Scientific Journal of Impact Factor (SJIF): 4.72

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 4, Issue 7, July -2017

MODELING THE PRECIPITATION AMOUNTS DYNAMICS FOR DIFFERENT TIME SCALES IN SRINAGAR CITY, J&K-INDIA USING MULTIPLE LINEAR REGRESSION APPROACH

Mehnaza Akhter

Department of Civil Engineering, NIT Srinagar, J&K, India

Abstract: *Water resources are very important for ecosystems and water deficit may cause serious social and economic issues. The aim of this study is to analyze the performances ofprediction procedure based on Multiple Linear Regression Model (MLRM), for the precipitation amounts for yearly and monthly time scales, in Srinagar City located in Jammu and Kashmir, India. For this purpose we have used as predictand monthly amounts of precipitation and as predictors Mean Sea level pressure(mslpas),Mean temperature at 2m(tempas),Specific humidity at 2m (humas), Geopotential height at 500 hPa (p500-as), Surface zonal velocity(u) and Surface meridional velocity(v). The selection of predictors is based on the backward elimination or predicative (p-test) analysis.All data sets used in this study are gridded data with a spatial resolution of 2º x 2º lat/lon and are obtained from Canadian Global Climate model version 3(CGCM3) .The analysis is made for a period of 45 years between1970 and 2015, the period 1970– 2010 being used to build the MLRM and the period 2011 – 2015 for testing the prediction performances of the MLRM. Using MLRM we have obtained some good correlation between predicted and measured precipitation amount. The correlation coefficient varies between 0.57 and 0.84, with the smallest values in winter and the greatest values in spring. The total annual precipitation shows a decrease of 41.42% at the end of 21st century for Srinagar city.*

Keywords: *Precipitation amount prediction, Multiple linear regression,*

1. Introduction: Precipitation is a principal element of the hydrological cycle, and changes inprecipitation pattern is important because it may lead to floods or droughts events that may cause economic loss and casualties. Frequency of precipitation plays an important role in the management of agriculture, water resources and ecosystems, and the time scales of the precipitation variability varies from months to years or decades. Identifying and understanding the influence of the large-scale air circulation patterns which produce temporal and spatial variations of precipitation in Srinagar city is therefore of a great importance. In addition to global warming, on regional scale, several other factors may determine future climate change such as variation in air circulation and topography [1]. Variations in air circulation influence the climate in large areas both on interannual and longtime scales while the topography modifies the effects of air circulation on local scale [2]. Precipitation studies have been made for various time periods and on various spatial scales: global [3], hemispheric [4], regional [5] and local [6]. Various methods are used to analyze spatial and temporal variability of precipitation, for example: trend and change point analysis [8, 4], cluster analysis [9, 10], Empirical Orthogonal Function(EOF) analysis [11,6], canonical correlation analysis [11], and multiple linear regressions [12, 13, 14]. The multiple linear regression approach was developed and successfully applied several decades ago to the specification of surface temperature from midtroposphere circulation, mainly for the purpose of weather prediction, by Klein [15]. The multiple linear regression model was also used in several studies of atmospheric physics, such as pollutant dynamics [16, 17], the estimation and prediction of atmospheric concentrations of the natural occurring radionuclides,radon and thoron [18], and to estimate and predict the radiative forcing of the clouds [19].The aim of this paper is to develop a rainfall predictive model on a yearly and monthly time scale. In order to achieve this, the multiple linear regression model(MLRM) is used. The MLRM was build up by using as predictors Mean sea level pressure(mslpas), Mean temperature at 2m (tempas),Specific humidity at 2m(humas), Geopotential height at 500 hPa (p500-as), Surface zonal velocity(u) and Surface meridional velocity(v). Backward elimination or predicative analysis (p-test) was done between individual predictors and predictand to identifypredictors that are most correlated with precipitation. After that, the selected predictors were used to build up the MLRM for a period of 45 years between 1970 and 2015. Therefore the analytical equation obtained by building up the MLRM was used to predict the precipitation amount for yearly and monthly time scales for 10 years, during the period 2010–2015. The performances of the MLRM were tested by statistical parameters R^2, Root Mean Square Error (RMSE) and Mean Absolute Deviation (MAD).

