British Journal of Environment & Climate Change
7(1): 26-42, 2017; Article no.BJECC.2017.003
ISSN: 2231–4784

SCIENCEDOMAIN _international_
www.sciencedomain.org

Seasonally Separated Logistic Models to Assess the Impact of Climate Variables on Occurrence of Rainfall over the Bagmati River Basin of Nepal

Rajendra Man Shrestha[1*] and Srijan Lal Shrestha[2]

[1]Padmakanya Multiple Campus, Kathmandu, Nepal.
[2]Central Department of Statistics, Tribhuvan University, Kirtipur, Nepal.

Authors' contributions

This work is carried out in collaboration of both authors. Author RMS managed the literature review, designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Author SLS contributed in data analysis, writing and editing the manuscript. Both the authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJECC/2017/30696

Original Research Article

Received 25th November 2016
Accepted 16th February 2017
Published 30th March 2017

ABSTRACT

Aims: This paper aims to develop prediction models for forecasting rainfall occurrence over the Bagmati river basin of Nepal based upon climate related predictor variables.
Study Design: Time series design with statistical downscaling of large scale daily climate data and observed rainfall data.
Place and Duration of Study: Study was conducted at Central Department of Statistics, Tribhuvan University, Kirtipur, Nepal, between 2013 and 2015.
Methodology: A day is considered as a wet day if area weighted daily rainfall (AWDR) is more than 1 mm. Extreme rainfall is determined by the 98thpercentile of AWDR. Binary logistic regression models are built with 13 possible principal components (PCs) of 7 climate related predictor variables using daily data for 1981-2000 period. Thereafter, built models are validated for 2001-2008 period.
Results: Nine separate seasonal logistic models are fitted with Hosmer-Lemeshow tests having at least 0.207 *p*-values. The first PC of Air surface temperature has the greatest influence with odds ratio (OR) of 4.757 in predicting a wet day during post-monsoon across four models. It is followed by the first PC of Relative humidity with OR (4.112) in winter, first PC of Relative humidity with OR (3.443) in pre-monsoon and second PC of Relative humidity with OR (3.601) in monsoon. Similarly, second PC of Relative humidity has the highest contribution with OR (7.395) in predicting extreme rainfall in post-monsoon across all five models. It is followed by the first PC of Air surface

---

*Corresponding author: E-mail: rajendramanshrestha65@gmail.com;

temperature with OR (7.194) in monsoon, first PC of Relative humidity in winter with OR (6.820) and pre-monsoon with OR (5.076), and second PC of Relative humidity with OR (3.186) for the non-seasonal model.

Conclusion: The developed logistic regression models are applicable in forecasting rainfall occurrence seasonally in the Bagmati river basin of Nepal.

Keywords: Bagmati basin; climate predictors; climate models; logistic regression; rainfall; statistical downscaling.

1. INTRODUCTION

Rainfall is a natural process and very important hydrological phenomenon that affects human life and environment in both positive and negative perspectives. Globally, it is found that climate change has significant effect on rainfall and its cycle with increased climatic variability [1]. The circulation and distribution of availability of water under climate projection has become more complex and difficult [2,3]. Climate change has affected not only the overall magnitude of rainfall but also its seasonal distribution and inter-annual variability worldwide [4]. The Intergovernmental Panel on Climate Change (IPCC) stated that climate hazards, including changes in precipitation cycles, reduced crop yields due to extreme weather event and changing local temperature, are likely to decrease the food security of vulnerable population [5]. It is stated that climate change is not expected to be homogeneous globally. Substantial differences in precipitation trends at the regional level will occur [6]. In context of Nepal, it is found that there is a huge fluctuation in runoff from season to season [7]. For instance, a difference in runoff between 400m^3/s to 4300 m^3/s from February to August were estimated in Sapta-Koshi [7] which induced the risk of flooding. Further, landslide and sedimentation are likely to occur due to intense precipitation in monsoon and water shortage in dry season [8]. In Nepal, seasonal prediction of rainfall is required by many sectors such as agricultural sector and for hydropower generation [9]. Nepal Electricity Corporation has reported that, from past some years, Nepal has been experiencing prolong load-shedding problems as a consequence of insufficient water availability despite the fact that the power shortage is mainly due to a less number of hydropower projects. Naturally, the load-shedding schedule is affected seasonally due to seasonal variation in rainfall.

In order to understand the seasonal rainfall patterns given climate change scenario, many studies have shown that General Circulation Models (GCMs) are currently the most credible tools and also provide estimates of climate variables (e.g. air temperature, relative humidity etc.) on a global scale. However, direct use of GCM outputs is not suitable to assess the climate change impact at regional level due to their coarse resolution in finer scale [10]. Therefore, one of common methods applicable to solve this problem is the empirical statistical-downscaling technique. This technique is widely employed to downscale climate information from the global scale [11]. Yarnal [12] mentioned that the goal of the downscaling is to describe the relationship between atmospheric circulations and the surface environment, with attention being focused more on model parsimony and accuracy, rather than understanding the relationship between them. Under the fundamental assumptions of the statistical downscaling techniques, the relationships observed between predictor(s) and the response is time-invariant [12]. Consequently, this paper has attempted to use the statistical-downscaling technique to explain the rainfall occurrence for separate seasons. This technique assumes that the relationship between large scale variables and local variables should explain a large part of the observed variability at basin and the expected changes in the mean climate condition should lie within the natural variability [13].

Several studies were found applying different downscaling techniques for the study of climate system. For example, Wigena [14] has applied statistical downscaling model to predict the rainfall in Indramayu. The analysis was done to determine the best domain output by using projection pursuit regression. Furthermore, Cavazos and Hewitson [15] showed the performance of NCEP/NCAR output to find the potential combination of response variables by using artificial neural network. Sahar et al. [16] performed genetic programming for the downscaling of extreme rainfall events on the east coast of Peninsular Malaysia. They found that models derived using this technique can predict both annual and seasonal extreme rainfall indices more accurately compared to artificial neural network and statistical downscaling model (SDSM) [17].

Some climate change impact studies have also been conducted in Nepalese region [18,19]. Mishra et al. [10] used quantile-based bias correction method for climate projection downscaling and impact assessment on precipitation over upper Bagmati River basin in Nepal. Babel et al. [20] used a statistical-downscaling technique to study about climate change and water resources in the upper Bagmati River basin, Nepal. Likewise, Shrestha et al. [21] studied about impact of climate change on River flow and hydropower production in Kulekhani hydropower project of Nepal using SDSM. Shrestha et al. [22] studied the rainfall occurrence using logistic model as a statistical downscaling technique for the Bagmati River basin in Nepal but without accounting seasonal variability. This paper attempts to address the seasonal effects of the climate predictor variables on the rainfall occurrence over the Bagmati River basin [22] in Nepal.

