



Quality of local scale surface weather analogs over the north-west Himalaya (NWH), India

DAN SINGH^{1,*}, ASHAVANI KUMAR² and M S SHEKHAR¹

¹*Research and Development Centre, Snow and Avalanche Study Establishment, Sector 37 A, Chandigarh 160 036, India.*

²*Department of Physics, National Institute of Technology Kurukshetra, Kurukshetra, India.*

*Corresponding author. e-mail: dan_@rediffmail.com

MS received 18 August 2018; revised 22 April 2019; accepted 26 April 2019

An analog ensemble system was developed for the realisation of local-scale surface meteorological variables for independent test data (test data) at six stations over the north-west Himalaya (NWH), India. Extreme values (the maximum value and the minimum value) and the mean value in 10 analog days (the analog mean) and the climatological mean of each surface meteorological variable were compared with its corresponding observed values on the same day (d_0 , lead time 0 hour (h)), d_1 ($d_0 + 1$, lead time 24 h), d_2 ($d_0 + 2$, lead time 48 h) and d_3 ($d_0 + 3$, lead time 72 h) of test data. Pearson correlation coefficients (CCs), Mean Absolute Differences (MADs) and Root Mean Square Differences (RMSDs) of the extreme values in analog days, and the analog mean and climatological mean of each meteorological variable on d_0 with its corresponding observed values on d_0 , d_1 , d_2 and d_3 of test data were computed at six stations over the NWH. CCs of extreme values in analog days and the analog mean of each meteorological variable on d_0 with its observed values on d_0 , d_1 , d_2 and d_3 were found to be higher than the CCs of the climatological mean of each meteorological variable on d_0 with its observed values on d_0 , d_1 , d_2 and d_3 . MADs (RMSDs) of extreme values in analog days and the analog mean of each meteorological variable on d_0 with its observed values on d_0 , d_1 , d_2 and d_3 were found to be lesser than the MADs (RMSDs) of the climatological mean of each meteorological variable on d_0 with its observed values on d_0 , d_1 , d_2 and d_3 . However, the MADs (RMSDs) of the extreme values of each meteorological variable in analog days were found to be higher than the MADs (RMSDs) of its analog mean. Results show that the analog mean of each meteorological variable holds better predictive skill than the extreme values in analog days and its climatological mean. MADs (RMSDs) of different surface meteorological variables in surface weather analogs comparable to Mean Absolute Errors (MAEs) and Root Mean Square Errors (RMSEs) for their prediction with the help of different types of weather forecast models show that the surface weather analogs hold good promise for the local-scale prediction of surface meteorological variables over the NWH.

Keywords. Weather; weather forecasting; analog ensemble system; north-west Himalaya.

1. Introduction

Ensemble forecasting is operationally being used for a variety of meteorological, hydrological and climatological applications (Toth and Kalnay 1993;

Tracton and Kalnay 1993; Hamill and Colucci 1997; Buizza *et al.* 1999; Zorita and Storch 1999; Whitaker *et al.* 2006; Jaun and Ahrens 2009; Raje and Mujumdar 2011). The analog method is one of the methods employed to develop ensemble

forecasting systems (Hamil and Whitaker 2007; Daoud *et al.* 2011; Charles *et al.* 2013; Delle Monache *et al.* 2013 and references herein). A variety of analog ensemble (AE) systems have been developed primarily following two approaches, i.e., searching for analogs corresponding to a single meteorological field (or variable) of interest and searching for analogs corresponding to multiple meteorological fields (variables) of interest (Lorenz 1969; Radinovic 1975; Barnett and Preisendorfer 1978; Van Den Dool 1989; Hamil and Whitaker 2006; Gibergans-Baguena and Llasat 2007; Bannayan and Hoogenboom 2008; Hall *et al.* 2010; Singh *et al.* 2010; Nagarajan *et al.* 2015).

AE systems have been developed for the prediction of various weather elements such as temperature, precipitation amount, wind speed and solar radiation, etc., statistical downscaling and/or post-processing of predictions of various meteorological variables with the help of numerical weather prediction (NWP) models (Van Den Dool 1989; Zorita and Storch 1999; Sievers *et al.* 2000; Obled *et al.* 2002; Hamil and Whitaker 2006; Bannayan and Hoogenboom 2008; Singh *et al.* 2008; Delle Monache *et al.* 2013; Junk *et al.* 2015; Nagarajan *et al.* 2015). AE systems developed as statistical downscaling and/or post-processing tools for the NWP model predictions are found to show better skills than parent NWP models for various meteorological variables like temperature, precipitation, wind speed, etc. (Hamil and Whitaker 2007; Delle Monache *et al.* 2013; Junk *et al.* 2015; Nagarajan *et al.* 2015). However, applications of the AE system as the post-processing and/or statistical downscaling technique demand historical archives of NWP predictions from the frozen state-of-the-art NWP models and corresponding surface meteorological observations (Hamil and Whitaker 2006, 2007; Delle Monache *et al.* 2013; Junk *et al.* 2015). Such requirements of the AE system and other statistical downscaling techniques may limit their effective and efficient use for data-sparse mountainous regions like the north-west Himalaya (NWH) (or elsewhere) (Das 2005; Hatwar *et al.* 2005; Singh and Ganju 2005; Dimri 2006; Dimri and Niyogi 2013; Dimri *et al.* 2013). This can be attributed to many factors such as the lack of adequate surface and an upper air observational network, spatially-temporally inhomogeneous meteorological observations and the lack of a long historical archive of the NWP model predictions at a high spatial resolution and corresponding surface meteorological observations over the NWH (Das

et al. 2003; Mohanty and Dimri 2004; Dimri and Mohanty 2007; Dimri and Niyogi 2013; Menegoz *et al.* 2013; Palazzi *et al.* 2013; Fiddes and Gruber 2014; Terzago *et al.* 2014; Tiwari *et al.* 2014).

Large spatio-temporal variability in various surface weather elements over the NWH demands real-time local-scale weather forecasts (Mohanty and Dimri 2004; Dimri and Mohanty 2007; Singh *et al.* 2008, 2010, 2019; Srinivasan *et al.* 2010; Joshi and Ganju 2012; Joshi and Ganju 2013). The AE system developed utilising local-scale surface meteorological observations can be of great help for the real-time local-scale prediction of various surface meteorological variables for data-sparse mountainous regions due to its ability to find reasonably good analogs, capability to capture a flow-dependent error growth and address forecast uncertainty (Barnett and Preisendorfer 1978; Van Den Dool 1989; Bannayan and Hoogenboom 2008; Delle Monache *et al.* 2013). However, this can be ensured only via a systematic study of whether the AE system developed utilising the local-scale surface meteorological observations is capable of realising different surface meteorological variables and their temporal evolutions in analog days, and whether it is adaptable to the local-scale climatic conditions of the forecast area and the differences (errors) between observed surface meteorological variables on independent test days (test data) and their corresponding values in analog days are reasonable at a local scale (specific site) over the NWH. Such study is possible by finding the resemblance between observed surface meteorological variables and their temporal evolutions on independent test days with the corresponding surface meteorological variables in analog days. A good resemblance between observed surface meteorological variables and their temporal evolutions on test days with the corresponding surface meteorological variables in analog days can provide evidence that the AE system holds promise for local-scale prediction of various surface meteorological variables over the NWH (or elsewhere). Hence it is extremely important to study the resemblance between local-scale surface meteorological variables and their temporal evolutions on independent test days with the corresponding surface meteorological variables in analog days to gain an insight into and an understanding of the quality of local-scale surface weather analogs over the NWH and develop an AE system for the prediction of various surface meteorological variables at the local scale over the NWH.

AE systems have been developed for qualitative weather and quantitative categorical snowfall forecasts and prediction of a few surface meteorological variables at a specific site over the NWH (Singh *et al.* 2008, 2010). Singh *et al.* (2008) focus on the prediction of a few surface meteorological variables at a single station location over the NWH. However, it does not provide a comprehensive study on the quality of surface weather analogs and errors associated with various surface meteorological variables in analog days at different stations over the NWH. Hence the objectives of this study are to find whether the AE system developed utilising local-scale surface meteorological observations is capable of realisation of various surface meteorological variables and their temporal evolutions in analog days for independent test data, whether it is adaptable to the local climatic conditions of the forecast area and examine and compare the differences (errors) associated with the local-scale surface weather analogs and climatological mean of various surface meteorological variables at six stations over the NWH. This study focuses and addresses the predictability associated with the surface weather analogs (the AE system) at six stations spread across different mountain ranges over the NWH. The configuration of the AE system and a set of variables taken to conduct this study are significantly different from an earlier study (Singh *et al.* 2008). The findings of this study can be useful to gain an insight into and an understanding of local-scale surface weather analogs and to develop a real-time local-scale weather forecast model for high-altitude remote areas of the NWH.

2. Study area and data

This study was conducted at six stations falling within the altitude range 2192–3800 m above mean sea level and these stations are located in the Shamsawari range, the Pir Panjal range and the Great Himalaya range in the NWH (figure 1). A major component of winter precipitation (November–April) over the NWH is in solid form, i.e., snowfall (Dimri and Das 2011; Singh *et al.* 2015). Heavy snowfall during winter causes interruptions in necessary services and supplies, electricity and communication failures and generates many weather-related natural hazards such as severe cold temperatures, snow avalanches and landslides, etc. over the NWH (Rangachary and Bandyopadhyay

1987; Das *et al.* 2003; Singh and Ganju 2005). Therefore, local-scale weather forecasts play a very important role in planning day-to-day human activities and the mitigation of weather-generated natural hazards over the NWH during winter (Rangachary and Bandyopadhyay 1987; Hatwar *et al.* 2005; Srinivasan *et al.* 2005; Singh *et al.* 2008). Large spatio-temporal variability in various surface meteorological variables over the NWH demands real-time local-scale weather forecasts (Mohanty and Dimri 2004; Dimri and Mohanty 2007; Singh *et al.* 2008). The AE system developed based on *in-situ* surface meteorological observations can be of great help for the local-scale prediction of surface meteorological variables over the NWH. However, a systematic study assessing the quality of local-scale surface weather analogs over the NWH is lacking.

