded as complex random number generators, the output of which resembles daily weather data at a particular location. There are three fundamental types of daily weather generators, based on the approach to modeling the daily precipitation occurrence: the Markov chain approach (Hughes et al. 1999), the spell-length approach (Wilks 1999) and weather types (Conway and Jones 1998). In the Markov chain approach, a random process is constructed which determines a day at a station as rainy or dry, conditional upon the state of the previous day, following given probabilities. If a day is determined as rainy, then the rainfall amount is drawn from yet another probability distribution. In case of the spell-length approach, instead of simulating rainfall occurrences day by day, the models operate

by fitting probability distribution to observed relative frequencies of wet and dry spell lengths.

Transfer functions are a conceptually simple means for representing linear and nonlinear relationships between the predictors and predictands. Therefore, a diverse range of statistical downscaling methods using transfer functions have been developed in the recent past. Examples include transfer functions based on linear and nonlinear regression, artificial neural networks, canonical correlation analysis, principal component analysis, Support vector machine (Tripathi et al. 2006; Anandhi et al. 2008; Anandhi et al. 2009) and Relevant vector machine (Ghosh and Mujumdar 2008). Transfer function based downscaling methods are sensitive to the subjective choices made in their design, viz., the type of transfer function used, choice of predictors and how well they are simulated by GCM, type of predictand, calibration period, timescale of downscaling (e.g., annual, seasonal, monthly, or daily), and temporal variation of the relationship between the predictors and predictand. However, transfer function methods have generally not been subjected to careful evaluation as the other downscaling techniques (Winkler et al. 1997; Anandhi et al. 2008; Anandhi et al. 2009). In spite of this, transfer functions are most commonly used for downscaling due to relative ease of their application. Individual downscaling schemes differ according to the choice of mathematical transfer function, predictor variables or statistical fitting procedure (Conway et al. 1996; Schubert and Henderson-Sellers, 1997).

Regression-based downscaling methods rely on the direct quantitative relationship between the local scale climate variable (predictand) and the variables containing the larger scale climate information (predictors) through some form of regression function (Karl et al. 1990; Wigley et al. 1990). The main advantage of the regression-based downscaling methods is the relative ease of their application. However, these models often explain only a fraction of the observed climate variability as they are unable to capture the extremes, especially when the predictand is precipitation (Wilby et al. 2004; Tripathi et al. 2006; Anandhi et al. 2008). Downscaling future extreme hydrologic events using regression based models may be problematic, because these events usually tend to be situated at the margins or beyond the range of the extremes in the calibration data set (Wilby et al. 2002).

Artificial Neural Network (ANN) based downscaling techniques have gained wide recognition owing to their ability to capture nonlinear relationships between predictors and predictand (Tatli et al. 2005). Mathematically, an ANN is often viewed as a universal approximator. The ability to generalize a relationship from given patterns makes it possible for ANNs to solve large-scale complex problems such as pattern recognition, nonlinear modeling and classification. The ANNs have been extensively used in a variety of physical science applications, including hydrology (ASCE Task Committee 2000; Govindaraju and Rao 2000).

Despite a number of advantages, the traditional neural network models have several drawbacks including possibility of getting trapped in local minima and

subjectivity in the choice of model architecture (Suykens 2001). (Vapnik 1995; Vapnik 1998) pioneered the development of a novel machine learning algorithm, called support vector machine (SVM), which provides an elegant solution to these problems. The SVM has found wide application in the field of pattern recognition and time series analysis. Introductory material on SVM is available in a number of books (Cortes and Vapnik 1995; Vapnik 1995; Schölkopf et al. 1998; Vapnik 1998; Cristianini and Shawe-Taylor 2000; Haykin 2003; Sastry 2003). Most of the traditional neural network models seek to minimize the training error by implementing the empirical risk minimization principle, whereas the SVMs implement the structural risk minimization principle, which attempts to minimize an upper bound on the generalization error, by striking a right balance between the training error and the capacity of the machine (i.e., the ability of the machine to learn any training set without error). The solution of traditional neural network models may tend to fall into a local optimal solution, whereas global optimum solution is guaranteed in SVM (Haykin 2003). Further, the traditional ANNs have considerable subjectivity in model architecture, whereas for SVMs the learning algorithm automatically decides the model architecture (number of hidden units). Moreover, traditional ANN models do not give much emphasis on generalization performance, while SVMs seek to address this issue in a rigorous theoretical setting. The flexibility of the SVM is provided by the use of kernel functions that implicitly map the data to a higher, possibly infinite, dimensional space. A linear solution, in the higher dimensional feature space, corresponds to a non-linear solution in the original lower dimensional input space. This makes SVM a plausible choice for solving a variety of problems in hydrology, which are non-linear in nature.

12.2.3 Least-Square Support Vector Machine (LS-SVM)

The Least-Square Support Vector Machine (LS-SVM) provides a computational advantage over standard SVM (Suykens 2001). This subsection presents the underlying principle of the LS-SVM and is extracted from Anandhi et al. (2008) and Tripathi et al. (2006).

Consider a finite training sample of N patterns $\{(\mathbf{x}_i, y_i), i = 1, \dots, N\}$, where \mathbf{x}_i representing the “ i -th” pattern in n -dimensional space (i.e., $\mathbf{x}_i = [x_{i1}, \dots, x_{in}] \in \mathcal{R}^n$) constitutes the input to LS-SVM, and $y_i \in \mathcal{R}$ is the corresponding value of the desired model output. Further, let the learning machine be defined by a set of possible mappings $\mathbf{x} \mapsto f(\mathbf{x}, \mathbf{w})$, where $f(\cdot)$ is a deterministic function which, for a given input pattern \mathbf{x} and adjustable parameters \mathbf{w} ($\mathbf{w} \in \mathcal{R}^n$), always gives the same output. The training phase of the learning machine involves adjusting the parameter \mathbf{w} . These parameters are estimated by minimizing the cost function $\Psi_L(\mathbf{w}, e)$.

$$\Psi_L(\mathbf{w}, e) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{1}{2} C \sum_{i=1}^N e_i^2$$

subject to the equality constraint

$$y_i - \hat{y}_i = e_i \quad i = 1, \dots, N \quad (\text{Eq. 12.1})$$

$$\hat{y}_i = \mathbf{w}^T \phi(x) + b \quad (\text{Eq. 12.2})$$

where C is a positive real constant, and \hat{y}_i is the actual model output. The first term of the cost function represents weight decay or model complexity-penalty function. It is used to regularize the weight sizes and to penalize the large weights. This helps in improving the generalization performance (Hush and Horne 1993). The second term of the cost function represents penalty function.

The solution of the optimization problem is obtained by considering the Lagrangian as

$$L(\mathbf{w}, b, \mathbf{e}, \boldsymbol{\alpha}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{1}{2} C \sum_{i=1}^N e_i^2 - \sum_{i=1}^N \alpha_i \{ \hat{y}_i + e_i - y_i \} \quad (\text{Eq. 12.3})$$

where α_i are Lagrange multipliers, and b is the bias term defined in eq. 2. The conditions for optimality are given by

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^N \alpha_i \phi(\mathbf{x}_i) = 0 \\ \frac{\partial L}{\partial b} = \sum_{i=1}^N \alpha_i = 0 \\ \frac{\partial L}{\partial e_i} = \alpha_i - C e_i = 0 \quad i = 1, \dots, N \\ \frac{\partial L}{\partial \alpha_i} = \hat{y}_i + e_i - y_i = 0 \quad i = 1, \dots, N \end{array} \right. \quad (\text{Eq. 12.4})$$

The above conditions of optimality can be expressed as the solution to the following set of linear equations after elimination of \mathbf{w} and \mathbf{e}_i .

$$\begin{bmatrix} 0 & \bar{\mathbf{1}}^T \\ \bar{\mathbf{1}} & \boldsymbol{\Omega} + C^{-1} \mathbf{I} \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \quad (\text{Eq. 12.5})$$

$$\text{where } \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}; \bar{\mathbf{1}} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{N \times 1}; \boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix}; \mathbf{I} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}_{N \times N}$$

In Eq. 12.5, $\boldsymbol{\Omega}$ is obtained from the application of Mercer's theorem.

$$\Omega_{i,j} = K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j) \quad \forall i, j \quad (\text{Eq. 12.6})$$

where $\phi(\cdot)$ represents nonlinear transformation function defined to convert a non-linear problem in the original lower dimensional input space to linear problem in a higher dimensional feature space.

The resulting LS-SVM model for function estimation is:

$$f(\mathbf{x}) = \sum \alpha_i^* K(\mathbf{x}_i, \mathbf{x}) + b^* \quad (\text{Eq. 12.7})$$

where α_i^* and b^* are the solutions to Eq. 12.5 and $K(\mathbf{x}_i, \mathbf{x})$ is the inner product kernel function defined in accordance with Mercer's theorem (Mercer 1909; Courant and Hilbert 1970) and b^* is the bias. There are several choices of kernel functions, including linear, polynomial, sigmoid, splines and Radial basis function (RBF). The linear kernel is a special case of RBF (Keerthi and Lin 2003). Further, the sigmoid kernel behaves like RBF for certain parameters (Lin and Lin 2003). In this study RBF is chosen to map the input data into higher dimensional feature space, which is given by:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma}\right) \quad (\text{Eq. 12.8})$$

where, σ is the width of RBF kernel, which can be adjusted to control the expressivity of RBF. The RBF kernels have localized and finite responses across the entire range of predictors.

The advantage with RBF kernel is that it maps the training data non-linearly into a possibly infinite-dimensional space, and thus, it can effectively handle the situations when the relationship between predictors and predictand is nonlinear. Moreover, the RBF is computationally simpler than polynomial kernel, which requires more parameters. It is worth mentioning that developing LS-SVM with RBF kernel involves a judicious selection of RBF kernel width σ and parameter C.

