INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **33**: 312–325 (2013)
Published online 4 January 2012 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.3424

RMetS
Royal Meteorological Society

Influence of climate on incidences of malaria in the Thar Desert, northwest India

Deepak Jhajharia,[a] Surajit Chattopadhyay,[b]* Rahul R. Choudhary,[c] Vas Dev,[d] Vijay P. Singh[e] and Shankar Lal[f]

[a] *Department of Agricultural Engineering, North Eastern Regional Institute of Science and Technology, Nirjuli, Itanagar–791109, Arunachal Pradesh, India*

[b] *Department of Computer Application, Pailan College of Management and Technology, Bengal Pailan Park, Kolkata–700104, West Bengal, India*

[c] *Department of Electronic Instrumentation and Control Engineering, Engineering College Bikaner, Karni Industrial Area, Bikaner–334004, Rajasthan, India*

[d] *National Institute of Malaria Research (Field Station), Chachal, Guwahati-781022, Assam, India*

[e] *Department of Biological and Agricultural Engineering & Department of Civil & Environmental Engineering, Texas A & M University, 2117 TAMU Scoates Hall, College Station, TX 77843-2117, USA*

[f] *Smt. Durgadevi Birla Memorial Government Hospital Chirawa, Chirawa–333026, Jhunjhunu, Rajasthan, India*

ABSTRACT: Climatic variability and rise in temperature are considered as the key determinants to the transmission of malaria. In the present study, the trends in the cases of malaria caused by *Plasmodium falciparum* and *Plasmodium vivax* were investigated by using the nonparametric Mann-Kendall test after removing the effect of significant lag-1 serial correlation from the time series of cases of malaria incidence by pre-whitening in annual, seasonal, and monthly time scales at Bikaner, located in the Thar Desert of Rajasthan, in northwest India. Multi-collinearity within the datasets under consideration was investigated by means of correlation matrix, the Bartlett sphericity test, and the Kaiser-Meyer-Olkin measure of sampling adequacy, subsequent to which it was removed by using principal component analysis. Finally, artificial neural network models were employed to predict cases of malaria incidence caused by *P. falciparum* and *P. vivax* at various scales. During the last 34 years from 1975 to 2008, *P. falciparum* malaria incidence cases have been found to increase significantly corresponding to monthly (April and September) and seasonal (monsoon) time scales over Bikaner. On the other hand, no significant trends were observed in *P. vivax* malaria cases at Bikaner. Concomitant increases in *P. falciparum* cases of malaria incidence and observed temperature increases at Bikaner hint that *P. falciparum* malaria may have grown significantly under the warming climate of the Thar Desert. Copyright © 2012 Royal Meteorological Society

KEY WORDS climate change; malaria; *Plasmodium vivax*; *Plasmodium falciparum*; Bikaner; Thar Desert; Rajasthan; principal component analysis; artificial neural network

Received 29 June 2011; Revised 27 September 2011; Accepted 29 November 2011

1. Introduction

The global mean surface temperature has increased by 0.6 °C over the last 100 years, with 1998 being the warmest year, and most of the increase in global mean temperature has been observed since 1970 (IPCC, 2007). Several researchers (Hingane *et al*., 1985; Srivastava *et al*., 1992; Rupa Kumar *et al*., 1994; Kothyari and Singh, 1996; Yadav *et al*., 2004; Bhutiyani *et al*., 2007; Singh *et al*., 2008; Jhajharia *et al*., 2009; Jhajharia and Singh, 2011; Chattopadhyay *et al*., 2011) have reported a rise in temperature at different time scales over various parts of India, ranging from northeast India to the western and northwestern Himalayas, and from the Ganges basin to the river basins of northwest and central India. These studies in different parts of India are in general agreement with the observed global warming over the rest of the world, but differ in the trends in the diurnal temperature range (DTR), i.e. difference between maximum and minimum temperatures (Srivastava *et al*., 1992; Rupa Kumar *et al*., 1994; Roy and Balling, 2005, Jhajharia and Singh, 2011).

Global warming and climate change are likely to cause an increase in the incidences of vector-borne diseases, such as malaria, yellow fever, and dengue fever (Patz *et al*., 2000). About 300–500 million malaria cases are reported annually, of which about 2 million deaths, mostly children below 5 years of age, occur annually all over the globe (Sturchler *et al*., 1990; Makler *et al*., 1998). About 2.4 million malaria cases are reported annually from south Asia, of which 75% cases are reported in India alone (WHO, 2002; Jain *et al*., 2008). In the

*Correspondence to: S. Chattopadhyay, Department of Computer Application, Pailan College of Management and Technology, Bengal Pailan Park, Kolkata–700104, West Bengal, India.
E-mail: surajit_2008@yahoo.co.in

year 1998, approximately 2000 people and an estimated 577 000 disability-adjusted life years (DALYs) were lost owing to malaria in India (IPCC, 2001). Malaria, an acute infectious disease, is caused by parasites, *Plasmodium spp* and spread by vector, the female Anopheles mosquito. Climatic variables affect the incubation rate of *Plasmodium* parasites and the breeding of Anopheles, and are important environmental contributors to the transmission of malaria (McMichael and Martens, 1995). The malaria vectors are poikilothermic and a change of temperature by 1 °C in the range of 18–26 °C can change a mosquito's life by more than a week (Jepson *et al*., 1947). A rise in the average global temperature in the range of 1.0–3.5 °C, increases the likelihood of many vector-borne diseases in new areas (Githeko *et al*., 2000).

Since the 1970s, Bikaner, has undergone much change in the landscape and demography due to the extensive spread of irrigation canal networks and urbanisation. Choudhary *et al*. (2009) reported increasing trends in minimum, maximum, and mean temperatures mainly during monsoon and post-monsoon seasons in the range of 0.1–0.7 °C/decade over Bikaner, a malaria-prone area. The reported increases in temperature in Bikaner provided the encouragement to determine if there was a relationship between incidences of malaria cases and global warming. Long-term changes in CO_2 in the atmosphere, changes in local land-use patterns (like more areas being brought under intensive canal irrigation), and the rise in urbanisation may be responsible for the long-term rise in temperature over Bikaner. Little information is available on the trends in malaria and the trends in climatic parameters, i.e. temperature, rainfall, relative humidity, etc. in the Bikaner region. The objective of this study, therefore, was to investigate trends in the incidences of malaria (*Plasmodium vivax* and *Plasmodium falciparum*) and the climatic parameters (temperature, rainfall, rainy days, and relative humidity) at annual, seasonal (winter, pre-monsoon, monsoon, and post-monsoon) and monthly (January–December) time scales over Bikaner.

2. Material and methods

2.1. Hydrometeorological and malaria data

Bikaner (longitude 71°54′E 74°12′E, latitude 27°11′N 29°3′N, and altitude 237 m a.m.s.l.) is situated in the northwestern part of Rajasthan, India, near the Pakistan border. Figure 1 shows the location of Bikaner. The temperature in Bikaner and in the Thar Desert ranges from a minimum of 2 °C in winter to a maximum of 48 °C in summer. The region is characterized by hot summer, cold winter, and with a low annual rainfall of 250 mm (Kochar *et al*., 2007; Choudhary *et al*., 2009). The Bikaner area has always been regarded as a hypoendemic area for malaria (Kochar *et al*., 2006). The monthly malaria epidemiological data, cases of *P. vivax* and *P. falciparum*, were obtained from the Bikaner Malaria Centre, Government of Rajasthan (India), for 34 years from 1975 to 2008. Monthly climatic data of various meteorological variables, namely, rainfall (*Rain*), rainy days (*Rdays*), minimum temperature (*Tmin*), maximum temperature (*Tmax*), morning relative humidity (*RHmor*), afternoon relative humidity (*RHeve*), mean relative humidity (*RHmn*), and mean wind speed (*MWS*) for Bikaner were obtained from the India Meteorological Department (Pune) for the similar duration of 34 years from 1975 to 2008. The values of *DTR* were obtained by subtracting the minimum temperature from the maximum temperature on different time scales at Bikaner. Monthly datasets were used to compute seasonal and annual time series, and the four seasons were defined as: winter (January–February), pre-monsoon (March–June); monsoon (July–September), and post-monsoon (October–December). It may be noted that the seasonal and annual time series were derived from the monthly datasets by averaging over the said seasons and years, respectively.

2.2. Trends in malaria incidences and climatic variables

Trends in data can be identified by using parametric or nonparametric methods. Nonparametric tests are more suitable for non-normally distributed, censored data, missing values, and are less influenced by the presence of outliers in the data. Recently, Kahya and Kalayci (2004), Partal and Kahya (2006), Jhajharia *et al*. (2009), Chattopadhyay and Chattopadhyay (2010), Chattopadhyay *et al*. (2011), and Jhajharia and Singh (2011) used the nonparametric Mann-Kendall (MK) test (Mann, 1945; Kendall, 1975) for trend analysis of various hydrometeorological variables. One of the main problems in testing and interpreting trends is the effect of serial dependence. If there is a positive correlation (persistence) in the time series, then the nonparametric test will suggest a more significant trend in the time series than

Figure 1. Map showing the location of Bikaner situated in Rajasthan in northwest India. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

specified by the significance level (Zhang *et al*., 2001). First, the significance of lag-1 serial correlation (r_1) for all *P. vivax*/*P. falciparum* malaria incidences and hydrometeorological time series was tested to eliminate the effect of serial correlation. If the absolute value of r_1 was less than the significance level then the time series was subjected to the original MK test. Otherwise, pre-whitening was carried out before applying the MK test in order to eliminate the influence of serial correlation. Readers can refer to Kahya and Kalayci (2004), Kumar *et al*. (2009), Dinpashoh *et al*. (2011) and Jhajharia *et al*. (2011) for further details of pre-whitening and the detailed formulae of the MK test.

