Journal of Hydrology (2006) 330, 621–640

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jhydrol

Downscaling of precipitation for climate change scenarios: A support vector machine approach

Shivam Tripathi [a], V.V. Srinivas [a,*], Ravi S. Nanjundiah [b]

[a] *Department of Civil Engineering, Indian Institute of Science, Bangalore 560 012, Karnataka, India*
[b] *Centre for Atmospheric and Oceanic Sciences, Indian Institute of Science, Bangalore 560 012, Karnataka, India*

Received 27 April 2005; received in revised form 22 March 2006; accepted 14 April 2006

KEYWORDS
Precipitation;
Downscaling;
Climate change;
General circulation model (GCM);
Support vector machine;
Neural network;
Hydroclimatology;
India

Summary The Climate impact studies in hydrology often rely on climate change information at fine spatial resolution. However, general circulation models (GCMs), which are among the most advanced tools for estimating future climate change scenarios, operate on a coarse scale. Therefore the output from a GCM has to be downscaled to obtain the information relevant to hydrologic studies. In this paper, a support vector machine (SVM) approach is proposed for statistical downscaling of precipitation at monthly time scale. The effectiveness of this approach is illustrated through its application to meteorological sub-divisions (MSDs) in India. First, climate variables affecting spatio-temporal variation of precipitation at each MSD in India are identified. Following this, the data pertaining to the identified climate variables (predictors) at each MSD are classified using cluster analysis to form two groups, representing wet and dry seasons. For each MSD, SVM- based downscaling model (DM) is developed for season(s) with significant rainfall using principal components extracted from the predictors as input and the contemporaneous precipitation observed at the MSD as an output. The proposed DM is shown to be superior to conventional downscaling using multi-layer back-propagation artificial neural networks. Subsequently, the SVM-based DM is applied to future climate predictions from the second generation Coupled Global Climate Model (CGCM2) to obtain future projections of precipitation for the MSDs. The results are then analyzed to assess the impact of climate change on precipitation over India. It is shown that SVMs provide a promising alternative to conventional artificial neural networks for statistical downscaling, and are suitable for conducting climate impact studies.

Introduction

Recently, there is growth in scientific evidence that global climate has changed, is changing and will continue to

* Corresponding author. Tel.: +91 80 2293 2641; fax: +91 80 2360 0404.
E-mail addresses: vvs@civil.iisc.ernet.in, vvsrinivas@yahoo.com (V.V. Srinivas).

0022-1694/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2006.04.030

change (NRC, 1998). In this scenario, there is a need to improve our understanding of the global climate system to assess the possible impact of a climate change on hydrological processes.

General Circulation Models (GCMs), which describe the atmospheric process by mathematical equations, are the most adapted tools for studying the impact of climate change at regional scale. These climate models have been evolving steadily over the past several decades. Recently fully coupled Atmosphere–Ocean GCMs (AOGCMs), along with transient methods of forcing the concentration of greenhouse gases, have brought considerable improvement in the climate model results.

The resolution of the present state-of-the-art GCMs is coarser than 2° for both latitude and longitude, which is of the order of a few hundred kilometers between grid-points. In other words, GCM provides output at nodes of grid boxes, which are tens of thousands of square kilometers in size, whereas the scale of interest to hydrologists is of the order of a few hundred square kilometers. In the past decade, to deal with this problem of mismatch of spatial scales several downscaling methodologies have been developed.

More recently, downscaling has found wide application in hydroclimatology for scenario construction and simulation/prediction of (i) regional precipitation (Kim et al., 2004); (ii) low-frequency rainfall events (Wilby, 1998) (iii) mean, minimum and maximum air temperature (Kettle and Thompson, 2004); (iv) soil moisture (Georgakakos and Smith, 2001; Jasper et al., 2004); (v) runoff (Arnell et al., 2003) and streamflows (Cannon and Whitfield, 2002); (vi) ground water levels (Bouraoui et al., 1999); (vii) transpiration (Misson et al., 2002), wind speed (Faucher et al., 1999) and potential evaporation rates (Weisse and Oestreicher, 2001); (viii) soil erosion and crop yield (Zhang et al., 2004); (ix) landslide occurrence (Buma and Dehn, 2000; Schmidt and Glade, 2003) and (x) water quality (Hassan et al., 1998).

The approaches, which have been proposed for downscaling GCMs could be broadly classified into two categories: dynamic downscaling and statistical downscaling. In the dynamic downscaling approach a Regional Climate Model (RCM) is embedded into GCM. The RCM is essentially a numerical model in which GCMs are used to fix boundary conditions. The major drawback of RCM, which restricts its use in climate impact studies, is its complicated design and high computational cost. Moreover, RCM is inflexible in the sense that expanding the region or moving to a slightly different region requires redoing the entire experiment (Crane and Hewitson, 1998).

The second approach to downscaling, termed statistical downscaling, involves deriving empirical relationships that transform large-scale features of the GCM (Predictors) to regional-scale variables (Predictands) such as precipitation, temperature and streamflow. There are three implicit assumptions involved in statistical downscaling (Hewitson and Crane, 1996). Firstly, the predictors are variables of relevance and are realistically modeled by the host GCM. Secondly, the empirical relationship is valid also under altered climatic conditions. Thirdly, the predictors employed fully represent the climate change signal.

A diverse range of statistical downscaling methods has been developed in recent past. Among them Artificial Neural Network (ANN) based downscaling techniques have gained wide recognition owing to their ability to capture non-linear relationships between predictors and predictand (e.g., Cavazos, 1997; Crane and Hewitson, 1998; Wilby et al., 1998; Trigo and Palutikof, 1999; Sailor et al., 2000; Snell et al., 2000; Mpelasoka et al., 2001; Schoof and Pryor, 2001; Cannon and Whitfield, 2002; Crane et al., 2002; Olsson et al., 2004; Shivam, 2004; Solecki and Oliveri, 2004; Tatli et al., 2004). The concept of ANNs came into being approximately 60 years ago (McCulloch and Pitts, 1943) inspired by a desire to understand the human brain and emulate its functioning. 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, non-linear modeling and classification. It has been extensively used in a variety of physical science applications, including hydrology (Govindaraju and Rao, 2000; ASCE Task Committee on Artificial Neural Networks in Hydrology, 2000b).

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). Recently, Vapnik (1995, 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. Readers are referred to Vapnik (1995, 1998), Cortes and Vapnik (1995), Schölkopf et al. (1998), Cristianini and Shawe-Taylor (2000), Haykin (2003) and Sastry (2003) for introductory material on SVM.

The research presented in this paper is motivated by a desire to explore the potential of the SVM in downscaling future climate projections provided by GCMs. The SVM would be ideally suited for the downscaling task owing to its ability to provide good generalization performance in capturing non-linear regression relationships between predictors and predictand, despite the fact that it does not incorporate problem domain knowledge.

The following objectives have been set for this paper. Firstly, to investigate the potential of SVM in downscaling GCM simulations by comparing its performance with multilayer back-propagation neural network based downscaling model, and secondly to assess the impact of climate change on hydrological inputs to meteorological sub-divisions in India using simulations from the second generation Coupled Global Climate Model (CGCM2) for IS92a scenario (IPCC, 1992).

The study region has been selected because gaining an understanding of plausible effects of climate change on water resources is of great significance in Indian context owing to its agro-based economy. With the inherent scarcity of water in several parts of India and projected changes in climate for the coming decades, acute water shortages or critical droughts are imminent on Indian sub-continent. The evidence of climatic link with hydrology of Indian sub-continent (IPCC, 2001) necessitates development of effective strategies for regional hydrologic analysis to cope with critical water shortages in future. To progress towards this goal, it is necessary to develop efficient downscaling strategy to interpret climate change signals at regional scale.

While a plethora of statistical downscaling techniques have been used in different parts of the globe (Wilby and Wigley, 1997; Xu, 1999), there is a paucity of such studies over Indian sub-continent.

The remainder of this paper is structured as follows: First, the fundamental principle of SVM and its formulation are presented along with a brief description of multi-layer back-propagation neural network. Following this, details of the study region are provided and the methodology proposed for downscaling of precipitation is presented. Finally, a set of conclusions is drawn following discussion on results obtained from the downscaling models.

Support vector machine

In the past few decades, traditional neural networks such as multi-layer back-propagation neural network and radial basis function networks have been extensively used in a wide range of engineering applications including hydrology (Govindaraju and Rao, 2000; Maier and Dandy, 2000; Poulton, 2002; Meireles et al., 2003). More recently, many new training algorithms have been proposed to overcome the drawbacks of traditional neural networks and to increase their reliability (Bianchini and Gori, 1996; Neocleous and Schizas, 2002). In this paradigm, one of the significant developments is a class of kernel based neural networks called Support Vector Machines (SVMs), the principle of which is rooted in the statistical learning theory and method of structural risk minimization (Haykin, 2003). Support Vector Machines (SVMs) have found wide application in several areas including pattern recognition, regression, multimedia, bio-informatics and artificial intelligence. Very recently SVMs are gaining recognition in hydrology (Dibike et al., 2001; Asefa et al., 2004; Khadam and Kaluarachchi, 2004). Dibike et al. (2001) applied SVM for classification of digital remote sensing image data, and for rainfall-runoff modeling. Asefa et al. (2004) used SVM to design groundwater head monitoring networks in order to reduce spatial redundancy, while Khadam and Kaluarachchi (2004) applied SVM to reconstruct short streamflow record with significant gaps.

In this section the theoretical background of SVM is provided following a brief mention of the motivation for its use. Following this, the standard formulation of SVM for regression is presented. Subsequently the Least Square Support Vector Machine (LS-SVM), which has been used in this study, is described.

Motivation for the use of support vector machine (SVM)

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 machine (i.e., the ability of 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 for 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 feasible choice for solving a variety of problems in hydrology, which are non-linear in nature.

Theoretical background of support vector machine

Turing (1950) proposed the idea of learning machines in 1950. An important feature of a learning machine is that its teacher will often be very largely ignorant of quite what is going on inside, although he may still be able to some extent predict his pupil's behavior (Turing, 1950). Vapnik (1995) argued that to control the generalization ability of a learning machine one has to control two different factors: the error-rate on the training data and the capacity of the learning machine as measured by its Vapnik–Chervonenkis (VC) dimension, named in honor of its originators, Vapnik and Chervonenkis (1971). The VC dimension, which is a non-negative integer, measures the expressive power of the family of classification functions realized by the learning machine (Haykin, 2003, p. 94).

Consider a finite training sample of N patterns $\{(\mathbf{x}_i, y_i),\ i = 1, \ldots, N\}$, where $\mathbf{x}_i$ denotes the ''ith'' pattern in n-dimensional space (i.e., $\mathbf{x}_i = [x_{1i}, \ldots, x_{ni}] \in \Re^n$) and y_i ($y_i \in \{-1, +1\}$) represents the class label of that pattern. 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 $\mathbf{w}$ ($\mathbf{w} \in \Re^n$), always gives the same output $f(\mathbf{x}; \mathbf{w})$. Training phase of the learning machine involves adjusting the parameters $\mathbf{w}$.