Till the present research work, no literature was found that assessed the occurrence of wet (or dry) day and occurrence of extreme (or higher) rainfall day accounting seasonal effects using the probabilistic projection method (or a binary logistic regression) as a statistical downscaling technique in Nepal. Thus, in order to fill this research gap, the present study is conducted for the investigation of the rainfall occurrence pattern given the effect of some possible climate predictor variables, which are found well simulated by GCMs [10]. Further, extensive reviews of the published papers concerning the downscaling techniques and use of them in projection of the rainfall at their own target locations have motivated us to study about the rainfall occurrence under seasonal effects by developing predictive models with the following research questions. (i) Are the selected climate predictor variables well fitted to the target model to project rainfall occurrence with seasonal separation? (ii) What is the extent of effects of the predictors on rainfall occurrence separated under seasonal changes? With these research questions, the present study thus has the first objective set to develop the predictive models of the rainfall occurrence (a dependent variable) as statistical downscaling techniques using binary logistic regression on the basis of the reanalysis data for climate predictor variables available from NCEP/NCAR project considering seasonality. Further, it also has the second objective set to assess the effect of the different predictors on the rainfall occurrence through those models. Therefore, it is hypothesized that the selected predictors have significant effects on the rainfall occurrence at a given confidence level.

Review of the papers pertaining to the rainfall pattern with use of logistic regression model as statistical downscaling technique, a binary logistic model, a member of generalized linear models (GLM) [23], has been frequently applied to model climatological data. Some examples are evident from papers published by Chandler and Wheater [23], Chandler [24]; Yan et al. [25]; Fealy et al. [11]; Prasad et al. [26]; Filho et al. [27]; Shrestha et al. [22].

In the research papers, it is found that Nadja (2005) used GLM with logit link (logistic model) to simulate daily rainfall at Heathrow, Birmingham and Manchester airports, United Kingdom. The results were that all of the models projected a decrease in mean daily rainfall in summer and an increase in winter at Heathrow [28]. In addition, Prasad et al. [26] used a logistic regression approach for monthly rainfall forecasts in meteorological subdivision of India based on DEMTER retrospective forecasts. The model showed good performance in capturing extreme rainfall years and appeared to perform better than the direct model forecasts of total precipitation in such years. A study used quantile regression as statistical downscaling technique to estimate extreme monthly rainfall at station Bangkir Indonesia. The results showed that at 95th percentile, the pattern of forecasted rainfall in January to December 2008 was similar to actual rainfall with correlation 0.98 and the forecasted rainfall (843 mm) in February 2008 was considered as the extreme rainfall month which confirms well to the highest actual rainfall (727 mm) with probability 0.99 [29].

Therefore, this study attempts to address some key issues related to downscaling non-normally distributed climate variables like rainfall [11] in a topographically adverse location like the Bagmati River basin in Nepal. It is also expected that the models so built would be applicable to study about seasonally separated rainfall phenomena in future in the Bagmati River basin under the impact of climate change when the GCM outputs with future emission scenarios are available.

2. MATERIALS AND METHODOLOGY

2.1 Study Area

The Bagmati River basin (BRB) is located within the middle mountain of Nepal. The Fig. 1 is a map of the BRB located within Nepal. It extends from 26° 45'N-27° 49' N and 85° 02' E-85°27' E and has a catchment area of 3, 750 square kilometers in Nepal. However, the area considered for the study is 3604.44 square kilometer based upon location of 25 stations and derived using of Thiessen polygon method in GIS software [22]. The Bagmati River originates from the Shivapuri hills of the Mahabharata range in the Kathmandu Valley and drains out of Nepal across the Indian state-Bihar.

Fig. 1. Map of Nepal and the Bagmati River Basin with its tributaries and 25 stations

It reaches the River Ganges after passing through the inner Mahabharata range and the plain of Terai. It is mentioned that the elevation of the Bagmati River basin ranges from about less than 80 m in Terai, its southern part to 2900 m in the Mahabharata range, its northern part [30]. Its length is about 51 km in Nepal. Its main tributaries are Manohara, Bishnumati, Kulekhani, Kokhajor, Marin, Chandi, Jhanjh and Manusmara. The Kathmandu valley comprises 15% of the basin area in Nepal. Main source of water in the Bagmati River basin is rain and natural springs [30]. There are tributaries and sub-tributaries of the Bagmati River shown in the map along with precipitation, climatological and agro-meteorological stations indicated by black triangles.

2.2 Climate of the Study Area

Nepal is ecologically divided into three belts namely Himalayan, Mountain and Terai belts and surrounded by India (West, East and South) and China (North). The nearest ocean is the Indian Ocean and near to the Bay of Bengal at the south-east side. The climatic condition of the Bagmati River basin is therefore, quite changing due to the intrinsic local topography and influenced by global climatic activities due to the atmosphere and Ocean. Rainfall and temperature patterns differ according to seasons in this area. Basically there are four seasons in Nepal, namely, winter (December to February), pre-monsoon (March to May), monsoon or summer (June to September) and post-monsoon (October to November). Temperature generally decreases with elevation and becomes low in winter and high in summer. More specifically, the climate changes naturally from cold temperature in higher mountains via warm temperature at mid-elevation levels to subtropical region in the southern low land (Terai). Thus, the whole Bagmati River basin is actually divided into three climatic zones.

Cool temperature zone lies between 2000-3000 m which covers only about 5 percent of the basin with mean annual temperature varying between 10℃ to 15℃. The warm temperature zone lies between 1000-2000 m which covers about 60 percent part of the basin with mean annual temperature varying between 15℃ to 20℃. Lastly, the sub-tropical zone lies below 1000 m which covers southern part of the basin with the Siwaliks and Terai with mean annual temperature ranging between 20℃ to 30℃. The mean relative humidity of the basin varies between 70% and 86%. Annual rainfall is about 1800 mm with 80 percent of the total rain occurring in the monsoon season [30]. Rainfall occurrence in the basin is mainly due to the south east monsoon, generally starting from June and ending at September. In this course, the humid monsoon air stream blows from the Bay of Bengal and rises till it meets the Himalaya. Then ultimately, rainfall occurs heavily on some section of the southern Himalayan slopes. It also occurs heavily along the Chure range. As mentioned in a report of Department of Hydrology and Meteorology in Nepal, the area close to the Indian boarder receives about 1500 mm rain annually. It rises up to 2000 mm at the foot hills of the Chure but it diminishes at the northern part of the Chure. It is also noted that rainfall reduces due to the rain shadow effect. Furthermore, the rainfall pattern also changes due to orographic effect in this region [10]. From the statements, it is clear that the study area, although small in size compared to Nepal has spatially varied climatic situation.