12.3 Study Region and Data Used

The study region is the catchment of Malaprabha River, upstream of Malaprabha reservoir in India. The region covers an area of 2093.46 km² situated between 15° 30' N and 15° 56' N latitudes, and 74° 12' E and 75° 8' E longitudes. It lies in the extreme western part of the Krishna River basin in India, and includes parts of Belgaum, Bagalkot and Dharwad districts of North Karnataka (Figure 12.2). Analysis of temporal variation of rainfall showed that, in general, the climate of the study region is dry, except in monsoon months (June–September) when warm winds blowing from Indian Ocean cause copious amount of rainfall. Isohyetal map prepared for the region showed considerable variation in spatial distribution of annual rainfall. Heavy rainfall (more than 3000 mm) is recorded at gauging stations in the upstream reaches of the Malaprabha catchment, which forms a part of the western Ghats. In

contrast, the average annual rainfall in the reservoir command area (i.e., downstream of the dam) is 576 mm. The average annual rainfall in the basin is 1051 mm. It may be noted that the Malaprabha River originates in a region of high rainfall, and it is the main source of surface water for arid and semi-arid regions downstream of Malaprabha reservoir.

The data adopted for this study consists of monthly mean atmospheric variables simulated by Canadian Center for Climate Modeling and Analysis's (CCCma) third generation Coupled Global Climate Model (CGCM3). The data comprised of the 20th century simulations (20C3M) for the period of 1971–2000, and future simulations forced by four SRES scenarios namely, A1B, A2, B1 and COMMIT for the period of 2001–2100. Reanalyzed data of the monthly mean atmospheric variables prepared by National Centers for Environmental Prediction (NCEP) for the period 1971–2000 were used. The data on observed precipitation were obtained from the Department of Economics and Statistics, Government of Karnataka, India, for the period of 1971–2000. The data on observed temperature were obtained from India Meteorological Department (IMD) for the period of 1978–2000. The details of the data are furnished in Table 12.1. For the sake of analysis, the GCM data were re-gridded to NCEP grid using Grid Analysis and Display System (GrADS) (Doty and Kinter 1993).

12.4 Methodology

The development of a downscaling model begins with the selection of probable predictors, followed by their stratification (which is optional and variable dependant), and training and validation of the model. The developed model is subsequently used to obtain projections of predictand for simulations of GCM.

12.4.1 Selection of Probable Predictors

The selection of appropriate predictors is one of the most important steps in a downscaling exercise (Fowler et al. 2007). The choice of predictors could vary from region to region depending on the characteristics of the large-scale atmospheric circulation and the predictand to be downscaled. Any type of variable can be used as predictor as long as it is reasonable to expect that there exists a relationship between the variable and the predictand. Often, in climate impact studies, only such variables are chosen as predictors that are: (i) reliably simulated by GCMs and are readily available from archives of GCM output and reanalysis data sets; (ii) strongly correlated with the predictand; and (iii) based on previous studies. The number of probable predictors is referred to as m_1 in this chapter.

12.4.2 Stratification of Predictors

For the sake of stratification of predictors, the m_2 climate variables (potential predictors), which are realistically simulated by the GCM, were selected from the m_1

probable predictors, by specifying a threshold value (T_{ng1}) for correlation between the probable predictor variables in NCEP and GCM data sets. For the estimation of correlation, product moment correlation (Pearson 1896), Spearman's rank correlation (Spearman 1904a and b) and Kendall's tau (Kendall 1951) were considered.

Table 12.1 The details of the meteorological data used in the study

Data type	Source of data	Period	Details	Time scale
Observed data of precipitation	Dept. of Economics & Statistics, Government of Karnataka (GOK), India	1971–2000	Data at 11 gauging stations are used to arrive at representative values of precipitation for the basin	Daily
Observed data of temperature	India Meteorological Department (IMD)	1978–2000	Data at 2 gauging stations namely Santhebathewadi and Gadag	Daily
CGCM3 T/47 data on atmospheric variables	http://www.cccma.bc.ec.gc.ca/cgi-bin/data/cgcm3	1971–2100; baseline: 20C3M (1971–2000); future: SRES A1B, A2, B1 & COMMIT (2001–2100).	12 grid points for atmospheric variables, with grid spacing $\approx 3.75^\circ$. Latitudes range: 9.28°N to 20.41°N . Longitudes range: 71.25°E to 78.75°E	Monthly
NCEP re-analysis data of atmospheric variables	Kalnay et al. (1996)	1971–2000	9 grid points for atmospheric variables, with grid spacing 2.5° . Latitudes range: 12.5°N to 17.5°N . Longitudes range: 72.5°E to 77.5°E	Monthly
NCEP re-analysis data of atmospheric fluxes	Kalnay et al. (1996)	1971–2000	16 grid points for atmospheric fluxes with grid spacing 1.9° . Latitudes range: 12.3°N to 20.0°N longitude range : 71.6°E to 77.5°E	Monthly

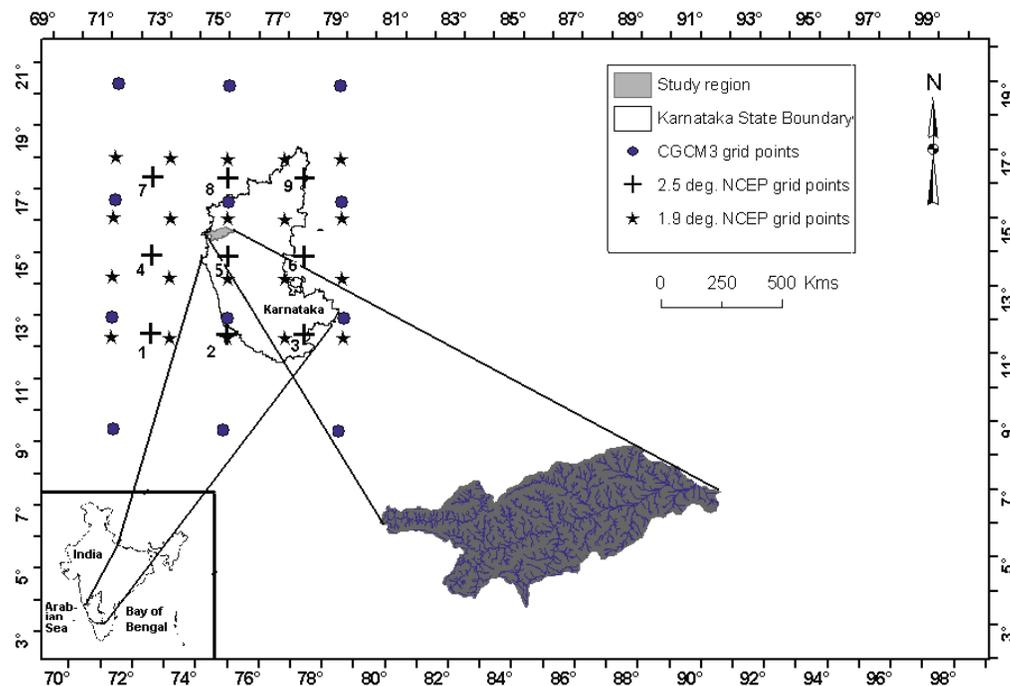


Figure 12.2. Location of the study region in Karnataka State, India. The latitude, longitude and scale of the map refer to Karnataka State. The data extracted at CGCM3 and 1.9° NCEP grid points are re-gridded to the nine 2.5° NCEP grid points. Among the nine grid points 1, 4 and 7 are on Arabian Sea, and the remaining points are on land

Depending on the predictand variable to be downscaled, the stratification of the corresponding potential predictors was carried out in space (land and ocean) or in time (e.g., wet and dry seasons). When precipitation was considered as predictand, the stratification of the predictors was carried out in time domain to form clusters corresponding to wet and dry seasons. When maximum and minimum temperatures were considered as predictands, the stratification of predictors was carried out in space domain. The following part of this subsection outlines finer details on the procedure suggested for stratification of potential predictors in the context of downscaling precipitation and temperature.

Stratification of Predictors for Downscaling Precipitation. The climate of a region can be broadly classified into seasons for analyzing precipitation. The predictor variables for downscaling a predictand could vary from season to season. Further the relationship between the predictor variables and the predictand varies seasonally because of the seasonal variation of the atmospheric circulation (Karl et al. 1990). Hence seasonal stratification has to be performed to select the appropriate predictor variables for each season to facilitate development of a separate downscaling model for each of the seasons. The seasonal stratification can be carried out by defining the seasons as either conventional (fixed) seasons or as “floating” seasons. In fixed season stratification, the starting dates and lengths of seasons remain the same for every year. In contrast, in “floating” season stratification, the date of onset and duration of each season is allowed to change from year to year. Past studies have shown that floating seasons are better than the fixed seasons, as they reflect ‘natural’ seasons, especially under altered climate conditions (Winkler et al. 1997). Therefore identification of the floating seasons under altered climate conditions helps to effectively model the relationships between predictor variables and predictands for each season, thereby enhancing the performance of the downscaling model. Hence, for the case study presented in this chapter, the floating method of seasonal stratification is considered to identify dry and wet seasons in a calendar year for both NCEP and GCM data sets. In the floating method of seasonal stratification, the NCEP data are partitioned into two clusters depicting wet and dry seasons by using the K-means clustering method (MacQueen 1967), whereas the GCM data are partitioned into two clusters by using the nearest neighbor rule (Fix and Hodges 1951).

From NCEP data on the m_2 variables, n principal components (PCs), which preserve more than 98% of the variance, are extracted using principal component analysis (PCA). The PCs corresponding to each month are used to form a feature vector for the month. The PCs are also extracted from GCM data, but along the principal directions obtained for the NCEP data. They are used to form feature vectors for GCM data. Each feature vector (representing a month) can be visualized as an object having a specific location in multidimensional space, whose dimensionality is defined by the number of PCs.