Study area

The study area comprises of Srinagar city with gps coordinates of 34° 5' 1.1616" N (latitude) and 74° 47' 50.5356" E(longitude). Srinagar is the largest city and the summer capital of the Indianstate of Jammu and Kashmir. It lies in the Kashmir Valley on the banks of the Jhelum River, a tributary of the Indus, and Dal and Anchar lakes. The city is famous for its gardens, waterfronts and houseboats. It is also known for traditional Kashmiri handicrafts and dried fruits. Srinagar has a humid subtropical climate, much cooler than what is found in much of the rest of India, due to its moderately high elevation and northerly position. The valley is surrounded by the Himalayas on all sides. Winters are cool, with daytime temperature averaging to 2.5 °C (36.5 °F), and drops below freezing point at night. Moderate to heavy snowfall occurs in winter and the highway connecting Srinagar with the rest of India faces frequent blockades due to icy roads and avalanches. Summers are warm with a July daytime average of 24.1 °C (75.4 °F). The average annual rainfall is around 720 millimetres (28 in). Spring is the wettest season while autumn is the driest. The highest temperature reliably recorded is 38.3 °C (100.9 °F) and the lowest is −20.0 °C (−4.0 °F)[31].

Figure 1. Location map of Srinagar city

Data

The data for the present study has been obtained from the Central Research Institute Metrology, Pune and Regional Meteorological Station, Srinagar which is centrally located. The study is based on 45 years (1970-2015). This data, which was available for a period of 45 years (1970-2015) for Srinagar city, included: Monthly precipitation data at Srinagar metrological station of Kashmir valley. To project the future climate of the valley GCM predictor data was required which was made available from Canadian Climate Data and Scenarios (CCDS) website http://ccds-dscc.ec.gc.ca/ and www.cccsn.ec.gc.ca/.

All data are built up for a period of 45 years, during 1970 – 2015. Data weredivided into two groups: the first group is a period of 40 years between 1970 and 2010 and the second is a period of 5 years between 2011 and 2015. The firstgroup was used to build MLRM, and the second to test the prediction performancesof the MLRM. For this study we have been used the same data base source forpredictors and predictands to be consistent with the multiple linear regressionmethod.

Methodology

In order to predict the annual precipitation amount the multiplelinear regression model (MLRM) has been used. MLRM is a statistical techniquethat is used to model a linear relationship between a dependent variable(predictand), a continuous variable and one or more independent variables(predictors) and assumes a linear relationship between variables [24].

The estimation procedure is made by using the analytical expression of themultiple linear regression with selected predictors. Therefore, the predictand y(t) isassociated with the predictors $\{ x_i(t)\}_{i=1,2,\ldots,n}$ by the following relationship [25]:

$y(t) = \beta_0 + \beta_1 x_1(t) + \beta_2 x_2(t) + \ldots + \beta_n x_n(t)$ (1)
where β0 is the regression constant, β1…βn are the regression coefficients.
The performances of the MLRM are quantified by the following statisticparameters:
• Multiple correlation coefficient (R) – is a measure of correlation betweenpredictand and predictors;
• Squared multiple correlation coefficient (R^2) – quantifies the proportionof variance of the predictand that is explained by the predictors;

•Root Mean Square Error (RMSE) – Alternatively known as the Root Mean Square Deviation (RMSD), the RMSE is commonly used to measure the difference between the values predicted by the model and the values observed.
•Mean Absolute Deviation (MAD) – The Mean Absolute Deviation (MAD) of a set of data is the average distance between each data value and the mean.
• p-value – characterizes the significance level so that the null hypothesisto be rejected or accepted against the alternative hypothesis. The p-valueis computed by using the t-Student distribution for the regressioncoefficients and F distribution for the F-Test for the overall model. Thep-value less than 0.05 indicate that the null hypothesis may be rejectedand the alternative one has to be accepted.
The model having minimum RMSE, minimum MAD and maximum R^2 is considered as the best model.In the MLR model CGCM3 was used for projection of precipitation in Srinagar city upto 21st century. General Circulation Models (GCMs) discretise the equations for fluid motion and energy transfer and integrate these over time. Unlike simpler models, GCMs divide the atmosphere and/or oceans into grids of discrete "cells," which represent computational units. Unlike simpler models which make mixing assumptions, processes internal to a cell—such as convection—that occur on scales too small to be resolved directly are parameterised at the cell level, while other functions govern the interface between cells.