2.3 Data

The present study considered only 25 stations installed at different parts of the basin due to unavailability of complete time-series of daily rainfall data. The data, available from the Department of Hydrology and Meteorology, Kathmandu, Nepal covers 28 years of daily data between January 1981 and December 2008. The daily rainfall of all stations is aggregated into single time-series using area weighted mean. Specifically, the Thiessen polygon method in Arc GIS 9.3 [22] is used for the purpose. Then, the resulted time series of the rainfall is named as area weighted daily rainfall (AWDR) and hence rainfall pattern of the BRB.

2.3.1 Outcome variable

For statistical modeling, there are two outcome variables representing status related to rainfall. The first outcome variable represents a day either as wet day with code 1 when AWDR is more than 1 mm or as dry day with code 0 otherwise.

The second outcome variable represents a day either as extreme/higher rainfall day with code 1 when AWDR is equal to or more than it's the 98th percentile, or no extreme/lower rainfall day with code 0 otherwise. The past papers have demonstrated various percentile levels (for example, 90 or 95 or 98) to represent the extreme/higher rainfall event. In this study, the 98th percentile model showed the better model meeting all possible criteria that are necessary. Both the outcome variables may be treated as rainfall occurrence or not.

2.3.2 Predictors

A number of predictors required for modeling are unavailable from meteorological stations. The study is not only limited to build a functional relationship between a dependent variable and a set of independent variables but also to forecast the scenario of the dependent variable in future. Hence, it is necessary to have information on past, present and future time-series data of the predictors. To overcome this difficulty, data are obtained from the National Center for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis project (website: http [31]. In this study, NCEP/NCAR reanalysis data resemble the observed data, usually used in hydro-climatology. Furthermore, in absence of observed data, it is not possible to check the validity of this data [32].

Studies in the past have found that rainfall phenomenon is very complex system on the earth. This paper considers only seven climate predictors despite there may be more than these to explain the rainfall event in a region. Geopotential height (GPH) (m) at 850 hpa, Relative humidity (RH)(%) at 850 hpa , Air surface temperature (AST) in Kelvin , Sea level pressure (pa), U- component of wind (Zonal wind)(UW) (m/s) at 850 hpa , V - component of wind (Meridional wind) (VW) (m/s) at 850 hpa and Precipitation flux (PF) (kg m^{-2} s^{-1}) are selected for spatial resolution of 25°N - 30°N in latitude and 82.5°E – 87.5°E in longitude for the period 1981-2008.These grids encapsulate the Nepal and hence the BRB.

For each predictor each with 9 grids, principal component analysis (PCA) is applied in order to reduce dimensionality and eliminate multicollinearity effect in the model [22]. The reduction in dimensionality and the principal component analysis coefficients are obtained in the study carried out by Shrestha et al. [22]. However, PCs of Precipitation flux were not included in the study. The following Table 1 shows the remaining PCs for Precipitation flux only.

2.4 Methods

2.4.1 Rainfall occurrence model

A binary logistic regression model, as Statistical downscaling method, is used to establish a functional relationship between a outcome variable and a set of 13 predictor variables (Table 3 in [22]) using the1981-2000 as calibration period. The fitted model is validated by the period 2001-2008 [33]. Moreover, there are four separate models for winter, pre-monsoon, monsoon and post-monsoon seasons for the first outcome variable. Additionally, there are five models for four season and year for the second outcome variable. Here, symbolically an outcome variable, say, Y_i (defined in section 2.3.1) assumes two possible values 1 and 0 with probability of rainfall occurrence or success (p_i), and 1- p_i, probability of no rainfall occurrence or failure respectively. Then, an odd of success is expressed as $\left(\frac{p_i}{1-p_i}\right)$ the ratio of the probability of success to the probability of failure. The logistic regression model [34] is expressed:

$$Log\left(\frac{p_i}{1-p_i}\right)=\beta_0+\sum_{i=1}^{k}\beta_i x_i \quad (1)$$

p_i can be computed by:

$$p_i=\frac{e^{\beta_0+\sum_{i=1}^{k}\beta_i x}}{1+e^{\beta_0+\sum_{i=1}^{k}\beta_i x}}=\frac{1}{1+e^{-\left(\beta_0+\sum_{i=1}^{k}\beta_i x\right)}} \quad (2)$$

β_0, β_1 = coefficients of the model estimated from the data

x_i= i^{th} predictor variable, I = 1, 2, …, k.

The maximum likelihood estimation (MLE) method is employed to estimate parameters within the model. The predicted values can be interpreted as probability p_i with range (0, 1) evaluated by the equation 2.

2.4.2 Model fit and diagnostics

There can be 8191 (or 2^{13} -1) possible binary logistic models with or without significant 13 predictors. Selecting the best one manually is very difficult task. The forward and backward likelihood ratio or purposeful selection method are considered for the selection of significant predictors. A goodness of fit is assessed by Hosmer-Lemeshow (HL) test and Deviance Statistic.

Standardized Deviance residuals are examined to detect the outlier(s) against linear predictor. The residuals falling outside range ±3 are regarded as outliers. Index of concordance (‘c’) as area under curve (AUC) is applied to assess the capability of accurate classification of a model. The bigger value between 0 and 1 suggests a better overall performance of the model as capability of best prediction.

3. RESULTS AND DISCUSSION

Four separate seasonal models are calibrated with the data for the period 1981-2000. The models are used to compare the effects of each of 13 predictors included in the separate models irrespective of whether they are statistically significant or not at 5% level. Table 2 shows the results of all models with goodness of fit test. Table 2 depicts that all four models are not equally fitted on the basis of HL test. Models 1 and 2 are well fitted as they show higher insignificant *P*-values but Models 3 and 4 are not well-fitted to the data. However, Deviance statistics show that all the fitted models have significant results which are contradictory to the HL test. However, Table 2 is constructed basically to compare the influence of each of all the 13 predictors present in the model on the outcome variable.

Geopotential height1 has positive effect on the outcome variable except for Model 3 for monsoon season. But the coefficients are insignificant in all four models. It has relatively more impact in post-monsoon season irrespective of sign of the coefficient.

Relative humidity has two components with positive coefficients significant for all the models. This indicates that it has a very important role in predicting a day as wet day in all seasons throughout a year.