The feature vectors of the NCEP data are partitioned into two clusters (depicting wet and dry seasons) using the K-means cluster analysis. The clustering

should be such that the feature vectors within each cluster are as close to each other as possible in space, and are as far as possible in space from the feature vectors of the other clusters. The distance between each pair of feature vectors in space is estimated using Euclidian measure. Subsequently, each feature vector of the NCEP data is assigned a label that denotes the cluster (season) to which it belongs. Following this, the feature vectors prepared from GCM data (past and future) are labeled using the nearest neighbor rule to get the past and future projections for the seasons. As per this rule, each feature vector formed using the GCM data is assigned the label of its nearest neighbor from among the feature vectors formed using the NCEP data. To determine the nearest neighbors for this purpose, the distance between each pair of NCEP and GCM feature vectors is computed using Euclidean measure. Comparison of the labels of contemporaneous feature vectors formed from NCEP and GCM past data is useful in checking if the GCM simulations represent the regional climate fairly well, during the past period.

Optimal T_{ng1} is identified as a value for which the wet and dry seasons formed for the study region using NCEP data are well correlated with the possible true seasons for the region. For this analysis, the plausible true wet and dry seasons in the study region are identified using a method based on truncation level (TL). In this method, the dry season is considered as consisting of months for which the estimated Theissen Weighted Precipitation (TWP) values for the region are below the specified TL, whereas the wet season is considered as consisting of months for which the estimated TWP values are above the TL. Herein, two options have been used to specify the TL. In the first option, the TLs are chosen as various percentages of the observed mean monthly precipitation (MMP) (70 to 100% of MMP at intervals of 5%). In the second option, the TL is chosen as the mean monthly value of the actual evapotranspiration in the river basin. The actual evapotranspiration is obtained for Krishna basin from Gosain et al. (2006). The potential predictors corresponding to optimal T_{ng1} are noted.

Stratification of Predictors for Downscaling Surface Temperature. The surface temperature in a region is dominated by local effects such as evaporation, sensible heat flux and vegetation in the region. Therefore the potential predictor variables influencing surface temperature in the study region are stratified based on the location of grid points (land and/or ocean) corresponding to the variables, to assess the impact of their use on downscaled temperature. Out of the nine 2.5° NCEP grid points considered in the study region, six are above land and the remaining three are over sea. As there are no distinct seasons based on temperature, seasonal stratification as in the case of precipitation is not relevant.

12.4.3 SVM Downscaling Model

For downscaling the predictand, the m_1 probable predictors at each of the NCEP grid points will be considered as probable predictors. Thus, there are m_3 ($= m_1 \times$ number of NCEP grid points) probable predictor predictors. The potential predictors (m_4) are selected from the m_3 probable predictor variables. For this

purpose, the cross-correlations are computed between the probable predictor variables in NCEP and GCM data sets, and the probable predictor variables in NCEP data set and the predictand. A pool of potential predictors is then identified for each season by specifying threshold values for the computed cross-correlations. The threshold value for cross-correlation between variables in NCEP and GCM data sets is denoted hereafter by T_{ng2} , whereas the same between NCEP variables and predictand is depicted as T_{np} . The T_{np} should be reasonably high to ensure choice of appropriate predictors for downscaling the predictand. Similarly, T_{ng2} should also be reasonably high to ensure that the predictor variables used in downscaling are realistically simulated by the GCM in the past, so that the future projections of the predictand obtained using GCM data would be acceptable.

The downscaling model is calibrated to capture the relationship between NCEP data on potential predictors and the predictand. The data on potential predictors is first standardized for each season or location separately for a baseline period. Such standardization is widely used prior to statistical downscaling to reduce systemic bias (if any) in the mean and variance of the predictors in the GCM data, relative to those of the same in the NCEP reanalysis data (Wilby et al. 2004). This step typically involves subtraction of mean and division by the standard deviation of the predictor for the baseline period. The standardized NCEP predictor variables are then processed using PCA to extract such PCs which are orthogonal and which preserve more than 98% of the variance originally present in them. A feature vector is formed for each month using the PCs. The feature vector forms the input to the SVM model, and the contemporaneous value of predictand is its output. The PCs account for most of the variance in the input and also remove the correlations, if any, among the input data. Hence, the use of PCs as input to a downscaling model helps in making the model more stable and at the same time reduces the computational load.

To develop the SVM downscaling model, the feature vectors formed are partitioned into a training set and a testing set. The partitioning was initially carried out using multifold cross-validation procedure, which was adopted from Haykin (2003) in an earlier work (Tripathi et al. 2006). In this procedure, about 70% of the feature vectors are randomly selected for training the model, and the remaining 30% are used for validation. However, in this study the multifold cross-validation procedure is found to be ineffective because the time span considered for analysis is small and there are more extreme events in the past decades than in the recent decade. Therefore, the feature vectors formed from approximately first 70% of the available data are chosen for calibrating the model and the remaining feature vectors are used for validation. The 'normalized mean square error' is used as an index to assess the performance of the model. The training of SVM involves selection of the model parameters σ and C . The width of RBF kernel σ gives an idea about the smoothness of the derived function. Smola et al. (1998), in their attempt to explain the regularization capability of RBF kernel, have shown that a large kernel width acts as a low-pass filter in frequency domain. It attenuates the higher order frequencies, resulting in a smooth function. Alternately, RBF with a small kernel width retains most of the higher order frequencies leading to an approximation of a complex

function by the learning machine. In this study, grid search procedure (Gestel et al. 2004) is used to find the optimum range for each of the parameters. Subsequently, the optimum values of the parameters are obtained from within the selected ranges, using the stochastic search technique of genetic algorithm (Haupt and Haupt 2004).

The feature vectors prepared from GCM simulations are processed through the validated SVM downscaling model to obtain future projections of the predictand, for each of the four emission scenarios considered (i.e., SRES A1B, A2, B1 and COMMIT). Subsequently, for each scenario, the projected values of the predictand are chronologically divided into five parts (2001–2020, 2021–2040, 2041–2060, 2061–2080 and 2081–2100) to determine the trend in the projected values of the predictand. The procedure is illustrated in the flowchart in Figure 12.3.

12.5 Results

The results of the downscaled precipitation, maximum and minimum temperatures are discussed in this section.

12.5.1 Predictor Selection

For downscaling precipitation, the predictor variables are screened on the twin basis that monsoon rain is dependent on dynamics through advection of water from the surrounding seas and thermodynamics through effects of moisture and temperature, both of which can modify the local vertical static stability. In a changed climate scenario, both the thermodynamic and dynamic parameters may undergo changes. Therefore in the present study, only such probable predictor variables, which incorporate both the effects, are chosen. Winds during south-west monsoon season advect moisture into the region while temperature and humidity are associated with local thermodynamic stability and hence are useful as predictors. Zonal wind is the response to heating in the monsoon trough in the North India. Meridional wind has more local effects, and together the winds are responsible for convergence of moisture and hence related to precipitation. Temperature affects the moisture holding capacity and the pressure at a location. The pressure gradient affects the circulation which in turn affects the moisture brought into the place and hence the precipitation. Higher precipitable water in the atmosphere means more moisture, which in turn causes statically unstable atmosphere leading to more vigorous overturning, resulting in more precipitation. Lower pressure leads to more winds and so more precipitation. At 925 mb pressure height, the boundary layer (near surface effect) is important. The 850 mb pressure height is the low level response to regional precipitation. The 200 mb pressure level depicts the global scale effects. Temperature at 700 mb and 500 mb represent the heating process of the atmosphere due to monsoonal precipitation which is maximum at mid-troposphere on a constant pressure height. Geopotential height represents the pressure variation, which reflects the flow, based on which the moisture changes. Due to these reasons, fifteen probable predictors are extracted from the NCEP reanalysis and CGCM3 data sets. They are the air temperature at 925 mb

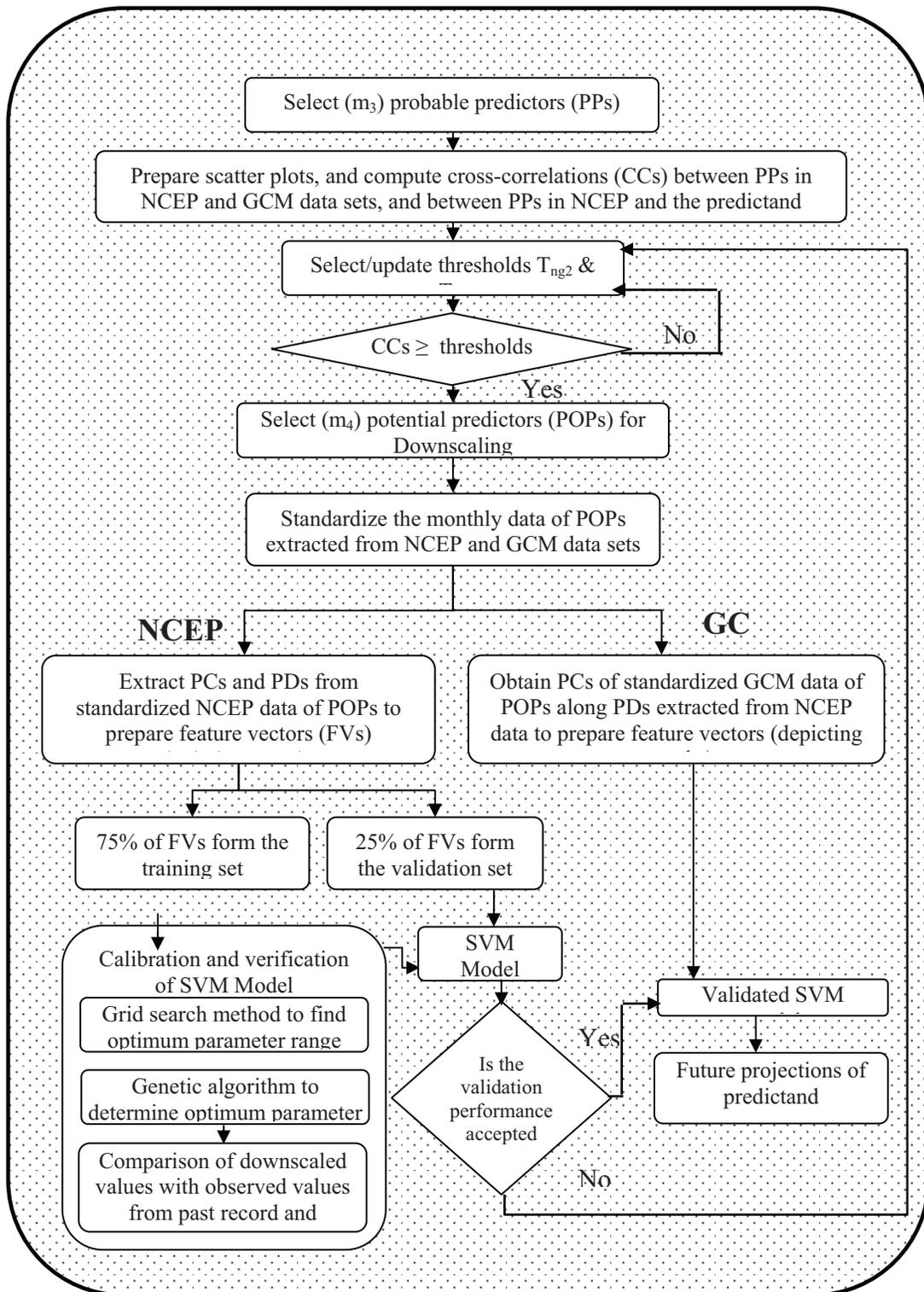


Figure 12.3. Methodology followed for SVM downscaling. PCs and PDs denote principal components and principal directions, respectively. T_{ng2} is the threshold between predictors in NCEP and GCM data sets. T_{np} denotes the threshold between predictors in NCEP data and the predictand