2.2.1. *Theil-Sen estimator*

A bootstrap-based test for linear trend is the bootstrap slope test, based on the Theil-Sen estimator of the slope (Bonaccorso *et al*, 2005). The slope of n pairs of data points, x, was estimated using Theil-Sen's estimator given as

$$\beta = Median\left(\frac{x_j - x_l}{j - l}\right) \quad \forall\ 1 < l < j \tag{1}$$

where, β is the Theil-Sen estimator (Theil, 1950; Sen, 1968). The slope computed by Theil-Sen's estimator is a robust estimate of the magnitude of a trend. Yue *et al*. (2002), Dinpashoh *et al*. (2011) and Jhajharia *et al*. (2011) have used the Theil-Sen estimator in identifying the slope of a trend line in the hydrological time series.

2.3. Principal component analysis

The objective of principal component analysis (PCA) is to reduce the number of predictive variables and transform them into new variables, called principal components (PC). These new variables are independent linear combinations of the original data and retain the maximum possible variance of the original set (Webster, 2001; Sousa *et al*., 2007). PCA is also useful since it yields a smaller set of variables which makes the model easier to interpret. To better elucidate the influence of each original variable in the PCs, a rotational algorithm, such as varimax rotation, is generally applied to obtain the rotated factor loadings that represent the contribution of each variable to a specific PC. The applicability of PCA to the datasets under study was verified by applying the modified Bartlett's sphericity test, which assumes as null hypothesis that the correlation matrix is an identity matrix. The alternative hypothesis assumes the negation of the null hypothesis. Under the assumption that the null hypothesis is true, the following chi-square statistic was constructed (Sousa *et al*., 2007):

$$\chi_k{}^2 = \left[n - k - \frac{2(p-k) + 7 + 2/(p-k)}{6}\right] + \sum_{j=1}^{k}\left(\frac{\overline{\lambda}}{\lambda_j - \overline{\lambda}}\right)^2 \times \left[-\ln \Pi_{j=k+1}^{p}\lambda_j + (p-k)\ln\overline{\lambda}\right] \tag{2}$$

where p is the number of components, λ_j represents the eigenvalue for the k th component, n is the number of observations in the sample, and $\overline{\lambda}$ was obtained by the following equation:

$$\overline{\lambda} = \sum_{j=k+1}^{p}\frac{\lambda_j}{p-k}. \tag{3}$$

Implementation procedure and relevance of PCA to the problem under consideration would be detailed in Section 3.3 of the present paper. For a thorough theoretical discussion on PCA, the readers may refer the work of Jolliffe (2002). Some significant applications of PCA in climate research include Yu *et al*. (1997), Monahan (2000), Hsieh (2001), Cannon *et al*. (2002), Camberlin and Diop (2003), Giannini *et al*. (2003), Astel *et al*. (2004), and Muñoz-Díaz and Rodrigo (2004).

2.4. Artificial neural networks

Artificial neural networks (ANN) are highly potent in generating predictive models because of the following abilities of ANN (Shiva Nagendra and Khare, 2006): (1) An ANN requires a known input dataset without any assumptions, (2) It exhibits rapid information processing and is able to develop a mapping of input and output variables, and (3) A multilayer neural network can approximate any smooth, measurable function between input and output vectors by selecting a suitable set of connecting weights and transfer functions. Shiva Nagendra and Khare (2006) summarized the steps for implementing an ANN: (1) selection of the model architecture, (2) selection of the best activation function, (3) selection of the optimum learning rate and momentum rate, (4) initialisation of the network weights, (5) training and testing of the model, and (6) evaluation of the model.

A detailed theoretical description of ANN is available in Rojas (1996) and Haykin (2001). Some remarkable applications to meteorological studies include Gardner and Dorling (1998), Maier and Dandy (2000), Ramirez *et al*. (2005), and Chattopadhyay *et al*. (2011). The present paper applied a multilayer perceptron (MLP) ANN. Its suitability in atmospheric modelling is well established by Gardner and Dorling (1998) and Hsieh and Tang (1998). In MLP, each network consists of several simple processors called neurons, or cells, which are highly interconnected and arranged in several layers. There are three basic types of layers: input layer, hidden layer(s), and output layer. Input and output layers are connected through a hidden layer(s) (Gardner and Dorling, 1998). There may be one to several hidden layers between input and output layers. The number of neurons in the hidden layer is obtained by training several networks and estimating the corresponding errors on the test dataset. A few neurons in the hidden layer produce high training and testing errors due to under-fitting and statistical bias (Shiva Nagendra and Khare, 2006). The type of nonlinearity in the ANN depends upon the type of activation function chosen to train the ANN. Gardner

and Dorling (1998) have shown that the sigmoid activation function ($f(x) = (1 + e^{-x})^{-1}$) performs better than other types of activation functions.

The relationship between incidences of malaria and climatic parameters, like temperature, rainfall, and wind speed, was also explored using an ANN-a mathematical analog of the central nervous system. In the present study, land use changes in an area are not considered because of the non-availability of the land use databases over the desert site with monthly temporal resolution. We have tried to observe the responses of malaria transmission through the changes in climatic variables in the arid region of the Thar Desert. Over the last few decades, ANNs have received significant attention in the health science area (e.g. Baxt, 1990; Tourassi, 1997; Ciampi and Zhang, 2002). In order to provide reliable data for the development of preventive strategies, a meteorological factor-based predicative model for malaria forecast was investigated by Gao *et al*. (2003) who reported that the ANN might be used as a new tool for malaria forecasting based on meteorological parameters, like air temperature, relative humidity, monthly maximum air temperature, minimum air temperature, rainfall, rainy day, average monthly pressure, evaporation, and sunshine hours. Seman *et al*. (2008) applied a MLP ANN for classifying the malaria parasite into three species, namely, *P. Falciparum, P. Vivax*, and *P. Malariae*, and reported the potential of back-propagation learning over the Levenberg-Marquardt and Bayesian Rule algorithms. Jayavanth and Singh (2003) implemented an ANN based on feed forward-error back-propagation for malaria severity analysis using aggregate sedimentation velocity (ASV), effective number of cells (ENC), process completion time (PCT), and mean filtration time (MFT) of blood samples as a measure of severity of malaria. Boulle *et al*. (2001) revealed the supremacy of ANN over logistic regression in classifying the cause of death from verbal autopsy and found that ANN produced less sensitivity in the case of malaria than in the cases of tuberculosis, AIDS, meningitis, cardiovascular disorders, diarrhea, and tetanus.

3. Results and discussion

Trends in incidences of malaria, both *P. vivax* and *P. falciparum*, and in climatic variables, i.e. *Tmin*, *Tmax*, relative humidity, and rainfall were obtained through the Mann-Kendall nonparametric test in Bikaner for different durations: annual; seasonal: winter, pre-monsoon, monsoon, and post-monsoon; and monthly: January–December. The values of test statistics obtained through the MK test for the incidence cases of *P. vivax* malaria and *P. falciparum* malaria are shown in Table I. Table I also shows the magnitudes of trends of *P. vivax* malaria and *P. falciparum* malaria in the number of incidences/decade obtained by using the nonparametric Thiel-Sen's test at annual, seasonal, and monthly time scales. Table II also shows the magnitudes of trends of various climatic parameters, i.e. temperatures, rainfall, relative humidity, etc., obtained by using the Thiel-Sen's test at different time scales over Bikaner site using the

Table I. Trends in cases of malaria incidence obtained through Mann-Kendall test and Thiel-Sen's test at Bikaner located in the Thar Desert region of Rajasthan, India (1975–2008), where PF and PV denote *Plasmodium falciparum* and *Plasmodium vivax*, respectively.

	Plasmodium falciparum		*Plasmodium vivax*	
Different Time scales	No. of cases/decade	Values of Z statistic obtained through M-K test	No. of cases/decade	Values of Z statistic obtained through M-K test
JANUARY	0	−0.07	−3	−0.57
FEBRUARY	0	−0.54	−5	−1.07
MARCH	0	−0.28	−3	−0.39
APRIL	1[a]	1.78	−6	−0.70
MAY	2	1.52	−5	−0.26
JUNE	0	0.59	−12	−0.86
JULY	0	0.99	−8	−0.50
AUGUST	1	0.95	−6	−0.11
SEPTEMBER	7[a]	1.72	40	0.56
OCTOBER	4	1.13	48	1.19
NOVEMBER	3	1.41	14	0.83
DECEMBER	0	−0.10	4	0.77
ANNUAL	25	1.27	471	1.10
WINTER	0	−0.03	−9	−1.04
PRE-MONSOON	3	1.20	−20	−0.36
MONSOON	10[a]	1.90	40	0.33
POST-MONSOON	6	0.76	68	1.05

[a] Value denotes the statistically significant trend at 10% level of significance obtained through the Mann-Kendall test after carrying out the pre-whitening.