Sea level pressure1 shows a negative impact in all models except in Model 3 (monsoon season). Like Geopotential height1, it has greater impact in post-monsoon season with insignificant negative coefficient. But it has a negative but significant effect on the outcome variable in pre-monsoon season (Model 2).

Table 1. Spatial grid loadings of predictor variables

Predictor variables	Lon82.5_Lat_25 (G1)	Lon85._Lat_25 (G2)	Lon87.5_Lat_25 (G3)	Lon82.5_ Lat_27.5 (G4)	Lon85_ Lat_27.5 (G5)	Lon87.5_ Lat_27.5 (G6)	Lon82.5 _Lat_30 (G7)	Lon85_ Lat_30 (G8)	Lon87.5_Lat_30 (G9)
PF1	0.184	0.306	0.414	0.451	0.737	0.715	0.698	0.9	0.857
PF2	0.865	0.91	0.729	0.736	0.52	0.428	0.313	0.242	0.214

PF1 = 0.184(PF1_G1) + .306 (PF1_G2) + .414 (PF1_G3) + .451(PF1_G4) + .737(PF1_G5) + .715(PF1_G6) + .698(PF1_G7) + .900(PF1_G8) + .857(PF1_G9) and so on

Table 2. Coefficients and Wald Test in seasonally separated binary logistic models for a rainfall occurrence: a wet day/dry day

Predictor variables	Parameter Estimates (B Coefficients)			
	Model 1 (Winter)	**Model 2 (Pre-monsoon)**	**Model 3 (Monsoon)**	**Model 4 (Post-monsoon)**
Geopotential height1	1.789	2.372	-2.001	4.084
Relative humidity1	1.525*	1.425*	.996*	1.302*
Relative humidity 2	1.010*	1.181*	1.122*	1.697*
Sea level pressure1	-2.274	-3.978*	1.844	-5.254
Precipitation flux 1	.834*	-.214	.277*	.049
Precipitation flux 2	.090	.261	.297*	.192
Air surface temperature1	-1.556	-1.437*	.729	-.997
Air surface temperature2	-.811	.114	.604	-1.068
U-wind1	.721*	-.496*	-.068	.067
U-wind2	.377*	-.167	-.223*	.014
V-wind1	.270*	.009	-.077	-.007
V-wind 2	.305	.596*	-.025	.426*
V-wind 3	-.170	.129	-.251*	.073
Constant	-3.237*	-1.093*	-.871*	-1.503*
HL Test/df	5.677/8	6.102/8	16.912/8	20.074/8
(p-value)	(.683)	(.636)	(.031)	(.010)
Deviance (-2LL)	627.817	1757.067	1681..249	688.265
Df	1796	1821	2426	1206

The procedure models treated Dry day as the reference category (Test-statistic is Wald test with significant p-values: $^{*}P< 0.01$ $^{**}P< 0.05$ and $^{***}P< 0.10$)

Precipitation flux with two components has positive impact on the outcome variable except in model 2 along with negative insignificant impact of its first PC. Its two PCs seem changing in their significance with respect to different seasons. However, the first component has relatively a greater influence in the winter season.

Air surface temperature with two PCs has insignificant negative impact in all models except in Model 2. But its second PC has insignificant positive impact in some models (Table 2). This may be due to its spatial variation or presence of multicollinearity in the models. However, its first PC has a greater influence in the winter season. It reflects that a day has more likely to be a dry day. Naturally, this is true and common result in the winter.

Likewise, U-wind has two components and V-wind has three components in all models. But most of them are insignificant with varying impact (positive or negative) on the outcome variable. This variability in the wind may be because of seasonality effect or spatial variation or multicollinearity effect.

When the coefficients of all 13 predictors are compared across all models, they have a varying impact with different magnitudes with varying test results seen from one model to other. These varying outputs thus show the seasonal effect of the predictors on the outcome variable. However, these models seem to violate restrictive assumption of regression, for example, multicollinearity.

Diagnosis of all models shows that there are multicollinearity effects detected by variance inflation factor (VIF) of some predictors, for example Geopotential height1 (VIF>139), Sea level pressure1 (VIF>247), Air surface temperature1 (VIF>14), and Air surface temperature2 (VIF>19).Results are not shown here because of limited space. In order to observe the significant effect of the predictors with less influence of multicollinearity, all the above four models are re-calibrated. The models so re-calibrated may exclude some predictors due to multicollinearity effect. This is true for both the rainfall occurrence models.

Tables 3(a) and 3(b) demonstrate the results of a goodness-of-fit test for both the rainfall occurrence variables. There are nine separate best fitted final models with these two outcome variables. All models show highly insignificant HL statistic (*P* values ≥ .20) with minimum AIC or BIC.

3.1 Adequacy and Validation of Model

All the models have highly significant index of concordance (Tables 3(a)-(b)) with a value ('c') more than 80 percent for their training periods. These indices are also more than 80 percent for validation period too. This evidences that the models are adequately validated. Examination of the assumption of the log-odds linearly related with the linear predictor is performed including the linear predictor and its square term with other predictors in each model. But the results showed insignificant result of the square of a linear predictor in each model.

Checking of over-dispersion in each model reveals absence of over-dispersion of models since the results of Deviance divided by degrees of freedom (mean deviance) in Tables 3(a)-3(b) are all less than one.

Scatter diagrams (Figs. 2(a), (b), (c), (d)) of standardized deviance residual against a value of linear predictors show that there may be few outliers in winter and post-monsoon but not in monsoon and pre-monsoon seasons. However, these outliers have no serous effect on the coefficients and their standard errors in the model when the models are re-run without them.

Tables 3(a) and 3(b) revealed that all 9 separate binary logistic models are well fitted to the data since the H-L statistic tests (Chi-square values with degrees of freedom (df) of 8) ranged from 3.327 to 11.258 with corresponding insignificant *p*-values from 0.912 to 0.188. The fitting of the models are also supported by the Deviance tests with the Chi-square's values ranging from 970.125 (df =7297) to 1612.34 (df=1828) each with the significant *p*-value < 0.01. Along with the goodness of fit test, test of adequacy and validation presented in section 3.1 have supported that all the models are well fitted and validated. Hence, the first objective is achieved and that all the nine separated binary logistic models are well developed and applicable for projecting the rainfall occurrence over the study area.