(Ta 925), 700 mb (Ta 700), 500 mb (Ta 500) and 200 mb (Ta 200) pressure levels, geo-potential height at 925 mb (Zg 925), 500 mb (Zg 500) and 200 mb (Zg 200) pressure levels, specific humidity at 925 mb (Hus 925) and 850 mb (Hus 850) pressure levels, zonal (Ua) and meridional wind velocities (Va) at 925 mb (Ua 925, Va 925) and 200 mb (Ua 200, Va 200) pressure levels, precipitable water (prw) and surface pressure (ps).

For downscaling temperature, large scale atmospheric variables, namely air temperature, zonal and meridional wind velocities at 925 mb, which are often used, are considered as predictors. Surface flux variables, namely latent heat, sensible heat, shortwave radiation and longwave radiation fluxes can also be considered for downscaling temperature as they control the temperature of the earth's surface. The incoming solar radiation heats the surface, while latent heat flux, sensible heat flux, and longwave radiation cool the surface. Due to these reasons, seven probable predictors are extracted from the NCEP reanalysis and CGCM3 data sets to downscale temperature. They are air temperature, zonal, and meridional wind velocities at 925 mb, and four fluxes: latent heat (LH), sensible heat (SH), shortwave radiation (SWR), and longwave radiation (LWR).

12.5.2 SVM Downscaling Models

From the selected potential predictors for each season, principal components are extracted to form feature vectors. These feature vectors are provided as input to develop SVM downscaling model following the procedure described in Section 12.4. For obtaining the optimal range of each of the SVM parameters (kernel width σ , and penalty term C), the grid search procedure is used. Typical results of the domain search performed to estimate the optimal ranges of the parameters for wet and dry seasons are shown in Figure 12.4. From this figure, the range of σ and C having the least NMSE (Normalized Mean Square Error) is selected as the optimum parameter range. The NMSE values are indicated in the bar code provided close to the two parts of the figure. Using Genetic algorithm, the optimum parameter is selected from the optimum parameter range. The optimal values of SVM parameters C and σ thus obtained are 550 and 50 for wet season, and 850 and 50 for dry season, respectively. For maximum temperature the optimal values of SVM parameters C and σ are 2050 and 50 while for minimum temperature 1050 and 50 were the optimal values of SVM parameters. The results of downscaling are compared with observed variables and showed in figure 12.5 The details of the downscaled variables were elaborated in Anandhi et al. (2008, 2009).

12.5.3 Projected Future Scenarios

The future projections of three meteorological variables (precipitation, maximum and minimum temperatures) were obtained for each of the four SRES scenarios (A1B, A2, B1 and COMMIT) using the developed SVM downscaling models. The projections were subsequently divided into five 20-year intervals (2001–2020, 2021–2040, 2041–2060, 2061–2080, 2081–2100). The mean monthly values of

observed and projected precipitation for the study area were estimated using the Theissen method. For each of the four SRES scenarios, average of the mean monthly values of Theissen weighted precipitation, maximum and minimum temperatures are presented as bar plots, for all the five 20-year intervals in Figures 12.6, 12.7 and 12.8 respectively. These plots facilitate in assessing the projected changes in each meteorological variable across twenty-year intervals over the period of 2001–2100, with respect to the past (20C3M), for each SRES scenario. Secondly, for each of the five 20-year intervals, the average of the mean monthly values of the aforementioned variables are plotted individually, for all the five scenarios (20C3M, SRES A1B, A2, B1 and COMMIT) in Figures 12.9, 12.10 and 12.11 respectively. These plots facilitate comparison of the past and projected mean monthly values of each meteorological variable across SRES scenarios, for each 20-year interval, and thus, help in assessing the changes in the variables across all the months in a year.

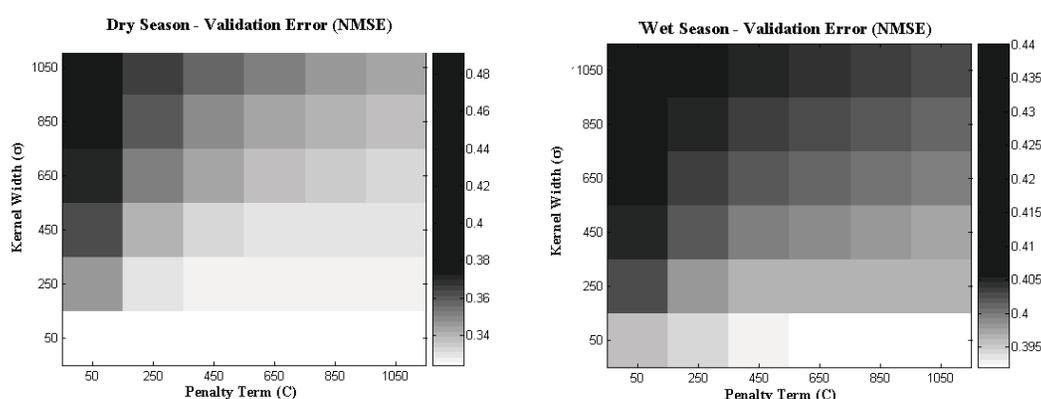


Figure 12.4. Illustration of the domain search performed to estimate optimal values of kernel width (σ) and penalty (C) for the SVM, for dry and wet seasons

From the figures it is observed that precipitation, and maximum and minimum temperatures are projected to increase in future for A1B, A2 and B1 scenarios, whereas no trend is discerned with the COMMIT. The projected increases are high for A2 scenario, whereas they are least for B1 scenario. This is because among the scenarios considered, the scenario A2 has the highest concentration of carbon dioxide (CO_2) equal to 850 ppm, while the same for A1B, B2 and COMMIT scenarios are 720 ppm, 550 ppm and ≈ 370 ppm respectively. Rise in the concentration of CO_2 in atmosphere causes the earth's average temperature to increase, which in turn causes increase in evaporation especially at lower latitudes. The evaporated water would eventually precipitate. In the COMMIT scenario, where the emissions are held the same as in the year 2000, no significant trend in the pattern of projected future precipitation could be discerned.

From a perusal of Figures 12.6, 12.7 and 12.8 it can be observed that, in general, for the meteorological variables, the change from past to future is gradual, and the change is more for A1B scenarios, while it is the least for B1 scenario. In A2 scenario the change is more and different from A1B. In the case of COMMIT no clear pattern change is visible.