Snow and Avalanche Study Establishment (SASE), India, has data archive of surface meteorological observations collected daily at 0300 and 1200 UTC for more than two decades over the NWH (Dimri and Das 2011; Singh *et al.* 2015). Surface meteorological observations on variables listed under table A1 at each station were obtained to conduct this study (table 1). More than 97% observations on different surface meteorological variables fall within the range mean \pm 3 standard deviations; hence the data were assumed to be suitable to conduct this study (Singh *et al.* 2015). Cumulative precipitation amount of a 24 h time interval was computed by adding the snow water equivalent of the 24 h time interval with the rainfall amount of the 24 h time interval.

Surface meteorological observations on variables listed under appendix A1 at each station were split into two mutually exclusive datasets: base dataset and independent test dataset (test data, test days). The base dataset was utilised to find analog days corresponding to each day of the test dataset. The base dataset consists of surface meteorological observations collected daily at 0300 and 1200 UTC of the past 11–28 winters and the test dataset consists of surface meteorological observations collected at 1200 UTC of the past 10 winters (winter 2006–2015) at six stations over the NWH (table 1).

A large spatial variability in mean maximum air temperature, mean minimum air temperature, mean cumulative precipitation amount in the 24 h time interval of precipitation days (precipitation amount \geq 0.1 mm) and their standard deviations, large differences in the number of records in the

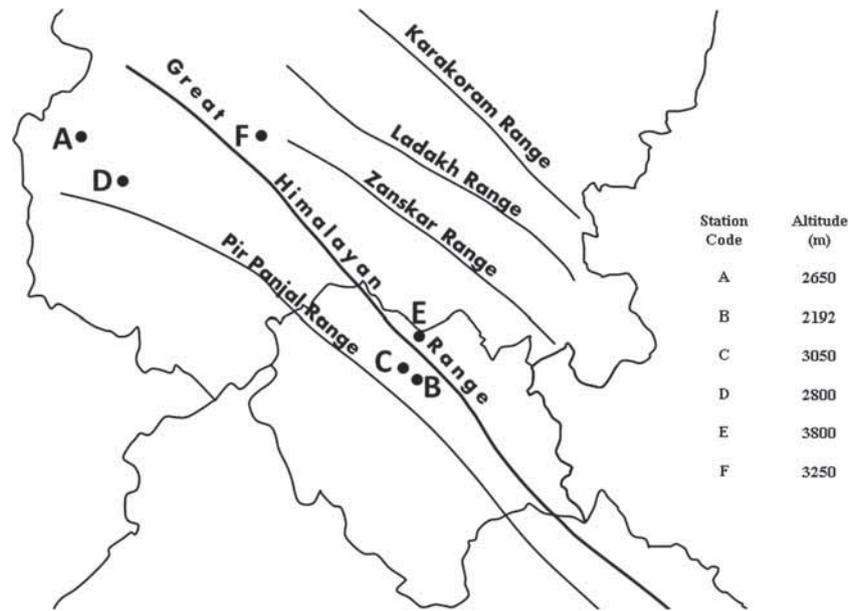


Figure 1. Station locations in different mountain ranges over the Northwest Himalaya (NWH).

Table 1. Mean and standard deviation of maximum air temperature, minimum air temperature, cumulative precipitation amount of 24 hours (h) time interval of precipitation days, number of records in base dataset (0300 and 1200 UTC) and number of days in independent test dataset (test data) (1200 UTC only) at six stations over the NWH (Mtn range = mountain range; Stn = station; Winters = winters to which base dataset belongs; S range = the Shamshawari range; P range = the Pir Panjal range; G range = the Great Himalaya range).

Mtn range	Stn	Winters	No. of records in base dataset	No. of days in test dataset	MX		MN		PT	
					μ	Ω	μ	Ω	μ	Ω
S range	A	1991–2005	4640	1711	6.0	4.8	-2.8	4.1	18.3	19.2
P range	B	1979–2005	8811	1476	14.8	5.0	1.8	3.1	16.0	17.6
	C	1989–2005	5825	1643	8.1	4.9	-1.2	3.8	24.3	27.6
	D	1995–2005	3483	1650	5.0	4.5	-5.4	4.4	15.7	17.6
G range	E	1983–2005	6808	1722	2.1	4.9	-11.9	5.4	8.7	10.2
	F	1995–2005	3612	1543	2.2	5.7	-13.3	7.3	8.0	8.4

base dataset, number of independent test days in the test dataset and temporal resolutions of various surface meteorological variables show that the AE system was tested for a variety of climatic regimes and meteorological conditions over the NWH (tables 1 and A1).

3. Materials and methods

Ten analog days (*hereafter analog days*) corresponding to each test day of the test dataset were searched from the base dataset using Euclidean distance between surface meteorological variables of a test day and surface meteorological variables of the candidate analog days at each station (Barnett and Preisendorfer 1978; Gutzler and Shukla 1984; Toth 1989; Van den Dool 1989; Bannayan and

Hoogenboom 2008; Singh *et al.* 2008, 2010). The following Euclidean distance metric was used to find the analog days from the base dataset corresponding to each test day of the test dataset (Singh *et al.* 2008, 2010; Delle Monache *et al.* 2011):

$$d_{lm} = \left[\sum_{k=1}^p (x_{l,k} - x_{m,k})^2 \right]^{1/2}, \quad (1)$$

where d_{lm} is the Euclidean distance between day l and day m and $x_{l,k}$ and $x_{m,k}$ are the vectors of k ($k = 1$ to p) measurements for day l and day m .

Surface meteorological variables may be partially correlated and the principal components are not used to find the analog days (Storch and Zwiers 1999; Xavier and Goswami 2007). This is due to the partially incomplete surface meteorological

observations over the NWH and a meteorological variable missing for any day is not considered for finding the analog days (Das 2005; Singh and Ganju 2005; Singh *et al.* 2010; Dimri 2006; Dimri and Niyogi 2013; Dimri *et al.* 2013). However, uncorrelated surface meteorological variables are likely to provide better analog days and less errors in various surface meteorological variables due to the removal of redundancies in the data.

All surface meteorological variables were normalised into the range 0.0–1.0 for computing Euclidean distances (Singh *et al.* 2010). Ten analog days (10 days with least Euclidean distances) corresponding to each test day were searched from the base dataset and dates and values of different surface meteorological variables in analog days were retained. Statistical characteristics of each surface meteorological variable such as the maximum value, minimum value, mean value and standard deviation in analog days were computed using the following expressions:

$$x_{k,\text{mx}} = \text{Max} \left(\bigcup_{m=1}^n x_{k,m} \right), \quad (2)$$

$$x_{k,\text{mn}} = \text{Min} \left(\bigcup_{m=1}^n x_{k,m} \right), \quad (3)$$

$$x_{k,\text{me}} = \text{Mn} \left(\bigcup_{m=1}^n x_{k,m} \right), \quad (4)$$

$$x_{k,\text{st}} = \text{St} \left(\bigcup_{m=1}^n x_{k,m} \right), \quad (5)$$

where $x_{k,m}$ is the value of the surface meteorological variable, k ($k = 1$ to p) in the m th analog day ($m = 1$ to n , $n = 10$), \bigcup is the union operator, $x_{k,\text{mx}}$, $x_{k,\text{mn}}$, $x_{k,\text{me}}$ and $x_{k,\text{st}}$ are the maximum value, minimum value, mean value and standard deviation, respectively, of a meteorological variable x_k in analog days. Max, Min, Mn and St are mathematical operators used to compute the maximum value, minimum value, mean value and standard deviation of a meteorological variable x_k in analog days, respectively.

Resemblance measures such as Pearson correlation coefficients (CCs), Mean Absolute Differences (MADs) and Root Mean Square Differences (RMSDs) between observed values of different surface meteorological variables on test days and corresponding maximum values, minimum values and mean values of them in analog days (analog means)

are computed to assess the quality of local-scale surface weather analogs at six stations over the NWH (Lorenz 1969; Radinovic 1975; Gutzler and Shukla 1984; Suranjana and Van Den Dool 1988; Van Den Dool 1989). CCs provide an insight into the joint distributions of the observed values of different surface meteorological variables on test days and their corresponding extreme values (the maximum value, the minimum value) and analog means in analog days as summary measures (Murphy and Winkler 1987; Murphy 1988; Jensenius 1990; Kalnay *et al.* 1990). However, CCs cannot identify the differences and biases between the observed values of different surface meteorological variables on test days and their corresponding extreme values and analog means in analog days (Murphy and Winkler 1987; Barnston 1992; Livezey 1994). Therefore, MADs and RMSDs are used as resemblance measures in addition to CCs to assess the quality of local-scale surface weather analogs at six stations over the NWH.

CCs, MADs and RMSDs of the observed values of different surface meteorological variables on test days with the corresponding extreme values (the maximum value, the minimum value) in analog days and their analog means for the same day ($d0$, lead time 0 hour (h)) are computed. Here, $d0$ refers to a test day for which analog days are searched from the base dataset. CCs, MADs and RMSDs of the observed values of different surface meteorological variables on $d1$ ($d0 + 1$, lead time 24 h), $d2$ ($d0 + 2$, lead time 48 h) and $d3$ ($d0 + 3$, lead time 72 h) of the test days with their analog means on $d0$ are computed to get an insight into whether surface weather analogs are capable of capturing temporal evolutions of different surface meteorological variables on test days at a local scale over the NWH. Reasonable values of resemblance measures between the observed values of different surface meteorological variables on test days (test data) with their corresponding extreme values and analog means in analog days can provide evidence that the AE system holds good promise for the real-time prediction of various surface meteorological variables at a local scale over the NWH.