Tables 4(a) and 4(b) show different number of predictors included in the separate models developed. The Wald statistics in each model demonstrate that the coefficients of each of the predictors in each model are all highly significant at 5 percent level of significance. The Wald statistics (chi-square values each with df =1) are found to range from 5.347 to 434.178 with *p*-values less than 0.01. These results thus depict that the coefficients of the predictors in all models are significantly different from zero. Hence, the hypothesis of statistically significant predictors in all the models is satisfied with each coefficient different from zero.

While assessing the effect of each predictor in each model, interpretations are made in terms of odds ratio (OR). Relative humidity and Air surface temperature have demonstrated more contribution to the wet day compared to other predictors. The first component, Relative humidity1 has more significant positive effect on the wet day with OR of 4.112 in winter. Its effect seems gradually decreasing in pre-monsoon (OR=3.443), monsoon (OR=3.241) and post-monsoon (OR=2.092). More precisely, it has a greater impact on dry season than in wet season. However, the second component, Relative humidity2 has a greater effect (OR=3.601) on the wet day in monsoon. Further, this component has also more impact in winter (OR=2.857) and pre-monsoon (OR=3.213). However, it has no influence in post-monsoon. The changing impact of the Relative humidity seems to be due to (a) seasonal change or (b) its spatial variation due to difference in functional forms of principal components or (c) the inclusion of other climate predictors or all of them in different models. Relative humidity2 has demonstrated a greater influence on the wet in the monsoon where Air surface temperature is absent. But post-monsoon shows that Relative humidity1 has an effect on the wet day little less (OR=2.092) since there is the presence of Air surface temperature with its two components (OR=4.757 and OR=2.73). This fact evidences that Relative humidity has changing impact because of presence of second more influencing factor Air surface temperature along with others. Further, this also evidences that Air surface temperature is also giving significant effect differently from season to season. Its first component, Air surface temperature1 has less effect on the wet day in winter (OR=.343) and pre-monsoon (.559) but a greater effect on post-monsoon (OR=4.757). It also reveals that the spatial impact on wet day is due to presence of second component in post-monsoon (OR=2.73). Like Relative humidity, Precipitation flux has its one component giving a significant effect on the wet day in winter (OR=1.973) and post-monsoon (OR=1.595). Its spatial impact is not much evident like Relative humidity and Air surface temperature. Moreover, the less difference in OR for two seasons indicates that its impact on the wet day is almost homogeneous. It means that the Precipitation flux has a less seasonal impact on the wet day. Other factors like Geopotential height, U-wind and V-wind have their components too. U-wind with its two components (OR=.588 in pre-monsoon and OR=.759 in monsoon) has less impact on the wet day and shows relatively similar effect on both seasons. But V-wind has little different nature of its impact. Like Air surface temperature, V-wind shows positive impact (OR=1.707) with its second component, V-wind2 but negative impact (OR=.831) in monsoon. This behavior of V-wind indicates that there are both spatial and seasonal variations and impact due to the presence of other factors in the same model. Geopotential height demonstrates the negative impact on the wet day in monsoon (OR=.410) and post-monsoon (OR=.509) with its single component. Sea level pressure with its single component has least effect in pre-monsoon (OR=.296) and in winter (OR=.422) across all the factors. The presence of Relative humidity seems to dominate Sea level pressure in winter and pre-monsoon.

Likewise, coefficients of predictors included in 5 different binary logistic models for extreme/higher rainfall are also examined and assessed. It shows different odds ratios from one season to another. Like the models for wet day, Relative humidity also showed dominant character in predicting the extreme rainfall day in all models except for monsoon. Relative humidity2, the second component of the Relative humidity has the greatest effect (OR=7.395) followed by Relative humidity1, the first component (OR=4.435) in post-monsoon. The coefficients seem changing from season to season. Relative humidity1 has the second greatest effect on the extreme rainfall day (OR=6.82) in winter. But it has the least effect with Relative humidity2 (OR=2.77) in pre-monsoon. Annual model demonstrates its character little differently with (Relative humidity1 (OR=2.781) and Relative humidity2 (OR=3.186). Post-monsoon shows its highest influence because there is absence of most influencing factor Air surface temperature along with other predictors in this season. Such behavior of Relative humidity indicates that it has significant seasonal variation and more explanatory power for predicting the extreme rainfall if other more influencing factors are absent. It is also supported by the fact that it is absent in monsoon. At the same time, Air surface temperature1, the first component has the greatest effect (OR=7.194) on the extreme rainfall day. Other seasons demonstrate that Air surface temperature has little less effect on it as it is seen in pre-monsoon (OR=2.74) in the presence of two components of Relative humidity. Its similar pattern can also be seen in annual model (OR=2.01 for Air surface temperature1). Thus, both Relative humidity and Air surface temperature have relatively more dominant effect on the extreme rainfall and show changing nature from season to season with the presence of spatial variation. Again like in the models for wet day, Precipitation flux has two components present in all models except in post-monsoon. Its odds ratios are ranging from 1.215 in monsoon to 2.479 in winter. So, it indicates relatively homogeneous positive effect for the extreme rainfall over the basin for all seasons. This could be due to the presence of the most dominant factor, Relative humidity or Air surface temperature. Sea level pressure with single component has the smallest effect on the extreme rainfall and has almost similar negative effect on it in winter (OR=.322) and post-monsoon (OR=.210). It means that it is changing slowly from one season to another. U-wind with its two components has almost homogenous positive effect on the extreme rainfall with odds ratios ranging from 1.713 (for annual) to 2.632 (for monsoon). The seasonal variation of this predictor seems smaller compared to Relative humidity and Air surface temperature. But the spatial variation is found to be strong at its components present in the models. Similarly, V-wind has three components with odds ratio ranging from 1.602 for annual model to 2.082 for monsoon. This predictor has similar character as it is for U-wind. However, it show a significant negative effect on the extreme rainfall (OR=.599) in post-monsoon. This may be due to strong seasonal effect.