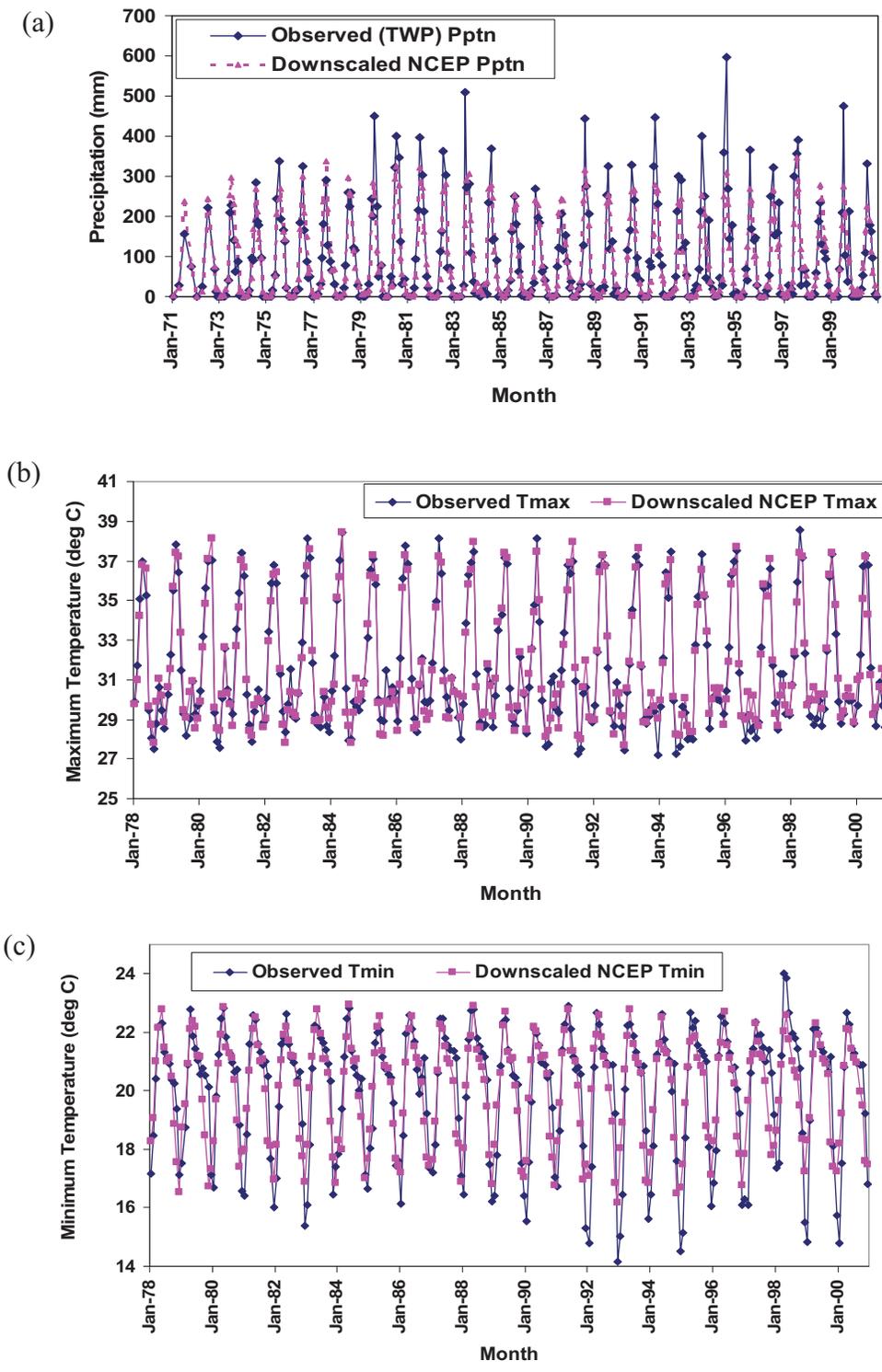


Figure 12.5. Comparison of the monthly observed meteorological variable with the corresponding simulated variable using SVM downscaling model for NCEP data (a) Thiessen weighted precipitation (TWP) (b) maximum temperature (Tmax) (c) minimum temperature (Tmin)

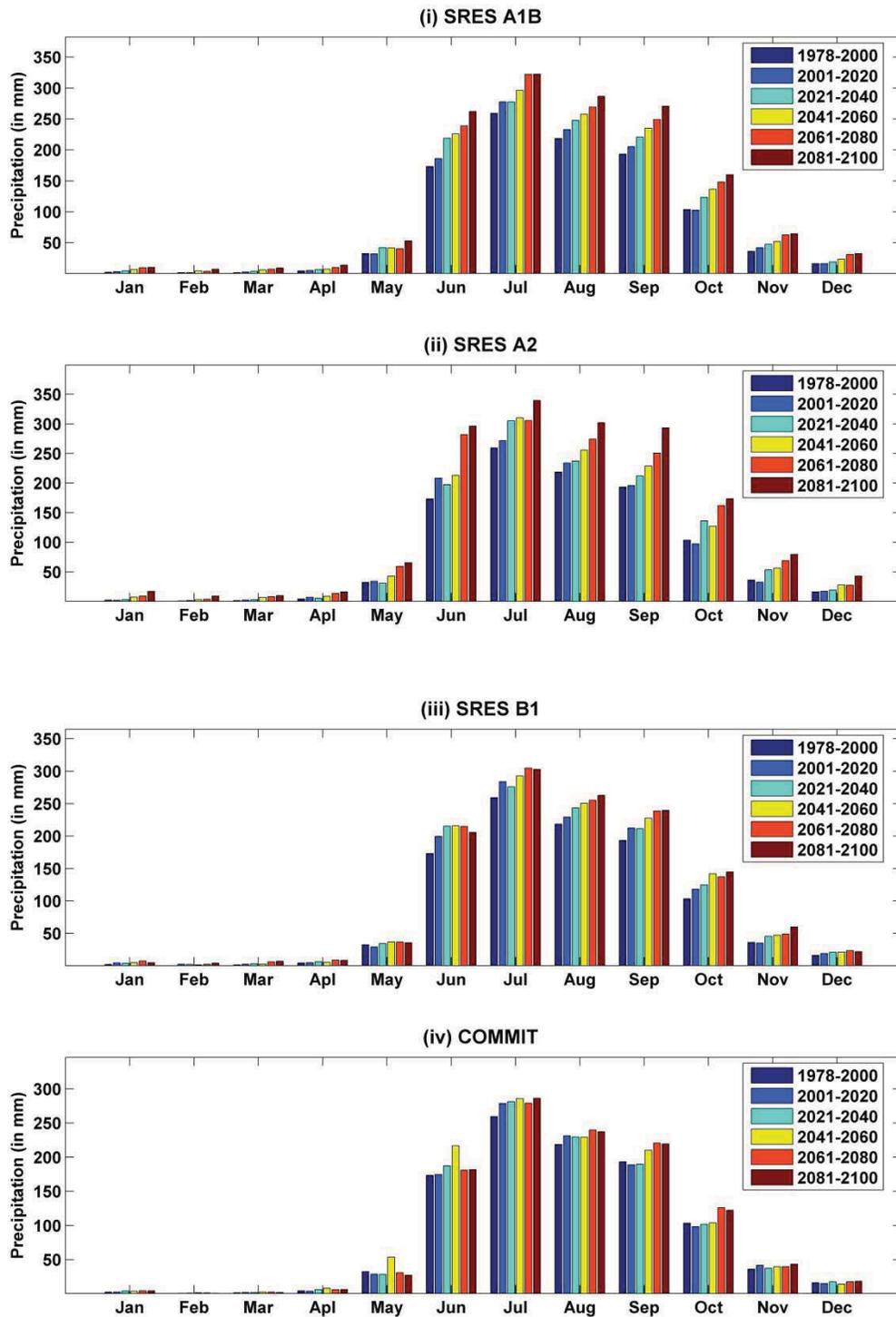


Figure 12.6. Mean monthly precipitation in the study region for the period 1971–2100, for the four scenarios considered

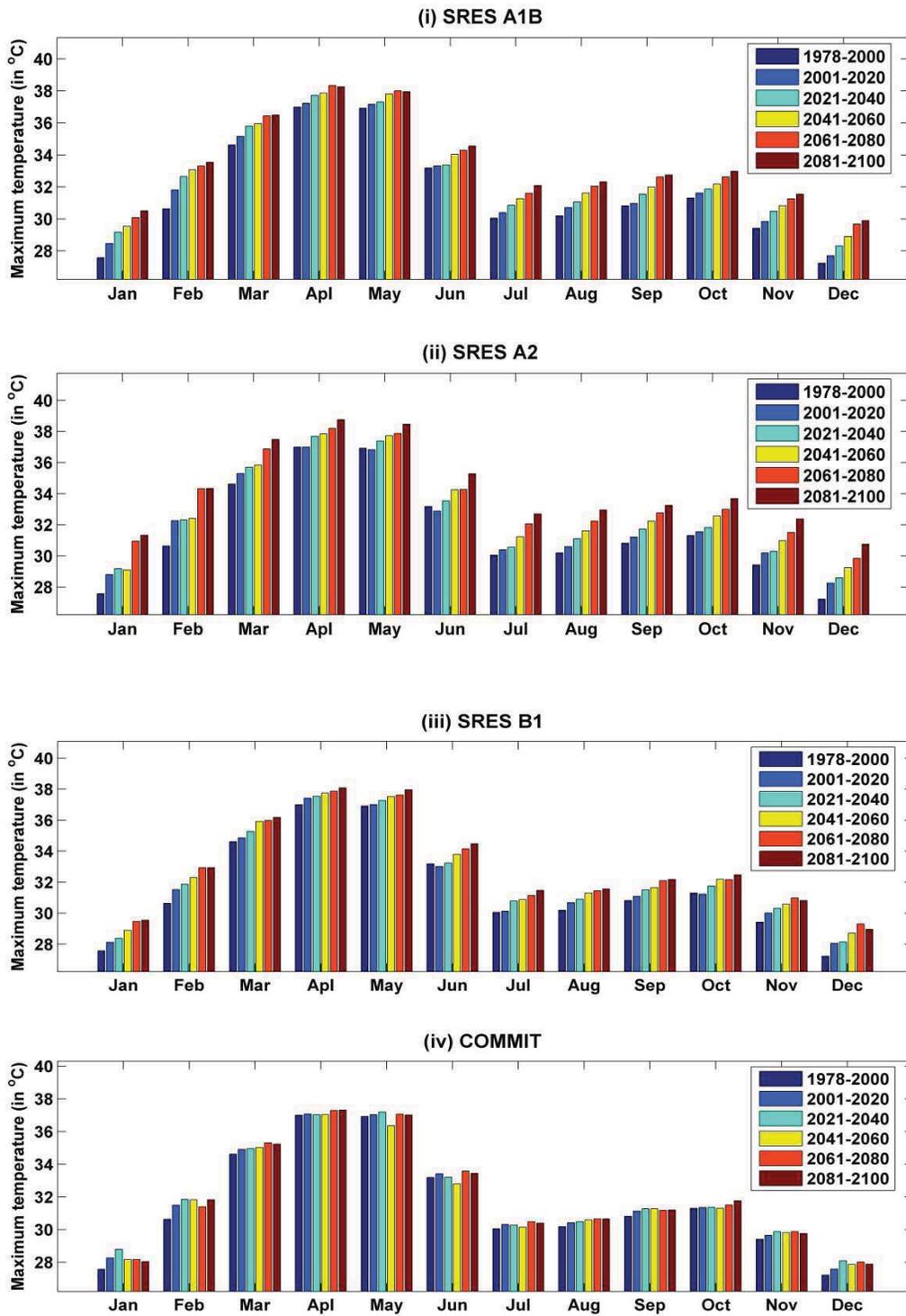


Figure 12.7. Mean monthly maximum temperatures in the study region for the period 1978–2100, for the four scenarios considered