MADs and RMSDs between the observed values of a surface meteorological variable, x_k , on $d0$, $d1$, $d2$ and $d3$ of test days with extreme values in analog days and its analog mean on $d0$ of test days were computed using the following expression:

$$X_{k,\text{MADs},d0,d1,d2,d3} = \frac{\sum_{i=1}^{n_{d0,d1,d2,d3}} \text{abs}(x_{k,o,i,d0,d1,d2,d3} - x_{k,a,i,d0})}{n_{d0,d1,d2,d3}}, \quad (6)$$

$$X_{k,\text{RMSDs},d0,d1,d2,d3} = \sqrt{\frac{\sum_{i=1}^{n_{d0,d1,d2,d3}} (x_{k,o,i,d0,d1,d2,d3} - x_{k,a,i,d0})^2}{n_{d0,d1,d2,d3}}} \quad (7)$$

where $X_{k,\text{MADs},d0,d1,d2,d3}$ and $X_{k,\text{RMSDs},d0,d1,d2,d3}$ are MADs and RMSDs, respectively, of the observed values of a variable, x_k , on $d0$, $d1$, $d2$ and $d3$ of test days from extreme values in analog days ($d0$ only) and its analog means on $d0$. $X_{k,o,i,d0,d1,d2,d3}$ is the observed value of a variable, x_k , on the i th test day corresponding to $d0$, $d1$, $d2$ and $d3$ and $X_{k,a,i,d0}$ is the extreme value in analog days ($d0$ only) and analog mean of a variable, x_k , on the i th test day corresponding to $d0$. $n_{d0,d1,d2,d3}$ is the total number of test days corresponding to $d0$, $d1$, $d2$ and $d3$.

CCs of the climatological means of different surface meteorological variables on $d0$ with their observed values on $d0$, $d1$, $d2$ and $d3$ of test days and MADs (RMSDs) of climatological means of different surface meteorological variable on $d0$ with their observed values on $d0$, $d1$, $d2$ and $d3$ of test days were computed using expressions (6 and 7). This was done to gain an insight into the resemblance measures of the climatological means of different surface meteorological variables on $d0$ with their corresponding observed values on $d0$, $d1$, $d2$ and $d3$ of test days. Resemblance measures (CCs, MADs, RMSDs) of different surface meteorological variables in analog days with their corresponding observed values on test days and resemblance measures of the climatological means of different surface meteorological variables with their corresponding observed values on test days are compared to know whether the surface weather analogs exhibit some skill for the prediction of different meteorological variables relative to their climatological means.

4. Results and discussion

The developed AE system was run for independent test days falling in the range of 1476–1722 days for the past 10 winters (winter 2006–2015) at six stations over the NWH (table 1). The maximum value, minimum value, analog mean and standard deviation of each surface meteorological variable in analog days corresponding to $d0$ of test data were computed. CCs of the observed values of different surface meteorological variables on $d0$ of test days with extreme values (the maximum value, minimum value) in analog days and their analog means

on $d0$ were computed. CCs between the observed values of different surface meteorological variables on $d1$, $d2$ and $d3$ of test days and their analog means on $d0$ were computed. Statistically significant positive CCs between the observed values of different surface meteorological variables on test days and their values in analog days will indicate a good resemblance between the observed surface meteorological variables on test days and the corresponding surface meteorological variables in analog days. This will indicate that the AE system is capable of finding reasonably good analogs and it can capture the observed dynamics (tendency) of different surface meteorological variables for $d0$ through $d3$ at a local scale over the NWH.

Statistically significant high positive CCs (>0.85) between the observed values of different surface meteorological variables on $d0$ with the corresponding extreme values and their analog means on $d0$ show that surface meteorological variables in analog days resemble well with their corresponding observed values on $d0$ of test days at a local scale over the NWH (table 2). However, lower values of CCs between the observed surface atmospheric pressure on $d0$ of test days with the surface atmospheric pressure in analog days at a few stations compared to the other stations over the NWH can be attributed to the large variability in observations on surface atmospheric pressure at a few stations over the NWH (table 2). The temporal resolution of observations on variables AM, AVG and P was 9 hour (9 h) and the temporal resolution of observations on variables MX, MN, DAVG, PT and SSH was 24 h (appendix, table A1). Comparable values of CCs between the observed values of MX, MN, AM, AVG, P , DAVG, PT and SSH on $d0$ of test days with their corresponding values in analog days show that the AE system is capable of handling surface meteorological variables of different temporal resolutions simultaneously at a local scale over the NWH (table 2).

CCs between the observed values of different surface meteorological variables on $d0$ of test days with their climatological means were found to be significantly lesser than CCs between the observed values of different surface meteorological variables on $d0$ of test data with the corresponding extreme values in analog days and analog means of them at six stations over the NWH (table 2). This suggests that different surface meteorological variables in analog days resemble their observed values on $d0$ of test days better than their climatological means

Table 2. Pearson correlation coefficients (CCs) between the observed values of different surface meteorological variables on d_0 (same day, lead time 0 hour (h)) of independent test days and corresponding maximum values, minimum values and analog mean of them in 10 analog days (analog days).

Variables	A			B			C		
	MX	MN	M(CM)	MX	MN	M(CM)	MX	MN	M(CM)
MX(O)	0.97*	0.97*	0.98* (0.64*)	0.97*	0.97*	0.98* (0.55*)	0.98*	0.97*	0.98* (0.66*)
MN(O)	0.97*	0.97*	0.98* (0.74*)	0.95*	0.95*	0.97* (0.65*)	0.96*	0.97*	0.97* (0.65*)
AM(O)	0.97*	0.97*	0.98* (0.70*)	0.97*	0.97*	0.98* (0.54*)	0.97*	0.97*	0.98* (0.71*)
AVG(O)	0.99*	0.99*	0.99* (0.13*)	0.98*	0.98*	0.99* (0.24*)	0.99*	0.99*	0.99* (0.12*)
DAVG(O)	0.94*	0.94*	0.96* (0.17*)	0.89*	0.87*	0.91* (0.20*)	0.86*	0.85*	0.88* (0.18*)
P(O)	0.32*	0.37*	0.42* (0.24*)	0.96*	0.96*	0.98* (0.23*)	0.97*	0.96*	0.98* (0.30*)
PT(O)	0.96*	0.94*	0.98* (0.07*)	0.93*	0.86*	0.94* (0.13*)	0.94*	0.92*	0.96* (0.18*)
SSH(O)	0.99*	0.99*	0.99* (0.12*)	0.98*	0.98*	0.99* (0.17*)	0.99*	0.99*	0.99* (0.13*)

Variables	D			E			F		
	MX	MN	M(CM)	MX	MN	M(CM)	MX	MN	M(CM)
MX(O)	0.97*	0.96*	0.97* (0.63*)	0.97*	0.97*	0.98* (0.75*)	0.97*	0.97*	0.98* (0.79*)
MN(O)	0.97*	0.97*	0.98* (0.72*)	0.97*	0.97*	0.98* (0.69*)	0.97*	0.97*	0.98* (0.70*)
AM(O)	0.96*	0.96*	0.97* (0.75*)	0.97*	0.97*	0.98 (0.77*)	0.96*	0.94*	0.97* (0.79*)
AVG(O)	0.98*	0.98*	0.99* (0.08*)	0.97*	0.98*	0.99* (0.10*)	0.99*	0.98*	0.99* (0.20*)
DAVG(O)	0.93*	0.95*	0.95* (0.08*)	0.92*	0.93*	0.95* (0.03)	0.95*	0.93*	0.96* (0.17*)
P(O)	-0.09*	0.19*	0.07* (0.09*)	0.95*	0.95*	0.97* (0.40*)	-0.05	0.27*	0.18* (0.03)
PT(O)	0.94*	0.87*	0.96* (0.02)	0.89*	0.88*	0.94* (0.04)	0.92*	0.86*	0.95* (0.11*)
SSH(O)	0.99*	0.98*	0.99* (0.11*)	0.99*	0.99*	0.99* (0.09*)	0.99*	0.99*	0.99* (0.09*)

CCs between the observed values of different surface meteorological variables on d_0 of test data and climatological means of them on d_0 are given within brackets (MX(O) = observed maximum air temperature, MX = maximum value of maximum air temperature in 10 analog days; MN = minimum value of maximum air temperature in 10 analog days; M = mean value of maximum air temperature in 10 analog days; CM = climatological mean, A, B to F = station; * = CCs significant at 1% significance level; see text for details).

(table 2). The analog means and climatological means of different surface meteorological variables on d_0 show statistically significant high positive CCs with their observed values on d_1 , d_2 and d_3 of test days at six stations over the NWH (tables 3 and 4). However, CCs of the climatological means of different surface meteorological variables on d_0 of test days with their corresponding observed values on d_1 , d_2 and d_3 are found to be lesser than the CCs of analog means of different surface meteorological variables on d_0 of test days with their corresponding observed values on d_1 , d_2 and d_3 (tables 3 and 4).

CCs between analog means of air temperatures (the maximum air temperature, the minimum air temperature and the ambient air temperature) on d_0 with the observed values of air temperatures on d_1 , d_2 and d_3 were found comparable to CCs between the observed air temperature and predicted air temperature over the NWH and/or other places (Maini *et al.* 2003; Kumar and Krishnamurti 2006; Eccel *et al.* 2007; Shank *et al.* 2008; Joshi and Ganju 2012; Tiwari *et al.* 2014). CCs of analog

mean and climatological mean of each surface meteorological variable on d_0 with its observed values on d_0 , d_1 , d_2 and d_3 decrease with increasing lead time, i.e., d_0 through d_3 (tables 2–4). This shows that the resemblance between surface meteorological variables in analog days and their observed values on d_1 , d_2 and d_3 decreases with increasing lead time from d_0 through d_3 (tables 3 and 4). Similarly, the resemblance between the climatological means of different surface meteorological variables on d_0 with their observed values on d_1 , d_2 and d_3 decreases with increasing lead time, i.e., d_0 through d_3 (tables 3 and 4). The analog mean and the climatological mean of surface atmospheric pressure on d_0 with its observed values on d_1 , d_2 and d_3 at stations D and F show lower values of CCs than the CCs between analog mean and climatological mean of surface atmospheric pressure on d_0 with its observed values on d_1 , d_2 and d_3 at other stations over the NWH (tables 2–4). This may be due to large variations in the measurements of surface atmospheric pressure at stations D and F. The analog mean of the

Table 3. *CCs between the mean values of different meteorological variables in analog days on d0 of the test data and their corresponding observed values on d1 (d1 = d0 + 1, lead time 24 hours (h)) and d2 (d2 = d1 + 1, lead time 48 h) of test days.*