Table II. Magnitude of trends obtained through the Thiel-Sen's test for various climatic parameters over Bikaner (1975–2008), where T_{max}, T_{min}, *Rdays*, *Rain*, RH_{mn}, RH_{mor}, RH_{eve} and *DTR* denote maximum temperature, minimum temperature, rainy days, rainfall, mean, morning and evening relative humidity and diurnal temperature range, respectively.

	T_{max}	T_{min}	*Rdays*	*Rain*	RH_{mn}	RH_{mor}	RH_{eve}	*DTR*
	°C/dec	°C/dec	No/dec	mm/dec	%/dec	%/dec	%/dec	°C/dec
JANUARY	0.04	0.39[a]	0.00	−0.22	1.15	1.11	0.95	**−0.79**
FEBRUARY	0.50	**0.86**	0.00	1.65[a]	2.50	2.22	2.11	−0.64[a]
MARCH	0.56[a]	**0.74**	0.00	−0.29	0.00	0.00	0.38	−0.46[a]
APRIL	0.30	0.50[a]	0.00	0.00	0.32	0.32	0.00	−0.36[a]
MAY	−0.11	0.25	0.00	0.00	0.88	1.10	0.91	**−0.38**
JUNE	−0.27	0.05	0.00	−4.40	0.96[a]	**1.79**	0.49	−0.27
JULY	**0.59**	0.37[a]	**−1.54**	**−26.92**	−1.25	−0.91	−2.00	0.20[a]
AUGUST	0.33	0.22[a]	**−1.11**	−11.9[a]	−1.25	−0.70	−2.03	0.29[a]
SEPTEMBER	0.37	**0.48**	−0.59[a]	−1.75	0.96	1.09	1.10	−0.04
OCTOBER	0.00	0.45[a]	0.00	0.00	0.37	0.00	0.00	−0.21
NOVEMBER	0.25	**0.93**	0.00	0.00	−1.67	−1.98	−1.76[a]	−0.20
DECEMBER	−0.14	**0.85**	0.00	0.00	0.00	0.00	0.00	**−0.91**
ANNUAL	0.23[a]	**0.47**	**−3.64**	**−58.0**	−0.04	0.24	−0.02	**−0.31**
WINTER	0.18	**0.80**	0.00	2.22	1.43	1.67	1.61	**−0.60**
PRE-MONSON	0.19	**0.47**	−0.77	−8.48	0.78	1.09	0.50	**−0.36**
MONSOON	**0.50**	**0.42**	**−3.48**	**−54.0**	−1.48	−0.91	−1.85	0.13
POST-MONSN	0.08	**0.87**	0.00	−0.11	−0.24	−0.48	0.00	**−0.63**

[a] Values, and bold, denote the magnitude of statistically significant trend at 10 and 5% level of significance, respectively, obtained through the Mann-Kendall test after carrying out the pre-whitening.

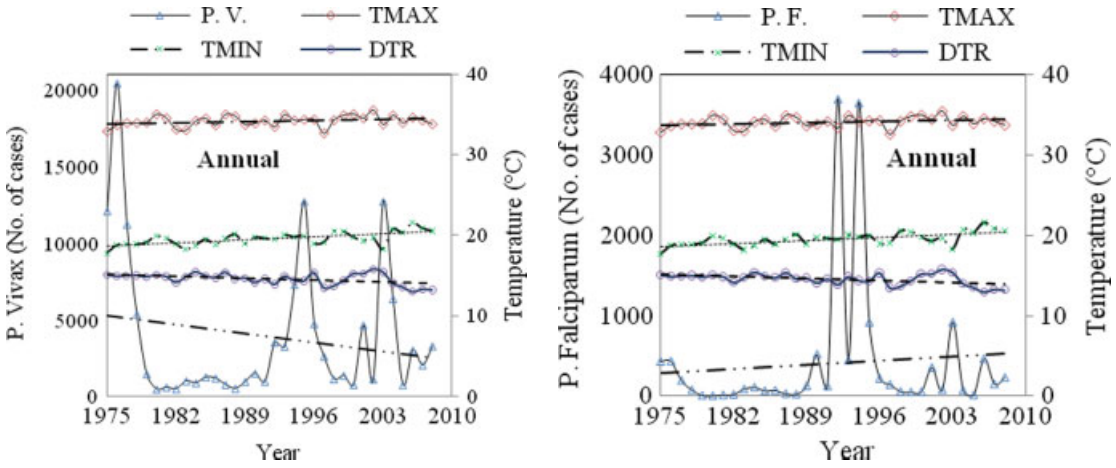

Figure 2. Time series of *P. falciparum* malaria, *P. vivax* malaria and temperature over Bikaner corresponding to the annual duration. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

data from 1975 to 2008. A brief summary of results follows.

3.1. Trends in malaria

Cases of malaria incidence were witnessed throughout the year with marked seasonal variation over Bikaner. Mainly two *Plasmodium* species are encountered, namely, *P. vivax* and *P. falciparum* in the Thar Desert of Rajasthan. It is generally observed that incidences of malaria are likely to increase (decrease) in the post-rainy season (the winter season) in Bikaner. The population of this region has increased from nearly 0.6 million in 1975 to almost 2.4 million in 2006 (Kochar *et al*., 2007). In the context of the present study, it is worthwhile mentioning that no relationships are found between the population of Bikaner (Rajasthan) and cases of incidence of *P. vivax* malaria and *P. falciparum* malaria as indicated by the small Pearson correlations −0.013 and 0.051, respectively, in the Thar Desert region. Figures 2, 3, and 4 show the time series of *P. falciparum* malaria, *P. vivax* malaria and maximum temperature, minimum temperature, and *DTR* corresponding to annual duration, different seasons (winter, pre-monsoon, and monsoon), and monthly times scales (April and September) of Bikaner. The linear trend lines of various time series are shown in Figures 2–4 as dashed lines. It is interesting to note that cases of *P. falciparum* malaria incidence have witnessed an increase in the

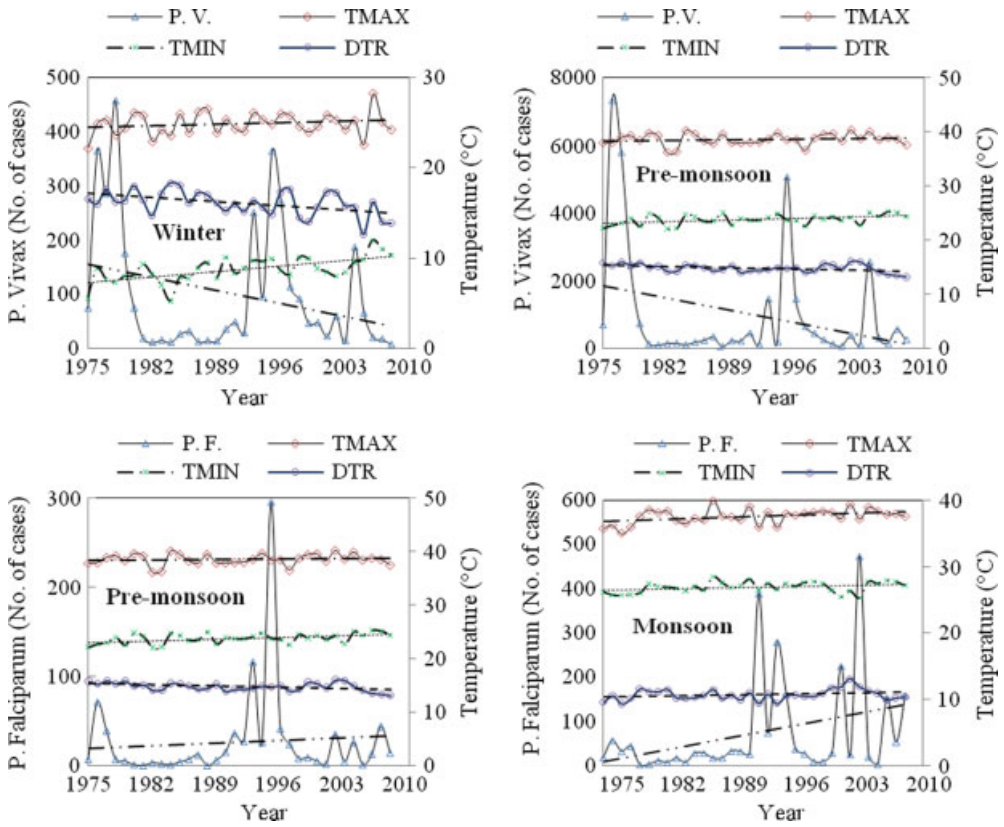

Figure 3. Time series of *P. falciparum* malaria, *P. vivax* malaria and temperature over Bikaner in different seasons. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