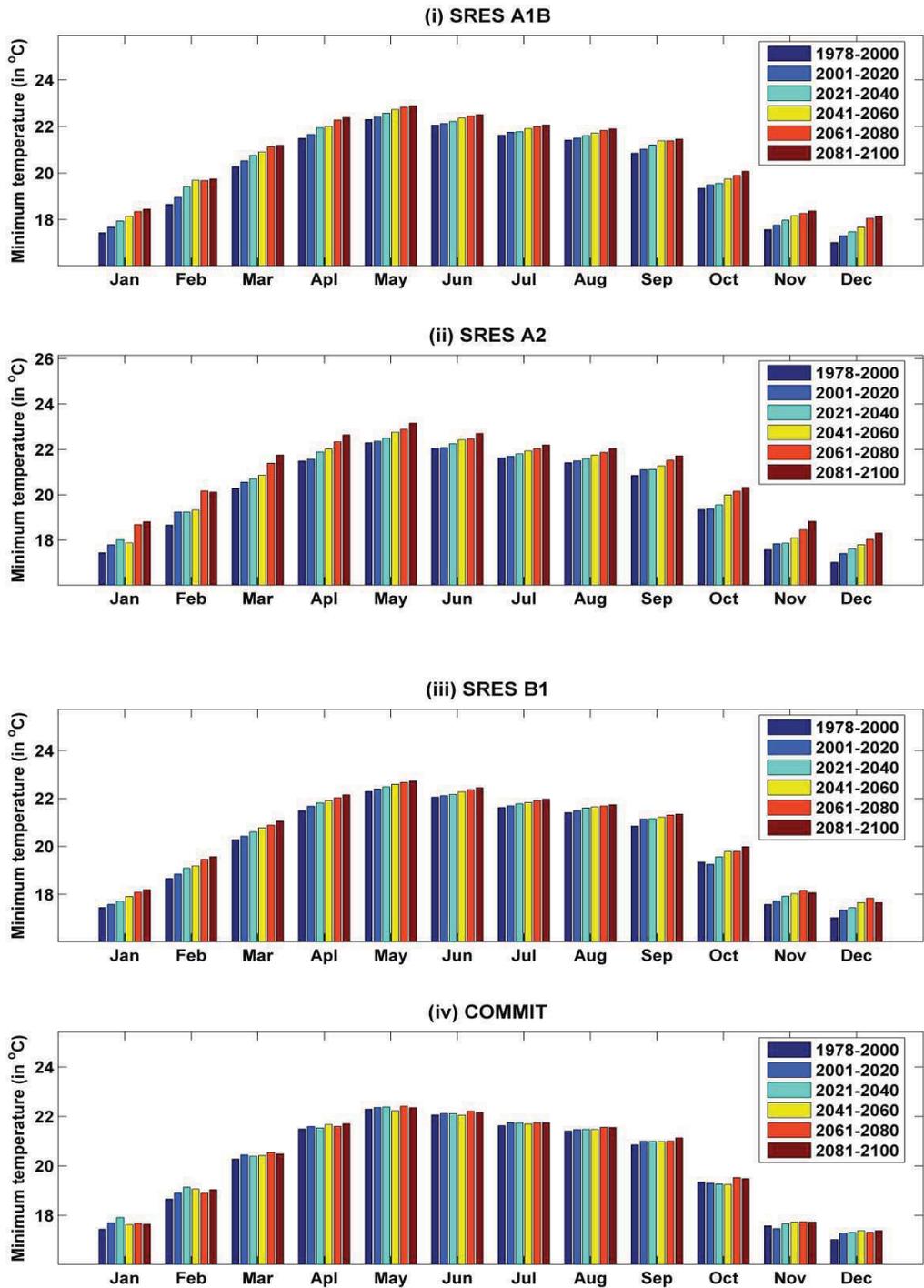


Figure 12.8. Mean monthly minimum temperatures in the study region for the period 1978–2100, for the four scenarios considered

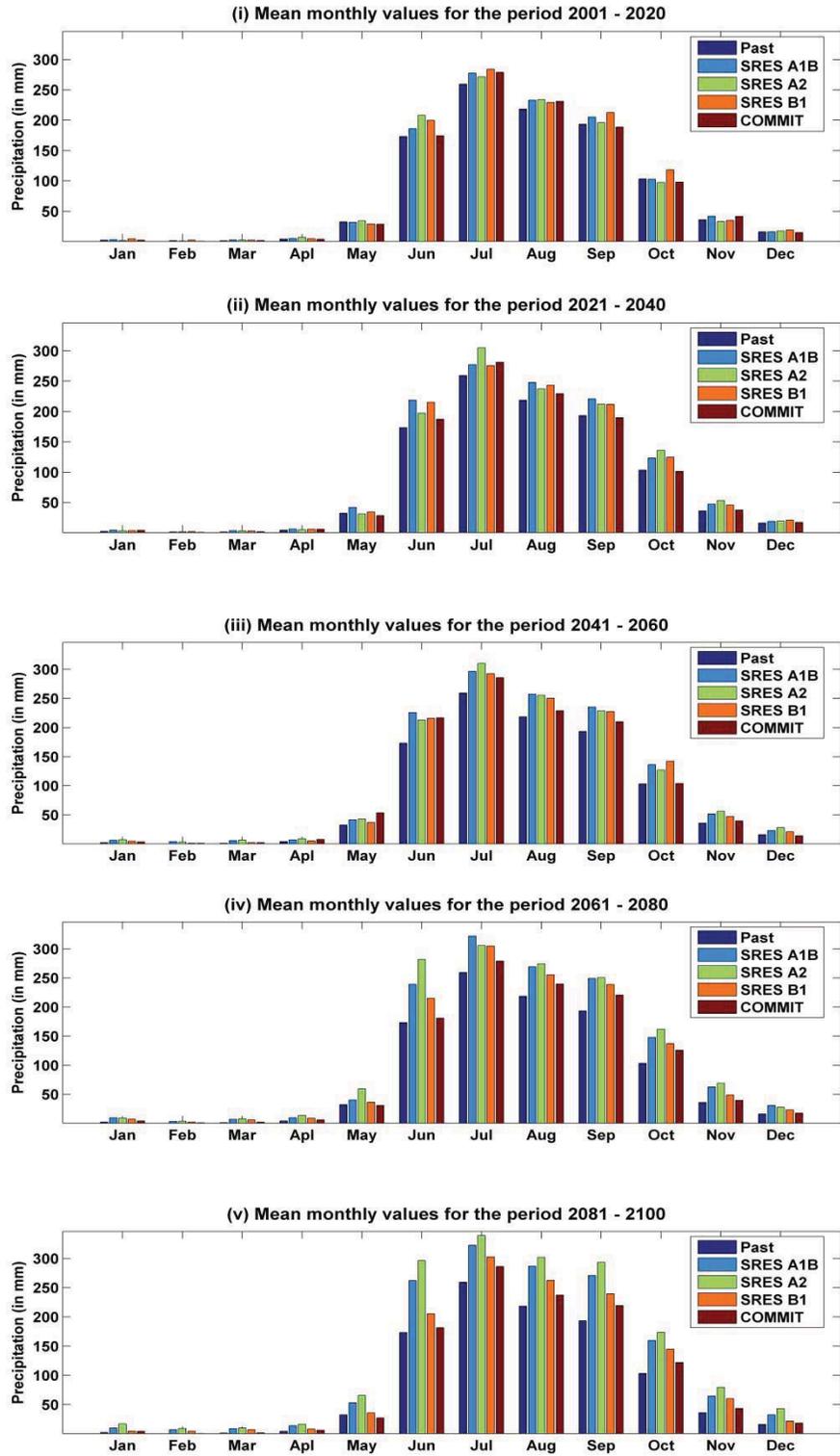


Figure 12.9. Projections obtained for ‘mean monthly precipitation’ in the study region for the four scenarios are compared with the past (20C3M) value of the statistic, for different future periods

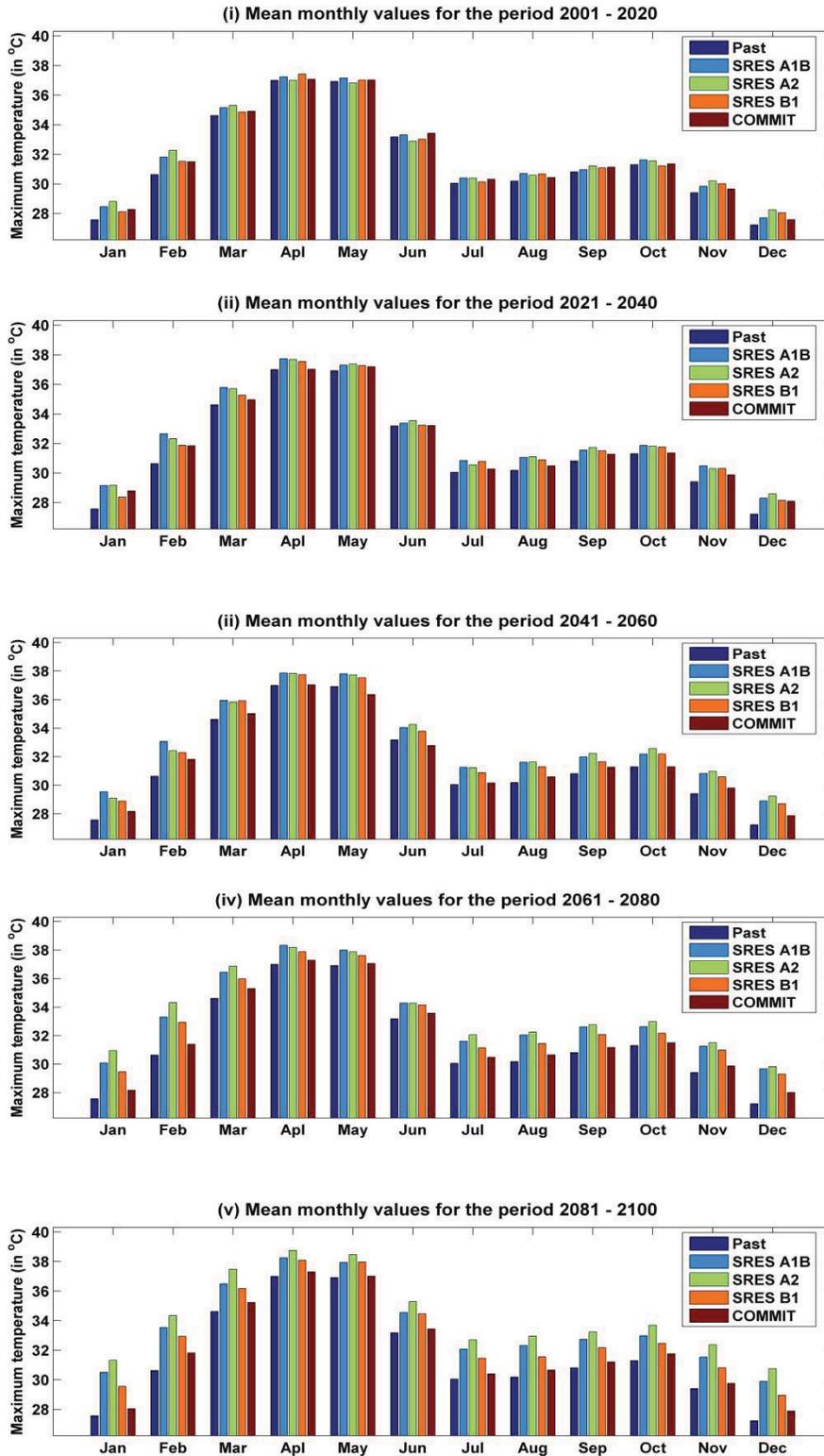


Figure 12.10. Projections obtained for ‘mean monthly maximum temperature’ in the study region for the four scenarios are compared with the past (20C3M) value of the statistic, for different future periods