Let the training error for the trained machine be denoted by $v(\mathbf{w})$. Then, according to the principle of structural risk minimization (Vapnik, 1992, 1998) there exists a bound for the probability of classification errors on the test set, $P(\mathbf{w})$, expressed as:

$$P(\mathbf{w}) \leqslant v(\mathbf{w}) + \varepsilon_1(N, h, \eta, v) \tag{1}$$

$$\text{In Eq. (1)},\ v(\mathbf{w}) = \frac{1}{2N}\sum_{i=1}^{N} |y_i - f(\mathbf{x}_i, \mathbf{w})| \tag{2}$$

$$\text{and } \varepsilon_1(N, h, \eta, v) = 2\varepsilon_0^2(N, h, \eta)\left(1 + \sqrt{1 + \frac{v(\mathbf{w})}{\varepsilon_0^2(N, h, \eta)}}\right) \tag{3}$$

$$\text{In Eq. (3)},\ \varepsilon_0(N, h, \eta) = \sqrt{\frac{h}{N}\left[\ln\left(\frac{2N}{h}\right) + 1\right] - \frac{1}{N}\ln\left(\frac{\eta}{4}\right)} \tag{4}$$

The inequality in Eq. (1) is valid with a probability of $(1 - \eta)$. The $\varepsilon_0(N, h, \eta)$, which represents the capacity of the learning machine, is called VC confidence interval. The value of this confidence interval depends on the number of training patterns N, the VC dimension of the learning

machine h and the value of η. For a small training error $v(\mathbf{w})$ close to zero Eq. (1) reduces to Eq. (5), whereas for a large training error $v(\mathbf{w})$ close to unity Eq. (1) reduces to Eq. (6) (Haykin, 2003, p. 100).

$$P(\mathbf{w}) \lesssim v(\mathbf{w}) + 4\varepsilon_0^2(N, h, \eta) \tag{5}$$

$$P(\mathbf{w}) \lesssim v(\mathbf{w}) + \varepsilon_0(N, h, \eta) \tag{6}$$

Most of the traditional neural network models, which implement empirical risk minimization principle, seek to minimize only training error, $v(\mathbf{w})$. However, this does not result in good generalization performance because drop in $v(\mathbf{w})$ alone does not guarantee reduction in test error, as evident from Eqs. (5) and (6). On the other hand, SVM seeks to minimize an upper bound on the generalization error by striking a right balance between the training error $v(\mathbf{w})$ and the VC confidence interval $\varepsilon_0(N, h, \eta)$ by using the principle of structural risk minimization (Haykin, 2003, pp. 98–100).

Standard support vector machine for regression

Herein the basic ideas of the SVM for the case of function approximation are reviewed. Consider the finite training sample pattern $(\mathbf{x}_i, y_i)$, where $\mathbf{x}_i \in \Re^n$ is a sample value of the input vector $\mathbf{x}$ consisting of N training patterns (i.e., $\mathbf{x} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$) and $y_i \in \Re$ is the corresponding value of the desired model output. A non-linear transformation function $\phi(\cdot)$ is defined to map the input space to a higher-dimension feature space, $\Re^{n_h}$ (Fig. 1). According to Cover's theorem (Cover, 1965) a linear function, $f(\cdot)$, could be formulated in the high dimensional feature space to look for a non-linear relation between inputs and outputs in the original input space, as shown below.

$$\hat{y} = f(\mathbf{x}) = \mathbf{w}^T\phi(\mathbf{x}) + b \tag{7}$$

In Eq. (7) $\hat{y}$ denotes the actual model output. The coefficients $\mathbf{w}$ and b are the adjustable model parameters ($\mathbf{w} \in \Re^{n_h}$ and $b \in \Re$). In the SVM one aims at minimizing the empirical risk, R_{emp}, defined as:

$$R_{\text{emp}} = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|_\varepsilon \tag{8}$$

where $|y_i - \hat{y}_i|_\varepsilon$ is the Vapnik's ε-insensitive loss function (shown as thick line in Fig. 2(a)) defined as

$$|y_i - \hat{y}_i|_\varepsilon = \begin{cases} 0 & \text{if } |y_i - \hat{y}_i| \leqslant \varepsilon \\ |y_i - \hat{y}_i| - \varepsilon & \text{otherwise} \end{cases} \tag{9}$$

Following regularization theory (Haykin, 2003), the parameters $\mathbf{w}$ and b are estimated by minimizing the cost function $\psi_\varepsilon(\mathbf{w}, \xi, \xi^*)$.

$$\psi_\varepsilon(\mathbf{w}, \xi, \xi^*) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{N}(\xi_i + \xi_i^*)$$

subjected to the constraints

$$\begin{aligned} y_i - \hat{y}_i &\leqslant \varepsilon + \xi_i & i &= 1, 2, \ldots, N \\ -y_i + \hat{y}_i &\leqslant \varepsilon + \xi_i^* & i &= 1, 2, \ldots, N \\ \xi_i &\geqslant 0 & i &= 1, 2, \ldots, N \\ \xi_i^* &\geqslant 0 & i &= 1, 2, \ldots, N \end{aligned} \tag{10}$$

where ξ_i and ξ_i^* are positive slack variables and C is a positive real constant. The first term of the cost function, which represents weight decay (or model complexity-penalty function), is used to regularize weight sizes and to penalize large weights. This helps in improving generalization performance (Hush and Horne, 1993). The second term of the cost function, which represents penalty function, penalizes deviations of $\hat{y}$ from y larger than $\pm\varepsilon$ using Vapnik's ε-insensitive loss function. The constant C determines the amount up to which deviations from ε are tolerated. Deviations above ε are denoted by ξ_i, whereas deviations below $-\varepsilon$ are denoted by ξ_i^*.

The constrained quadratic optimization problem given in Eq. (10) can be solved using the method of Lagrangian multipliers (Haykin, 2003, p. 323), from which one can obtain $\mathbf{w}$ as:

$$\mathbf{w} = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)\phi(\mathbf{x}_i) \tag{11}$$

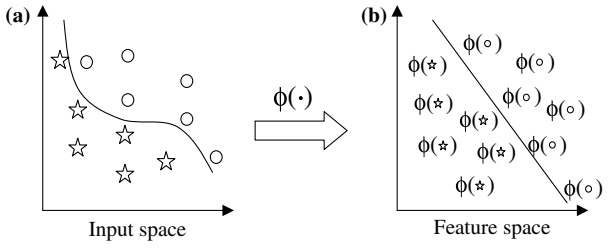

Figure 1 A nonlinear transformation function $\phi(\cdot)$ defined to convert a non-linear problem in the original lower dimensional input space (a) to linear problem in a higher dimensional feature space (b). The stars and circles shown in Fig. 1(a) denote data points.

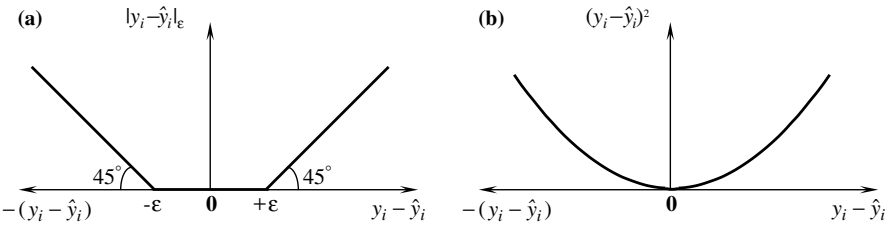

Figure 2 Loss functions. (a) denotes Vapnik's ε-insensitive loss function used by standard support vector machine (SVM), while (b) is quadratic loss function used by least square SVM.

where α_i and α_i^* are the Lagrange multipliers, which are positive real constants. The data points corresponding to non-zero values for $(\alpha_i - \alpha_i^*)$ are called support vectors. Finally, Eq. (7) takes the form of Eq. (12), which represents the SVM for non-linear function estimation. The architecture of SVM is shown in Fig. 3.

$$\hat{y} = f(\mathbf{x}) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}) + b \tag{12}$$

where $K(\mathbf{x}_i, \mathbf{x})$ is the inner product kernel function defined in accordance with Mercer's theorem (Mercer, 1909 and Courant and Hilbert, 1970) and b is the bias.

$$K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j) \tag{13}$$

There are several possibilities for the choice of kernel function, 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 used to map the input data into higher dimensional feature space, which is given by:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \tag{14}$$

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 non-linearly maps the training data into a possibly infinite-dimensional space, thus it can effectively handle the situations when the relationship between predictors and predictand is non-linear. Moreover, the RBF is computationally simple than polynomial kernel, which has more parameters.

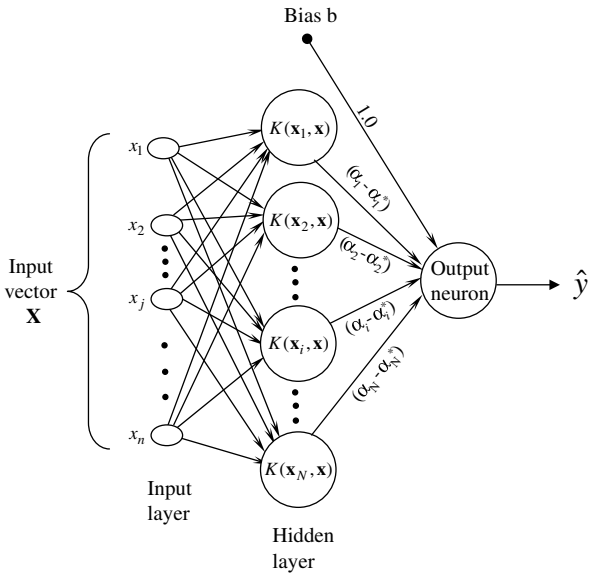

Figure 3 Architecture of support vector machine.

Least square support vector machine

The Least Square Support Vector Machine (LS-SVM), which has been used in this study, provides a computational advantage over standard SVM by converting quadratic optimization problem into a system of linear equations (Suykens, 2001). The LS-SVM optimization problem for function estimation is formulated by minimizing the cost function $\psi_L(\mathbf{w}, e)$.

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

subjected to the equality constraint

$$y_i - \hat{y}_i = e_i \quad i = 1, \ldots, N \tag{15}$$

Important differences with standard SVMs are the equality constraints and the quadratic loss term e_i^2, which greatly simplifies the problem. The quadratic loss function is shown as thick line in Fig. 2(b). The solution of the optimization problem is obtained by considering the Lagrangian as

$$L(\mathbf{w}, b, \mathbf{e}, \boldsymbol{\alpha}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}C\sum_{i=1}^{N} e_i^2 - \sum_{i=1}^{N} \alpha_i\{\hat{y}_i + e_i - y_i\} \tag{16}$$

where α_i are Lagrange multipliers. The conditions for optimality are given by

$$\begin{cases} \frac{\partial L}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^{N} \alpha_i \phi(\mathbf{x}_i) = 0 \\ \frac{\partial L}{\partial b} = \sum_{i=1}^{N} \alpha_i = 0 \\ \frac{\partial L}{\partial e_i} = \alpha_i - Ce_i = 0 \qquad i = 1, \ldots, N \\ \frac{\partial L}{\partial \alpha_i} = \hat{y}_i + e_i - y_i = 0 \qquad i = 1, \ldots, N \end{cases} \tag{17}$$

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

$$\begin{bmatrix} 0 & \vec{1}^T \\ \vec{1} & \boldsymbol{\Omega} + C^{-1}\mathbf{I} \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \text{ where } \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix};\ \vec{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{N\times 1};$$

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix};\ \mathbf{I} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}_{N\times N} \tag{18}$$

In Eq. (18), $\boldsymbol{\Omega}$ is obtained from the application of Mercer's theorem.

$$\Omega_{i,j} = K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j) \quad \forall i, j \tag{19}$$

The resulting LS-SVM model for function estimation is:

$$f(\mathbf{x}) = \sum \alpha_i^* K(\mathbf{x}_i, \mathbf{x}) + b^* \tag{20}$$

where α_i^* and b^* are the solutions to Eq. (18). It is worth mentioning that developing LS-SVM with RBF kernel involves selection of RBF kernel width σ and parameter C, the details of which are provided elsewhere in the paper.

Multi-layer back-propagation neural network

The Multi-layer back-propagation neural network is chosen as the base ANN model to compare the performance of SVM. Readers are referred to Bishop (1995) and Haykin (2003) for introductory material and statistical interpretation of the ANN model. The ANN is composed of a number of simple, highly interconnected processing elements commonly referred to as neurons, nodes or units. The nodes are organized into several layers: the input layer, the output layer and one or more hidden layers in between. At each node in a layer the information is received, stored, processed, and communicated further to nodes in the next layer. Developing the ANN model involves training, testing and validation phases. The training consists of two phases: A forward pass, during which the processing of information occurs from the input layer to the output layer; and a backward pass, when the error from the output layer is propagated back to the input layer and the weights associated with interconnections are modified. A detailed description of the algorithm, which has been implemented in this paper, is found in ASCE Task Committee on Artificial Neural Networks in Hydrology (2000a, pp. 121—122). Methods of designing and training the ANN model can be found in several books addressing the topic. Hence effort is not devoted to provide minute details of implementation of the same in this paper.