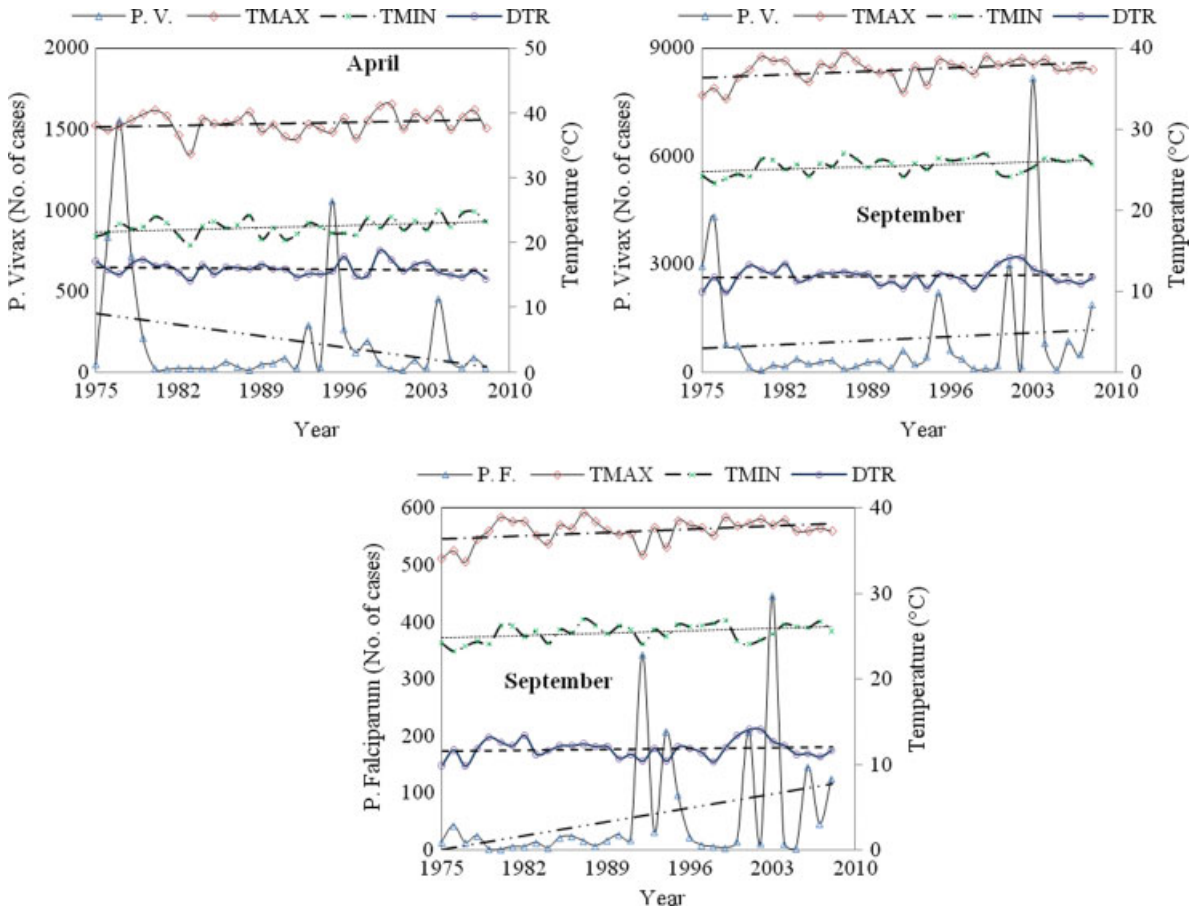

Figure 4. Time series of *P. falciparum* malaria, *P. vivax* malaria, and temperature over Bikaner in April and September. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

number of cases in all the time series of *P. falciparum* malaria over Bikaner (Figures 2–4).

Statistically significant increasing trends in *P. falciparum* malaria incidence cases were witnessed by using the MK test after applying the pre-whitening in order to eliminate the influence of serial correlation from the time series of malaria incidences in the months of April and September, and in the monsoon season at Bikaner located in the Thar Desert region of India. However, no trends in *P. falciparum* malaria incidences were observed at any of the remaining time scales, i.e. annual duration; monthly: all 12 but 2; and seasonal: winter, pre-monsoon and post-monsoon. On the other hand, no trends in the incidence cases of *P. vivax* malaria were observed at any of the time scales in Bikaner. However, it is interesting to note that most of the trends in the incidences of *P. vivax* malaria are found to be increasing, although statistically non-significant even at a 10% level of significance over this desert site of northwestern Rajasthan.

Dev and Dash (2007) reported that the focal outbreaks of *P. Falciparum* malaria frequently occur, particularly during the rainy season (April–September) in northeast India. The results of the present study, i.e. increases in *P. Falciparum* malaria incidence cases in the month of April and September at Bikaner, support the findings of Dev and Dash (2007). Dev and Dash (2007) mentioned that any rise in the minimum temperature caused by global warming might increase the incidence of *P. falciparum* malaria substantially. In the present study, we have witnessed significant increasing trends in minimum temperature in the months of April and September (in monsoon season) at the rate of about 0.5 and 0.48 °C/decade (about 0.42 °C/decade), respectively, over the Thar Desert (Tables I and II). It is interesting to note that cases of *P. falciparum* malaria incidence have increased significantly corresponding to the months of April and September (monsoon season) over the Thar Desert along with the occurrence of the nighttime warming in these two months (in this rainy season) in Bikaner. Bhattacharya *et al*. (2006) also indicate the dominant role of temperature and relative humidity in the transmission and growth of malaria under the climate change conditions over India. However, Dev (2009) reported no significant association between absolute rainfall and inter-annual variation in cases of malaria in Assam, northeast India. Lindsay and Martens (1998) have mentioned that a rise in minimum temperature would tend to extend the duration of each malaria-transmission season, amplify entomological inoculation rates, and enlarge the geographical distribution of malaria to regions previously at low risk for malaria, such as the highland areas of Africa.

3.2. Trends in climatic parameters

Statistically significant increasing trends, mostly at a 5% significance level, were observed in T_{min} corresponding to different time scales: annual; seasonal: all the four seasons; and monthly: January–April and September–December in the range of 0.22–0.93 °C/decade (Table II), and no trends in the remaining 4 months. However, no trends in T_{max} were observed at any of the time scales except annual, monsoon season, and the months of March and July in the range of 0.23–0.56 °C/decade. Decreases in rainy days and rainfall were observed in the range of 1–4 days/decade and 12–58 mm/decade in the annual duration; the monsoon season; and the months of July–August. However, no trends were observed at almost any of the time scales in the morning and afternoon relative humidity, and the mean relative humidity at Bikaner.

DTR decreases were observed at most of the time scales: annual; seasonal: all the four seasons except monsoon; and monthly: January–May, and December, caused mainly due to the increases witnessed in *Tmin*. On the other hand, the *DTR* increase was also observed in the months of July and August. It is interesting to note that most of the *DTR* decreases (increase) are (is) caused due to comparatively stronger *Tmin* (*Tmax*) increases than *Tmax* (*Tmin*) in Bikaner. Concomitant increases in maximum temperature and minimum temperature, and decreases in rainfall and rainy days observed in September and monsoon season are serious causes for concern for the already stressed agriculture in the arid region of northwest India. Also, increase in temperature observed in Bikaner is a serious concern for the spread of malaria because temperature, being one of the main climatic parameter, plays a vital role in the development of malaria parasites and the survival of the mosquito vector species that contribute to the intensity of malaria transmission in areas, like the Thar Desert region located in northwest India.

3.3. Implementation of PCA

First, we considered ten variables, including *P. falciparum* malaria, *Tmax*, *Tmin*, *MWS*, *Rdays, Rain, RHmean, RHmor, RHeve*, and *DTR* at the annual time scale. The correlation matrix obtained for annual *P. falciparum* malaria incidences is given in Table III. The null hypothesis that the correlation matrix was an identity matrix was tested against the alternative hypothesis that assumed negation. The Bartlett sphericity test showed that under the null hypothesis, the value of χ^2 was 609.454 with 45° of freedom. Thus, the null hypothesis was rejected. Simultaneously, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy turned out to be 0.61, which was greater than 0.5. The KMO statistic reflects the degree to which it is likely that common factors explain the observed correlations among the variables and is calculated as the sum of the squared simple correlations between pairs of variables divided by the sum of squared simple correlations plus the sum of squared partial correlations (Dilalla and Dollinger, 2006). The small value of KMO indicates that correlations between pairs of variables cannot be accounted for by common factors. A high KMO (maximum 1.0, minimum acceptable 0.5) indicates that data are likely to factor well since correlations between pairs of variables can be explained

Table III. Correlation matrix obtained for the annual PF cases. The upper triangle displays the correlation values and the lower triangle displays the corresponding p-values.