Predicative test analysis

The selection of predictors is an iterative process based partly on user'ssubjective judgment. The two main requirements for predictors are: a goodrelationship with predictant, and a reasonable temporal extension. In order to selectthe independent predictors to build the MLRM thepredicative (p-test) analysis was performed.

Backward elimination or predicative analysis (p-test) was done between individual predictors and predictand to identifypredictors that are most correlated with temperature and precipitation. Table 1 shows the list of predictors that were selected for CGCM3. When there are many possible predictors, we need some strategy to select the best predictors to use in a regression model.

A common approach is to do a multiple linear regression on all the predictors and disregard all variables whose *p*-values are greater than 0.05.

Table 1: The regression coefficients and probability of factor (p value) influencing the output

Predictor	Coefficient	P
Constant	0.766936	0.00000
X_1	0.20463	0.21233
X_2	0.653565	0.00010
X_3	0.0566	0.081377
X_4	0.090785	0.028717
X_5	0.08047	0.018681
X_6	0.02558	0.567739
X_7	0.047642	0.264882
X_8	0.595344	0.000010
X_9	0.01239	0.766736
X_{10}	0.151002	0.001972
X_{11}	0.545084	0.000209

Thus predictors X_2,X_4,X_5,X_8,X_{10} and X_{11} which correspond to tempas,humas,u,v,p500-as and mslpas were choosen .Remaining predictors were discarded as they gave p value more than 0.05.Vorticity,divergence, Strength of the resultant flow, Surface airflow strength and Surface wind direction which correspond to X_1,X_3,X_6,$X_7$and X_9 respectively were eliminated as they showed p value exceeding 0.05.

Table 2: Selected predictor variables.

Model	Predictand	Predictor variables	Symbol
CGCM3	Precipitation	Mean temperature at 2m	Tempas
		Geopotential height at 500 hPa	p500-as
		Surface zonal velocity	U
		Surface meridional velocity	V
		Mean sea level pressure	Mslpas
		Specific humidity at 2m	Humas

Results and Discussions

Annual precipitation amounts were estimated by using multiplelinear regression software develop by Wessa [29]. The regression coefficients aredetermined by using the least square method. The estimation of annual amount of precipitation is performed by buildup the MLRM for 45 yearsduring the period 1970–2015 using as predictors Mslpas, tempas, humas, u, v and p500-as. Therefore the analytical expressions obtained were used to estimate theprecipitation amount for yearly time scales for the same period thatwas used to build up the MLRM.