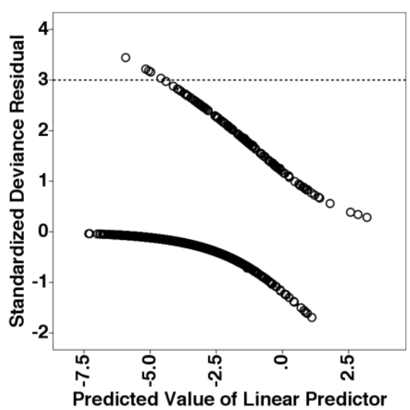

Fig. 2a. For winter

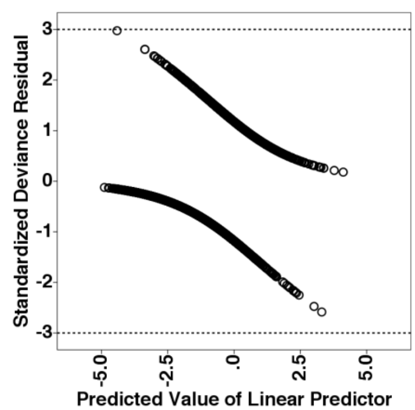

Fig. 2b. For pre-monsoon

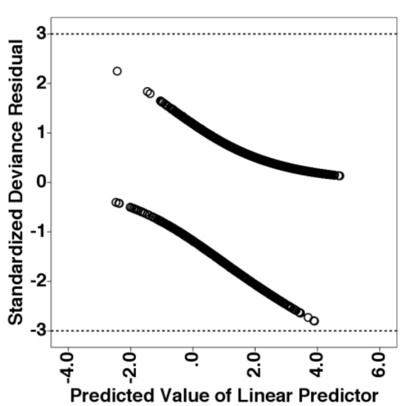

Fig. 2c. For monsoon

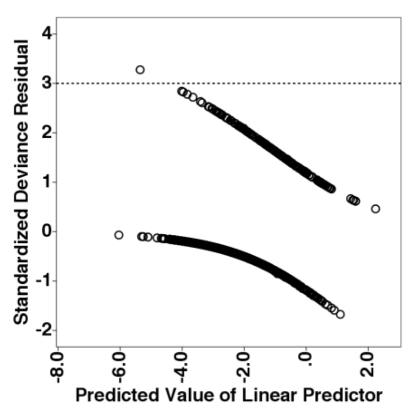

Fig. 2d. For post-monsoon

Fig. 2. Scatter-plot of Standardized Deviance Residual against predicted value of linear predictor

Table 3a. A goodness of Fit Test for a rainfall occurrence: Wet day/Dry day

Model	No. of predictor	Deviance		HL test		AIC	BIC	'c' for training period	'c' for validation period
		-2LL	df	χ^2 (df =8)	Sig				
Winter	5	659.36	1804	7.824	.451	671.36	704.36	.839***	.876***
Pre-monsoon	6	1612.34	1828	11.258	.188	1626.34	1664.95	.847***	.827***
Monsoon	6	1669.41	2433	7.312	.503	1683.41	1724.01	.812***	.720***
Post-monsoon	6	794.59	1214	3.916	.865	806.59	837.23	.807***	.842**

Note: df stands for degrees of freedom

Table 3b. A goodness of Fit Test for a rainfall occurrence: Extreme/higher Rainfall Day

Model	No. of predictor	Deviance		HL test		AIC	BIC	% of accurate classification for periods	
		-2LL	df	χ^2 (df=8)	Sig			Training	Validation
Winter	4	219.899	1805	6.447	.597	229.899	257.404	98.3***	98.1***
Pre-monsoon	4	294.664	1830	7.089	.527	304.664	332.238	98.0***	98.0***
Monsoon	5	373.668	2434	10.910	.207	385.668	440.467	98.0***	97.3***
Post-monsoon	4	138.455	1215	7.825	.451	148.455	173.988	98.1***	97.3***
Annual	7	970.125	7297	3.327	.912	986.125	1041.30	98.0***	97.1***

*Note: 'c' stands for index of concordance or measure of area under curve (AUC) in ROC analysis. HL refers to Hosmer and Lemeshow test. *P< 0.05, **P< 0.01 and ***P< 0.001 for Table 3(a) and (b).*
Note: df stands for degrees of freedom

Table 4a. Coefficient table for a rainfall occurrence: [*]Wet / Dry day

Predictors	Winter		Pre-monsoon		Monsoon		Post-monsoon	
	B (S.E.)	**OR**	**B (S.E.)**	**OR**	**B (S.E.)**	**OR**	**B (S.E.)**	**OR**
Geopotential height1	---	---	---	---	-.891*** (.104)	.410	-.676*** (.203)	.509
Relative humidity1	1.414*** (.161)	4.112	1.236*** (.111)	3.443	1.176*** (.098)	3.241	.738*** (168)	2.092
Relative humidity2	1.050*** (.203)	2.857	1.167*** (.114)	3.213	1.281*** (.132)	3.601	---	---
Sea level pressure1	-.862*** (.228)	.422	-1.217*** (.145)	.296	---	---	---	---
Air surface temperature1	-1.069*** (.333)	.343	-0.581*** (.181)	.559	---	---	1.560*** (.241)	4.757
Air surface temperature 2	---	---	---	---	---	---	1.005*** (.319)	2.73
Precipitation flux1	.680*** (.193)	1.973	---	---	---	---	.467*** (.149)	1.595
U-wind1	---	---	-0.531*** (.097)	.588	---	---	---	---
U- wind2	---	---	---	---	-.276*** (.082)	.759	---	---
V- wind2	---	---	.535*** (.103)	1.707	---	---	---	---
V- wind3	---	---	---	---	-.185*** (.062)	.831	---	---
Constant	-2.220*** (.448)	.109	-0.544*** (.159)	.580	-.721*** (.140)	.486	-.867*** (.200)	.420

[*]The procedure models treated Dry day as the reference category.
OR stands for Odds ratio

Table 4b. Coefficient table for a rainfall occurrence: *Extreme/higher Rainfall day

Predictors	Annual		Winter		Pre-monsoon		Monsoon		Post-monsoon	
	B (S.E.)	**OR**	**B (S.E.)**	**OR**	**B (S.E.)**	**OR**	**B (S.E.)**	**OR**	**B (S.E.)**	**OR**
Relative humidity1	1.023*** (.2696)	2.781	1.920*** (.281)	6.820	1.624*** (.346)	5.076	---	---	1.490** (.482)	4.435
Relative humidity2	1.159*** (.3296)	3.186	---	---	1.019*** (.295)	2.770	---	---	2.001*** (.492)	7.395
Sea level pressure1	---	---	-1.134** (.400)	.322	---	---	---	---	-1.560* (.684)	.210
Air surface temperature1	.698* (.3018)	2.010	---	---	---	---	1.973*** (.549)	7.194	---	---
Air surface temperature2	---	---	---	---	1.008*** (.334)	2.740	---	---	---	---
Precipitation flux1	.152** (.0548)	1.165	1.367*** (.234)	2.479	---	---	.195** (.075)	1.215	---	---
Precipitation flux2	---	---	---	---	.705** (.308)	2.024	---	---	---	---
U-wind1	.765*** (.0851)	2.149	.665** (.314)	1.944	---	---	.968*** (.149)	2.632	---	---
U-wind2	.538*** (.1497)	1.713	---	---	---	---	.699** (.270)	2.011	---	---
V- wind1	.471*** (.1163)	1.602	---	---	---	---	.733*** (.207)	2.082	---	---
V- wind3	---	---	---	---	---	---	---	---	-.512* (.266)	.599
Constant	-6.098*** (.3154)	0.002	-3.903*** (.458)	.020	-4.160*** (.458)	.016	-7.329*** (.803)	.001	-3.941*** (.524)	.019
98^{th} percentile	36.66		6.35		17.38		57.11		16.75	