	PFannual	*Tmax*	*Tmin*	*MWS*	*Rdays*	*Rain*	*RHmean*	*RHmor*	*RHeve*	*DTR*
PFannual	1.000	−0.028	0.003	−0.227	−0.182	−0.196	0.057	0.076	0.023	−0.041
Tmax	0.875	1.000	0.702	0.432	−0.620	−0.626	−0.575	−0.576	−0.574	0.482
Tmin	0.987	0.000	1.000	0.217	−0.417	−0.419	−0.320	−0.255	−0.353	−0.286
MWS	0.197	0.011	0.218	1.000	−0.327	−0.363	−0.226	−0.215	−0.214	0.313
Rdays	0.303	0.000	0.014	0.059	1.000	0.875	0.680	0.667	0.636	−0.321
Rain	0.267	0.000	0.014	0.035	0.000	1.000	0.689	0.698	0.640	−0.327
RHmean	0.749	0.000	0.065	0.199	0.000	0.000	1.000	0.968	0.957	−0.379
RHmor	0.669	0.000	0.146	0.222	0.000	0.000	0.000	1.000	0.870	−0.461
RHeve	0.897	0.000	0.041	0.224	0.000	0.000	0.000	0.000	1.000	−0.337
DTR	0.818	0.004	0.101	0.072	0.064	0.059	0.027	0.006	0.051	1.000

by the other variables (Leung *et al*, 2005). Thus, the existence of collinearity within the dataset was established and the necessity for PCA was felt. Table IV gives results of the eigen values and the variances explained by principle components extracted using the variables related to *P. falciparum* malaria. It can be seen that all the variance in the data was represented by 10 PC. However, approximately 88% of the total variance in the data can be represented by the first 4 PCs and for all of the four PCs the eigen values were greater than 1. The individual contributions of the remaining 6 PCs were small and their total contribution was only 12% of the total variance. The eigen vectors with the largest eigen values represented the dimensions with the strongest correlation in the dataset. This dataset showed that the highest correlations were also in the first four PCs. Thus, the first four PCs were selected for the PC matrix. The loadings, which indicated the influence of variables in these four PCs, are presented in Table V. A loading close to 1 indicates a very strong correlation. In Table V, the new components showed how the variables were separated between the columns according to their characteristics or to the properties that they represented. This table shows that the first rotated PC was mostly related to the *RHmean*, *RHmor*, and *RHeve*. The highest loading was observed for *RHmean* (0.980), which was followed by *RHmor* (0.957) and *RHeve* (0.940). The second rotated PC was mostly related to *MWS* (0.864). The third rotated PC was mostly related to *Tmin* (0.875). The fourth rotated PC was related to *P. falciparum*$_{\text{annual}}$ (0.947). Therefore, the six variables extracted by PCA were *P. falciparum*$_{\text{annual}}$, *Tmin, MWS, RHeve, RHmor* and *RHmean* T_{min}.

Table IV. Eigen values and variances explained by principle components extracted using the variables related to PF under study at annual scale.

	Initial eigen values		
Component	Total	% of variance	Cumulative %
1	5.156	51.559	51.559
2	1.387	13.872	65.431
3	1.250	12.499	77.929
4	1.016	10.161	88.091
5	0.580	5.797	93.887
6	0.377	3.766	97.653
7	0.134	1.342	98.995
8	0.097	0.966	99.961
9	0.004	0.039	100.000
10	0.000	0.000	100.000

Table V. Factor loadings of the variables after rotating the principal component matrix using varimax with Kaiser normalisation.

	Component			
	1	2	3	4
PFannual	0.029	−0.138	−0.004	0.947
Tmax	−0.532	0.671	0.273	0.068
Tmin	−0.283	0.333	0.875	0.044
MWS	−0.060	0.864	−0.017	−0.180
Rdays	0.689	−0.416	−0.139	−0.373
Rain	0.694	−0.435	−0.131	−0.376
RHmean	0.980	−0.116	−0.014	0.051
RHmor	0.957	−0.146	0.075	0.048
RHeve	0.940	−0.102	−0.064	0.039
DTR	−0.368	0.493	−0.709	0.038

The study described just above had been carried out at the annual time scale. A similar approach was now adopted at the seasonal time scale as well, and the factor loadings for the rotated PCs are listed in Table VI. It was observed that in all the four seasons, four PCs had been extracted considering their contributions to the explained variance. In the case of winter season, the correlation matrix failed to be positive and consequently Bartlett's sphericity test and Kaiser-Meyer-Olkin Measure were not possible. Therefore, the suitability of PCA in this case was not well understood. In all the remaining three seasons, i.e. pre-monsoon, monsoon and post-monsoon, Kaiser-Meyer-Olkin Measure of Sampling Adequacy came out to be greater than 0.5 and the chi-square obtained for Bartlett's sphericity test was found to exceed the tabular value. Thus, PCA was applied to pre-monsoon, monsoon, and post-monsoon seasons. However, we applied PCA to winter also, although its

Table VI. Factor loadings of variables after rotating the principal component matrix using varimax with Kaiser Normalisation. The variables related to PF are at the seasonal time scale.

	PC (Winter)				PC (Pre-monsoon)			
	1	2	3	4	1	2	3	4
PF	0.120	0.073	0.348	0.638	0.029	−0.138	−0.004	**0.947**
Tmax	−0.447	−0.285	0.709	0.255	−0.532	0.671	0.273	0.068
Tmin	0.240	0.163	**0.926**	−0.143	−0.283	0.333	**0.875**	0.044
MWS	0.143	−0.050	0.241	**−0.816**	−0.060	**0.864**	−0.017	−0.180
Rdays	0.315	**0.820**	−0.137	0.131	0.689	−0.416	−0.139	−0.373
Rain	0.167	**0.914**	0.130	0.013	0.694	−0.435	−0.131	−0.376
RHmean	**0.970**	0.193	0.011	−0.003	**0.980**	−0.116	−0.014	0.051
RHmor	**0.951**	0.164	0.003	−0.035	**0.957**	−0.146	0.075	0.048
RHeve	**0.942**	0.217	0.018	0.035	**0.940**	−0.102	−0.064	0.039
DTR	−0.659	−0.431	−0.310	0.383	−0.368	0.493	−0.709	0.038
	PC (Post-monsoon)				PC (Monsoon)			
	1	2	3	4	1	2	3	4
PF	0.060	−0.068	−0.045	**0.989**	0.192	−0.011	0.135	**0.937**
Tmax	−0.623	0.062	−0.561	0.014	−0.765	0.371	−0.272	−0.188
Tmin	0.072	**0.937**	−0.114	−0.009	−0.420	−0.024	**−0.842**	−0.223
MWS	−0.167	**0.881**	0.008	−0.083	−0.114	**0.950**	0.021	−0.019
Rdays	0.497	0.210	0.631	−0.176	0.747	−0.436	0.313	−0.198
Rain	0.032	−0.054	**0.912**	0.017	0.773	−0.341	0.276	−0.151
RHmean	**0.980**	0.046	0.124	0.040	**0.940**	−0.062	0.126	0.252
RHmor	**0.966**	0.107	0.172	0.014	**0.898**	0.000	0.158	0.290
RHeve	**0.964**	−0.033	0.059	0.070	**0.944**	−0.107	0.098	0.216
DTR	−0.503	−0.747	−0.303	0.016	−0.630	0.528	0.457	−0.037

suitability in this season was not tested. The results are listed in Table IV, where the mostly related variables to each factor are displayed in bold. In the case of winter, PCA removed the *P. falciparum* malaria as it was not highly related to any of the PCs. In future ANN modelling of *P. falciparum* malaria, we therefore may not consider winter for the development of predictive model for the number of malaria cases.

In the next phase, a similar approach was adopted for *P. Vivax* cases for annual and seasonal time scales. In the annual case, the correlation matrix for *P. Vivax* cases is given in Table VII. Clearly, significant correlations existed among predictors. The value of χ^2 was 981.878 with 45° of freedom. Thus, the null hypothesis was rejected. Simultaneously, the Kaiser-Meyer-Olkin Measure of Sampling Adequacy turned out to be 0.516, which was greater than 0.5. Thus, the existence of collinearity within the dataset was established and the necessity for PCA arose. Table VIII, which lists the eigen values, shows that the first four PCs had very high eigen values and they explained 92.125% of the variance. The rotated component matrix presented in Table IX shows that *RHmean* (0.983), *RHmor* (0.963) and *RHeve* (0.917) were mostly related to the first PC. The second component was mostly related to *DTR* (−0.866), the third component was mostly related to *Rain* (0.914), *Rdays* (0.912),

Table VII. Correlation matrix obtained for the annual PV cases. The upper triangle displays the correlation values and the lower triangle displays the corresponding *p*-values.

	PV	*Tmax*	*Tmin*	*MWS*	*Rdays*	*Rain*	*RHmean*	*RHmor*	*RHeve*	*DTR*
PV	1.000	0.176	0.159	0.144	0.348	0.301	0.147	0.013	0.318	−0.046
Tmax	0.319	1.000	0.983	0.470	0.439	0.416	−0.565	−0.639	−0.411	−0.606
Tmin	0.369	0.000	1.000	0.467	0.457	0.443	−0.483	−0.559	−0.333	−0.743
MWS	0.416	0.005	0.005	1.000	−0.076	−0.093	−0.353	−0.329	−0.352	−0.308
Rdays	0.044	0.009	0.007	0.669	1.000	0.892	0.150	0.022	0.309	−0.375
Rain	0.084	0.014	0.009	0.601	0.000	1.000	0.153	0.037	0.296	−0.397
RHmean	0.407	0.001	0.004	0.041	0.397	0.388	1.000	0.971	0.945	0.030
RHmor	0.000	0.000	0.001	0.057	0.902	0.835	0.000	1.000	0.839	0.088
RHeve	0.066	0.016	0.054	0.041	0.075	0.089	0.000	0.000	1.000	−0.054
DTR	0.796	0.000	0.000	0.076	0.029	0.020	0.866	0.621	0.762	1.000

Table VIII. Eigen values and variances explained by principle components extracted using the variables related to PV under study at the annual scale.

Component	Initial eigen values		
	Total	% of variance	Cumulative %
1	4.187	41.869	41.869
2	2.997	29.971	71.839
3	1.054	10.543	82.382
4	0.974	9.743	92.125
5	0.431	4.313	96.439
6	0.177	1.767	98.205
7	0.104	1.036	99.241
8	0.076	0.759	100.000
9	0.000	0.000	100.000
10	0.000	0.000	100.000

Table IX. Loadings of variables in the principle component matrix for the first four principle components related to PV under study at the annual scale.