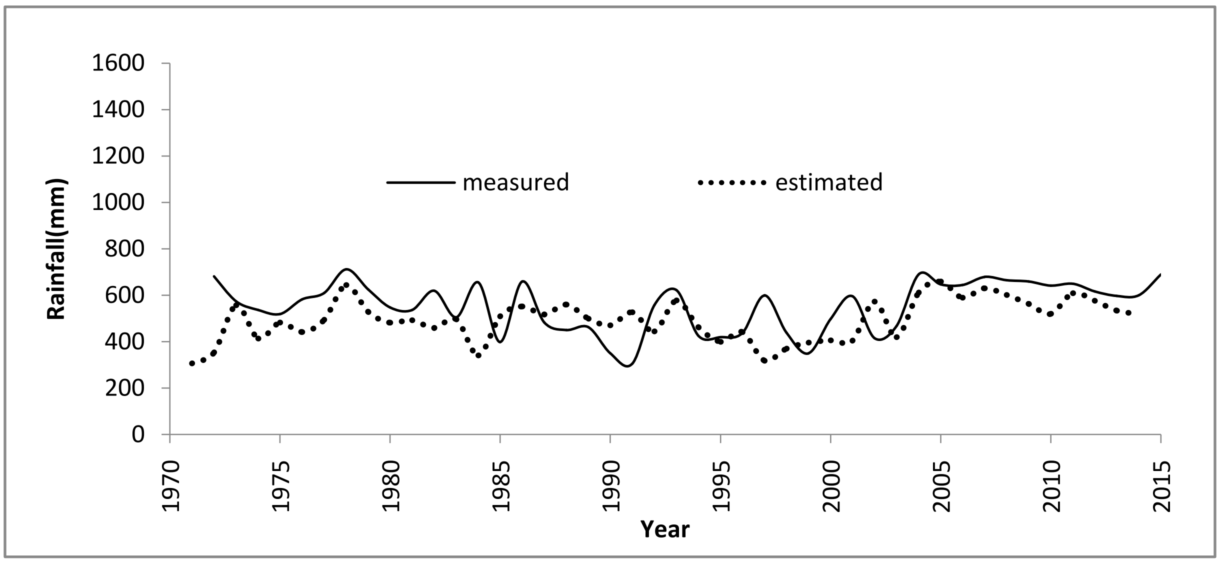


Figure 2. Estimated and measured annual amount of precipitation (PP) for the period 1970-2015 for Srinagar city.

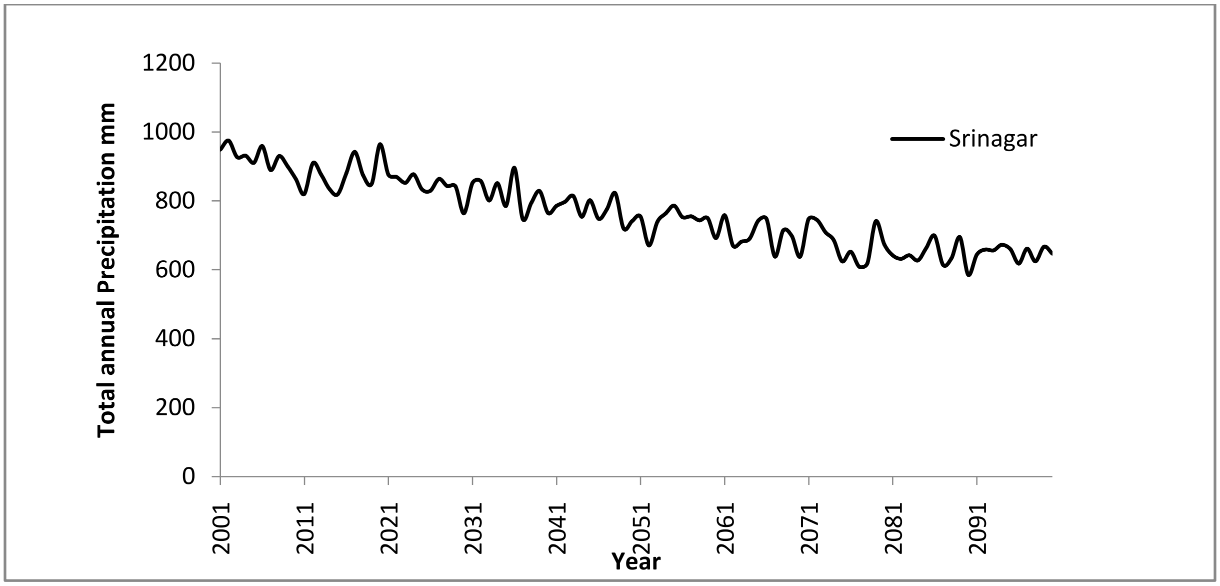


Figure 3. Variation of total annual precipitation for 21st century for Srinagar city using CGCM3 model.