Study region

The study region India, which is located between 8°4′ N and 37°6′ N latitudes and 68°7′ E and 97°25′ E longitudes, receives average annual rainfall of 120 cm. It has a tropical monsoon climate where most of the precipitation is confined to a few months of the monsoon season. The south-west (summer) monsoon has warm winds blowing from Indian Ocean over almost entire country and causing copious amount of rainfall during June—September months. About 75% of annual rainfall in India is due to the south-west monsoon, however the spatial distribution of monsoon rainfall shows significant variation across the country. For the remaining part of the year, except for coastal-strip in the south-eastern peninsular India comprising coastal Andhra Pradesh and Tamil Nadu meteorological subdivisions (MSDs), the rest of country receives practically no rainfall. The precipitation in India varies from about 10 cm in western Rajasthan (located in north-western part of the country) to over 900 cm in Meghalaya located in the north-eastern part (Sharma et al., 2003). The distribution of annual rainfall over MSDs in India is shown in Fig. 4.

Anthropogenic causes are known to have increased carbon-dioxide and other green house gases in the atmosphere. As a consequence the global mean temperature has shown a rising trend. For India, Rupa Kumar et al. (1994) have shown that the countrywide mean maximum temperature has risen by 0.6 °C in a study covering the period 1901—1987. For the region, the projected area-averaged annual mean warming is 1.6 ± 0.2 °C in the 2020s, 3.1 ± 0.3 °C in the 2050s, and 4.6 ± 0.4 °C in the 2080s as a result of increase in the atmospheric concentration of green house gases (IPCC, 2001).

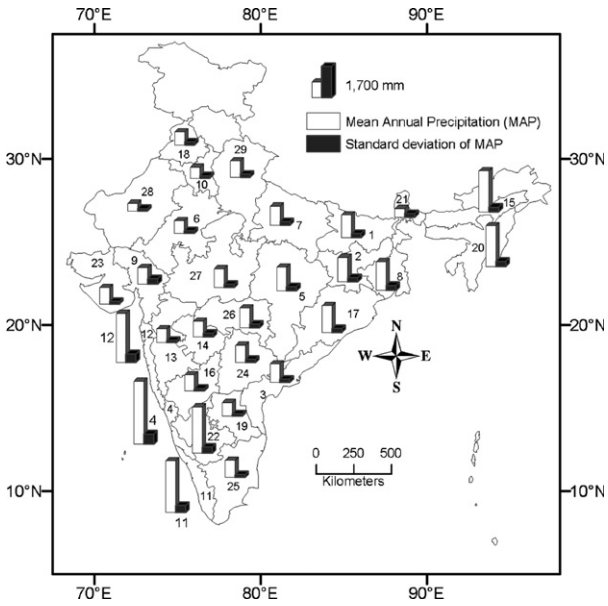

Figure 4 Distribution of rainfall over meteorological subdivisions of India. The number shown against each subdivision denotes its serial number in Table 1.

Parthasarathy et al. (1993) found no systematic trend in the all India rainfall in a study covering the period 1871—1990. However, they reported noting large interannual and decadal variations. On the other hand, the results of Rupa Kumar et al. (1992) showed long-term changes in the Indian monsoon rainfall on regional and local scales. Increasing trend in the monsoon seasonal rainfall is reported along the west coast, north Andhra Pradesh and north-west India, while decreasing trend in the same is reported over east Madhya Pradesh and adjoining areas, north-east India and parts of Gujarat and Kerala.

Inter-Governmental Panel on Climate Change (IPCC, 2001) evaluation of the results of a number of AOGCMs, indicates that the mean monsoon precipitation over India will intensify with increase in carbon-dioxide content. Lal et al. (1995) predicts a decrease in South Asian monsoon rainfall owing to radiative cooling induced by sulphate aerosols. It is worth mentioning that to the best of our knowledge statistical downscaling models have not been used to assess the impact of climate change on regional precipitation of India.

Methodology

Development of ANN/SVM based downscaling model involves identification and screening of climate variables affecting spatio-temporal variation of precipitation at various MSDs in India. The data pertaining to screened climate variables and precipitation for each MSD is used to develop (train and validate) a downscaling model. The validated downscaling model is subsequently applied to projected climate information from GCM to obtain future scenarios of precipitation for the sub-division. This section outlines the procedure involved in collection and processing of data,

model development and application of the model to the output from the second generation Coupled General Circulation Model (CGCM2) of the Canadian Center for Climate Modelling and Analysis (CCCma).

Data extraction

For the study region, grid point climate data at monthly time scale, prepared by National Center for Environmental Prediction (NCEP) (Kalnay et al., 1996), is extracted from the web site http://www.cdc.noaa.gov/, maintained by National Oceanic and Atmospheric Administration and Cooperative Institute for Research in Environmental Sciences Climate Diagnostics Center, Boulder, CO, USA. The data, spanning the period from January 1948 to December 2002 is extracted for 182 grid points whose latitude ranges from 5° N to 37.5° N and longitude ranges from 67.5° E to 97.5° E at a spatial resolution of 2.5° covering entire India. The variables extracted from the NCEP reanalysis dataset include air temperature, relative humidity, specific humidity, geo-potential height, zonal, vertical and meridional wind velocities at various pressure levels and sea level pressure.

The 55-year long dataset may be small to delineate inter-decadal oscillation of Indian summer monsoon (Kailas and Narasimha, 2000). However, from Parthasarathy et al. (1993) it is seen that all India summer monsoon rainfall is largely stationary for the past 120-years and there appears to be no long-term trends. It must also be borne in mind that datasets which are internally consistent with longer periods are not available.

The monthly area weighted rainfall data of 29 MSDs in India (Parthasarathy et al., 1994), which extends from January 1948 to December 2002, is extracted from Indian Institute of Tropical Meteorology, Pune, web site http://www.tropmet.res.in. Primary source of the data is India Meteorological Department. The list of the MSDs is provided in Table 1.

Simulated monthly climate data of the second generation Coupled General Circulation Model (CGCM2) for the scenario IS92a, which extends from January 1948 to December 2100, is extracted from Canadian Center for Climate Modelling and Analysis (CCCma) web site http://www.cccma.bc.ec.gc.ca/. The extracted data pertains to 99 grid points whose latitude ranges from 1.86° N to 38.97° N and longitude ranges from 67.5° E to 97.5° E. The CGCM2 grid is uniform along the lon-

Table 1 List of meteorological subdivisions in India and the parameters of the SVM downscaling model (σ and C) developed for the same

Subdivision number	Subdivision Name	Parameters of SVM model (σ, C)	
		Wet	Dry
1	Bihar Plains	496.47, 42.94	—
2	Bihar Plateau	732.55, 19.58	—
3	Coastal Andhra Pradesh	210.00, 50.86	560.78, 391.76
4	Coastal Karnataka	472.94, 179.41	—
5	East Madhya Pradesh	594.12, 134.24	—
6	East Rajasthan	594.12, 134.24	—
7	East Uttar Pradesh	930.98,145.88	—
8	Gangetic West Bengal	814.90, 231.29	—
9	Gujarat	582.75, 246.82	—
10	Haryana	275.29, 72.12	—
11	Kerala	551.37, 91.53	—
12	Konkan and Goa	896.47, 538.00	—
13	Madhya Maharastra	417.65, 91.53	—
14	Marathwada	901.00, 313.00	—
15	North Assam	850.00, 365.00	—
16	North Interior Karnataka	796.80, 151.00	—
17	Orissa	872.00, 168.00	—
18	Punjab	701.18, 161.80	—
19	Rayalaseema	760.05, 134.50	—
20	South Assam	839.92, 874.28	—
21	Sub-Himalayan West Bengal	380.00, 110.00	—
22	South Interior Karnataka	578.50, 325.64	—
23	Saurashtra and Kutch	833.00, 204.00	—
24	Telangana	415.87, 308.00	—
25	Tamil Nadu	674.00, 142.00	680.00, 40.00
26	Vidarbha	771.43, 327.14	—
27	West Madhya Pradesh	678.00, 418.00	—
28	West Rajasthan	625.00, 18.00	—
29	West Uttar Pradesh	987.30, 115.90	—

The geographic location of the subdivisions in India is shown in Fig. 4.

gitude with grid box size of 3.75° and nearly uniform along the latitude (approximately 3.75°). The variables provided by CGCM2 include air temperature, specific humidity, geopotential height, zonal and meridional wind velocities at various pressure levels and sea level pressure.

The CGCM2 data is interpolated to the same $2.5° \times 2.5°$ grid as used by NCEP data. The interpolation is performed with a linear inverse square interpolation procedure using spherical distances (Willmott et al., 1985). The utility of this interpolation algorithm was examined in previous downscaling studies (Hewitson and Crane, 1992; Shannon and Hewitson, 1996; Crane and Hewitson, 1998).

Seasonal stratification

Seasonal variation of the atmospheric circulation and precipitation is the distinguishing feature of the monsoon regions of the world, including India. Often the date of onset of monsoon and its duration has been observed to vary from one-year to another in the study region (Chhabra et al., 1999). Elsewhere, recent studies have shown that conventional climatological seasons, such as June–September or December–March, may not reflect 'natural' seasons contained in the climate data and so alternative delimitations may be required (Winkler et al., 1997). Furthermore, rigid classification of data using seasonal definitions based on present climate behavior may not be valid under altered climate conditions. Hence the seasonal stratification of data prior to calibration of downscaling model is necessary. In this study clustering algorithm is used to identify dry and wet seasons in a calendar year for NCEP and GCM data sets. Clustering is a process by which a set of feature vectors is divided into clusters or groups such that the feature vectors within a cluster are as similar as possible and the feature vectors of different clusters are as dissimilar as possible.

Clustering analysis is widely recognized for recovering natural groups present in the data. Out of various clustering algorithms reported in the literature, *K*-means clustering algorithm, which is widely recognized for its simplicity and computational efficiency (MacQueen, 1967), has been used in this study. The *K*-means algorithm is an iterative procedure which produces clusters by implicitly minimizing a square error criterion.

From the *N*-months long NCEP reanalysis records the climate variables, which represent atmospheric circulation over the study region and are also realistically simulated by GCM, are selected. Following Wilby and Wigley (2000), the spatial domain of each climate variable is chosen as 36-grid points surrounding each MSD. From the above selected variables the principal components, which preserved more than 95% of the variance, are extracted to form feature vectors for seasonal stratification.

Suppose that the given set of *N*-feature vectors (in *n*-dimensional space) have been partitioned into *K* clusters $\{C_1, C_2, \ldots, C_K\}$ such that cluster C_k has N_k feature vectors and each feature vector is in exactly one cluster, so that $\sum_{k=1}^{K} N_k = N$.

The centroid of cluster C_k is defined as:

$$\mathbf{z}^{(k)} = \left(\frac{1}{N_k}\right) \sum_{i=1}^{N_k} \mathbf{x}_i^{(k)} \quad k = 1, \ldots, K. \tag{21}$$

In Eq. (21), $\mathbf{x}_i^{(k)}$ is the *i*th feature vector belonging to cluster C_k. The square-error for the cluster C_k is the sum of the squared Euclidean distances between each feature vector in C_k and its centroid $\mathbf{z}^{(k)}$, which is given by Eq. (22)

$$e_k^2 = \sum_{i=1}^{N_k} (\mathbf{x}_i^{(k)} - \mathbf{z}^{(k)})^T (\mathbf{x}_i^{(k)} - \mathbf{z}^{(k)}) \tag{22}$$

The square-error criterion function for the *K*–means algorithm is given by Eq. (23):

$$E_K^2 = \sum_{k=1}^{K} e_k^2 \tag{23}$$

The objective of *K*-means algorithm is to find a partition containing *K* clusters that minimizes E_K^2 for chosen *K*. The *K*-means algorithm starts with a given initial partition and keeps reassigning feature vectors to clusters based on similarity between the feature vectors and cluster centers until a convergence criterion is met. One criterion for convergence is to stop the iterations when the change in the value of E_K^2 between two successive iterations becomes sufficiently small. An alternative is to terminate the algorithm when there is no reassignment of any feature vector from one cluster to another during successive iterations.