**The procedure models treated No Extreme rainfall day as the reference category*
*Test-statistic is Wald test with significant p-values: *P< 0.05, **P< 0.01 and ***P< 0.001 for Table 4(a) and (b).*
OR stands for Odds ratio, which is given by Exp(B) in the above table

With consideration of the effect of the Relative humidity with its two components on the rainfall occurrence, review of literatures show a significant positive correlation between the precipitation and the relative humidity. For instance, Gerapetritis H [35] reported the positive correlation between them while using the FRH/FRHT data. It is also agreed by the report of Roy et al. [36], which revealed the positive effect of afternoon relative humidity on the precipitation event provided other predictors in logistic model. Like Relative humidity, Precipitation flux has also the same nature in influencing the rainfall occurrence along with its two components in all the seasons. On examining its two components, Air surface temperature2 mostly belong to latitude of 25° N along longitude of 82.5°– 87.5° E while its first component belonging to latitude of 27.5° – 30°N along longitude of 82.5°– 87.5° E. The former predictor reveals higher loadings to lower part (southern) of Nepal, especially Terai region, and the latter predictor shows higher loadings to upper/middle (northern) part of Nepal, especially Hill or Mountain regions. According to the report of Department of Hydrology and Meteorology in Nepal, the region closer to Indian boarder receives about 1500 mm rain in a year [37]. This fact verifies that the rainfall happens more in Terai than in Mountain. This is a strong evidence of spatial variation of this predictor affecting the rainfall occurrence mostly in monsoon season.

While reviewing the literatures, Gerapetritis H, also [35] evidenced that there was significant negative correlation with sea level pressure using the FRH/FRHT data. Furthermore, Roy et al. [36] reported in his logistic regression model that there was negative effect of maximum temperature and positive effect of minimum temperature with the precipitation event. Filho et al. [27] showed that the rainfall pattern was highly positively and significantly associated with relative humidity, maximum temperature and V-Wind component in the logistic regression model used in the northern Brazil. All these literatures agree with results of the study. Similarly, Khalil et al. had developed relationship between Precipitation and Temperature over the 80-year period from 1905 to 1984 at nearly 1000 stations in United States. He found that over most of the United States, summer precipitation and temperature were negatively correlated with indication of warm summers tended to be dryer in the central and southern Great Plains and a significant positive correlation between them over the area south of the Great Lakes covering the eastern portion of the Corn Belt in winter [30]. This supports the fact that the relation between the precipitation and temperature may change with season or place. Further, Trenberth et al. [38] found negative correlation between precipitation and surface temperature over land during summer and positive correlation at high latitudes in winter. He also added that ocean conditions drive the atmosphere with higher surface air temperature positively associated with precipitation. Kutiel et al. [39] shows that relationship between rainfall in Turkey and the regional sea level pressure is large in winter and non-existing in summer. Pressure patterns associated with dry conditions, showed usually positive departure, whereas, pressure associated with wet conditions showed negative sea level pressure. Prasad et al. (2010) demonstrated that the rainfall had significant positive correlation with V-Wind and negative correlation with U-wind when the logistic regression model was used to relate them on the basis of realizations obtained for whole India [26]. All these literatures are found to have more or less similar findings as obtained in the present study and the similar pattern of relationship of both rainfall occurrences with the predictors present in the fitted and validated models. Therefore, the major findings of this study are the significant effects of the predictors on the rainfall occurrence with presence of both spatial and seasonal variation over the basin.

4. CONCLUSION

This study has developed nine separate binary logistic regression models for the two rainfall occurrence (i) wet/dry day and (ii) the extreme/higher rainfall day given different components of seven climate predictor variables. All the models are good fit to the binary logistic models tested from Hosmer-Lemeshow statistics and Deviance statistic. These models have only included a number of significant predictors tested by Wald statistic. ROC and AUC have validated that all the models perform well in prediction of the rainfall occurrence. Components of Relative humidity are found to show a greater influence in predicting the rainfall occurrence mostly on winter for the wet day and on post-monsoon for the extreme rainfall day although the effect is found greatest due to Air surface temperature in monsoon for the extreme rainfall. Further Geopotential height and Sea level pressure depict negative association with the rainfall. But V-wind and U-wind show mostly positive association in most the models and negative association in few models. The degree and direction of impact of those predictors are found to vary from season to season. This evidenced that their impacts may vary from one season to another in predicting the rainfall occurrence. Finally, it is concluded that all the nine separate logistic models are applicable to downscale the rainfall occurrence for different seasons or year given those predictors in future. Therefore, the models can be applied as one of statistical downscaling techniques to assess the impact of climate change on the rainfall occurrence in the future at the Bagmati River basin in Nepal based on the outputs of general circulation models.