Station	MX (CM)	MN (CM)	AM (CM)	AVG (CM)	DAVG (CM)	P (CM)	PT (CM)	SSH (CM)
<i>Surface meteorological variables and CCs for D1</i>								
A	0.83* (0.62*)	0.92* (0.73*)	0.86* (0.69*)	0.63* (0.12*)	0.65* (0.15*)	0.29* (0.23*)	0.38* (0.08*)	0.51* (0.14*)
B	0.75* (0.53*)	0.82* (0.61*)	0.73* (0.53*)	0.35* (0.20*)	0.45* (0.19*)	0.67* (0.23*)	0.33* (0.14*)	0.46* (0.18*)
C	0.81* (0.65*)	0.87* (0.65*)	0.84* (0.71*)	0.40* (0.09*)	0.48* (0.18*)	0.77* (0.30*)	0.39* (0.14*)	0.41* (0.13*)
D	0.86* (0.62*)	0.88* (0.70*)	0.86* (0.74*)	0.53* (0.08*)	0.56* (0.09*)	0.03 (0.06*)	0.35* (0.01)	0.55* (0.11*)
E	0.87* (0.75*)	0.80* (0.68*)	0.87* (0.76*)	0.58* (0.08*)	0.61* (0.0)	0.81* (0.38*)	0.24* (0.07*)	0.34* (0.04)
F	0.86* (0.77*)	0.86* (0.68*)	0.88* (0.77*)	0.64* (0.15*)	0.67* (0.13*)	0.17* (0.03)	0.35* (0.09*)	0.41* (0.05*)
<i>Surface meteorological variables and CCs for D2</i>								
A	0.72* (0.61*)	0.84* (0.73*)	0.79* (0.68*)	0.57* (0.11*)	0.61* (0.15*)	0.22* (0.23*)	0.10* (0.08*)	0.26* (0.14*)
B	0.65* (0.64*)	0.70* (0.64*)	0.69* (0.70*)	0.24* (0.09*)	0.28* (0.17*)	0.48* (0.28*)	0.02 (0.10*)	0.13* (0.12*)
C	0.82* (0.77*)	0.71* (0.66*)	0.82* (0.78*)	0.57* (0.15*)	0.57* (0.10*)	0.09* (0.05)	0.10* (0.06*)	0.17* (0.07*)
D	0.77* (0.61*)	0.77* (0.70*)	0.80* (0.73*)	0.43* (0.08*)	0.48* (0.11*)	0.03 (0.08*)	0.09* (0.02)	0.33* (0.08*)
E	0.77* (0.61*)	0.77* (0.70*)	0.80* (0.73*)	0.43* (0.08*)	0.48* (0.11*)	0.03 (0.08*)	0.09* (0.02)	0.33* (0.07*)
F	0.82* (0.77*)	0.71* (0.66*)	0.82* (0.78*)	0.57* (0.15*)	0.57* (0.10*)	0.09* (0.05)	0.10* (0.06*)	0.17* (0.08*)

CCs between the climatological means of different surface meteorological variables on d0 and their corresponding observed values on d1 and d2 of test days are given within brackets (MX = CCs between the observed values of maximum air temperature on d1/d2 and the mean values of maximum air temperature in analog days, on d0, CM = CCs between the observed values of the maximum air temperature on d1/d2 and the climatological means of the maximum temperature on d0, A, B, to F = station; * = CCs significant at 1% significance level).

Table 4. *Same as table 3 but for d3 (d3 = d2+1, lead time 72 h).*

Surface meteorological variables and CCs for D3								
Station	MX (CM)	MN (CM)	AM (CM)	AVG (CM)	DAVG (CM)	P (CM)	PT (CM)	SSH (CM)
A	0.65* (0.60*)	0.77* (0.72*)	0.73* (0.67*)	0.57* (0.11*)	0.59* (0.16*)	0.15* (0.22*)	0.02 (0.08*)	0.15* (0.13*)
B	0.53* (0.52*)	0.64* (0.59*)	0.54* (0.51*)	0.32* (0.20*)	0.34* (0.18*)	0.41* (0.22*)	0.01 (0.11*)	0.18* (0.17*)
C	0.65* (0.64*)	0.70* (0.64*)	0.69* (0.70*)	0.24* (0.09*)	0.28* (0.17*)	0.48* (0.28*)	0.02 (0.10*)	0.13* (0.12*)
D	0.70* (0.60*)	0.69* (0.70*)	0.75* (0.72*)	0.44* (0.06*)	0.46* (0.08*)	0.08 (0.08*)	0.04 (0.05)	0.21* (0.12*)
E	0.77* (0.75*)	0.60* (0.67*)	0.77* (0.75*)	0.48* (0.09*)	0.48* (0.01)	0.57* (0.35*)	0.01 (0.04)	0.09* (0.11*)
F	0.80* (0.76*)	0.61* (0.66*)	0.78* (0.78*)	0.52* (0.15*)	0.51* (0.11*)	0.03 (0.06**)	-0.04 (0.07**)	0.08* (0.10*)

** = CCs significant at 5% significance level.

precipitation amount on d0 do not show statistically significant CCs with its observed values on d3 of test data (table 4). This can be attributed to the very less number of precipitation days and a large temporal variability in the precipitation amount over the NWH.

MADs and RMSDs of extreme values (the maximum value, the minimum value) in analog days, analog means and climatological means of different surface meteorological variables on d0 from their observed values on d0, d1, d2 and d3 were computed. This was done to gain an insight into the differences (errors) associated with local-scale surface weather analogs over the NWH. MADs (RMSDs) of extreme values in analog days and analog means of different surface meteorological variables on d0 from their observed values on d0, d1, d2 and d3 were compared with MADs

(RMSDs) of the climatological means of different surface meteorological variables on d0 from their observed values on d0, d1, d2 and d3.

The mean and standard deviation of extreme values in analog days, analog mean and climatological mean of different surface meteorological variables on d0 were computed to gain an insight into whether surface weather analogs are capable of reproducing (represent) the observed mean and variability (standard deviation) of different meteorological variables at a local scale over the NWH. The mean and standard deviation of the observed maximum air temperature, extreme values of maximum air temperature in analog days, the analog mean of maximum air temperature and the climatological mean of maximum air temperature on d0 of test days are given in table 5 at six stations over the NWH.

Table 5. *Statistics (mean, standard deviation) of the observed maximum air temperature (M(O)), the maximum value of the maximum air temperature (MMX(A)), the minimum value of the maximum air temperature (MMN(A)) and mean value of the maximum air temperature (M(A)) in analog days and the climatological mean of the maximum air temperature (M(C)) on d0 of test data at six stations over the NWH (M = mean of six stations over the NWH).*

Station	Mean					Standard deviation				
	MX(O)	MMX(A)	MMN(A)	M(A)	M(C)	MX(O)	MMX(A)	MMN(A)	M(A)	M(C)
A	6.2	7.7	4.1	5.8	6.2	4.7	4.5	4.4	4.4	3.4
B	15.8	17.6	13.5	15.5	14.6	4.9	4.4	4.7	4.5	2.7
C	8.6	10.3	6.5	8.4	7.8	5.2	4.8	4.7	4.7	3.1
D	5.5	7.1	3.2	5.1	4.7	4.6	4.3	4.0	4.1	3.2
E	2.3	4.7	0.4	2.6	1.9	4.9	4.4	4.8	4.5	3.5
F	2.3	4.9	-0.1	2.4	1.9	5.7	5.4	5.2	5.2	4.7
M	6.8	8.7	4.6	6.6	6.2	5.0	4.6	4.6	4.5	3.4

The analog mean of maximum air temperature on *d0* was found to represent the observed mean of maximum air temperature on *d0* of test days better than the extreme values of maximum air temperature in analog days and the climatological mean of maximum air temperature at each station over the NWH (table 5). This indicates that the AE system (surface weather analogs) is capable of capturing local-scale features of maximum air temperature at six stations over the NWH and it reproduces the observed mean of maximum air temperature better than the climatological mean (table 5). The extreme values in analog days and the analog mean of the maximum air temperature are found to represent the observed variability of maximum air temperature better than the climatological mean of the maximum air temperature on *d0* of test days at six stations over the NWH (table 5). This suggests that the surface weather analogs get tuned to the local climatic conditions of each station over the NWH (table 5).

The mean value of the observed means, extreme values in analog days, analog means and the climatological means of each surface meteorological variable at six stations over the NWH were computed. The analog means of different surface meteorological variables were found to represent their observed means better than the extreme values in analog days and the climatological means of different meteorological variables on *d0* of test days (figure 2). Extreme values of different surface meteorological variables in analog days were found to differ largely from their observed values on *d0* of test days (figure 2). The extreme values in analog days and the analog means of different surface meteorological variables were found to represent their observed variability better than their climatological means on *d0* of test days (figure 3). The

precipitation amount shows a large variability (standard deviation) as compared to other surface meteorological variables over the NWH (figure 3). The observed variability of the precipitation amount on *d0* was poorly reproduced by extreme values in analog days and its climatological mean as compared to the analog mean of the precipitation amount over the NWH (figure 3). This suggests that the analog mean is capable of reproducing (representing) a large variability in any surface meteorological variable better than the extreme values in analog days and its climatological mean (figure 3).