A major problem with the *K*-means algorithm is that the final partition provided by it is sensitive to the selection of the initial partition and may result in the convergence to local minima of the criterion function (Eq. (23)). As no single procedure for selecting the initial partition has been theoretically proven to yield a global minimum value for the criterion function, several methods of initialization are in use. Wiltshire (1986) randomly partitioned data to initiate the clustering algorithm. Bhaskar and O'Connor (1989) considered initial cluster seeds as feature vectors that are separated by at least a specified minimum distance (Bhaskar and O'Connor, 1989, p. 795). Burn (1989) suggested choosing *K* of the *N* feature vectors as the starting centroids to ensure that each cluster has at least one member (Burn, 1989, p. 569). In the present study, all the above methods of initializing *K*-means algorithm have been tried. It is found that initializing *K*-means algorithm by randomly partitioning the data results in minimum value of the criterion function. Nevertheless, it is worth mentioning here that a number of random initializations (herein 20) have to be tried before arriving at the optimal result.

Each feature vector (representing a month) of the NCEP data is assigned a label that denotes the cluster (season) to which it belongs. Following this, GCM simulations (past and future) are labeled using nearest neighbor rule. Fix and Hodges (1951) introduced the nearest neighbor rule whose applicability was studied by Cover and Hart (1967). Following this rule each feature vector of the GCM data is assigned the label of its nearest neighbor among the *N* feature vectors of the NCEP data. Herein, to determine the neighbors, the distance is computed between NCEP and GCM feature vectors using Euclidean measure.

Selecting predictors

In literature, several attempts have been made to develop a plethora of empirical downscaling techniques using a wide

variety of predictors. The choice of predictors could vary from one region to another. Since there are no general guidelines for selection of predictors in different parts of the world, a comprehensive search of predictors is necessary. Following the suggestions of Wilby and Wigley (2000) the spatial domain of each predictor is chosen as 36 grid points surrounding each MSD (example, see Fig. 5). At each grid point the value of a predictor varies with change in pressure level. In general, the values of the climate variables at earth's surface (which corresponds to approximately 1000 mb), 850 mb, 500 mb and 200 mb pressure levels are found to be representative of circulation pattern in the study region (example, Maini et al., 2004). The screening of predictors is done for each MSD to identify the climate variables (in terms of grid locations and pressure levels), which have high correlation with precipitation for each season. For this analysis, the threshold value of cross-correlation between predictors and precipitation is sensibly chosen in the range 0.4—0.8 for each subdivision and for each season. It is worth mentioning here that the selection of threshold value is purely subjective. The value of a threshold is varied to ensure that a reasonable number of predictors are filtered for each subdivision for further analysis. This is because just one or two predictor variables do not reflect circulation patterns in a subdivision. The analysis is also performed for dry season for those subdivisions where precipitation is considerable during both seasons of the water year. Herein, it is to be mentioned that tens (or hundreds) of variables screened for each MSD are not listed in the paper due to lack of space.

Effective alternatives to linear cross-correlation measure for selecting predictors from a pool of potential predictors include CART (Classification and Regression Trees, Faucher et al., 1999) and graphical sensitivity analysis (Cannon and McKendry, 2002). They could provide a better predictor set for downscaling, particularly when non-linear relationship between predictors and the predictand is more pronounced.

Different climate variables have different characteristic spatial scales. Grid-scale processes are realistically simulated by GCMs at coarse spatial resolution, while the sub-grid-scale processes such as precipitation are taken into account in GCMs by means of parameterizations, that is, by semi-empirical methods that are tuned to reproduce the net effect of the considered process on the global-scale. Therefore, precipitation provided by GCM is usually not considered as robust information at the regional and grid scales (e.g., Osborn and Hulme, 1997; Trigo and Palutikof, 1999). This justifies downscaling the climate variables such as winds, temperature, specific humidity, which are realistically simulated by GCM, to precipitation.

Development of SVM and ANN downscaling models

The predictors, which are selected following the foregoing analysis, are standardized. Standardization is widely used prior to statistical downscaling to reduce systematic bias (if any) in the mean and variance of GCM predictors relative to NCEP reanalysis data (Wilby et al., 2004). The procedure typically involves subtraction of mean and division by the standard deviation of the predictor for a pre-defined baseline period. In this study standardization is done for each season separately for a baseline period extending from 1948 to 2002.

Most of the predictor variables, which are screened from a pool of possible predictors by using their cross-correlations with predictand, are highly correlated and convey similar information. Therefore, to obtain relevant predictors as

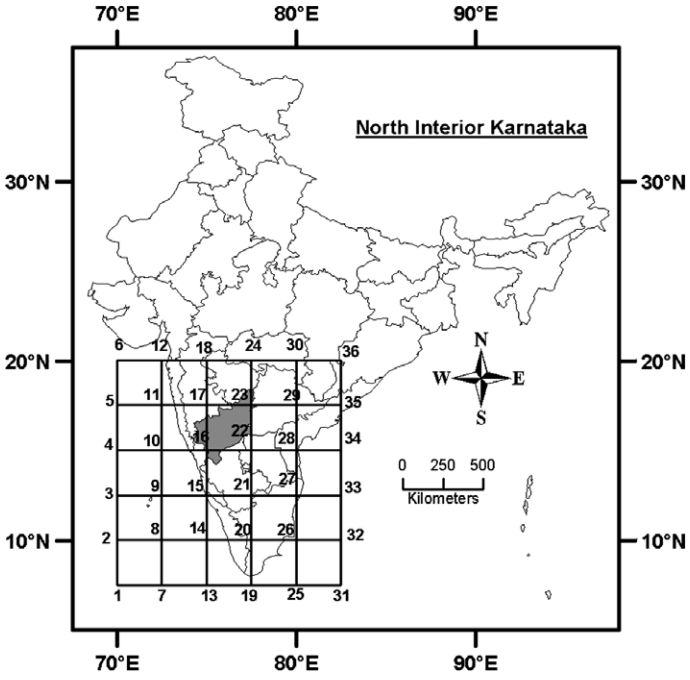

Figure 5 Thirty-six NCEP grid points considered for identification of predictor variables influencing precipitation in the North Interior Karnataka subdivision (shown in gray colour) of India.

input to the SVM and the ANN downscaling models, the standardized NCEP predictor variables are then processed using principal component analysis (*PCA*) to extract principal components (*PCs*) which are orthogonal and which preserve more than 98% of the variance present in them. Precipitation (predictand) constitutes the output from the downscaling models. The use of *PCs* as input to a downscaling model helps in making the model more stable and at the same time reduces its computational burden.

To develop the SVM downscaling model, the available data set is randomly partitioned into a training set and a test set following the multifold cross-validation procedure documented in Haykin (2003, pp. 213–218). About 70% of the available record is randomly selected for training while the remaining 30% is used for testing. The training set is further partitioned into *K* disjoint sets. The model is trained, for a chosen set of parameters, on all the subsets except for one and the validation error is measured on the subset left out. The procedure is repeated for a total of *K* trials, each time using a different subset for validation. Average of the normalized mean squared error under validation over all the trials of the experiment is used to assess the performance of the model.

The training of SVM involves selection of the model parameters σ and *C*. In this study, grid search procedure (Gestel et al., 2004) is used to find the optimum range for the parameters. The optimum values of parameters are then obtained from the selected range using stochastic search technique of genetic algorithm (Haupt and Haupt, 2004). The model parameters σ and *C* are encoded in a binary format representing a chromosome. Initial population of chromosome is randomly generated from the range of parameters obtained through grid search. Fitness of each member of the population is then evaluated using average of the normalized mean squared error (under validation) as fitness function. Based on the value of fitness function, genetic operations (mutation and cross-over) are performed on initial population to yield offspring for the next generation. These offsprings constituted a new population. The procedure of evaluating fitness of population and performing genetic operations is repeated till the improvement in the value of fitness function between two subsequent generations is sufficiently small. Readers are referred to Feng et al. (2004), Pai and Hong (2005) and Zheng and Jiao (2004) for further details on using genetic algorithms for updating parameters of SVM.

The ANN model is trained following the procedure described in ASCE Task Committee on Artificial Neural Networks in Hydrology (2000a, pp. 121–122). The architecture of ANN is decided by trial and error procedure. A comprehensive search of ANN architecture is done by varying the number of hidden layers from 1 to 3 and varying the number of nodes in hidden layer(s). The network is trained using back-propagation algorithm. Logistic sigmoid activation function has been used in hidden layer(s), whereas linear activation function has been used in the output layer. The network error is computed by comparing the network output with the target or the desired output. Mean square error is used as an error function. The model has been subjected to the aforementioned cross-validation procedure. The generalization performance of the SVM and the ANN downscaling models is measured on the test set which is different from the validation subset.

Downscaling GCM simulations to precipitation

The GCM simulations are run through the calibrated and validated SVM downscaling model to obtain future simulations of predictand. The future values of simulated predictand are divided into 5 periods, each 20 years long (2000–2019, 2020–2039, 2040–2059, 2060–2079, and 2080–2099). This is done to determine the trend in projected values of precipitation.

Results

Typical results of seasonal stratification performed using *K*-means algorithm are presented in Fig. 6. It is evident from the figure that the onset of wet season and its duration has varied from year to year in the past. This is in agreement with the findings of India Meteorological Department. It is to be recalled that labels are assigned to feature vectors prepared from NCEP and GCM data during seasonal stratification to denote the season to which they belong. Comparison of the labels of contemporaneous feature vectors of NCEP and GCM past data indicates that the GCM simulations represent the regional climate of the sub-divisions fairly well, during the period 1948–2002. Furthermore, the results indicate that the duration of wet season increases in future over several MSDs in the study region.

The predictor variables, which have maximum influence on the precipitation of each MSD, are identified following the procedure described earlier in this paper. For brevity, the predictors identified for North-Interior Karnataka subdivision are shown in Table 2 for wet season. The selected predictors for each MSD are then used to develop a SVM downscaling model. The optimal values of parameters C and σ are obtained using cross-validation procedure described in foregoing section. In this study cross-validation of the model is carried out with ten ($K = 10$) folds. Typical results of the same are provided in Fig. 7 for Coastal Andhra Pradesh and East Madhya Pradesh sub-divisions for wet season. The parameters of the SVM downscaling model (σ and *C*) developed for each MSD are listed in Table 1. In dry season, precipitation is reasonable only for coastal Andhra Pradesh and Tamil Nadu subdivisions. Therefore results are presented only for these subdivisions.

The width of RBF kernel σ used in this work can give 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, attenuating higher order frequencies and thus resulting in a smooth function. Alternatively, RBF kernel with small kernel width retains most of the higher order frequencies leading to an approximation of a complex function by learning machine.