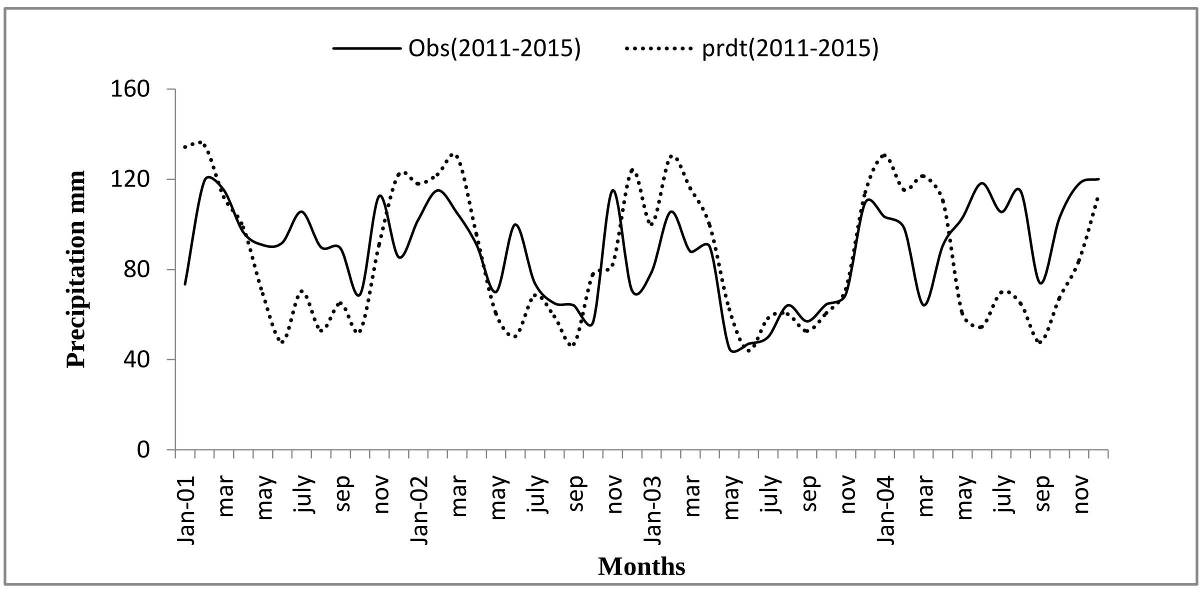


Figure 4. Validation of monthly total precipitation of Srinagar city for the period 2010-2015 using MLR.

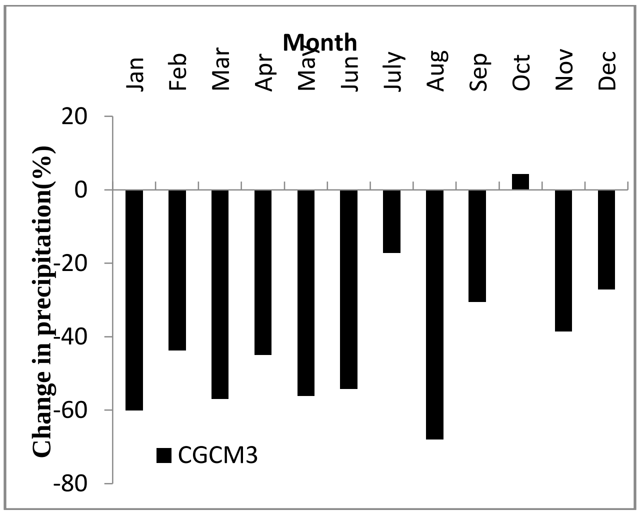
Figure 2presents the estimated and measured precipitation amount from 1970-2015 for Srinagar city. From Figure 2 one can clearly see that there is strong correlation between estimated and measured values of precipitation amount. Figure 3 shows variation of total annual precipitation from 2001 to 2100 for Srinagar city using CGCM3 model. From this figure it is observed that total annual precipitation for Srinagar city is going to decrease by 41.42% at the end of 21st century.

From Figure 4 it is clear that the observed and predicted values of precipitation varied in the same direction throughout the validation period.

In Table 3 the regression statistics for the estimation of annual precipitation amount for Srinagar city is presented. One may notice that the multiple correlation coefficients (R^2) have the closer values for all months. The R value is the greatest for summer months (0.84), and the smallest value of R (0.57) was obtained for the month of December.

Table 3: Statistical parameters of MLR validation for precipitation using CGCM3 model

Month	R^2	MSE	RMSE	MAD
January	0.64	4.658125	2.158269	-1.73750
February	0.67	8.467225	2.909850	-2.76750
March	0.70	3.735528	1.932751	-1.92655
April	0.76	8.556900	2.925218	-2.73500
May	0.72	6.361152	2.522132	-2.49370
June	0.77	6.936034	2.633635	1.03087
July	0.84	0.284225	0.533128	-0.34250
August	0.75	9.275193	3.045520	0.22110
September	0.70	3.986950	1.996735	0.03140
October	0.64	10.76250	3.280625	-2.82500
November	0.61	5.405000	2.324866	-2.20000
December	0.57	9.685000	3.112073	-2.65000