ACKNOWLEDGEMENTS

The authors are grateful and thank University Grants Commission, Sanothimi, Bhaktapur for providing PhD fellowships to conduct the research work.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Hansen J, Ruedy R, Sato M, Lo K. Global surface temperature change. Rev. Geophysics. 2010;48:RG4004. DOI: 10.1029/2010RG000345
2. Pachauria RK. Global warming impacts and implications for South Asia. In: The regions warming: impacts and response strategies [Schmandt, J. & Clarkson (eds.)], Oxford University Press, New York. 1992;79-90.
3. Chun KP. Statistical downscaling of climate model outputs for hydrological extremes. Imperial College London; 2010. DOI: 10.1029/2009RG000314
4. (a) Easterling DR. Climate extremes: Observations, modeling, and impacts. Science. 2000;289:2068-2074.
   (b) Trenberth KE, et al. IPCC climate change 2007: The physical science basis (Eds Solomon, S. et al.) Ch. 3 (Cambridge Univ. Press); 2007.
   (c) Zeng N, Neelin J, Lau K, Tucker C. Enhancement of inter-decadal climate variability in the Sahel by vegetation interaction. Science. 1999;286:1537–1540.
5. Intergovernmental Panel on Climate Change (IPCC). Climate Change 2013: The physical science basis. Working Group I Contribution to the Fifth Assessment Report of The Intergovernmental Panel on Climate Change. Cambridge University Press: Cambridge, UK and New York, NY. Ireson R. 2006. Designing food security in North Korea: Designing; 2013.
6. Reiter A, Weidinger R, Mauser W. Recent climate change at the upper Danube – a temporal and spatial analysis of temperature and precipitation time series. Clim. Change. 2012;111:665–696. DOI: 10.1007/s10584-011-0173-y
7. Alam M, Regmi BR. Adverse Impacts of Climate Change on Development of Nepal: Integrating Adaptation into Policies and Activities. Bangladesh Centre for Advance Studies (BCAS). Working Paper 3; 2004.
8. Sharma RH, Shakya NM. Hydrological changes and its impact on water resources of Bagmati watershed, Nepal. Journal of Hydrology. 2006;327(3-4):315–322.
9. Wigena AH, Djuraidah A. Quantile regression in statistical downscaling to estimate extreme monthly rainfall. Science Journal of Applied Mathematics and Statistics. 2014;2(3):66-70.
10. Mishra BK, Herath S. Climate projections downscaling and impact assessment on precipitation over upper Bagmati River basin, Nepal. Third International Conference on Addressing Climate Change for Sustainable Development through Up-Scaling Renewable Energy Technologies; RETRUD-11 Conference Proceedings; Kathmandu, Nepal; 2011.
11. Fealy R, Sweeney J. Statistical downscaling of precipitation for a selection of sites in Ireland employing a generalised linear modeling approach. International Journal of Climatology. 2007;27:2083–2094. DOI: 10.1002/joc.1506
12. Yarnal B, Comrie AC, Frakes B, Brown DP. Developments and prospects in synoptic climatology. International Journal of Climatology. 2001;21:1923–1950. DOI: 10.1002/joc.675
13. Von Storch H, Zorita E, Cubasch U. Downscaling of global climate change estimates to regional scales: An application to Iberian Rainfall in wintertime. Journal of Climate. 1993;6:1161–1171.
14. Wigena AH. Modeling of statistical downscaling using projection pursuit

regression for forecasting monthly rainfall. Bogor Agricultural University, Indonesia; 2006.

15. Cavazos T, Hewitson BC. Performance of NCEP–NCAR reanalysis variables in statistical downscaling of daily precipitation. Climate Research. 2005;28: 95-107.
16. Sahar Hadi Pour, Sobri Bin Harun, Shamsuddin Shahid. Genetic programming for the downscaling of extreme rainfall events on the East Coast of Peninsular Malaysia. Atmosphere. 2014;5:914-936. DOI: 10.3390/atmos5040914
17. Wilby RL, Dawson CW, Barrow EM. SDSM — a decision support tool for the assessment of regional climate change impacts. Environmental and Modelling Software (in press). Available:http://co-public.lboro.ac.uk/cocwd/SDSM
18. Chaulagain NP. Impacts of climate change on water resources of Nepal: The physical and socio-economic dimensions. University of Flensburg, Germany; 2006.
19. Chaudhary P, Aryal KP. Global warming in Nepal: Challenges and politics imperatives. Journal of Forest and Livelihood. 2009;8(1):4-13.
20. Babel MS, Bhusal SP, Wahid SM, Agarwal A. Climate change and water resources in the Bagmati River basin, Nepal. Theoretical and Applied Climatology; 2013.
21. Shrestha S, Khatiwada M, Babel MS, Parajuli K. Impact of climate change on river flow and hydropower production in Kulekhani hydropower project of Nepal. Environmental Process. 2014;1:231-50.
22. Shrestha RM, Shrestha SL &Sthapit, A. Logistic model as a Statistical Downscaling Approach for Forecasting a Wet or Dry day in the Bagmati River Basin. The Open Atmospheric Science Journal. 2015;9:9-22.
23. McCullagh P, Nelder JA. Generalized linear models, 2nd edn. Chapman & Hall, London; 1989.
24. Chandler RE, Wheater HS. Analysis of rainfall variability using generalized linear models: A case study from the west of Ireland. Water Resource Research. 2002; 38:1192. DOI: 10.1029/2001WR000906
25. Yan Z, Bate S, Chandler RE, Isham V, Wheater HS. An analysis of daily maximum wind speed in North Western Europe using generalized linear models. Journal of Climatology. 2002;15:2073–2088.
26. Prasad K, Dash SK, Mohanty UC. A logistic regression approach for monthly rainfall forecasts in meteorological subdivisions of India based on DEMETER retrospective forecasts. International Journal of Climatology. 2009;30:1577-1588.
27. Filho WLFC, Lucio PS, Spyrides MHC. Precipitation extremes analysis over the Brazilian Northeast via logistic regression. Scientific Research. 2014;4:53-59.
28. Leith N. Using generalised linear models to stimulate daily rainfall under scenarios of climate change. Technical Report. London, UK: University College London; 2005.
29. Li X, Sailor D. Application of tree structured regression for regional precipitation prediction using general circulation model output. Climate Res 2000;16:17-30.
30. Khalil MAK, Zhao W. The relationship between precipitation and temperature over the contiguous United States. Journal of Climate. 1993;6:1232-36.
31. Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, et al. The NCEP/NCAR 40-year reanalysis project. Bulletin of the American Meteorological Society; 1996.
32. Ghosh S. Hydrologic impact of climate change: uncertainty modeling. Indian Institute of Science, Bangalore, India; 2007.
33. Kutner MH, Nachtsheim CJ, Neter J, Li William. Applied linear statistical models; McGraw-Hill Education India Edition Pty Ltd: New Delhi, India; 2013.
34. Agresti A. Categorical data analysis. John Wiley & Sons Inc: New Jersey, USA; 2002.
35. Gerapetritis H. A probability of precipitation equation for Columbia, South Carolina derived from logistic regression. Eastern Region Technical Attachment 1999;99:1.
36. Roy M, Imon A, Bhattacharjee S. Prediction of rainfall using logistic regression. Pakistan Journal of Statistics and Operation Research. 2012;8(3):655-667.
37. Shrestha MN. Hydrology and environmental perspective of Bagmati

River basin. Journal of Water, Sanitation, Health and Environment, Nepal; 2009.
38. Trenberth KE, Shea DJ. Relationships between precipitation and surface temperature. Geophysical Research Letters. 2005;32:L14703. DOI: 10.1029/2005GL022760
39. Kuitel H, Hirsch-Eshkol TR, Türkeş M. Sea level pressure patterns associated with dry or wet monthly conditions in Turkey. Theoretical and Applied Climatology. 2001; 69:39-67.