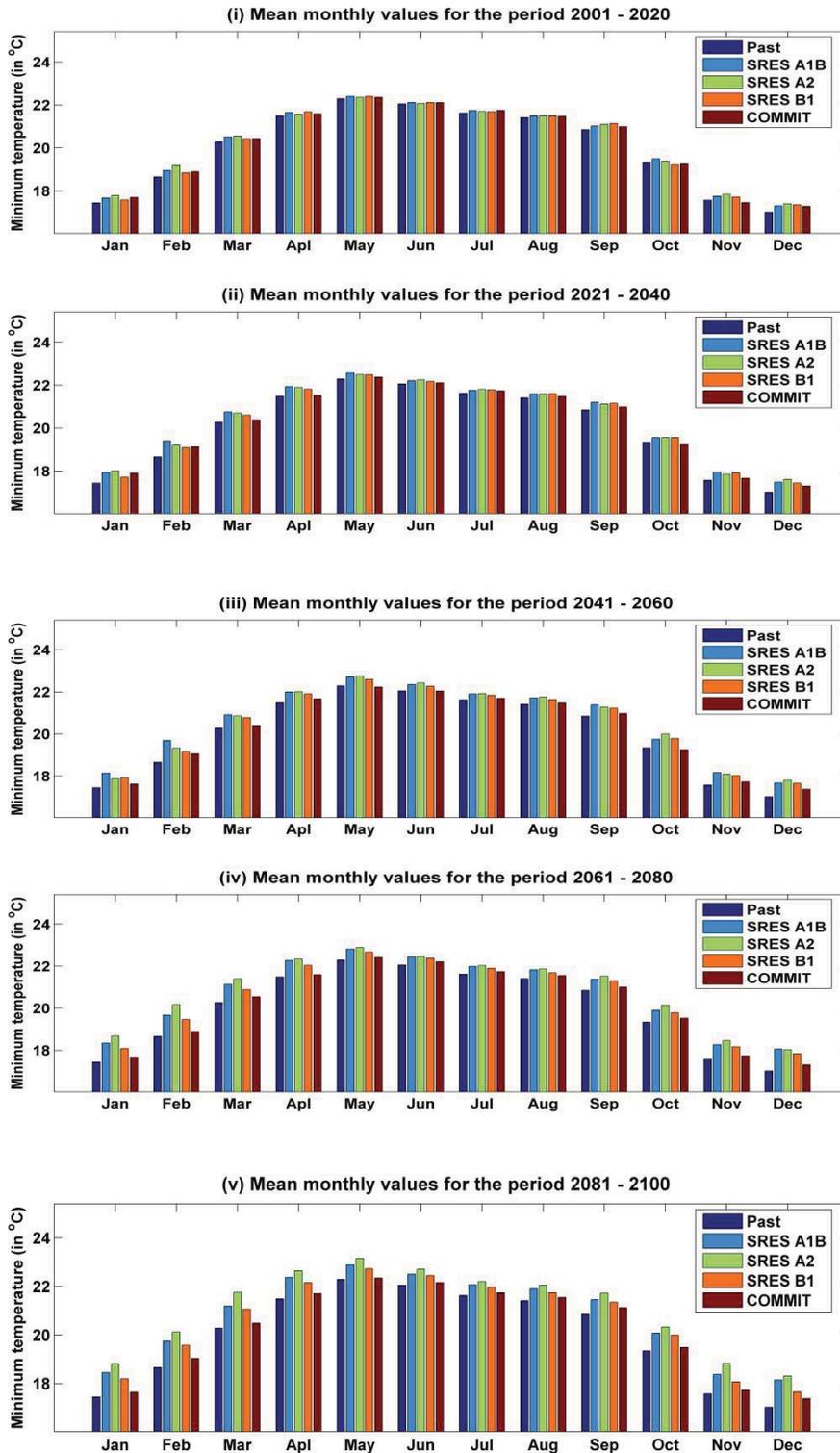


Figure 12.11. Projections obtained for ‘mean monthly minimum temperature’ in the study region for the four scenarios are compared with the past (20C3M) value of the statistic, for different future periods

From the Figures 12.9, 12.10 and 12.11 it can be inferred that the change in the variables is least in the first 20-year interval (2001–2020) and maximum in the last 20-year interval (2081–2100).

12.5.4 Impacts of Climate Change on Hydrology

The variables—precipitation and temperature play an important role in the hydrology of a river basin and are commonly used for impact studies. Some of the possible impacts of changes in the aforementioned variables are discussed in the following part of this subsection.

In general, changes in climate variables (precipitation and temperature) cause changes in the water balance, by changing the various components of hydrologic cycle such as runoff, evapotranspiration, soil moisture, infiltration and groundwater recharge.

Changes in precipitation and temperatures can affect the magnitude and timing of runoff, which in turn affect the frequency and intensity of hydrologic extremes such as floods and droughts. Changes in precipitation could be in the amount, distribution, intensity and frequency. Most of the precipitation in the region occurs in the monsoonal months (June to October). An increased precipitation amount, intensity and frequency during monsoon period could affect the frequency of floods while a decreased precipitation during the period could affect the frequency of drought. In general, the increase in surface temperatures modify the hydrologic cycle through changes in the volume, intensity, or type of precipitation (rain versus snow), and through shifts in the seasonal timing of stream flow (Regonda et al. 2005). In this region, with no snow cover, changes in temperature may not directly affect the runoff, but will cause changes in precipitation patterns and other climate variables and may also affect the evaporation and hence the runoff of the region. The changes in runoff affect the water resources infrastructure such as reservoirs. Reduced flow will mean less supply and potential economic damages, and increased flow may mean an under-designed reservoir or spillway with potential flood risk.

A change in temperature affects the evaporation, evapotranspiration, and desertification processes and is also considered as an indicator of environmental degradation and climate change. These changes affect soil moisture content. Apart from temperature, the other factors that affect the evaporative demand of the atmosphere include vapor-pressure deficit, wind speed and net radiation. Therefore implications of the change in all these factors on evaporative demand should be carefully analyzed. Increased temperature increases evaporation from the reservoirs and evapotranspiration from plants. Further, increased temperatures can cause warming of reservoir and rivers in the region which in turn will increase evaporation as well as will affect their thermal structure and water quality.

With changes in the various components of the hydrologic cycle, agriculture and the natural ecosystems in the river basins are affected. The growth of biological

pests and diseases increases as temperature and relative humidity levels increase with increase in precipitation. Natural ecosystems such as forests, pastures, deserts, mountain regions, lakes, streams, wetlands, coastal systems and oceans may face difficulties in adapting, and it is also possible to lose some of the flora and fauna.

With increase in population, the demand of freshwater for domestic, industrial and agricultural uses definitely increases. This situation makes it prudent to assess the sensitivity of hydrological processes to the potential future changes in climate and population to meet the requirements. Incident solar radiation, relative humidity and wind speed are other variables that are also worth analyzing owing to their significance in effecting hydrological processes.

12.6 Conclusions

The Support Vector Machine (SVM) based models are developed to downscale monthly sequences of hydrometeorological variables (precipitation, maximum and minimum temperatures) in Malaprabha river catchment (upstream of Malaprabha reservoir) of Krishna river basin, India. The large scale atmospheric variables simulated by the third generation coupled Canadian GCM for various IPCC scenarios (SRES A1B, SRES A2, SRES B1 and COMMIT) were used to prepare inputs to the SVM models.

The variables, which include both the thermodynamic and dynamic parameters, and which have a physically meaningful relationship with the precipitation, are chosen as the probable predictors for downscaling precipitation. For downscaling temperatures, large-scale atmospheric variables often used for downscaling maximum and minimum temperatures, and fluxes which control the temperature at the earth's surface are chosen as plausible predictor variables in this study.

Precipitation, maximum and minimum temperatures are projected to increase in future for A1B, A2 and B1 scenarios, whereas no trend is discerned with the COMMIT. The projected increase in predictands is high for A2 scenario and is least for B1 scenario.

The implications of climate change on monthly values of each of the hydrometeorological variables are assessed. The changes in the intensity, frequency of extreme values need to be considered. Further, the uncertainties in the projections to the choice of downscaling methods and GCMs should also be considered to draw reliable conclusions about the possible impacts of climate change in the study region, which would help policy makers for realistic assessment, management and mitigation of natural disasters, and for sustainable development. Investigating these uncertainties is a future scope of the study.

Acknowledgements

This work is partially supported by Dept of Science and Technology, Govt. of India, through AISRF project no. DST/INT/AUS/P-27/2009. Third author acknowledges support from Ministry of Earth Sciences, Govt. of India, through project no. MoES/ATMOS/PP-IX/09.