To investigate the potential of SVM in downscaling GCM simulations, its performance is compared with that of multi-layer back-propagation neural network based downscaling model. The quantitative evaluation of the performance of

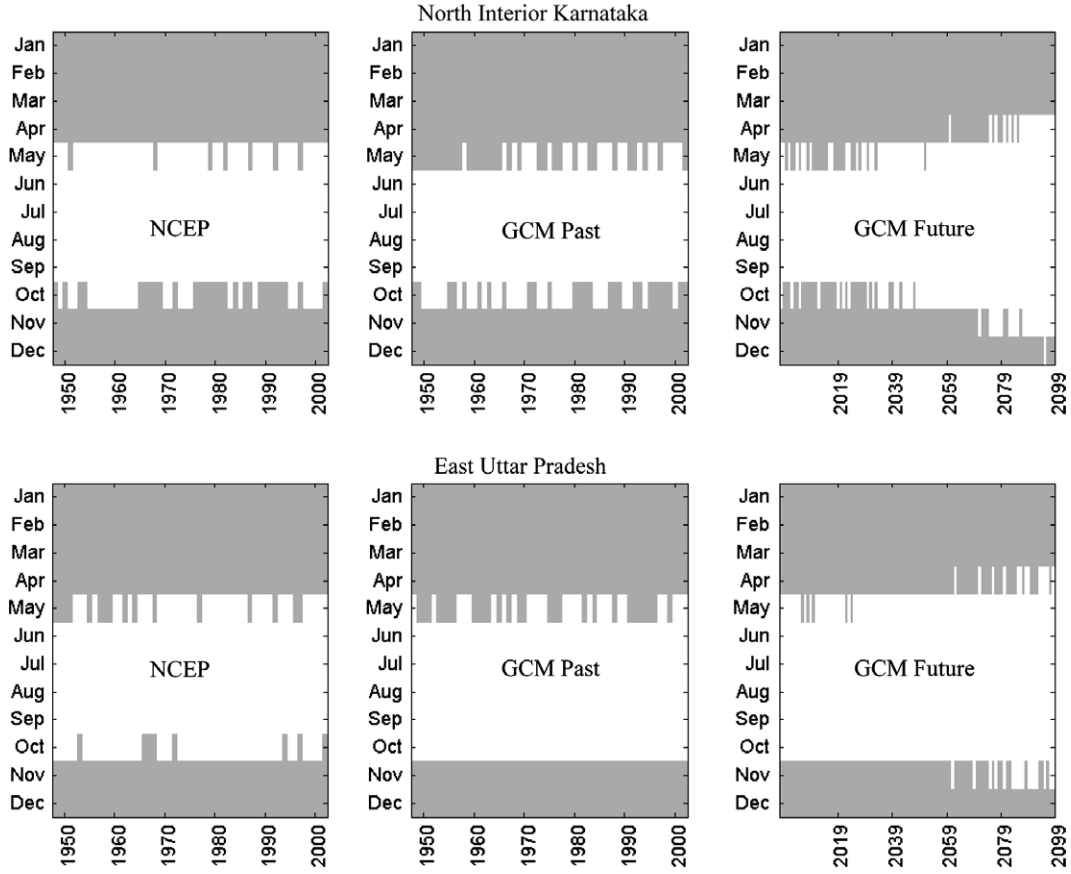

Figure 6 Results of seasonal stratification performed using cluster analysis for North Interior Karnataka and East Uttar Pradesh, India. Dry season is shown in gray shade, whereas the wet season is shown in white colour.

Table 2 The list of identified predictors and their domain for North Interior Karnataka sub-division of India for wet season

Predictors	Domain (grid point numbers)																
Air temperature at 500 mb	6	12	17	18	23	24	28	29	30	34	35	36					
Geopotential height at 200 mb	12	18	24	30													
Specific humidity at 500 mb	24	30	36														
Specific humidity at 850 mb	2	3	4	5	6	8	9	10	11	12	15	16	17	18	30	33	36
Zonal wind speed at 850 mb	1	2	3	7	8	13	14	19	20	25	26	31					

The geographic location of grid points listed in the table can be found in Fig. 5. Principal components that are extracted from the predictors at the specified grid points form input to the downscaling models.

ANN and SVM downscaling models is done using the following statistics:

(i) Relative bias in preservation of mean and standard deviation of the observed precipitation.
(ii) Normalized mean square error (NMSE, Zhang and Govidaraju, 2000) defined as the ratio of the mean square error in simulating observed precipitation to the variance of the observed (desired) precipitation.

$$\mathrm{NMSE} = \frac{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{(S_{\mathrm{obs}})^2} \tag{24}$$

where N represents the number of feature vectors prepared from the NCEP record, y_i and $\hat{y}_i$ denote the observed and the simulated precipitation respectively, and S_{obs} is the standard deviation of the observed precipitation. For an ideal model the relative bias in preservation of mean and standard deviation, and NMSE value must be zero. Typical results of this analysis are presented in Table 3 for Bihar Plateau, Kerala, Coastal Karnataka and Orissa sub-divisions for wet season. From the table it is evident that the SVM-based downscaling model outperforms the multi-layer back-propagation neural network based downscaling model.

In predicting values near the tails of the observed rainfall distribution, the performance of SVM is marginally better

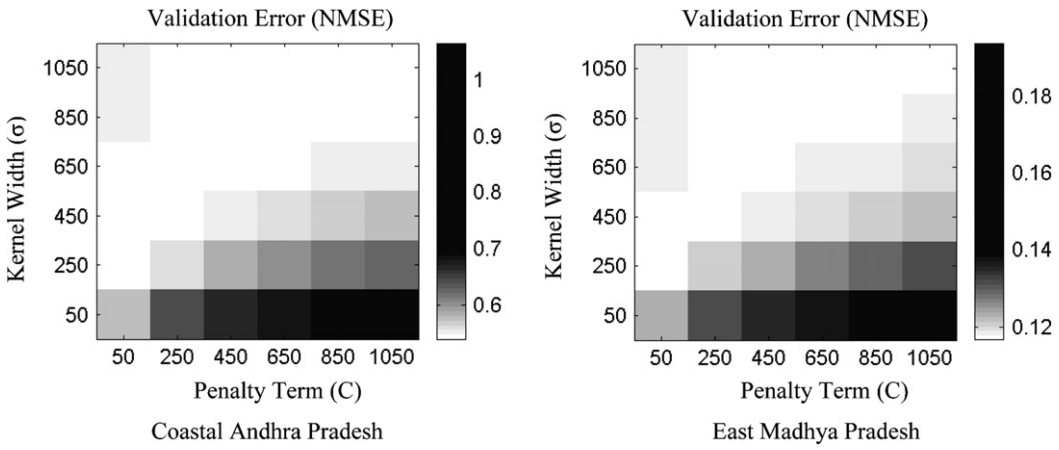

Figure 7 Illustration of the domain search performed to estimate optimal values of kernel width (σ) and penalty (C) for the SVM for two typical meteorological sub-divisions for wet season.

Table 3 Statistical comparison of ANN and SVM downscaling model simulations for four typical meteorological subdivisions of India for wet season (Stdev is the standard deviation of the observed/simulated precipitation; NMSE represents normalized mean square error; values in parenthesis denote percentage relative bias of simulated value with respect to observed value)

Sub-division name	Training/ testing	Observed		ANN simulated value			SVM simulated value		
		Mean (mm)	Stdev (mm)	Mean (mm)	NMSE	Stdev (mm)	Mean (mm)	NMSE	Stdev (mm)
Bihar Plateau	Training	1153.62	196.23	1177.83 (2.10%)	0.59	135.25 (31.08%)	1154.17 (0.05%)	0.56	132.37 (32.55%)
	Testing	1270.28	234.06	1242.95 (2.15%)	0.49	153.17 (34.56%)	1252.75 (1.38%)	0.46	160.41 (31.47%)
Kerala	Training	2244.22	486.56	2247.48 (0.15%)	0.47	336.53 (30.83%)	2244.22 (0.00%)	0.24	453.89 (6.71%)
	Testing	2319.19	442.48	2179.18 (6.04%)	0.59	246.38 (44.32%)	2259.62 (2.57%)	0.46	314.53 (28.92%)
Coastal Karnataka	Training	3259.57	663.73	3254.54 (0.15%)	0.62	327.56 (50.65%)	3259.57 (0.00%)	0.42	468.19 (29.46%)
	Testing	2951.23	396.91	3303.93 (11.95%)	1.51	331.07 (16.59%)	3152.60 (6.82%)	0.96	368.42 (7.18%)
Orissa	Training	1279.15	213.22	1312.61 (2.62%)	0.55	137.70 (35.42%)	1279.15 (0.00%)	0.51	146.27 (31.40%)
	Testing	1285.50	209.31	1355.36 (5.43%)	0.73	142.75 (31.80%)	1310.50 (1.94%)	0.27	193.09 (7.75%)

The architecture of ANN downscaling model for Bihar Plateau, Kerala, Coastal Karnataka and Orissa is 16:16:1, 17:16:1, 13:21:1, and 14:20:1 respectively (a:b:c indicates that the respective number of nodes in input, hidden and output layers of ANN are a, b and c). The number of nodes in the input layer of ANN downscaling model for a subdivision is equal to the number of principal components extracted for the subdivision.

than that of ANN (Fig. 8). However, it is clear that even SVM is not able to mimic the extreme rainfall observed in the record. Possibly this could be because regression based statistical downscaling models often cannot explain entire variance of the downscaled variable (Wilby et al., 2004). Exploration of a wider range of predictor variables and a much longer validation phase could possibly provide more insight into this problem. However, in the present study, investigation in this direction is constrained by the paucity of data.

Typical results from the SVM-based downscaling model for wet and dry seasons are presented in Figs. 9–11 using boxplots. The span of the box represents the interquartile range of the simulated (or observed) rainfall. The whiskers extend from the box to 5% and 95% quantiles on the lower and the upper side of the box, respectively.

The statistical significance of the results is assessed using null hypothesis considering three significance levels (1%, 5% and 10%). For the null hypothesis test it is assumed that the variances of past and projected precipitation are unknown

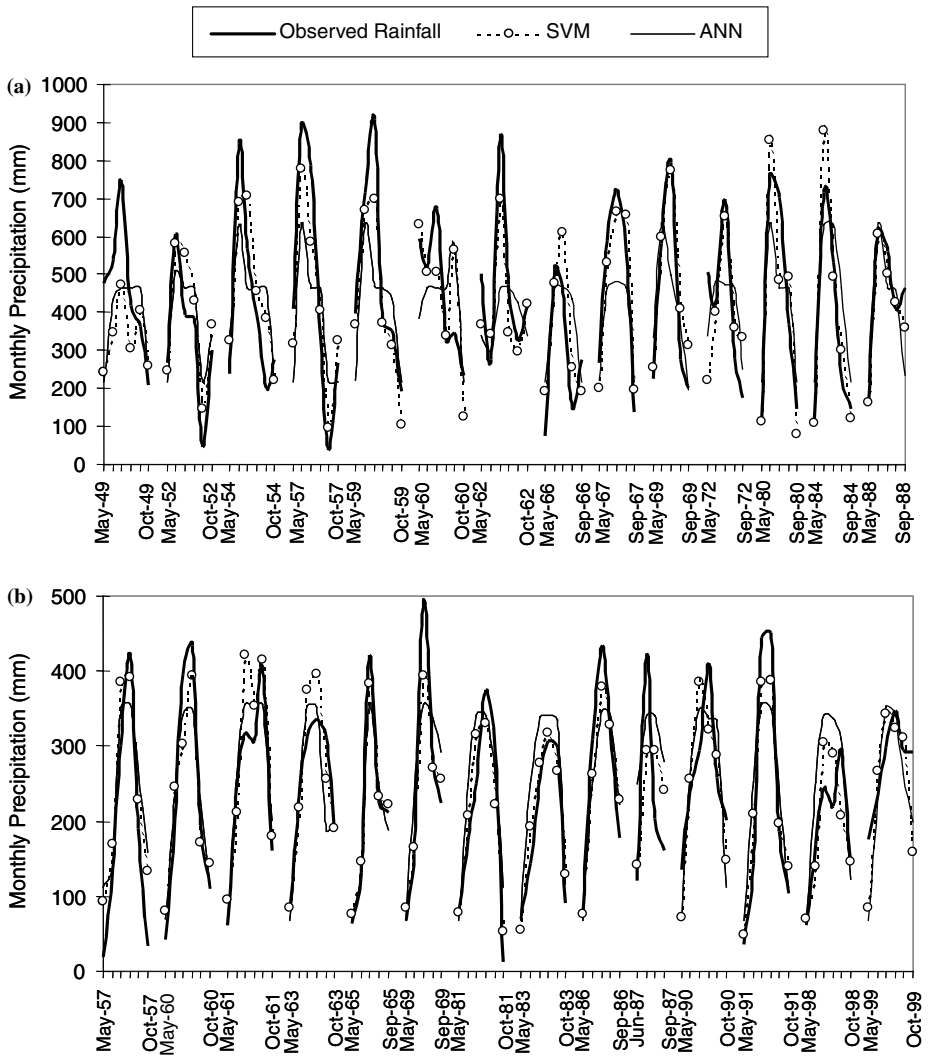

Figure 8 Comparison of the observed wet season monthly rainfall with that simulated using SVM and ANN downscaling models for (a) Kerala and (b) Orissa meteorological sub-divisions for the testing phase.

and unequal. The test statistic, T, is computed using Eq. (25) (Kottegoda and Rosso, 1998), which has an approximate t-distribution with υ degrees of freedom given by Eq. (26).

$$T = \frac{(\bar{X}_{obs} - \bar{X}_{sim})}{\sqrt{(S^2_{obs}/N) + (S^2_{sim}/N)}} \tag{25}$$

$$\upsilon = \frac{(S^2_{obs}/N) + (S^2_{sim}/N)}{[(S^2_{obs}/N)^2/(N-1)] + [(S^2_{sim}/N)^2/(N-1)]} \tag{26}$$

where $\bar{X}_{obs}$ and $\bar{X}_{sim}$ are estimated means of the observed and the simulated precipitation respectively, S_{obs} and S_{sim} are respectively the standard deviations of the observed and simulated precipitation.