Statistics (mean, standard deviation) of different surface meteorological variables does not provide any insight into the differences (errors) between the observed values of different surface meteorological variables on test days and their climatological means and extreme values and analog means. Hence MADs (RMSDs) of extreme values in analog days, analog mean and climatological mean of each surface meteorological variable on *d0* from its observed values on *d0*, *d1*, *d2* and *d3* of test days were computed. MADs (RMSDs) of extreme values of minimum air temperature in analog days, analog mean of the minimum air temperature and the climatological mean of minimum air temperature on *d0* from its observed values on *d0*, *d1*, *d2* and *d3* of test data at six stations over the NWH are given in table 6. MADs (RMSDs) of the analog mean of minimum air temperature on *d0* from the observed values of minimum air temperature on *d0*, *d1*, *d2* and *d3* were found to be lesser than the MADs (RMSDs) of extreme values of minimum air temperature in analog days and the climatological mean of minimum air temperature on *d0* from the observed values of minimum air temperature on *d0*, *d1*, *d2*

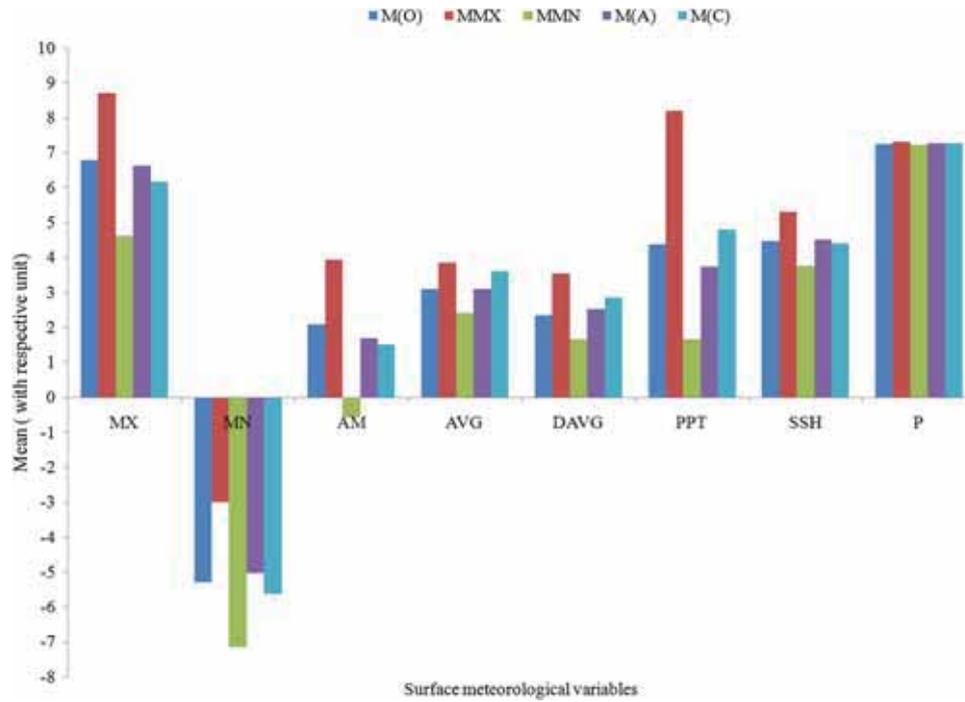


Figure 2. Observed means ($M(O)$), climatological means ($C(M)$) and mean of maximum values (MMX) in analog days, mean of minimum values (MMN) in analog days and mean of analog means ($M(A)$) of different surface meteorological variables on $d0$ of test data at six stations over the Northwest Himalaya (NWH) (mean value of each surface meteorological variable is computed based on values of it at six stations over the NWH).

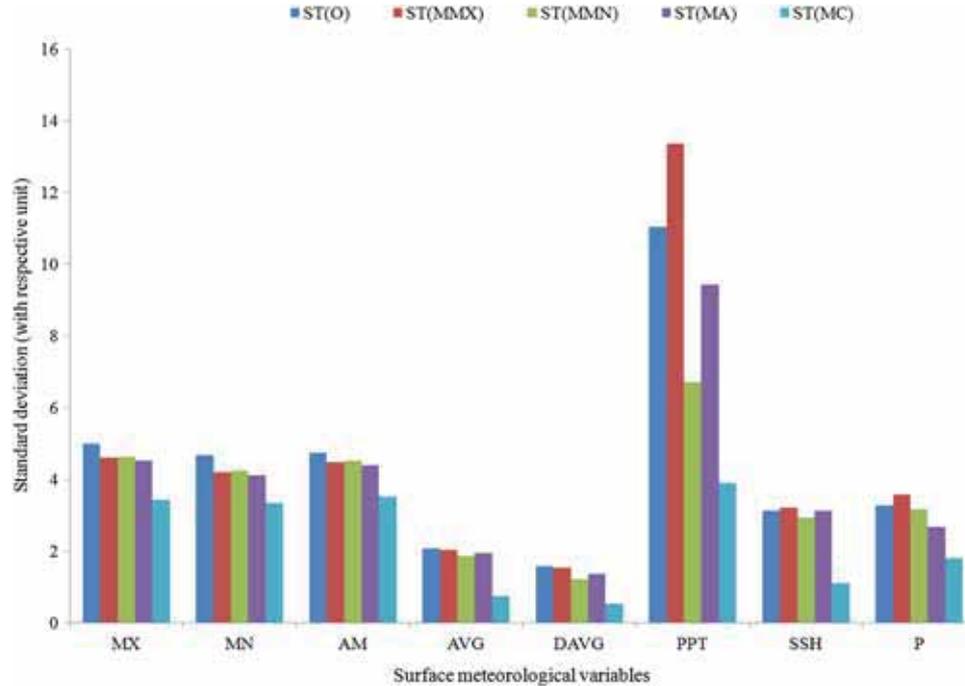


Figure 3. Observed standard deviations ($ST(O)$), climatological standard deviations ($ST(MC)$), standard deviations of maximum values ($ST(MMX)$), standard deviations of minimum values ($ST(MMN)$) and standard deviations of mean values ($ST(MA)$) in analog days of different surface meteorological variables on $d0$ of test data at six stations over the NWH (standard deviation of each surface meteorological variable represents mean value of it at six stations over the NWH).

Table 6. Mean Absolute Differences (MADs) and Root Mean Square Differences (RMSDs) of the maximum value (MX(A)), the minimum value (MN(A)) and the mean value (M(A)) of the minimum air temperature in analog days and the climatological mean of the minimum air temperature (M(C)) on d0 from its observed values on d0, d1, d2 and d3 of test days at six stations over the NWH (RMSDs (°C) are given within brackets; M = mean of six stations over the NWH).

Station	MAD (RMSD) (°C)									
	d0 (same day, lead time 0 hour (h))				d1 (d1 = d0 + 1, lead time 24 h)		d2 (d2 = d1 + 1, lead time 48 h)		d3 (d3 = d2 + 1, lead time 72 h)	
	MX(A)	MN(A)	M(A)	M(C)	M(A)	M(C)	M(A)	M(C)	M(A)	M(C)
A	2.2 (2.5)	1.4 (1.7)	0.9 (1.2)	2.4 (3.1)	1.4 (1.8)	2.4 (3.0)	1.8 (2.4)	2.4 (3.0)	2.1 (2.7)	2.4 (3.1)
B	1.9 (2.2)	1.3 (1.6)	0.7 (0.9)	1.9 (2.4)	1.2 (1.7)	1.8 (2.3)	3.6 (4.8)	4.4 (5.6)	1.8 (2.4)	1.9 (2.4)
C	1.9 (2.2)	1.6 (1.8)	0.7 (1.0)	2.4 (3.0)	1.4 (1.9)	2.3 (3.0)	1.9 (2.5)	2.3 (3.0)	2.2 (2.9)	2.4 (3.0)
D	2.3 (2.6)	1.8 (2.1)	0.8 (1.1)	2.5 (3.2)	1.6 (2.2)	2.5 (3.1)	2.1 (2.9)	2.5 (3.1)	2.5 (3.3)	2.5 (3.1)
E	2.4 (2.7)	2.2 (2.5)	0.8 (1.1)	3.2 (4.0)	2.4 (3.3)	3.2 (4.0)	3.2 (4.2)	3.2 (4.0)	3.5 (4.6)	3.2 (4.0)
F	3.2 (3.6)	3.2 (3.6)	1.2 (1.5)	4.3 (5.4)	2.5 (3.4)	4.3 (5.5)	3.6 (4.8)	4.4 (5.6)	4.2 (5.5)	4.4 (5.6)
Mean (M)	2.3 (2.6)	1.9 (2.2)	0.9 (1.1)	2.8 (3.5)	1.7 (2.4)	2.7 (3.5)	2.7 (3.6)	3.2 (4.0)	2.7 (3.6)	2.8 (3.5)

and d3 at most of the stations over the NWH (table 6).

Mean values of MADs (RMSDs) of extreme values in analog days, analog means and climatological means of different surface meteorological variables (MX, MN, AM, AVG, DAVG, P, PT and SSH) at six stations over the NWH were computed. Mean MADs (RMSDs) of extreme values of different surface meteorological variables in analog days and mean MADs (RMSDs) of climatological means of different surface meteorological variables on d0 from their observed values on d0 of test days were found to be higher than the mean MADs (RMSDs) of the analog means of different surface meteorological variables on d0 from their observed values on d0 (figures 4 and 5). Mean MADs (RMSDs) of the climatological means of most of the surface meteorological variables on d0 from their observed values on d0 were found to be higher than the mean MADs (RMSDs) of extreme values of most of the surface meteorological variables in analog days on d0 from their observed values on d0 (figures 4 and 5). This implies that extreme values in analog days and climatological means of different surface meteorological variables are less accurate to represent their observed values on d0 of test days than the analog means of different surface meteorological variables over the NWH (figures 4 and 5).

Mean MADs (RMSDs) of the climatological means of different surface meteorological variables on d0 from their observed values on d1 and d2 were found to be higher than the mean MADs (RMSDs) of the analog means of different surface meteorological variables on d0 from their observed values on d1 and d2 (figures 6 and 7). However, the mean MADs (RMSDs) of the climatological means of different

surface meteorological variables on d0 from their observed values on d3 were found to be marginally higher and/or comparable to the mean MADs (RMSDs) of the analog means of different surface meteorological variables on d0 from their observed values on d3 (figures 6 and 7). The mean RMSDs of the climatological means of the precipitation amount and sunshine duration on d0 from their observed values on d1, d2 and d3 were found to be lesser than the mean RMSDs of the analog means of the precipitation amount and sunshine duration on d0 from its observed values on d1, d2 and d3 (figure 7). This suggests that the analog means of the precipitation amount and sunshine duration on d0 differ largely from their observed values on d1, d2 and d3 for a few test days as compared to the climatological means of the precipitation amount and sunshine duration on d0 from their observed values on d1, d2 and d3 (figure 7).