	Component			
	1	2	3	4
PV	0.104	0.031	0.245	0.929
Tmax	−0.543	0.681	0.401	0.121
Tmin	−0.436	0.775	0.410	0.066
MWS	−0.251	0.695	−0.365	0.367
Rdays	0.088	0.180	0.912	0.173
Rain	0.098	0.184	0.914	0.112
RHmean	0.983	−0.133	0.087	0.039
RHmor	0.963	−0.155	−0.042	−0.062
RHeve	0.917	−0.089	0.255	0.174
DTR	−0.089	−0.866	−0.310	0.155

and the fourth component was mostly related to *P. vivax* (0.929). When considered at the seasonal scale (Table X), it was observed that the first 3 PCs explained 92.87% of the total variance in winter. From factor loadings of the rotated PCs, the variables extracted were *P. Vivax* (0.937), *Tmax* (0.914), *Tmin* (0.930), *Rdays* (0.915), *Rain* (0.937), *RHmean* (0.995), *RHmor* (0.955), and *RHeve* (0.972). In pre-monsoon, the first 3 components explained 82% of the total variance, and accordingly the variables extracted were *Tmin* (−0.911), *Tmax* (−0.767), *Rdays* (0.777), *DTR* (−0.883) and PV (0.970). For monsoon, the first 3 components explained 86.61% of the total variance and, accordingly, the variables extracted were *RHmean* (0.966), *RHmor* (0.958), *RHeve* (0.953), *DTR* (−0.823), *Tmax* (−0.846), *Rdays* (0.801), *Rain* (0.844), and PV (0.939). In the case of post-monsoon the extracted variables in this case were *Tmin* (0.792), *DTR* (−0.826), *RHmean* (0.985), *RHmor* (0.970), *RHeve* (0.964), *Rdays* (0.826), *Rain* (0.888), *MWS* (0.897) and PV (0.991). Variables extracted in this manner were used to develop the proposed ANN models explained in the next section.

Table X. Factor loadings of variables after rotating the principal component matrix using varimax with Kaiser normalisation. The variables are related to PV. Here the variables are considered at the seasonal scale.

	PC(Winter)			PC(Pre-monsoon)		
	1	2	3	1	2	3
PV	**0.937**	−0.072	0.192	0.129	0.018	0.970
Tmax	**0.914**	−0.275	0.148	−0.767	−0.443	−0.120
Tmin	0.930	−0.015	0.296	−0.911	0.051	−0.243
MWS	0.307	0.006	**0.940**	−0.166	−0.130	−0.056
Rdays	**0.915**	0.054	0.102	0.777	0.412	−0.070
Rain	**0.937**	−0.008	0.121	0.779	0.427	−0.034
RHmean	−0.082	**0.995**	0.014	0.679	0.669	0.191
RHmor	−0.210	**0.955**	0.017	0.642	0.710	0.138
RHeve	0.082	**0.972**	0.009	0.676	0.631	0.233
DTR	−0.528	−0.482	−0.436	−0.133	−0.883	0.119
	PC (Monsoon)			PC (Post-monsoon)		
	1	2	3	1	2	3
PV	0.193	0.179	0.939	0.022	0.079	0.060
Tmax	0.136	−0.846	−0.282	−0.597	0.257	0.549
Tmin	0.584	−0.690	−0.382	−0.041	0.792	0.567
MWS	−0.045	−0.286	0.034	−0.165	0.897	0.257
Rdays	0.398	0.801	0.051	0.145	0.472	0.826
Rain	0.414	0.844	−0.058	0.084	0.276	0.888
RHmean	0.966	0.148	0.160	0.985	0.065	0.066
RHmor	0.958	0.087	0.159	0.970	0.153	0.093
RHeve	0.953	0.195	0.158	0.964	−0.047	0.032
DTR	−0.823	0.042	0.275	−0.389	−0.826	−0.322

3.4. Development of ANN model

Prior to the development of an ANN model, we executed a cross-correlation analysis for the various pairs of parameters obtained from the entire dataset under consideration. In all cases, it was observed that there was no specific pattern in cross-correlations. This indicates that there was no stability in mutual relationships among the various pairs of parameters. This hinted at a highly complex and nonlinear association between predictors (i.e. meteorological parameters) and predictands (incidences of *P. vivax* malaria and *P. falciparum* malaria). Considering the complexity reflected by the cross-correlation analysis, we also explored the relationship between cases of malaria and changes in temperature and other climatic parameters using ANN. The purpose of generating an ANN model in the present study was to forecast the incidences of *P. vivax* and *P. falciparum* malaria for a given year $(t+1)$ using the meteorological parameters of the previous year (t) as predictors. Thus, in mathematical form it can be written as

$$PVorPF(t+1) = f(Tmax(t), Tmin(t), MWS(t), Rdays(t), Rain(t), RHmean(t), RHmor(t), RHeve(t), DTR(t)) \quad (4)$$

Table XI. Predictors extracted through PCA for various ANN models.

Type of Malaria	Period	Predictors	Predictand	Name of the model
P. Falciparum (PF)	Annual	*Tmin, MWS, RHeve, RHmor, RHmean.*	Number of PF cases	$ANNPF_{annual}$
	Winter	The correlation matrix is not positive definite		No ANN proposed
	Pre-monsoon	*RHeve, RHmor, RHmean, MWS, Tmin*	Number of PF cases	$ANNPF_{premonsoon}$
	Monsoon	*RHeve, RHmor, RHmean, MWS, Tmin*	Number of PF cases	$ANNPF_{monsoon}$
	Post-monsoon	*RHeve, RHmor, RHmean, MWS, Tmin, Rain*	Number of PF cases	$ANNPF_{postmonsoon}$
P. Vivax (PV)	Annual	*RHeve, RHmor, RHmean, DTR, Rain, Rdays*	Number of PV cases	$ANNPV_{annual}$
	Winter	*Tmax*, *Tmin*, *Rdays, Rain, RHmean, RHmor, RHeve*	Number of PV cases	$ANNPV_{winter}$
	Pre-monsoon	*Tmin*, *Tmax*, *Rdays, DTR*	Number of PV cases	$ANNPV_{premonsoon}$
	Monsoon	*RHmean, RHmor, RHeve, DTR, Tmax, Rdays, Rain*	Number of PV cases	$ANNPV_{monsoon}$
	Post-monsoon	*Tmin, DTR, RHmean, RHmor, RHeve, Rdays, Rain, MWS*	Number of PV cases	$ANNPV_{postmonsoon}$

The form of function f was determined by the type of activation function. If f is linear, then it becomes a multiple linear regression. In the present paper, single hidden-layer MLPs were generated for the predictive models with predictors selected by the PCA described earlier. For different scales (annual and seasonal), the MLP models were proposed based on the PCA results, as given in Table XI.

The 9 ANN models were trained and tested by dividing the relevant dataset in the ratio 7 : 3. After testing, the model was validated over the entire dataset. The minimisation of root mean squared error was chosen as the stopping criterion for the models which were run up to 3000 epochs. The batch mode of learning was adopted. The conjugate gradient (Rojas, 1996) learning rule was used to train MLPs as well. The models were assessed statistically using the Pearson Correlation Coefficient (PCC) between the actual and predicted values of *P. falciparum* malaria and *P. vivax* malaria incidences at annual and seasonal time scales. The PCC values for the various models are presented in Figures 5(a) and (b). From Figure 5(a), we find that the proposed ANN model can predict the number of *P. falciparum* malaria cases at the annual time scale most efficiently. This is indicated by the maximum value (0.85) of the PCC generated by the predictions from the ANN T the annual scale. The second highest PCC (0.76) pertains to the post-monsoon season. Thus, it may be stated that with the given set of climate variables, it is possible to predict the number of *P. falciparum* malaria cases using ANN in the post-monsoon season. However, despite removing the multicolinearity, the proposed ANN fails to predict the number of *P. falciparum* malaria cases in the other seasons over the study area. However for *P. vivax* malaria, ANN performs better than in the case of *P. falciparum* malaria. It is noted that for post-monsoon the ANN develops the best prediction indicated by highest PCC (0.87). In pre-monsoon, the PCC is 0.64 that indicates significant goodness of fit of the proposed model to the actual dataset. At annual time scale and in monsoon season, the PCCs are in the neighbourhood of 0.50. However, the prediction is not good enough in winter, when PCC is 0.31.