The projected scenarios of precipitation for 2000–2019, 2020–2039, 2040–2059, 2060–2079, and 2080–2099 are displayed in Fig. 12. Significant increase in precipitation is projected for Konkan and Goa, Coastal Karnataka, Gujarat, Saurashtra and Kutch along west coast of India, Coastal Andhra Pradesh along east coast, Telangana and Rayalaseema in peninsula India, Punjab and Haryana in the north-west, east Uttar Pradesh, west Uttar Pradesh plains and Bihar plains in the north, and north Assam and south Assam in the north-east India. Drop in precipitation is projected for Kerala and East Madhya Pradesh, while mixed trend in precipitation is projected for the remaining parts of the country by the SVM downscaling model for the simulations of CGCM2 model under IS92a scenario.

The results presented in this section strongly support SVM downscaling model as a feasible and potential alternative to multi-layer back-propagation neural network based downscaling model for climate impact studies in hydrology.

Summary and discussion

In this paper support vector machine (SVM) approach has been introduced for statistical downscaling of precipitation at monthly time scale from simulations of GCM. The effectiveness of the proposed approach is illustrated through its application to future climate projections provided by CGCM2 over India. The proposed model is shown to be statistically superior compared to the multi-layer back-propagation

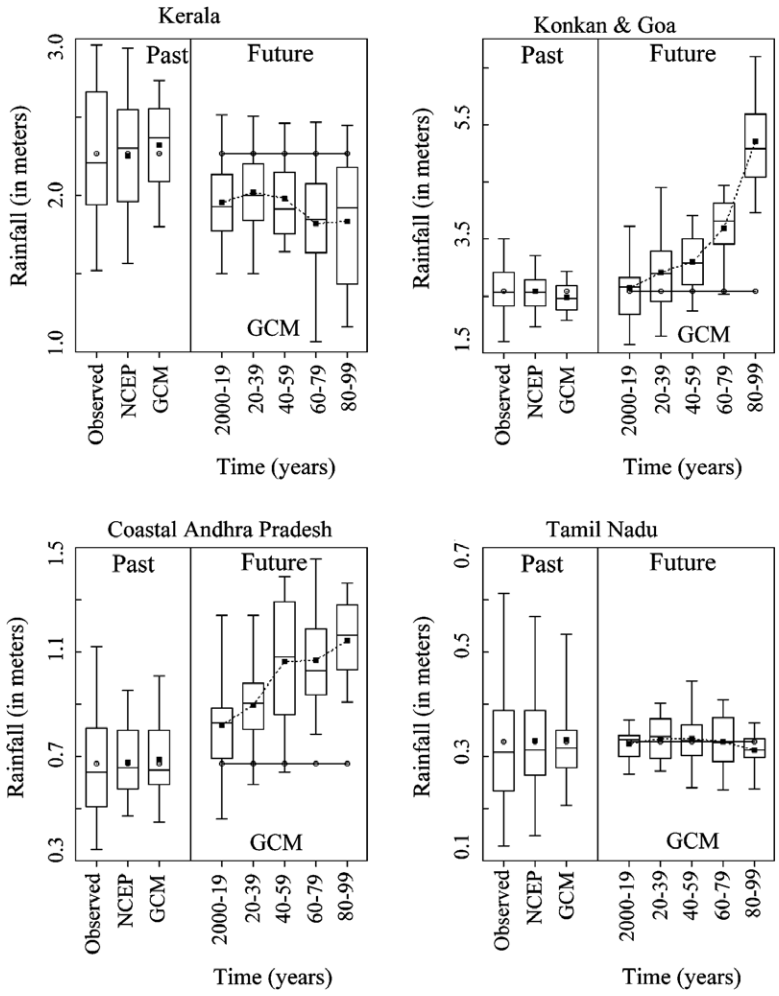

Figure 9 Typical results from the SVM based downscaling model for wet season for the south peninsula-India graphed using box plots. The horizontal line in the middle of the box represents median. The circle denotes the mean value of observed rainfall and the darkened square represents the mean value of simulated rainfall. In the left part of figure for each sub-division, the gap between darkened square and circle denotes bias in the rainfall simulated by the downscaling model for NCEP and GCM data sets. Whereas in the right part of the figure the solid line that joins the circles indicates the historical trend of rainfall, while the dotted line connecting the solid squares depicts the mean trend simulated by GCM.

neural network based downscaling model in vogue in literature. The results suggest that SVM offers considerable prospects for improvements in climate impact studies in hydrology.

Some broad contiguous areas showing statistically significant trends in precipitation have been identified in the study region. Decreasing trends in the wet season rainfall are found over East Madhya Pradesh and Kerala, while increasing trends in the same are found over several parts of the country.

Nevertheless, it is worth mentioning that the future projections of hydrologic variables provided by a downscaling model for a given climate change scenario depend on the capability of GCM to simulate future climate. A realistic simulation by GCM could yield pragmatic solution of predictand, while an inconsistent simulation could result in absurd values of the predictand. Hence, it is necessary to use simulations from more than one GCM for a given climate change scenario to test the robustness of the result projected by the downscaling model. However, in this paper we devoted our efforts to present SVM as a potential alternative to ANN downscaling model. Hence we confine ourselves to CGCM2 model for the IS92a scenario.

The choice of predictor variables can significantly affect the result of a downscaling model. Since the climate variables affecting the precipitation vary across time and space, there is a need to identify/develop robust framework for selection of predictor variables in different parts of the world.

Besides this, there are uncertainties associated with predictions of future climate change scenario and the assumption that the empirical relationships developed for the current state of atmosphere remains valid in the future. In spite of these uncertainties/assumptions downscaling remains the most popular tool for hydrologists to assess the impact of climate change on hydrological processes of a region.

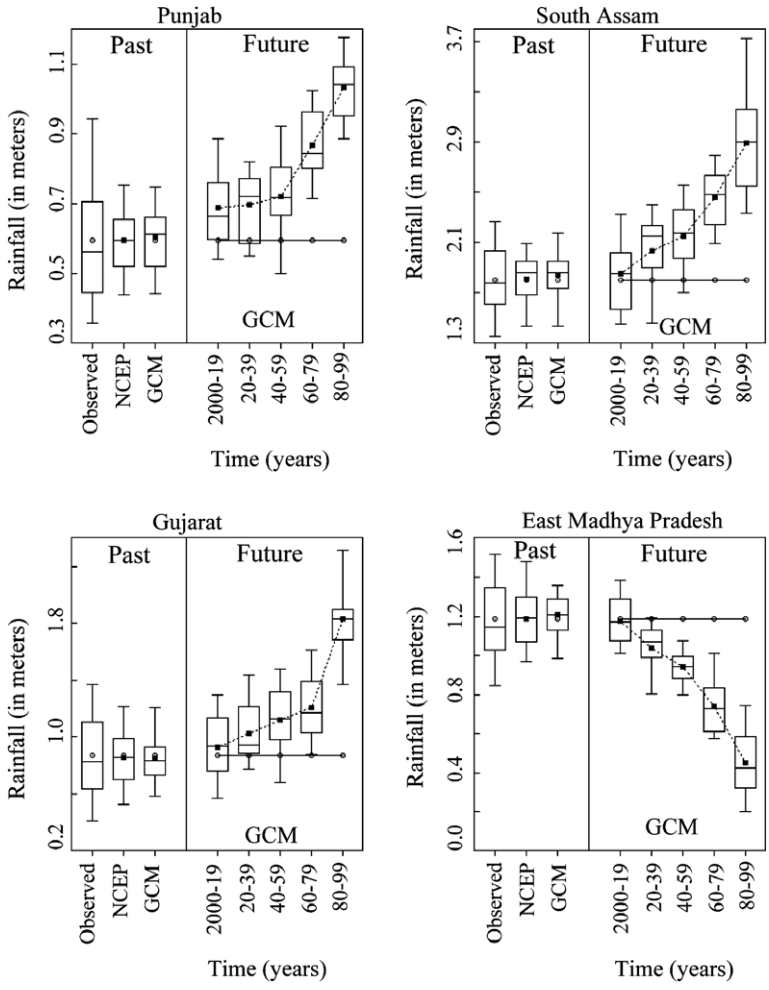

Figure 10 Typical results from the SVM based downscaling model for wet season for the north, central and north-eastern India.

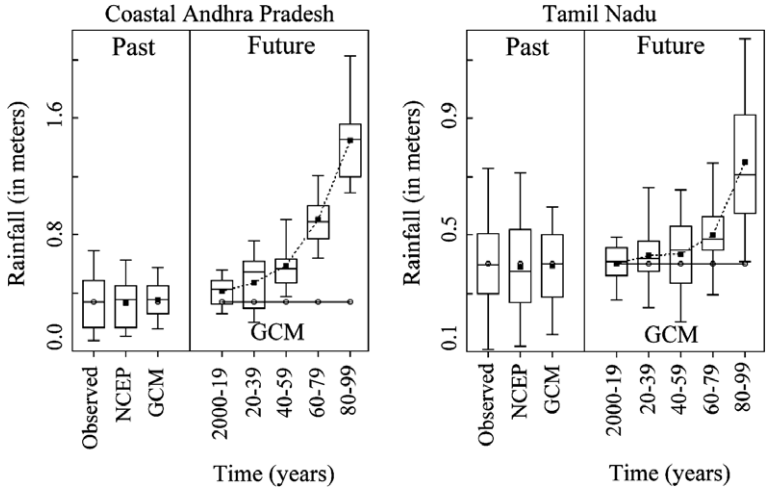

Figure 11 Typical results from the SVM based downscaling model for dry season.

Statistical downscaling is popular in literature because its computational overheads are almost insignificant compared to dynamic downscaling. However, developing a statistical downscaling model for finer time scales (such as daily or hourly) often becomes a challenging task due to high memory requirements and slow convergence associated with modeling large data sets.