The mean MADs (RMSDs) of extreme values in analog days, analog means and climatological means of different surface meteorological variables will represent the Mean Absolute Errors (MAEs) and Root Mean Square Errors (RMSEs) for their prediction if the developed AE system (surface weather analogs) and the climatological means are utilised to predict them at a local scale over the NWH. However, the AE system for the prediction of different surface meteorological variables may have different configurations than the AE system developed to conduct this study (Woodcock 1980; Bannayan and Hoogenboom 2008; Singh et al. 2008, 2010). Large values of MADs (RMSDs) of climatological means of different surface meteorological variables on d0 from their observed values on d0, d1, d2 and d3 as compared to the MADs

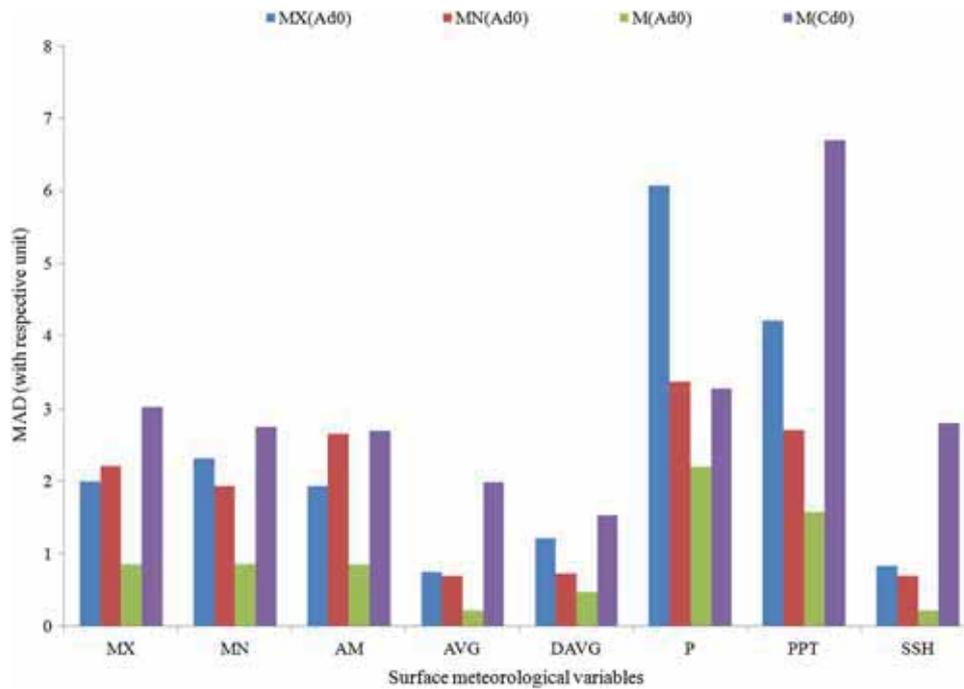


Figure 4. Mean Absolute Differences (MADs) of maximum values (MX(*Ad0*)), minimum values (MN(*Ad0*)) and mean values (M(*Ad0*)) in 10 analog days and climatological means (M(*Cd0*)) of different surface meteorological variables from observed values of them on *d0* (same day, lead time 0 h) of test data at six stations over the NWH (MAD for each surface meteorological variable represents mean MAD of six stations over the NWH).

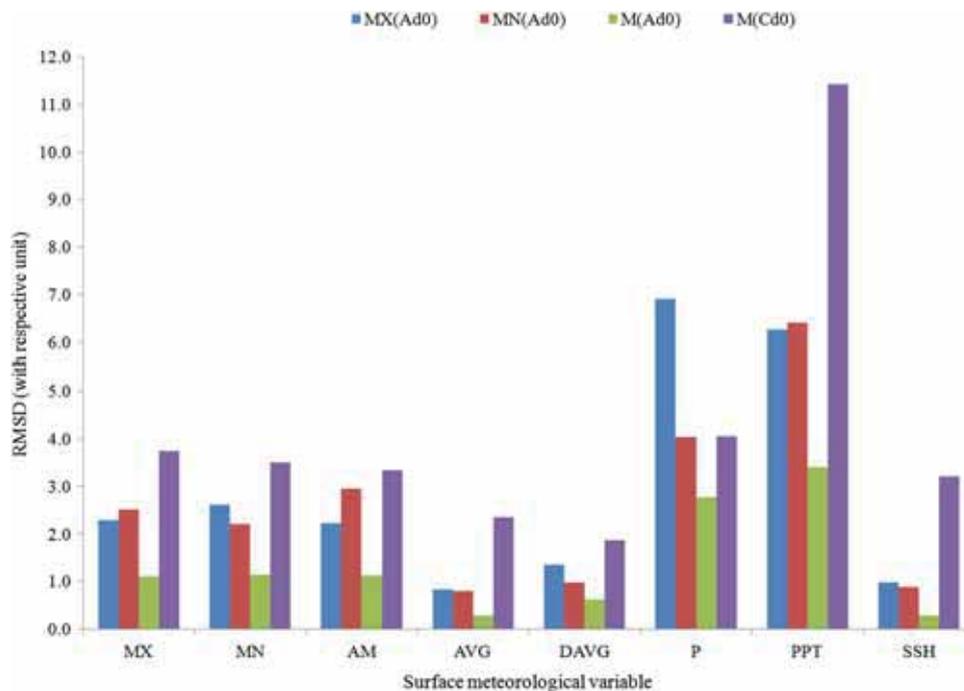


Figure 5. Root Mean Square Differences (RMSDs) of maximum values (MX(*Ad0*)), minimum values (MN(*Ad0*)) and mean values (M(*Ad0*)) in 10 analog days and climatological means (M(*Cd0*)) of different surface meteorological variables from observed values of them on *d0* (same day, lead time 0 h) of test data at six stations over the NWH (RMSD for each variable represents mean RMSD of all six stations).

(RMSDs) of extreme values (the maximum value, the minimum value) in analog days and the analog means of different surface meteorological variables

on *d0* from their observed values on *d0*, *d1*, *d2* and *d3* show that the surface weather analogs exhibit lesser errors (MAEs, RMSEs) than the

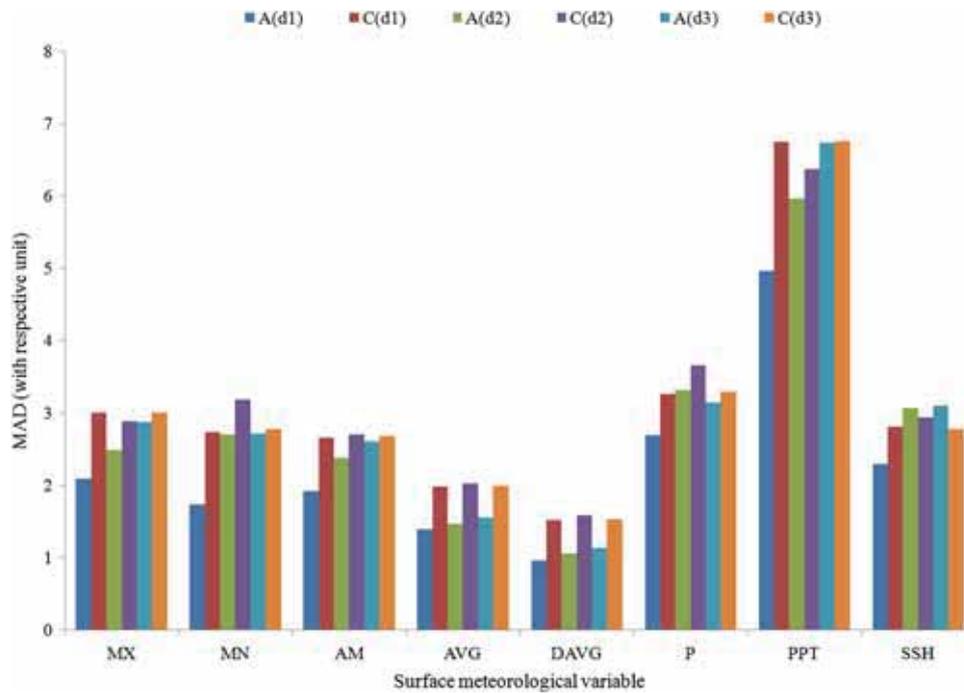


Figure 6. Mean Absolute Differences (MADs) of climatological means and analog means of different surface meteorological variables on d_0 from observed values of them on d_1 ($d_0 + 1$, lead time 24 h), d_2 ($d_0 + 2$, lead time 48 h) and d_3 ($d_0 + 3$, lead time 72 h) at six stations over the NWH (MAD for each variable represents mean MAD of six stations; $A(d_1, d_2, d_3)$ = MAD of analog means for d_1, d_2 and d_3 ; $C(d_1, d_2, d_3)$ = MADs of climatological means for d_1, d_2 and d_3).

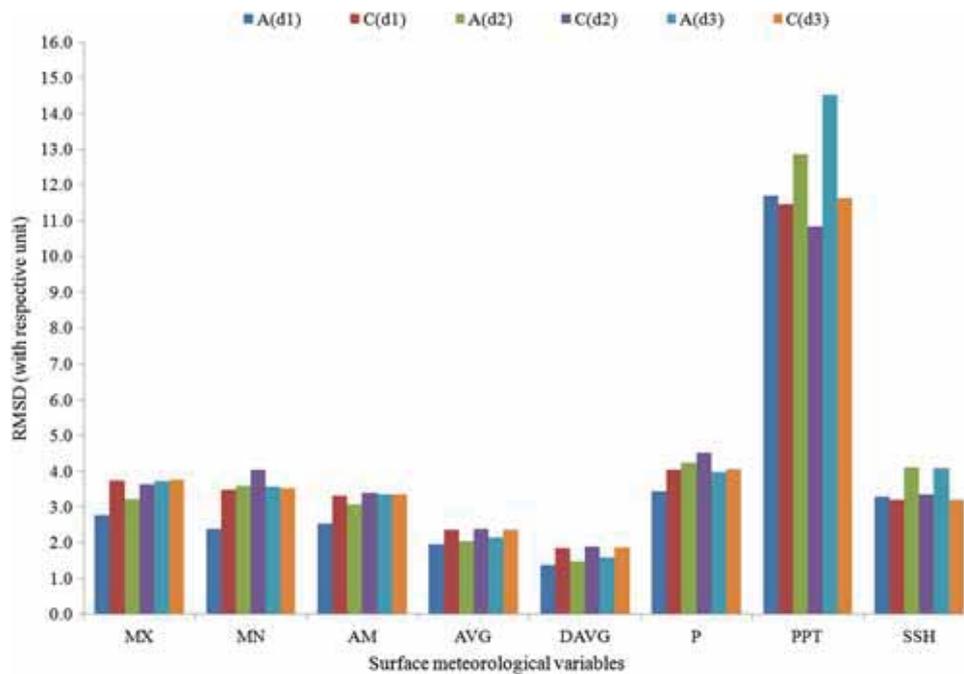


Figure 7. Root Mean Square Differences (RMSDs) of climatological means and analog means of different surface meteorological variables on d_0 from observed values of them on d_1 ($d_0 + 1$, lead time 24 h), d_2 ($d_0 + 2$, lead time 48 h) and d_3 ($d_0 + 3$, lead time 72 h) at six stations over the NWH (RMSD for each variable represents mean RMSD of six stations; $A(d_1, d_2, d_3)$ = RMSD of analog means for d_1, d_2 and d_3 ; $C(d_1, d_2, d_3)$ = RMSDs of climatological means for d_1, d_2 and d_3).

climatological means for d_0, d_1, d_2 and d_3 for a local-scale prediction of different surface meteorological variables over the NWH (figures 4–7).