12.7 References

- Anandhi, A. (2010). "Assessing impact of climate change on season length in Karnataka for IPCC Scenarios." *Journal of Earth System Science*, 119(4), 447-460.
- Anandhi, A., Srinivas, V. V., Kumar, D. N., and Nanjundiah, R. S. (2009). "Role of Predictors in Downscaling Surface Temperature to River Basin in India for IPCC SRES Scenarios using Support Vector Machine." *International Journal of Climatology*, 29(4), 583-603.
- Anandhi, A., Srinivas, V. V., Nanjundiah, R. S., and Kumar, D. N. (2008). "Downscaling Precipitation to River Basin in India for IPCC SRES Scenarios using Support Vector Machine." *International Journal of Climatology*, 28(3), 401-420.
- ASCE Task Committee (2000). "ASCE Task Committee, Artificial neural network in hydrology." *Journal of Hydrologic Engineering*, 5(2), 124-144.
- Breiman, L., Friedman, J., Ohlsen, R., and Stone, J. (1984). *Classification and Regression Trees*. Wadsworth and Brooks, 358.
- Brinkmann, W. (1999). "Application of non-hierarchically clustered circulation components to surface weather conditions: Lake Superior basin winter temperatures." *Theoretical and Applied Climatology*, 63, 41-56.
- 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.
- Conway, D., Wilby, R., and Jones, P. (1996). "Precipitation and air flow indices over the British Isles." *Climate Research*, 7, 169-183.
- Cortes, C., and Vapnik, V. (1995). "Support vector networks." *Machine Learning*, 20, 273-297.
- Courant, R., and Hilbert, D. (1970). *Methods of Mathematical Physics*, vols. I and II. Wiley Interscience, New York, USA.
- Cristianini, N., and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, Cambridge.
- Doty, B., and Kinter, J. I., (1993). "The Grid Analysis and Display System (GrADS): a desktop tool for earth science visualization." American Geophysical Union 1993 Fall Meeting, San Fransico, CA, 6-10 December, 1993.
- Enke, W., and Spekat, A. (1997). "Downscaling climate model outputs into local and regional weather elements by classification and regression." *Climate Research*, 8, 195-207.

- Fix, E., and Hodges, J. (1951). *Discriminatory analysis: Nonparametric discrimination: Consistency properties*. In: *USAF School of Aviation Medicine* (Editor). Project 21-49-004.
- Fowler, H. J., Blenkinsop, S., and Tebaldi, C. (2007). "Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling." *International Journal of Climatology*, 27(12), 1547–1578.
- Frakes, B., and Yarnal, B. (1997). "A procedure for blending manual and correlation-based synoptic classifications." *International Journal of Climatology*, 17, 1381–1396.
- Gestel, T., Suykens, J., Baesens, B., Viaene, S., Vanthienen, J., Dedene, G., Moor, B., and Vandewalle, J. (2004). "Benchmarking least squares support vector machine classifiers." *Machine Learning*, 54(1), 5–32.
- Ghosh, S., and Mujumdar, P. P. (2008). "Statistical downscaling of GCM simulations to streamflow using relevance vector machine." *Advances in Water Resources*, 31, 132–146.
- Gosain, A., Sandhya, R., and Debajit, B. (2006). "Climate change impact assessment on hydrology of Indian river basins." in special issue on Climate change and India, *Current Science*, 90(3), 346–353.
- Govindaraju, R., and Rao, A. (Editors) (2000). *Artificial neural networks in hydrology*. Kluwer Academic Publishers, Holland, 348.
- Haupt, R., and Haupt, S. (2004). "Practical Genetic Algorithms." *John Wiley and Sons*, New Jersey, 253.
- Haykin, S. (2003). *Neural Networks: A comprehensive foundation*. Fourth Indian Reprint. *Pearson Education*, Singapore.
- Hess, P., and Brezowsky, H. (Editors) (1969). "Katalog der Grosswetterlagen Europas." *Selbstverlag des Deutschen Wetterdienstes* 113(15). Offenbach am Main, Germany, 56.
- Hughes, J., Guttorpi, P., and Charles, S. (1999). "A non-homogeneous hidden Markov model for precipitation occurrence." *Applied Statistics*, 48, 15–30.
- Hush, D. R., and Horne, B. G. (1993). "Progress in supervised neural networks: What's new since Lippmann?" *IEEE Signal Processing Magazine*, 10, 8–39.
- Huth, R., Nemesova, I., and Klimperova, N. (1993). "Weather categorization based on the average linkage clustering technique: An application to European mid-latitudes." *International Journal of Climatology*, 13, 817–835.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, 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 40-year reanalysis project." *Bulletin of the American Meteorological Society*, 77(3), 437–471.
- Karl, T., Wang, W., Schlesinger, M., Knight, R., 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.
- Keerthi, S. S., and Lin, C.J. (2003). "Asymptotic behaviors of support vector machines with Gaussian kernel." *Neural Computation*, 15(7), 1667–1689.

- Kendall, M. G. (1951). *Regression, Structure And Functional Relationship. Part I. Biometrika*, 38, 11–25.
- Kidson, J. (1994). "An automated procedure for the identification of synoptic types applied to the New Zealand region." *International Journal of Climatology*, 14, 711–721.
- Lamb, H., (1972). *British Isles weather types and a register of daily sequence of circulation patterns, 1861–1971. Geophysical Memoir*, 116. HMSO, London.
- Lin, H. T., and Lin, C. J. (2003). "A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods." Department of Computer Science and Information Engineering, *National Taiwan University*, Taiwan.
- MacQueen, J. (Editor) (1967). "Some methods for classification and analysis of multivariate observation." *Proceedings of the fifth Berkeley Symposium on mathematical statistics and probability. University of California Press*, Berkeley, 281-297.
- Mercer, J. (1909). "Functions of positive and negative type and their connection with the theory of integral equations." *Philosophical Transactions of the Royal Society*, London, A 209, 415–446.
- Nakicenovic, N., Alcamo, J., Davis, G., de Vries, B., Fenhann, J., Gaffin, S., Gregory, K., Grübler, A., Jung, T., Kram, T., La Rovere, E., Michaelis, L., Mori, S., Morita, T., Pepper, W., Pitcher, H., Price, L., Raihi, K., Roehrl, A., Rogner, H., Sankovski, A., Schlesinger, M., Shukla, P., Smith, S., Swart, R., van Rooijen, S., Victor, N., and Dadi, Z. (2000). *IPCC Special Report on Emissions Scenarios*, Cambridge, United Kingdom and New York, NY, USA.
- Pearson, K. (1896). "Mathematical contributions to the theory of evolution III regression heredity and Panmixia." *Philosophical Transactions of the Royal Society of London Series B*, 187, 253–318.
- Regonda, S. K., Rajagopalan, B., Clark, M., and Pitlick, J. (2005). "Seasonal cycle shifts in hydroclimatology over the western United States." *Journal of Climate*, 18(2), 372–384.
- Sastry, P. (2003). "An Introduction to Support Vector Machines." In *Computing and information sciences: Recent trends*. Narosa Publishing House: New Delhi.
- Schölkopf, B., Burges, C., and Smola A., (1998). *Advances in Kernel Methods: Support Vector Learning*. MIT Press, 386.
- Schubert, S., and Henderson-Sellers (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.
- Smola, A., Scholkopf, B., and Muller, K. (1998). "The connection between regularization operators and support vector kernels." *Neural Networks*, 11(4), 637–649.
- Spearman, C. (1904a). "'General intelligence' objectively determined and measured." *American Journal of Psychology*, 5, 201–293.
- Spearman, C. (1904b). "Proof and measurement of association between two things." *American Journal of Psychology*, 15, 72–101.

- Suykens, J. A. K. (2001). "Nonlinear modelling and support vector machines." IEEE Instrumentation and measurement technology conference, Budapest, Hungary, pp. 287–294.
- Tatli, H., Dalfes, H., and Mentés, S. (2005). "Surface air temperature variability over Turkey and its connection to large-scale upper air circulation via multivariate techniques." *International Journal of Climatology*, 25, 161–180.
- Tripathi, S., Srinivas, V. V., and Nanjundiah, R. S. (2006). "Downscaling of precipitation for climate change scenarios: A support vector machine approach." *Journal of Hydrology*, 330, 621–640.
- Vapnik, V. (1995). *The nature of statistical learning theory*. Springer, New York.
- Vapnik, V. (1998). *Statistical learning theory*. Wiley, New York.
- Wang, Y., Leung, L. R., McGregor, J. L., Lee, D. K., Wang, W. C., Ding, Y., and Kimura, F. (2004). "Regional climate modeling: progress, challenges, and prospects." *Journal of the Meteorological Society of Japan*, 82, 1599–1628.
- Wetterhall, F., Halldin, S., and Xu, C. (2005). "Statistical precipitation downscaling in central Sweden with the analogue method." *Journal of Hydrology*, 306, 136–174.
- Wigley, T., Jones, P., Briffa, K., and Smith, G. (1990). "Obtaining subgrid scale information from coarse-resolution general circulation model output." *Journal of Geophysical Research-Atmospheres*, 95, 1943–1953.
- Wilby, R. L., Charles, S. P., Zorita, E., Timbal, B., Whetton, P., and Mearns, L.O. (2004). *The guidelines for use of climate scenarios developed from statistical downscaling methods*. Supporting material of the Intergovernmental Panel on Climate Change (IPCC), prepared on behalf of Task Group on Data and Scenario Support for Impacts and Climate Analysis (TGICA). <http://ipccddc.cru.uea.ac.uk/guidelines/StatDown_Guide.pdf> (accessed April 2011).
- Wilby, R. L., Dawson, C. W., and Barrow, E. M. (2002). "SDSM—a decision support tool for the assessment of regional climate change impacts." *Environmental Modelling & Software*, 17(2), 147–159.
- Wilks, D. S. (1999). "Multi-site downscaling of daily precipitation with a stochastic weather generator." *Climate Research*, 11, 125–136.
- Winkler, J., Palutikof, J., Andresen, J., and Goodess, C. (1997). "The simulation of daily temperature time series from GCM output Part II: Sensitivity analysis of an empirical transfer function methodology." *Journal of Climate*, 10(10), 2514–2532.