The proposed SVM approach to downscaling is computationally more intensive than the ANN method adopted in the present study. Training of standard SVM with interior

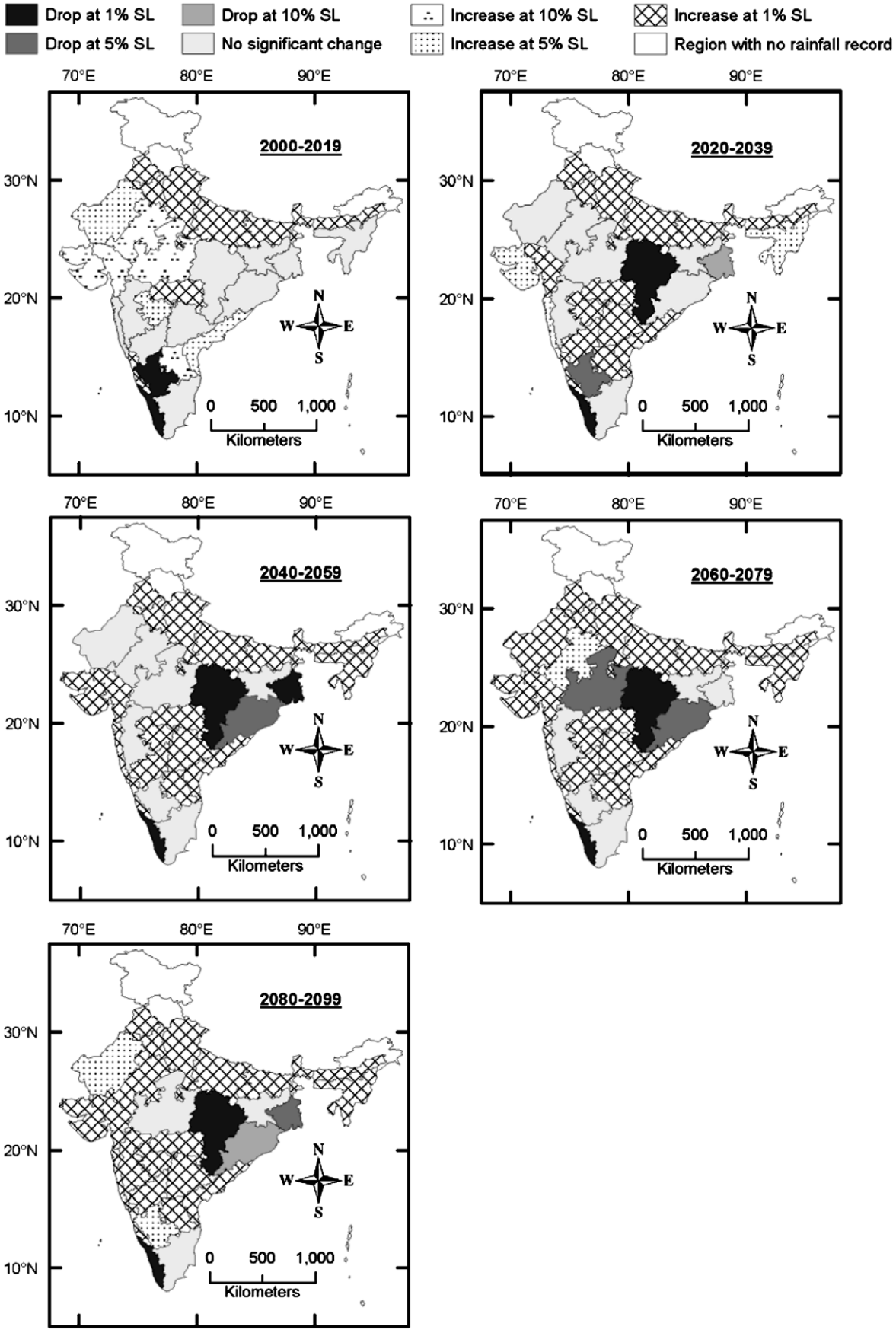

Figure 12 Precipitation in meteorological subdivisions of India projected for wet season by CGCM2 model for IS92a scenario — result from SVM downscaling model.

point methods and training of LS-SVM with elimination method has $O(N^3)$ time and $O(N^2)$ space complexities where N is the training set size. These methods are suitable for relatively small data sets ($N \approx 2000$, depending upon the computer memory). For large data sets one can train standard SVM and LS-SVM by using decompositions methods (e.g., chunking, sequential minimum optimization) or iterative methods (e.g., successive over-relaxation, conjugate gradient). These techniques have been successfully applied on massive data sets with millions of data points (Mangasarian and Musicant, 1999; Suykens et al., 2002).

On the other hand, the complexity of the developed multi-layer back-propagation neural network depends on the number of layers and the number of nodes in the network. The standard back propagation algorithm scales badly to large problems. That is, run times increase quickly with problem size. The conjugate gradient back propagation and the standard back-propagation algorithms were found to have near equal median time complexities, approximately $O(N^4)$, where N denotes the size of the problem (Stone and Lister, 1994; Lister and Stone, 1995).

In the context of downscaling large data sets, ensemble methods like bagging (Breiman, 1996; Breiman, 1998) and boosting (Freund, 1995; Freund and Schapire, 1997; Schwenk and Bengio, 2000) that combine the outputs of several artificial neural networks could be feasible alternatives to SVM. These methods could, in a different context, achieve better accuracy at faster speeds (Hernandez-Espinosa et al., 2004; Pasquariello et al., 2002). However, these newly developed statistical learning techniques are yet to find their way to downscaling applications and the comparison of their relative performance is still an open research issue.

Several avenues should be explored to further refine this attempt to statistically downscale GCM simulations. The proposed approach to statistical downscaling of precipitation can be readily extended to downscaling of a variety of variables of interest to hydrologists including temperature, wind speed, soil moisture, evapo-transpiration, runoff, and streamflows. Extended research work in this direction is underway.

Acknowledgements

The authors express their sincere thanks to the two anonymous reviewers for their constructive comments and suggestions on the earlier draft of the paper. The authors express their gratitude to Prof. M. Narasimha Murty, Department of Computer Science and Automation, Indian Institute of Science, Bangalore, for his valuable input on Support Vector Machines. The support provided by the Department of Science and Technology, India (DO No. SR/FTP/ETA-25/2003), to the second author is also duly acknowledged.

References

Arnell, N.W., Hudson, D.A., Jones, R.G., 2003. Climate change scenarios from a regional climate model: Estimating change in runoff in southern Africa. Journal of Geophysical Research — Atmospheres 108 (D16), AR 4519.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000a. Artificial neural networks in hydrology, I: Preliminary concepts. Journal of Hydrologic Engineering, ASCE 5 (2), pp. 115—123.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000b. Artificial neural networks in hydrology, II: Hydrologic applications. Journal of Hydrologic Engineering, ASCE 5(2), 124—137.

Asefa, T., Kemblowski, M.W., Urroz, G., McKee, M., Khalil, A., 2004. Support vectors-based groundwater head observation networks design. Water Resources Research 40 (11), W11509. doi:10.1029/2004WR003304.

Bhaskar, N.R., O'Connor, C.A., 1989. Comparison of method of residuals and cluster analysis for flood regionalization. Journal of Water Resources Planning and Management 115 (6), 793—808.

Bianchini, M., Gori, M., 1996. Optimal learning in artificial neural networks: A review of theoretical results. Neurocomputing 13 (2—4), 313—346.

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Oxford University Press, New York.

Bouraoui, F., Vachaud, G., Li, L.Z.X., Le Treut, H., Chen, T., 1999. Evaluation of the impact of climate changes on water storage and groundwater recharge at the watershed scale. Climate Dynamics 15, 153—161.

Breiman, L., 1996. Bagging predictors. Machine Learning 24 (2), 123—140.

Breiman, L., 1998. Arcing classifiers. Annals of Statistics 26 (3), 801—824.

Buma, J., Dehn, M., 2000. Impact of climate change on a landslide in South East France, simulated using different GCM scenarios and downscaling methods for local precipitation. Climate Research 15 (1), 69—81.

Burn, D.H., 1989. Cluster analysis as applied to regional flood frequency. Journal of Water Resources Planning and Management 115 (5), 567—582.

Cannon, A.J., McKendry, I.G., 2002. A graphical sensitivity analysis for statistical climate models: Application to Indian monsoon rainfall prediction by artificial neural networks and multiple linear regression models. International Journal of Climatology 22, 1687—1708.

Cannon, A.J., Whitfield, P.H., 2002. Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models. Journal of Hydrology 259 (1), 136—151.

Cavazos, T., 1997. Downscaling large-scale circulation to local winter rainfall in north-eastern Mexico. International Journal of Climatology 17 (10), 1069—1082.

Chhabra, B.M., Hatwar, H.R., Das Gupta, M., 1999. Onset of southwest monsoon — A diagnostic study. In: Gupta, R.K., Reddy, S.J. (Eds.), Advanced Technologies in Meteorology, Symposium Proceedings. Tata McGraw-Hill, New Delhi, India (Chapter 5).

Cortes, C., Vapnik, V., 1995. Support vector networks. Machine Learning 20, 273—297.

Courant, R., Hilbert, D., 1970. Methods of Mathematical Physics, vols. I and II. Wiley Interscience, New York.

Cover, T.M., 1965. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. IEEE Transactions on Electronic Computers EC-14, 326—334.

Cover, T.M., Hart, P.E., 1967. Nearest neighbor pattern classification. IEEE Transactions on Information Theory IT-13, 21—27.

Crane, R.G., Hewitson, B.C., 1998. Doubled CO_2 precipitation changes for the Susquehanna Basin: Down-Scaling from the Genesis General Circulation Model. International Journal of Climatology 18, 65—76.

Crane, R.G., Yarnal, B., Barron, E.J., Hewitson, B., 2002. Scale interactions and regional climate: Examples from the

Susquehanna River Basin. Human and Ecological Risk Assessment 8 (1), 147–158.
Cristianini, N., Shawe-Taylor, J., 2000. An Introduction to Support Vector Machines and other Kernel-based Learning Methods. Cambridge University Press, Cambridge.
Dibike, Y.B., Velickov, S., Solomatine, D., Abbott, M.B., 2001. Model induction with support vector machines: Introduction and applications. Journal of Computing in Civil Engineering 15 (3), 208–216.
Faucher, M., Burrows, W.R., Pandolfo, L., 1999. Empirical-statistical reconstruction of surface marine winds along the western coast of Canada. Climate Research 11 (3), 173–190.
Feng, X.T., Zhao, H.B., Li, S.J., 2004. A new displacement back analysis to identify mechanical geo-material parameters based on hybrid intelligent methodology. International Journal for Numerical and Analytical Methods in Geomechanics 28 (11), 1141–1165.
Fix, E., Hodges, J.L., 1951. Discriminatory analysis: Nonparametric discrimination: Consistency properties. USAF School of Aviation Medicine, Project 21-49-004, Report No. 4, Randolph Field, TX, pp. 261–279.
Freund, Y., 1995. Boosting a weak learning algorithm by majority. Information and Computation 121 (2), 256–285.
Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences 55 (1), 119–139.
Georgakakos, K.P., Smith, D.E., 2001. Soil moisture tendencies into the next century for the conterminous United States. Journal of Geophysical Research – Atmospheres 106 (D21), 27367–27382.
Gestel, T.V., Suykens, J.A.K., Baesens, B., Viaene, S., Vanthienen, J., Dedene, G., Moor, B.D., Vandewalle, J., 2004. Benchmarking least squares support vector machine classifiers. Machine Learning 54 (1), 5–32.
Govindaraju, R.S., Rao, A.R. (Eds.), 2000. Artificial Neural Networks in Hydrology. Kluwer Academic Publishers, Holland, pp. 329.
Hassan, H., Hanaki, K., Matsuo, T., 1998. A modeling approach to simulate impact of climate change in lake water quality: Phytoplankton growth rate assessment. Water Science and Technology 37 (2), 177–185.
Haupt, R.L., Haupt, S.E., 2004. Practical Genetic Algorithm. Wiley, New Jersey, pp. 253.
Haykin, S., 2003. Neural Networks: A comprehensive foundation. Fourth Indian Reprint, Pearson Education, Singapore, pp. 842.
Hernandez-Espinosa, C., Fernandez-Redondo, M., Torres-Sospedra, J., 2004. Some experiments on ensembles of neural networks for hyperspectral image classification, Knowledge-Based Intelligent Information and Engineering Systems. PT 1 Proceedings, Lecture Notes in Computer Science 3213, 677–684.
Hewitson, B.C., Crane, R.G., 1992. Large-scale atmospheric controls on local precipitation in tropical Mexico. Geophysical Research Letters 19 (18), 1835–1838.
Hewitson, B.C., Crane, R.G., 1996. Climate downscaling: Techniques and application. Climate Research 7, 85–95.
Hush, D.R., Horne, B.G., 1993. Progress in supervised neural networks: What's new since Lippmann? IEEE Signal Processing Magazine 10, 8–39.
IPCC, 1992. Leggett, J., Pepper W.J., Swart, R., (Eds.) Climate change 1992, the supplementary report to the IPCC scientific assessment, Intergovernmental Panel on Climate Change IPCC, Cambridge University Press, Cambridge, UK, (Chapter 3).
IPCC, 2001. McCarthy, J.J., Canziani, O.F., Leary, N.A., Dokken, D.J., White, K.S. (Eds.), Climate Change 2001: Impacts, Adaptation and Vulnerability, Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK.
Jasper, K., Calanca, P., Gyalistras, D., Fuhrer, J., 2004. Differential impacts of climate change on the hydrology of two alpine river basins. Climate Research 26 (2), 113–129.
Kailas, S.V., Narasimha, R., 2000. Quasi cycles in monsoon rainfall by wavelet analysis. Current Science 78, 592–595.
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., Joseph, D., 1996. The NCEP/NCAR 40-year reanalysis project. Bulletin of the American Meteorological Society 77 (3), 437–471.
Keerthi, S.S., Lin, C.-J., 2003. Asymptotic behaviors of support vector machines with Gaussian kernel. Neural Computation 15 (7), 1667–1689.
Kettle, H., Thompson, R., 2004. Statistical downscaling in European mountains: verification of reconstructed air temperature. Climate Research 26 (2), 97–112.
Khadam, I.M., Kaluarachchi, J.J., 2004. Use of soft information to describe the relative uncertainty of calibration data in hydrologic models. Water Resources Research 40 (11), W11505. doi:10.1029/2003WR002939.
Kim, M.K., Kang, I.S., Park, C.K., Kim, K.M., 2004. Superensemble prediction of regional precipitation over Korea. International Journal of Climatology 24 (6), 777–790.
Kottegoda, N.T., Rosso, R., 1998. Statistics, Probability, and Reliability for Civil and Environmental Engineers. McGraw-Hill, Singapore.
Lal, M., Cubasch, U., Voss, R., Waszkewitz, J., 1995. Effect of transient increase in greenhouse gases and sulfate aerosols on monsoon climate. Current Science 69, 752–763.
Lin, H.-T., Lin, C.-J., 2003. A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. Technical report, Department of Computer Science and Information Engineering, National Taiwan University.
Lister, R., Stone, J.V., 1995. An empirical study of the time complexity of various error functions with conjugate gradient back propagation. In: Proceedings of IEEE International Conference on Neural Networks, Perth, Australia, pp. 237–241.
MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations. In: Le Cam L.M., Neyman J. (Eds.), Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, University of California Press, Berkeley, pp. 281–297.
Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environmental Modelling and Software 15 (1), 101–124.
Maini, P., Kumar, A., Singh, S.V., Rathore, L.S., 2004. Operational model for forecasting location specific quantitative precipitation and probability of precipitation over India. Journal of Hydrology 228, 170–188.
Mangasarian, O.L., Musicant, D.R., 1999. Successive overrelaxation for support vector machines. IEEE Transactions on Neural Networks 10 (5), 1032–1037.
McCulloch, W.S., Pitts, W., 1943. A logic calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics 5, 115–133.
Meireles, M.R.G., Almeida, P.E.M., Simoes, M.G., 2003. A comprehensive review for industrial applicability of artificial neural networks. IEEE Transactions on Industrial Electronics 50 (3), 585–601.
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.
Misson, L., Rasse, D.P., Vincke, C., Aubinet, M., Francois, L., 2002. Predicting transpiration from forest stands in Belgium for the 21st century. Agricultural and Forest Meteorology 111 (4), 265–282.
Mpelasoka, F.S., Mullan, A.B., Heerdegen, R.G., 2001. New Zealand climate change information derived by multivariate statistical