Site-specific (local scale) prediction of maximum air temperature, minimum air temperature and ambient air temperature over the NWH with the

help of empirical modelling utilising surface meteorological observations and/or statistical downscaling of MM5 model outputs have been carried out in the past (Srinivasan *et al.* 2010; Joshi and Ganju 2012; Joshi *et al.* 2017). Mean RMSDs of air temperatures (the maximum air temperature, the minimum air temperature, the ambient air temperature) in surface weather analogs for $d1$ through $d3$ were found to be comparable to RMSEs of the prediction of temperatures at different stations over the NWH (Srinivasan *et al.* 2010; Joshi and Ganju 2012; Joshi *et al.* 2017). However, surface weather analogs can be utilised simultaneously to predict multiple surface meteorological variables at a specific site (local scale) compared to the prediction of the maximum air temperature and the minimum air temperature at a specific site over the NWH (Srinivasan *et al.* 2010; Joshi and Ganju 2012; Joshi *et al.* 2017). Mean MADs and mean RMSDs of air temperatures (the maximum air temperature, the minimum air temperature, the ambient air temperature) in surface weather analogs over the NWH were found to be comparable to the MAEs and RMSEs of temperature prediction (Maini *et al.* 2003; Eccel *et al.* 2007; Shank *et al.* 2008; Novak *et al.* 2014). Mean RMSD of precipitation amount in surface weather analogs over the NWH was found to be comparable to the RMSE of the raw MM5 model output and statistically downscaled precipitation amount for $d1$ through $d3$ at different stations over the NWH (Shekhar *et al.* 2014). However, the mean RMSD of the precipitation amount in surface weather analogs for $d1$ were found to be marginally higher than the downscaled precipitation amount using surface meteorological variables and MM5 precipitation amount employing artificial neural network (Joshi and Ganju 2013). Mean MADs (RMSDs) of the precipitation amount in the local-scale surface weather analogs over the NWH for a short time interval were found to be comparable to MAEs (RMSEs) of the quantitative precipitation forecast (Kumar and Krishnamurti 2006; Roy Bhowmik and Durai 2008). Mean MADs (RMSDs) of the surface atmospheric pressure and the average wind speed in surface weather analogs were found to be comparable to MAEs (RMSEs) of the prediction of surface atmospheric pressure and the prediction of average wind speed (Das *et al.* 2002, 2008; Kumar and Krishnamurti 2006; Dimri and Chevuturi 2014).

MADs (RMSDs) of different surface meteorological variables in surface weather analogs

comparable to MAEs (RMSEs) of NWP models, and empirical modelling using surface meteorological observations and statistical downscaling of the NWP model outputs (MM5 model) show that the surface weather analogs hold good promise for real-time local-scale simultaneous prediction of surface meteorological variables in remote areas of the NWH with least requirement of resources (human and computational).

5. Conclusions and limitations

Weather forecasts play an extremely important role in planning various activities and mitigation of weather-generated natural hazards due to heavy snowfall during winter over the NWH. Large spatio-temporal variability in various surface meteorological variables over the NWH demands their real-time local-scale prediction. Surface meteorological observations can be of great help in generating real-time local-scale forecasts of various surface meteorological variables over the NWH. A detailed study to assess the quality of surface weather analogs at a local scale over the NWH was conducted. The results of this study show that the surface weather analogs are capable of realising different surface meteorological variables and their temporal evolutions in analog days and the AE system is adaptable to the local climatic conditions of each station over the NWH. Surface meteorological variables in analog days better resemble the corresponding observed surface meteorological variables of independent test days than their climatological values. MADs and RMSDs of extreme values in analog days and analog means of different surface meteorological variables from their corresponding observed values on independent test days comparable to the MAEs (RMSEs) for the prediction of various surface meteorological variables by different weather forecast models show that the surface weather analogs hold good promise for local-scale prediction of surface meteorological variables over the NWH. However, the configuration of the AE system for the prediction of various surface meteorological variables may be different from the AE system employed to conduct this study.

This study was conducted using short length and partially incomplete surface meteorological observations of the winter period over the NWH. The quality of surface weather analogs was assessed for a set of available surface meteorological variables

over the NWH. However, a similar study on long homogeneous surface meteorological observations for a complete year over the NWH (or elsewhere) may provide better insight into the quality of local-scale surface weather analogs.

Acknowledgements

The authors are thankful to the director, SASE, for his constant encouragement during this study. The invaluable suggestions from the anonymous reviewers on improving the quality of the manuscript is gratefully acknowledged. Efforts made by the Scientists and Technical Staff of SASE towards the collection of surface meteorological observations from remote areas of the NWH under harsh climatic conditions are acknowledged.

Appendix

Table A1. List of surface meteorological variables utilised to assess the quality of local-scale surface weather analogs over the NWH.

Sl. no.	Variable (abbreviation)	Unit	Temporal resolution of observation (h)
1	Maximum air temperature (MX)	°C	24:00
2	Minimum air temperature (MN)	°C	24:00
3	Ambient air temperature (AM)	°C	9:00
4	Average wind speed (AVG)	km/h	9:00
5	Average wind speed in a day (DAVG)	km/h	24:00
6	Surface atmospheric pressure (P)	hPa	9:00
7	Precipitation amount (PPT)	mm	24:00
8	Sunshine hours (SSH)	h	24:00

References

Bannayan M and Hoogenboom G 2008 Weather analogue: A tool for real time prediction of daily weather data realization based on a modified k-nearest neighbor approach; *Environ. Model Softw.* **23** 703–713.

Barnston A G 1992 Correspondence among the correlation, RMSE and Heidke forecast verification measures: Refinement of the Heidke score; *Wea. Forecasting* **7** 699–709.

Barnett T P and Preisendorfer R W 1978 Multifield analog prediction of short-term climate fluctuations using a climate state vector; *J. Atmos. Sci.* **35** 1771–1787.

Buizza R, Hollingsworth A, Lalaurette F and Ghelli A 1999 Probabilistic predictions of precipitation using the ECMWF ensemble prediction system; *Wea. Forecasting* **14** 168–189.

Charles A, Timbal B, Fernandez E and Hendon H 2013 Analog downscaling of seasonal rainfall forecasts in the Murray Darling Basin; *Mon. Weather Rev.* **141** 1099–1117.

Daoud A B, Sauquet E, Lang M, Bontron G and Oblot C 2011 Precipitation forecasting through an analog sorting technique: A comparative study; *Adv. Geosci.* **29** 103–107.

Das S 2005 Mountain weather forecasting using MM5 modelling system; *Curr. Sci.* **88** 899–905.

Das S, Mitra A K, Iyengar G R and Singh J 2002 Skill of medium-range forecasts over the Indian monsoon region using different parameterizations of deep convection; *Wea. Forecasting* **17** 1194–1210.

Das S, Singh S V, Rajagopal E N and Gall R 2003 Mesoscale modeling for mountain weather forecasting over the Himalayas; *Bull. Am. Meteor. Soc.* **84**(9), <https://doi.org/10.1175/bams-84-9-1237>.

Das S, Ashrit R, Iyengar G R, Mohandas S, Gupta M D, George J P, Rajagopal E N and Dutta S K 2008 Skills of different meso scale models over Indian region during monsoon season: Forecast errors; *J. Earth Syst. Sci.* **117** 603–620.

Delle Monache L, Nipen T, Liu Y, Roux G and Stull R 2011 Kalman filter and analog schemes to post process numerical weather predictions; *Mon. Weather Rev.* **139** 3554–3570.

Delle Monache L, Eckel F A, Rife D L, Nagarajan B and Searight K 2013 Probabilistic weather prediction with an analog ensemble; *Mon. Weather Rev.* **141** 3498–3516.

Dimri A P 2006 Surface and upper air fields during extreme winter precipitation over the western Himalayas; *Pure Appl. Geophys.* **163** 1679–1698.

Dimri A P and Mohanty U C 2007 Location-specific prediction of maximum and minimum temperature over the Western Himalayas; *Meteorol. Appl.* **14** 79–93.

Dimri A P and Das S K 2011 Wintertime climatic trends in the western Himalayas; *Clim. Change* **111** 775–800.

Dimri A P and Chevuturi A 2014 Model sensitivity analysis study for western disturbances over the Himalayas; *Meteorol. Atmos. Phys.* **123**(3–4) 155–180, <https://doi.org/10.1007/s00703-013-0302-4>.

Dimri A P and Niyogi D 2013 Regional climate model application at subgrid scale on Indian Winter Monsoon over the Western Himalayas; *Int. J. Climatol.* **33** 2185–2205.

Dimri A P, Yasunari T, Wiltshire A, Kumar P, Mathison C, Ridley J and Jacob D 2013 Application of regional climate models to the Indian winter monsoon over the Western Himalayas; *Sci. Total Env.* **468** S36–S47, <http://dx.doi.org/10.1016/j.scitotenv.2013.01.040>.

Eccel E, Ghielmi L, Granitto P, Barbiero R, Grazzini F and Cesari D 2007 Prediction of minimum temperatures in an alpine region by linear and non-linear post-processing of meteorological models; *Nonlinear Process. Geophys.* **14** 211–222.