4. Conclusions

The present work investigates trends in malaria incidences (*P. falciparum* and *P. vivax*) at annual, monthly, and seasonal time scales in Bikaner by using the non-parametric Mann-Kendall test after removing the effect of significant lag-1 serial correlation by pre-whitening. Increasing trends are observed in *P. falciparum* malaria in the months of April and September and the monsoon season over Bikaner. However, no trends are observed in *P. vivax* malaria incidences corresponding to all the time scales in the Thar Desert region. Thus, the observed increases in *P. falciparum* malaria over Bikaner is a serious health issue because *P. falciparum* malaria is a life-threatening illness and may cause death if not treated well in time (desert areas usually have poor health infrastructure). Also, *P. falciparum* infects all stages of red blood cells and multiplies at much higher rates than *P. vivax* malaria. Thus, *P. falciparum* malaria spreads rapidly under suitable climatic conditions in a given locality. Since climate variables affect the transmission of malaria, trends in the maximum and minimum temperatures, *DTR*, rainfall, number of rainy days, and relative humidity are also examined. Concomitant increases in maximum temperature, minimum temperature, and decreases in rainfall and rainy days are observed together with the increasing trends in the incidence cases in *P. falciparum* malaria in the month of September and the monsoon season in Bikaner. Peng *et al*. (2003) report the highest positive

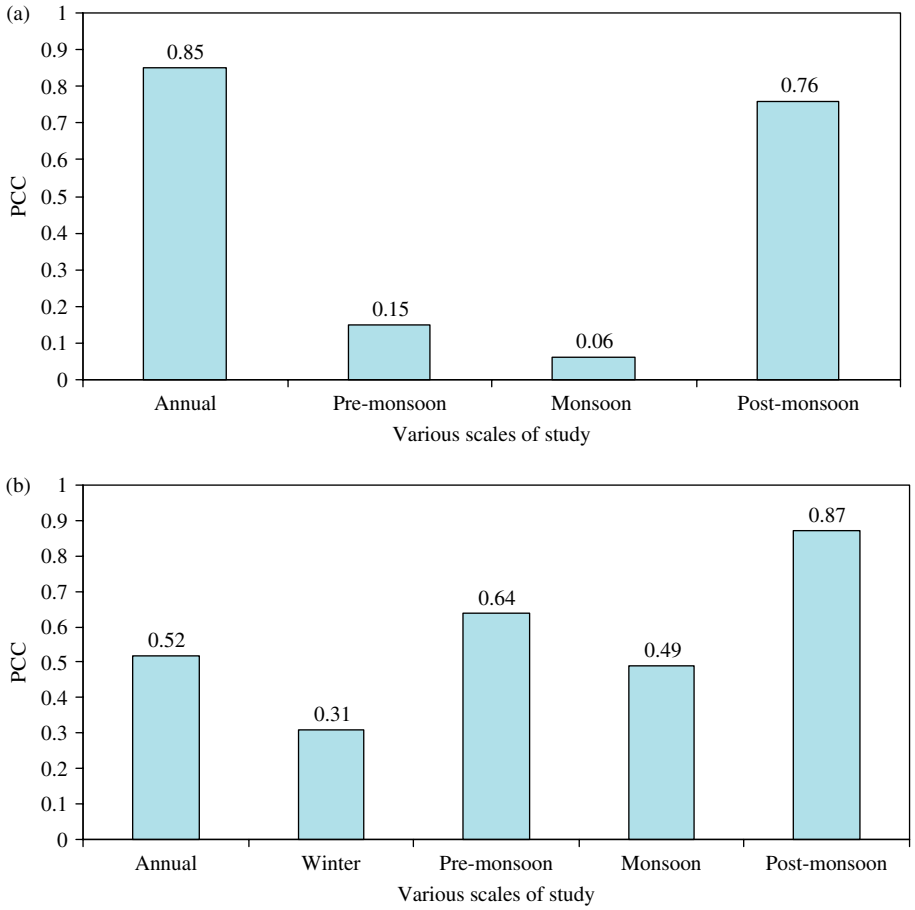

Figure 5. Pearson correlation coefficients (PCC) between predicted *P. falciparum* malaria (a), and *P. vivax* malaria (b) cases by ANN, and the actual number of cases at various scales of study. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

correlation between monthly incidence of malaria and monthly mean minimum temperature in Shuchen County, China. Devi and Jauhari (2006) also report that the monthly incidence of malaria is strongly correlated with T_{mean}, T_{min} and rainfall over Dehradun, India. However, no association between rainfall and annual incidences of malaria in Assam has been observed in recent years using the data of 1986–2003 (Dev and Dash, 2007). Results of this study, i.e. increases in PF in monsoon season under the observed warming witnessed over Bikaner, support the changes in the distribution of *P. falciparum* malaria predicted by Rogers and Randolph (2000), i.e. would mainly be restricted to the tropics and sub-tropics, and therefore, large-scale planning would be needed for adaptation measures for health and economic problems that may be caused due to the spread and the changing pattern of malaria, which is influenced by the anthropogenic induced climate changes over desert and other poor infrastructure sites of India and other developing countries.

After studying the trends, prediction models based on ANNs have been generated to examine the predictability of *P. falciparum* and *P. vivax* malaria. Prior to this model generation, the multicollinearity within the datasets under consideration is investigated by means of correlation matrix, the Bartlett sphericity test, and the Kaiser-Meyer-Olkin measure of sampling adequacy and existence of multi-collinearity and suitability of PCA for this study are strongly established. Subsequently, a variable selection method is done for ANN by means of computing factor loadings. Different neural network models are generated for various scales of study separately for *P. falciparum* and *P. vivax* malaria using the predictors extracted by PCS. The neural network models have been characterized by sigmoid nonlinearity in the hidden layer. Assessing the performance of the neural network by means of the Pearson correlation coefficient, it has been observed that it is possible to predict the number of cases of *P. falciparum* malaria using an ANN during the post-monsoon and at the annual scale. A better overall prediction by the ANN has been observed in the case of *P. vivax* than *P. falciparum* malaria. For *P. vivax* malaria the proposed neural network model has generated best predictions for pre-monsoon and post-monsoon seasons.

Acknowledgements

The data provided by the India Meteorological Department (Pune), and the Malaria Centre, Bikaner (Rajasthan) is gratefully acknowledged. We thank Mr Hanuman Godara and Mr Deepak Purohit, Dept of EI&CE,

E. C. B., Bikaner (Rajasthan) for their assistance in typing the data. The authors are thankful to the editor and the anonymous reviewers for their incisive comments which helped improve the quality of the manuscript.

References

Astel A, Mazerski J, Polkowska Z, Namienik J. 2004. Application of PCA and time series analysis in studies of precipitation in Tricity (Poland). *Advances in Environmental Research* **8**: 337–349.

Baxt WG. 1990. Use of an Artificial Neural Network for Data Analysis in Clinical Decision-Making: The Diagnosis of Acute Coronary Occlusion. *Neural Computation* **2**(4): 480–489.

Bhattacharya S, Sharma C, Dhiman RC, Mitra AP. 2006. Climate change and malaria in India. *Current Science* **90**(3): 369–375.

Bhutiyani MR, Kale VS, Pawar NJ. 2007. Long-term trends in maximum, minimum and mean annual air temperatures across the Northwestern Himalaya during the twentieth century. *Climatic Change* **85**: 159–177.

Bonaccorso B, Cancelliere A, Rossi G. 2005. Detecting trends of extreme rainfall series in Sicily. *Advances in Geosciences* **2**: 7–11.

Boulle A, Chandramohana D, Wellerb P. 2001. A case study of using artificial neural networks for classifying cause of death from verbal autopsy. *International Journal of Epidemiology* **30**: 515–520.

Camberlin P, Diop M. 2003. Application of daily rainfall principal component analysis to the assessment of the rainy season characteristics in Senegal. *Climate Research* **23**: 159–169.

Cannon AJ, Whitfield PH, Lord ER. 2002. Synoptic Map-Pattern Classification Using Recursive Partitioning and Principal Component Analysis. *Monthly Weather Review* **130**: 1187–1206.

Chattopadhyay S, Chattopadhyay G. 2010. Univariate modelling of summer-monsoon rainfall time series: Comparison between ARIMA and ARNN. *Comptes Rendus Geoscience* **342**: 100–107.

Chattopadhyay S, Jhajharia D, Chattopadhyay G. 2011. Univariate modelling of monthly maximum temperature time series over northeast India: neural network versus Yule–Walker equation based approach. *Meteorological Applications* **18**: 70–82.

Choudhary RR, Jhajharia D, Lal M, Jain JK, Lunayach A, Choudhary MK. 2009. Climate and its variations over Bikaner since 1951–2008. *Journal Indian Geological Congress* **1**(2): 79–86.

Ciampi A, Zhang F. 2002. A new approach to training back-propagation artificial neural networks: empirical evaluation on ten data sets from clinical studies. *Statistics in Medicine* **21**: 1309–1330.

Dev V. 2009. Climatic determinants and malaria transmission in Assam, northeast India; a search for an early warning indicator for impending disease outbreaks. *Parassitologia* **51**: 101–107.

Dev V, Dash AP. 2007. Rainfall and malaria transmission in north–eastern India. *Annals of Tropical Medicine & Parasitology* **101**: 457–459.

Devi NP, Jauhari RK. 2006. Climatic variables and malaria incidence in Dehradun, Uttaranchal, India. *Journal of Vector Borne Diseases* **43**: 21–28.

Dilalla DL, Dollinger SJ. 2006. Cleaning up data and running preliminary analysis. In *The psychology research handbook: a guide for graduate students and research assistants*, Leong FTL, Austin JT (eds). Sage Publications Inc.: California.

Dinpashoh Y, Jhajharia D, Fakheri-Fard A, Singh VP, Kahya E. 2011. Trends in reference crop evapotranspiration over Iran. *Journal of Hydrology* **399**(3–4): 422–433, DOI:10.1016/j.jhydrol.2011.01.021.

Gao CY, Xiong HY, Yi D, Chai GJ, Yang XW, Liu L. 2003. Study on meteorological factors-based neural network model of malaria. *Zhonghua Liuxingbingxue Zazhi* **24**(9): 831–834.

Gardner MW, Dorling Sr. 1998. Artificial neural networks: the multilayer perceptron: a review of applications in atmospheric sciences. *Atmospheric Environment* **32**: 2627–2636.

Giannini A, Saravanan R, Chang P. 2003. Oceanic Forcing of Sahel Rainfall on Interannual to Interdecadal Time Scales. *Science* **302**: 1027–1030.