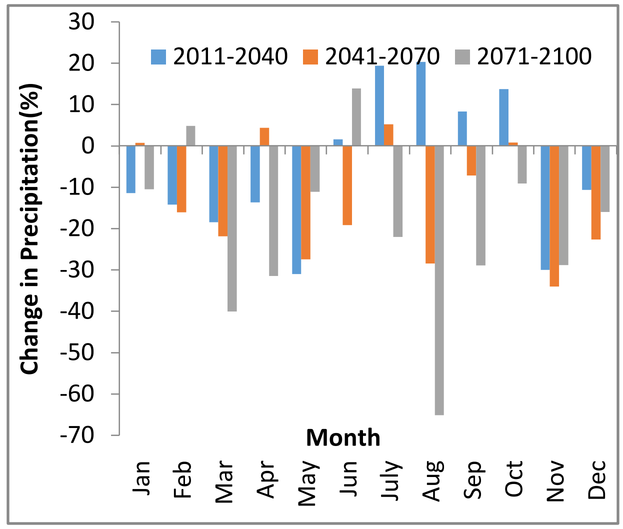


Figure 5. Percentage change in monthly precipitation for Srinagar cityat the end of 21st century from 2010using MLR.

Conclusions

1) The MLR modeling result showed good agreement between the observed data and predicted values with a good coefficient of determination, R^2=0.57-0.84 for precipitation. Hence GCMs predict precipitation to a much good accuracy.
2) The total annual precipitation decreased by 41.42% for CGCM3 model over 21st century in Srinagar city using MLR technique.

REFERENCES

[1] J. H. Christensen, B. Hewiston, A. Busuioc, A. Chen, X. Gao, I. Held, R. Jones, R. K. Kolli,W. T. Kown, R. Laprise., V. Magana Rueda, L. Mearns, C. G. Menéndez., J. Räisänen,A Rinke, A. Saar and P. Whetton, Regional Climate Projections. In: Climate Change 2007:The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Reportof the Intergovernmental Panel on Climate Change [Solomon S., Qin D., Manning M., ChenZ., Marquis M., Averyt K.B., Tignor M. and Miller H.L. (eds.)]. Cambridge University Press,Cambridge, United Kingdom and New York, NY, USA (2007).

[2] R. Bojariu and F. Giorgi, The North Atlantic Oscillation signal in a regional climate simulation forthe European region. Tellus, **57A**:641–653 (2005).

[3] H. F. Diaz, R.S. Bradley and J.K. Eischeid, Precipitation Fluctuation Over Global Land Areas Since the Lates 1800's, J. Geoph. Res. **94D1:**1195–1210 (1989).

[4] R. S. Bradley, H. F. Diaz, J. K. Eischeid, P.D. Jones, P. M. Kelly and C. M. Goodess, PrecipitationFluctuation over Northern Hemisphere Land Areas Since the Mid-19th Century, Science(Reprint Series), **237**:171–175 (1987).