and artificial neural networks approaches. International Journal of Climatology 21 (11), 1415–1433.

Neocleous, C., Schizas, C., 2002. Artificial neural network learning: A comparative review. Methods and Applications of Artificial Intelligence 2308, 300–313.

NRC, 1998. Decade-to-Century-Scale Climate Variability and Change: A Science Strategy. National Resources Council, Panel on Climate Variability on Decade-to-Century Time Scales, National Academy Press, Washington, DC.

Olsson, J., Uvo, C.B., Jinno, K., Kawamura, A., Nishiyama, K., Koreeda, N., Nakashima, T., Morita, O., 2004. Neural networks for rainfall forecasting by atmospheric downscaling. Journal of Hydrologic Engineering 9 (1), 1–12.

Osborn, T.J., Hulme, M., 1997. Development of a relationship between station and grid-box rainday frequencies for climate model evaluation. Journal of Climate 10, 1885–1908.

Pai, P.F., Hong, W.C., 2005. Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms. Electric Power Systems Research 74 (3), 417–425.

Parthasarathy, B., Rupa Kumar, K., Munot, A.A., 1993. Homogeneous Indian monsoon rainfall: Variability and prediction. Proceedings of the Indian Academy of Sciences (Earth and Planetary Sciences) 102 (1), 121–155.

Parthasarathy, B., Munot, A.A., Kothawale, D.R., 1994. All India monthly and seasonal rainfall series: 1871–1993. Theoretical and Applied Climatology 49, 217–224.

Pasquariello, G., Ancona, N., Blonda, P., Tarantino, C., Satalino, G., D'Addabbo, A., 2002. Neural network ensemble and support vector machine classifiers for the analysis of remotely sensed data: A comparison. In: IEEE International Geoscience and Remote Sensing Symposium, Toronto, Canada, pp. 509–511.

Poulton, M.M., 2002. Neural networks as an intelligence amplification tool: A review of applications. Geophysics 67 (3), 979–993.

Rupa Kumar, K., Pant, G.B., Parthasarathy, B., Sontakke, N.A., 1992. Spatial and subseasonal patterns of the long-term trends of Indian summer monsoon rainfall. International Journal of Climatology 12 (3), 257–268.

Rupa Kumar, K., Krishna Kumar, K., Pant, G.B., 1994. Diurnal asymmetry of surface temperature trends over India. Geophysical Research Letters 21, 677–680.

Sailor, D.J., Hu, T., Li, X., Rosen, J.N., 2000. A neural network approach to local downscaling of GCM output for assessing wind power implications of climate change. Renewable Energy 19 (3), 359–378.

Sastry, P.S., 2003. An introduction to support vector machines. In: Misra, J.C. (Ed.), Computing and Information Sciences: Recent Trends. Narosa Publishing House, New Delhi.

Schmidt, M., Glade, T., 2003. Linking global circulation model outputs to regional geomorphic models: a case study of landslide activity in New Zealand. Climate Research 25 (2), 135–150.

Schölkopf, B., Burges, C., Smola, A. (Eds.), 1998. Advances in Kernel Methods – Support Vector Learning. MIT Press.

Schoof, J.T., Pryor, S.C., 2001. Downscaling temperature and precipitation: A comparison of regression-based methods and artificial neural networks. International Journal of Climatology 21 (7), 773–790.

Schwenk, H., Bengio, Y., 2000. Boosting neural networks. Neural Computation 12 (8), 1869–1887.

Shannon, D.A., Hewitson, B.C., 1996. Cross-scale relationships regarding local temperature inversions at Cape Town and global climate change implications. South African Journal of Science 92 (4), 213–216.

Sharma, C., Roy, J., Rupa kumar, K., Chadha, D.K., Singh, R.N., Saheb, S.P., Mitra, A.P., 2003. Impacts of climate change on water resources in India. In: Muhammed A. (Ed.), Proceedings of Year End Workshop on Climate Change and Water Resources in South Asia, Kathmandu, Nepal, Asianics Agro Development International publishers, Islamabad, Pakistan, pp. 61–90 (chapter 3).

Shivam, T., 2004. Downscaling of General Circulation Models to Assess the Impact of Climate Change on Rainfall of India, ME thesis, Indian Institute of Science, Bangalore, India.

Smola, A.J., Scholkopf, B., Muller, K.R., 1998. The connection between regularization operators and support vector kernels. Neural Networks 11 (4), 637–649.

Snell, S.E., Gopal, S., Kaufmann, R.K., 2000. Spatial interpolation of surface air temperatures using artificial neural networks: Evaluating their use for downscaling GCMs. Journal of Climate 13 (5), 886–895.

Solecki, W.D., Oliveri, C., 2004. Downscaling climate change scenarios in an urban land use change model. Journal of Environmental Management 72 (1–2), 105–115.

Stone, J.V., Lister, R., 1994. On the relative time complexities of standard and conjugate gradient back-propagation. In: Proceedings of IEEE International Conference on Neural Networks, Orlando, FL, USA, pp. 84–87.

Suykens, J.A.K., 2001. Nonlinear modelling and support vector machines. In: Proceedings of IEEE Instrumentation and measurement technology conference, Budapest, Hungary, pp. 287–294.

Suykens, J.A.K., Gestel, T.V., Brabanter, J.D., Moor, B.D., Vandewalle, J., 2002. Least squares support vector machines. World Scientific, River Edge, NJ, pp. 294.

Tatli, H., Dalfes, H.N., Mentes, S., 2004. A statistical downscaling method for monthly total precipitation over Turkey. International Journal of Climatology 24 (2), 161–180.

Trigo, R.M., Palutikof, J.P., 1999. Simulation of daily temperatures for climate change scenarios over Portugal: a neural network model approach. Climate Research 13 (1), 45–59.

Turing, A.M., 1950. Computing Machinery and Intelligence. Mind 59 236, 433–460.

Vapnik, V.N., 1992. Principles of risk minimization for learning theory. Advances in Neural Information Processing Systems, vol. 4. San Mateo, Morgan Kaufmann, CA, pp. 831–838.

Vapnik, V.N., 1995. The Nature of Statistical Learning Theory. Springer Verlag, New York.

Vapnik, V.N., 1998. Statistical Learning Theory. Wiley, New York.

Vapnik, V.N., Chervonenkis, A.Ya., 1971. On the uniform convergence of relative frequencies of events to their probabilities. Theoretical Probability and its applications 17, 264–280.

Weisse, R., Oestreicher, R., 2001. Reconstruction of potential evaporation for water balance studies. Climate Research 16 (2), 123–131.

Wilby, R.L., 1998. Modelling low-frequency rainfall events using airflow indices, weather patterns and frontal frequencies. Journal of Hydrology 213 (1–4), 380–392.

Wilby, R.L., Wigley, T.M.L., 1997. Downscaling general circulation model output: a review of methods and limitations. Progress in Physical Geography 21 (4), 530–548.

Wilby, R.L., Wigley, T.M.L., 2000. Precipitation predictors for downscaling: observed and General Circulation Model relationships. International Journal of Climatology 20 (6), 641–661.

Wilby, R.L., Wigley, T.M.L., Conway, D., Jones, P.D., Hewitson, B.C., Main, J., Wilks, D.S., 1998. Statistical downscaling of general circulation model output: A comparison of methods. Water Resources Research 34, 2995–3008.

Wilby, R.L., Charles, S.P., Zorita, E., Timbal, B., Whetton, P., 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://ipcc-ddc.cru.uea.ac.uk/guidelines/StatDown_Guide.pdf).

Willmott, C.J., Rowe, C.M., Philpot, W.D., 1985. Small-scale climate map: a sensitivity analysis of some common assumptions

associated with the grid-point interpolation and contouring. American Cartographer 12, 5—16.

Wiltshire, S.E., 1986. Regional flood frequency analysis II Multivariate classification of drainage basins in Britain. Hydrological Sciences Journal 31, 335—346.

Winkler, J.A., Palutikof, J.P., Andresen, J.A., Goodess, C.M., 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.

Xu, C.Y., 1999. From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. Progress in Physical Geography 23 (2), 229—249.

Zhang, B., Govidaraju, R.S., 2000. Prediction of watershed runoff using bayesian concepts and modular neural network. Water Resources Research 36 (3), 753—762.

Zheng, C., Jiao, L., 2004. Automatic parameters selection for SVM based on GA. In: IEEE Conference Proceedings of the Fifth World Congress on Intelligent Control and Automation. Hangzhou, PR China, pp. 1869—1872.

Zhang, X.C., Nearing, M.A., Garbrecht, J.D., Steiner, J.L., 2004. Downscaling monthly forecasts to simulate impacts of climate change on soil erosion and wheat production. Soil Science Society of America Journal 68 (4), 1376—1385.