Fiddes J and Gruber S 2014 TopoSCALE v.1.0: Downscaling gridded climate data in complex terrain; *Geosci. Model Dev.* **7** 387–405.

Gibergans-Baguena J and Llasat M C 2007 Improvement of the analog forecasting method by using local

- thermodynamic data. Application to autumn precipitation in Catalonia; *Atmos. Res.* **86**(3–4) 173–193, <https://doi.org/10.1016/j.atmosres.2007.04.002>.
- Gutzler D S and Shukla J 1984 Analogs in the wintertime 500 mb height field; *J. Atmos. Sci.* **41** 177–189.
- Hall T J, Thessin R N, Bloy G J and Mutchler C N 2010 Analog sky condition forecasting based on a k-NN algorithm; *Wea. Forecasting* **25** 1463–1478.
- Hamill T M and Colucci S J 1997 Verification of Eta–RSM short range ensemble forecasts; *Mon. Weather Rev.* **125** 1312–1327.
- Hamil T M and Whitaker J S 2006 Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application; *Mon. Weather Rev.* **134** 3209–3229.
- Hamil T M and Whitaker J S 2007 Ensemble calibration of 500-hPa geopotential height and 850-hPa and 2-m temperatures using reforecasts; *Mon. Weather Rev.* **135** 3273–3280.
- Hatwar H R, Yadav B P and Rama Rao Y V 2005 Prediction of western disturbances and associated weather over Western Himalayas; *Curr. Sci.* **88** 913–920.
- Jaun S and Ahrens B 2009 Evaluation of a probabilistic hydrometeorological forecast system; *Hydrol. Earth Syst. Sci.* **13** 1843–1877.
- Jensenius J S 1990 A statistical comparison of the forecasts produced by the NGM and LFM for the 1987/88 cool season; *Wea. Forecasting* **5** 116–127.
- Joshi P and Ganju A 2012 Maximum and minimum temperature prediction over Western Himalaya using artificial neural network; *Mausam* **63** 283–290.
- Joshi P and Ganju A 2013 Downscaling of MM5 model output using artificial neural network over western Himalaya; *Mausam* **64** 221–230.
- Joshi J C, Kumar T, Srivastava S and Sachdeva D 2017 Optimisation of hidden Markov model using Baum–Welch algorithm for prediction of maximum and minimum temperature over Indian Himalaya; *J. Earth Syst. Sci.* **126**, <https://doi.org/10.1007/s12040-016-0780-0>.
- Junk C, Delle Monache L and Alessandrini S 2015 Analog-based ensemble model output statistics; *Mon. Weather Rev.* **143** 2909–2917, <https://doi.org/10.1175/mwr-d-15-0095.1>.
- Kalnay E, Kanamitsu M and Baker W E 1990 Global numerical weather prediction at NMC; *Bull. Am. Meteorol. Soc.* **71** 1410–1428.
- Kumar T S V V and Krishnamurti T N 2006 High resolution numerical weather prediction over the Indian subcontinent; *J. Earth Syst. Sci.* **115** 529–555.
- Livezey R E 1994 The evaluation of forecasts; In: *Analysis of climate variability: Application of statistical techniques* (eds vonStorch H and Navarra A, Springer-Verlag, pp. 177–196.
- Lorenz E N 1969 Atmospheric predictability as revealed by naturally occurring analogues; *J. Atmos. Sci.* **26** 636–646.
- Maini P, Kumar A, Rathore L S and Singh S V 2003 Forecasting maximum and minimum temperatures by statistical interpretation of numerical weather prediction model output; *Wea. Forecasting* **18** 938–952.
- Menegoz M, Gallee H and Jacobi H W 2013 Precipitation and snow cover in the Himalaya: From reanalysis to regional climate simulations; *Hydrol. Earth Syst. Sci.* **17** 3921–3936.
- Mohanty U C and Dimri A P 2004 Location specific prediction of probability of occurrence and quantity of precipitation over Western Himalayas; *Wea. Forecasting* **19** 520–533.
- Murphy A H 1988 Skill scores based on mean square error and their relationship to the correlation coefficient; *Mon. Weather Rev.* **116** 2417–2424.
- Murphy A H and Winkler R L 1987 A general framework for forecast verification; *Mon. Weather Rev.* **115** 1330–1338.
- Nagarajan B, Delle Monache L D, Hacker J P, Rife D L, Searight K, Knierul J C and Nipen T N 2015 An evaluation of analog-based postprocessing methods across several variables and forecast models; *Wea. Forecasting* **30** 1623–1643.
- Novak D R, Bailey C, Brill K F, Burke P, Hogsett W A, Rausch R and Schichte M 2014 Precipitation and temperature forecast performance at the Weather Prediction Center; *Wea. Forecasting* **29** 489–504.
- Obled C, Bontron G and Garçon R 2002 Quantitative precipitation forecasts: A statistical adaptation of model outputs through an analogues sorting approach; *Atmos. Res.* **63** 303–324.
- Palazzi E, Von Hardenberg J and Provenzale A 2013 Precipitation in The Hindu-Kush Karakoram Himalaya: Observations and future scenarios; *J. Geophys. Res.* **118** 85–100.
- Radinovic D 1975 An analogue method for weather forecasting using 500/1000 mb relative topography; *Mon. Weather Rev.* **103** 639–649.
- Raje D and Mujumdar P P 2011 A comparison of three methods for downscaling daily precipitation in the Punjab region; *Hydrol. Process.* **25** 3575–3589, <https://doi.org/10.1002/hyp.8083>.
- Rangachary N and Bandyopadhyay B K 1987 An analysis of the synoptic weather pattern associated with extensive avalanching in Western Himalaya; *Avalanche formation, movement and effects*, IAHS Publ no. 162, pp. 311–316.
- Roy Bhowmik S K and Durai V R 2008 Multi-model ensemble forecasting of rainfall over Indian monsoon region; *Atmosfera* **21** 225–239.
- Shank D B, Hoogenboom G and McClendon R W 2008 Dew point temperature prediction using artificial neural networks; *J. Appl. Meteorol. Clim.* **47** 1757–1769, <https://doi.org/10.1175/2007JAMC1693.1>.
- Shekhar M S, Kumar M S, Joshi P and Ganju A 2014 Mountain weather research and forecasting over western and central Himalaya by using mesoscale models; *Int. J. Earth Atmos. Sci.* **1** 71–84.
- Sievers O, Fraedrich K and Raible C 2000 Self-adapting analog ensemble predictions of tropical cyclone tracks; *Wea. Forecasting* **15** 623–629.
- Singh D and Ganju A 2005 Expert system for prediction of avalanches; *Curr. Sci.* **94** 1076–1081.
- Singh D, Dimri A P and Ganju A 2008 An analogue method for simultaneous prediction of surface weather parameters at a specific location in the Western Himalaya in India; *Meteorol. Appl.* **15** 491–496.
- Singh D, Srinivasan K, Ganju A and Snehmani 2010 Comparative study of performance of different weather forecast models at specific sites in northwest Himalaya in India; *Meteorol. Atmos. Phys.* **107**(3) 137–147, <https://doi.org/10.1007/s00703-010-0071-2>.
- Singh D, Sharma V and Juyal V 2015 Observed linear trend in few surface weather elements over the northwest Himalayas (NWH) during winter season; *J. Earth Syst. Sci.* **124** 553–565.
- Singh D, Kumar A and Shekhar M S 2019 Spatiotemporal variability of binary weather patterns and precipitation

- amounts of short time intervals during winter period over the Northwest Himalaya (NWH); *J. Earth Syst. Sci.* (in press).
- Srinivasan K, Ganju A and Sharma S S 2005 Usefulness of meso-scale weather forecast for avalanche forecasting; *Curr. Sci.* **88** 921–926.
- Srinivasan K, Kumar A, Verma J and Ganju A 2010 Statistical downscaling of MM5 model output to better assess avalanche threats; *Ann. Glaciol.* **51** 14–18.
- Storch H and Zwiers F W 1999 *Statistical analysis in climate research*; Cambridge University Press, ISBN 0521450713.
- Suranjana S and Van Den Dool H M 1988 A measure of the practical limit of predictability; *Mon. Weather Rev.* **116** 2522–2526.
- Terzago S, Hardenberg J V, Palazzi E and Provenzale A 2014 Snowpack changes in The Hindu Kush–Karakoram–Himalaya from CMIP5 Global Climate Models; *J. Hydrometeorol.* **15** 2293–2313, <https://doi.org/10.1175/jhm-d-13-0196.1>.
- Tiwari P R, Kar S C, Mohanty U C, Dey S, Sinha P, Raju P V S and Shekhar M S 2014 Dynamical downscaling approach for wintertime seasonal scale simulation over the western Himalayas; *Acta Geophys.* **62**(4) 930–952, <https://doi.org/10.2478/s11600-014-0215-8>.
- Toth Z 1989 Long-range weather forecasting using an analog approach; *J. Clim.* **2** 594–607.
- Toth Z and Kalnay E 1993 Ensemble forecasting at NMC: The generation of perturbations; *Bull. Am. Meteorol. Soc.* **74** 2317–2330.
- Tracton M S and Kalnay E 1993 Operational ensemble prediction at the National Meteorological Center: Practical aspects; *Wea. Forecasting* **8** 379–398.
- Van Den Dool H M 1989 A new look at weather forecasting through analogues; *Mon. Weather Rev.* **117** 2230–2247.
- Whitaker J S, Wee X and Vitart F 2006 Improving week-2 forecasts with multimodel reforecast ensembles; *Mon. Weather Rev.* **34** 2279–2284.
- Woodcock F 1980 On use of analogues to improve regression forecasts; *Mon. Weather Rev.* **108** 292–297.
- Xavier P K and Goswami B N 2007 An analog method for real-time forecasting of summer monsoon subseasonal variability; *Mon. Weather Rev.* **135** 4149–4160.
- Zorita E and Storch H V 1999 The analog method as a simple statistical downscaling technique: Comparison with more complicated methods; *J. Clim.* **12** 2474–2489.

Corresponding editor: N V CHALAPATHI RAO