[5] C. D. Schönwiese, J. Rapp, T. Fuchs and M. Denhard, Observed climate trends in Europe 1891–1990.Meteorol Z NF **3**:22–28 (1994).
[6] A. Busuioc and H. von Storch, Changes in the winter precipitation in Romania and its relation tothe large scale circulation. Tellus **48A**:538–552 (1996).
[7] L. Buffoni, M. Maugeri and T. Nanni, Precipitation in Italy from 1833 to 1996. Theor. Appl. Climatol. **63**:33–40 (1999).
[8] A. N. Pettitt, A non-parametric approach to the change-point problem. Applied Statistics.126–135(1979).
[9] G. Galliani and F. Filippini, Climate clusters in a small area. J. Clim. **3**:47–63 (1985).
[10] C. Cacciamani, S. Tibaldi and S. Nanni, Mesoclimatology of winter temperature and precipitationin the Alpine Region of Northern Italy. Int. J. Climatol. **14**:777–814 (1994).
[11] H. von Storch, Spatial patterns: EOFs and CCA. In: von Storch H, Navarra A (eds) Analysis ofclimate variability: applications of statistical techniques. Springer, Heidelberg, 227–258(1995).
[12] M. C. Valverde Ramíres, N. J. Ferreira and H. F. de Campos Velho, Linear and Nonlinear Statistical Downscaling for Rainfall Forecasting over Southeastern Brazil. Weather andForecasting. **21**:969–989 (2006).
[13]M. D. Mizanur Rahman, M. Rafiuddin and M. D. Mahbub Alam, Seasonal forecasting of Bangladesh summer monsoon rainfall using simple multiple regression model. J. Earth Syst. Sci. **122(2):**551–558 (2013).
[14] R. Chifurira and D. Chikobvu, A Weighted Multiple Regression Model to Predict Rainfall Patterns: Principal Component Analysis approach. Mediterranean Journal of Social Sciences, **5(7):**34–42 (2014).
[15]W. H. Klein, Specification of monthly mean surface temperatures from 700 mb heights. J. Appl.Meteor. **1**:154–156 (1962).
[16]M. C. Hubbard and W. G. Cobourn, Development of a regression model to forecast ground-levelozone concentration in Louisville, KY. Atmos. Environ. **32(14–15)**:2637–2647 (1998).
[17]A. Vlachogianni, P. Kassomenos, A. Karppinen, S. Karakitsios and J. Kukkonen, Evaluation of amultiple regression model for the forecasting of the concentrations of NOx and PM10 in Athens and Helsinki. Science of the Total Environment. **409:**1559–1571 (2011).
[18]F. Simion, V. Cuculeanu, E. Simion and A. Geicu, Modeling the ^{222}Rn and ^{220}Rn progeny concentration in atmosphere using multiple linear regression with meteorological variables as predictors. Rom. Rep. Phys. **65**, 524–544 (2013).
[19]V. Cuculeanu, I. Ungureanu and S. Stefan, Study of the relationship among radiative forcing, albedo and cover fraction of the clouds. Rom. Journ. Phys. **58**, 987–999 (2013).
[20]A. Busuioc, Large-scale mechanisms influencing the winter Romanian climate variability,Detecting and Modelling Regional Climate Change and Associated Impacts, M. Brunet andD. Lopez eds., Springer-Verlag, 333–343 (2001).
[21] R. A. Houze, Cloud dynamics. Academic Press, Inc, San Diego (1993).
[22] G. P. Compo, Whitaker J.S., Sardeshmukh P.D, Matsui N., Allan R.J., Yin X., GleasonB.E., Vose R.S., Rutledge G., Bessemoulin P.,Brönnimann S., Brunet M., CrouthamelR.I., Mok H.Y.,Nordli O., Ross T.F., Trigo R.M., Wang X.L.,Woodruff S.D. andWorley S.J.,The Twentieth Century Reanalysis Project, Q. J. R. Meteorol. Soc. **137:**1–28. doi:10.1002/qj.776 (2011).
[23] G. P. Compo, J. S. Whitaker and P. D. Sardeshmukh, Feasibility of a 100 year reanalysis usingonly surface pressure data. Bull. Am. Meteorol. Soc. **87:**175–190, doi:10.1175/BAMS-87-2-175 (2006).
[24] P. Aksornsingchai and C. Srinilta, Statistical Downscaling for Rainfall and TemperaturePrediction in Thailand, Proceedings of the International MultiConference of Engineers andComputer Scientists, 2011, March 16 – 18, 2011, Hong Kong (2011).Grant A.N., Groisman P.Y., Jones P.D., Kruk M.C.,Kruger A.C., Marshall G.J.,MaugeriM.,
[25] D.S. Wilks, Forecast verification. Statistical Methods in the Atmospheric Sciences, AcademicPress, p467 (1995).
[26]K. Pearson, Mathematical contributions to the theory of evolution. III. Regression, heredity, andpanmixia. Philosophical Transactions of the Royal Society Ser. A **187**:253–318 (1896).
[27]C. E. Spearman, The proof and measurement of association between two things. Am. J. Physiol.**15**:72–101 (1904).
[28]P. A. Rogerson, A Statistical Method for the Detection of Geographic Clustering. Geogr Anal **33(3):**215–227 (2001).
[29]P. Wessa, Free Statistics Software, Office for Research Development and Education, version 1.1.23-r7, URL http://www.wessa.net/ (2014).
[30]M. Rajeevan, D.S. Pai and R. Anil Kumar, New statistical models for long range forecasting ofsouth-west monsoon rainfall. Climate Dynamic. **28(7–8)**:813–828, doi:10.1007/s00382-006-0197-6 (2006).
[31]www.indpune.gov.in.