Githeko AK, Lindsay SW, Confalonieri UE, Patz JA. 2000. Climate change and vector-borne diseases: a regional analysis. *Bulletin World Health Organisation* **78**(9): 1136–1147.

Haykin S. 2001. *Neural Networks: A Comprehensive Foundation*, 2nd edn. Pearson Education Inc.: New Delhi, India.

Hingane LS, Kumar KR, Murty BVR. 1985. Long term trends of surface air temperature in India. *International Journal of Climatology* **5**(5): 521–528.

Hsieh WW. 2001. Nonlinear principal component analysis by neural networks. *Tellus A* **53**: 599–615.

Hsieh WW, Tang B. 1998. Applying Neural Network Models to Prediction and Data Analysis in Meteorology and Oceanography. *Bulletin of the American Meteorological Society* **79**: 1855–1870.

IPCC. 2001. *Climate Change 2001: The scientific basis. Contribution of working group I to the third assessment report of the Intergovernmental Panel on Climate Change (IPCC)*. Cambridge University Press: Cambridge, UK.

IPCC. 2007. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press: Cambridge, United Kingdom and New York, NY, USA, 996.

Jain V, Nagpal AC, Joel PK, Shukla M, Singh MP, Gupta RB, Dash AP, Mishra SK, Udhayakumar V, Stiles JK, Singh N. 2008. Burden of cerebral malaria in Central India (2004–2007). *American Journal of Tropical Medicine and Hygiene* **79**(4): 636–642.

Jayavanth S, Singh M. 2003. Artificial neural network analysis of malaria severity through aggregation and deformability parameters of erythrocytes. *Clinical Hemorheology and Microcirculation* **29**: 457–468.

Jepson WF, Moutia A, Courtois C. 1947. The malaria problem in Mauritius: The bionomics of Mauritian anophelines. *Bulletin of Entomological Research* **38**: 177–208.

Jhajharia D, Shrivastava SK, Sarkar D, Sarkar S. 2009. Temporal characteristics of pan evaporation trends under the humid conditions of northeast India. *Agricultural and Forest Meteorology* **149**: 763–770.

Jhajharia D, Singh VP. 2011. Trends in temperature, diurnal temperature range and sunshine duration in northeast India. *International Journal of Climatology* **31**: 1353–1367, DOI: 10.1002/joc.2164.

Jhajharia D, Dinpashoh Y, Kahya E, Singh VP, Fakheri-Fard A. 2011. Trends in reference evapotranspiration in the humid region of northeast India. *Hydrological Processes*. DOI: 10.1002/hyp.8140.

Jolliffe IT. 2002. *Principal component analysis.* Springer: New York.

Kahya E, Kalayci S. 2004. Trend analysis of stream?ow in Turkey. *Journal of Hydrology* **289**: 128–144.

Kendall MG. 1975. *Rank Correlation Methods*, 4th edn. Charles Griffin: London.

Kochar DK, Kochar SK, Agrawal RP, Sabir M, Nayak KC, Agrawal TD, Purohit VP, Gupta RP. 2006.. The changing spectrum of severe falciparum malaria: a clinical study from Bikaner (northwest India). *Journal of Vector Borne Diseases* **43**: 104–108.

Kochar DK, Sirohi P, Kochar SK, Budania MP, Lakhotia JP. 2007. Dynamics of malaria in Bikaner, Rajasthan, India (1975–2006). *Journal of Vector Borne Diseases* **44**: 281–284.

Kothyari UC, Singh VP. 1996. Rainfall and temperature trends in India. *Hydrological Processes* **10**: 357–372.

Kumar S, Merwade V, Kam J, Thurner K. 2009. Streamflow trends in Indiana: Effects of long term persistence, precipitation and subsurface drains. *Journal of Hydrology* **374**(1): 171–183.

Lindsay SW, Martens WJ. 1998. Malaria in the African highlands: past, present and future. *Bulletin of the World Health Organization* **76**: 33–45.

Leung TF, Wong GWK, Ko FWS, Lam CWK, Fok TF. 2005. Clinical and atopic parameters and airway inflammatory markers in childhood asthma: a factor analysis. *Thorax* **60**: 822–826.

Maier HR, Dandy GC. 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. *Environmental modeling and Software* **15**: 101–124.

McMichael AJ, Marten WJM. 1995. The health impact of global climate changes: grasping with scenarios, predictive models and multiple uncertainties. *Ecosystem Health* **1**: 23–33.

Makler MT, Palmer CJ, Ager AL. 1998. A review of practical techniques for the diagnosis of malaria. *Annals of Tropical Medicine and Parasitology* **92**: 419–434.

Mann HB. 1945. Non-parametric tests against trend. *Econometrica* **33**: 245–259.

Monahan AH. 2000. Nonlinear Principal Component Analysis by Neural Networks: Theory and Application to the Lorenz System. *Journal of Climate* **13**: 821–835.

Muñoz-Díaz D, Rodrigo F. 2004. Spatio-temporal patterns of seasonal rainfall in Spain (1912–2000) using cluster and principal component analysis: comparison. *Annales Geophysicae* **22**: 1435–1448.

Partal T, Kahya E. 2006. Trend analysis in Turkish precipitation data. *Hydrological Processes* **20**: 2011–2026.

Patz JA, Engelberg D, Last J. 2000. The effects of changing weather on public health. *Annual Review of Public Health* **2**: 271–307.

Peng Bt, Tong S, Donald K, Parton KA, Jinfa N. 2003. Climatic variables and transmission of malaria: A 12-year data analysis in Shuchen County, China. *Public Health Report* **118**: 65–71.

Ramirez MCV, de Campos Velho HF, Ferreira NJ. 2005. Artificial neural network technique for rainfall forecasting applied to the Sao Paulo region. *Journal of Hydrology* **301**: 146–162.

Rogers DJ, Randolph SE. 2000. The Global Spread of Malaria in a Future Warmer World. *Science* **289**: 1763–1766.

Rojas R. 1996. *Neural networks: a systematic introduction*. Springer-Verlag New York, Inc.: New York, USA.

Roy SS, Balling RC. 2005. Analysis of trends in maximum and minimum temperature, diurnal temperature range, and cloud cover over India. *Geophysical Research Letters* **32**(12): L12702, DOI: 10.1029/2004GL022201.

Rupa Kumar K, Krishna Kumar K, Pant GB. 1994. Diurnal asymmetry of surface temperature trends over India. *Geophysical Research Letters* **21**(8): 677–680.

Seman NA, Isa NAM, Li LC, Mohamed Z, Ngah UK, Zamli KZ. 2008. Classification Of Malaria Parasite Species Based On Thin Blood Smears Using Multilayer Perceptron Network. *International Journal of the Computer, the Internet and Management* **16**: 46–52.

Sen PK. 1968. Estimates of the regression coefficients based on Kendall's tau. *Journal of the American Statistical Association* **63**: 1379–1389.

Shiva Nagendra SM, Khare M. 2006. Artificial neural network approach for modelling nitrogen dioxide dispersion from vehicular exhaust emissions. *Ecological Modelling* **190**: 99–115.

Singh P, Kumar V, Thomas T, Arora M. 2008. Basin-wide assessment of temperature trends in northwest and central India. *Hydrological Sciences Journal* **53**(2): 421–433.

Sousa SIV, Martins FG, Alvim-Ferraza MCM, Pereira MC. 2007. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environmental Modelling and Software* **22**: 97–103.

Sturchler D, Naef U, Fernex M, Mittelholzer ML, Reber RM, Steffen R. 1990. Malaria and mosquitoes: how often for how long? *Transactions of the Royal Society of Tropical Medicine and Hygiene* **84**: 780, DOI: 10.1016/0035-9203(90)90075-P.

Srivastava HN, Dewan BN, Dikshit SK, Rao PGS, Singh SS, Rao KR. 1992. Decadal trends in climate over India. *Mausam* **43**: 7–20.

Theil H. 1950. A rank invariant method of linear and polynomial regression analysis. Part 3. *Netherlands Akademie van Wettenschappen, proceedings* **53**: 1397–1412.

Tourassi GD. 1997. The Effect of Data Sampling on the Performance Evaluation of Artificial Neural Networks in Medical Diagnosis. *Medical Decision Making* **17**: 186–192.

Webster TJ. 2001. A principal component analysis of the U.S. News & World Report tier rankings of colleges and universities. *Economics of Education Review* **20**: 235–244.

WHO. 2002. *The malaria problem in South East Asia region*. Available at: http://www.searo.who.int/en/Section10/Section21/Section340_4018.htm. Accessed on 6 December 2007.

Yadav RR, Park WK, Singh J, Dubey B. 2004. Do the western Himalaya defy global warming? *Geophysical Research Letters* **31**: L17201, DOI: 10.1029/2004GL020201.

Yu Z-P, Chu P-S, Schroeder T. 1997. Predictive Skills of Seasonal to Annual Rainfall Variations in the U.S. Affiliated Pacific Islands: Canonical Correlation Analysis and Multivariate Principal Component Regression Approaches. *Journal of Climate* **10**: 2586–2599.

Yue S, Pilon P, Phinney B, Cavadias G. 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrological Processes* **16**: 1807–1829.

Zhang X, Harvey KD, Hogg WD, Yuzyk TR. 2001. Trends in Canadian streamflow. *Water Resources Research* **37**(4): 987–998.

